Beating the Clock: Using Year-end Changes to Identify Intensive Margin Labor Supply Responses to Taxation

Matthew Unrath [*]

November 20, 2020
Click here for the latest version.

Abstract

Identifying the effect of tax policy on the labor supply of individuals who would work regardless has been a longstanding empirical challenge. This paper proposes a new strategy for identifying workers' intensive-margin labor supply elasticity using within-year variation in anticipated year-end tax rates. I modify the standard non-linear budget set approach to include uncertainty about future employment. With uncertainty, households must forecast their annual income in order to anticipate the average and marginal tax rates that apply to their earnings. Workers only learn their true tax rates as the year progresses and they realize their employment history. Using survey and administrative data, I conclude that low-income households' labor supply responds more to expected tax rates at the end of the year, when certainty about annual income is greatest. I use the excess sensitivity to tax incentives near the end of the year, relative to other periods, to estimate an intensive margin labor supply elasticity between .08 and .18. This response is identified largely from non-linearity in the Earned Income Tax Credit (EITC) schedule and implies a larger intensive margin response to this program than previous estimates.

[*]Goldman School of Public Policy, UC Berkeley, unrath@berkeley.edu. I thank Jesse Rothstein for his invaluable guidance. I also thank Hilary Hoynes, David Card, and participants at multiple seminars at UC Berkeley for their comments and feedback. I am also indebted to Daniel Feenberg at NBER, as well as staff at the California Franchise Tax Board (FTB) and California Department of Social Services (CDSS), in particular: Sean McDaniel, Monica Trefz, Ian Kiltz, Chad Angaretis, Silvano Guitierrez, Bob Schlie, and Xudong Chen, and especially Allen Prohofsky and Julie Moreno at FTB; and Alexis Fernandez, Kim McCoy-Wade, Jianjun Chen, Brittney Gossard, and Akhtar Khan at CDSS. I also thank Charles Davis, Konrad Franco, and John Iselin for their assistance with completing the match between the CDSS and FTB records.

1 Introduction

Identifying the effect of taxes on the labor supply of people who would work regardless has been a longstanding empirical challenge. Despite the incredible policy relevance of this parameter, the micro literature lacks reliable estimates of the intensive margin response to tax incentives. Identification is challenging, in part, because it is difficult to isolate exogenous variation in marginal tax rates within non-linear tax schedules, and in part because extensive margin responses to tax policies create sample selection that biases standard difference-in-difference analyses. Further, even credible average estimates mask important heterogeneous responses. I would expect larger responses to short-term variation in tax rates that can be avoided via intertemporal substitution than to longer-term variation, and it is not clear which are identified by many existing studies.

Concretely, understanding the impact of the Earned Income Tax Credit (EITC) on labor supply is an important topic for policy purposes. The EITC is the largest means-tested cash assistance program in the US and largely shapes the tax policy facing lower-income workers. Many studies demonstrate the program's substantial effects on workers' extensive margin labor supply decision,[1] but few find any effect on the intensive margin. This result is not wholly convincing, however, given the large changes to the composition of the workforce that reforms to the EITC induce (Nichols and Rothstein, 2015).

A possible explanation for the lack of a response, and an important unresolved issue in the labor supply literature, relates to workers' information structure. Labor supply models of non-linear budget sets (Hausman, 1982) estimate individuals' responses to their estimated tax rates, but it is not clear that workers are aware of the rates that they face. This is a particular concern with non-linear, annual tax schedules: The marginal tax rate on an individual's labor supply one day depends on their total earnings throughout the year, and for many periods in a year, those earnings have not yet been realized. It seems likely that many low-wage workers, whose earnings are disproportionately volatile and whose income tax schedule is highly non-linear, have trouble forecasting their annual income and thus their average and marginal tax rates.

In this paper, I propose a new strategy for identifying short-term labor supply responses that exploits this uncertainty. At the beginning of the year, workers' forecasts of their annual earnings and their marginal tax rates are imprecise, but they gain more information with each work day, so they can make more accurate estimates of the tax rate that will ultimately apply to their earnings. As a result, we might expect workers' labor supply to depend more strongly on their expected annual tax rates in the fourth quarter of the tax year, when their employment history is nearly realized, than in the first.

I use this idea to obtain a new measure of the intensive margin Frisch elasticity. I use the difference in awareness of true tax incentives between the beginning and end of the tax year

[1] One notable exception to this consensus is Kleven (2019), who argues, in part, that welfare reform waivers were responsible for labor supply increases observed in the early 1990s, as opposed to the EITC expansion. If true, these waivers would also confound studies of intensive margin responses to reforms of the EITC, even if those reforms did not have an extensive margin effect. See Schanzenbach and Strain (2020) for a response to Kleven.

to distinguish workers' independent response to tax policy from standard serial correlation in earnings. I interpret the excess sensitivity of earnings to likely tax incentives in the fourth quarter, relative to other calendar quarters, to reflect workers' intentional reallocation between labor and leisure in response to those tax incentives.

I evaluate whether workers make these year-end adjustments, and I measure the size of this response, using two data sources: the Survey of Income and Program Participation (SIPP) and administrative earnings records for lower-income Californians enrolled in the Supplemental Nutrition Assistance Program (SNAP). Both datasets contain within-year earnings and information needed to identify households' tax rates. I construct likely tax units from SIPP households and SNAP cases, measure each workers' total earnings through multiple within and cross-year periods, and identify predicted marginal and average tax rates for each tax unit in each of those periods using National Bureau of Economic Research's (NBER) TAXSIM program. To measure households' labor supply response to their expected tax rates, I relate earnings in each period to the average tax rate that would apply to that period's predicted earnings. I distinguish apparent year-end responses to tax policy from other serial earnings patterns by separately identifying this response in each calendar quarter, and evaluating whether the relationship appears strongest in the final quarter of the tax year. To account for any lingering omitted variable bias concerns, I identify this response within the same workers and households over multiple consecutive years.

I conclude that household labor supply is indeed more sensitive to expected tax incentives at the end of the tax year. For a 10 percentage point increase in their predicted net of tax wage rate (i.e., a 10 percentage point decrease in their predicted tax rate), households increase earnings in the fourth quarter by 1 to 2 percent on average. This response is largely driven by households facing the most negative tax rate at year's end. When the same household expects to have especially modest annual earnings, such that they would expect to be on the phase-in segment of the EITC schedule, they tend to increase earnings in the following quarter. This response is most pronounced at the end of the tax year, suggesting that this adjustment is not due simply to mean reversion or an effort to achieve a minimal level of earnings. At the same time, I find no evidence that households facing steeply positive marginal tax rates – those with earnings on the phase-out segment of the EITC – reduce their labor supply due to those incentives.

My primary empirical strategy assumes that labor supply choices early in the year are not made based on worker's expectation of their end-of-year tax rate. While uncertainty about annual earnings makes this likely to be close to accurate, it may not hold exactly. To assess the importance of this assumption, I estimate an alternative model in which I test whether households adjust earnings earlier in the year based on their expectation of year-end tax incentives. I find that my main result still holds; earnings responses are more sensitive to forecasted tax incentives nearer the end of the tax year. This supports the interpretation of my main estimates, in that they reflect a real response to tax policy rather than bias.

My approach boasts several advantages over other strategies for measuring labor supply elasticity. First, I overcome classic econometric issues that plague most other empirical approaches (See Keane, 2011, for a summary). I address the "taste for work" bias problem by controlling for to-date earnings and measuring adjustments within the same household over multiple years. I address the issue of simultaneity by instrumenting for workers' current net wage using their predicted year-end after-tax income. This instrument also addresses concerns about mismeasurement of hours and earnings. Second, my approach overcomes the selection and composition issues that bias difference-in-difference evaluations (Nichols and Rothstein, 2015); I use panel data to identify responses within the same households within and across tax years. Third, I address external validity concerns inherent to investigations of unique tax reforms by studying a common and regular setting faced by many different workers in a variety of circumstances over many different years. Fourth, I do not rely on a structural model with strong parametric assumptions to interpret cross-sectional data.

To make this approach more tractable, I propose a finite, multi-period model of labor supply in which a representative agent, who aims to maximize utility over consumption and leisure, decides at the start of each period how much to work. The agent makes this decision given her to-date earnings, uncertainty about being employed in this and future periods, a non-linear tax schedule, and parameters governing preferences and probability of employment. In the first period, the agent chooses a preferred bundle of work and leisure based on their expectation of future employment. As she realizes her employment history and gains certainty about her expected tax rate, her optimal labor supply choice becomes more sensitive to her previous earnings. I show that households facing a less positive tax rate tend to increase labor supply in the final period in order to maximize post-tax income, and workers facing a more positive tax rate will work less. These adjustments result in greater bunching in proposed annual income as the tax year progresses. Uncertainty about employment, plus a reasonable compensated elasticity, yields earnings densities that resemble those observed empirically.

This paper makes a number of important contributions to several literatures in labor and public economics.

First, this paper contributes to another extensive literature measuring labor supply and taxable income elasticity using non-linear budget sets (e.g., Blundell et al., 2009, 2000; Blundell and MaCurdy, 1999; Burtless and Hausman, 1978; Hausman, 1985; Keane and Moffitt, 1998; MaCurdy et al., 1990; Moffitt, 1990), and a related literature using "bunching" at kink points in tax schedules to measure the same elasticities (Blomquist and Newey, 2017; Kleven and Waseem, 2013; Saez, 2010). The non-linear budget set is the starting point for my model; I expand it by permitting agents to adjust their labor supply choice over time. I provide important context to Saez's (2010) well-known result, arguing that individuals try to move toward kink points but are precluded from perfectly bunching by numerous frictions.

Second, this paper adds to a newer literature measuring the size and relevance of those frictions. Chetty et al. (2011) and Gelber et al. (2020) both measure the importance of adjustment frictions in mitigating labor supply responses to tax and transfer policy. Liebman and Zeckhauser (2004) and Rees-Jones and Taubinsky (2019) study how workers respond to marginal versus average tax rates. Chetty et al. (2013), Feldman et al. (2016), and Miller et al. (2015) study how complexity and salience of tax policies matter to workers' labor supply choices. I contribute to this literature by measuring the relevance of uncertainty about future employment – an oft-cited but under-studied type of an optimization friction – to labor supply choices. I also document that workers appear more responsive to their expected average tax rate, as opposed to their expected marginal tax rate.

Third, this paper adds to the literature investigating labor supply effects of the EITC. Much of the non-linearity in tax incentives that I study is due to the EITC's structure. Accordingly, this paper can largely be understood as a study of households' response to this program. Most work finds that the EITC draws individuals into the workforce (Eissa and Liebman, 1996; Gelber and Mitchell, 2011; Grogger, 2003; Hotz and Scholz, 2006; Meyer and Rosenbaum, 2001), but there is limited evidence of an intensive-margin response, despite clear theoretical predictions that it should have such an effect (Eissa and Hoynes, 2006; Hotz, 2003; Nichols and Rothstein, 2015; Saez, 2010).

There are two notable exceptions to this near consensus finding. Using apparent variation in local knowledge about the EITC schedule from tax return data, Chetty et al. (2013) identify how labor supply changes in response to the birth of a child across areas with more or less awareness of the EITC's incentives; they estimate an average intensive margin earnings elasticity of .14 in the phase-out region and .31 in the phase-in region. Second, Chetty and Saez (2013) use an experiment in which they inform some taxpapers about the EITC's structure to test how awareness of the EITC's incentives affects labor supply. They find a small (3 percent) increase in average earnings among treated subjects, implying a labor supply elasticity of .075 (Nichols and Rothstein, 2015). Like these studies, I use differences in workers' awareness of their tax incentives to identify a labor supply response to the EITC. I also find a non-zero intensive margin response to the EITC, similar to these studies but in contrast to the rest of the literature. My preferred estimates are also quite similar to these authors' results. Unlike Chetty et al. (2013), I find that this effect is driven entirely by households with the lowest earnings. I find no evidence in labor supply reduction by households predicted to be in the phase-out part of the EITC schedule.

