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

John Iselin
UMD College Park

Taylor Mackay
UC Irvine

Matthew Unrath*
UC Berkeley

July 13, 2021
Click here for the latest version.

Abstract

The Earned Income Tax Credit (EITC) is the country's largest cash-based means-tested transfer program, 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 but do not claim it. We find that nearly half a million households who receive SNAP benefits and who were eligible for the state EITC in 2017 did not receive the credit. This includes approximately 42,000 eligible households who claimed the federal EITC but not the state credit; 110,000 eligible households who filed a state tax return but did not claim the state credit; and 290,000 eligible households who did not file a state tax return. The corresponding take-up rate for the CalEITC among eligible SNAP-enrolled households was 53%. Altogether, these households left on the table a total of $75 million in state EITC funds. If received, these credits would have increased incomes among these households by 2.6% and increased total state EITC outlays by 38.8%.

*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. 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, 2016; 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]). The EITC's take-up rate exceeds that of many other means-tested programs. Nevertheless, 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. 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.[2] As such, there exists no estimates of take-up of the California EITC or any other state supplements.

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. Importantly, 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

[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.

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 440,000 households who received CalFresh benefits and who appeared eligible for the state EITC did not receive the credit in 2017. This includes roughly 42,000 who claimed the federal EITC but not the state credit; 110,000 who filed a state tax return but did not claim the state credit; and 290,000 who did not file a state tax return. The corresponding take-up rate for these households was 53%. Altogether, these households left on the table a total of almost $76 million in state EITC funds. If received, these credits would have increased incomes among these households by 2.6% and increased total state EITC outlays by 38.8%.

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 $80 on average. Still, we show that tens of thousands of tax units with dependents and non-filing households with children failed to claim the CalEITC. Unclaimed amounts for these households was much higher: an average of $259 for filers and $650 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. 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.

The paper proceeds as follows. In Section 2, we describe the federal and California EITC more fully and summarize previous work estimating EITC participation and potential explanations for incomplete take-up. Section 3 describes the unique linked data that make our project possible. Section 4 describes our methods for simulating EITC eligibility for three populations of CalFresh recipients and presents estimates of participation for each population. Section 5 presents additional results and estimates of take-up for various subgroups. Section 6 discusses the relevance of our finding to strategies aimed at increasing take-up. Section 7 concludes.

2 Background on the EITC

2.1 Structure of the federal and state credits

EITC eligibility depends on a tax unit's filing status, earned income, and number of qualifying children. At low earnings levels, the EITC amount increases with each dollar of earnings until it reaches a maximum value. The credit is then stable across a range of earnings, before eventually beginning to decline as earnings rise. This structure encourages labor force participation for low earners, who cannot receive the credit unless they 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, 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, is close to $6,000. Filers with earnings above or below this range qualify for smaller credits, so long as their earnings are positive and do not exceed $45,007.[3]

In 2015, California joined 27 other states and the District of Columbia in supplementing the federal EITC. Most states with supplements simply offer a partial match for the federal credit, but California adopted its own schedule. The gold area in Figure 1 shows the California schedule in tax year 2017. California's schedule is targeted towards households with the lowest earnings. The largest credit is available to families with earnings around $7,500, though households with earnings up to $25,000 can still qualify for small credits. In 2017, approximately 1.5 million tax units claimed the CalEITC, receiving a total of $351 million (Davis and White, 2019).

EITC eligibility depends on the tax unit's Adjusted Gross Income (AGI), investment income, and earned income. A tax unit must have investment income below a set threshold ($3,450 in 2017), and both the tax unit's AGI and earned income must be in the eligible range. The calculation of the actual EITC amount generally depends on the tax unit's earned income and in some cases its AGI.[4]

The EITC uses a different count of children than do other components of the tax code. An EITC qualifying child must pass have a valid SSN; must be under 19, or under 24 if a full-time student; 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, as well as overly withheld income taxes being returned, creates an incentive to do so. Prior to the introduction of the CalEITC, there was little reason for many families to file state returns, even if they filed federal returns in order to claim the federal credit. This possibility raises concerns that many eligible families still might not file state returns or, if they do, not know to claim the CalEITC. This concern is relevant to all state supplements, but the the concentration of CalEITC benefits at very low earnings levels heightens these concerns in this setting.

2.2 Estimates of take-up rates

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

A number of studies use survey data to simulate eligibility. 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 million to 2.0 million eligible taxpayers failed to claim the credit. In 2001, using multiple household surveys, the Government Accountability Office (GAO) estimated EITC

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

[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 using their earned income and their AGI, and claim the lesser of the two amounts.

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 to be nearly 90% among legally obligated filers for tax year 1988 , but only 30 to 40% among those not legally obligated to file. Among all eligible households in that year, they estimated take-up to be roughly 70%. 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, which is prone to error (Meyer and Mittag, 2015).[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 identified 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 data from Florida, linked with IRS records, to validate EITC eligibility. The authors find that SNAP data could be useful in evaluating which children are qualified children under the EITC, improve enforcement of improper claiming, and to aid with outreach efforts.

[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.

2.3 Explanations for incomplete take-up

In addition to the narrower 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, 2002; Smeeding, Phillips, and O'Connor, 2000; Maag et al., 2015). Awareness also appears to vary across different communities (Maag et al., 2005; 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, Manoli, Schafer, and Sebastiani (2016) studies the effect of sending mailings to eligible, non-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, Homonoff, Javaid, and Schafer (2021) find that a single letter from the IRS sent to non-filers increased tax filing a .7 percentage points.

There is limited evidence about 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 calendar year 2017.

Our first database contains individual- and case-level CalFresh enrollment records. In 2017,

5.6 million unique individuals enrolled in CalFresh, representing 50.5 million person-months. Forty-eight percent of individuals ever enrolled in CalFresh during 2017 were enrolled for all 12 months, while the other 52% were enrolled for only part of the year. CalFresh cases are defined as groups of people who prepare meals together, and an individual can participate in different cases over the course of a year. Our sample includes 2.9 million distinct cases, and 95% of individuals in our sample are associated with only one case during 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 provide key information for determining EITC eligibility that tends to be unavailable in survey and tax data. 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 test, 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.[6]

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.[7]

On the FTB side, our linkage is limited to individuals who appear on a state, as opposed to a federal, 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 these households, we can observe a federal return, a federal EITC claim if there is one, and whether dependents are likely qualifying

[6]Given the observed lower incomes of paper filers, and the fact that those who file paper returns are more likely to self-file as opposed to using either paid preparers or volunteer-run tax assistance sites, 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.

[7]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. 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 identifiers.

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 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 file returns. A central piece of our analysis will involve constructing tax units for CalFresh recipients who do not appear on a state return. Overall, of the 5.6 million individuals receiving CalFresh in 2017, 3.1 million (or 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.

One important sample restriction is that we exclude from our analysis the approximately 1 million dependents who were claimed on a tax return in which neither the head nor the spouse was a CalFresh recipient. Since we do not observe these filers' casefiles 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 these other adults. After this restriction and limiting ourselves only to e-filers, our primary sample includes 2.4 million CalFresh recipients who appear on a tax return in our sample. Of all CalFresh recipients who appear on a state tax return in 2017, 58% (32% of all recipients) include claims for the federal EITC and 38% (21% of the total) include claims for the state credit. Of all 766,000 tax units in 2017 containing a CalFresh-enrolled head or spouse, 69% claimed the federal EITC and 45% claimed the CalEITC (Table 3). Their average federal and state EITC claims were $3,081 and $396, respectively.

The Californians who elect to 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 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.

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 file a state return and claim the federal EITC; (2) CalFresh participants who file a state return but do not claim the federal EITC; and (3) CalFresh participants who do not file a state return.

For those who claim the federal EITC, we can observe nearly all relevant income,[8] 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 children's age and residency test. For those who do not file, we transform CalFresh cases into likely tax units, and measure eligibility using the casefiles, recipients' ages, and merged earnings records.

