



Tax Talk Thursdays with Arkansas Asset Builders & Center on Budget and Policy Priorities

Using data sources to inform your free tax filing program

July 18, 2024

Presented by the Center on Budget and Policy Priorities' Get It Back Campaign

Presenters:

[Joyia Yorgey](#) – Executive Director, Arkansas Asset Builders

[Claire Zippel](#) – Senior Research Analyst, Center on Budget and Policy Priorities

Data Sources for Free Tax Filing Programs

Claire Zippel, Center on Budget and Policy Priorities

CBPP's Program Participation Dashboard

Relevant info you can find on the dashboard:

- Estimated characteristics of people eligible for the EITC
 - Includes common industries and occupations, common languages spoken, education, race and ethnicity, and other demographics
 - Listed by state and for selected counties and metropolitan areas
- Potential uses
 - Targeting outreach and services to communities with many potential EITC filers
 - Educating partners and funders about the population eligible for the EITC
- Available for download as Excel spreadsheets with sub-state data for a more granular analysis
- Caveats
 - Figures are estimates based on Census Bureau household survey data
 - Shouldn't be compared with IRS data on actual EITC claims
 - Currently available data is for 2018 (as of time of session)

- EITC and CTC claims
 - Dashboard includes data on EITC and CTC claims by state, 118th congressional district, and state legislative districts (state house & state senate districts)
 - The IRS publishes tax return data by state, county, metropolitan area, 117th congressional district, and ZIP code
- Potential uses
 - Educating policymakers about how many of their constituents are helped by the EITC and CTC and how much money/ economic activity flows to their district thanks to the EITC and CTC

Census Bureau Data

Data on low-income workers and families

- American Community Survey data from data.census.gov
 - Numbers of individuals and families below the poverty line, number of households by income
 - Available down to the census tract (neighborhood) level
 - Can give a rough sense of where people potentially eligible for the EITC or ACTC live
 - Can display as a map format, being able to select specific geographic regions for analysis
- Employment statistics from [OnTheMap](#)
 - Number of low-paid workers by place of work
 - Available down to the block level, and by industry
 - Can give rough sense of where workers potentially eligible for the EITC are employed

Tools for Geographic Analysis of Client Data

- [Census geocoder](#)
 - Find the county, city, congressional/state legislative district, metro area, and urban/rural status of an address
- [Geocorr zip code crosswalk](#)
 - Find the county, city, congressional/state legislative district, metro area, and urban/rural status of a zip code

Joyia Yorgey, Arkansas Asset Builders

What are the questions we want data to answer?

- Who needs service?
 - Who are you already serving? Who may qualify for free tax help that you are NOT already serving?
- Where is the need?
 - Where are your current clients located? Which areas near you are NOT being effectively served by a free tax site?
- What is the impact?
 - What are the actual outputs and outcomes of your VITA program? What are the long-term impacts you seek? What baseline or benchmarks are you comparing against?

Who gets free tax help?

Current Population

- **Internal data sources**, such as: [TaxSlayer](#), other custom intake forms, demographic surveys, client management tool/apps, social media audience
 - Population may include people who booked an appointment but never showed up!
- **External data sources**: [ALICE \(asset limited, income constrained, employed\) state reports](#), [FPL \(federal poverty level\)](#), SNAP or other benefit eligibility charts, [FFEIC income stats](#)
 - ALICE represents a household's ability to pay for goods and services to survive
 - FFEIC income stats can be used as a comparative tool to measure whether the clients you serve are above or below certain income medians
 - Banks use this measure to validate their Community Reinvestment Act (CRA) requirements

Possible Population

- **External data sources**, such as: U.S. Census (American Community Survey), CBPP Program Participation Dashboard, [IRS Statistics of Income \(SOI\) - Form 1040](#)
 - IRS Statistics of Income represents data on largest local employment sectors or specific lower-wage employers; can be segmented by income level
 - Includes data on tax credit claims, use of paid preparers & free tax prep, and all items on Form 1040 by income level

Where is the need?

Available Geography

- **TaxSlayer:** Client street address, city, state, ZIP code, location of your tax site(s)
 - Custom intake forms may collect county or other relevant geographies

Possible Geography

- [Census place designations \(e.g. tracts, blocks\)](#)
- **Political designations**, such as: congressional districts, state legislative districts, county voting districts, city limits
 - Example: counting how many clients are from each voting district in order to support a legislation
- Utilities service areas
 - Utility service companies may support funding or sharing information if many clients are within their service area.
- Urban/rural designations: [USDA](#), [Census](#), and [OMB](#) definitions
 - These tools can be used to define addresses by urban/rural; however, there are multiple definitions to consider
- To map street addresses onto other geographic designations, we either need to use “crosswalks” or tools to geocode the data

What impact is being made?