Fourth, this paper contributes to an extensive literature studying both the extensive and intensive margin Frisch elasticity (Altonji, 1986; Angrist, 1991; Blundell et al., 2016; Card and Hyslop, 2005; Keane, 2011; Kimball and Shapiro, 2008; Laitner and Silverman, 2005; MaCurdy, 1981; Manoli and Weber, 2016; Pencavel, 1986; Pistaferri, 2003) and how transitory changes in wages and tax policy affect labor supply choices (Camerer et al., 1997; Crawford and Meng, 2011; Farber, 2008; Fehr and Goette, 2007; Martinez et al., ming; Powell, 2015; Stafford, 2015). Much of this work focuses on the relevance of those shocks in life-cycle models (Heckman and MaCurdy, 1980; Keane and Rogerson, 2012; Keane, 2011). I study how idiosyncratic wage shocks affect workers short-run labor supply behavior. My estimate of the Frisch elasticity is similar to those in the micro-literature summarized by Reichling and Whalen (2012).

Fifth, this paper extends a small literature that documents how workers' labor supply can vary within the tax year. Yang (2018) and Powell (2020) both study labor supply effects of receiving a lump sum cash payment (i.e., EITC disbursement and 2008 stimulus payments). Each of these authors document an important extensive margin response due to the income effect. In contrast, this paper studies the relevance of the substitution effect on households' within year labor supply behavior. More similar to this study, Wilson (2020) uses the panel-nature of the CPS to document how EITC expansions decrease workforce exits and increase overall months worked among single mothers. Looney and Singhal (2006) study how losing a dependent exemption from aging children affects households' labor supply behavior in the short-term.

Finally, this paper makes an important contribution to policy-relevant discussions about potential reforms to tax-based means-tested programs like the EITC. Policymakers continually express interest in reforming the EITC so that it subsidizes households' earnings throughout the tax year instead of via one lump sum (Jones, 2010; Maag, 2019). If household earnings and tax liabilities fluctuate enough, this raises important concerns about the viability of such reforms. This study provides evidence about how households exhibit large changes in expected EITC eligibility across and within years. I also demonstrate for which households these predictions are likely to be most wrong.

The paper proceeds as follows. In Section 2, I describe my model. In Section 3, I describe the SIPP and the California administrative data, how I convert households in these data to tax units, and how I estimate households' tax rates using TAXSIM. In Section 4, I use these data to show the extent of non-linearity in the income tax schedule facing low-income households as well as variability in household income and tax rates over the year. In Section 5, I describe my empirical analysis. In Section 6, I summarize my results. In Section 7, I conclude.

2 Model

Consider the quasi-linear and isoelastic labor supply utility model proposed by Saez (2010). A worker with ability n aims to maximize utility over consumption c subject to a cost from working, $f(z)$. z denotes workers' pre-tax earnings, which are the product of the worker's wage w and labor supply h. e indexes the worker's compensated elasticity. Consumption c equals net of tax earnings: $(z - \tau(z)$. The worker faces a single-kinked tax schedule: $\tau(z) = \tau_0 \min(z, T) + \tau_1 \min(0, z - T)$.[2]

$$\begin{aligned} u(c, z) &= c - f(z) \\ &= c - \frac{n}{1+\frac{1}{e}}\left(\frac{z}{n}\right)^{1+\frac{1}{e}} \end{aligned}$$

[2] Saez (2010) shows that, with a continuous ability parameter n, the size of the mass in the earnings distribution clustered around z^* is determined by the compensated elasticity e. Saez (1999) shows that if workers actually earn $z + \epsilon$, this bunching would appears more as a dispersed mass as opposed to an atom, but the model still predicts taxpayers cluster at the kink T.

Saez's (2010) model assumes each worker makes a single labor supply choice h^* to maximize $u(c,z)$. I modify the model to incorporate sequential choices of labor supply and uncertainty about future earnings. Suppose a worker chooses a labor supply level h in each of D periods that comprise the tax year, and in each period, the worker is actually employed at h with probability p. The worker earns pre-tax income $z_d = wh_d$ in period d if employed and 0 if not. She only pays the cost of working, $f(h)$ if she's employed. She faces the same single kinked tax schedule, $\tau(z)$, but the tax is applied on total earnings across all D periods, meaning her consumption is a function of total earnings across all D periods: $c = \sum_{d=1}^{D} z_d - \tau\big(\sum_{d=1}^{D} z_d\big)$.

The worker chooses h_d^* in each period d to maximize expected total utility across all D periods, treating h_t as fixed for periods $t < d$, and knowing that she will adjust h_t for $t > d$. Given values for parameters p, τ_0, τ_1, n, and e, I can identify the agent's labor supply choice h for any level of y_{D-1} via dynamic programming. I solve the agent's labor supply decision recursively, beginning in period D and ending with period 1.

In period D, the agent chooses h_D^* to maximize total consumption given the same cost of working $f(h)$ from above:

$$\begin{aligned}\max_{h_D} E\big[U(c, h_1^*, h_2^*, \ldots h_{D-1}^*, h_D)\big] = \;& \underbrace{c(h_1^* + h_2^* + \ldots + h_{D-1}^* + h_D)}_{\text{total consumption}} \\ & - \underbrace{f(h_1^*) - \ldots - f(h_{D-2}^*) - f(h_{D-1}^*) - f(h_D)}_{\text{cost from working in each period}}\end{aligned}$$

This yields $h_D^*(h_1 + h_2 + ... + h_{D-1})$. The agent chooses h_D^* to maximize utility given to-date hours. The agent solves this final period maximization problem in nearly the same way an agent would solve the single period problem.

Now, I move back one period to $D-1$. The agent maximizes h_{D-1} given to-date earnings and an expectation of her choice in the following period.

$$\begin{aligned}\max_{h_{D-1}} E\big[U(c, h_1^*, h_2^*, \ldots h_{D-2}^*, h_{D-1}, h_D)\big] = \;& \underbrace{E\big[c(h_1^* + h_2^* + \ldots + h_{D-1} + h_D)\big]}_{\text{expected total consumption}} \\ & - \underbrace{f(h_1^*) - \ldots - f(h_{D-2}^*) - f(h_{D-1}) - E\big[f(h_D)\big]}_{\text{cost from working in each period}}\end{aligned}$$

The agent chooses h_{D-1}^* to maximize expected utility. Both optimal labor supply and expected utility can be expressed as functions of to-date earnings: $h_d^* = g(h_{D-1})$, and $Eu_d(h_{D-1})$. This gives a general expression for expected utility EU_{d+1} in any period d given optimal labor supply choices in subsequent periods. The agent chooses h_d^* to maximize EU_{d+1} for every period d.

Since EU_{d+1} is non-linear, I solve this model using grid search for a particular set of values

$D, p, \tau_0, \tau_1, z^*$, and e.[3] Figure 3 presents the labor supply choices in each period d for an agent with $n = 1$. In each period, and for any hours worked to date, the agent has a unique optimal labor supply choice h^*. In the second period, the worker is fairly insensitive to hours worked in the first. As time progresses though, the worker becomes more sensitive to average hours worked to date. If average hours worked is low, meaning the worker experienced a number of unemployed periods and her earnings will likely be low enough that she won't face the higher tax, she proposes to work her maximum number of hours in the next period. If average hours worked is high, the worker scales back her hours choice.

Panel A in Figure 4 presents the same results but for three agents with different levels of ability n. This means these agents incur different costs from work and will have different optimal labor supply choices. Though levels and sensitivities vary, the overall pattern remains the same. Regardless of $h_{d=1}$, workers choose roughly the same $h^*_{d=2}$ given an expectation of future employment, but labor supply choices adjust in each subsequent period as their uncertainty is resolved. All three agents choose to work their h_{max} if average hours to date are sufficiently low, and if to-date hours are sufficiently high, they work their optimal number of hours given the higher tax rate.

Panel B in Figure 4 presents results for three agents with the same ability n but three different elasticities e. Changes in e reflect workers' sensitivity to the change in the cost of work. Though levels and sensitivities vary, the overall pattern is the same. Workers choose $h^*_{d=2}$ that reflects an expectation of future employment, but labor supply choices change in each period as that uncertainty is resolved. All three agents choose to work the maximum number of hours if average hours to date are low, meaning they have a low chance of facing tax rate τ_1.

Next, I solve the model for 500 levels of n.[4] For each n, I solve the full dynamic programming model, saving all possible earnings and choices in each period. I then select a single earnings sequence for each agent. I select a sequence by, first, identifying each agent's $h^*_{d=1}$. For this period and all others, the agent's actual $h_d = h^*_d$ with probability p, and $h_d = 0$ with probability $1-p$. I recover each worker's proposed hours choice and actual hours in each subsequent period given realized earnings to-date. This simulation yields one employment history for all levels of n. I calculate the agent's predicted annual earnings as of each period d, where $d < D : \hat{Y}_d = \frac{D}{d}\sum_{t=1}^{d} y_t$. I repeat this process for the same levels of e used above.

Figure 5 presents the kernel densities for $\hat{Y}_4$ and $\hat{Y}_8$, as well as actual total earnings. There are three main takeaways from this simulation. The first is that agents exhibit greater bunching in predicted annual earnings at the end of the year than earlier in the year. As of the fourth period, the distribution exhibits limited mass at the kink point,[5] but as time progresses and labor

[3] I present results from a model in which I use $D = 12$, $p = .8$, $\tau_0 = 0$, $\tau_1 = .3$, and $e = .5$, but results are qualitatively similar for alternative values of each parameter. For each period, I construct a grid of discrete levels of possible to-date earnings and hours choices, $z \times h$, where $z = [0, d \times wh_{\text{max}}]$ and $h = [0, wn(1-\tau_0)^e]$. Specifically, I use 200 possible hours choices and 5000 earnings possibilities. Using a coarser or richer grid would not affect my main results.

[4] I sample random levels of n from $\sim \mathcal{N}(1, .01^2)$.

[5] The distribution is bi-modal because some workers will proposed to work more early in the year, anticipating

supply decisions move toward each agent's extreme, bunching increases. The second takeaway is that the amount of bunching is affected by workers' elasticity. Workers with higher elasticity are more sensitive to tax policy, and exhibit greater bunching, as predicted. The third takeaway is that, with a low elasticity, bunching is not conspicuous enough in any period to be identified from the standard bunching estimator. However, the earnings distribution at year's end[6] is distinguishable from the distribution of predicted annual earnings as of period 4, and the difference between these distributions is due to the agent's sensitivity to tax policy.

3 Data

3.1 SIPP

I use the 1996, 2001, 2004 and 2008 panels of the Survey of Income and Program Participation (SIPP).[7] The SIPP is a nationally representative sample of approximately 35,000 primarily low-income households administered by the Census Bureau over the course of three to four years. Respondents are interviewed about their employment, hours worked, wages, earned and unearned income, household composition, and participation in government programs, among much else, every four months. Most important for my purposes, the SIPP asks respondents to report employment and earnings information at the time of the survey and recall this information for each of the three previous months. The SIPP is the only survey that captures individual and household earnings at multiple continuous periods across multiple calendar years.

I identify likely tax units from SIPP survey units and narrow my sample to households for whom I can credibly identify tax liabilities, who are likely to face non-linear tax policy and who are firmly attached to the labor force. First, I limit my analysis to households that contain only one family, following Yang (2018). I further restrict to households for whom I have complete information about earnings, employment and household configuration throughout each calendar year for all household members. This restriction ensures that my estimates of tax liabilities are not confounded by either changes in household composition or missing earnings information. I also restrict to households that include at least one working-age member between the ages of 25 and 55, and I restrict to households where only the household head and spouse, as opposed to siblings or adult children, have earnings, because I cannot distinguish whether these households represent a single tax unit or multiple. Finally, I restrict to households who have non-zero earnings in every calendar quarter, and whose total earnings in any set of three consecutive quarters

that they might become unemployed at some point, but in fact are employed in each period. This adjustment captures a negative intensive margin response to τ_1. An alternative set of parameter values could yield a mass of workers with low proposed earnings who increase labor supply as the year progresses.

[6] One can think of this distribution as the earnings researchers observe in annual tax data.

[7] I do not use the most recent SIPP panel, which was first fielded in 2014. In the newly redesigned SIPP, respondents are asked to recall employment, earnings, and program participation for each month in the calendar year. I find that this reform tends to worsen cross-calendar year seam bias, which poses a unique threat to my approach. I find that employment rates exhibit a distinct discontinuity within households between December and January that I do not observe in other panels.

is between $2,000 and $75,000. The purpose of these restrictions is to focus only on the intensive margin response and limit attention to households who are more likely to face some uncertainty in their tax rates.

My final sample includes earnings information for approximately 12,000 unique households and 18,000 unique tax units (i.e., households by tax year). Table 1 summarizes mean values for key demographic characteristics in my sample, and how my restrictions affect the composition of my sample.

