Measuring Take-up of the California EITC with State Administrative Data

John Iselin
UMD College Park

Taylor Mackay
UC Irvine

Matthew Unrath[1]
UC Berkeley

October 13, 2022

Abstract

The Earned Income Tax Credit (EITC) is the largest cash-based means-tested transfer program in the United States, delivering $63 billion to 25 million households in 2019. Twenty-eight states supplement the federal credit. Though an estimated one-fifth of eligible households fail to claim the federal credit, little is known about take-up of these state supplements. We use administrative data from California on the population of Supplemental Nutrition Assistance Program (SNAP) recipients linked to state tax records to estimate the number of households who are eligible for California's supplement to the federal EITC (CalEITC) but do not claim it. We find that over 400,000 households who receive SNAP benefits and who were eligible for the state EITC in 2017 did not receive the credit. This includes approximately 40,000 eligible households who claimed the federal EITC but not the state credit; nearly 98,000 eligible households who filed a state tax return but did not claim the state or federal credit; and roughly 270,000 eligible households who did not file a state tax return. The corresponding take-up rate for the CalEITC among eligible SNAP-enrolled households was 54%. Altogether, these households left on the table a total of $71 million in state EITC funds. If received, these credits would have increased incomes among these households by 2.7% and increased total state EITC outlays by 20%.

Keywords: Administrative Data, California Earned Income Tax Credit, Social Safety Net, Take-up.
JEL Classification: H240, H710, I380

[1]Contact: jiselin@umd.edu; tmackay@uci.edu; unrath@berkeley.edu. This research was conducted through the California Policy Lab at UC Berkeley. We thank Evan White, Aparna Ramesh, Jesse Rothstein, and Johanna Lacoe for their support and guidance. At the California Franchise Tax Board, we thank Sean McDaniel, Monica Trefz, Ian Kiltz, Chad Angaretis, Silvano Guitierrez, Bob Schlie, and Xudong Chen, and especially Allen Prohofsky and Julie Moreno. At the California Department of Social Services, we thank Alexis Fernandez, Kim McCoy-Wade, Jianjun Chen, and Brittney Gossard. We also thank Katherine Meckel, David Splinter, Jacob Mortensen, Katherine Meckel and National Tax Association meeting participants, as well as Tatiana Homonoff and the participants at the Online Public Finance Seminar, for their helpful feedback. Support for this research was provided in part by University of Wisconsin Institute for Research on Poverty, the National Institute on Aging, and the Robert Wood Johnson Foundation's Policies for Action program. The views expressed here do not necessarily reflect the views of the Foundation.

1 Introduction

The Earned Income Tax Credit (EITC) is the largest means-tested cash transfer program in the United States. In 2019, 25 million households received about $63 billion, with an average benefit of approximately $2,500 (IRS, n.d.*b*). One-fifth of all tax units, and nearly one-half of tax units with children, claimed the EITC. California introduced its own EITC in 2015, joining 27 other states and the District of Columbia that supplement the federal EITC. In 2017, approximately 1.5 million tax units claimed this state supplement, known as the CalEITC, receiving a total of $351 million (Davis and White, 2019).

Numerous studies have documented the EITC's beneficial effects on work, income, and poverty; children's educational achievement and attainment; and adult and infant health (see reviews in Hoynes and Rothstein, 2017; Nichols and Rothstein, 2016). For eligible households to realize the EITC's many beneficial effects, they must file a tax return and claim it. However, the IRS estimates that one in five households who are eligible for the EITC do not receive it (IRS, n.d.*a*). While the EITC's take-up rate exceeds that of many other means-tested programs, this level of non-participation means millions of households fail to receive critical financial assistance that is available to them.

Exactly how many eligible households fail to claim the EITC remains a disputed statistic. The unofficial, but most commonly cited, measure of take-up relies on matching responses from the Current Population Survey's Annual Social and Economic Supplement (CPS ASEC) and the American Community Survey (ACS) to IRS records.[2] While a significant improvement over past approaches, this method relies upon imputations of EITC eligibility from survey data that are prone to error (Jones and Ziliak, 2019). Moreover, this approach cannot be used to assess take-up of state-level EITCs, since the IRS only shares copies of federal returns with Census, and these federal returns do not include indicators for claiming of state tax credits. As such, no estimate of take-up of the California EITC or any other state supplement exists.

To provide a comprehensive look at CalEITC participation and address limitations of the current approach to measuring EITC take-up, we propose a new method to measure eligibility and participation that relies solely on state administrative data. We use enrollment records for the CalFresh program (the state's instantiation of the Supplemental Nutrition Assistance Program [SNAP], or food stamp program) linked to administrative earnings records and California state income tax returns. CalFresh program records provide detailed family composition information, which we use to identify likely eligible tax units even for those families who do not file tax returns. We use these program records for all CalFresh participants in 2017, matched to tax return data for that year, to measure eligibility and participation across both filers and non-filers in our CalFresh universe. For each population, we use available information to measure three components of eligibility (filing status, earned income, and number of qualifying children) and then identify which of the seemingly eligible tax units claimed the state credit.

Importantly, these data contain two pieces of information critical to measuring EITC eligibility that tend not to be available in tax records or household surveys. First, we observe the date of birth for each CalFresh recipient, which allows us to identify the ages of dependents on tax returns and identify likely child dependents among non-filers. Second, CalFresh enrollment records provide monthly snapshots of household composition, which enable us to observe whether adults reside with children for a certain number of months in the tax year and to group non-filing adults and children into likely tax units. The novelty of this information and our data match enable us to make important progress on measuring eligibility among filing non-claimants and the larger population of non-filers.

We find that more than 400,000 households who received CalFresh benefits and who appeared eligible for the state EITC did not receive the credit in 2017. This includes over 40,000 who claimed the federal EITC but not the state credit; nearly 100,000 who filed a state tax return but did not claim either the federal or state credit; and roughly 270,000 who did not file a state tax return. The corresponding take-up rate for these households was 54%. Altogether, these households left on the table a total of over $71 million in state EITC funds. If received, these

[2]The ACS has a larger sample size and could be used to produce state-level estimates of take-up, but doing so would mean sacrificing precision in the estimation of income and family structure.

credits would have increased incomes among these households by 2.7% and increased state EITC outlays by 20%. We also replicate our results for tax year 2016 in Appendix G, finding a take-up rate of 43%. The increase in the take-up rate between 2016 and 2017 was driven by a sharp growth in the number of eligible filers.

Two-thirds of non-claiming is attributable to eligible households not filing a state return. Among filers and non-filers, the majority of non-claimants are single individuals without dependents; these non-claimants appeared eligible for about $85, on average. At the same time, we show that tens of thousands of tax units with dependents and non-filing households with children failed to claim the CalEITC, and unclaimed amounts for these households was much higher – an average of $251 for filers and $596 for non-filers. We also present estimated participation rates by individuals' race and filers' preparation method. We find that eligible Black and Hispanic filers were less likely to claim the credit, and that Black, Hispanic, and American Indian or Alaskan Native non-filers were more likely to be eligible. Overall, eligible Asian adults were most likely to claim the CalEITC, and eligible American Indian and Alaskan Native adults were least likely to claim. Finally, we also show that eligible households who claimed the federal EITC were much less likely to also claim the state credit if they filed using a paid preparer.

In the preceding paragraphs and throughout the paper, we refer to actual and simulated tax units as "households". We recognize that this label is not precise. Tax units and households are not synonymous, as discussed below and at length in the paper. However, referring to simulated tax units (i.e., a collection of SNAP enrollees who we predict would appear on a return together if they were to file one) as "tax units", per se, risks even more confusion, since these individuals do not appear on a tax return together. Since the paper is unique in its focus on this concept, we are especially sensitive to that source of confusion. Instead, we use the more commonly understood language of "household" to refer to immediate families that could appear on a return together and likely reside together. Ultimately, like the phenomenon we're studying, the language is necessarily nebulous. Indeed, a principal contribution of this paper is to document the extent to which the same households or families present differently with respect to a safety net program and tax policy.

This paper makes several important contributions to the literature on participation in means-tested programs, and the EITC in particular. First, we produce the first-ever estimate of take-up of a state-level EITC. We also demonstrate how non-participation is due to non-claiming among eligible filers and non-filing among eligible households. Our approach also allows us to estimate take-up by detailed demographic characteristics. These estimates will provide valuable insights to both economists and policymakers. For the former group, understanding the take-up of state EITCs can improve the understanding of the administrative burdens implicit in running social safety net programs through the tax code. Further, if researchers want to use state-EITCs as a source of variation in the study of the effect of policy on labor market outcomes, they will need to accurately gauge the take-up of these programs. For the policymakers, knowing how many and which households are eligible for and not claiming the CalEITC is important in shaping both outreach and efforts to reduce the barriers faced by tax filers.

Second, we contribute to a narrow literature studying shifting household composition interacts with social and tax policy. Jones and O'Hara (2016) and Splinter, Larrimore and Mortenson (2017) find that adults residing in the same households strategically claim children to maximize after-tax income. Larrimore, Mortenson and Splinter (2021) use information returns to group individuals into shared addresses and multi-tax-unit households to compare income and inequality measures at the household-level, as opposed to the tax unit level. We extend this work by studying household presentations in the one of the largest safety net programs, documenting the extent to which SNAP cases overlap with tax units. For example, we find that 42 percent of the individuals who both filed taxes and claimed CalFresh in 2017 appeared on a return where at least one individual did not receive CalFresh, and one out of five children who was enrolled in CalFresh and appeared on a return was claimed by an adult who was not on the child's primary CalFresh case.

The paper proceeds as follows. In Section 2, we describe the federal and California EITC and summarize previous work estimating EITC participation, as well as potential explanations for incomplete take-up. In Section 3, we describe our unique linked data that make our project possible. In Section 4, we describe how we identify

CalEITC-eligible CalFresh recipients, and we present participation estimates. Section 5 presents additional results and estimates of take-up for various subgroups. In Section 6, we discuss the relevance of our findings to strategies for increasing take-up. In Section 7, we conclude.

2 Background

2.1 Structure of the federal and state credits

At low earnings levels, the amount of the federal EITC for which households are eligible increases with each dollar of earnings until it reaches a maximum value. The credit is then stable across a range of earnings until earnings reach a particular level, after which the value declines as earnings rise, eventually reaching zero. This structure encourages labor force participation among those who would not otherwise work (Nichols and Rothstein, 2016). The basic shape of the schedule is the same for all households, but the quantitative parameters differ. Families with children qualify for much larger credits, and at higher earnings levels, than do families without children. Married couples can have higher earnings before the credit begins to phase out than can single filers.

The blue area in Figure 1 shows the federal EITC schedule for a single filer with two children in tax year 2017. The maximum credit, available to filers with earnings between $14,040 and $18,340, was close to $6,000. Filers with earnings above or below this range qualified for smaller credits, so long as their earnings were positive and did not exceed $45,007.[3]

In 2015, California joined 27 other states and the District of Columbia in offering eligible taxpayers a supplement to the federal credit. Unlike most other states which offer a match equal to some percentage of the federal credit, California adopted its own schedule, targeting assistance at households with the lowest earnings. The gold area in Figure 1 shows the California schedule in tax year 2017. The largest credit is available to families with earnings around $7,500, though households with earnings up to $25,000 can still qualify for smaller credits.

To be eligible for the federal or California EITC, taxpayers must file as single, head of household or married filing jointly, have positive earned income and adjusted gross income (AGI) below a certain maximum threshold, which varies by household size, filing status, and tax year.[4] Tax units must also have investment income below a set threshold ($3,450 in 2017).

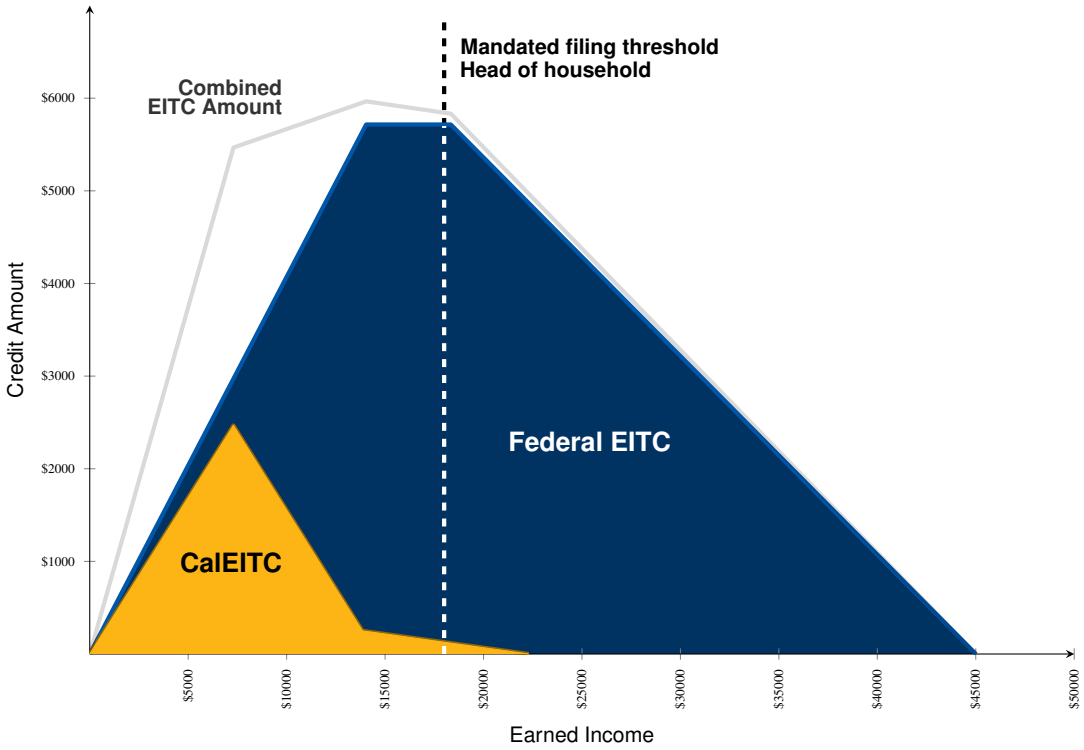
Tax units are eligible for much larger credits if they have qualifying children. EITC eligibility depends on a different count of children than do other components of the tax code. An EITC qualifying child must have a valid SSN; must be under 19, or be under 24 if a full-time student, or be any age if the child has a total or permanent disability; must reside in the household for at least half the year; cannot be claimed by a different filer; and must be a near relation or an adopted or foster children of the filer.

Families with very low earnings levels are typically not required to file tax returns, and only their eligibility for the EITC and other tax credits, or the refund of overly withheld income taxes, creates an incentive to do so. Prior to the introduction of the CalEITC, there was little reason for the lowest-income families in California to file a state return, even if they filed a federal return in order to claim the federal credit. This possibility raises concerns that many EITC-eligible families still might not file a state return or, if they do, not know to claim the CalEITC. This issue is relevant to all state supplements, but the the concentration of CalEITC benefits at very low earnings levels heightens the concern in this setting.

[3] This schedule assumes the family does not have investment or other unearned income, which can cause a family to lose eligibility.

[4] Tax units with income from sources other than earnings and whose AGI is above a certain level are instructed to calculate their EITC amount according to both their earned income and their AGI, and claim the lesser of the two amounts.

Figure 1: Federal and California EITC schedules for a single-parent family with two children, tax year 2017



Notes. Figure 1 illustrates the federal (blue) and state (gold) EITC schedules for a head of household with two children. The gray line illustrates the combined value of the EITC for a filer who claims both credits. The dotted line denotes the filing threshold for a head of household in tax year 2018, which was $18,000; families with incomes below this threshold are generally not required to file returns.

2.2 Estimates of take-up rates

An accurate estimate of EITC participation requires a sample of the eligible population (including information on household composition, earnings, citizenship status, and more) and a measure of EITC receipt. Few administrative or survey datasets contain even proxies for each of these variables, let alone measures that match the IRS definitions.

A number of studies use survey data to simulate eligibility and receipt. Using the Survey of Income and Program Participation (SIPP), Scholz (1994) estimated that take-up in tax year 1990 was between 80 and 86%, and that 1.3 to 2.0 million eligible taxpayers failed to claim the credit. In 2001, using multiple household surveys, the Government Accountability Office (GAO) estimated EITC participation to be 75% and dollar participation to be almost 90% in tax year 1999. In 2002, IRS researchers estimated that about 15% of eligible households in the CPS and SIPP did not file a tax return, suggesting a maximum take up rate of 85%. Using the CPS, Blumenthal, Erard and Ho (2005) estimated EITC participation in tax year 1988 to be nearly 90% among legally obligated filers, only 30 to 40% among those not legally obligated to file, and roughly 70% among all eligible taxpayers. A major drawback to the survey-only evidence, however, is that these data typically do not have direct measures of EITC receipt.

In 2004, the IRS and Census Bureau completed the first exact match of tax records and a household survey to estimate EITC take-up. Plueger (2009) describes the match of households in the Individuals Return Tax file to households interviewed for the Current Population Survey Annual Social and Economic Supplement (CPS ASEC) for tax year 2005. With these merged data, researchers estimated take up of the EITC to be 75%. Jones (2014) uses the same match to measure EITC eligibility and participation between 2009 and 2014 and identify factors that predict take-up. The IRS continues to partner with the Census to produce these estimates and to refine this merge. Their results continue to be the most commonly cited measure of take-up.

Though an improvement over previous efforts to measure take-up, this match still suffers from important data challenges. One limitation is that the CPS ASEC collects information about household composition at the time the household is surveyed, but key tax variables depend on households' circumstances at year's end or over the course of the entire year. For example, it is not possible to determine for how long an ASEC child lived with a parent in the relevant tax year, a key consideration when determining whether a dependent is a qualifying child for the purposes of the EITC. Marital relationships may also change between the end of the tax year and the ASEC survey. Further, the matched CPS-IRS data suffers from selection and measurement issues. For example, annual income is measured via respondents' own self-reports or via imputation, both of which are prone to error (Meyer and Mittag, 2019; Jones and Ziliak, 2020; Meyer et al., 2020).[5] In addition to standard selection concerns from survey non-response issues, researchers were forced to drop almost half of CPS households from the exact match process because they had incomplete survey information, did not complete the ASEC, or did not agree to have information matched to the IRS (Jones and Ziliak, 2019).

Other research on EITC take-up has found that non-filers account for most of the unclaimed dollars. Of households that qualify for the federal EITC but do not claim it, approximately two-thirds do not file any tax return (as cited in Cranor, Kotb and Goldin, 2019). In contrast to non-filers, a large majority (91.5%) of EITC-eligible households that do file a return claim the credit. However, these estimates are based on federal data, where the IRS has developed sophisticated tools for inferring EITC eligibility for those who do not claim it and conducts substantial outreach to non-claiming households (Bhargava and Manoli, 2015). There is no available evidence about take-up of state credits.

Similar to this work, Maag et al. (2015) use SNAP enrollment records from Florida linked to IRS records to validate whether dependents claimed on a federal return appear to be qualifying children according to the SNAP caseload data. Our purpose is to measure all components of CalEITC eligibility across both filers and non-filers.

2.3 Explanations for incomplete take-up

In addition to the narrow literature investigating EITC participation, our paper also contributes to a broader literature investigating the extent of and the reasons for incomplete participation in an array of means-tested programs. Researchers point to three broad explanations for why eligible households may fail to take up benefits for which they're eligible: learning, compliance, and stigma costs (Currie, 2006). In the context of the EITC, households may be unaware of the credit or that they're eligible; they may know the EITC exists but find the process of filing and claiming it too costly or burdensome; or the stigma costs or stresses of tax filing may outweigh the financial benefit of claiming.

There is some evidence about the relative importance of each of these factors in the context of the EITC. Several surveys of eligible households suggest many workers are unaware of the EITC or its structure (Liebman, 1998; Phillips, 2001; Romich and Weisner, 2000; Smeeding, Phillips and O'Connor, 2000; Maag et al., 2015). At the same time, other qualitative evidence suggests many eligible taxpayers are highly aware of the benefits of tax filing, which is due largely to the EITC (Edin, Tach and Halpern-Meekin, 2014; Halpern-Meekin et al., 2015). Awareness also appears to vary across different communities (Chetty, Friedman and Saez, 2013). Some research suggests that providing information to tax filers about the EITC and their likely eligibility can increase EITC claims (Bhargava and Manoli, 2015). Chetty, Friedman and Saez (2013) study the effect of tax preparers explaining the EITC's incentive to taxpayers already claiming the credit and find that taxpayers who received this information are likely to claim higher EITC amounts in the next year.