In the following three sections, we describe these processes more fully and present estimates of non-claiming for each population. Table 4 provides an overview of the information we use to estimate each component of eligibility across these three populations.

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 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) who reside with the filer for at least six months in the tax year and are near relations or adopted or foster child of the filer. Federal claimants do report their count of QCs on the Schedule EIC, but we do not observe this information as part of the federal return. 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.

We present two pieces of evidence that help 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

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

who claimed the CalEITC (Appendix Table 6).[9] 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 F 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 state EITC in tax year 2017. Of the approximately 755,000 e-filed state tax returns containing a head or spouse who enrolled in CalFresh and a federal EITC claim, 71% were eligible for the CalEITC (Table 6). Of these eligible units, 8%, or about 42,000 CalFresh tax units, did not claim the CalEITC. These tax units received an average federal EITC benefit of $3,794, but did not claim an additional $233 from the CalEITC. The forgone CalEITC amount for this group totaled nearly $9.8 million, and if received, would have raised annual incomes in this population by 1.6%.[10]

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 dependents forwent an average state credit of $81, and households with dependents forwent an average of $262 to $306, depending on the number of dependents in the household.

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 $571), 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 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.

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

[9] 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.

[10] We also identify a small number of tax units (0.1% 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 $267.

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.[11] As mentioned above, QCs must be under 19 (under 24 for full-time students), must reside with the 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. Since we cannot verify this information for dependents who do not appear in our CalFresh records, we assume these dependents are qualifying children.

We are unable to identify students, so we only allow dependents age 18 and under to be potential QCs, provided that they satisfy the residency test.[12] 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.[13]

To simulate the residency test, we observe the number of months that dependents appeared on the same CalFresh case as the filer and/or spouse on their return. 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 the head or spouse as a proxy for residential arrangements. A dependent who shared a CalFresh case with the tax unit head or spouse 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, months in which neither was on CalFresh are allowed to count toward this six-month threshold, on the assumption that the child and parent lived together in this period as well. 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 throughout the year and meets the residency test. The effect of these rules is to disqualify dependents as possible QCs if we can observe that they were in different CalFresh cases than the head and/or spouse on their return for more than six months in in the tax year.

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 rule 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. For 81% of tax units they are identical. In most of the remainder, we assign one fewer QC than the

[11] 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.

[12] Of the 167,240 dependents under age 24 on a return with a CalFresh-enrolled head or spouse and without a federal EITC claim, 4,364 (2.6%) are between the ages of 19 and 24.

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

number of dependents. Of the 43,221 dependents who we deem not to be qualifying children, 29,780 (18% of all dependents) fail the residency test, 13,043 (8% of all dependents) fail the age test, and 11,203 (7% 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 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.

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 disqualify 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.

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. Nearly 350,000 tax units contain a head or spouse who was enrolled in CalFresh in 2017 and did not claim the federal EITC. Of these units, nearly 113,000 (33%) were eligible for the CalEITC, and among these eligible units, nearly 110,000 (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 $84, and totaled nearly $9.3 million. If received, these benefits would have raised annual incomes in this population by 1.1%.

Single filers without QCs were more likely to be eligible than married filers or filers with QCs, which reflects their much lower average earnings. Eligible non-claimants in this group missed out on $82, on average. Though there are far fewer eligible non-claimants among tax units with QCs (just over 3,000), their average forgone credit was higher ($216). 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 do not appear on any California tax return. For these participants, we must construct simulated tax units from CalFresh casefiles and simulate eligibility using those casefiles and linked earnings records.

This is the most complex part of our entire 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 a part of the tax unit,

and individuals in the tax unit (e.g., dependents of non-custodial parents) not on the CalFresh case. We do not observe household relationships, which would allow us to associate filers with their spouses and children and vice versa.[14] 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.

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 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 income.

4.3.1 Reference cases

First, we assign each CalFresh recipient to a single CalFresh case. For the 95% of individuals who appear on only a single CalFresh case in 2017, this is straightforward. For the remaining 5%, we assign individuals to the case they appeared on most frequently. In the rare event of ties, we pick the most recent of the cases. Hereafter, we refer to the assigned CalFresh cases as the individual's *reference case*.

We construct simulated tax units from these reference cases. Some reference cases will include filers and non-filers. Following our general principle that we defer to submitted tax returns when possible, we always respect the tax unit reported on the submitted tax return and construct one or more tax units from the remaining non-filing members of the case, as detailed below.

Some reference cases may only contain 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 heads, spouses, and dependents

We assign each non-filing individual on a CalFresh reference case to be a filer, spouse, or dependent. We assign everyone under the age of 18 or over the age of 80 to be dependents.[15] For those aged between 18 and 80, we predict whether they should be a dependent or a 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 prior year status. Specifically, we use the CalFresh observations who do file returns to train a prediction model to classify observations as likely filers/spouses or dependents using these variables.

[14] 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.

[15] In 2017, 99.4% of child CalFresh recipients (aged 17 or younger) and 77.9% 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 not an obvious break point at 81, but the size of the population at that age is fairly small, making reliable imputation of an individual's tax status challenging.

We use cross-validation to select a threshold predicted probability of being a filer/spouse that maximizes out-of-sample accuracy. See Appendix D for a full description of this process.

4.3.3 Filing status

We then identify which of the recipients assigned to be filers would file as a single adult or married couple. This step is only relevant for the subset of reference cases with multiple adults, as adults 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 tax year 2016 returns, 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 individuals (using the distribution among tax filers) and if the age difference between them is between the 10th and 90th percentiles of within-couple age-differences (also among filers).[16] 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.[17] Overall, we impute that 13% of the heads or spouses in our non-filing sample are married.

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.[18] When there is only one such tax unit, the assignment is straightforward. When multiple candidate tax

[16]The 10th percentile of age among married heads and spouses 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.

[17]We anticipate that this set of marriage rules will underestimate the actual number of married filing jointly tax units that could exist within the non-filing population. Applied over all CalFresh-enrolled filers, our marriage assignment process results in our incorrectly assigning 39,000 individuals who filed a single return to a married couple, reflecting 4% of all single returns with a CalFresh-enrolled head. However, that same process incorrectly assigns the primary taxpayer of 106,000 married tax units to a single return, reflecting 39% of married tax units. We correctly set as married filing jointly only 59% of married tax units. In 90% of cases where we incorrectly label an individual as single, they are a member of mixed tax return, meaning we do not observe their spouse in the CalFresh records. If we assume that, among non-filing CalFresh recipients,a there is a similar share of mixed-CalFresh status married couples, then we should expect a similar undercount of married couples there as well. Under-simulating married units on EITC eligibility 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.

[18]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.

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.[19] In the case of adult dependents with multiple candidate tax units, we assign them to the unit with the highest earnings.

4.3.5 Qualifying children

We determine which of the dependents assigned to each imputed tax unit might be a qualifying child using the same procedure described in Section 4.2. We start with the list of dependents assigned to the the simulated tax unit and assess which dependents pass the age and residency test. Table 9 compares the number of dependents in each imputed tax unit to the number of simulated QCs. The share of dependents who appear to qualify as QCs is similar to those reported for filing non-claimers of the federal EITC (Table 7). Of the 197,375 imputed dependents on a reference case who we deem not to be qualifying children, 122,757 (7% of all imputed dependents) fail the residency test and 81,197 (8% of all imputed dependents) fail the age test.

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 earnings, like those from self-employment, for this population. This omission means we understate EITC-qualifying earnings. This might lead us to overestimate eligibility, for example, if non-filing households have both self-employment and wage earnings, and the combination makes some ineligible. 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. See Appendix E for more information about self-employment earnings among CalFresh tax filers.