For each tax unit, I identify each head and spouses' total earned income in each month from both wages and self-employment. I sum both sources of earnings within each wave and quarter for each tax year. I also identify each households' unearned income each month (e.g., Social Security and unemployment insurance benefits).

SIPP households also complete various supplemental interviews in each panel. Two of these topical modules ask respondents about variables particular to tax filing, including: property tax bill, amount of itemized deductions, retirement account contributions and deductions, capital gains and losses, and child and dependent care expenses. When available, I associate each adult in each tax unit with tax-relevant variables from these topical modules.

3.2 California Administrative Data

I start with program rosters for California's instantiation of the Supplemental Nutrition Assistance Program (SNAP), known as CalFresh, between 2014 and 2017. These records capture every recipients' per month enrollment in the program, the cases in which they were enrolled, and their demographic characteristics. In 2017, I observe approximately 5.6 million unique individuals across 2.9 million unique SNAP cases. Of these 5.6 million individuals, 2.5 million were younger than 18. Total caseloads are fairly constant over the four years in my sample.

I associate each adult enrolled in SNAP with their quarterly employer-reported earnings records. These records are collected by the California Employment Development Department (EDD), which administers the state's unemployment insurance program. I observe the earnings of each individual for six quarters prior to their enrollment in SNAP, every quarter in which they're enrolled, and 18 months after their last month enrolled. This means that even if an adult is only enrolled for a handful of months in 2016, I still observe their earnings for most, if not all of, 2015 and 2017.

I then match each individual in the SNAP program rosters to the universe of California state tax returns between 2015 and 2017. For each return, I observe basic information about the composition of the tax unit, as well as all variables on the primary state tax form (Form 540). For e-filed returns, I also observe all variables on the Form 1040. I also observe select variables from individuals' information returns, including total wages reported on the W2. Together, these forms allow me to observe all the relevant tax information (i.e., unearned income, deductions, capital gains and losses, etc) necessary to identify income tax rates.

Between 33 and 34 million individuals appear on a state tax return in each tax year. Of the

5.6 million individuals enrolled in SNAP in 2017, about 3.7 million appeared across 1.9 million unique state returns in tax year 2017. Roughly 38 percent of those 3.7 million were a head or a spouse on a return, and the remainder were dependents. These counts and fractions are fairly stable over the three years in my sample. Of the 1.4 million individuals who enrolled in SNAP and appear on a return as a head or spouse, 77 percent have positive EDD wages.

For my primary sample, I implement similar restrictions to those I applied to my SIPP sample, again, in order to focus attention on working-age households firmly attached to the labor force and for whom I can confidently estimate likely tax rates. I limit to households with an adult between the ages of 25 and 55, who have non-zero earnings in all quarters in each tax year, and whose earnings in any consecutive sequence of three quarters is between $2,000 and $75,000. I further restrict to California tax units in which all members were enrolled in SNAP for at least one month in the respective tax year and whose reported AGI matches their total quarterly earnings. I limit my analysis to these households in order to ensure that I can infer changes in true tax incentives based on changes in quarterly earnings. If tax units contained other adults with earnings or earnings from non-UI covered employment, then I could not rely only on quarterly earnings to predict future tax incentives. Table 2 summarizes how these restrictions affect the characteristics of my sample.

3.3 TAXSIM

I use NBER's TAXSIM program to identify households' average and marginal tax rates (Feenberg and Coutts, 1993). TAXSIM allows users to input key tax-related information for a given household, and returns income tax calculations using federal and state income tax policies for any year between 1960 and 2023. I input filing status, state, number of dependents, ages, earned and unearned income, and a variety of possible deductions. The program returns federal and state income tax liabilities and marginal tax rates (inclusive and exclusive of FICA taxes), federal and state EITC amounts, and more for all households. I use TAXSIM to identify how households' likely tax rates and liabilities change over the course of the tax year by summing household earnings over various periods and inputting these sums into TAXSIM. For example, if I want to identify the likely tax rate on a household's fourth quarter earnings, I sum that household's earnings through the first three quarters of a tax year, and input this sum, as opposed to a households' true annual income.

4 Motivation

4.1 Non-linear Tax Policy

I use output from TAXSIM to document cross-sectional variation in average and marginal tax rates for households in the SIPP and SNAP samples.

Figure 6 illustrates how average and marginal tax rates vary by household income in 1997 versus 2012 for married SIPP households with 0, 1, 2 or 3+ dependents. I group households into bins of $2,500, and within each bin, I calculate the average marginal and average income tax rates (combining federal and state income taxes as well as payroll taxes) that all households within that bin face on their annual earnings.

Households with children and very low earnings tend to face a steeply negative marginal and average tax rate on annual income. In 2012, the average household with children and annual earnings below $10,000 faced a marginal tax rate between negative 30 and 50 percent. Income taxes boosted these households' net income by 30 to 40 percent. When pre-tax household income eclipses about $45,000, the average tax rate settles to around 30 percent, regardless of the household's number of children.

In Figure 7, I plot the marginal and average tax rates by number of dependents and annual income for single and married SNAP households in California in 2017. The patterns are roughly the same as in the SIPP sample. For households with children and incomes below roughly $10,000, marginal and average tax rates are steeply negative. Thanks to California's supplement to the federal EITC, they are even lower than the national averages. Households with children and income in the phase-in portion of the EITC face both a marginal and average tax rate between negative 50 and 75 percent.

Both figures make clear how important the EITC is for eligible households: Negative marginal tax rates align with the phase-in part of the EITC schedule, and the highest marginal tax rates align with the phase-out range. When household income exceeds the maximum eligible income for the EITC, average and marginal tax rates appear to converge and hold steady at around 25 to 30 percent, regardless of household type. Figure 6 illustrates the impact of reforms that increased the maximum credit amount. Households with dependents and incomes below about $10,000 benefited from program expansions in 2001 and 2009, as well as the introduction of numerous state supplements.

The key takeaway, which is clear in both figures, is that similar households within a fairly narrow income range can face starkly different tax incentives as a function of their to-date earnings. As earnings rise from around $15,000 to $30,000, households quickly face steep positive marginal tax rates and positive tax liabilities.

These different tax rates will only impact labor supply if households appreciate that tax policy is non-linear. Surveys and interviews of low-income workers suggest widespread awareness that tax filing is often associated with receiving a refund (Edin et al., 2014; Halpern-Meekin et al., 2015; Smeeding et al., 2000), but only half can recognize the EITC by name (Bhargava and Manoli, 2015) and a minority are aware of the program's benefit structure (Chetty et al., 2013; Smeeding et al., 2000). Still, these surveys and other ethnographic evidence document an appreciation among EITC-eligible taxpayers that income tax policy boosts their income (Halpern-Meekin et al., 2015). For income volatility to imply tax uncertainty, households must only grasp that these benefits are reduced when earnings exceed some level.

4.2 Income variation

Next, I illustrate how pre- and post-tax income varies significantly across the tax year for a substantial share of households. In the SIPP, I identify household income through each month of the calendar year, and in the California administrative data, through every quarter. Using this observed income, I project forward what household income might be for the entire year as of the end of each period. Using these income projections, I also identify each household's predicted year-end average tax rate through each period.

In Table 3, I report the share of SNAP and SIPP households for whom the absolute difference between their predicted annual income, predicted average and marginal income tax rates, and predicted total EITC amounts as of the end of each calendar quarter are more than particular units away from their year-end values. For one-third of households, predicted annual income as of the end of the first quarter is more than $5,000 from their actual annual income. By the end of September, however, this share falls to just 5 percent. Only three percent of households have a total EITC amount that is more than $1,000 different than their predicted EITC value as of the end of the third quarter. Table 3 also reports the average standard deviation for each variable within each tax year across all SNAP households. Figure 11 plots the distribution of standard deviations in predicted earnings within tax years across all SNAP households.

Table 4 also reports similar shares but compares differences between, as opposed to within, tax years. Cross-year variation is more significant than within-year variation, which reflects both how volatile earnings can be over longer time periods and how households gain clarity about likely earnings within tax years. The various panels in Figure 8 plot the distributions of these differences. The red dotted lines indicate the median of the absolute value of all the differences. Half of all SNAP households experience a year-over-year change in wage earnings of at least $4,500. This corresponds to half of households experiencing an average tax rate one year that is more than seven percentage points different than what they faced the year before.

This variation does not necessarily imply unexpected volatility for all workers. Workers might be able to anticipate future spikes and dips in work hours and wages, which a researcher cannot observe. Workers might also choose to substitute when and how much they work in response to changes in the personal opportunity cost of work, unrelated to tax incentives. Still, it is reasonable to expect that, especially for lower-income households, these variations do reflect some volatility. Unemployment spells can be unanticipated, both in their occurrence and their length, which is why we have a large social program to insure against those risks. Firms also exert significant control over many workers' schedules, which translates into volatile hours worked and total earnings (Gerstel and Clawson, 2018; Golden, 2015; Maag et al., 2017; Schneider and Harknett, 2019). Even if wages are stable, year-end wage bonuses or unexpected dividends can also affect annual income and tax liability (Saez, 1999). I cannot distinguish which of the households I observe in the SIPP or SNAP data are those whose change in earnings over time represents unanticipated volatility or intentional reallocations between labor and leisure. I rely on others' work demonstrating that volatility is common enough that a significant share of workers in both samples experience these idiosyncratic wage shocks.

Figure 9 provides suggestive evidence that households do shift in the direction of maximizing their after-tax income. I plot the distribution of predicted annual earnings among SNAP households with two dependents in 2017 as of the end of quarter 1 and quarter 3, alongside their actual earnings. The distributions look fairly similar, reflecting that average distribution of quarterly earnings in the SNAP sample is fairly constant over time. However, note that a greater mass of SNAP households are predicted to have very modest earnings at the beginning of the year than later in the year. This mass appears to shift towards the center of the distribution, where households would maximize their total EITC receipt, by the end of the year. This shift mirrors the predictions from my model. It is suggestive of an increase in labor supply over the course of the tax year on the part of households facing a tax incentive to work more. Further, a greater mass of taxpayers appears to be near the second kink point of the EITC schedule as of the end of Q3, but that mass decreases in the following quarter. There is limited if any evidence of any shift on the part of households with earnings predicted to be in the phase-out portion of the EITC, however.

Figure 10 provides additional evidence that the change in the distribution in annual earnings reflects a shift along the EITC schedule. For all SNAP households, I identify their predicted state and federal EITC amounts as of the end of the third quarter, assuming that fourth quarter earnings equal the average of the first three. I subtract this amount from the households' actual EITC amount, and plot this difference over predicted annual income as of the end of the third quarter. Households whose predicted earnings would place them on the phase-in part of the EITC schedule exhibit the greatest difference between their actual and predicted EITC amounts.

Of course, the shifts identified in Figure 9 and Figure 10 may be due to mean reversion: households with low past earnings bounce back to a more normal earnings level over time. They might also reflect households with very low past earnings increasing their labor supply in order to achieve a minimal level of earned income. They might also be noise. I distinguish these explanations from an intentional response to tax incentives in my empirical analysis, which I detail in the following section.

5 Empirical Framework

Consider the simple cross-sectional estimation of the elasticity of labor supply with respect to net of tax wages:

$$\underbrace{l_i}_{\substack{\text{labor}\\\text{supply}}} = \alpha + \underbrace{\beta w_i}_{\substack{\text{net of tax}\\\text{wage rate}}} + \underbrace{\gamma y_i}_{\substack{\text{non-labor}\\\text{income}}} + \underbrace{X_i' \delta}_{\substack{\text{demographic}\\\text{controls}}} + \varepsilon_i \tag{1}$$

β captures uncompensated wage effects on labor supply l, and γ captures income effects. Estimates of β using cross-sectional variation in w are biased due to omitted variable and simultaneity: "taste for work" is correlated with w, and changes in l affect w via the tax rate the agent faces. To overcome both concerns, recent work tends to estimate β by measuring employment responses to tax reforms that affect similar workers differently and where changes in w are plausibly exogenous. These analyses regularly use repeated cross-sectional surveys to identify how employment rates change between treated and untreated workers. Since tax reforms can affect the composition of these groups of workers, one generally cannot rely on these settings to identify intensive margin effects. That these reforms might also affect the equilibrium wage raises additional identification concerns (Rothstein, 2010).

I proposed to identify β by measuring how earnings change in response to changes in expected tax rates faced by similar workers within the same tax year and the same worker across subsequent tax years. I estimate this response across multiple specifications, but each approach borrows from Equation 1 in that I regress a measure of labor supply, y, in a given period on an observation's expected net of tax earnings rate for that period, ω, along with a variety of controls. The models differ in how I define and instrument for ω and how I test for varied earnings responses across the tax year.