Evidence on the effect of sending information about the EITC to potentially eligible non-filers is more mixed. Jones (2010) finds small participation effects of an employer-run program designed to increase take-up of an advance on eligible workers' likely EITC amount. Guyton et al. (2016) studies the effect of sending mailings to eligible, non-

[5] Of the CPS respondent households who can be matched to a return, the difference between the CPS reported-AGI and that on the tax return exceeds $10,000 for 14% of households.

filing, eligible households, finding small but statistically significant effects on EITC take-up. Cranor, Kotb and Goldin (2019) find that mandating employers to inform employees about their potential eligibility of the EITC had no effect on EITC take-up. Linos et al. (2020) find that a variety of informational nudges sent by both public and non-profit messengers to over one million potentially eligible California households encouraging them to take-up the EITC had no effect on tax filing or EITC claiming. Goldin et al. (2021) find that a single letter from the IRS sent to non-filers increased tax filing by 0.7 percentage points.

There is limited evidence on the effect of simplifying the tax filing process on EITC claiming. One exception is Kopczuk and Pop-Eleches (2007) who find that the introduction of electronic tax filing increased EITC claims. Goldin and Liscow (2018) argues that efforts to increase take-up should focus on the complexity of tax filing rather than on awareness.

3 Data

We draw on data from several California administrative data systems to identify the family structure, income, and tax filing status of households that participated in the CalFresh program (the state's version of the Supplemental Nutrition Assistance Program, or SNAP, formerly known as food stamps) in 2017.

Our first database contains individual- and case-level CalFresh enrollment records. In 2017, 5.3 million unique individuals enrolled in CalFresh, representing 48.4 million person-months.[6] In 2017, 49.2% of individuals ever enrolled in CalFresh were enrolled for all 12 months, while the other 51.8% were enrolled for only part of the year. CalFresh cases are defined as groups of people who prepare and eat meals together. Our sample includes 2.7 million distinct cases. An individual can participate in different cases over the course of a year, though 95% of individuals in our sample are associated with only one case over the course of the year. We link adults' CalFresh enrollment records with their quarterly wage records filed with the Employment Development Department (EDD), which administers the state's unemployment insurance program. This allows us to measure wage and salary earnings (though not cash or self-employment earnings) for all adult CalFresh participants.

As mentioned above, the CalFresh records contain information key to determining EITC eligibility. First, we observe the date of birth for each CalFresh recipient. Second, CalFresh enrollment records provide monthly snapshots of household composition. We use this information to identify which CalFresh-enrolled dependents on tax returns, as well as non-filing children, would pass the EITC's age and residency tests, and to group non-filers into likely tax units according to their ages and common CalFresh cases.

Our second database consists of all California resident personal income tax returns filed with the California Franchise Tax Board for tax year 2017. Roughly 33 million of the 39 million individuals residing in California in 2017 are represented on these returns. This dataset contains all tax returns that could have claimed the CalEITC. For the 87% of California filers who submit their returns electronically (known as "e-filing"), we can associate their state returns with their federal returns. Given the importance of federal tax information to our analysis, all statistics presented below are restricted to this population of e-filers, for whom we can observe a federal return.[7]

For this project, we completed the first-ever individual linkage between these two datasets. We used a fuzzy linking algorithm, identifying both exact and near matches on names, birth dates, and Social Security numbers.[8]

[6] We tally CalFresh enrollment using a dataset of individual-level enrollment records compiled by the California Department of Social Services. These data originate in a data system maintained by the California Department of Health Care Services to determine and track Medicaid eligibility. Based on guidance from staff at CDSS, we identify an individual as enrolled in CalFresh if s/he is recorded as enrolled in both data systems. The first version of this paper, published in July 2021, did not apply this restriction, leading us to possibly overstate CalFresh enrollment and the number of eligible non-claimants in this population.

[7] Since paper filers have lower average incomes and are more likely to self-file, we anticipate that excluding paper filers from our analysis will bias our estimate of non-claiming downwards. We explore this issue further in Appendix B.

[8] The datasets used for the linkage were more expansive than those used for our analysis: All tax returns from tax years 2015-2018, and all individuals enrolled between 2012-2019 in any safety net programs administered by the California Department of Social Services (CDSS). The inclusion of additional years and observations helps us avoid false positive matches in our focal samples. The research team never had access to personally identifying information from either dataset; we used a hashed linking algorithm to identify exact and near matches using hashed

On the FTB side, our linkage is limited to individuals who appear on a state tax return. The absence of federal-only filers does not bias our estimates of CalEITC participation; CalFresh recipients who do not match to a state tax return cannot have claimed the state EITC. However, we cannot confidently estimate take-up of the federal EITC among CalFresh recipients who do not appear on a state tax return, because some of these individuals may have appeared on a federal return and may have claimed the federal EITC. In Section 5, we present take-up estimates for the federal EITC among the narrower population of e-filed tax returns containing a CalFresh-enrolled head or spouse, for whom we can observe a federal return, a federal EITC claim if there is one, and whether dependents are likely qualifying children.

Table 1 presents summary statistics for the FTB tax filer population. Overall, 17% of tax returns include a claim for the federal EITC, and 9% include a claim for the CalEITC. We also present each statistic broken out by number of dependents, as the presence of children is a key criterion for EITC eligibility. Both federal and state EITC claiming shares are larger for returns with dependents than for those without.

Table 1: Summary statistics for tax filer sample

	By number of dependents				
	0	**1**	**2**	**3+**	**Total**
Count of Individuals	13,020,867	6,570,080	7,591,456	6,623,982	33,806,385
Count of Tax Units	10,310,603	2,802,932	2,171,327	1,340,842	16,625,704
Count of Tax Units That E-Filed	8,888,248	2,489,056	1,943,459	1,200,415	14,521,178
% E-Filed	86%	89%	90%	90%	87%
Statistics for e-filers (tax unit level)					
Filing Status					
Single	72%	11%	7%	6%	47%
Married Filing Jointly	26%	42%	61%	64%	37%
Married Filing Separately	1%	1%	1%	0%	1%
Head of Household	0%	45%	32%	29%	15%
EITC Claiming					
Fraction claiming Federal EITC	7%	34%	33%	32%	17%
Mean Federal EITC claim (if positive)	$335	$2,315	$3,514	$3,836	$2,334
Fraction claiming CalEITC	7%	17%	13%	10%	9%
Mean CalEITC claim (if positive)	$76	$290	$527	$573	$266
Fraction claiming either EITC	8%	34%	33%	32%	17%
Mean total EITC (if positive)	$394	$2,452	$3,719	$4,014	$2,462
Income					
Mean earnings	$49,370	$73,348	$101,551	$83,189	$63,259
Mean AGI	$76,431	$90,430	$126,030	$106,207	$87,930

Notes. Universe is all state tax returns in tax year 2017. Other than row 1, all statistics are at the tax-unit level.

Table 2 presents statistics for the CalFresh population. We distinguish between three groups of CalFresh cases: those for which all members appear on tax returns, providing complete information about tax unit structure; those for which some but not all members appear on returns, providing partial information; and those for which none of the members appear on a return. Overall, of the 5.6 million individuals receiving CalFresh in 2017, 3.1 million (or

identifiers.

66% of CalFresh recipients) appeared on a state tax return. Of these, 1.3 million were heads or spouses and 2.4 million were dependents. Hereafter, we refer to heads and spouses collectively as *primary filers*.

Table 2: Summary statistics for CalFresh sample

	By case-level filing status			
	Everyone files	**Some file**	**No one files**	**Total**
Number of individuals	2,895,771	1,007,485	1,404,030	5,307,286
Row percent	55%	19%	26%	100%
Mean case size				
Total	2.2	3.6	1.4	2.0
Adults	1.1	1.4	1.0	1.1
Children	1.1	2.2	0.4	1.0
Share of adults w. linked wages	67%	52%	28%	51%
Mean total EDD wages (if positive)	$15,614	$13,315	$8,944	$13,932
Tax return linkage				
Linked to tax return	100%	49%	0%	56%
Linked to return with federal EITC claim	58%	33%	0%	33%
Mean federal EITC claim amount (if positive)	$3,047	$3,735	.	$3,122
Linked to return with CalEITC claim	38%	21%	0%	21%
Mean CalEITC claim amount (if positive)	$383	$515	.	$397

Notes. Universe is all CalFresh recipients in tax year 2017. Statistics are reported at the reference-case level; see Section 4.3 for more information about how these are defined. Column 1 reports means for cases in which every member appears on a 2017 state tax return. Column 2 reports means for cases in which at least one member, but not all, appear on a 2017 state tax return. Column 3 reports means for cases in which no members appear on a 2017 tax return.

We exclude from our analysis the approximately 1 million dependents who were claimed on a tax return where no primary filer was a CalFresh recipient. Since we do not observe these primary filers' case files or earnings records in our CalFresh data, we are unable to verify their eligibility for the CalEITC. We also choose not to associate these dependents with other adults who appear on their CalFresh cases, since they were already claimed by the primary filer(s) on their tax return. See Appendix F for a greater discussion of how SNAP cases misalign with tax units, and how many children are claimed by adults other than those who appear in their SNAP cases.

Our final sample includes 2.4 million CalFresh recipients who appear on a state tax return in 2017. Of all CalFresh recipients who appear on a state tax return in 2017, 51% (38% of all recipients) include claims for the federal EITC and 34% (24% of the total) include claims for the state credit. Of all 1,035,623 tax units in 2017 containing a CalFresh-enrolled primary filer, 71% claimed the federal EITC and 47% claimed the CalEITC (Table 3). Their average federal and state EITC claims were $3,109 and $402, respectively.

The Californians who enroll in CalFresh are a subset of the population eligible for the CalEITC. An obvious question is to what extent the CalFresh population is representative of that broader low-income population in the state. In Appendix A, we use ACS data to compare a variety of demographic characteristics of households who report receiving SNAP to non-recipients and the wider low-income population in California. We conclude that families in households with at least one SNAP recipient are more likely to be eligible for and to claim the CalEITC due to their lower incomes, larger household sizes, and willingness to interact with government programs. This suggests that our estimates might overstate eligibility, but understate the take-up gap among all Californians.

Table 3: Claiming of federal and state EITC among all CalFresh tax units

	Total tax units	Fed EITC Claimants			CalEITC Claimants		
	Count	Count	Share	Amount	Count	Share	Amount
For all filers							
Total	1,035,623	732,832	71%	$3109	485,562	47%	$402
By filing status and number of dependents							
Single							
No dependents	314,584	105,804	34%	$340	101,538	32%	$88
1 dependent	224,343	205,409	92%	$2,673	153,468	68%	$357
2 dependents	166,266	159,059	96%	$4,142	105,799	64%	$629
3+ dependents	99,134	94,597	95%	$4,613	53,862	54%	$671
Married							
No dependents	37,928	15,409	41%	$393	10,842	29%	$80
1 dependent	46,045	36,028	78%	$2,658	18,940	41%	$287
2 dependents	66,506	53,869	81%	$4,121	21,997	33%	$524
3+ dependents	80,817	62,657	78%	$4,382	19,116	24%	$555
By total earnings (thousands)							
$0-$5	144,482	85,170	59%	$1,016	78,084	54%	$481
$5-$10	179,489	137,680	77%	$2,148	130,975	73%	$743
$10-$15	211,213	174,030	82%	$3,620	159,837	76%	$304
$15-$20	158,678	111,553	70%	$4,675	87,220	55%	$122
$20-$25	116,237	83,079	71%	$4,213	29,240	25%	$31
$25-$30	78,715	59,202	75%	$3,525	100	0%	$374
$30+	146,809	82,118	56%	$2,268	106	0%	$444

Notes. Universe is e-filed tax returns with at least one primary filer who claimed CalFresh. Column 1 reports the total number of tax units in each cell. Column 2 reports the count of those tax units that claimed the federal EITC. Column 3 reports what share of all returns in each cell claimed the federal EITC. Column 4 reports the average claimed amount of the federal EITC for each cell. Column 5 through 7 report the same statistics but for the state EITC.

4 Eligibility and participation

Our primary methodological challenge is to measure EITC eligibility for an individual who does not claim the credit. This entails measuring three components of eligibility: the composition of the filing unit on which the individual would appear if he/she filed a return; the unit's AGI and total earned income, and the unit's number of EITC qualifying children (QC).

We measure the extent of eligibility and non-claiming across three distinct populations for whom we have different information about these three components: (1) CalFresh participants who appear on a state return and claim the federal EITC; (2) CalFresh participants who appear on a state return but do not claim the federal EITC; and (3) CalFresh participants who do not appear on a state return.

For those who claim the federal EITC, we can observe nearly all relevant income,[9] and we can infer their number of qualifying children using only information from the tax unit's own return. For those who file but do not claim the federal EITC, we supplement tax unit composition, as well as income and earnings reported on their return, with CalFresh records. Specifically, we use the CalFresh casefiles to test whether dependents pass the qualifying

[9]The only relevant income sources we do not observe are interest and dividend income, which CalFresh-enrolled filers rarely have.

children's age and residency test. For those who do not file a state return, we transform CalFresh cases into likely tax units and measure eligibility using the casefiles, recipients' ages, and merged earnings records. Table 4 provides an overview of the information we use to estimate each component of eligibility across these three populations.

In the following three sections, we describe these processes more fully and present estimates of non-claiming for each population.

Table 4: Data sources for measuring three components of eligibility across three populations

	Eligibility components		
	Filing Status	**Earned income**	**Qualifying children**
Filers			
Fed EITC claimants	Filing status (540, Lines 1-5)	CA wages (540, Line 12) AGI (540, Line 13) Investment (1040, Lines 13, 14, 17) Self-employment (1040, Lines 12, 17, 27)	CA wages (540, Line 12) AGI (540, Line 13) Fed EITC amt (1040, Line 66a) Fed EITC amt (Sch 3514, Line 3)
Non Fed EITC claimants	Filing status (540, Lines 1-5)	CA wages (540, Line 12) AGI (540, Line 13) Investment (1040, Lines 13, 14, 17) Self-employment (1040, Lines 12, 17, 27)	CA wages (540, Line 12) AGI (540, Line 13) # of deps (540, Line 10) CalFresh casefiles
Non-Filers			
	CalFresh casefiles Participants' ages	EDD wages	CalFresh casefiles Participants' ages

Notes. Table 4 summarizes the data sources we use to measure the three components of eligibility across our three populations. Forms and line numbers are applicable to 2017 returns. "CalFresh casefiles" refers to our ability to observe individuals sharing CalFresh cases with each other for a certain number of months in the tax year.

4.1 Among filers who claim the federal EITC

For those who file a state return and claim the federal EITC, we observe nearly all of the information needed to simulate eligibility for the CalEITC, including AGI, earned income, and filing status, directly from the filers' state and federal returns.

4.1.1 Qualifying children

The only variable necessary to identify CalEITC eligibility that we do not observe directly is each unit's number of qualifying children (QCs). As discussed above, a qualifying child for the purposes of the EITC are dependents under the age of 19 (or 24 if the dependent is a full-time student or any age if the dependent has a total or permanent disability) who reside with a primary filer for at least six months in the tax year and are near relations or adopted or foster child of the filer. Federal EITC claimants do report their count of QCs on the Schedule EIC, but we do not observe this information in our copies of filers' federal returns. However, since there is a unique number of QCs that can rationalize the unit's federal EITC claim given their filing status, AGI, and earned income, we can use these variables from the unit's return to infer their number of claimed QCs. In Table 5, we compare the number of dependents claimed on the tax return to the count of qualifying children inferred from the federal EITC amount and reported earned income. In the vast majority of cases (97.6.%), the number of QCs from the federal EITC claim and the number of dependents are the same. In just 1.5% of cases, the number of QCs is smaller than the number of dependents on the tax return, and in less than 1% of cases the number of QCs is larger than the number of dependents.

Table 5: Relationship between the number of dependents and number of qualifying children among CalFresh households claiming the federal EITC

	Dependents claimed on tax return			
	0	**1**	**2**	**3+**
Number of EITC qualifying children				
0	98.9%	2.1%	0.4%	0.1%
1	0.8%	97.4%	7.7%	1.6%
2	0.2%	0.4%	91.7%	8.7%
3+	0.1%	0.1%	0.3%	89.6%
N	118,820	237,766	209,638	155,130

Notes. Universe is all tax units that e-filed their returns, included at least one primary filer who claimed CalFresh, and had a positive federal EITC claim. Cells represent column percentages.

We present two pieces of evidence that help to demonstrate that this inference yields trustworthy estimates of each unit's number of qualifying children. First, we compare the count of QCs *inferred* from the federal claim amount to the *actual* number of QCs reported by tax units who claimed the CalEITC. [10] For nearly all units, the inferred number of QCs exactly matches the number reported on the Schedule 3514. Second, we compare actual CalEITC claims to predicted claim amounts, using the process described here, for the subset of federal EITC claimers who also claimed the CalEITC. We correctly predict the exact CalEITC claim amount for over 97% of CalEITC claimants. See Appendix E for more information about this analysis.

4.1.2 Results

With each unit's inferred number of qualifying children, plus their filing status and reported California earned income, we can identify which appeared eligible for the CalEITC in tax year 2017. Of the 721,354 e-filed state tax returns containing a primary filer who enrolled in CalFresh, and included a federal EITC claim, 71% were eligible for the CalEITC (Table 6). Of these eligible units, 8%, or 40,366, did not claim the CalEITC. These tax units received an average federal EITC benefit of $3,794, but did not claim an average additional benefit of $235 from the CalEITC. The forgone CalEITC amount for this group totaled nearly $9.5 million, and if received, would have raised annual incomes in this population by 1.7%.[11]

Table 6 also summarizes eligibility and take-up of the CalEITC among federal EITC claimants by tax unit type and income levels. A greater share of eligible households with dependents (8-10%), than those without (5%), failed to claim the CalEITC. Eligible households with no qualifying children forwent an average state credit of $65. Households with qualifying children forwent an average of $264.

The share of eligible households not claiming increases with total earnings: Only 5% of eligible participants with $5,000 to $10,000 in annual earnings did not claim the credit (average credit amount of $579), while 16% of participants in the $20,000-$25,000 income bracket did not claim the credit (average unclaimed credit amount of $28). This difference may be due to eligibility being less certain and the expected return being lower. While take-up was higher in the lower-income categories, there were still many tax units who missed out on a credit for which they

[10] We can observe the actual number of qualifying children claimed by these units on the form they submit to claim the state EITC (Schedule 3514). As noted above, we cannot use these counts for all tax units in this population, because we only observe them for those who claimed the CalEITC.

[11] We also identify a small number of tax units (0.05% of all CalFresh units who claim the federal EITC) who do not appear to be eligible for the CalEITC, based on their reported state earnings and federal qualifying children, but who nevertheless claim it. This could indicate an incorrect claim but could also reflect inaccuracies in our simulation. The apparently ineligible claimants have relatively high earnings and an average CalEITC claim of $283.

were eligible, and the average amounts at stake were substantially larger. Households with incomes below $10,000 and who failed to claim the CalEITC left an estimated $400 to $600 on the table.

Table 6: Eligibility and take-up of the CalEITC among CalFresh filers who claim the federal EITC

	Total tax units	**CalEITC eligible**		**Eligible non-claimants**		
	Count	Count	Share	Count	Share	Amount
For all filers						
Total	721,354	514,469	71%	40,366	8%	$235
By filing status and number of qualifying children						
Single						
0 QCs	107,181	106,242	99%	5,198	5%	$64
1 QC	213,410	169,477	79%	13,641	8%	$223
2 QCs	153,180	110,195	72%	9,799	9%	$327
3+ QCs	82,232	52,613	64%	5,215	10%	$311
Married						
0 QCs	16,121	11,775	73%	596	5%	$71
1 QCs	37,830	21,062	56%	1,775	8%	$152
2 QCs	53,719	23,547	44%	2,182	9%	$248
3+ QCs	57,681	19,558	34%	1,960	10%	$229
By total earnings (thousands)						
$0-$5	80,007	75,524	94%	2,490	3%	$424
$5-$10	135,092	133,955	99%	6,315	5%	$579
$10-$15	172,809	169,616	98%	11,598	7%	$250
$15-$20	110,678	100,740	91%	14,255	14%	$122
$20-$25	82,424	34,634	42%	5,708	16%	$28
$25-$30	58,817	0	0%	0	.	.
$30+	81,527	0	0%	0	.	.

Notes. Universe is e-filed tax returns linked to at least one primary filer who claimed CalFresh, and that included a claim for a non-zero federal EITC. The number of qualifying children in each tax unit was calculated using the process described in Section 4.1.1. Column 1 reports the total number of tax units that meet those criteria. Column 2 reports the count of those tax units that were eligible for the California EITC. Column 3 reports what share of all returns were eligible for the CalEITC. Column 5 reports the number of eligible returns that did not claim the CalEITC, and Column 6 reports the share of eligible units that did not claim. Column 7 reports the average imputed amount among those non-claimers for each cell.

4.2 Among filers who do not claim the federal EITC

Next, we estimate eligibility among CalFresh-participating tax units who filed a state return but did not claim the federal EITC. As in the previous section, we observe each tax unit's filing status and relevant earned income. However, we do not observe these tax units' number of EITC qualifying children directly, nor can we infer the number from a federal EITC claim. Instead, we impute this number for each unit using other information available on the tax unit's return and in the CDSS records. We describe that process below.