[19] Our process of assigning imputed dependents to potential tax units results in 732,000 imputed dependents (out of a total of 1,067,000) being unassigned to any tax unit. Essentially all of these dependents were assigned to a reference case without an adult who was likely a filer. Thirty-six percent of these unassigned dependents are children who do not appear on a CalFresh case with any adult. We do not observe these children's parents or guardians in our CalFresh records; we cannot make any progress in understanding whether they might be eligible for the CalEITC, so we assume they are not. These children likely reflect CalFresh enrollees whose parents or guardians are undocumented adults, who are not eligible for CalFresh, the federal EITC, or, in 2017, the CalEITC. Parents of young children with ITINs became eligible for the CalEITC and certain other tax credits in 2020. We discuss these issues further in Appendix C.

Sixty-one percent of the unassigned dependents are adults. One concern is that our process of predicting whether CalFresh recipients are likely to be filers or dependents should have predicted these lone dependents to actually be single filers. However, 96.5% of these adults have no earnings, suggesting they would not file a return by themselves. We discuss this issue and other concerns about this prediction procedure in Appendix D. As a robustness check, we reclassify these remaining adult non-filers to be single filers and find that slightly under 15,000 would be eligible for the CalEITC (with a mean credit amount of $71).

We also do not observe investment income for non-filers, which might disqualify some families from the EITC. Few CalFresh recipients have investment income, however. According to the 2017 American Community Survey (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 6.9 and 2.0% for non-recipients, respectively). Among actual tax-units with a CalFresh recipient, only 1.3% had positive investment income in tax year 2017, and 0.6% were disqualified from receiving the EITC because their investment income was too high.

We discuss the accuracy of our earnings imputation further in Appendix E.

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 10 summarizes our estimates of eligibility for this population. We identify nearly 804,000 potential tax units from the non-filing CalFresh population, and we estimate that nearly 290,000 of these households (36%) were eligible for the CalEITC but did not claim it. Their average forgone credit was $196. The total forgone credit for this population was $56.6 million and, if received, would have increased annual income for this population by 3.9%.

Table 10 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 38% 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 of these in our data. 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 between $489 and $920. 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.

It is important to recall that if these eligible units represent actual non-filers, as opposed to federal-only filers, their forgone CalEITC amounts understate the possible 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 Bringing it all together

Table 11 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, and among eligible tax units, 53% claimed the credit. This means 441,575 (47%) eligible tax units did not claim the CalEITC, forgoing $172 on average and almost $75 million in total. Among filers, take-

up rates were fairly high: 78% of eligible tax units claimed the CalEITC. Nevertheless, roughly 141,000 eligible filers failed to claim the CalEITC. The average forgone credit among filers with children was $259, and for those without children was $81. Two-thirds of non-claimants were non-filers. Among non-filing CalFresh households, 36%, or over 289,000 tax units, were eligible for the CalEITC, forgoing an average of $192. The average forgone credit among non-filers with children was $650, and for those without children was $87.

5.2 Take-up of the federal EITC

We are not able to develop a comprehensive estimate of federal EITC claiming among CalFresh households, because some families may have filed federal returns including EITC claims but not state returns, and we do not observe those federal returns. 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 12 summarizes our estimates of eligibility and participation among all state returns with CalFresh enrollees. We split participation by those who claimed the state EITC and those who did not. Mirroring our count from Table 1, about 1.1 million returns contained a CalFresh filer or spouse. Among those, 78% were eligible to receive the federal EITC. Of those, 87% 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 109,000 tax units who were eligible did not claim the federal EITC, forgoing $423 on average.

5.3 By tax preparation method

An important question raised by our results is why any eligible filer would fail to claim the credit. We begin to explore this question by investigating how eligibility and participation varies across three methods of tax preparation we observe in our FTB records: self-prepared, prepared by a paid professional, and prepared through the Volunteer Income Tax Assistance (VITA) program (Table 13). Among the 42,000 tax units who claimed the federal EITC and were eligible for but did not claim the CalEITC, returns filed by paid preparers are over-represented. While returns filed by paid preparers make up only 60% our sample of tax returns who claimed the federal EITC, they make up 92% of non-claimants. 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 (33% for the federal EITC and 32% for the CalEITC) than among returns prepared by VITA or paid preparers (between 23 and 28%). 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. The slightly higher levels of non-claiming among self-preparers that did not claim the federal EITC could be explained by a lack of experience with or information about the tax system, but it is harder to pinpoint what might be causing a higher error rate among paid

preparers when it comes to unclaimed CalEITC dollars among tax units who claimed the federal EITC. Certain preparers might have limited experience with a new tax program. Fees charged by preparers might also dissuade eligible filers from claiming the state credit. More research on the role played by the tax preparation industry in accurately and efficiently distributing tax-based benefits to households is needed.

5.4 By race

Table 14 presents estimates of eligibility and participation by the race of individuals in our pool of CalFresh recipients.[20]

Presenting results at the tax unit level would require that we ascribe results only to the actual or simulated head of each tax unit. Instead, we present results at the individual level.[21] We also report our results separately for heads and spouses and for dependents.

Among filers, eligible Hispanic heads and spouses are less likely to claim the CalEITC than filers from other racial groups represented in our CalFresh sample. Take-up is about 5 percentage points lower for Hispanic filing heads and spouses than White heads and spouses for example. Hispanic dependents in eligible tax units are also 3 to 4 percentage points less likely to receive the CalEITC. Across other racial groups, share of eligible households not claiming the CalEITC is roughly equal (between 19% and 20%).

A greater share of Black, Hispanic, and Native Hawaiian and Pacific Islander (NHPI) non-filing imputed heads and spouses appear eligible for the CalEITC than other imputed heads and spouses. Forty-three percent of Black, 39% of Hispanic, and 36% of NHPI non-filing imputed heads and spouses appeared eligible for the CalEITC, compared to 30% of White heads and spouses. 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 (20% of heads/spouses and 22% of dependents).

[20] 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, the California Department of Social Services (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 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.

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

Altogether, non-participation rates are fairly comparable between Black, NHPI, and White households. Eligible Hispanic and Asian households are more likely to claim. Non-claiming is highest among American Indian and Alaska Native households.

6 Discussion

By disaggregating participation rates between 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, so that the state could send them their owed credit.

Increasing take-up among eligible filers who do not claim the federal credit would be slightly 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 owed credit amounts to those households. The benefits to taxpayers would surely exceed this outreach's cost. Successful outreach to these eligible filers might also help them claim the more valuable federal credit.

Those who do not file returns at all are the hardest to reach. We estimate that there were 290,000 potential tax units among the CalFresh enrollees who did not file a state return in 2017 but could have claimed the state EITC if they did, and that this group passed up an opportunity to receive $56 million from the California EITC. We expect that our imputation process yields reliable estimates of average eligibility and participation, but we are not able to ensure accuracy of this imputation at the individual tax unit level. These recipients would need to be engaged to confirm their family structure and earnings.

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. The CalFresh population tends to have lower incomes and larger families relative to the entire low-income population in the state, suggesting they may have higher rates of eligibility (See Appendix A). These households have already demonstrated their willingness to interact with government programs, meaning they may be more likely to claim benefits for which they are eligible. 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 require data-sharing between the state Department of Social Services and the IRS.

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 290,000 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.

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 providing researchers a roadmap to measuring 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, pp. 3489–3529.

Blumenthal, Marsha, Brian Erard, and Chih-Chin Ho (2005). "Participation and compliance with the earned income tax credit". *National Tax Journal*, pp. 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, pp. 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, pp. 1–8.

Currie, Janet (2006). "The take-up of social benefits". *Poverty, the Distribution of Income, and Public Policy*. Ed. by David Card Alan Auerbach and John Quigley. Russell Sage, pp. 80–148.

Davis, Charles and Evan White (2019). *Who receives the Earned Income Tax Credit in California?* Tech. rep. California Policy Lab.