First, I define my key independent variable: ω, the net of tax wage rate, or the share of the next period's projected earnings that the household expects to retain after taxes. It is a function of household i's earnings from the three previous quarters, z, and tax policy in year y and state s.[8] A positive one unit increase in this ratio equals a 100 percentage point increase in the observation's net of tax predicted earnings.

$$\omega_{iyq} = \underbrace{\left(\frac{1}{3}z_{iyq}\right)^{-1}}_{\text{average to-date earnings}} \Bigg(\underbrace{f(\frac{4}{3}z_{iyq})}_{\text{post-tax predicted quarter's earnings}} - \underbrace{f(z_{iyq})}_{\text{post-tax previous three quarter's earnings}} \Bigg),$$

$$\text{where } f(y) = y - \tau_{ys}(y) \text{ and}$$

$$z_{iyq} = \sum_{q=-3}^{-1} y_{iyq}$$

Figure 12 summarizes the distribution of $\omega_{q=4}$ for all SIPP and SNAP tax units.

5.1 Effect in Q4

In the first model, I identify how each households' earnings in the fourth quarter vary with their predicted net of tax earnings rate in the fourth quarter. I control for each household's earnings from the three preceding quarters, z, as well as state-by-year, household type and the demographic fixed effects.[9] The parameter β is a measure of households' earnings elasticity.

[8] Note that I use y to denote quarterly earnings, as well as the tax year. When y is in a subscript, it denotes a tax year. Otherwise, it represents earnings.

[9] Household type is the interaction between filing status (single versus married) and number of dependents (0, 1, 2 and 3+). Demographic fixed effects are the interaction of the household head's race, binned age levels (25-34, 35-44,

$$\underbrace{\ln y_{iy,q=4}}_{\text{log Q4 earnings}} = \beta \underbrace{\omega_{iy,q=4}}_{\substack{\text{predicted}\\ \text{net of tax}\\ \text{wage rate}}} + \gamma \underbrace{\ln z_{iy,q=4}}_{\substack{\text{log Q1-Q3}\\ \text{earnings}}} + \underbrace{\alpha_i}_{\substack{\text{household}\\ \text{fixed effect}}} + \underbrace{\theta_{ys} + \theta_h + \theta_x}_{\substack{\text{year} \times \text{state,}\\ \text{household type,}\\ \text{demographic}\\ \text{fixed effects}}} + \varepsilon_{iy,q=4} \tag{2}$$

This approach addresses the simultaneity problem, because I identify how earnings in a given period vary with predicted tax incentives in that period. I also account for unobserved and fixed characteristics of each household by measuring this response within the same household over subsequent tax years. However, since ω is a function of previous earnings and despite controlling for z, my estimate of β is confounded by relationship between earnings and lagged earnings. Households with lower past earnings are likely to have lower future earnings. Moreover, the correlation between past and future earnings likely varies over levels of z and ω. For example, households with especially low z or high ω might exhibit higher future earnings if they are reverting back to average earnings level or working to achieve a certain minimal level of earnings.
par

5.2 "Rolling Window" Approach

I distinguish households' response to tax policy from standard serial correlation in earnings by comparing the response in the actual tax year to those that in simulated tax years that end in the first, second or third calendar quarter. I construct simulated tax years comprised of all possible sequences of four consecutive quarters, and I identify ω for each sequence.[10] Figure 1 illustrates how these sequences are constructed.

I stack these sequences for each household, and estimate the following variation of Equation 2.

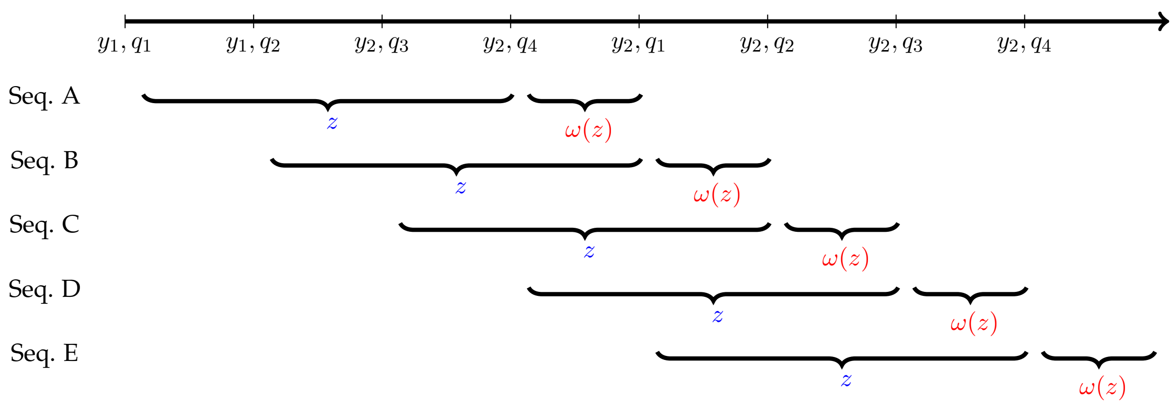
$$\underbrace{\ln y_{iyq}}_{\text{log earnings}} = \beta\omega_{iyq} + \pi_q \underbrace{\omega_{iyq} \cdot Q}_{\substack{\text{predicted net of}\\ \text{tax wage rate} \times\\ \text{quarter dummy}}} + \gamma \underbrace{\ln z_{iyq}}_{\substack{\text{log sum of last}\\ \text{three quarters'}\\ \text{earnings}}} + \underbrace{\alpha_i}_{\substack{\text{household}\\ \text{fixed effect}}} + \underbrace{\theta_{ys} + \theta_h + \theta_x}_{\substack{\text{year} \times \text{state,}\\ \text{household type,}\\ \text{demographic}\\ \text{fixed effects}}} + \varepsilon_{iyq} \tag{3}$$

I interact ω with an indicator for calendar quarter, Q. The coefficient on this interaction, π, captures the unique relationship between $\ln y$ and ω, or the excess sensitivity of $\ln y$ to ω, in calendar quarters 2, 3 and 4 relative to quarter 1. If serial correlation in earnings does not have a systematic seasonal pattern, then each π_q should equal zero. If households adjust earnings to

45-54), language, and gender.

[10] This means that I have four measures of ω and z for each tax year, one for each quarter, for every tax unit. If I observe a household for three tax years, I have twelve observations for that household.

Figure 1: Constructing sequences of four consecutive quarters



Notes. Figure 1 illustrates how I construct simulated tax years. Each sequence of four consecutive quarters with earnings information is a simulated tax year, and I use the sum of the first three quarters' earnings in this simulated tax year z to predict earnings in the next quarter, and I identify ω as a function of z. Note that Seqs. A and E align with true tax years.

predicted tax incentives when those incentives are more binding at year's end, then $\pi_{q=4}$ will be positive. Indeed, my preferred estimate of labor supply elasticity will be my estimate of $\pi_{q=4}$ from this model. This coefficient identifies the relationship between log earnings and predicted net of tax wage rate in the fourth quarter, netting out the standard period-to-period relationship between past and future earnings captured by β.

I then estimate four additional variations of Equation 3.

First, the linear parameters π_q mask important heterogeneity in the response across both earnings and values of ω. To identify whether responses estimated in Equation 3 are driven by households facing particular levels of ω, I re-estimate Equation 3, but replace the continuous measure of ω with binned levels of ω. I group households into those with ω less than .55, greater than 1.1, and increments of .05 in between.

$$\ln y_{iyq} = \beta\Omega_{iyq} + \pi_{q\omega}\Omega_{iyq} \cdot Q + \gamma \ln z_{iyq} + \alpha_i + \theta_{ys} + \theta_h + \theta_x + \epsilon_{iyq} \tag{4}$$

Next, to illustrate how this response varies with income as opposed to tax rates, I replaced binned values of ω with binned values of predicted annual income, ζ, where predicted annual income equals earnings from the previous three quarters, z, plus predicted income in the next quarter, $\frac{4}{3}z$. I group households into $2,000 bins. I interact values of ζ with indicators for calendar quarters 2, 3 and 4.

$$\ln y_{iyq} = \beta\zeta_{iyq} + \pi_{q\zeta}\zeta_{iyq} \cdot Q + \alpha_i + \theta_{ys} + \theta_h + \theta_x + \epsilon_{iyq} \tag{5}$$

To evaluate whether households respond to their expected average tax rate or their expected marginal tax rate, I re-estimate Equation 3 but replace ω with households' predicted marginal tax rate on the last predicted dollar earned in the subsequent quarter.

Finally, to test whether Equation 3 identifies a change in the amount worked, as opposed to earned, I use log hours as the outcome variable, as opposed to log earnings. I only observe reported hours in the SIPP, meaning I can only estimate this response in that sample.[11]

5.3 Instrumenting for $\omega_{q=4}$

Thus far, I have assumed workers understand their final period's net of tax wage rate only after realizing their likely annual income. In actuality, households have more information about what their true tax rate is likely to be even if they are ignorant of their future earnings. Households can make reasonable estimates of their annual income and their likely net of tax wage rate on January 1st using other information, some of which is observed by the researchers (e.g., household composition, state, year, occupation, age) and some of which is not (e.g., full work histories, preferences, understandings of their local labor market, or agreements with employers).

I test whether households exhibit this behavior by estimating the relationship between labor supply in each calendar quarter, y_q, and households' expected year-end net of tax wage rate $\omega_{q=4}$ as of each quarter. I instrument for each household's $\omega_{q=4}$ with similar households' actual $\omega_{q=4}$. I regress each households' actual $\omega_{q=4}$ on a vector of household characteristics (age, race, gender, state, year) and their predicted year-end income given their to-date earnings within the tax year, $\tilde{z}$, as of four different periods: the start of Q1, the end of Q1, the end of Q2 and the end of Q3.[12] Figure 2 illustrates how I construct these sequences.

I stack each sequence for each tax unit, and regress each tax unit's actual $\omega_{q=4}$ on earnings as of the start of each period as well as various fixed effects.

$$\underbrace{\omega_{iy,q=4}}_{\substack{\text{household } i\text{'s}\\ \text{true } \omega_{q=4}\\ \text{in year } y}} = \underbrace{\tilde{z}_{iyq}}_{\substack{\text{earnings as of}\\ \text{the start of}\\ \text{qtr } q \text{ in year } y}} + \underbrace{\theta_{ys} + \theta_h + \theta_x}_{\substack{\text{year} \times \text{state,}\\ \text{household type,}\\ \text{demographics}}} + \epsilon_{iyq} \tag{6}$$

Given estimates from Equation 6, I can predict each tax unit's year-end $\hat{\omega}$ as of every sequence. I then estimate Equation 7.

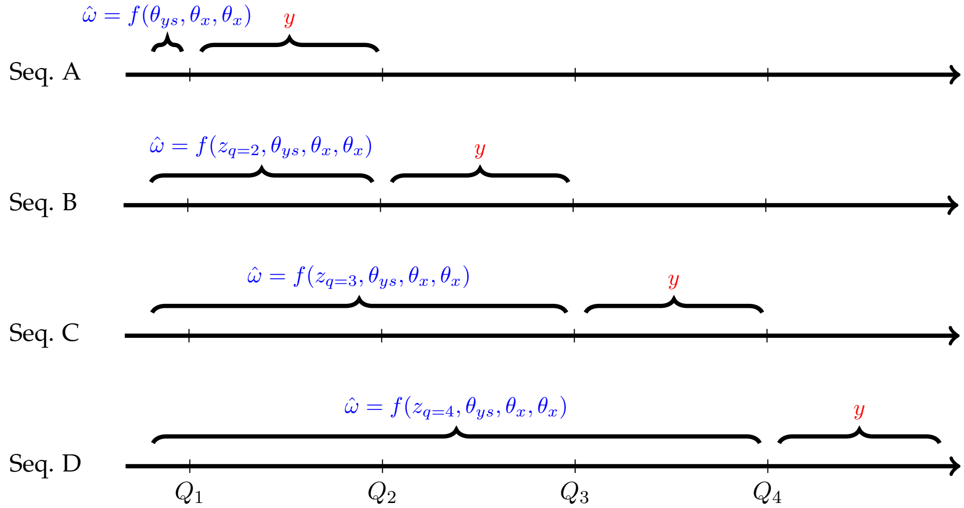
$$\ln y_{iyq} = \beta\hat{\omega}_{iy,q=4} + \pi_q\hat{\omega}_{iy,q=4} \cdot Q + \gamma \ln z_{iyq} + \alpha + \theta_{ys} + \theta_h + \theta_x + \varepsilon_{iyq} \tag{7}$$

As in Equation 3, I control for the log of the three previous quarter's earnings to account

[11]SIPP respondents only report average hours worked in a representative week in the calendar month.