4.2.1 Qualifying children

Beginning with the list of dependents on the return, we interrogate which dependents might be a qualifying child for purposes of the EITC.[12] As mentioned above, QCs must be under 19 (under 24 for full-time students or any age for dependents with a total or permanent disability), must reside with a primary filer for at least six months of the year, must have a valid Social Security number (SSN), and the child can be claimed as a QC on only one return. We observe some of this information in our CalFresh records, meaning we can verify which dependents might in fact be qualifying children.

We are unable to identify students or observe disability status, so we only allow dependents age 18 and under to be potential QCs, provided that they satisfy the residency test.[13] We are able to observe whether a dependent has an SSN or an Individual Tax Identification Number (ITIN), and we use this to disqualify children without a valid SSN. See Appendix C for more information about ITIN filers.[14]

To simulate the residency test, we observe the number of months in which dependents appeared on the same CalFresh case as a given primary filer on their return.[15] For tax units with single or head of household filing status, or where both members of a married couple appear in the CalFresh records, we use the number of months a child appeared on the same CalFresh case with a primary filer as a proxy for residential arrangements. A dependent who shared a CalFresh case with the a primary filer for at least six months is counted as meeting the residency test. If the child and parent were on CalFresh for only part of the year, we assume they reside together in months in which neither was enrolled and we allow these months to count toward this six-month threshold. For married tax units in which only one member appears in the CalFresh data, we cannot track the residency of the non-CalFresh spouse, so we assume the child lives with that spouse and meets the residency test. Similarly, we assume that dependents who were not enrolled in CalFresh are also qualifying children, since we cannot test whether they fail the residency test.[16] The effect of these rules is to only disqualify dependents as possible QCs if we can observe that they were in different CalFresh cases than any primary filer on their return for more than six months in in the tax year.[17]

Table 7 compares the number of imputed EITC qualifying children to the number of dependents claimed on the tax return among units that did not claim the Federal EITC. Since our imputation method begins with the number of dependents and removes those who do not appear to be qualifying children, the number of QCs is never greater than the number of dependents. We find that the number of qualifying children equals the number of dependents for 56% of tax units with at least one dependent. For most units in which those are not equal, we assign just one fewer QC than the number of dependents. Of the 42,227 dependents who we deem not to be qualifying children, 27,036 (18.6% of all dependents) fail the residency test, 17,913 (12.3% of all dependents) fail the age test, and 10,187 (7.0% of all dependents) fail the SSN test. A small number of dependents fail more than one test.

To assess the accuracy of our imputation method, we rerun our CalFresh-based imputation process on the set

[12] Although in principle a parent may have a QC who does not qualify as a dependent — for example, a child may be the dependent of a non-custodial parent who provides substantial child support but would be claimable as a QC by the custodial parent – we expect that this is rare and we do not attempt to model it. This rare circumstance might explain the few cases in which there are more inferred QCs in a tax unit than observed dependents.

[13] Of the 145,783 dependents under age 24 on a return with a CalFresh-enrolled head or spouse and without a federal EITC claim, 5,691 (3.9%) are between the ages of 19 and 24.

[14] In some cases, returns claim more dependents than the number of SSNs we observe. Since we cannot link these additional unlisted dependents to CalFresh participants, we do not allow them to count as QCs.

[15] Recall that we do not perform this exercise over the CalFresh-enrolled dependents claimed by a primary filer who was not enrolled in CalFresh in 2017.

[16] If a dependent was enrolled in Medicaid, in any program administered by CDSS, or in CalFresh in any other year between 2005 and 2020, we are able to observe their age.

[17] Ultimately, data limitations introduce a trade-off in deciding which dependents might be qualifying children. We acknowledge that by allowing months in which a dependent is not enrolled in CalFresh to count towards the residency test, we take the less conservative approach. In Appendix Table 8, we report the same comparison as in Appendix Table 7, except that we do not allow dependents who did not enroll in CalFresh to be qualifying children. This results in our disqualifying many dependents who are claimed as qualifying children. Though allowing non-CalFresh dependents to be QCs risks overstating EITC eligibility, restricting our analysis only to dependents observed in CalFresh seems to severely understate the number of actual QCs. All that said, these assumptions do not meaningfully affect our results, because, as discussed below, most eligible tax units who fail to claim both the federal and state EITC are single adults without children, for whom this imputation does not matter.

of tax units that claimed both the federal EITC and state EITC, and for whom we can observe actual number of qualifying children. We then compare each tax unit's estimated number of qualifying children according to our CalFresh records against the number reported on those unit's 3514 (Appendix Table 7). In three-quarters of the cases, our inferred number matches the reported number exactly. When our imputation errs, it most often does so by underestimating the number of QCs, and therefore the family's EITC eligibility and/or credit.

Table 7: Relationship between number of dependents and imputed number of qualifying children among CalFresh filers that did not claim federal EITC

	Dependents claimed on tax return			
	0	**1**	**2**	**3+**
Number of EITC qualifying children				
0	100.0%	51.8%	21.3%	8.9%
1	0.0%	48.2%	18.7%	8.1%
2	0.0%	0.0%	59.9%	17.9%
3+	0.0%	0.0%	0.0%	65.2%
N	263,865	32,597	22,267	25,659

Notes. Universe is all tax units that e-filed their returns, included at least one primary filer who claimed CalFresh, and did not claim the federal EITC. Cells represent column percentages.

4.2.2 Results

Confident in the imputation method for determining EITC qualifying children, we return to estimating eligibility for and take-up of the CalEITC for this population. Table 8 summarizes our results. Over 300,000 tax units contain a primary filer who was enrolled in CalFresh in 2017 and did not claim the federal EITC. Of these units, just over 100,000 (33%) were eligible for the CalEITC, and among these eligible units, nearly 98,000 (97%) did not claim it. Very few eligible households claimed the state EITC but not the federal EITC. The average forgone CalEITC amount for this group was $85, and totaled over $8.3 million. If received, these benefits would have raised annual incomes in this population by 1.1%.

The vast majority of eligible non-claimants were single filers without qualifying children. This fact help explains why these units did not claim the more valuable federal EITC; they were likely eligible, but only for a small credit amount. This fact also helps address any concerns about how we address qualifying children for these units; because it is such a small population, it has a small effect on our overall participation rates.

Though there are far fewer eligible non-claimants among tax units with QCs (just short of 1,500), their average forgone credit was much higher. Participation rates tended not to vary too significantly with either tax unit composition or total tax unit earnings. Take-up ranged between 2 and 9%. Again, since most non-claimants were adults without QCs, the average forgone credit was fairly low regardless of income level.

4.3 Among non-filers

Lastly, we turn to our third population: CalFresh participants who did not appear on a 2017 California tax return. For these participants, we construct simulated tax units from CalFresh casefiles and simulate eligibility using those

Table 8: Eligibility and take-up of the CalEITC among CalFresh filers who did not claim the federal EITC

	Total tax units	CalEITC eligible		Eligible non-claimants		
	Count	Count	Share	Count	Share	Amount
For all filers						
Total	302,791	100,881	33%	97,999	97%	$85
By filing status and number of qualifying children						
Single						
0 QCs	223,357	96,644	43%	94,385	98%	$82
1 QC	9,480	1,292	14%	991	77%	$203
2 QCs	4,183	258	6%	153	59%	$366
3+ QCs	1,752	70	4%	51	73%	$439
Married						
0 QCs	29,170	2,317	8%	2,196	95%	$73
1 QC	10,035	133	1%	110	83%	$291
2 QCs	11,778	103	1%	72	70%	$493
3+ QCs	7,588	36	0%	23	64%	$710
By total earnings (thousands)						
$0-5	59,312	32,563	55%	31,568	97%	$140
$5-10	41,809	37,916	91%	36,990	98%	$81
$10-15	37,183	29,414	79%	28,585	97%	$29
$15-20	47,125	701	1%	594	85%	$84
$20-25	33,158	287	1%	262	91%	$23
$25-30	19,513	0	0%	0	.	.
$30+	64,691	0	0%	0	.	.

Notes. Universe is e-filed tax returns linked to at least one primary filer who claimed CalFresh, and that did not include a federal EITC claim. The number of qualifying children in each tax unit was calculated using the process described in Section 4.2.1. Column 1 reports the total number of tax units that meet those criteria. Column 2 reports the count of those tax units that were eligible for the California EITC. Column 3 reports what share of all returns were eligible for the CalEITC. Column 5 reports the number of eligible returns that did not claim the CalEITC, and Column 6 reports the share of eligible units that did not claim. Column 7 reports the average imputed amount among those non-claimers for each cell.

casefiles and linked earnings records.

This is the most complex part of our analysis. CalFresh cases represent groups of individuals who eat and prepare meals together, while tax units generally reflect immediate families. These two types of households might not coincide. For example, there may be individuals in the CalFresh case (e.g., extended family or unrelated roommates) who are not part of the tax unit, and individuals in a tax unit (e.g., dependents of non-custodial parents) who do not share a CalFresh case. We do not observe household relationships, which would allow us to associate filers with their spouses and children and vice versa.[18] That said, measuring eligibility among non-filers (or non-participants across a range of public programs) is not a new challenge, and our administrative data boasts several advantages over commonly used survey data, as discussed in Section 2. See Appendix F for a fuller discussion of the misalignment between households observed in our case records and those observed in our tax records.

The following section details how we construct these simulated filing units, that is, groups of CalFresh recipients who would likely appear on a tax return together if a return had been filed. We proceed in several steps: We disambiguate CalFresh cases, classify individuals as primary filers or dependents, identify married couples, assign dependents to filers, test which of those dependents might be a qualifying child, and construct a measure of earned

[18] Using merged IRS and SNAP data from Florida, Maag et al. (2015) show that, among federal EITC claimants, 99% of all claimed qualifying children pass the relationship test and 77% of qualifying children appear to pass the residency test. These findings suggest that the relationship test may be less of a concern for determining EITC eligibility within this population.

income.

For each step, we discuss how accurate our simulation is likely to be, We measures accuracy by applying constructing simulated tax units among CalFresh recipients who do appear on on a state return, and then comparing the results of our simulation to the composition of, and income reported on, those observation's actual returns. Table 9 summarizes our conclusions about the accuracy of each step in our simulation process.

Table 9: Accuracy of tax unit imputation process for tax filing CalFresh recipients

Step in Imputation Process	Accuracy Rate	Definition of Accuracy
Assigning primary filers and dependents		
Aged 0-17	99.4%	Percent of CalFresh recipients who filed a tax return and who were used for cross-validation who were correctly assigned as either a primary filer or dependent.
Aged 18-80	92.3%	
Aged 81+	84.6%	
Filing status		
Correctly set as single	89.0%	Percent of CalFresh recipients who were primary filers on a state tax return who were correctly assigned to their appropriate filing status.
Correctly set as married	71.2%	
Correctly set overall	83.4%	
Assigning dependents to any imputed tax unit		
Non-filers	31.4%	Percent of CalFresh-enrolled dependents who are assigned to an imputed tax return.
Filers	59.4%	
Assigning dependents to their primary filer		
All primary filers	74.5%	Percent of assigned CalFresh-enrolled dependents matched with an adult on their tax return.
At least one primary filer	84.4%	
Qualifying children		
0 QC	69.4%	Percent of state tax returns with the correct number of imputed qualifying children.
1 QC	76.5%	
2 QC	79.9%	
3+ QC	93.0%	
Total	74.3%	
Earned income (wages)		
Perfect match	14.7%	Percent of state tax returns whose imputed earnings are within X dollars of the wages reported on their 540.
Within $200	64.8%	
Within $500	68.9%	
Within $1,000	74.1%	
Within $10,000	94.1%	
Earned income (AGI)		
Perfect match	2.8%	Percent of state tax returns whose imputed earnings are within X dollars of the AGI reported on their 540.
Within $200	36.1%	
Within $500	40.9%	
Within $1,000	56.4%	
Within $10,000	77.7%	

Notes. Table 9 summarizes the accuracy of our process of imputing the structure and income of tax returns. For more detailed descriptions of the measures of accuracy references in this table, please refer to Section 4.3.

4.3.1 Reference cases

First, we assign each CalFresh recipient to a single representative CalFresh case. For the 95% of individuals who appear on only a single CalFresh case in 2017, this assignment is straightforward. For the remaining 5%, we assign individuals to the case they appeared on most frequently. In the rare event of a tie, we use the most recent case. Hereafter, we refer to CalFresh cases to which recipients are assigned as the their *reference case*.

We construct simulated tax units from these reference cases. Some reference cases will include filers and non-filers. We construct one or more tax units from the non-filing members in each reference case. Some reference

cases may contain only children, who cannot file a tax return by themselves. Other reference cases may only contain non-filers, but include more individuals than could plausibly appear on a tax return together. Below, we discuss how we address these issues and transform these reference cases into likely tax units.

4.3.2 Assigning primary filer and dependent status

We assign each non-filing individual on a CalFresh reference case to be a primary filer or dependent. We assign everyone under the age of 18 or over the age of 80 to be dependents. For those aged between 18 and 80, we predict whether they should be a dependent or a primary filer using other available information, including their earnings, age, sex, race, language spoken, disability status, number of months on CalFresh, participation in other safety net programs, and whether they were observed as a primary filer on a tax return from the prior year. Specifically, we use the CalFresh observations who do file returns to train a prediction model to classify observations as likely primary filers or dependents using these variables. We use cross-validation to select a threshold predicted probability of being a primary filer that maximizes out-of-sample accuracy. See Appendix D for a full description of this process.

Applying these same rules to CalFresh recipients who do appear on a 2017 tax return, we find these procedures yield highly accurate assignments. In 2017, 99.4% of child CalFresh recipients (aged 17 or younger) and 84.6% of elderly CalFresh recipients (aged 81 years or older) who appeared on a tax return were dependents. Appendix Figure 2 shows the share of dependents by single-year-of-age among tax filers. There is no obvious break at 81, but the size of the population at that age is fairly small, making reliable imputation of an individual's tax status challenging. For those filers between 18 and 80 years old, we can assess the accuracy of the machine learning model we employ by testing how often we correctly assign individuals to the role of primary filer versus dependent. We restrict our focus to tax filers not used to create the machine learning model but who are used to find the optimal threshold predicted probability. We are able to correctly predict the status of 92.3% of individuals in this sample.

4.3.3 Filing status

We then predict which recipients assigned to be primary filers would file as a single adult or married couple. This step is only relevant for the subset of reference cases with multiple adults predicted to be a primary filer, as those adults who are alone on a reference case must be single filers. To identify likely married couples, we first look to tax returns from the prior tax year. If two individuals on the same 2017 CalFresh case filed as married filing jointly in tax year 2016, we assume they are still married and would still file as part of the same tax unit. For those remaining adults who do not appear on a tax year 2016 return, we use the relative age of each adult to decide whether the pair is likely a married couple. We marry two individuals in a reference case if they are each older than the 10th percentile of ages among married filers and if the age difference between them is between the 10th and 90th percentiles of within-couple age-differences (also among filers).[19] When there are multiple pairings that would satisfy this rule, we pair the adults who are closest in age. The effect of these rules is that anyone who can be paired to another reference case member of a plausible age is assumed to use a "married" filing status. Anyone who remains unpaired is assumed to use a "single" filing status. Overall, we impute that 13% of the primary filers in our non-filing sample are married.

Applied to the population of CalFresh-enrolled primary filers, these rules correctly assign 89.0% of single primary filers to the marital status of single on their imputed tax return. Among the remaining 11% of single CalFresh-enrolled primary filers, we incorrectly predicted 7% of them would be dependents, and the remainder would file as part of a married return. Our assignment rules performs worse in predicting married returns: only 71% of married filing jointly filers are assigned to be a primary filer of married imputed return. Most incorrectly-assigned married

[19]The 10th percentile of age among married primary filers is 26 and 27 for women and men, respectively. For women, the 10th and 90th percentile of within-couple age differences is 6 and 11 years, meaning women are permitted to marry individuals in their reference case who are 6 years younger or 11 years older than them. The percentiles for men are the inverse. Men are permitted to marry individuals in their reference case who are 11 years younger and 6 years older than them.

primary filers (27%) are set as single, with only 2% set as dependents. In 81% of cases where we incorrectly label a married primary filer as single we do not observe their spouse in the CalFresh records. If we assume there is a similar share of mixed-CalFresh status married couples among non-filing CalFresh recipients, then we should expect a similar under-count of married couples there. Under-counting married units likely biases our measure of EITC eligibility upwards, since some number of simulated eligible single filers might become ineligible if we combined their earnings with a spouse. It is possible that eligibility for the EITC could increase, but that would likely be a function of the qualifying children that are connected to the other adult, and not the marriage itself.

4.3.4 Assigning dependents

The next step is to assign dependents to imputed tax units. We only consider as candidate tax units those that contain an adult in the dependent's CalFresh reference case and contain an adult with whom the dependent might have resided for at least six months in the tax year. To measure how long adults and children resided together, we count the number of unique months they shared a case in the calendar year, aggregate across multiple cases where needed, and also count months in which the dependent was not enrolled. The effect of this restriction is to rule out candidate tax units containing only adults with whom we can confidently infer from our CalFresh records the child did not reside for more than half the year.

When there is only one such tax unit, the assignment is straightforward. When multiple candidate tax units satisfy these criteria, we assign dependents to tax units containing adults with whom the dependent appeared on the prior year's tax return. If there are still multiple candidate tax units, we assign child dependents to tax units with adults who are at least least 16 years older than the child, meaning they could plausibly be the child's parent. If there are still multiple such units, we assign child dependents to the unit with the youngest plausible parent, and in the event of further ties, to the unit with the highest earnings. In the case of adult dependents with multiple candidate tax units, we assign them to the unit with the highest earnings.

We are unable to assign imputed dependents to any imputed tax unit if their reference case lacks a plausible adult primary filer. This is not a trivial issue. Since these children could be claimed on some return and, if claimed, could deliver more significant refunds to EITC-eligible filers, failing to assign these children to any imputed tax unit is major limitation of our approach. Among non-filers, we are unable to assign 669,681 imputed dependents (out of a total of 986,358) to any imputed tax unit. Of these unassigned dependents, 58% are adults (who we predict would be adult dependents) and the remaining 42% are children.[20] For 99.9% of the unassigned dependents who are children, there are no potential tax filers on their reference case, and for 93%, there are no adults. The expectation is that many of these "child-only" reference cases represent households in which parents or guardians are not eligible for CalFresh (or either the CalEITC or the Federal EITC) due to their immigration status.[21]

To test these hypotheses, We study the individuals who appeared on a reference case without a plausible adult filer according to our rules, but who were actually claimed on a state return. Of the more than 2.2 million individuals assigned to be dependents, we are unable to assign 894,379 dependents (698,840 of which are children) to an imputed tax unit – again, because 90% if these individuals are on reference cases without a likely primary filer. Over 92% of these unassigned dependents appear on tax returns in which no primary filer enrolled in CalFresh. Of the dependent children on reference cases without a likely primary filer, 42% appear on a tax return with at least one individual with an ITIN. In other words, as expected, a large share of child-only reference cases represent households in which an adult is ineligible for CalFresh because they do not have a valid SSN. In Appendix F, we report characteristics of the children from child-only reference cases who are not claimed by a filer with an ITIN. [22]

[20] A reasonable concern is whether our process of determining if CalFresh recipients are primary filers or dependents should have labeled these lone adult dependents as single filers. However, 96% of these adults have no earnings, suggesting they would not file a return by themselves anyway. As a robustness check, we reclassify these remaining adult non-filers to be single filers and find that 12,246 would be eligible for the CalEITC (with a mean credit amount of $73).

[21] Parents of young children with ITINs became eligible for the CalEITC and certain other tax credits in 2020. We discuss this issue further in Appendix C. We also discuss how many of these children were claimed on other returns in Appendix F.

[22] Hispanic children are over-represented in both the count of unassigned children and the count of unassigned children with ITINs present on

Even when we can match a child to a non-filing adult in their reference case, it's possible that that adult is not the child's parent or guardian, or might not claim the child on their return if they were to file one. To identify how common this type of error is, we again turn to the CalFresh recipients who do appear on a state tax return. We focus on the 1,278,052 dependents who (a) appear on a CalFresh case, (b) are imputed to be dependents according to the process described in Appendix D, and (c) who are assigned to an imputed tax return. We calculate the share of these dependents who we assign an imputed tax unit containing a CalFresh-enrolled adult who appeared on that child's actual tax return. We find that our procedures assign 75% of these dependents to the correct combination of primary filers on their return (i.e., to the correct single primary filer if they appeared as a dependent on a single return or the correct pair of primary filers if they were claimed on a MFJ return). For a further 10% of these dependents, we correctly match them to one but not both of the primary filers on their tax return. We fail to match a dependent to any primary filer who claimed them in 15% of cases. Similar to the issue of accurately matching dependents to any imputed tax return, the largest errors emerge when no primary filer was enrolled CalFresh at any point in 2017. When we consider only tax units in which at least one primary filer was a CalFresh recipient, we match dependents to at least one of them 97% of the time. Overall, for households where most or all individuals claim CalFresh, we are very accurate in assigning dependents to the correct imputed tax unit. However, there are many dependents who will never be matched, as their appropriate primary filer is not present in the CalFresh data.