Goldin, Jacob, Tatiana Homonoff, Rizwan Javaid, and Brenda Schafer (2021). *Tax Filing and Take-Up: Experimental Evidence on Tax Preparation Outreach and EITC Participation*. Tech. rep. National Bureau of Economic Research.

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

Guyton, John, Dayanand S Manoli, Brenda Schafer, and Michael Sebastiani (2016). *Reminders & recidivism: evidence from tax filing & eitc participation among low-income nonfilers*. Tech. rep. National Bureau of Economic Research.

Hoynes, Hilary and Jesse Rothstein (2016). "Tax policy toward low-income families". *NBER Working Paper*. URL: https://doi.org/10.3386/w22080.

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.

— (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, pp. 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*.

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, pp. 1351–1367.

Liebman, Jeffrey B (1998). "The impact of the earned income tax credit on incentives and income distribution". *Tax policy and the economy* 12, pp. 83–119.

Linos, Elizabeth, Allen Prohofsky, Aparna Ramesh, Jesse Rothstein, and Matt Unrath (2020). *Can Nudges Increase Take-up of the Earned Income Tax Credit?: Evidence from Multiple Field Experiments*. National Bureau of Economic Research.

Maag, Elaine et al. (2005). "Disparities in Knowledge of the EITC". *Tax Notes* 106.11, p. 1323.

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". *Washington, DC: Urban Institute*.

Meyer, Bruce D and Nikolas Mittag (2015). *Using Linked Survey and Administrative Data to Better Measure Income*.

Nichols, Austin and Jesse Rothstein (2016). "The earned income tax credit". *Economics of Means-Tested Transfer Programs in the United States, Volume 1*. University of Chicago Press, pp. 137–218.

Phillips, Katherin Ross (2001). *Who Knows About the Earned Income Tax Credit?* Tech. rep. The Urban Institute.

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

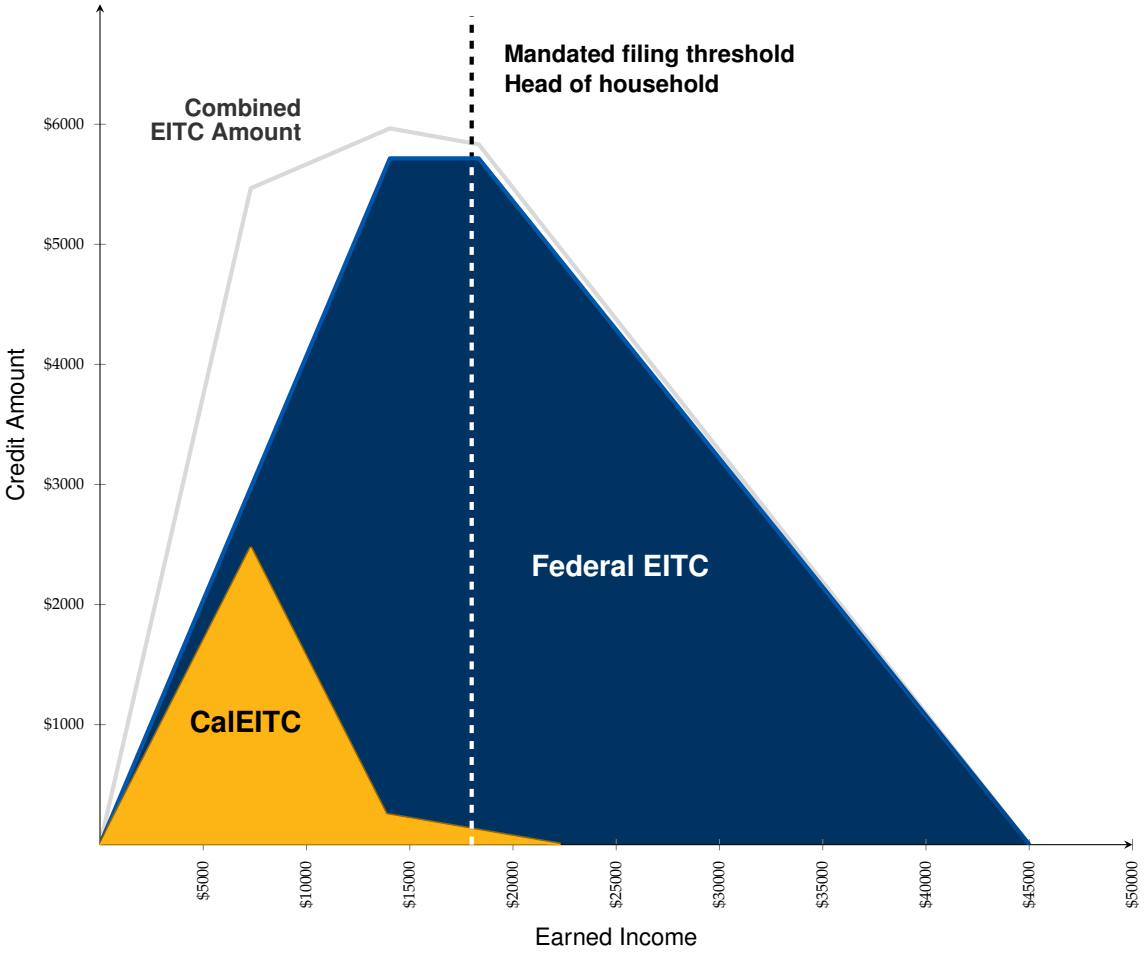
Romich, Jennifer L and Thomas S Weisner (2002). "How families view and use the Earned Income Tax Credit: Advance payment versus lump-sum delivery". *making Work Pay: The Earned Income Tax Credit and Its Impact on America's Families*, pp. 366–92.

Scholz, John Karl (1994). "The earned income tax credit: Participation, compliance, and antipoverty effectiveness". *National tax journal*, pp. 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*, pp. 1187–1209.

Figures

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



Notes. This diagram 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.

Tables

Table 1: Summary statistics for tax filer sample

	By number of dependents				
	0	**1**	**2**	**3+**	**Total**
Count of Individuals	13,071,385	6,610,757	7,633,885	6,672,745	33,988,772
Count of Tax Units	10,355,163	2,817,791	2,181,257	1,347,900	16,702,111
Count of Tax Units That E-Filed	8,926,909	2,502,564	1,952,585	1,206,975	14,589,033
% E-Filed	86%	89%	90%	90%	87%
Statistics for e-filers (tax unit level)					
Filing Status					
Single	72%	12%	7%	6%	48%
Married Filing Jointly	26%	42%	60%	64%	37%
Married Filing Separately	1%	1%	1%	0%	1%
Head of Household	0%	45%	32%	29%	15%
EITC Claiming					
% claiming federal EITC	7.4	34.0	32.8	32.1	17.4
Mean federal EITC claim (if positive)	$335	$2,317	$3,516	$3,835	$2,335
% claiming CalEITC	7%	17%	13%	10%	9%
Mean CalEITC claim (if positive)	$76	$290	$526	$573	$266
% claiming either EITC	8%	34%	33%	32%	18%
Mean total EITC (if positive)	$394	$2,455	$3,721	$4,012	$2,464
Income					
Mean earnings	$49,218	$73,065	$101,194	$82,917	$63,053
Mean AGI	$76,165	$90,063	$125,567	$105,820	$87,614

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: Summary statistics for CalFresh sample

	By case-level filing status			
	Everyone files	**Some file**	**No one files**	**Total**
Number of individuals	3,032,005	1,079,949	1,505,795	5,617,749
Row percent	54%	19%	27%	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	0.3	0.9
Share of adults w. linked wages	68%	53%	28%	51%
Mean total EDD wages (if positive)	$15,919	$13,803	$9,197	$14,229
Tax return linkage				
Linked to tax return	100%	50%	0%	55%
Linked to federal EITC tax return	57%	33%	0%	32%
Mean federal EITC amount (if positive)	$3,017	$3,698	.	$3,094
Mean CalEITC amount (if positive)	$378	$509	.	$392

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.