[12]Since households have no earned income as of the beginning of the tax year, I do not include a control for earnings when predicting ω at the start of the tax year.

Figure 2: Instrumenting for $\hat{\omega}$ as of the start of each quarter



Notes. Figure 2 illustrates the sequences for which I identify $\hat{\omega_q}$. In Seq. A, I use non-earnings information observable before the tax year begins, including indicators for filing status by number of dependents, year by state, and demographic characteristics. I identify the household's likely $\omega_{q=4}$ from these variables, and then identify whether $y_{q=1}$ positively covaries with this prediction. In Seq. B, I use the same information as in Seq. A and add in predicted earnings information from quarter 1, and so on.

for dynamic earnings processes. The coefficient, π_q, on the interaction of ω and the quarter Q, identifies whether there is a unique relationship between one's expected year-end net of tax wage rate and one's earnings in the following quarter. If workers can forecast their year-end $\omega_{q=4}$ accurately, we might expect each value of π_q to be zero. Households can respond to their year-end tax incentives but that response does not have to principally occur at year's end.

6 Results

6.1 Main Results

Table 5 and Table 6 summarize estimates of β and π across different versions of Equation 3 among SIPP and SNAP households, respectively. When I do not account for previous earnings, I find that, in both samples, households with lower previous earnings tend to have low earnings in the final quarter as well. When I control for z, earnings response is positive in the SNAP sample, but statistically insignificant in the SIPP sample. When I include the household fixed effect, I estimate an earnings elasticity in the SNAP sample to be .07; the counterpart estimate in the SIPP sample is negative, but highly insignificant.

These estimates are likely biased because Equation 2 does not account for serial correlation issues. Table 7 and Table 8 summarize estimates of β and π across different versions of Equation 3 among SIPP and SNAP households, respectively, which aims to address this concern. Though magnitudes differ, overall patterns are similar in both samples. When I do not account for z, I find that households with higher ω have lower earnings in the subsequent quarter, which captures the serial correlation issue raised above: Households with lower past earnings are expected to continue to have low earnings, despite their high ω. If I include z but only account for heterogeneity across households by controlling for demographics, household type, and state-by-year effects, θ_x, θ_h, and θ_{ys}, I find almost no relationship between short-term tax incentives and earnings. When I include a household fixed effect, α_i, I recover the predicted pattern. When the same household faces a higher ω, they tend to earn more in the next period, and this relationship is strongest at year's end when ω is most binding. Recall that my preferred estimate of the intensive margin labor supply elasticity is $\pi_{q=4}$; this parameter identifies the excess sensitivity of earnings to ω at the end of the real tax year, netting out the period-to-period relationship between earnings levels captured by β. In the SIPP sample, I estimate an intensive margin labor supply elasticity of .08. In the SNAP sample, I estimate the same elasticity to be .18.

Figure 13 plots estimates of $\pi_{q\omega}$ from Equation 4. The positive relationship between earnings and the predicted net of tax wage rate is driven largely by households facing an ω greater than one, and this relationship is strongest in the fourth quarter. This response is observable in both samples, though it's clearest in the larger SNAP sample, where earnings are measured with greater accuracy and there are a greater share of households with especially low earnings.

Figure 14 plots estimates from Equation 5 and further illustrates that this response is largely driven by incentives created by the EITC. I overlay these estimates on the EITC schedule for single filers with two dependents in 2017. For Panel B, I use the combined federal and state EITC schedule. The non-linear relationship between $\ln y$ and levels of predicted income, ζ, suggests that households who expect to be in the phase-in range of the EITC are especially likely to increase earnings in the following quarter. That this effect is particularly pronounced in the fourth quarter suggests it is a response to tax incentives and not due to other serial patterns.

Table 9 and Table 10 summarize estimates from Equation 7. Results are similar across the four models, because ω is a function of these controls. Despite the alternative definition of ω, my results are quite similar to those summarized in Table 7 and Table 8. Households still appear to increase subsequent labor supply when they expect to face a positive ω and this response is strongest at the end of the year.

6.2 Supplementary Results

6.2.1 Response to Marginal Tax Rate

I estimate another version of Equation 3 where I use households' predicted combined federal and state marginal income tax rate as my independent variable. Results from this estimation are summarized in Table 11 and Table 12. In the SIPP, I find no relationship between earnings and marginal tax rate. In every model, the effect is near zero. This is likely due to my SIPP sample having too few households with low enough earnings to face steeply negative tax marginal rates. My preferred estimate in the SNAP sample, however, recovers the expected negative relationship, and the effect is again clearest at the end of the year. The response is much more limited, which suggests that households are more responsive to their expected average tax rate.

6.2.2 Change in Hours Worked

Table 13 summarize estimates of Equation 3 when I use log hours as my outcome variable in Equation 3, as opposed to log earnings. Results are similar to those summarized in Table 7, but I cannot rule out that the fourth quarter response is the same as the response in the second or third quarter. I interpret these findings to suggest that labor supply plays a role in the earnings elasticity reported in Table 7, but measurement issues in the SIPP make it difficult to draw a strong conclusion.

6.2.3 Subgroups

Next, I test whether these responses vary by filing status and presence of dependents. Estimates from analyses within these subgroups are summarized in Table 14 and Table 15. I find that married households are more sensitive to their predicted net of tax wage at year's end, which is consistent with previous literature. Results differ between the SIPP and administrative data for households with or without children.[13]

Table 16 summarizes results from additional estimates of Equation 3 restricting to the following subgroups in the SIPP.

- **Hourly workers.** We might expect this response to be clearest among hourly workers, since they are more subject to scheduling volatility and have greater flexibility in adjusting shifts and schedules in a particular quarter. Column 1 in Table 16 presents estimates of Equation 3 limited to households in which the head or spouse is an hourly worker. Though the effect is slightly higher, the estimates are not dramatically different.

- **Self-employed workers.** Saez (2010) shows that self-employed workers exhibit much more significant bunching at the first kink-point in the EITC schedule than wage-earners, which implies they have a higher taxable income elasticity. To test whether this bunching is due to a true earnings response or just tax manipulation, I estimate my model among the small number of households in my SIPP sample in which a head or spouse reports having some self-employment earnings in the tax year. I find that the response to year-end tax incentives is indeed much higher among households with a self-employed workers. My estimate of $\pi_{q=4}$ is .24, which is three times larger than the main result reported in Table 7. However, when I test whether that response is driven by changes in self-employed earnings, the effect disappears (Column 3). Together, these estimates paint a mixed picture about whether greater bunching among self-employed workers reflects a real labor supply response. That

[13] Recall that ω accounts for the significantly different tax incentives facing parents versus single adults. This analysis is comparing the earnings response between parents and non-parents who are facing similar tax incentives, and not those with similar levels of earnings.

said, my sample size is small and there are reasonable concerns about measurement of self-employment earnings in the SIPP.

- **Excluding retail workers.** One concern with my approach is that there are significant changes in labor demand and supply around national holidays at the end of the calendar year. For example, if lower-wage retail employees are more likely to work overtime, receive higher wages, or work more frequent shifts in the holiday season independent of tax incentives, this increase in earnings could bias my result. By controlling for occupational fixed effects in the SIPP, as well as household and worker fixed effects in both samples, I should account for this concern. To the extent those concerns remain, however, I also re-estimate Equation 3 in the SIPP excluding retail workers entirely from the sample. My results do not change.

- **Addressing seam bias.** Responses in the SIPP suffer from well-recognized seam bias. Changes in hours worked and wages are reported at the time of each survey. Respondents tend to project their current employment situation backwards, instead of accurately recalling each month's unique values. I address this concern, in part, by restricting my sample to households whose waves align with the tax year, meaning a wave does not stretch from Q4 of one year to Q1 of the next. Results from this model are reported in Column 5 of Table 16. Results are nearly identical to those from Table 7, suggesting this concern is not a significant problem.

6.2.4 Dynamic Panel Bias

My approach raises standard concerns with studies involving dynamic panels and lagged dependent variables. Even though neither of my models include an actual lagged value of y as a regressor, both ω and z are functions of lagged values of y. My preferred version of Equation 3 uses panel fixed effects, which risks introducing a mechanical relationship between the lagged dependent variable and the error term, biasing my estimated coefficients on ω and $\ln(z)$ (Nickell, 1981).

Alternative estimation strategies involving dynamic panel data account for heterogeneity across households without using fixed effects. Estimating these models in my setting presents some challenges. First, available tools for implementing the GMM estimators assume the endogenous term is a single lagged dependent variable. My approach involves five separate endogenous regressors: ω and the interactions with each calendar quarter, plus z, the log of the three quarters' earnings. Instrumenting for each of these variables with their lagged levels and differences yields not only a proliferation of instruments, but also involves differencing lagged values across tax years, as opposed to subsequent quarters. Second, even though I have information about earnings for up to twelve periods for some households, I am interested in the unique effect in particular calendar quarters. This means I only have a maximum of three periods for each household. GMM estimators are more useful when more past and future realization of the dependent variable are available.

Notwithstanding these issues, I estimate the Anderson-Hsiao estimator, which does not require using multiple lagged and future values of the dependent variable (Anderson and Hsiao, 1982).[14] I instrument for $\Delta\omega_{iyq}$ and $\Delta \ln z_{iyq}$ – the difference between each household's value of ω and $\ln z$ in the final tax year and the value from the penultimate tax year – with the furthered lagged values of ω_q and $\ln z$. I then relate the first difference between $\ln y_{iyq}$ in the final and previous period with the instrumented values of $\Delta\omega_{iyq}$ and $\Delta \ln z_{iyq}$.

$$\Delta \ln y_{iyq} = \pi_q \tilde{\omega}_{iyq} + \gamma \tilde{\ln} z_{iyq} + \theta_{ys} + \theta_h + \theta_x + v_{iyq} \tag{8}$$

Table 17 summarizes the results from this estimation among SIPP and SNAP households. The results differ from those presented in Table 7 and Table 8. Among SNAP households, despite noisy estimates, it remains the case that even instrumented earnings tend to increase as the tax year progresses, though the largest response now appears in the third quarter as opposed to the fourth. Among SIPP households, my estimates are dramatically different. The earnings response are substantially lower in the fourth quarter, the opposite of my main result. However, none of my estimates are statistically significant. In addition to the lack of statistical significance, and all of the issues about applying this estimator in my setting, there are other reasons to be cautious in interpreting these results. First, the true values of each parameter should fall between the biased estimates from the OLS and the fixed effects models (Bond, 2002), but all of my estimates in the SIPP sample, and all but my estimate of γ in the SNAP sample, are outside these bounds, suggesting my instruments are not valid. Second, I fail the over-identification test in both versions.

7 Conclusion

This paper studies the impact of tax policy on household labor supply using differences in uncertainty about annual income and tax incentives. Using survey and administrative data, I document significant within and cross year variation in household earnings and implied tax rates on those earnings. I use the fact that uncertainty about tax rates is resolved over the course of the tax year to identify the effect of tax policy on labor supply. I relate household earnings in the final quarter of the tax year to the share of the households' predicted earnings they expect to retain after taxes. I distinguish this response to tax incentives from standard serial correlation in earnings by comparing this response in other quarters, as though those quarters were the end of a tax year. I also use household fixed effects to account for omitted variable bias. I interpret households' excess sensitivity to their predicted net of tax wage rate in the fourth quarter as a

[14] The "Difference GMM" and "Systems GMM" estimators promise increased efficiency over the Anderson-Hsiao estimator by leveraging additional information about the evolution of the lagged dependent variable from additional past of future values of that variable. Given my short panel, I am unable to implement these more popular estimators.

measure of their labor supply elasticity. I conclude that households exhibit a small but non-zero intensive margin response to tax policy. My preferred estimate of the intensive margin labor supply elasticity is between .08 using the SIPP and .18 in the SNAP sample. Finally, I conclude that this effect is driven largely by the steeply negative tax rates created by the phase-in part of the EITC.