4.3.5 Qualifying children

We determine which of the dependents assigned to each simulated tax unit might be a qualifying child using the same procedure described in Section 4.2. Table 10 compares the number of dependents in each imputed tax unit to the number of simulated QCs. The share of dependents who are QCs according to our procedure is similar to those reported for filing non-claimers of the federal EITC (Table 7). Of the 40,828 non-filing dependents on an imputed tax unit who we deem not to be qualifying children, 19,863 (6%) fail the residency test and 22,136 (7%) fail the age test.

We described in Section 4.2.1 how we assess the accuracy of these age and residency inferences. Appendix Table 7 reports, among CalFresh-enrolled who appeared on a state return with a CalEITC claim, how many were claimed as QC according to the Form 3514 versus how many appeared to be a QC using our procedures. As discussed above, our method perfectly matches the number of QCs in 74% of tax units. In the cases where we incorrectly impute the number of tax units, we more often underestimate the number of QCs and therefore the family's EITC eligibility and likely credit amount.

4.3.6 Earned income

Within each imputed tax unit, we sum all adults' EDD wage earnings over the tax year, and we assume that this total reflects both the AGI and earned income that the tax unit would report on their return if they filed. We do not observe any other form of income, such as self-employment and investment income, for this population. This omission means we understate EITC-qualifying earnings. This might lead us to overestimate eligibility if non-filing households have both self-employment and wage earnings, and the combination pushes some households above the maximum eligible earnings limit. It is also possible, and perhaps more likely, that this omission results in our underestimating eligibility and overestimating take-up, because we assume many households with no wage earnings are ineligible, even though they might have some positive self-employment earnings which could make them eligible.

In Appendix G, we investigate this issue by replicating our analysis using data from tax year 2016, when filers

their tax returns. Overall, 65% of children enrolled in CalFresh and claimed on a return are identified as Hispanic, but Hispanic children comprise 78% of all CalFresh-enrolled children claimed on a tax return who we cannot assign to an imputed tax unit. Further, 49% of Hispanic children who are not assigned to an imputed tax unit have at least one individual with an ITIN in their actual tax unit, while the average across all other races is 9.6%. It appears more likely that we will not be able to match a Hispanic child with their correct imputed tax unit because of issues related to documentation status, relative to children of other races.

Table 10: Relationship between number of dependents and imputed number of qualifying children among CalFresh imputed tax units that did not file return

	Dependents claimed on imputed tax return			
	0	**1**	**2**	**3+**
Number of EITC qualifying children				
0	100.0%	17.9%	6.3%	6.1%
1	0.0%	82.1%	9.0%	1.1%
2	0.0%	0.0%	84.6%	6.9%
3+	0.0%	0.0%	0.0%	85.9%
N	579,456	99,427	49,456	33,763

Notes. Universe is simulated tax returns including only non-filing CalFresh participants. These tax units are constructed using the process described in Section 4.3. Cells represent column percentages.

were not permitted to claim the CalEITC using self-employment earnings. By comparing the number of claimers across the two years, we can infer how many households became newly eligible for the CalEITC thanks to their self-employment earnings, rather than ineligible due to the combination of their wage and self-employment earnings. We find that excluding self-employment income likely underestimates the number of eligible non-filers. Between 2016 and 2017, the number of eligible tax units who claimed the CalEITC and federal EITC and had earnings between \$1 and \$15,000[23] doubled, while the same pool of eligible non-filers grew at half that rate. Assuming that the filing and non-filing pool are reasonably similar in terms of the distribution of self-employment income, this implies that we are missing a number of non-filers whom we would impute to be eligible for the CalEITC if we could observe their self-employment income. We conclude that our estimates are an upper-bound on the take-up rate, at least with respect to not observing all eligible income, since observing more eligible non-filers would push down the take-up rate.

Few CalFresh recipients have investment income. According to the 2017 ACS 5-year sample, only 2% of 18-64-year-old adults with below-median income enrolled in SNAP had positive investment income, and less than 1% had investment income above the eligibility threshold (compared to 7% and 2% for non-recipients, respectively). Among actual tax-units with a CalFresh recipient, just above 1% had positive investment income in tax year 2017, and just below 1% were disqualified from receiving the EITC because their investment income was too high.

To test our assumption that the sum of EDD (or UI-covered) wages is a reliable measure of non-filers' true total earned income, we compare these sums to reported earned income on state returns among those who filed. For actual tax units containing only CalFresh-enrolled primary filers (i.e., excluding married filing jointly households with only one primary filer enrolled in CalFresh), we compare the sum of EDD wages for all primary filers to their total reported California wages and AGI. Appendix Figure 3a presents the distribution of differences between California wages and total EDD wages at the tax unit-level. The median difference is \$0, and the mean difference is \$392. Appendix Figure 3b presents the distribution of differences between AGI and total EDD wages at the tax unit-level. The median difference is \$0, and the mean difference is \$2,048.

In Table 9, we break down these distributions by the share of tax units with EDD wages within a set range of the wages or AGI reported on their state tax returns. For 14.9% of these tax returns, the sum of EDD wages perfectly matches the wages reported on the 540 return. For 65%, the EDD wages are within \$100 of the actual wages reported on the tax return. By the time we reach the \$10,000 band, we capture 94% of tax returns. When we look at the difference between EDD wages and AGI, the differences are larger. We perfectly match for only 3% of

[23] This restriction is made to exclude the effect of a simultaneous expansion of the CalEITC eligible earnings range.

tax returns, are within $1,000 for 49%, and are within $10,000 for 82% of tax units.[24]

Among the 18% of households in which the difference between AGI and total earned income is greater than $10,000, 28% report only self-employment income and have no reported EDD earnings. Our inability to measure earned income correctly for self-employed workers does not risk our overestimating EITC eligibility, since we would assume that these households have no earnings and we would classify all of them as ineligible. [25]

Table 11: Eligibility and take-up of the CalEITC among non-filers

	Total tax units	**CalEITC eligible**		**Eligible non-claimants**		
	Count	Count	Share	Count	Share	Amount
For all filers						
Total	762,102	268,113	35%	268,113	100%	$199
By filing status and number of qualifying children						
Single						
0 QCs	569,710	208,204	37%	208,204	100%	$87
1 QC	80,701	27,830	34%	27,830	100%	$489
2 QCS	38,805	12,823	33%	12,823	100%	$820
3+ QCs	23,595	7,168	30%	7,168	100%	$921
Married						
0 QCs	32,764	6,467	20%	6,467	100%	$85
1 QC	5,746	1,900	33%	1,900	100%	$436
2 QCs	5,365	1,909	36%	1,909	100%	$805
3+ QCs	5,416	1,812	33%	1,812	100%	$869
By total earnings (thousands)						
$0	425,775	0	0%	0	.	.
$1-$5	166,408	163,723	98%	163,723	100%	$182
$5-$10	60,286	59,152	98%	59,152	100%	$314
$10-$15	37,615	36,962	98%	36,962	100%	$113
$15-$20	24,627	5,712	23%	5,712	100%	$118
$20-$25	17,546	2,564	15%	2,564	100%	$29
$25-$30	10,647	0	0%	0	.	.
$30+	19,198	0	0%	0	.	.

Notes. Universe is simulated tax returns including only non-filing CalFresh participants. These tax units are constructed using the process described in Section 4.3. The number of qualifying children in each imputed tax unit was calculated using the process described in Section 4.3.5. Column 1 reports the total number of tax units that meet those criteria. Column 2 reports the count of those tax units that were eligible for the California EITC. Column 3 reports what share of all returns were eligible for the CalEITC. Column 5 reports the number of eligible returns that did not claim the CalEITC, and Column 6 reports the share of eligible units that did not claim. Column 7 reports the average imputed amount among those non-claimers for each cell.

[24] EDD wages are more likely to exceed than W2 pages is likely due to the fact that gross earnings recorded for purposes of UI includes employee contributions to employer-sponsored health insurance plans, while the net income reported reported on W2's does not.

[25] Appendix Figure 4 plots the distribution of reported self-employment income among actual CalFresh tax units. Seventeen percent of CalFresh tax units reported positive self employment income in tax year 2017. Among those with self-employment earnings, the average amount was $9,236 and the median was $8,814. 10% (or more than half of those with any self-employment earnings) reported self-employment income less than $10,000. Ninety-five percent of these tax units had self-employment income below $20,500. The two clear masses of filers in Appendix Figure 4 correspond to the kink points in the federal EITC schedule for filers with qualifying children in 2017.

4.3.7 Results

After assigning a filing status, number of qualifying children, and earned income to each simulated tax unit, we can finally test how many appear eligible for the state EITC. Table 12 summarizes our estimates of eligibility for this population. We identify over 760,000 potential tax units from the non-filing CalFresh population, and we estimate that nearly 270,000 of these households (35%) were eligible for the CalEITC but did not claim it. Their average forgone credit was $199. The total forgone credit amount for this population was $53.3 million and, if received, would have increased annual income for this population by 4%.

Table 12 also presents estimates of eligibility and participation by filing status, number of qualifying children, and earnings levels. A large majority of the non-filers would likely be single filers without QCs if they filed return. Approximately 37% of these units are eligible for the CalEITC, but estimated CalEITC amounts are fairly small – just $87 on average. The numbers are similar for married couples, although there are far fewer. For tax units with QCs, CalEITC eligibility rates and amounts were notably higher. About two-fifths of single filers with QCs were eligible for the CalEITC and failed to claim on average $643. Nearly all households with very low earnings are estimated to be eligible for the CalEITC; only imputed tax units with heads outside the eligible age range are assumed to be ineligible. Most imputed tax units are ineligible because they have no observed earned income.

Recall that if these eligible units represent actual non-filers, as opposed to federal-only filers, their forgone CalEITC amounts likely understate the benefits of filing. Many of these units were likely eligible for the federal EITC and other tax credits, not to mention a refund on overly withheld income taxes.

5 Additional results

5.1 Summary

Table 12 brings together our estimates for all three populations – filers who claimed the federal EITC, filers who did not claim the federal EITC, and non-filers. Overall, 49% of CalFresh-participating tax units — either real or imputed -– were eligible for the CalEITC. Among eligible tax units, 406,478 (46%) did not claim the CalEITC, forgoing $175 on average and over $71 million in total.

Over 600,000 tax units with a CalFresh-enrolled primary filer appeared eligible for the CalEITC (60% of all such units), and among those households, over 138,000 (22%) did not receive the CalEITC. These units forwent an average credit of $129. The other two-thirds of non-claimants were non-filers. Among non-filing CalFresh households, 35%, or over 268,000 imputed tax units, were eligible for the CalEITC, forgoing an average of $199.

Table 13 presents the distribution of eligible non-claimants by bins of estimated credit amounts, broken out by units with zero or one or more qualifying children. This table clarifies a few key results which may not have been obvious in the preceding tables. First, three-quarters of non-claiming units have zero qualifying children, and among these non-claimants, two-thirds are eligible for less than $100. Second, most non-claimers of the Federal EITC are single filers without qualifying children, which, as mentioned above, helps to explain why these units did not claim the Federal EITC and to downplay concerns about how we infer the number of qualifying children for these filers. Third, most non-claimants, especially those eligible for more than $500, are non-filers. The average forgone credit was $264 for filers with children and $81 for filers without children. The average forgone credit was $649 for non-filers with children and $87 for non-filers without children.

Table 12: Summing up CalEITC take-up among CalFresh recipients

	Total	CalEITC eligible		Eligible non-claimants			
	Count	Count	Share	Count	Share	Mean Amount	Total Amount
Filers							
Fed EITC claimants	721,354	514,469	71%	40,366	8%	$235	$9,503,865
Non Fed EITC claimants	302,791	100,881	33%	97,999	97%	$85	$8,305,715
Non-Filers							
All	762,102	268,113	35%	268,113	100%	$199	$53,333,651
Total	1,786,247	883,463	49%	406,478	46%	$175	$71,143,231

Notes. Table 12 compiles information from earlier tables; see those tables for details. The addition is the final column which reports the total unclaimed dollars for each population. Cells represent column percentages.

Table 13: Number of eligible non-claimants by estimated credit amount and number of qualifying children

	Non-filers	Filers		Total
		Fed EITC Claimers	Non-Fed EITC Claimers	
Zero qualifying children in unit				
$1-99	139,366	4,829	70,701	214,896
$100-199	56,831	793	20,964	78,588
$200-499	12,924	172	4,916	18,012
Total	209,121	5,794	96,581	311,496
One or more qualifying child in unit				
$1-99	12,681	11,891	626	25,198
$100-199	10,347	11,237	406	21,990
$200-499	11,902	7,357	188	19,447
$500-999	10,169	1,716	85	11,970
$100-1999	11,006	1,858	95	12,959
$2000+	2,887	513	0	3,400
Total	58,992	34,572	1400	94,966

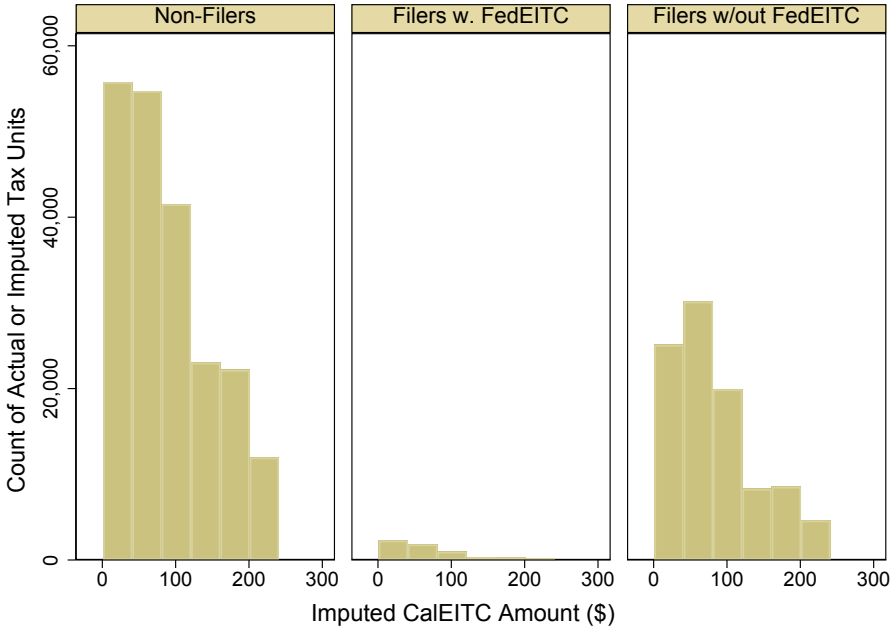
Notes. Table 13 tallies eligible non-claimants by number of qualifying children and expected credit amount. Column 1 reports counts for non-filers. Column 2 reports counts for filing units that claimed the Federal EITC. Column 3 reports counts for filing units that did not claim the Federal EITC. Column 4 reports row totals.

Figures 2 and 3 present the same results, plotting the frequency of eligible non-claiming tax units by estimated credit amount and earned income. Panel A in each figure presents counts for units with zero qualifying children, and

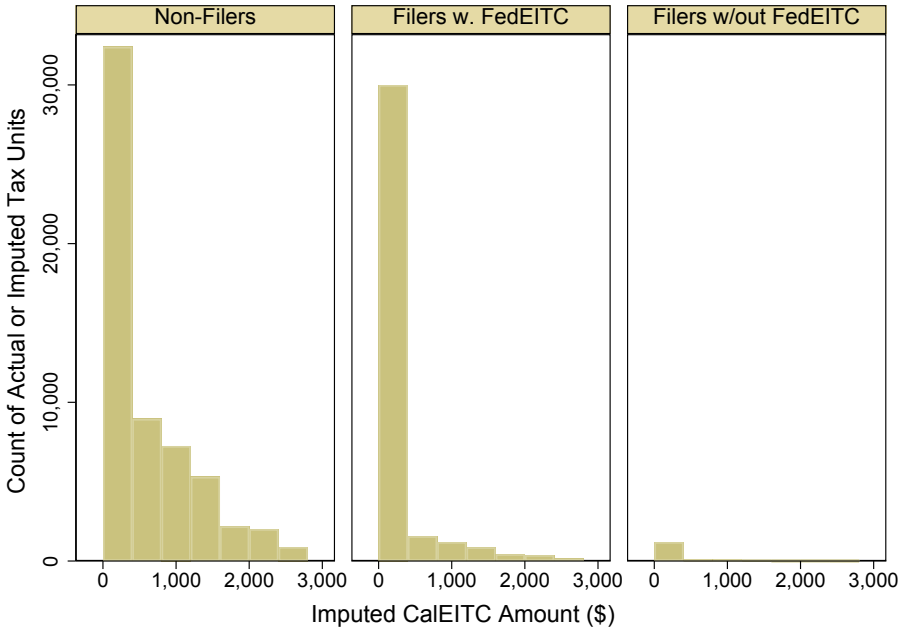
Panel B in each figure presents counts for units with one or more qualifying children.

Figure 2: Distribution of eligible non-claimants by estimated credit amount, TY 2017

(a) 0 qualifying children



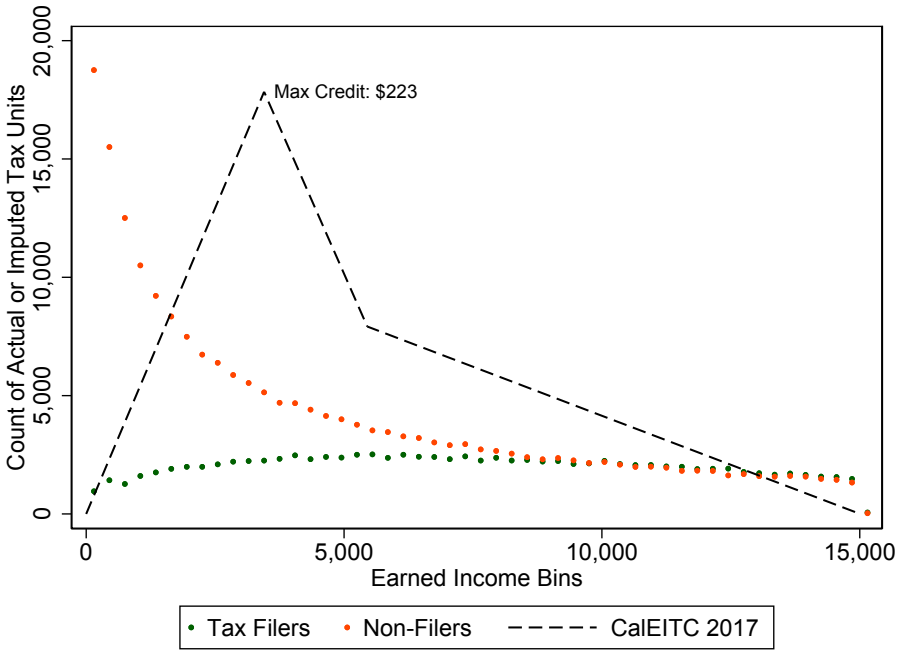
(b) 1+ qualifying children



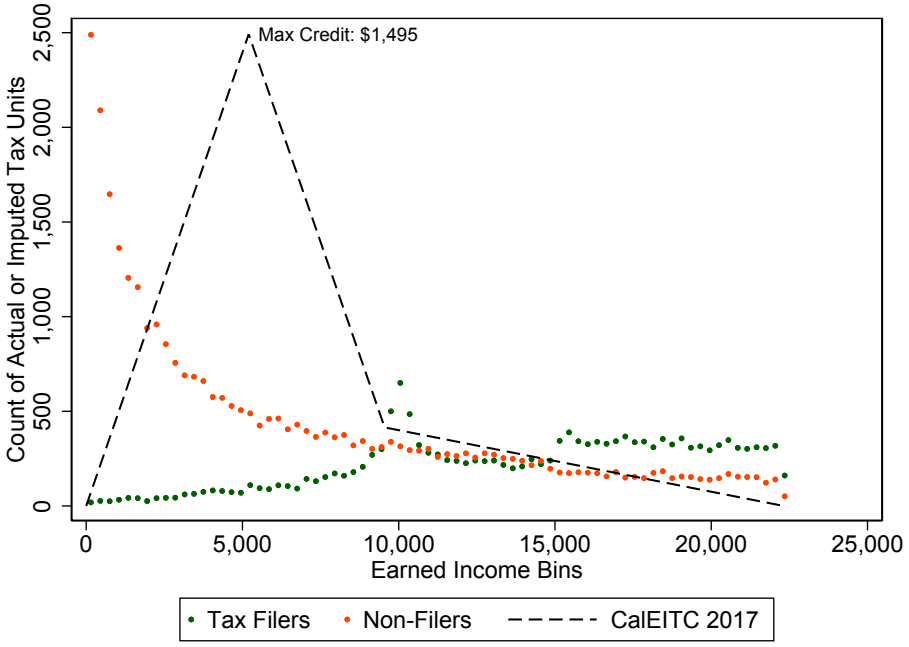
Notes. Figure 2 plots the frequency of of eligible non-claiming tax units by estimated credit amount. Panel A presents results for units with zero qualifying children. Panel B presents counts for units with one or more qualifying children.