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	1,110,984	766,596	69%	$3,081	504,230	45%	$396
By filing status and number of dependents							
Single							
0 dependents	351,882	113,314	32%	$340	108,696	31%	$88
1 dependent	236,994	215,584	91%	$2,665	159,290	67%	$355
2 dependents	174,204	166,082	95%	$4,124	108,989	63%	$625
3+ dependents	103,867	98,713	95%	$4,581	55,122	53%	$668
Married							
0 dependents	41,221	15,924	39%	$391	11,151	27%	$80
1 dependent	48,269	37,082	77%	$2,649	19,304	40%	$287
2 dependents	69,532	55,387	80%	$4,098	22,306	32%	$523
3+ dependents	85,015	64,510	76%	$4,350	19,372	23%	$555
By total earnings (thousands)							
$0-5	155,215	88,617	57%	$1,004	81,202	52%	$474
$5-10	189,622	143,054	75%	$2,127	136,061	72%	$732
$10-15	222,500	180,767	81%	$3,582	165,994	75%	$300
$15-20	168,681	115,493	68%	$4,659	90,214	53%	$121
$20-25	124,913	87,031	70%	$4,190	30,542	24%	$31
$25-30	85,532	63,023	74%	$3,494	107	0%	$354
$30+	164,521	88,611	54%	$2,240	110	0%	$430

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 5s through 7 report the same statistics but for the state EITC.

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 our ability to observe individuals sharing CalFresh cases with each other for a certain number of months in the tax year.

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%	7.9%	1.7%
2	0.2%	0.4%	91.5%	8.8%
3+	0.1%	0.1%	0.3%	89.4%
N	126,642	249,089	218,151	161,068

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

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	754,950	534,406	71%	42,024	8%	$233
By filing status and number of qualifying children						
Single						
0 QCs	114,737	113,701	99%	5,607	5%	$64
1 QC	224,558	176,198	78%	14,289	8%	$220
2 QCs	159,830	113,405	71%	10,134	9%	$326
3+ QCs	85,611	53,829	63%	5,347	10%	$310
Married						
0 QCs	16,676	12,125	73%	620	5%	$71
1 QCs	38,970	21,459	55%	1,817	8%	$151
2 QCs	55,223	23,873	43%	2,218	9%	$247
3+ QCs	59,345	19,816	33%	1,992	10%	$230
By total earnings (thousands)						
$0-$5	83,266	78,551	94%	2,609	3%	$416
$5-$10	140,411	139,194	99%	6,594	5%	$571
$10-$15	179,530	176,178	98%	12,078	7%	$247
$15-$20	114,685	104,297	91%	14,787	14%	$121
$20-$25	86,407	36,186	42%	5,956	16%	$28
$25-$30	62,638	0	0%	0	.	.
$30+	88,013	0	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 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.

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%	43.7%	17.8%	7.2%
1	0.0%	56.3%	17.5%	7.1%
2	0.0%	0.0%	64.6%	16.9%
3+	0.0%	0.0%	0.0%	68.8%
N	263,865	32,597	22,267	25,659

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

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	344,388	112,881	33%	109,786	97%	$84
By filing status and number of qualifying children						
Single						
0 QCs	252,388	107,442	43%	105,012	98%	$82
1 QC	12,724	2,196	17%	1,869	85%	$175
2 QCs	5,159	337	7%	228	68%	$318
3+ QCs	2,983	96	3%	70	73%	$434
Married						
0 QCs	31,550	2,493	8%	2,367	95%	$73
1 QC	11,341	162	1%	136	84%	$267
2 QCs	13,560	107	1%	73	68%	$491
3+ QCs	14,683	48	0%	31	65%	$574
By total earnings (thousands)						
$0-5	66,598	35,861	54%	34,814	97%	$140
$5-10	46,568	42,203	91%	41,197	98%	$81
$10-15	41,733	33,136	79%	32,238	97%	$30
$15-20	53,188	1,164	2%	1,048	90%	$87
$20-25	37,882	517	1%	489	95%	$23
$25-30	22,509	0	0%	0	.	.
$30+	75,910	0	0%	0	.	.

Notes. Universe is e-filed tax returns linked to at least one CalFresh head or spouse 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.

Table 9: 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%	19.8%	7.1%	6.9%
1	0.0%	80.2%	9.5%	1.3%
2	0.0%	0.0%	83.4%	7.3%
3+	0.0%	0.0%	0.0%	84.6%
N	610,644	104,958	52,178	35,924

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.

Table 10: 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	803,724	289,765	36%	289,765	100%	$196
By filing status and number of qualifying children						
Single						
0 QCs	600,388	226,703	38%	226,703	100%	$87
1 QC	83,587	29,018	35%	29,018	100%	$489
2 QCS	40,473	13,431	33%	13,431	100%	$827
3+ QCs	24,691	7,455	30%	7,455	100%	$920
Married						
0 QCs	37,222	7,264	20%	7,264	100%	$85
1 QC	6,030	2,012	33%	2,012	100%	$434
2 QCs	5,645	1,990	35%	1,990	100%	$799
3+ QCs	5,688	1,892	33%	1,892	100%	$873
By total earnings (thousands)						
$0	435,797	0	0%	0	.	.
$1-$5	178,585	175,666	98%	175,666	100%	$180
$5-$10	65,929	64,693	98%	64,693	100%	$305
$10-$15	41,383	40,677	98%	40,677	100%	$109
$15-$20	27,449	6,009	22%	6,009	100%	$117
$20-$25	19,771	2,720	14%	2,720	100%	$29
$25-$30	12,141	0	0%	0	.	.
$30+	22,669	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.

Table 11: 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	754,950	534,406	71%	42,024	8%	$233	$9,791,892
Non Fed EITC claimants	344,388	112,881	33%	109,786	97%	$84	$9,275,297
Non-Filers							
All	803,724	289,765	36%	289,765	100%	$196	$56,651,396
Total	1,903,062	402,646	49%	441,575	47%	$171	$75,718,585

Notes. Table 11 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 12: 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	504,230	501,232	99%	3,069	1%	$714	$2,190,355
Non CalEITC claimant	606,754	369,209	61%	106,433	29%	$415	$44,122,653
Total	1,110,984	870,441	78%	109,502	13%	$423	$46,313,008

Notes. Table summarizes rates of eligibility for the federal EITC among e-filed tax returns with a CalFresh-enrolled head or spouse. Results are separated between units that included a CalEITC claim and those that did not.

Table 13: CalEITC take-up among CalFresh recipients by tax preparation method

	Tax Preparation Method			
	Paid	**Self**	**VITA**	**Total**
Number of Tax Returns	671,335	377,819	48,006	1,097,160
Claimed the Federal EITC				
Number of tax returns	474,790	251,232	27,647	753,669
CalEITC				
% eligible	69%	74%	79%	71%
% non-claiming among eligible	12.1%	1.6%	0.4%	7.9%
Mean unclaimed amount	$232	$252	$145	$233
Total unclaimed amount	$8,998,985	$757,609	$12,789	$9,769,383
Did Not Claim the Federal EITC				
Number of Tax Returns	196,545	126,587	20,359	343,491
CalEITC				
% eligible	31%	35%	32%	33%
% non-claiming among eligible	99%	94%	100%	97%
Mean unclaimed amount	$85	$83	$90	$84
Total unclaimed amount	$5,156,249	$3,504,335	$594,014	$9,254,598
Federal EITC				
% eligible	30%	36%	26%	32%
% non-claiming among eligible	100%	100%	100%	100%
Mean unclaimed amount	$425	$426	$372	$423
Total unclaimed amount	$24,871,421	$19,410,396	$1,940,375	$46,222,192

Notes. Universe is e-filed tax returns linked to at least one CalFresh participant. 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.