This study makes an important contribution to the academic literature studying labor supply response to tax policy. The most common approaches to identifying intensive margin labor supply elasticity suffer from important identification challenges, which my approach overcomes. Leveraging unique panel data, I identify how the same household adjust earnings when tax incentives change within and across tax years. I conclude labor supply elasticity is small, consistent with the rest of the micro literature.

My findings also provide useful guidance to policymakers on two policy issues related to the EITC. For policymakers interested in reforming the EITC to pay out benefits in advance of the tax filing season, I provide evidence about how likely those forecasts are to be wrong. For policymakers interested in encouraging eligible non-filers to claim the EITC or interested in pre-filling their tax returns, I provide evidence about how well within-year earnings can be used to identify likely eligible households and predict their year-end EITC amounts.

Second, I provide evidence about the EITC's effect on intensive margin labor supply. The negative marginal tax rate created by the EITC's phase-in is supposed to increase households hours choice via the substitution effect, and minimize negative labor supply distortions driven by the program's income effect. If the program does not have the intensive margin effect, then the consequence of the program's structure – that the lowest-income households receive limited to no assistance – is less justified. This paper provides evidence that the EITC's phase-in does increase labor supply, as intended. Whether this pro-work effect warrants limiting redistribution to the lowest income households, and whether a basic credit could be efficiently incorporated into current policy, remains an open question. Finally, I find limited evidence that households facing steeply positive marginal tax rates reduce their labor supply. This finding suggests that concerns about significant work disincentives created by some means-tested programs for a small subset of workers is not warranted.

References

Altonji, J. G. (1986). Intertemporal substitution in labor supply: Evidence from micro data. *Journal of Political Economy*, 94(3, Part 2):S176–S215.

Anderson, T. W. and Hsiao, C. (1982). Formulation and estimation of dynamic models using panel data. *Journal of Econometrics*, 18(1):47–82.

Angrist, J. D. (1991). Grouped-data estimation and testing in simple labor-supply models. *Journal of Econometrics*, 47(2-3):243–266.

Bhargava, S. and Manoli, D. (2015). Psychological frictions and the incomplete take-up of social benefits: Evidence from an irs field experiment. *American Economic Review*, 105(11):3489–3529.

Blomquist, S. and Newey, W. (2017). The bunching estimator cannot identify the taxable income elasticity. Technical report, National Bureau of Economic Research.

Blundell, R., Brewer, M., Haan, P., and Shephard, A. (2009). Optimal income taxation of lone mothers: an empirical comparison of the UK and Germany. *The Economic Journal*, 119(535):F101–F121.

Blundell, R., Duncan, A., McCrae, J., and Meghir, C. (2000). The labour market impact of the Working Families' Tax Credit. *Fiscal Studies*, 21(1):75–104.

Blundell, R. and MaCurdy, T. (1999). Labor supply: A review of alternative approaches. In *Handbook of labor economics*, volume 3, pages 1559–1695. Elsevier.

Blundell, R., Pistaferri, L., and Saporta-Eksten, I. (2016). Consumption inequality and family labor supply. *American Economic Review*, 106(2):387–435.

Bond, S. R. (2002). Dynamic panel data models: a guide to micro data methods and practice. *Portuguese economic journal*, 1(2):141–162.

Burtless, G. and Hausman, J. A. (1978). The effect of taxation on labor supply: Evaluating the Gary negative income tax experiment. *Journal of Political Economy*, 86(6):1103–1130.

Camerer, C., Babcock, L., Loewenstein, G., and Thaler, R. (1997). Labor supply of New York City cabdrivers: One day at a time. *The Quarterly Journal of Economics*, 112(2):407–441.

Card, D. and Hyslop, D. R. (2005). Estimating the effects of a time-limited earnings subsidy for welfare-leavers. *Econometrica*, 73(6):1723–1770.

Chetty, R., Friedman, J. N., Olsen, T., and Pistaferri, L. (2011). Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from danish tax records. *The Quarterly Journal of Economics*, 126(2):749–804.

Chetty, R., Friedman, J. N., and Saez, E. (2013). Using differences in knowledge across neighborhoods to uncover the impacts of the EITC on earnings. *American Economic Review*, 103(7):2683–2721.

Chetty, R. and Saez, E. (2013). Teaching the tax code: Earnings responses to an experiment with EITC recipients. *American Economic Journal: Applied Economics*, 5(1):1–31.

Crawford, V. P. and Meng, J. (2011). New York City cab drivers' labor supply revisited: Reference-dependent preferences with rational-expectations targets for hours and income. *American Economic Review*, 101(5):1912–32.

Edin, K., Tach, L., and Halpern-Meekin, S. (2014). Tax code knowledge and behavioral responses among EITC recipients: Policy insights from qualitative data. *Journal of Policy Analysis and Management*, 33(2):413–439.

Eissa, N. and Hoynes, H. W. (2006). Behavioral responses to taxes: Lessons from the EITC and labor supply. *Tax policy and the economy*, 20:73–110.

Eissa, N. and Liebman, J. B. (1996). Labor supply response to the Earned Income Tax Credit. *The Quarterly Journal of Economics*, 111(2):605–637.

Farber, H. S. (2008). Reference-dependent preferences and labor supply: The case of New York City taxi drivers. *American Economic Review*, 98(3):1069–82.

Feenberg, D. and Coutts, E. (1993). An introduction to the taxsim model. *Journal of Policy Analysis and management*, 12(1):189–194.

Fehr, E. and Goette, L. (2007). Do workers work more if wages are high? evidence from a randomized field experiment. *American Economic Review*, 97(1):298–317.

Feldman, N. E., Katuščák, P., and Kawano, L. (2016). Taxpayer confusion: Evidence from the child tax credit. *American Economic Review*, 106(3):807–35.

Gelber, A. M., Jones, D., and Sacks, D. W. (2020). Estimating adjustment frictions using nonlinear budget sets: Method and evidence from the earnings test. *American Economic Journal: Applied Economics*, 12(1):1–31.

Gelber, A. M. and Mitchell, J. W. (2011). Taxes and time allocation: Evidence from single women and men. *The Review of Economic Studies*, 79(3):863–897.

Gerstel, N. and Clawson, D. (2018). Control over time: Employers, workers, and families shaping work schedules. *Annual Review of Sociology*, 44:77–97.

Golden, L. (2015). Irregular work scheduling and its consequences. *Economic Policy Institute Briefing Paper*, (394).

Grogger, J. (2003). The effects of time limits, the eitc, and other policy changes on welfare use, work, and income among female-headed families. *Review of Economics and Statistics*, 85(2):394–408.

Halpern-Meekin, S., Edin, K., Tach, L., and Sykes, J. (2015). *It's not like I'm poor: How working families make ends meet in a post-welfare world*. Univ of California Press.

Hausman, J. A. (1982). Labor supply. Technical report, National Bureau of Economic Research.

Hausman, J. A. (1985). Taxes and labor supply. In *Handbook of public economics*, volume 1, pages 213–263. Elsevier.

Heckman, J. J. and MaCurdy, T. E. (1980). A life cycle model of female labour supply. *The Review of Economic Studies*, 47(1):47–74.

Hotz, V. J. (2003). The Earned Income Tax Credit. In *Means-tested transfer programs in the United States*, pages 141–198. University of Chicago press.

Hotz, V. J. and Scholz, J. K. (2006). Examining the effect of the Earned Income Tax Credit on the labor market participation of families on welfare. Technical report, National Bureau of Economic Research.

Jones, D. (2010). Information, preferences, and public benefit participation: Experimental evidence from the advance EITC and 401(k) savings. *American Economic Journal: Applied Economics*, 2(2):147–63.

Keane, M. and Moffitt, R. (1998). A structural model of multiple welfare program participation and labor supply. *International economic review*, pages 553–589.

Keane, M. and Rogerson, R. (2012). Micro and macro labor supply elasticities: A reassessment of conventional wisdom. *Journal of Economic Literature*, 50(2):464–76.

Keane, M. P. (2011). Labor supply and taxes: A survey. *Journal of Economic Literature*, 49(4):961–1075.

Kimball, M. S. and Shapiro, M. D. (2008). Labor supply: Are the income and substitution effects both large or both small? Technical report, National Bureau of Economic Research.

Kleven, H. (2019). The EITC and the extensive margin: A reappraisal. Technical report, National Bureau of Economic Research.

Kleven, H. J. and Waseem, M. (2013). Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from pakistan. *The Quarterly Journal of Economics*, 128(2):669–723.

Laitner, J. and Silverman, D. (2005). Estimating life-cycle parameters from consumption behavior at retirement. Technical report, National Bureau of Economic Research.

Liebman, J. B. and Zeckhauser, R. J. (2004). Schmeduling.

Looney, A. and Singhal, M. (2006). The effect of anticipated tax changes on intertemporal labor supply and the realization of taxable income. Technical report, National Bureau of Economic Research.

Maag, E. (2019.). Using the EITC to help fight an economic slowdown. https://www.taxpolicycenter.org/taxvox/using-eitc-help-fight-economic-slowdown. Accessed: 2020-11-01.

Maag, E., Peters, H. E., Hannagan, A., Lou, C., and Siwicki, J. (2017). Income volatility: New research results with implications for income tax filing and liabilities. *The Urban Institute*.

MaCurdy, T., Green, D., and Paarsch, H. (1990). Assessing empirical approaches for analyzing taxes and labor supply. *Journal of Human resources*, pages 415–490.

MaCurdy, T. E. (1981). An empirical model of labor supply in a life-cycle setting. *Journal of Political Economy*, 89(6):1059–1085.

Manoli, D. and Weber, A. (2016). Nonparametric evidence on the effects of financial incentives on retirement decisions. *American Economic Journal: Economic Policy*, 8(4):160–82.

Martinez, I. Z., Saez, E., and Siegenthaler, M. (forthcoming). Intertemporal labor supply substitution?: Evidence from the Swiss income tax holidays. Technical report.

Meyer, B. D. and Rosenbaum, D. T. (2001). Welfare, the Earned Income Tax Credit, and the labor supply of single mothers. *The Quarterly Journal of Economics*, 116(3):1063–1114.

Miller, B. M., Mumford, K., et al. (2015). The salience of complex tax changes: evidence from the child and dependent care credit expansion. *National Tax Journal*, 68(3):477–510.

Moffitt, R. (1990). The econometrics of kinked budget constraints. *Journal of Economic Perspectives*, 4(2):119–139.

Nichols, A. and Rothstein, J. (2015). The Earned Income Tax Credit. In *Economics of Means-Tested Transfer Programs in the United States, Volume 1*, pages 137–218. University of Chicago Press.

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica: Journal of the Econometric Society*, pages 1417–1426.

Pencavel, J. (1986). Labor supply of men: A survey. *Handbook of Labor Economics*, 1:3–102.

Pistaferri, L. (2003). Anticipated and unanticipated wage changes, wage risk, and intertemporal labor supply. *Journal of Labor Economics*, 21(3):729–754.

Powell, D. (2015). Do payroll taxes in the United States create bunching at kink points? *Michigan Retirement Research Center Research Paper*, (2015-327).

Powell, D. (2020). Does labor supply respond to transitory income? Evidence from the economic stimulus payments of 2008. *Journal of Labor Economics*, 38(1):1–38.

Rees-Jones, A. and Taubinsky, D. (2019). Measuring "Schmeduling". *The Review of Economic Studies*. rdz045.

Reichling, F. and Whalen, C. (2012). Review of estimates of the frisch elasticity of labor supply.

Rothstein, J. (2010). Is the EITC as good as an NIT? conditional cash transfers and tax incidence. *American Economic Journal: Economic Policy*, 2(1):177–208.

Saez, E. (1999). Do taxpayers bunch at kink points? Technical report, National Bureau of Economic Research.

Saez, E. (2010). Do taxpayers bunch at kink points? *American Economic Journal: Economic Policy*, 2(3):180–212.

Schanzenbach, D. W. and Strain, M. R. (2020). Employment effects of the earned income tax credit: Taking the long view. In *Tax Policy and the Economy, Volume 35*. University of Chicago Press.

Schneider, D. and Harknett, K. (2019). Consequences of routine work-schedule instability for worker health and well-being. *American Sociological Review*, 84(1):82–114.

Smeeding, T. M., Phillips, K. R., and O'Connor, M. (2000). The eitc: Expectation, knowledge, use, and economic and social mobility. *National Tax Journal*, pages 1187–1209.

Stafford, T. M. (2015). What do fishermen tell us that taxi drivers do not? an empirical investigation of labor supply. *Journal of Labor Economics*, 33(3):683–710.