Figure 3: Distribution of eligible non-claiming tax units by household earned income, TY 2017

(a) 0 qualifying children



(b) 1+ qualifying children



Notes. Figure 3 plots the frequency of eligible non-claiming tax units by earned income. Panel A presents results for units with zero qualifying children. Panel B presents counts for units with one or more qualifying children.

5.2 Take-up of the federal EITC

We are not able to produce a comprehensive estimate of federal EITC claiming among CalFresh households, because we do not observe federal returns or federal EITC claims for those who do not submit a state return or filed a paper return. However, we can measure take-up of the federal EITC among those who e-filed a state return.

We measure eligibility for the federal EITC within this population in the same way that we measured eligibility for the CalEITC among those that did not claim the federal EITC (summarized in Section 4.2). We use CalFresh records to test which dependents might be qualifying children, and we use filing status and earned income reported on the state return. Table 14 summarizes our estimates of eligibility and participation among all state returns with a CalFresh-enrolled head or spouse. We present results separately for units that claimed the state EITC and those that did not. Mirroring our count from Table 1, over 1 million returns contained a CalFresh filer or spouse. Among those, 80% were eligible to receive the federal EITC. Of those, 88% claimed the credit. Those who claimed the CalEITC were overwhelming likely to also claim the federal EITC. A large share of seemingly eligible households who failed to claim the state EITC also failed to claim the federal EITC. Altogether, over 97,000 tax units who were eligible did not claim the federal EITC, forgoing $385 on average and $37 million in total.

Table 14: Simulated federal EITC eligibility among CalFresh filers

	Tax units	**Fed EITC eligible**		**Eligible non-claimants**			
	Count	Count	Share	Count	Share	Mean Amount	Total Amount
Filers							
CalEITC claimant	485,562	482,509	99%	2,859	1%	$711	$2,034,108
Non CalEITC claimant	550,061	341,681	62%	94,261	28%	$375	$35,371,763
Total	1,035,623	824,190	80%	97,120	12%	$385	$37,405,871

Notes. Table 14 summarizes rates of eligibility for the federal EITC among e-filed tax returns with a primary filer who claimed CalFresh. Results are separated between units that included a CalEITC claim and those that did not.

5.3 By tax preparation method

A reasonable question raised by our results is why any eligible filer would not claim the CalEITC. One explanation we can explore with our data is the role of tax preparers. Table 15 reports estimates of eligibility and participation by the three methods of tax preparation that we observe in our FTB records: self-prepared, prepared by a paid professional, and prepared through the Volunteer Income Tax Assistance (VITA) program.

We find that units, among filers that claimed the federal EITC, non-claimants were disproportionately likely to have used a paid preparer. Among all returns in our sample containing a federal EITC claim, 63% were filed by a paid preparer. Among the 40,000 tax units who claimed the federal EITC and were eligible for but did not claim the CalEITC, 92% were filed by a paid preparer.

Among tax units on CalFresh who did not claim the federal EITC, we see a higher rate of eligible non-claiming among self-prepared returns (36% for the federal EITC and for the CalEITC) than among returns prepared by VITA or paid preparers (between 27 and 33%). The forgone credit amounts are similar across self-prepared and paid prepared returns, with the exception of forgone CalEITC dollars among federal EITC claimants who used VITA services, which were $100 less than the mean amount forgone among paid and self-preparers.

A lack of experience with or information about the tax system might explain the slightly higher levels of non-claiming among self-preparers. Self-preparers might decide that the extra time and attention needed to claim the

state EITC might not justify the small dollar amount for which they're eligible. It is harder to pinpoint what might be causing a higher non-claiming rate amount Federal EITC claimants who use paid paid preparers. Perhaps some preparers have limited experience with a new tax policy. More plausibly, the fees charged by preparers could dissuade eligible filers from claiming the state credit. The majority of non-claiming filers were eligible for less than $200, and the fees charged by preparers to claim credits on filers behalf can be in excess of this amount. More research on the role played by the tax preparation industry in accurately and efficiently distributing tax-based benefits to households is needed.

Table 15: CalEITC take-up among CalFresh recipients by tax preparation method

	Tax Preparation Method			
	Paid	**Self**	**VITA**	**Total**
Number of Tax Returns	630,424	346,893	44,821	1,022,138
Claimed the Federal EITC				
Number of tax returns	455,147	238,235	26,758	720,140
CalEITC				
% eligible	69%	74%	80%	71%
% non-claiming among eligible	11.8%	1.6%	0.4%	7.8%
Mean unclaimed amount	$234	$255	$139	$235
Total unclaimed amount	$8,736,498	$733,815	$11,813	$9,482,126
Did Not Claim the Federal EITC				
Number of Tax Returns	175,277	108,658	18,063	301,998
CalEITC				
% eligible	32%	36%	33%	33%
% non-claiming among eligible	99%	94%	100%	97%
Mean unclaimed amount	$85	$83	$91	$85
Total unclaimed amount	$4,696,586	$3,049,404	$540,389	$8,286,379
Federal EITC				
% eligible	30%	36%	27%	32%
% non-claiming among eligible	100%	100%	100%	100%
Mean unclaimed amount	$381	$396	$343	$385
Total unclaimed amount	$20,069,306	$15,613,368	$1,647,044	$37,329,718

Notes. Universe is e-filed tax returns linked to at least one primary filer who claimed CalFresh. The top panel is restricted to tax returns in which there was a positive federal EITC claim, and the bottom panel is restricted to returns with no federal EITC claim. For each population, we report the number of returns filed via each of the three preparation methods: paid preparer, self-prepared, or VITA. For each method, we report the share of returns that appeared eligible for the CalEITC (and the federal EITC for those that did not claim the Federal EITC), and among those deemed eligible, the share that did not claim. For eligible non-claimers, we also report the mean unclaimed amount and the total unclaimed dollars. The share of eligible non-claimants among tax units who claimed the federal EITC and used VITA is 0.4 percent, which we round down to 0%. A very small number do not claim the state eitc, and the average unclaimed amount for this group was $145 and the total was $12,789.

5.4 By race

Table 16 presents estimates of eligibility and participation by the race of CalFresh recipients.[26] Presenting results at the tax unit level would require that we ascribe results only to one of the actual or simulated primary filers of each

[26]The race/ethnicity variable we use comes from our CDSS data. This variable combines concepts of race and ethnicity. It is also a combination of self-reporting and social worker visual identification (applicants are asked to provide their self-identified race/ethnicity, but if they do not mark anything the eligibility worker may enter a value based on their own visual assessment). In February 2020, CDSS issued guidance to limit all reporting on race and ethnicity to be self-reported. The demographic distribution of race/ethnicity in the CDSS data is comparable to the

tax unit. Instead, we present results at the individual level.[27] We also report our results separately for primary filers and for dependents.

Among filers, eligible Hispanic primary filers are less likely to claim the CalEITC than primary filers from other racial groups represented in our CalFresh sample. Take-up is about 7 percentage points lower for Hispanic primary filers than White heads and spouses for example. Hispanic dependents in eligible tax units are also less likely to receive the CalEITC. Across other racial groups, share of eligible households not claiming the CalEITC is roughly equal (between 18% and 20% for primary filers and 7 to 9% for dependents).

A greater share of Black, Hispanic, and Native Hawaiian and Pacific Islander (NHPI) non-filing imputed primary filers appear eligible for the CalEITC than other imputed primary filers. Forty-two percent of Black, 38% of Hispanic, and 34% of NHPI non-filing imputed primary filers appeared eligible for the CalEITC, compared to 30% of White imputed primary filers. We see similar gaps in participation among imputed dependents. Black, Hispanic and NHPI non-filers are also eligible for higher imputed CalEITC credit amounts. Notably, a much lower share of Asian non-filers appear eligible for the CalEITC (25% of primary filers and 28% of dependents).

Altogether, among primary filers, non-claiming is highest among individuals who identify as American Indian and Alaskan Native (AIAN), followed by Black, NHPI, and White individuals, whose non-claiming rates are fairly comparable. Claiming is highest among Asian and Asian/Pacific Islander individuals. Among dependents, non-participation rates are highest among those who identify as AIAN, as well as NHPI. Claiming rates are fairly comparable between Black, White individuals. Claiming is again highest among individuals who identify as Asian or Asian/Pacific Islander.

6 Discussion

By separately estimating participation rates for filers and non-filers, and highlighting the components of eligibility that tax administrators can and cannot easily confirm, our analysis suggests multiple strategies for increasing take-up of the CalEITC among eligible families.

Increasing take-up among households who file a state return and claim the federal credit would be simplest. California tax administrators observe nearly all the information they need to confirm these households' eligibility and predict their correct CalEITC amount. The state tax agency could verify these households' eligibility on their own, and automatically send them their owed credit. If necessary, the state could also ask these households to attest to the information submitted on their return first.

Increasing take-up among eligible filers who do not claim the federal credit would be more complicated. Estimating likely EITC amounts for this group requires imputing the correct number of qualifying children. As we have shown, tax administrators can use CalFresh records to verify residential arrangements and identify likely eligible households. This process is imperfect, but largely accurate. The tax agency could use this process to focus on those tax units that are most likely eligible, and then reach out to those households to confirm their dependents are qualifying children. While some filers may not respond, others surely will, and the tax agency can then send

distribution of California households enrolled in SNAP by race/ethnicity from the American Community Survey (2019). We exclude the "two or more race" category due to small cell sizes.

Our analysis captures individuals in safety-net programs administered by CDSS who identify as American Indian and Alaska Native but do not live on tribal land and/or earn tribal income. Individuals who earn tribal income are exempt from state tax filing in California and may not appear as having received a payment automatically in our data. However, among those earners who qualify for safety-net programs, most are also likely eligible for tribal safety-net programs (such as the Food Distribution Program on Indian Reservations and Tribal TANF) and would not appear in the MEDS data.

CDSS reports nine ethnicities that are grouped by the US Census Bureau into an "Asian" category (Asian Indian, Cambodian, Chinese, Filipino, Japanese, Korean, Laotian, and Vietnamese), and three ethnicities that are grouped by the US Census Bureau as "Native Hawaiian and Other Pacific" (Guamanian, Hawaiian, and Samoan). Due to small cell sizes, we are unable to report each category uniquely, and use the US Census race/ethnicity categories to best capture the distinct take-up rates across all these categories. CDSS also has a separate category in the data named "Asian or Pacific Islander". We are unable to meaningfully distinguish between each community in that category and so we choose to report it separately.

[27] Accordingly, we do not total amounts of forgone EITC dollars to avoid double-counting.

Table 16: CalEITC eligibility and participation by race

	AIAN	Asian	Asian/PI	Black	Hispanic	NHPI	Other/ Unknown	White	Total
Within actual tax units									
Heads and Spouses									
Number of individuals	6,067	77,418	14,496	135,964	554,410	3,955	132,741	255,134	1,180,185
% eligible	59%	57%	53%	68%	56%	55%	56%	56%	58%
% non-claiming among eligible	18%	19%	18%	17%	25%	21%	19%	18%	21%
Mean unclaimed amount	$139	$108	$107	$118	$115	$127	$113	$113	$114
Dependents									
Number of individuals	4,412	51,566	10,878	130,346	614,925	3,977	116,712	179,740	1,112,556
% eligible	58%	53%	49%	69%	52%	48%	56%	55%	55%
% non-claiming among eligible	8%	7%	7%	7%	11%	9%	7%	7%	9%
Mean CalEITC amount	$320	$227	$183	$249	$221	$256	$253	$271	$233
Within imputed tax units									
Heads and Spouses									
Number of individuals	8,812	32,732	7,671	137,018	260,570	3,409	93,753	267,367	811,332
% eligible	30%	19%	25%	42%	38%	34%	33%	30%	35%
% non-claiming among eligible	100%	100%	100%	100%	100%	100%	100%	100%	100%
Mean unclaimed amount	$59	$40	$62	$87	$85	$89	$63	$57	$71
Dependents									
Number of individuals	2,864	8,599	2,955	56,260	135,730	1,494	33,808	74,951	316,661
% eligible	26%	23%	28%	37%	32%	36%	31%	29%	32%
% non-claiming among eligible	100%	100%	100%	100%	100%	100%	100%	100%	100%
Mean CalEITC amount	$160	$146	$192	$263	$227	$271	$209	$185	$219
Overall take-up gap									
Heads and Spouses									
% non-claiming among eligible	53%	29%	34%	49%	43%	49%	42%	47%	44%
Dependents									
% non-claiming among eligible	29%	13%	20%	25%	22%	29%	20%	24%	22%

Notes. The universe for the top panel is e-filed tax returns linked to at least one primary filer who claimed CalFresh. The universe for the bottom panel is simulated tax units only containing non-filing CalFresh participants. Within each race category, we report the share of heads/spouses and dependents (either as reported on their tax return for those who filed or as predicted via the process described in Appendix D) within each race category who were eligible for the CalEITC. Among those eligible, we also report the share that did not claim the CalEITC and the mean imputed amount for these eligible non-claimers. Column 1 reports statistics for American Indian and Alaskan Native enrollees. Column 2 reports statistics for Asian enrollees. Column 3 reports statistics for Asian and/or Pacific Islander enrollees. Column 4 reports statistics for Black enrollees. Column 5 reports statistics for Hispanic enrollees. Column 6 reports statistics for Native Hawaiian and/or Pacific Islander enrollees. Column 7 reports statistics for those whose race is reported as Other or Unknown. Column 8 reports statistics for White enrollees. We do not report statistics for the less than 1% of enrollees who are associated with different race categories across copies of our MEDS data. See Section 5.4 for more information about the data we use and how we group enrollees into these categories.

owed credit amounts to those households. The benefits to taxpayers would surely exceed the cost of this outreach. Successful outreach to these eligible filers might also help them claim the more valuable federal credit. That said, the large majority of non-claimants in this population are single filers without dependents, for whom imputing qualifying children would not be necessary. Tax administrators can estimate a likely credit amount for these filers using information contained within their tax returns.

Those who do not file a state return are the hardest to reach. We expect that our imputation process yields reliable estimates of average eligibility and participation, but we are not able to guarantee accuracy at the individual tax unit level. The tax agency would need to engage these recipients and encourage them to file a return, so as to confirm family structure and earnings.

7 Conclusion

In this paper, we use California administrative data to measure eligibility and participation in the state's Earned Income Tax Credit (CalEITC) program among households who enroll in the state's Supplemental Nutrition Assistance Program (CalFresh). We match state program enrollment data, as well as linked earnings records, to the universe of state tax returns in tax year 2017. We use CalFresh casefiles and earnings records to measure eligibility among non-filers, and these casefiles plus information from households' own returns to measure eligibility and participation among filers. We find that nearly 150,000 actual tax units failed to claim the state EITC, totaling almost $20 million in unclaimed state benefits. We also find almost 270,000 CalFresh households were eligible for the CalEITC but did not file a state return, forgoing an estimated $56 million dollars in CalEITC benefits. The overall household-level take-up rate of the California EITC within the CalFresh population is 53%. The average unclaimed benefit for both filers and non-filers is fairly low. Participation is lower for childless adults, for filers who use paid preparers, and among Black and Hispanic individuals. We also replicate these results for TY 2016, finding a lower take-up rate of 43%. The increase in the take-up rate between 2016 and 2017 is largely driven by the substantial increase in the number of eligible filers.

As mentioned, our study – like all those aiming to measure take-up – suffers from a few limitations. First, we examine only the population of CalFresh participants; we do not estimate eligibility or take-up rates for the much larger population of California families, many low income, who do not receive CalFresh benefits (See Appendix A). Second, we are forced to exclude paper filers. We expect that doing so results in our underestimating the size of the take-up gap (See Appendix B). Third, we estimate take-up only of the CalEITC. We provide limited evidence of non-claiming of the more valuable federal credit, but since we are unable to observe whether families who did not file state returns did file a federal return, we cannot produce a comprehensive measure of non-participation for the federal credit. Extension of our methods to incorporate federal-only returns would enable an analysis of this form, but this would likely require data-sharing between a state social service agency, like CDSS, and the IRS.

All of our estimates pertain only to the population of CalFresh participants. An important avenue for future work would be replicating this analysis using a fuller array of state administrative data – namely, Medicaid case files. Not only are many more households enrolled in Medicaid than SNAP, providing researchers a larger snapshot of the possibly eligible population, Medicaid cases tend to more closely resemble tax units. Using Medicaid case composition information would also improve upon our processes for associating filers with partners and dependents.

In addition to providing the first estimate of take-up of a state EITC program, a principal contribution of this paper is to provide a roadmap to other researchers so they can measure EITC participation in their own states.

References

Bhargava, Saurabh, and Dayanand Manoli. 2015. "Psychological Frictions and the Incomplete Take-up of Social Benefits: Evidence from an IRS Field Experiment." *American Economic Review*, 105(11): 3489–3529.

Blumenthal, Marsha, Brian Erard, and Chih-Chin Ho. 2005. "Participation and Compliance With the Earned Income Tax Credit." *National Tax Journal*, 189–213.

Chetty, Raj, John N Friedman, and Emmanuel Saez. 2013. "Using Differences in Knowledge across Neighborhoods to Uncover the Impacts of the EITC on Earnings." *American Economic Review*, 103(7): 2683–2721.

Cranor, Taylor, Sarah Kotb, and Jacob Goldin. 2019. "Does Informing Employees About Tax Benefits Increase Take-Up?: Evidence from EITC Notification Laws." *National Tax Journal*, 72(2): 1–8.

Currie, Janet. 2006. "The Take-up of Social Benefits." In *Poverty, the Distribution of Income, and Public Policy*. , ed. Alan Auerbach, David Card and John Quigley, 80–148. Russell Sage.

Davis, Charles, and Evan White. 2019. "Who Receives the Earned Income Tax Credit in California?" California Policy Lab.

Edin, Katharine, Laura Tach, and Sarah Halpern-Meekin. 2014. "Tax Code Knowledge and Behavioral Responses Among EITC Recipients: Policy Insights from Qualitative Data." *Journal of Policy Analysis and Management*, 33(2): 413–439.

Goldin, Jacob, and Zachary Liscow. 2018. "Tax Benefit Complexity and Take-Up: Lessons from the Earned Income Tax Credit." *Tax Law Review*, 72.

Goldin, Jacob, Tatiana Homonoff, Rizwan Javaid, and Brenda Schafer. 2021. "Tax Filing and Take-Up: Experimental Evidence on Tax Preparation Outreach and EITC Participation." NBER Working Paper No. 28398, https://doi.org/10.3386/w28398.

Guyton, John, Dayanand S Manoli, Brenda Schafer, and Michael Sebastiani. 2016. "Reminders & Recidivism: Evidence from Tax Filing & EITC Participation Among Low-Income Nonfilers." NBER Working Paper No. 21904, https://doi.org/10.3386/w21904.

Halpern-Meekin, Sarah, Kathryn Edin, Laura Tach, and Jennifer Sykes. 2015. *It's Not Like I'm Poor: How Working Familes Make Ends Meet in a Post-Welfare World.* Univ of California Press.

Hoynes, Hilary, and Jesse Rothstein. 2017. "Tax Policy Toward Low-Income Families." In *The Economics of Tax Policy*. , ed. Alan J Auerbach and Kent Andrew Smetters. Oxford University Press.

IRS. n.d.*a*. "EITC participation rates by state." *https://www.eitc.irs.gov/eitc-central/participation-rate/eitc-participation-rate-by-states*, Accessed: 2020-03-20.

IRS. n.d.*b*. "Statistics for Tax Returns with EITC Claim." *https://www.eitc.irs.gov/eitc-central /statistics-for-tax-returns-with-eitc/statistics-for-tax-returns-with-eitc*, Accessed: 2021-03-15.

Jones, Damon. 2010. "Information, Preferences, and Public Benefit Participation: Experimental Evidence from the Advance EITC and 401(k) savings." *American Economic Journal: Applied Economics*, 2(2): 147–63.

Jones, Maggie, and James Ziliak. 2019. "The Antipoverty Impact of the EITC: New Estimates from Survey and Administrative Tax Records." US Census Bureau, Center for Economic Studies.

Jones, Maggie R. 2014. "Changes in EITC Eligibility and Participation, 2005-2009." Center for Administrative Records Research and Applications Working Paper.

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, and James P Ziliak. 2020. "The antipoverty impact of the EITC: New estimates from survey and administrative tax records."

Kopczuk, Wojciech, and Cristian Pop-Eleches. 2007. "Electronic Filing, Tax Preparers and participation in the Earned Income Tax Credit." *Journal of Public Economics*, 91(7-8): 1351–1367.

Larrimore, Jeff, Jacob Mortenson, and David Splinter. 2021. "Household Incomes in Tax Data Using Addresses to Move from Tax-Unit to Household Income Distributions." *Journal of Human Resources*, 56(2): 600–631.