Table 14: 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	6,459	81,395	15,912	146,976	589,561	4,275	145,188	276,927	1,266,693
% eligible	58%	55%	51%	67%	55%	53%	55%	55%	56%
% non-claiming among eligible	19%	20%	19%	18%	25%	22%	20%	20%	22%
Mean unclaimed amount	$137	$108	$105	$115	$114	$136	$110	$111	$113
Dependents									
Number of individuals	4,654	53,691	11,500	138,062	652,631	4,287	124,972	191,880	1,181,677
% eligible	57%	52%	48%	67%	51%	47%	55%	53%	54%
% non-claiming among eligible	8%	7%	7%	8%	11%	9%	7%	7%	9%
Mean CalEITC amount	$312	$229	$181	$246	$220	$269	$251	$271	$232
Within imputed tax units									
Heads and Spouses									
Number of individuals	8,959	33,552	8,019	146,689	272,572	3,585	99,644	285,229	858,249
% eligible	31%	20%	26%	43%	39%	36%	34%	30%	35%
% non-claiming among eligible	100%	100%	100%	100%	100%	100%	100%	100%	100%
Mean unclaimed amount	$61	$42	$63	$86	$86	$92	$64	$57	$72
Dependents									
Number of individuals	3,027	9,094	3,117	59,143	142,773	1,610	36,920	79,721	335,405
% eligible	26%	22%	27%	37%	32%	35%	31%	29%	32%
% non-claiming among eligible	100%	100%	100%	100%	100%	100%	100%	100%	100%
Mean CalEITC amount	$160	$144	$188	$261	$226	$262	$204	$184	$217
Overall take-up gap									
Heads and Spouses									
% non-claiming among eligible	54%	30%	36%	50%	44%	50%	44%	49%	45%
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 CalFresh participant. 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.

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 indicate that families in SNAP households are more likely to be eligible for the federal EITC and CalEITC, given their lower income and higher number of children. Of course, many eligible families who 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 consider the approximately 13% of California tax units that 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.

Given the information we do have on paper filers,[22] we believe that excluding these tax units from our analysis results in our underestimating the share of tax units that are both 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, given the technology and information available to professional tax preparers. 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.

[22] 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.

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 from overpayment of income taxes. In our FTB data, 5.6% of our full sample – or 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, have 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 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 qualifying children. 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 400,000 CalFresh recipients who are on a tax return with an individual with an ITIN,[23] 94% appear in our data as Hispanic, even though Hispanic individuals represent just 56% 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 Unassigned dependents

As discussed in Section 4.3, we are unable to assign 732,000 non-filers (including 285,000 children) to an imputed tax unit. For 99.9% of these 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-

[23] 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.

only" reference cases represent households in which parents or guardians are not eligible for CalFresh due to their documentation status.

To test whether this expectation is accurate, we rerun our tax unit imputation procedure on the universe of CalFresh recipients who *do* appear on a state tax return. Of the 2,293,000 individuals assigned to be dependents, we are unable to assign 902,000 dependents (693,000 of which are children) to an imputed tax unit – again, because these individuals are on reference cases without a likely adult filer. Over 87% of these unassigned dependents appear on tax returns in which neither the head nor spouse enrolled in CalFresh. Of the dependent children on reference cases without a likely filer, 43.5% 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. Hispanic children are over-represented in both the count of unassigned children and the count of unassigned children with ITINs present on their tax returns. Overall, 61% of children enrolled in CalFresh are identified as Hispanic, but Hispanic children comprise 79% of those who are not assigned to an imputed tax unit. Further, 52% 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 11%. Overall, 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.

C.3 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, 611,607 returns included a filer or spouse with an ITIN (Appendix Table 4). Of these returns, we estimate that 67% would be eligible for the federal EITC and 25% for the CalEITC if not for the SSN test. The average federal EITC amount for these tax units would be $3,133 and $307 for the CalEITC. If these tax units were allowed to claim either credit, an additional 907,217 children would become eligible for the federal EITC and an additional 231,515 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 $1.3 billion from the federal EITC. 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.

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 208,335 such tax units. If these dependents became eligible, the number of qualifying children on the average affected tax unit would increase from .8 to 2.2. Of these returns, 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 179,404 children would become eligible for the federal EITC, and 39,236 for the CalEITC, and their tax units would be eligible for an additional $304 million from the federal EITC and $9.7 million from the CalEITC. The average federal EITC amount for these households would be over $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 heads/spouses 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 filer (a head or spouse) or a dependent on their unit's return if one were to 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 filer. While most of these working-age (or near working-age) individuals would likely be 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, and are being cared for by working-age children. Others 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 heads/spouses 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. We estimate the following logit on the observations in the training set.

$$\ln\left(\frac{p(Y_i=\text{Filer}|X_{ij}))}{1-p(Y_i=\text{Filer}|X_{ij})}\right) = \beta_0 + \beta_j \mathbf{X}_{ij}$$

X_i is a vector of characteristics we observe for each individual in the CalFresh data, including: 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 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 filers. 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 head or spouse on a tax return last year (a 4% cutoff), individuals who were a dependent last year (9%), 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 following paper.

E Accuracy of earnings information

For filers, we can use households' actual reported earnings and income to measure eligibility. For non-filers, our estimates of eligibility hinge on our assumption that the sum of EDD wages is a reliable measure of households' true total earned income. Since we do not observe tax information for non-filers, we cannot test this assumption. Instead, in the following section, we present evidence that EDD wages do often reflect households' total earned income among CalFresh filing households.

For all actual tax units containing only CalFresh-enrolled filers (ie, single filers enrolled in CalFresh or both head and spouse in married filing jointly households enrolled in CalFresh), we compare the sum of UI-covered wages for the head (and spouse, for married couples) to their total reported AGI and California wages. 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 \$300. For 2% of returns, the difference is greater than \$20,000 dollars. 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 \$1,800. For 5% of returns, the difference is greater than \$20,000 dollars.

Among the households in which the difference between AGI and total earned income is greater than \$20,000, 15% report only self-employment income and have no reported EDD earnings. Our inability to measure earned income correctly for this class of tax units 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.

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 $9,165 and the median was $8,689. Ten percent (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.

F Predicted versus actual CalEITC claim

In Appendix Table 6, we present estimates of CalEITC eligibility and participation among CalFresh tax filers who claimed the federal EITC. Since the number of qualifying children claimed for the purposes of the EITC is not available in our tax records, we use the value of each tax unit's federal EITC claim, plus their earned income, to infer their number of qualifying children. Using this inference, along with earnings information available in their tax return, we determine each tax unit's eligibility for the CalEITC, as well as an estimated credit amount for those units that are eligible. To further validate the reliability of this inference, the following section discusses how our predicted CalEITC amounts compare to actual claimed amount among tax units who claimed the credit.

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 begin 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.[24]

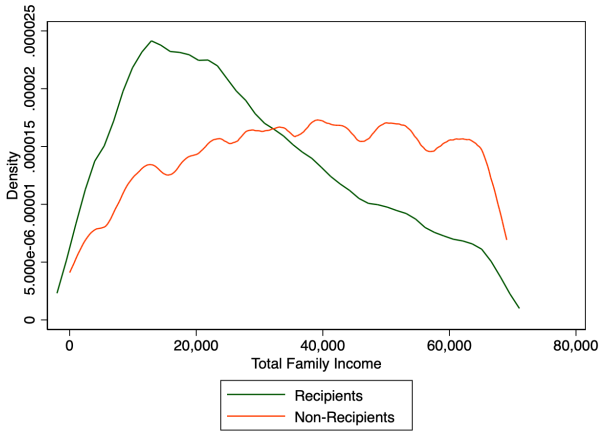
[24]One possible explanation for this pattern of results is that these tax units may have had dependents who did not reside in the state of California for the requisite 6 months required to satisfy the state residency requirement of the CalEITC to be counted as qualifying children. Using CalFresh records, 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 heads or spouses 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.

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.