Wilson, R. (2020). The EITC and employment transitions: Labor force attachment and annual exit. *National Tax Journal*, 73(1):11–46.

Yang, T.-T. (2018). Family labor supply and the timing of cash transfers: Evidence from the Earned Income Tax Credit. *Journal of Human Resources*, 53(2):445–473.

8 Tables and Figures

Table 1: Demographic characteristics for household head in SIPP sample

	Full SIPP Sample	Wages>0	Restricted Sample
	mean	mean	mean
Age	45.9	41.5	39.6
Female	0.52	0.49	0.41
Non-white	0.19	0.19	0.21
College grad	0.33	0.36	0.34
Married	0.45	0.49	0.49
Have kids	0.63	0.72	0.52
Annual wages (2017 $)	28,432	36,506	43,827
Observations	274,899	214,287	134,594

Notes. Table 1 summarizes the average value for select characteristics of the primary filer in each SNAP household/tax unit pooled over the three tax years in our sample, 2015-2017. Wages are reported in 2017 dollars.

Table 2: Demographic characteristics for primary taxpayer in SNAP sample

	SNAP + Tax Filer	Wages>0	Restricted Sample
	mean	mean	mean
Age	35.5	34.5	36.1
Female	0.57	.58	0.67
Non-white	0.78	.79	0.79
Married	0.19	.20	0.14
Have children	0.53	0.59	0.79
Annual wages ($)	11,700	22,358	21,223
Observations	2,102,483	1,227,227	106,636

Notes. Table 2 summarizes the average value for select characteristics of the primary filer in each SNAP household/tax unit pooled over the three tax years in my sample, 2015-2017. Wages are reported in 2017 dollars.

Table 3: Share of households whose predicted income, average and marginal tax rates, and EITC amounts as of each quarter differ from their year-end actual values by more than identified ranges

	$\lvert\hat{z} - z\rvert > \$5k$	$\lvert\hat{\tau} - \tau\rvert > 10pp$	$\lvert\widehat{\text{EITC}} - \text{EITC}\rvert > \$1k$	$\lvert\widehat{\text{MTR}} - \text{MTR}\rvert > 10pp$
			SIPP	
March	.28	.06	.13	.15
June	.18	.03	.03	.11
September	.06	.01	.01	.06
$\bar{\sigma}$	\$2,321	1.4	\$93	2.5
			SNAP	
March	.36	.25	.13	.32
June	.21	.15	.08	.22
September	.05	.06	.03	.12
$\bar{\sigma}$	\$2,651	4.1	\$240	7.1

Notes. Table 3 summarizes the share of households in each sample, as of the end of each quarter, whose: (1) predicted annual income z is more than \$10,000 from their year-end income, (2) predicted annual tax rate τ is more than 10 percentage points from their actual year-end average tax rate, (3) predicted total EITC refund is more than \$1,000 from their year-end amount, and (4) predicted marginal tax rate is more than 10 pp from their year-end marginal tax rate. I limit to households who have positive earnings and no more than \$75,000 in annual income through each listed quarter. The final row reports the mean standard deviation for each predicted value over all tax units.

Table 4: Share of households whose predicted income, average and marginal tax rates, and EITC amounts as of each quarter differ from their year-end actual values by more than identified ranges

	$\lvert\hat{z} - z\rvert > \$5k$	$\lvert\hat{\tau} - \tau\rvert > 10pp$	$\lvert\widehat{\text{EITC}} - \text{EITC}\rvert > \$1k$	$\lvert\widehat{\text{MTR}} - \text{MTR}\rvert > 10pp$
		SIPP		
Year over year	.46	.13	.08	.23
Min to max	.59	.19	.11	.31
$\bar{\sigma}$	\$6,298	5.70	\$210	5.35
		SNAP		
Year over year	.46	.41	.27	.44
Min to max	.53	.46	.31	.50
$\bar{\sigma}$	\$4,818	9.57	\$620	12.6

Notes. Table 4 summarizes the share of households whose earnings, average and marginal tax rates, and EITC amounts in one year are particular values different than in other years. In Row 1, I report the share of households whose (1) maximum annual income z is more than \$10,000 from their minimum year-end income, (2) maximum annual tax rate τ is more than 5 percentage points from their minimum year-end average tax rate, (3) maximum total EITC refund is more than \$1,000 from their lowest EITC amount, and (4) maximum marginal tax rate is more than 10 pp from their lowest marginal tax rate. In Row 2, I report the same shares but compare differences between subsequent years, meaning I count households that appear in multiple tax years more than once in the denominator. I limit to households who have at least positive earnings and no more than \$75,000 in annual income through each listed quarter. The final row reports the average standard deviation of each variable among households who appear for at least two tax years.

Table 5: Log earnings response in Q4 to predicted net of tax earnings, given Q1-Q3 earnings, SIPP sample

	(1)	(2)	(3)	(4)
	$\ln y$	$\ln y$	$\ln y$	$\ln y$
ω				
	-1.959***	-2.008***	0.023	-0.074
	(0.037)	(0.040)	(0.033)	(0.128)
z				
			0.950***	0.367***
			(0.008)	(0.044)
Observations	18309	18289	18289	10614
Households	12321	12312	12312	4649
Demographics		X	X	X
# of deps × marital status		X	X	X
State × year		X	X	X
Household FE				X
R^2	0.24	0.38	0.72	0.88

Notes. Table 5 summarizes estimates of Equation 2 in the SIPP sample. I limit to sequences that coincide with the actual tax year. Standard errors are clustered at the household-level, and I use household-level weights. The final row reports the p-value from the F-test that estimates of π_q are equal to each other.

Table 6: Log earnings response in Q4 to predicted net of tax earnings, given Q1-Q3 earnings, SNAP sample

	(1)	(2)	(3)	(4)
	$\ln y$	$\ln y$	$\ln y$	$\ln y$
ω				
	-1.368***	-1.491***	0.019	0.073**
	(0.007)	(0.007)	(0.012)	(0.022)
z				
			0.821***	0.432***
			(0.006)	(0.013)
Observations	175953	175951	175951	119804
Households	106638	106637	106637	50490
Demographics		X	X	X
# of deps × marital status		X	X	X
State × year		X	X	X
Household FE				X
R^2	0.22	0.35	0.45	0.76

Notes. Table 6 summarizes estimates of Equation 2 in the SNAP sample. I limit to sequences that coincide with the actual tax year. Standard errors are clustered at the household-level.

Table 7: Log earnings response in each quarter to predicted net of tax earnings, simulated using three previous quarters' earnings, SIPP sample

	(1)	(2)	(3)	(4)
	$\ln y$	$\ln y$	$\ln y$	$\ln y$
ω				
	-1.867***	-1.908***	0.027	-0.070
	(0.029)	(0.030)	(0.017)	(0.047)
$\omega\times$				
Q2				
	0.019***	0.019***	0.009	0.035***
	(0.005)	(0.005)	(0.005)	(0.004)
Q3				
	0.028***	0.027***	0.006	0.057***
	(0.006)	(0.006)	(0.007)	(0.006)
Q4				
	0.043***	0.043***	0.006	0.081***
	(0.006)	(0.006)	(0.007)	(0.006)
z				
			0.883***	0.073***
			(0.005)	(0.018)
Observations	73239	73239	73239	73239
Households	12321	12321	12321	12321
Demographics		X	X	X
# of deps $\times$ marital status		X	X	X
State $\times$ year		X	X	X
Household FE				X
R^2	0.24	0.38	0.70	0.81
P-value from F-test	0.00	0.00	0.79	0.00

Notes. Table 7 summarizes estimates of Equation 3 in the SIPP sample. Standard errors are clustered at the household-level, and I use household-level weights. The final row reports the p-value from the F-test that estimates of π_q are equal to each other.

Table 8: Log earnings response in each quarter to predicted net of tax earnings rate in that quarter, simulated using three previous quarters' earnings, SNAP sample

	(1)	(2)	(3)	(4)
	$\ln y$	$\ln y$	$\ln y$	$\ln y$
ω				
	-1.165*	-1.368	-0.001	-0.051***
	(0.004)	(0.004)	(0.007)	(0.009)
$\omega\times$				
Q2				
	0.039	0.034***	0.015***	0.105***
	(0.002)	(0.002)	(0.002)	(0.002)
Q3				
	0.034	0.024***	-0.008***	0.150***
	(0.002)	(0.002)	(0.002)	(0.002)
Q4				
	0.038	0.027***	-0.015***	0.180***
	(0.002)	(0.002)	(0.003)	(0.002)
z				
			0.667***	0.040***
			(.004)	(.006)
Observations	703,792	703,792	703,792	703,791
Households	106,636	106,636	106,636	106,635
Demographics		X	X	X
# of deps × marital status		X	X	X
State × year		X	X	X
Household FE				X
R^2	0.22	0.37	0.46	0.65
P-value from F-test	0.00	0.00	0.00	0.00

Notes. Table 8 summarizes results from estimations of Equation 3 in the SNAP sample. Standard errors are clustered at the household-level. The final row reports the p-value from the F-test that estimates of π_q are equal to each other.

Table 9: Log earnings response in each quarter to predicted net of tax earnings rate in final quarter, SIPP sample

	(1)	(2)	(3)	(4)
	$\ln y$	$\ln y$	$\ln y$	$\ln y$
$\hat{\omega}$				
	-4.032***	-6.236***	-1.525***	-0.074
	(0.041)	(0.045)	(0.057)	(0.041)
$\hat{\omega}$				
Q2				
	0.003	0.017**	0.019***	0.024***
	(0.005)	(0.006)	(0.004)	(0.003)
Q3				
	0.016**	0.032***	0.014**	0.037***
	(0.006)	(0.007)	(0.005)	(0.004)
Q4				
	0.003	0.057***	0.024***	0.055***
	(0.006)	(0.007)	(0.005)	(0.004)
z				
			0.660***	0.126***
			(0.008)	(0.010)
Observations	102191	102168	102052	99458
Households	17840	17826	17826	15238
Demographics		X	X	X
# of deps $\times$ marital status		X	X	X
State $\times$ year		X	X	X
Household FE				X
R^2	0.31	0.55	0.70	0.82
P-value from F-test	0.00	0.00	0.03	0.00

Notes. Table 9 summarizes estimates of Equation 7 in the SIPP sample. Standard errors are clustered at the household-level, and I use household-level weights. The final row reports the p-value from the F-test that estimates of π_q are equal to each other.

Table 10: Log earnings response in following quarter to predicted year-end net of tax earnings rate, SNAP sample

	(1)	(2)	(3)	(4)
	$\ln y$	$\ln y$	$\ln y$	$\ln y$
$\hat{\omega}$				
	-2.312	-2.941	-1.639***	-0.021*
	(0.007)	(0.008)	(0.009)	(0.010)
$\hat{\omega}\times$				
Q2				
	0.138	0.152	0.101***	0.122***
	(0.002)	(0.002)	(0.002)	(0.002)
Q3				
	0.198	0.220	0.109***	0.161***
	(0.002)	(0.002)	(0.002)	(0.002)
Q4				
	0.225	0.251	0.112***	0.179***
	(0.002)	(0.002)	(0.002)	(0.002)
z				
			0.352***	0.072***
			(.003)	(.002)
Observations	847,139	847,102	703,792	833,627
Households	135,871	135,856	135,856	125,926
Demographics		X	X	X
# of deps × marital status		X	X	X
State × year		X	X	X
Household FE				X
R^2	0.18	0.35	0.41	0.61
P-value from F-test	0.00	0.00	0.00	0.00

Notes. Table 10 summarizes results from estimations of Equation 3 in the SNAP sample. Standard errors are clustered at the household-level. The final row reports the p-value from the F-test that estimates of π_q are equal to each other.

Table 11: Log earnings response in each quarter to predicted federal and state marginal income tax rate, SIPP sample

	(1)	(2)	(3)	(4)
	$\ln y$	$\ln y$	$\ln y$	$\ln y$
MTR				
	0.016***	0.016***	-0.001***	-0.001
	(0.000)	(0.000)	(0.000)	(0.000)
MTR ×				
Q2				
	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Q3				
	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Q4				
	0.002***	0.002***	0.002***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
z				
			0.885***	0.089***
			(0.005)	(0.017)
Observations	73239	73239	73239	73239
Households	12321	12321	12321	12321
Demographics		X	X	X
# of deps × marital status		X	X	X
State × year		X	X	X
Household FE				X
R^2	0.19	0.33	0.70	0.81
P-value from F-test	0.00	0.00	0.00	0.00

Notes. Table 11 summarizes estimates from a version of Equation 3 in which I replace household's ω_q with the households' predicted marginal income tax rate in that quarter. Standard errors are clustered at the household-level, and I use household-level sampling weights. The final row reports the p-value from the F-test that estimates of π_q are equal to each other.