Outputs, Outcomes, and Impacts

- Potential data points to consider:
 - Number of tax returns prepared, number of taxpayers served, amounts of tax refunds committed to savings, volunteer hours served, other services offered by your program, annual and/or cumulative tax refunds, annual and/or cumulative prep fees saved, increase in client household savings or net worth, increase in client financial well-being over time, economic impact of tax refunds on service area, increase in overall population savings rates or other financial focus area (see [Prosperity Now Scorecard](#)), decrease in overall target population use of paid tax preparation services (see IRS SOI tax data), decrease in overall population use of instant refund or refund advance products (IRS SOI data), **increase in use of free tax filing services!** (IRS SOI data)

Resources

- [Center on Budget and Policy Priorities' Program Participation Dashboard](#)
- [Census data](#)
- [OnTheMap employment data](#)
- [Census geocoder](#)
- [Geocorr crosswalk tool](#)
- [TaxSlayer guide](#)
- [American Community Survey](#)
- [ALICE](#)
- [Federal poverty guidelines](#)
- [FFEIC income data](#)
- [IRS Statistics of Income \(SOI\) - Form 1040](#)
- [Census place designations](#)
- [Census urban-rural classification](#)
- [USDA urban-rural classification](#)
- [OMB urban-rural classification](#)
- [Prosperity Now Scorecard](#)

Q&A

- **What year is the CBPP Program Participation Dashboard being updated until? What year for the congressional and state legislative districts?**
 - The data on characteristics of people who are eligible for the EITC is currently for 2018. It will likely be updated to 2024, but the timing is TBD. The data by congressional and state legislative district reflects the districts after redistricting in 2020. The districts used are the ones from the 2022 election. In terms of the tax data, it is for the 2021 tax year (an atypical tax year). Plan to update the data this fall after the IRS releases the 2022 tax year data.
- **How would you use the dashboard information to calculate which geographic areas are over/underperforming in terms of EITC participation rates?**
 - It is not recommended that the dashboard be used for that purpose. Data such as how many people are eligible for the EITC and how many are claiming it cannot be derived from the dashboard. The dashboard can be useful to get a sense of what the EITC-eligible population is like and who is benefiting, or could benefit, from the credit.

- **Geographic areas are different in terms of average income or percentage of qualifying households. How would we account for this when comparing legislative districts, especially at more local levels?**
 - On the dashboard, there is the percentage of all filers who are claiming the EITC. It could be higher in certain districts due to other circumstances, but it cannot be determined precisely. One way you could resolve this is by using IRS data on AGI or what percentage of filers were below an income level and compare them between higher and lower income districts to give context on the number of EITC claims in the area.
- **Regarding the Geocorr tool, why would one small town with one ZIP code be shown with three different counties?**
 - Possible explanations may be that the ZIP code encompasses areas outside of the town, there is only one ZIP code in the town, and that there are people in surrounding areas who are also in that ZIP code. It could be a one-to-many relationship between a ZIP code and different geographies.
- **Even though AGI does not capture all income coming in (it excludes non-taxable SSA, etc.), do folks find that funders are okay with using AGI to determine, for example, if a client is below ALICE?**
 - AGI is not a full picture of the income that households receive. As long as you are clear that AGI is the metric you are using for household income, it should be okay for funders. In TaxSlayer, there is one report that includes an indicator of whether a household receives Social Security Income or not. You could filter your data on those households that are not receiving SSI to be more confident that client AGI more accurately reflects the household income.
- **What are the components of the financial well-being score? How is it measured, etc.?**
 - The [financial well-being score](#) is a standard that the Consumer Financial Protection Bureau uses. It has different models, one is a longer questionnaire, and one is a shorter questionnaire that is five questions. It has been found that the shorter questionnaire does accurately reflect, within a certain tolerance, the financial well-being that would be shown by the longer questionnaire. The questions reflect the person's feelings about their own financial situation, so it is subjective. An example is, "Does money control your life?" Responses range from "completely" to "sometimes" to "not at all." There is a client scoring sheet that is based on the client's age. You can derive numeric results based on the responses and age that can be compared to state and national benchmarks that the CFPB produces.
- **For the Program Participation Dashboard, are you able to explain how the total EITC amount and refundable portion of the EITC amount would be different as the EITC is**

fully refundable? Do you mean the portion that is left over to be refunded after accounting for taxes? If so, how do you segregate tax withheld and other refundable credits?

- On the dashboard, the refundable EITC refers to the amount that the person would have gotten as a refund after filing taxes. Total EITC amount includes the amount of the EITC that offsets the amount of taxes someone owes. If you hover over the names of the columns, there is explanatory text that may clarify this.
- **What definition of rural do you use, population or in relation to MSA, etc.? Different entities use different definitions of rural.**
 - There are multiple definitions of rural, which makes it extra complicated to identify and measure your organization's ability to serve rural taxpayers. The easiest definition to use is from [the USDA](#), which classified every county as either metropolitan (urban) or non-metropolitan (rural). You can use a basic crosswalk tool to map ZIP codes onto counties or create your own database where each ZIP code corresponds to a single county. You can then identify how many of your clients live in each county (approximately) and then total up all the clients in counties that meet the USDA rural definition.