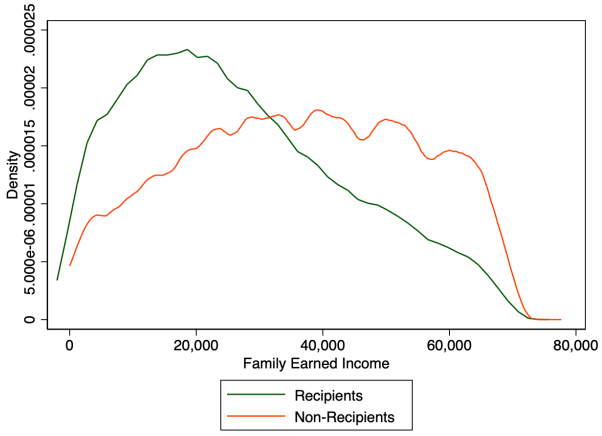
Following the above exercises, we are left with only 1,841 tax units (approximately 0.38 % of the full sample of tax units claiming the CalEITC) for which no combination of substitutions above yielded an exact match between predicted and actual credit amounts.

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

(a) Total family income



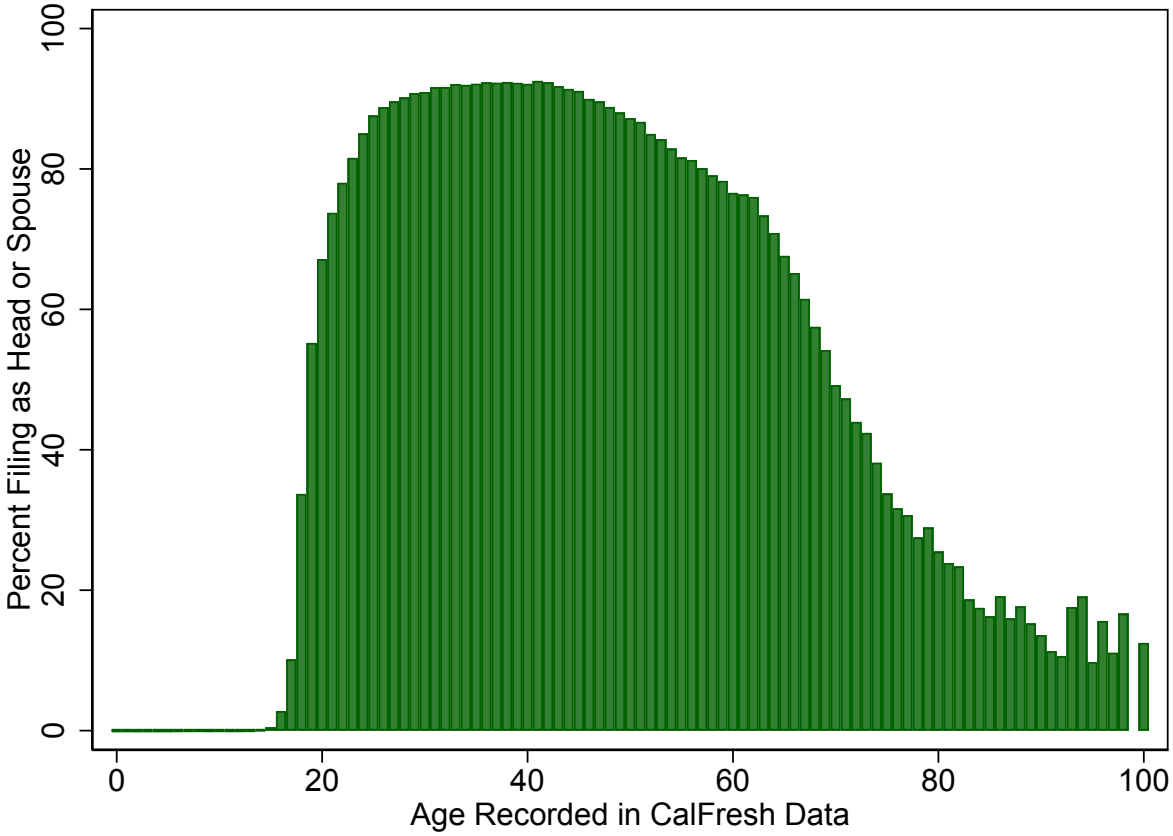
(b) Family earned income



(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.

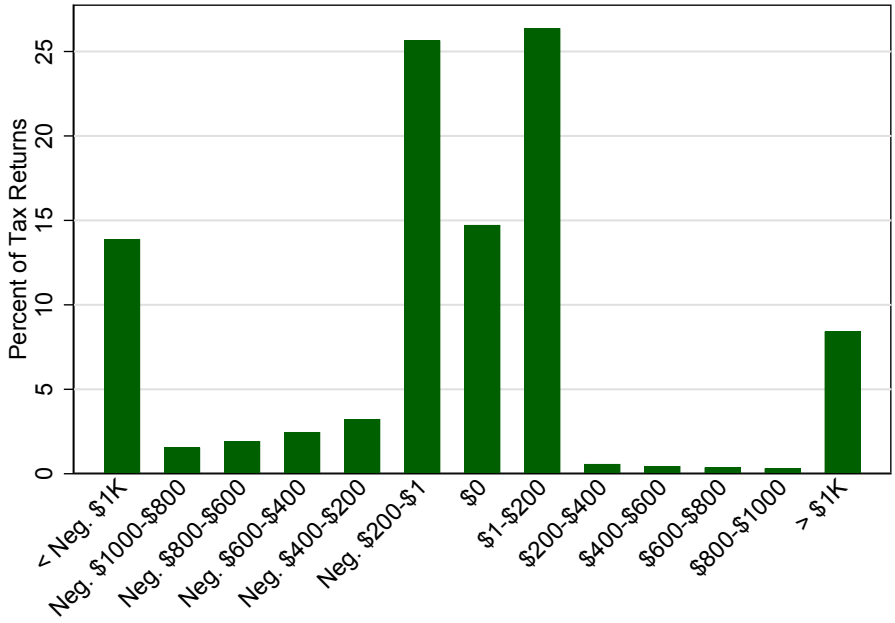
Appendix Figure 2: Distribution of ages among dependents on CalFresh tax units, TY 2017



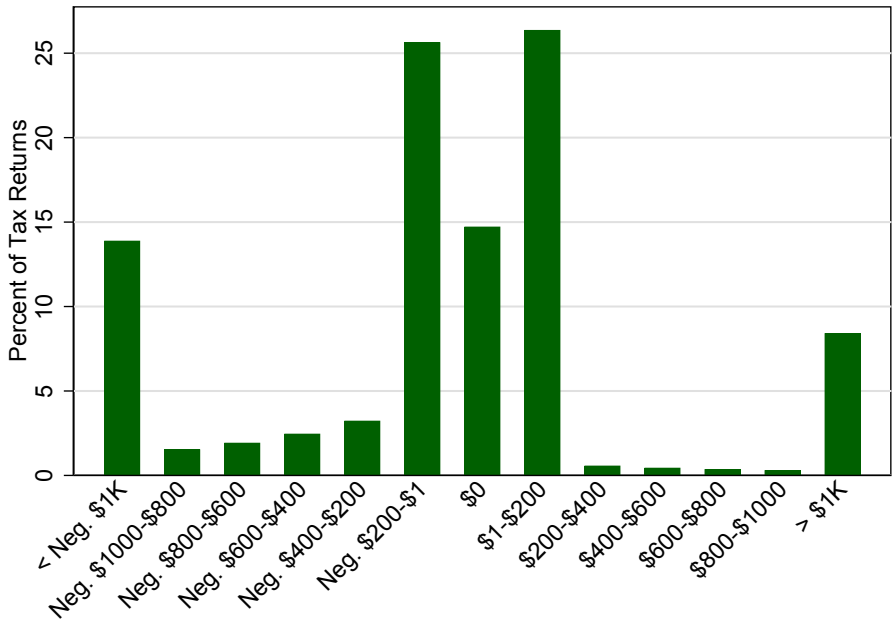
Notes. Universe is all individuals on tax year 2017 returns who also enrolled in CalFresh that year. We group ages into one year bins, according to enrollees' reported date of birth, and report the share in each bin who appear on their tax return as head/spouse as opposed to dependent.