Liebman, Jeffrey B. 1998. "The Impact of the Earned Income Tax Credit on Incentives and Income Distribution." *Tax Policy and the Economy*, 12: 83–119.

Linos, Elizabeth, Allen Prohofsky, Aparna Ramesh, Jesse Rothstein, and Matthew Unrath. 2020. "Can Nudges Increase Take-up of the Earned Income Tax Credit?: Evidence from Multiple Field Experiments." NBER Working Paper No. 28086, https://doi.org/10.3386/w28086.

Maag, Elaine, Michael Pergamit, Devlin Hanson, Caroline Ratcliffe, Sara Edelstein, and Sarah Minton. 2015. "Using Supplemental Nutrition Assistance Program Data in Earned Income Tax Credit Administration." Urban Institute.

Meyer, Bruce D, and Nikolas Mittag. 2019. "Using Linked Survey and Administrative Data to Better Measure Income: Implications for Poverty, Program Effectiveness, and Holes in the Safety Net." *American Economic Journal: Applied Economics*, 11(2): 176–204.

Meyer, Bruce D, Derek Wu, Grace Finley, Patrick Langetieg, Carla Medalia, Mark Payne, and Alan Plumley. 2020. "The Accuracy of Tax Imputations: Estimating Tax Liabilities and Credits Using Linked Survey and Administrative Data." National Bureau of Economic Research.

Nichols, Austin, and Jesse Rothstein. 2016. "The Earned Income Tax Credit." In *Economics of Means-Tested Transfer Programs in the United States, Volume 1*. 137–218. University of Chicago Press.

Phillips, Katherin Ross. 2001. "Who Knows About the Earned Income Tax Credit?" The Urban Institute.

Plueger, Dean. 2009. "The EITC Participation Rate for Tax Year 2005." *Internal Revenue Service Bulletin*.

Romich, Jennifer L, and Thomas S Weisner. 2000. "How Families View and Use the Earned Income Tax Credit: Advance Payment Versus Lump-Sum Delivery." *National Tax Journal*, 53(4.2): 1245–1265.

Scholz, John Karl. 1994. "The Earned Income Tax Credit: Participation, Compliance, and Antipoverty Effectiveness." *National Tax Journal*, 63–87.

Smeeding, Timothy M, Katherin Ross Phillips, and Michael O'Connor. 2000. "The EITC: Expectation, Knowledge, Use, and Economic and Social Mobility." *National Tax Journal*, 1187–1209.

Splinter, David, Jeff Larrimore, and Jacob Mortenson. 2017. "Whose child is this? shifting of dependents among eitc claimants within the same household." *National Tax Journal*, 70(4): 737–758.

A Representativeness of CalFresh population

We use the 2017 ACS 5-year sample to investigate how representative CalFresh recipients are of the low-income population in California, and to what extent our estimates of take-up in the CalFresh population might apply to all low-income households in the state. As discussed in Section 2, the ACS does not contain all the information needed to accurately estimate EITC eligibility, but we can compare the overall income, demographic, and household characteristics of SNAP enrollees (and those who reside with SNAP enrollees) to those who do not enroll in SNAP (and reside with no other SNAP enrollees).

First, we restrict the ACS sample to households who reside in California. Approximately 5.2 million individuals reside in a household with at least one SNAP recipient, versus 32.9 million in households without a SNAP recipient. We consider a subsample of these data, limited to those 18–64-year-old individuals with family income between $0 and $69,063 (the median family income among this population in 2017), excluding those in group-quarters. Within this group, we contrast those who live in a household with someone claiming SNAP to those who do not (Appendix Table 1). This sample includes 81% of those 18-64-year-olds in SNAP households, and 46% of those in non-SNAP households.

Overall, adults between 18-64 years old in households with a SNAP recipient (hereafter referred to as SNAP families) tend to belong to larger families than those in households without SNAP recipients (non-SNAP families). SNAP families contain an average of 4.1 individuals, compared to 2.8 in non-SNAP families. The composition of those families is also different. SNAP families tend to contain more children and fewer elderly individuals. Adults in SNAP families are more likely to be Hispanic relative to those in non-SNAP households and to a lesser degree more likely to be non-White.

SNAP families also tend to have lower incomes than non-SNAP families. Median total family income among SNAP families is $11,886 lower than for non-SNAP families. This trend is similar for earned income ($11,917 lower) and wage income ($11,376 lower). Non-SNAP families are slightly more likely to have a positive amount of investment income than SNAP families and are also more likely to have investment income over the EITC cutoff. The lower total, earned, and wage income at the family level among SNAP families suggests a larger share of these families could be eligible for the federal and California EITC than non-SNAP families.

In Appendix Figure 1, we compare the distribution of income among SNAP families versus non-SNAP families. The difference in the distribution of income between SNAP and non-SNAP families is similar across total, earned, and wage income. The lower means reported in Table A1 seems to be a product of this overall shift in the income distribution, although there is some small evidence of a slightly higher share of SNAP families having $0 in family income (though not wages).

These characteristics suggest that families in SNAP households are more likely to be eligible for the federal EITC and CalEITC, given their lower income and greater number of children. Of course, many eligible families do not appear in our CalFresh data. It is also likely we miss many eligible non-claimants, because there are many more low-income individuals and families in the non-SNAP population, and the same families who would choose not to enroll in a program like SNAP would also be less inclined to file a return and claim the CalEITC.

B Comparison of e-filers with paper and web filers

Our main analysis does not include the approximately 13% of California tax units who do not e-file. We exclude these paper filers because we are unable to observe their federal tax returns, and federal tax returns contain information necessary to determine both EITC eligibility and federal EITC receipt. Excluding paper filers does not impact our estimate of EITC eligibility among non-filers, because paper filers are included in the matching process between the tax and social service universe, meaning these filers are not inappropriately included in the non-filing population.

Based on the information we do have on paper filers,[28] we believe that excluding these tax units from our analysis leads us to underestimate the share of tax units that are eligible for but do not claim the CalEITC. Appendix Table 2 compares means of tax filing characteristics between paper and e-filers. Overall, e-filers and paper filers appear similar. E-filers are slightly more likely to file as married filing jointly or as a head of household and have slightly more dependents. Paper and e-filers claim the CalEITC in roughly equal proportions, and e-filers are eligible for slightly higher CalEITC amounts on average. E-filers are slightly more likely to be on CalFresh and are no-more likely to have an ITIN present on their tax return relative to paper filers.

However, there are two substantial differences that lead us to believe that the inclusion of paper and web filers would increase the rate of non-participation among tax filers. First, paper filers are much more likely to self-file tax returns, as opposed to filing with either a paid preparer or through a free tax preparation service like VITA. Among e-filers, 30% of tax units self-prepare their returns compared to 73% of paper or web filers. It seems likely that filing with the aid of either a paid or free tax preparation service increases the likelihood that a tax unit claims tax credits for which they are eligible. Second, paper filers have lower wages and adjusted gross income than e-filers, making it more likely that they have income in the EITC-eligible range. Median AGI is $38,009 and median wages are $24,732 for paper filers, versus $42,823 and $31,482 for e-filers. Even if we limit the sample to tax units with a head or spouse on CalFresh, these differences remain. These differences also remain, though the gap decreases, if we limit to the sub-population of tax units that claimed the CalEITC.

To test how these two factors combine to impact EITC non-claiming, we construct a measure of imputed CalEITC eligibility using only characteristics available from the primary state tax form (Schedule 540), which we can observe for all California filers. This measure of eligibility will be less accurate than our measure for e-filers alone. For example, it does not include non-wage earned income and investment income from the federal 1040 form and does not incorporate our efforts to accurately estimate the number of qualified children. But it allows us to roughly compare eligibility rates between e-filers and paper filers using the same information. We only consider CalEITC eligibility, because we cannot accurately determine which paper or filers claim the federal EITC without their 1040 information.

We find that roughly equal shares of e-filers (13%) and paper filers (14%) are likely to be eligible for the CalEITC, but paper filers are 13 percentage points less likely to claim the CalEITC, conditional on imputed eligibility. When limited to tax units on CalFresh, that difference falls to 8 percentage points. Given these results, we anticipate that excluding paper returns likely results in an underestimate of the share of eligible tax units that do not claim their CalEITC amounts.

C Individuals without Social Security Numbers

C.1 Valid SSN rule

Individuals without a valid Social Security Number (SSN) can use an Individual Tax Identification Number (ITIN) when filing a return. Though individuals with ITINs are generally not eligible for tax-based benefits, they can receive refunds of overpaid income taxes. In 2017, 6% of all individuals who appeared on a California state tax return (1.9 million individuals, including 1 million dependents and 900,000 filers or spouses) filed using an ITIN. Of the 5.1 million individuals on tax returns that include at least one CalFresh recipient, 7%, or 360,000, had an ITIN.

Only tax units in which both the head and spouse have a valid SSN can claim the federal EITC, and only dependents with a valid SSN can be a qualifying child. In 2017, eligibility for the California EITC was also restricted to tax units in which both head and spouse had a valid SSN, though ITIN filers became eligible for the CalEITC in 2020. Individuals with a valid SSN can enroll in CalFresh even if another household member does not have a

[28] Paper returns – or tax returns that are submitted by mail to the FTB and IRS – make up 94% of the filers that do not e-file. The remaining 6% are web filers or filers who make use of CalFile, a service where tax filers can submit their state tax return directly to the FTB. We refer to web and paper filers as paper filers in this section.

valid SSN, though the individual without a valid SSN cannot could toward the enrollees' household size and thus the calculation of their benefit amount.

We account for the 2017 SSN rules when assigning CalEITC eligibility to actual tax units. We disqualify any actual tax unit with a head or a spouse who has an ITIN, and we disregard dependents without a valid SSN when determining which might be a qualifying child. For non-filers, we assume that our simulated tax units would file returns according to the composition imputed to them. In other words, these simulated tax units would not include a head, spouse or dependent unobserved in the CalFresh records who might not have a valid SSN.

Of the 394,000 CalFresh recipients who are on a tax return with an individual with an ITIN,[29] 94% appear in our data as Hispanic, even though Hispanic individuals represent just 57% of all CalFresh recipients who appear on a 2017 state return. This suggests that that Hispanic tax filers may be more likely to lose eligibility for the EITC due to the documentation status of someone in their tax unit, even conditional on their receipt of CalFresh.

C.2 Impact of the modifying the SSN rule

In June 2020, California expanded eligibility for the CalEITC to ITIN filers with young children. To model the impact of this expansion, we consider the set of e-filers in 2017 (not limited to CalFresh recipients) and count the number of households who might be eligible for the CalEITC if not for a filer or spouse having an ITIN.

We take as given the actual tax unit composition and the income reported on the tax return, as we do with our imputations described in the main part of the paper. However, since ITIN filers cannot claim either the federal or state EITC, we cannot use information from those claims to infer which dependents on the return might be qualifying children. We also cannot rely on CalFresh casefiles for the residency test, since not all e-filers with an ITIN appear in our CalFresh records. Instead, we assume that all dependents are qualifying children.

In 2017, 611,570 returns included a filer or spouse with an ITIN (Appendix Table 4). Of these returns, we estimate that 71% would be eligible for the federal EITC and 18% for the CalEITC if not for the SSN test. On average, these tax units would qualify for, on average, $3,009 from the federal EITC and $288 from the CalEITC. We also report these statistics separately for single ITIN filers and married filers in which one or both filers have an ITIN. Eligibility rates for the federal and state EITC are highest for single filers, followed by married units in which both filers have an ITIN, and then mixed status couples. If these tax units were allowed to claim either credit, an additional 917,488 children would become eligible for the federal EITC and an additional 234,534 children would become eligible for the CalEITC. If these households were allowed to claim both credits, they would be eligible to receive nearly $47 million from the CalEITC and nearly $1.3 billion from the federal EITC.

The second group of tax units impacted by the repeal of either the federal or California SSN test are those in which all heads or spouses have SSNs, but at least one dependent has an ITIN. We observe 207,809 such tax units. If these dependents became eligible, 50% would be eligible for the federal EITC (up from 22%) and 13% would be eligible for the CalEITC (up from 8%). In total, an additional 176,179 children would become eligible for the federal EITC and 38,659 for the CalEITC. Their tax units would be eligible for an additional $302 million from the federal EITC and $9.7 million from the CalEITC. The average federal EITC amount for these households would be nearly $1,500 (up from $500), while the average CalEITC amount would be just $47 (up from $13).

D Assigning non-filing CalFresh recipients to be primary filers or dependents

As part of our construction of simulated tax units among non-filing CalFresh recipients, we predict whether a given recipient is likely to appear as a primary filer (a head or spouse) or a dependent on their unit's return if one were to

[29] We can use the racial data in our CalFresh records to study which recipients are most likely to appear on tax returns with an ITIN filer. We can only see the race of individuals who are enrolled in CalFresh, so we cannot observe the race of the individuals who have an ITIN.

be filed. Below, we describe how we use information in our CalFresh files to predict these roles for each individual.

First, we assign all individuals under the age of 18 or over the age of 80 to be dependents. Among CalFresh recipients who appeared on tax returns, 99.4% of individuals under the age of 18 are dependents. Second, we assign all individuals over the age of 80 to be dependents as well. We do this because (1) the vast majority (80%) of CalFresh recipients age 81 plus appear on a tax return as a dependent and (2) there are not enough individuals in this age bracket for our prediction method to produce reliable results.

Second, we take the pool of individuals between the ages of 18 and 80, and we predict whether each are a dependent or primary filer. While most of these working-age (or near working-age) individuals would likely be primary filers, a non-trivial share still appear as dependents in the tax records (Appendix Figure 2). Some of these adults might be too young to work, might still be in school, or are working but still residing with parents. Others could be older, unable to work, are being cared for by working-age children, and might have disabilities or other challenges which prevent them from working and qualify as a dependent.

We predict which non-filing adults might be dependents by identifying how characteristics observed in our CalFresh relate to whether filers appear as dependents or primary filers on their returns. We randomly assign 70% of individuals on CalFresh who appear on a tax year 2017 return to a training dataset and assign the remaining 30% to a test dataset. Using the observations in our training set, we estimate a logistic regression, with an indicator for filer/dependent status as our outcome variable and the following regressors: indicators for whether English is their primary language; whether they can be merged to EDD earnings records; whether they receive cash assistance or SSI; whether the individual is incarcerated, is a senior, is a non-resident; the size of the individual's reference case; and the individuals' age, interacted with their EDD wage income, number of unique CalFresh cases over the course of the tax year, whether the individual is disabled, a categorical race variable, number of months enrolled on CalFresh, a binary sex variable, and the number of persons on the individual's reference case.

Using estimates from this model, we constructed a predicted probability that each observation is a primary filer or dependent. We apply the predicted probabilities for each individual to the test set in order to select two probability cutoffs that we use for our final filer or dependent determination. All individuals with a predicted probability over the cutoff are set as primary filer. The first cutoff (50%) is the cutoff that minimizes error across the entire test dataset, with a prediction accuracy in the test set of 88%. The second set of cutoffs utilizes prior-year tax information for the test-set and allows the cutoff to vary across three groups: Individuals who were a primary filer on a tax return last year (a 4% cutoff), individuals who were a dependent last year (90%), and individuals who did not file last year (46%). With these cutoffs, we achieved an accuracy rate in the test set of 92%. This second version is used throughout the paper.

E Predicted versus actual CalEITC claim

To estimate eligibility for the CalEITC among CalFresh tax filers who claimed the federal EITC, we use the value of each tax unit's federal EITC claim, plus their earned income, to infer their number of qualifying children. In this section, we present evidence to further validate the reliability of this inference by considering the tax units that claimed both the federal and state EITC, and comparing our prediction of these units' CalEITC eligible claim amounts, using this inference procedure, to what they actually claimed.

Of the 752,597 tax units with a head or a spouse on CalFresh and who claimed the federal EITC, we observe 489,679 units claiming the CalEITC (489,262 eligible claimants and 410 apparently ineligible claimants). Of these, we fail to exactly predict the credit amount received for just 14,019 units, or 2.86% of all claiming units. On average, our predicted credit amount exceeds the credit amount such units actually received by roughly $235 dollars, while the median difference between predicted and actual credit values is $83.

In order to explain the source of these errors, we experimented with systematically varying the inputs to our CalEITC predictions. We began by substituting each tax unit's earned income with their AGI in our credit calculator.

The replacement of earned income with AGI allows us to match our predicted and the actual credit amount for 968 (7%) of the 14,019 tax units where we initially observed errors in our predictions. Though households are supposed to use their earned income to determine their correct eligible CalEITC amount, it is possible that a small number of preparers inputted the incorrect earnings variable.

For the remaining 13,051 units, we experimented with adjusting the number of qualifying children used in our credit amount predictions. We vary the possible number of qualifying children between 0 and 3 for each tax unit and calculate predicted credit amounts. This exercise allows us to recover the actual claimed amounts for an additional 10,925 units. For approximately 68% of these units, increasing or decreasing the count of QC by 1 child yielded accurate predictions of actual credit amounts.[30]

Finally, we experiment with both substituting AGI for earned income and varying the number of qualifying children for the 2,126 tax units for whom the above substitutions did not produce a predicted credit equal to their actual credit. In doing so, we are able to match predicted and actual credit amounts for an additional 285 tax units.

Following the above exercises, we are left with only 1,841 tax units (approximately 0.38% of the full sample of tax units claiming the CalEITC) for which no combination of substitutions above yielded an exact match between predicted and actual credit amounts.

F Alignment between SNAP cases and tax units

SNAP cases do not always align with tax units. Our process of constructing imputed tax units from SNAP records assumes that SNAP cases contain one or more tax units, and tax units do not stretch across SNAP cases. We believe this assumption is reasonable because, as discussed above, the rules for determining the bounds of a SNAP case (those who eat and prepare meals together) are arguably more flexible and expansive than the ones used to determine a tax unit. At the same time, this assumption could be mistaken. For example, by emphasizing shared residence and resources, the rules for determining a SNAP case could be stricter than those for a tax unit. Plus, SNAP records may only capture residential arrangements for parts of the tax year.

In this section, we present evidence on the extent to which SNAP case compositions disagree with tax unit composition, and the limitations of our disambiguation approach to constructing tax units. Since the bulk of the EITC's value comes from the claiming of a child, we're most concerned with correctly assigning children to their parents using the SNAP records, and this analysis is focused on how well those SNAP cases align with actual claiming. We present evidence for three mutually exclusive populations, covering all children enrolled in CalFresh in 2017: the children enrolled in a reference case with no adults; the children who appeared on a return and had at least one adult in their reference case; and he children who did not appear on a return and had at least one adult in their reference case.

Appendix Table 10 considers the nearly 700,000 children assigned to a reference case without an adult. Roughly one-quarter (28%) were not claimed on a return. [31] However, the vast majority of these children (72%) were claimed by an adult on a return, even though they were enrolled by themselves or with another child in their SNAP case. We hypothesized that many of these children must have resided with an adult who was ineligible for SNAP due to their documentation status, and we do find that roughly half were claimed on a return by someone who filed with an ITIN. However, it's less clear why the other half of children (over 250,000) were claimed by an adults who were not enrolled in the child's SNAP case in 2017. In the following rows, we present evidence for two candidate

[30] One possible explanation for this pattern of results is that these tax units may have had dependents who did not reside in the state of California for the requisite 6 months required to satisfy the state residency requirement of the CalEITC to be counted as qualifying children. Using CalFresh records, however, we find no evidence that dependents in this subset of tax units were systematically less likely to appear in CalFresh records, either on their own or matched with primary filers on their reference case, throughout the year than dependents on the broader sample of tax units included in this analysis. Likewise, such dependents were not more likely to fail at the age test requirement for the CalEITC as compared to the broader sample of dependents.

[31] Recall that these children were excluded from our analysis if they did not appear on a return or were claimed on a return not including a CalFresh-enrolled adult.

explanations. First, we consider the possibility that children might have enrolled with the claiming adult in a different SNAP case, and we misidentified their reference case. We do this by limiting our attention to children who enrolled in only one SNAP case throughout the year. Second, we consider the possibility that children were enrolled in SNAP for just a few months, and resided with the claiming adult for much of the year while not enrolled in SNAP. We do this by limiting our attention to the children who were enrolled in SNAP all 12 months. Neither restriction affects the share of children claimed on a return by a non-ITIN filer.

Appendix Table 11 considers the 1.25 million children enrolled in SNAP in California in 2017 who appeared on a state tax return and who had at least one adult on their reference case. We decompose these children into three groups: (1) children who appear on a return with all adults who on those children's reference case (Column 1); (2) children who appear on a return with at least one but not all adults who appeared on their reference case (Column 2); and (3) children who appear on a tax return containing no adults who appeared on the child's reference case, broken out by whether the return contained an adult who filed with an ITIN or not (Columns 3 and 4). Nearly two-thirds of filing children (roughly 800,000) appear on a return with all adults from their reference case. At the same time, one out of five children are claimed on a return by an adult who does not appear on their reference case. As with Appendix Table 10, we focus in on subgroups of these children to test explanations for this misalignment between reference cases and tax units. First, we consider children enrolled in just a single CalFresh case. Second, we consider children enrolled in a single CalFresh case for all 12 months of 2017. Third, we consider children enrolled in a single case including two or more adults. Fourth, we consider children enrolled in a single case with an adult who is in the most plausible age range to be the child's parent (25 to 34 years older). We find that children with multiple adults in their reference case are less likely to have all these adults appear on their return, and more likely to have at least one. Otherwise, these restrictions have limited effect to no effect on the distribution of children across these columns.

Appendix Table 12 considers the 360,000 children enrolled in SNAP in California in 2017 who did not appear on a state tax return but had an adult in their reference case. We decompose these children into three groups: (1) children for whom no adult in their reference case appeared on a tax return (Column 1); (2) children for whom at least one but not all adults on their reference case appeared on a return (Column 2); and (3) children for whom all adults on their reference case appeared on a return (Column 3). For nearly three-quarters of these children, no adults from their reference case appear on any return. However, for almost one-in-four of these children, at least one adult from their reference case does appear on a return. These shares remain unchanged after imposing the same restrictions that we used in Appendix Table 11. If claimed, these children might have allowed these adults to claim valuable (or even more valuable) refunds. It's unclear whether this non-claiming reflects incomplete take-up or the consequence of the rules governing qualifying children.

A limitation of our data is that we do not have information on intra-household relationships. Knowing whether an adult in a SNAP case is in fact a child's parent would improve our prediction of whether that adult and child should appear on a tax return together. An important area for future work is using more data than those available here to better describe the relationships between adults and children who alternatively do or do not appear together in a safety net program, and do or do not appear together on a tax return.

G Results for Tax Year 2016

Since we have the same CalFresh and tax data for tax years 2015 through 2017, we are able to replicate every stage of our procedure for tax year 2016. Note that since we rely on the prior year's tax filing information and we lack returns for tax year 2014, we cannot replicate our analysis for tax year 2015.

The CalEITC available to taxpayers in 2016 differed in two significant ways from the 2017 version we analyze in this paper. First, self-employment earnings were not counted as earned income for purpose of the CalEITC in 2016. Second, the maximum income tax units could report and still be eligible for the CalEITC increased from approximately $13,900 in 2016 to $22,300 in 2017. However, tax units in this expanded income range were only

eligible for fairly small benefit amounts as shown in Figure 1.

The number of individuals claiming the CalEITC was substantially lower in 2016 than in 2017. There were 376,000 CalEITC claims in 2016 compared to 1,463,000 claims in 2017 (Davis and White, 2019). We see a similar jump in claimants in our merged sample of CalFresh filers. As reported in Appendix Table 13, 172,092 CalFresh-related tax units claimed the CalEITC in 2016, compared to 485,562 in 2017. Around one-third of the increase seems to be due to the expansion in the eligible range of earnings. Another one-quarter of the increase is due to filers claiming using self-employment income.[32]

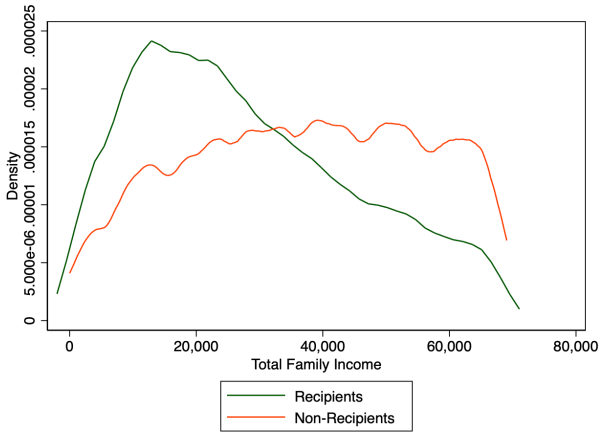
We present take-up estimates for 2016 separately by the same three populations considered above: claimers of the federal EITC (Appendix Table 14), federal EITC non-claimers (Appendix Table 15), and non-filers (Appendix Table 16). Appendix Table 17 aggregates estimates from the three tables. Overall, the take-up rate among CalFresh recipients in 2016 was 43%, over ten percentage points lower than in 2017. The number of eligible tax units more than doubled between 2016 and 2017 both for filers who claimed the federal EITC and those who did not. However, the share of eligible tax units who did not claim the credit was unchanged. Among non-filers, the count of eligible imputed tax returns also increased, but by a smaller amount, from 171,720 in 2016 to 268,113 in 2017. Finally, the mean CalEITC credit amount for 2016 was $295, over $100 higher than the mean amount for 2017. This is explained by the presence of more CalEITC recipients earning lower credit amounts on the expanded phase-out region in 2017.

The number of eligible filers grew faster over this period than the number of non-filers we impute to be eligible. The significant increase in the number of claimers with self-employment income suggests many more households became eligible for the CalEITC due to their self-employment income, rather than ineligible due to the combination of self-employment and wage income. Unless all of the newly eligible households filed and claimed, this suggests that there may be a number of non-filers with only self-employment income who could also be eligible which our estimate does not include. If that's the case, our estimate for the take-up rate in 2017 might be an upper bound.

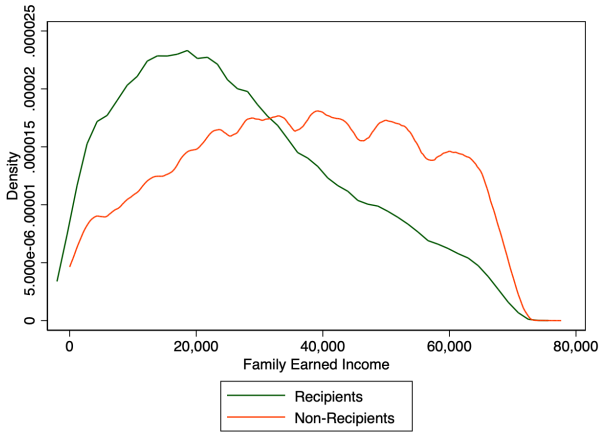
[32] We estimate this number by counting the number of claimants in 2017 for whom the majority of their earnings came from self-employment and whose total earnings was below $15,000 in order to avoid double-counting with the group above.

Appendix Figure 1: Distribution of income sources for SNAP and non-SNAP families in California, 2017 5-year ACS sample

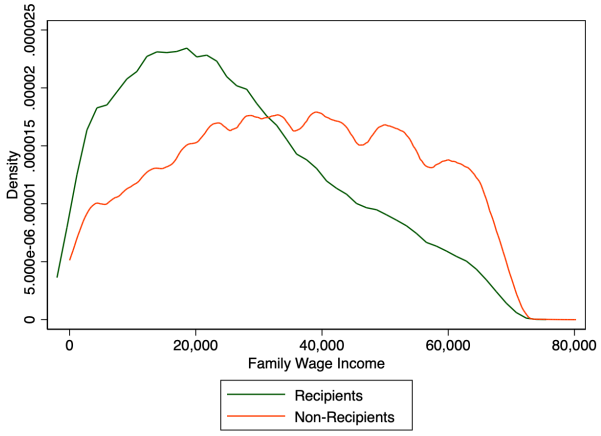
(a) Total family income



(b) Family earned income

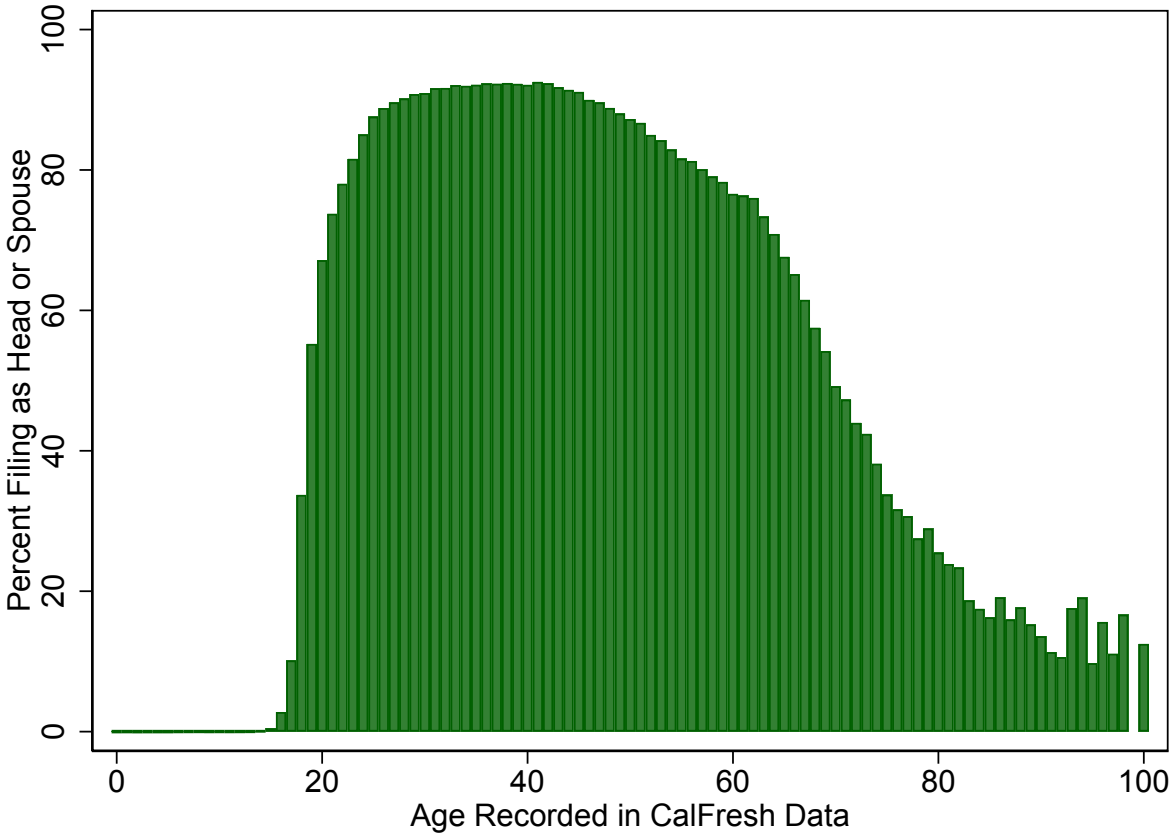


(c) Family wage income



Notes. Constructed using the 2017 ACS 5-Year Sample, restricted to 18–64-year-olds with total family income between $0 and $69,063, using ACS person-weights, excluding individuals in group quarters. Each figure excludes individuals with negative income of the graphed type.

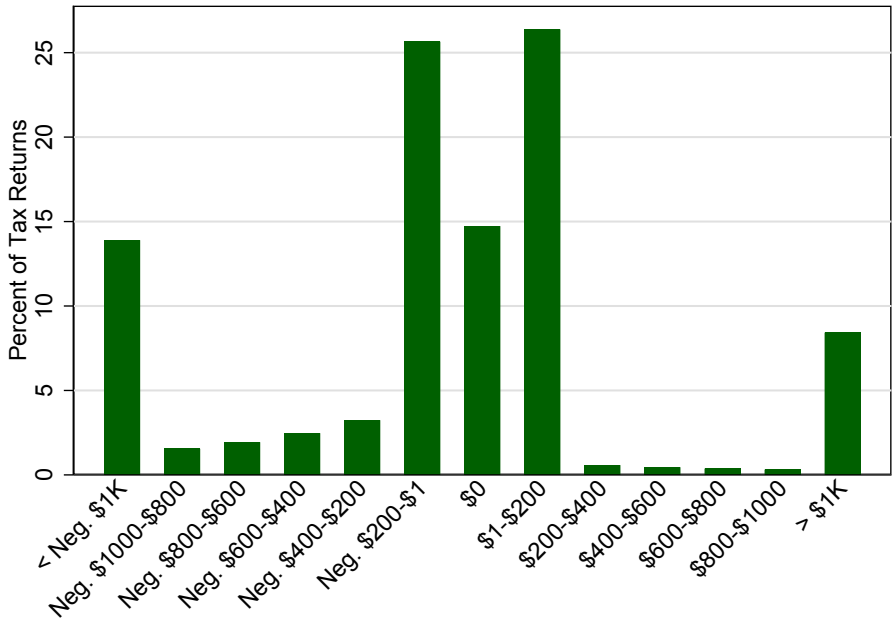
Appendix Figure 2: Distribution of ages among dependents on CalFresh tax units, TY 2017



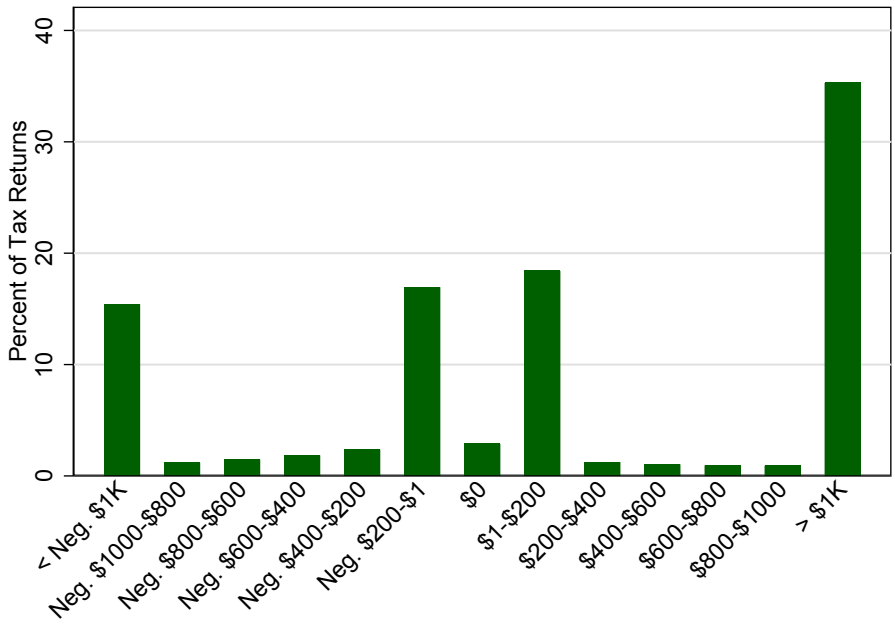
Notes. Universe is all individuals on tax year 2017 returns who also enrolled in CalFresh that year. We group ages into one year bins, according to enrollees' reported date of birth, and report the share in each bin who appear on their tax return as head/spouse as opposed to dependent.

Appendix Figure 3: Distribution of differences between total EDD earnings and reported California Income among tax units with all primary filers enrolled in CalFresh, TY 2017

(a) 540 wages versus EDD wages

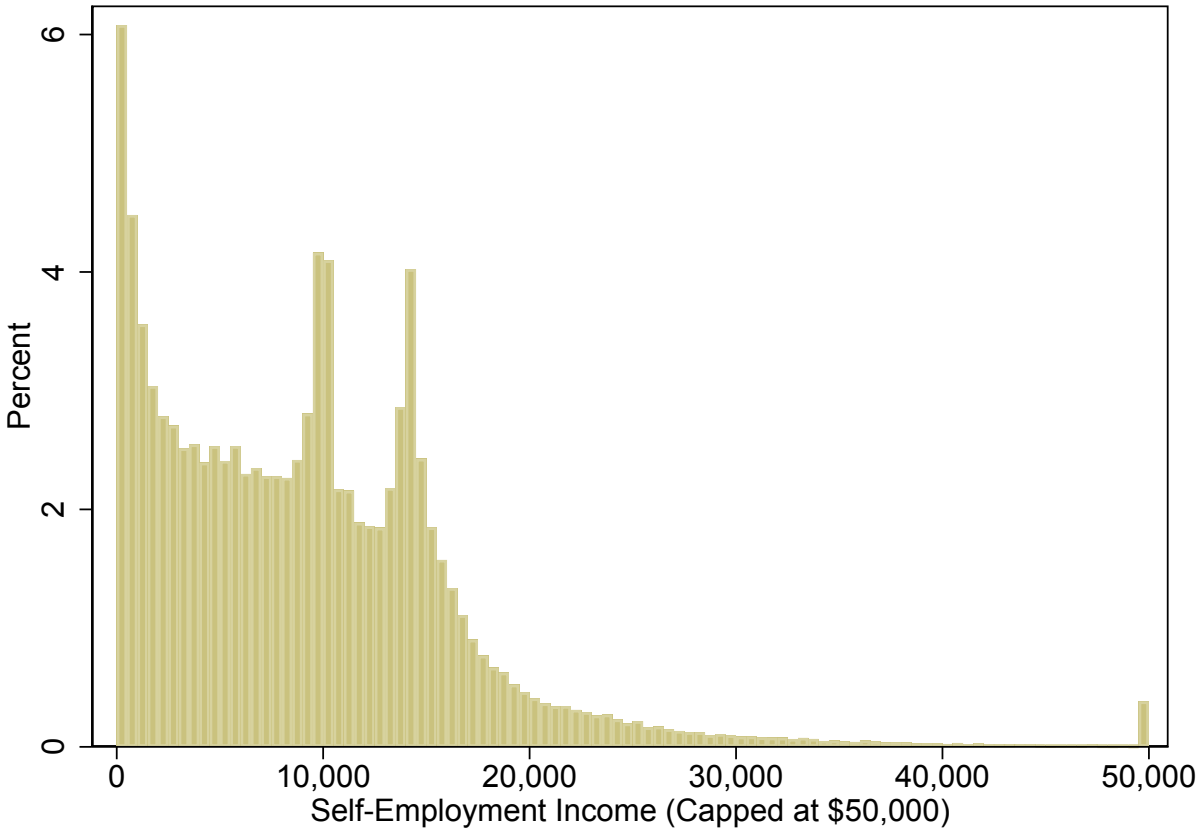


(b) AGI versus EDD wages



Notes. Universe is all tax returns with all primary filers are enrolled in CalFresh. EDD wages are total of all 2017 quarterly wage earnings for head and spouse (if present) on return.

Appendix Figure 4: Distribution of self-employment income among CalFresh-enrolled tax units with positive self-employment income, TY 2017



Notes. Universe is all tax returns with all primary filers enrolled in CalFresh and who report positive self-employment income.

Appendix Table 1: Comparison of means between SNAP recipient and non-recipient families

	Non-recipients	Recipients	Total
Sample Size	9,677,982	2,345,127	12,023,109
Family composition			
Family size	2.8	4.1	3.0
Count of adults	2.0	2.3	2.0
Count of children	0.6	1.6	0.8
Count of elderly	0.2	0.1	0.2
Demographics (percent)			
Married	38%	34%	37%
Non-white	41%	44%	42%
Black	6%	10%	7%
Hispanic	44%	58%	46%
Income (median)			
Total family income	$35,886	$24,000	$33,094
Family earned income	$31,566	$19,649	$29,474
Family wage income	$28,957	$17,581	$25,885
Family investment income	$0	$0	$0
Income (percent)			
Positive investment income	7%	2%	6%
Investment income over the EITC cap	0%	0%	0%
$0 total family income	5%	4%	5%
$0 family earned income	14%	17%	14%
$0 family wage income	19%	22%	20%

Notes. Constructed using the 2017 ACS 5-Year Sample. Restricted to 18–64-year-olds with total family income between $0 and $69,063, constructed using ACS person-weights, excluding individuals in group quarters.

Appendix Table 2: Comparison of means between e-filers and paper / web filers

	Tax Preparation Method		
	Paper or Web Filer	**E-Filer**	**Total**
Number of Tax Units	2,104,526	14,521,178	16,625,704
Filing status and number of dependents			
Share single	50%	47%	48%
Share head of household	12%	15%	14%
Share married filing jointly	36%	37%	37%
Mean number of dependents	0.6	0.7	0.7
EITC participation			
Share claiming CalEITC	8%	9%	9%
Mean CalEITC amount	$224	$266	$262
Share CalEITC eligible	13%	13%	13%
Share claiming CalEITC among eligible	41%	54%	52%
Income			
Mean Federal AGI	$64,855	$73,657	$72,543
Median Federal AGI	$38,218	$43,069	$42,478
Mean CA wages	$47,108	$55,522	$54,457
Median CA wages	$24,863	$31,644	$30867
Other characteristics			
Share on CalFresh	5%	7%	7%
Filed with Paid Preparer	26%	68%	62%
Self-Prepared	73%	30%	36%
Filed with VITA	0%	2%	2%
Share with ITIN on Return	6%	6%	6%

Note: Restricted to head filers on tax returns. For our EITC imputations, we use California wages and Federal AGI reported on the F540 to represent earned income and adjusted gross income, and we assume that no tax unit has investment income. In the income statistics reported above, we top code both California wages and Federal AGI at the 99th percentile (excluding $0s) to avoid the impact of potentially erroneous outliers. Including those outliers does not impact the median amounts but does increase the mean Federal AGI amount to $87,614 for e-filers and $155,117 for paper filers, and the mean California Wage amount to $62,631 for e-filers and $396,269,984 for paper-filers.