Appendix Figure 3: Distribution of differences between total EDD earnings and reported California wages among tax units with all CalFresh-enrolled heads and spouses, TY 2017

(a) 540 wages versus EDD wages

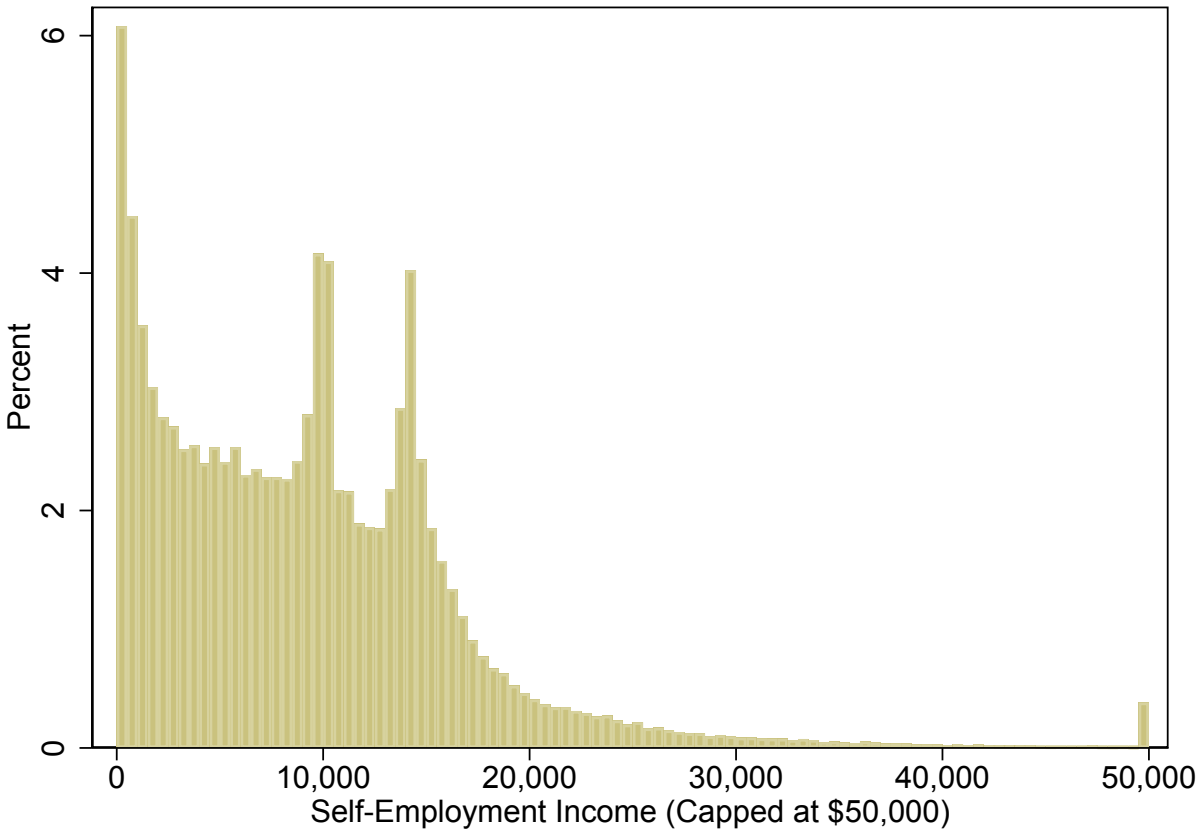


(b) AGI versus EDD wages



Notes. Universe is all tax returns with head and spouse (if present) 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 head and/or spouse (if present) enrolled in CalFresh and 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	2,113,078	14,589,033	16,702,111
Filing status and number of dependents			
Share single	50%	48%	48%
Share head of household	12%	15%	14%
Share married filing jointly	35%	37%	36%
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%	14%	14%
Share claiming CalEITC among eligible	41%	54%	52%
Income			
Mean Federal AGI	$64,644	$73,389	$72,282
Median Federal AGI	$38,009	$42,823	$42,239
Mean CA wages	$46,972	$55,341	$54,282
Median CA wages	$24,732	$31,482	$30,708
Other characteristics			
Share on CalFresh	6%	8%	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 that California wages and Federal AGI reported on 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	144,796	1,317,854	1,462,650
Income information			
Mean AGI	$9,862	$10,989	$10,877
Mean earned income	$9,781	$10,806	$10,705
Mean wage income	$6,885	$8,338	$8,194
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)	$1,883	$2,233	$2,198

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	611,607	611,607	611,607
Without SSN test			
Single filers			
% eligible	36%	77%	78%
Number eligible	108,774	232,806	233,702
Mean EITC amount	$289	$3,014	$3137
Total EITC amount	$31,394,302	$701,719,023	$733,113,325
Married filing jointly, one filer has ITIN			
% eligible	10%	43%	43%
Number eligible	15,245	64,009	64,185
Mean EITC amount	$313	$2,875	$2942
Total EITC amount	$4,775,348	$184,036,537	$188,811,885
Married filing jointly, both filers have ITIN			
% eligible	18%	71%	71%
Number eligible	29,114	113,543	113,670
Mean EITC amount	$370	$3,440	$3531
Total EITC amount	$10,768,346	$390,556,317	$401,324,663
Total			
% eligible	25%	67%	67%
Number eligible	153,133	410,358	411,557
Mean EITC amount	$307	$3,110	$3,215
Total EITC amount	$46,937,996	$1,276,311,877	$1,323,269,873

Notes. Universe is e-filed tax returns containing a 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	208,335	208,335	208,335
With SSN Test			
Mean count of QC	0.8	0.8	0.8
% eligible	8%	22%	23%
Mean EITC amount	$13	$480	$2,175
Total EITC amount	$2,752,525	$100,071,072	$102,823,597
Without SSN Test			
Mean count of QC	2.2	2.2	2.2
% eligible	13%	50%	50%
Mean EITC amount	$47	$1,462	$3,030
Total EITC amount	$9,769,810	$304,605,333	$314,375,143

Notes. Universe is e-filed tax returns containing a dependent with an ITIN, but 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	62.2%	0.0%	0.0%	0.0%
1	25.8%	76.9%	0.0%	0.0%
2	9.2%	19.0%	82.7%	0.0%
3+	2.8%	4.1%	17.3%	100%
Number of qualifying children reported on Sch 3514				
0	66.7%	4.6%	4.1%	3.5%
1	24.2%	77.4%	2.9%	1.6%
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	754,950	100%	68%	38,941	8%	$217
By filing status and number of qualifying children						
Single						
No QCs	114,737	15%	99%	5,616	5%	$81
1 QC	224,558	30%	73%	12,687	8%	$208
2 QCs	159,830	21%	68%	9,322	9%	$333
3+ QCs	85,611	11%	61%	5,103	10%	$354
Married						
No QCs	16,676	2%	73%	715	6%	$60
1 QC	38,970	5%	50%	1,468	8%	$115
2 QCs	55,223	7%	42%	2,075	9%	$173
3+ QCs	59,345	8%	33%	1,955	10%	$164
By total earnings (thousands)						
$0-$5	83,266	11%	94%	2,584	3%	$380
$5-$10	140,411	19%	99%	6,509	5%	$568
$10-$15	179,530	24%	97%	11,780	7%	$229
$15-$20	114,685	15%	79%	12,827	14%	$92
$20-$25	86,407	11%	37%	5,241	16%	$10
$25-$30	62,638	8%	0%	0	.	.
$30+	88,013	12%	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.