Appendix Table 3: Comparison of means between e-filers and paper filers in 2017, among tax units that claim CalEITC

	Tax Preparation Method		
	Paper or web filer	**E-filer**	**Total**
Count of tax units	143,236	1,300,779	1,444,015
Income information			
Mean AGI	$9,859	$10,982	$10,870
Mean earned income	$9,781	$10,804	$10,702
Mean wage income	$6,872	$8,319	$8,175
Mean investment income	$47	$39	$40
Other characteristics			
Mean number of qualifying children	0.7	0.8	0.8
EITC information			
Mean CalEITC amount (if positive)	$196	$239	$235
Federal EITC amount (if positive)	$1,881	$2,232	$2,197

Notes: Restricted to head filers on tax returns with positive CalEITC amounts reported on the Schedule 3514.

Appendix Table 4: Estimated EITC eligibility among tax units with a head or spouse with an ITIN

	CalEITC	Federal EITC	Either EITC
Count of tax units	611,570	611,570	611,570
Without SSN test			
Single filers			
% eligible	36%	77%	77%
Number eligible	108,637	232,432	233,328
Mean EITC amount	$288	$3,009	$3,132
Total EITC amount	$31,283,585	699,493,912	730,777,497
Married filing jointly, one filer has ITIN			
% eligible	10%	43%	43%
Number eligible	15,224	63,872	64,048
Mean EITC amount	$313	$2,870	$2,937
Total EITC amount	$4,762,703	$183,333,465	$188,096,168
Married filing jointly, both filers have ITIN			
% eligible	18%	71%	71%
Number eligible	29,091	113,342	113,469
Mean EITC amount	$369	$3,435	$3,526
Total EITC amount	$10,740,907	389,321,605	$400,062,512
Total			
% eligible	31%	70%	70%
Number eligible	143,952	409,646	410,845
Mean EITC amount	$325	$3,105	$3,210
Total EITC amount	$46,787,195	$1,272,148,982	$1,318,936,177

Notes. Universe is e-filed tax returns containing a head or spouse (if present) with an ITIN. We report four statistics (share and number eligible for either federal or state EITC, plus the mean and total amounts claimable) for three populations (single filers, married joint filers in which one spouse has an ITIN, and married joint filers in which both have an ITIN).

Appendix Table 5: Estimated EITC eligibility among tax units containing only a dependent with an ITIN

	CalEITC	Federal EITC	Either EITC
Count of tax units	207,809	207,809	207,809
With SSN Test			
Mean count of QC	1	0.8	1
% eligible	8%	22%	23%
Mean EITC amount	$13	$470	$2,168
Total EITC amount	$2,700,356	$97,652,210	$100,352,566
Without SSN Test			
Mean count of QC	2	2.2	2
% eligible	13%	50%	50%
Mean EITC amount	$47	$1,457	$3,025
Total EITC amount	$9,703,655	$302,808,748	312,512,403

Notes. Universe is e-filed tax returns containing a dependent with an ITIN, where both head and spouse (if present) have a valid SSN. Panel A reports average number of QCs, share eligible for either federal or state EITC, and the mean and total amounts claimed. Panel B reports same statistics assuming that dependents with an ITIN could be qualifying children. Dependents must still pass age test, but all are assumed to pass residency test.

Appendix Table 6: Relationship between number of inferred qualifying children from federal EITC claim and number of qualifying children reported on Schedule 3514 and total number of dependents, among CalFresh households claiming federal EITC and positive CalEITC amounts

	Number of qualifying children inferred from Federal EITC			
	0	1	2	3+
Dependents claimed on tax return				
0	95.4%	0.5%	0.2%	0.2%
1	3.9%	93.5%	0.6%	0.2%
2	0.6%	5.4%	94.4%	0.5%
3+	0.1%	0.6%	4.8%	99.1%
Number of qualifying children reported on Sch 3514				
0	99.9%	5.0%	4.7%	3.8%
1	0.1%	94.8%	2.5%	1.6%
2	0.0%	0.1%	92.6%	1.5%
3+	0.0%	0.0%	0.2%	93.1%
N	119,621	181,479	124,875	66,314

Notes. Universe is e-filed tax returns linked to at least one CalFresh participant and a non-zero CalEITC claim on the Schedule 3514. Columns represent number of EITC QCs inferred using process described in Section 4.1. Rows in Panel A are dependents reported on tax return. Rows in Panel B are number of QCs reported on Schedule 3514 for purposes of CalEITC claim. Cells represent column percentages.

Appendix Table 7: Relationship between number of inferred qualifying children from CalFresh records and number of qualifying children reported on Schedule 3514 and total number of dependents, among CalFresh households claiming federal EITC and positive CalEITC amounts

	Number of qualifying children inferred from CalFresh records			
	0	1	2	3+
Dependents claimed on tax return				
0	60.1%	0.0%	0.0%	0.0%
1	27.2%	75.6%	0.0%	0.0%
2	7.6%	16.3%	79.6%	1.6%
3+	2.0%	2.9%	14.0%	93.4%
Number of qualifying children reported on Sch 3514				
0	64.9%	4.3%	4.0%	3.5%
1	25.5%	76.5%	2.4%	1.5%
2	7.2%	15.3%	79.9%	1.9%
3+	1.9%	2.7%	13.2%	93.0%
N	185,820	165,375	96,723	44,371

Notes. Universe is e-filed tax returns linked to at least one CalFresh participant and a non-zero CalEITC claim on the Schedule 3514. Columns represent number of EITC QCs inferred using process described in Section 4.2. Rows in Panel A are dependents reported on tax return. Rows in Panel B are number of QCs reported on Schedule 3514 for purposes of CalEITC claim.

Appendix Table 8: Relationship between number of inferred qualifying children from CalFresh records (in which we disqualify dependents who are not observed in CalFresh records) and the number of qualifying children reported on Schedule 3514 and total number of dependents, among CalFresh households claiming federal EITC and positive CalEITC amounts

	Number of qualifying children inferred from CalFresh records			
	0	1	2	3+
Dependents claimed on tax return				
0	24.5%	0.0%	0.0%	0.0%
1	35.1%	53.2%	0.0%	0.0%
2	25.7%	31.7%	66.3%	0.0%
3+	14.7%	15.2%	33.7%	100%
Number of qualifying children reported on Sch 3514				
0	28.7%	8.3%	8.7%	6.2%
1	35.3%	54.3%	15.2%	7.0%
2	23.5%	26.2%	53.7%	10.8%
3+	12.5%	11.2%	22.4%	75.9%
N	471,715	17,866	2,339	369

Notes. Universe is e-filed tax returns linked to at least one CalFresh participant and a non-zero CalEITC claim on the Schedule 3514. Columns represent number of EITC QCs inferred using process described in Section 4.2, amended to disqualify any dependents who cannot be matched to our CalFresh records. Rows in Panel A are dependents reported on tax return. Rows in Panel B are number of QCs reported on Schedule 3514 for purposes of CalEITC claim.

Appendix Table 9: Simulated CalEITC eligibility using CalFresh-based QC imputation process among tax units with CalFresh-enrolled head or spouse and who claimed the federal EITC

	Tax units	CalEITC eligible		Eligible non-claimants		
	Count	Count	Share	Count	Share	Amount
For all filers						
Total	721,354	100%	68%	8%	37,098	$219
By filing status and number of qualifying children						
Single						
No QCs	184,853	153,411	83%	7,961	5%	$60
1 QC	196,874	151,583	77%	12,459	8%	$252
2 QCs	119,578	82,001	69%	7,399	9%	$427
3+ QCs	44,546	27,545	62%	2,678	10%	$402
Married						
No QCs	31,181	17,008	55%	822	5%	$37
1 QC	41,634	21,566	52%	1,955	9%	$131
2 QCs	48,739	19,845	41%	1,825	9%	$201
3+ QCs	29,375	9,607	33%	920	10%	$173
By total earnings (thousands)						
$0-$5	80,007	75,243	94%	2,465	3%	$382
$5-$10	135,092	133,232	99%	6,215	5%	$568
$10-$15	172,809	167,234	97%	11,278	7%	$229
$15-$20	110,678	86,419	78%	12,163	14%	$92
$20-$25	82,424	30,377	37%	4,977	16%	$10
$25-$30	58,817	8%	0%	0	.	.
$30+	81,527	11%	0%	0	.	.

Notes. Universe is e-filed tax returns linked to at least one CalFresh participant that included a claim for a non-zero federal EITC. The number of qualifying children for each tax unit is calculated using the process described in Section 4.2, as opposed to Section 4.1. To observe the effect of this alternative QC inference process, results repoted here can be compared against those reported in Table 5. Column 1 reports the total number of tax units that meet those criteria. Column 2 reports the count of those tax units that were eligible for the California EITC. Column 3 reports what share of all returns were eligible for the CalEITC. Column 5 reports the number of eligible returns that did not claim the CalEITC, and Column 6 reports the share of eligible units that did not claim. Column 7 reports the average imputed amount among those non-claimers for each cell.

Appendix Table 10: Distribution of children across tax returns by whether there is an adult in their CalFresh reference case

	All Children			Children w/o Adult in Reference Case			
	On tax return	Not on tax return	Total	On tax return w/ ITIN filer	On tax return w/o ITIN filer	Not on tax return	Total
All children							
Count	1,744,853	563,835	2,308,688	240,526	251,383	194,826	686,735
Percent	76%	24%	100%	35%	37%	28%	100%
Children enrolled in single CalFresh case							
Count	1,670,111	537,907	2,208,018	232,313	245,192	189,275	666,780
Percent	76%	24%	100%	35%	37%	28%	100%
Children enrolled in single CalFresh case for all 12 months							
Count	936,549	290,794	1,227,343	139,903	155,353	116,637	411,893
Percent	76%	24%	100%	34%	38%	28%	100%

Notes. Columns 1 and 2 report the number of all children enrolled in CalFresh in 2017 who appear on a tax return (Col 1) or not (Col 2). Columns 4 to 6 report the same statistic, but zoom in on the population of children for whom there was no adult on their CalFresh reference case. Among these children who also appear on a tax return, we report separately the number who appear on a return with an ITIN filer and those who do not.

Appendix Table 11: Distribution of SNAP-enrolled children who appear on a state tax return by whether adults in their SNAP reference case appear on their return

	All adults	Some adults	No adults		Total
			No ITINs	1+ ITIN	
All children					
Count	807,070	164,803	255,492	25,579	1,252,944
Percent	64%	13%	20%	2%	100%
Children enrolled in single CalFresh case					
Count	773,007	157,416	237,698	24,485	1,192,606
Percent	65%	13%	20%	2%	100%
Children enrolled in single CalFresh case all 12 months					
Count	399,169	83,464	144,064	14,596	641,293
Percent	62%	13%	22%	2%	100%
Children enrolled in single CalFresh case including 2+ adults in reference case					
Count	209,448	157,416	36,698	2,757	406,319
Percent	52%	39%	9%	1%	100%
Children enrolled in single CalFresh case including at least one adult 25-34 years older than child					
Count	356,239	86,223	92,512	4,794	539,768
Percent	66%	16%	17%	1%	100%

Notes. The universe in Appendix Table 11 is all SNAP-enrolled children in California in 2017 who appeared on a state tax return and who had at least one adult on their reference case. Column 1 reports the number and share of these children who appear on a tax return containing all adults who appeared on the child's reference case. Column 2 reports the number and share of these children who appear on a tax return containing at least one but not all adults who appeared on the child's reference case. Columns 3 and 4 report the number and share of these children who appear on a tax return containing no adults who appeared on the child's reference case, including cases containing no adults, broken out by whether the return contained an adult who filed with an ITIN or not.

Appendix Table 12: Distribution of SNAP-enrolled children who did not appear on a state tax return by whether adults in their SNAP reference case appear on any return

	No Adults	Some Adults	All Adults	Total
All children				
Count	282,875	19,220	66,914	369,009
Percent	77%	5%	18%	100%
Children enrolled in single CalFresh case				
Count	266,109	18,281	64,242	348,632
Percent	76%	5%	18%	100%
Children enrolled in single CalFresh case all 12 months				
Count	136,688	10,068	27,401	174,157
Percent	78%	6%	16%	100%
Children enrolled in single CalFresh case including 2+ adults in reference case				
Count	73,532	18,281	14,778	106,591
Percent	69%	17%	14%	100%
Children enrolled in single CalFresh case including at least one adult 25-34 years older than child				
Count	117,295	9,307	27,642	154,244
Percent	76%	6%	18%	100%

Notes. The universe in Appendix Table 12 is all SNAP-enrolled children in California in 2017 who did not appear on a state tax return in TY2017 and and who had at least one adult on their reference case. Column 1 reports the number and share of these children who did appear on a tax return containing no adults who appeared on the child's reference case. Column 2 reports the number and share of these children who appear on a tax return containing at least one but not all adults who appeared on the child's reference case. Column 3 reports the number and share of these children who did appear on a tax return and all adults who appeared on the child's reference case appeared on a return.

Appendix Table 13: Claiming of federal and state EITC among all CalFresh tax units in TY 2016

	Total tax units	Fed EITC Claimants			CalEITC Claimants		
	Count	Count	Share	Amount	Count	Share	Amount
For all filers							
Total	1,031,215	725,152	70%	$3097	172,092	17%	$738
By filing status and number of dependents							
Single							
No dependents	308,617	101,613	33%	$335	37,127	12%	$126
1 dependent	222,666	201,847	91%	$2649	52,001	23%	$663
2 dependents	163,374	154,660	95%	$4109	42,617	26%	$1086
3+ dependents	96,004	90,767	95%	$4569	20,079	21%	$1208
Married							
No dependents	37,236	15,327	41%	$392	2,699	7%	$116
1 dependent	48,393	38,048	79%	$2640	4,510	9%	$625
2 dependents	70,232	57,136	81%	$4094	7,263	10%	$1029
3+ dependents	84,693	65,754	78%	$4357	5,796	7%	$1157
By total earnings (thousands)							
$0-$5	242,254	177,319	73%	$2694	73,635	30%	$513
$5-$10	168,942	126,177	75%	$2363	73,100	43%	$987
$10-$15	163,746	122,940	75%	$3266	25,231	15%	$674
$15-$20	137,783	89,995	65%	$4378	57	0%	$890
$20-$25	109,370	77,633	71%	$4063	24	0%	$606
$25-$30	73,262	55,051	75%	$3436	0	0%	$0
$30+	135,858	76,037	56%	$2235	32	0%	$857

Notes. Universe is e-filed tax returns linked to at least one CalFresh participant. Column 1 reports the total number of tax units in each cell. Column 2 reports the count of those tax units that claimed the federal EITC. Column 3 reports what share of all returns in each cell claimed the federal EITC. Column 4 reports the average claimed amount of the federal EITC for each cell. Column 5 through 7 report the same statistics but for the state EITC.

Appendix Table 14: Eligibility and take-up of the CalEITC among CalFresh filers who claim the federal EITC in 2016

	Total tax units	CalEITC eligible		Eligible non-claimants		
	Count	Count	Share	Count	Share	Amount
For all filers						
Total	716,819	181,677	25%	13,805	8%	$684
By filing status and number of qualifying children						
Single						
0 QCs	103,365	37,303	36%	1,060	3%	$99
1 QC	211,184	56,738	27%	4,302	8%	$495
2 QCs	149,665	46,364	31%	4,665	10%	$853
3+ QCs	78,687	19,807	25%	2,010	10%	$957
Married						
0 QCs	16,135	2,790	17%	101	4%	$77
1 QCs	40,136	4,841	12%	347	7%	$442
2 QCs	57,247	7,888	14%	737	9%	$797
3+ QCs	60,400	5,946	10%	583	10%	$948
By total earnings (thousands)						
$0-$5	173,032	74,619	43%	3,757	5%	$466
$5-$10	124,585	78,370	63%	6,327	8%	$872
$10-$15	122,265	28,688	23%	3,721	13%	$583
$15-$20	89,407	0	0%	0	.	.
$20-$25	77,186	0	0%	0	.	.
$25-$30	54,794	0	0%	0	.	.
$30+	75,550	0	0%	0	.	.

Notes. Universe is e-filed tax returns linked to at least one primary filer who claimed CalFresh, and that included a claim for a non-zero federal EITC. The number of qualifying children in each tax unit was calculated using the process described in Section 4.1.1. Column 1 reports the total number of tax units that meet those criteria. Column 2 reports the count of those tax units that were eligible for the California EITC. Column 3 reports what share of all returns were eligible for the CalEITC. Column 5 reports the number of eligible returns that did not claim the CalEITC, and Column 6 reports the share of eligible units that did not claim. Column 7 reports the average imputed amount among those non-claimers for each cell.

Appendix Table 15: Eligibility and take-up of the CalEITC among CalFresh filers who did not claim the federal EITC in TY 2016

	Total tax units	CalEITC eligible		Eligible non-claimants		
	Count	Count	Share	Count	Share	Amount
For all filers						
Total	306,063	40,694	13%	39,836	98%	$127
By filing status and number of qualifying children						
Single						
0 QCs	222,884	39,617	18%	38,902	98%	$116
1 QC	11,135	354	3%	309	87%	$628
2 QCs	5,046	172	3%	158	92%	$1036
3+ QCs	1,998	29	1%	27	93%	$1289
Married						
0 QCs	28,989	371	1%	313	84%	$104
1 QC	10,786	56	1%	47	84%	$658
2 QCs	12,075	56	0%	48	86%	$989
3+ QCs	7,793	25	0%	22	88%	$1030
By total earnings (thousands)						
$0-5	64,935	28,611	44%	27,949	98%	$145
$5-10	42,765	11,974	28%	11,780	98%	$80
$10-15	40,806	109	0%	107	98%	$531
$15-20	47,788	0	0%	0	.	.
$20-25	31,737	0	0%	0	.	.
$25-30	18,211	0	0%	0	.	.
$30+	59,821	0	0%	0	.	.

Notes. Universe is e-filed tax returns linked to at least one primary filer who claimed CalFresh, and that did not include a federal EITC claim. The number of qualifying children in each tax unit was calculated using the process described in Section 4.2.1. Column 1 reports the total number of tax units that meet those criteria. Column 2 reports the count of those tax units that were eligible for the California EITC. Column 3 reports what share of all returns were eligible for the CalEITC. Column 5 reports the number of eligible returns that did not claim the CalEITC, and Column 6 reports the share of eligible units that did not claim. Column 7 reports the average imputed amount among those non-claimers for each cell.

Appendix Table 16: Eligibility and take-up of the CalEITC among non-filers in TY 2016

	Total tax units	CalEITC eligible		Eligible non-claimants		
	Count	Count	Share	Count	Share	Amount
For all filers						
Total	762,600	171,720	23%	171,720	100%	$303
By filing status and number of qualifying children						
Single						
0 QCs	538,831	115,863	22%	115,863	100%	$92
1 QC	93,453	25,140	27%	25,140	100%	$598
2 QCS	46,833	13,438	29%	13,438	100%	$964
3+ QCs	28,582	7,519	26%	7,519	100%	$1,071
Married						
0 QCs	34,243	4,733	14%	4,733	100%	$94
1 QC	7,096	1,554	22%	1,554	100%	$641
2 QCs	10,810	0	0%	0		
3+ QCs	6,837	1,739	25%	1,739	100%	$1,210
By total earnings (thousands)						
$0	398,424	0	0%	0	.	.
$1-$5	178,793	138,700	78%	138,700	100%	$215
$5-$10	66,537	28,911	43%	28,911	100%	$663
$10-$15	42,570	4,109	10%	4,109	100%	$737
$15-$20	28,071	0	0%	0	.	.
$20-$25	18,615	0	0%	0	.	.
$25-$30	10,647	0	0%	0	.	.
$30+	18,780	0	0%	0	.	.

Notes. Universe is simulated tax returns including only non-filing CalFresh participants. These tax units are constructed using the process described in Section 4.3. The number of qualifying children in each imputed tax unit was calculated using the process described in Section 4.3.5. Column 1 reports the total number of tax units that meet those criteria. Column 2 reports the count of those tax units that were eligible for the California EITC. Column 3 reports what share of all returns were eligible for the CalEITC. Column 5 reports the number of eligible returns that did not claim the CalEITC, and Column 6 reports the share of eligible units that did not claim. Column 7 reports the average imputed amount among those non-claimers for each cell.

Appendix Table 17: Summing up CalEITC take-up among CalFresh recipients in 2016

	Total	CalEITC eligible		Eligible non-claimants			
	Count	Count	Share	Count	Share	Mean Amount	Total Amount
Filers							
Fed EITC claimants	716,819	181,677	25%	13,805	8%	$684	9,436,792
Non Fed EITC claimants	306,063	40,694	13%	39,836	98%	$127	5,039,291
Non-Filers							
All	762,600	171,720	23%	171,720	100%	$303	52,070,267
Total	1,785,482	394,091	22%	225,361	57%	$295	66,546,350

Notes. Appendix Table 17 compiles information from earlier tables; see those tables for details. The addition is the final column which reports the total unclaimed dollars for each population. Cells represent column percentages.