Table 12: Log earnings response in each quarter to predicted federal and state marginal income tax rate, SNAP sample

	(1)	(2)	(3)	(4)
	$\ln y$	$\ln y$	$\ln y$	$\ln y$
MTR				
	0.867***	0.891***	-0.276***	-0.004
	(0.006)	(0.006)	(0.007)	(0.008)
MTR				
Q2				
	0.273***	0.274***	0.280***	0.012*
	(0.007)	(0.007)	(0.007)	(0.005)
Q3				
	0.450***	0.444***	0.407***	-0.016*
	(0.007)	(0.008)	(0.007)	(0.007)
Q4				
	0.521***	0.505***	0.425***	-0.020***
	(0.008)	(0.007)	(0.007)	(0.007)
z				
			0.689***	0.098**
			(0.003)	(0.007)
Observations	703815	703815	703815	703815
Households	106640	106640	106640	106639
Demographics		X	X	X
# of deps × marital status		X	X	X
State × year		X	X	X
Household FE				X
R^2	0.21	0.32	0.43	0.63
P-value from F-test	0.00	0.00	0.00	0.00

Notes. Table 12 summarizes estimates from a version of Equation 3 in which I replace household's ω_q with the households' predicted marginal income tax rate in that quarter. Standard errors are clustered at the household-level. The final row reports the p-value from the F-test that estimates of π_q are equal to each other.

Table 13: Log hours response in each quarter to predicted net of tax earnings, simulated using three previous quarters' earnings, SIPP sample

	(1)	(2)	(3)	(4)
	$\ln h$	$\ln h$	$\ln h$	$\ln h$
ω				
	-0.626***	-0.597***	-0.078**	-0.001
	(0.024)	(0.024)	(0.027)	(0.030)
$\omega\times$				
Q2				
	0.010*	0.010*	0.008	0.011**
	(0.004)	(0.004)	(0.004)	(0.004)
Q3				
	0.011	0.011*	0.006	0.015**
	(0.006)	(0.006)	(0.006)	(0.005)
Q4				
	0.005	0.006	-0.004	0.013*
	(0.005)	(0.005)	(0.005)	(0.005)
z				
			0.238***	0.046**
			(0.008)	(0.014)
Observations	70876	70876	70876	70817
Households	12224	12224	12224	12165
Demographics		X	X	X
# of deps $\times$ marital status		X	X	X
State $\times$ year		X	X	X
Household FE				X
R^2	0.05	0.25	0.30	0.70
P-value from F-test	0.39	0.52	0.06	0.58

Notes. Table 13 summarizes estimates of Equation 3 in the SIPP sample, in which I replace the outcome variable with the log of the average hours worked per week, summed over the respective quarter. Standard errors are clustered at the household-level, and I use household-level weights.The final row reports the p-value from the F-test that estimates of π_q are equal to each other.

Table 14: Earnings response in each quarter to predicted net of tax earnings rate, simulated using three previous quarters' earnings, by household type, SIPP sample

	Marital Status		Presence of Children	
	(1)	(2)	(3)	(4)
	Single	Married	No Kid	Have Kids
ω				
	-0.077	-0.044	0.101	-0.052
	(0.068)	(0.061)	(0.159)	(0.048)
$\omega\times$				
Q2				
	0.020**	0.054***	0.026***	0.044***
	(0.006)	(0.006)	(0.006)	(0.006)
Q3				
	0.035***	0.086***	0.049***	0.067***
	(0.008)	(0.008)	(0.007)	(0.009)
Q4				
	0.062***	0.105***	0.073***	0.091***
	(0.008)	(0.009)	(0.008)	(0.009)
z				
	0.046	0.095***	0.049	0.080***
	(0.028)	(0.021)	(0.030)	(0.022)
Observations	37539	35084	35751	37488
Households	6287	5925	6001	6420
Demographics	X	X	X	X
# of deps $\times$ marital status	X	X	X	X
State $\times$ year	X	X	X	X
Household FE	X	X	X	X
R^2	0.84	0.76	0.82	0.81
P-value from F-test	0.00	0.00	0.00	0.00

Notes. Table 14 summarizes results from estimations of Equation 3 in the SIPP sample. Standard errors are clustered at the household-level, and I use household-level weights. The final row reports the p-value from the F-test that estimates of π_q are equal to each other.

Table 15: Log earnings response in each quarter to predicted net of tax earnings rate, simulated using three previous quarters' earnings, by household type, SNAP sample

	Marital Status		Presence of Children	
	(1) Single	(2) Married	(3) No Kid	(4) Have Kids
ω				
	-0.058	-0.031	-0.194	-0.039
	(0.064)	(0.052)	(0.050)	(0.042)
$\omega\times$				
Q2				
	0.102***	0.125***	0.157***	0.097***
	(0.006)	(0.004)	(0.006)	(0.002)
Q3				
	0.145***	0.180***	0.205***	0.143***
	(0.007)	(0.005)	(0.007)	(0.002)
Q4				
	0.178***	0.192***	0.229***	0.174***
	(0.007)	(0.005)	(0.007)	(0.002)
z				
	0.041***	0.046***	-0.029*	0.048***
	(0.005)	(0.016)	(0.026)	(0.007)
Observations	602837	100692	149859	553956
Households	92599	13906	26984	82103
Demographics	X	X	X	X
# of deps $\times$ marital status	X	X	X	X
State $\times$ year	X	X	X	X
Household FE	X	X	X	X
R^2	0.62	0.64	0.62	0.62
P-val from F-test	0.00	0.00	0.00	0.00

Notes. Table 15 summarizes results from estimations of Equation 3 among married vs. single households and households with and without children in the SNAP sample. Standard errors are clustered at the household-level. The final row reports the p-value from the F-test that estimates of π_q are equal to each other.

Table 16: Earnings response in each quarter to predicted net of tax earnings rate, simulated using three previous quarters' earnings, by various subgroups, SIPP sample

	(1) Hourly workers	(2) Self-employed	(3) SE income	(4) Exclude retail	(5) Seam check
ω					
	-0.067	0.021	1.162	-0.063	0.039
	(0.044)	(0.270)	(2.060)	(0.048)	(0.080)
$\omega \times$					
Q2					
	0.042***	0.132**	0.053	0.035***	0.038***
	(0.005)	(0.048)	(0.398)	(0.005)	(0.009)
Q3					
	0.067***	0.173**	-0.123	0.059***	0.054***
	(0.007)	(0.064)	(0.492)	(0.006)	(0.012)
Q4					
	0.086***	0.261***	0.036	0.084***	0.078***
	(0.007)	(0.061)	(0.508)	(0.006)	(0.012)
z					
	0.039*	-0.366**	-0.872	0.069***	0.103**
	(0.020)	(0.130)	(0.894)	(0.018)	(0.031)
Observations	54811	832	832	71571	22116
Households	9485	178	178	12063	3375
Demographics	X	X	X	X	X
# of deps $\times$ marital status	X	X	X	X	X
State $\times$ year	X	X	X	X	X
Household FE	X	X	X	X	X
R^2	0.81	0.82	0.42	0.81	0.79
P-val from F-test	0.00	0.05	0.88	0.00	0.00

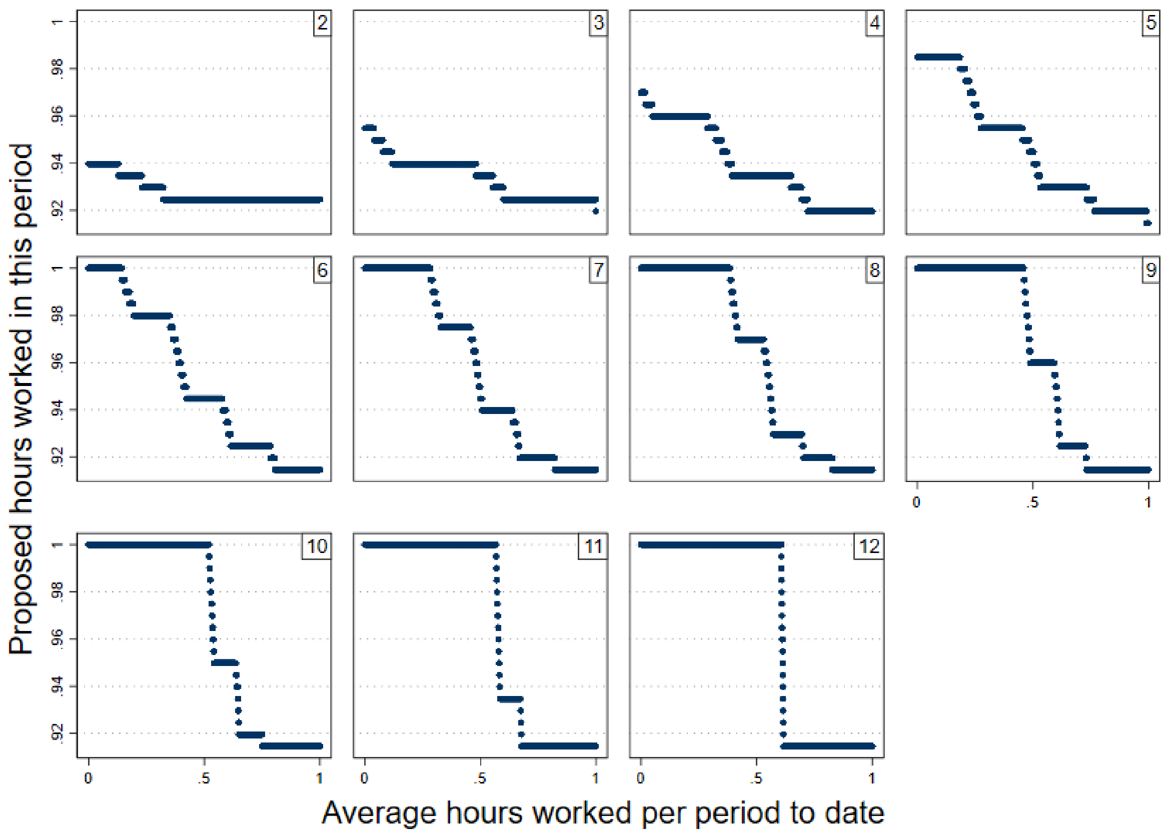
Notes. Table 16 summarizes estimates of Equation 3 for particular subsets of the SIPP sample. Column 1 reports estimated responses among households in which either the head or spouse report being an hourly worker. Column 2 reports estimates among households in which either the head or spouse report earnings any income from self employment in the tax year. Column 3 reports estimates among the same subset of self-employed workers, but I replace the outcome variable with the log of self-employment earnings, as opposed to all earned income. Column 4 reports estimates from the SIPP sample after I drop all households in in which either the head or spouse report working in the retail industry. Column 5 reports estimates from the SIPP sample after I restrict the sample to households whose survey waves align with the tax year. Standard errors are clustered at the household-level, and I use household-level weights. The final row reports the p-value from the F-test that estimates of π_q are equal to each other.

Table 17: Earnings response in each quarter to predicted net of tax earnings rate in that quarter, simulated using three previous quarters' earnings, SIPP and SNAP sample

	SIPP	SNAP
$\tilde{\omega}_{q=1}$		
	1.086*	-.381
	(1.466)	(0.111)
$\tilde{\omega}_{q=2}$		
	0.346	0.118***
	(1.436)	(0.126)
$\tilde{\omega}_{q=3}$		
	2.217	0.521***
	(4.193)	(0.113)
$\tilde{\omega}_{q=4}$		
	-2.161	0.126***
	(5.800)	(0.126)
$\tilde{z}$	.977	.630
	(.312)	(.067)
Observations	5,288	75,300
Households	1,322	18,825
Demographics	X	X
# of deps $\times$ marital status	X	X
State $\times$ year	X	X
Household FE		
R^2	0.22	0.03
P-value from F-test	0.87	0.00
Hansen J-stat	0.00	0.00

Notes. Table 17 summarizes results from Equation 8. In both models, I cluster standard errors at the household-level. In the SIPP, I apply household weights. The final row reports the p-value from the F-test that estimates of π_q are equal to each other.

Figure 3: Agent's predicted hours choice based on average hours worked to date

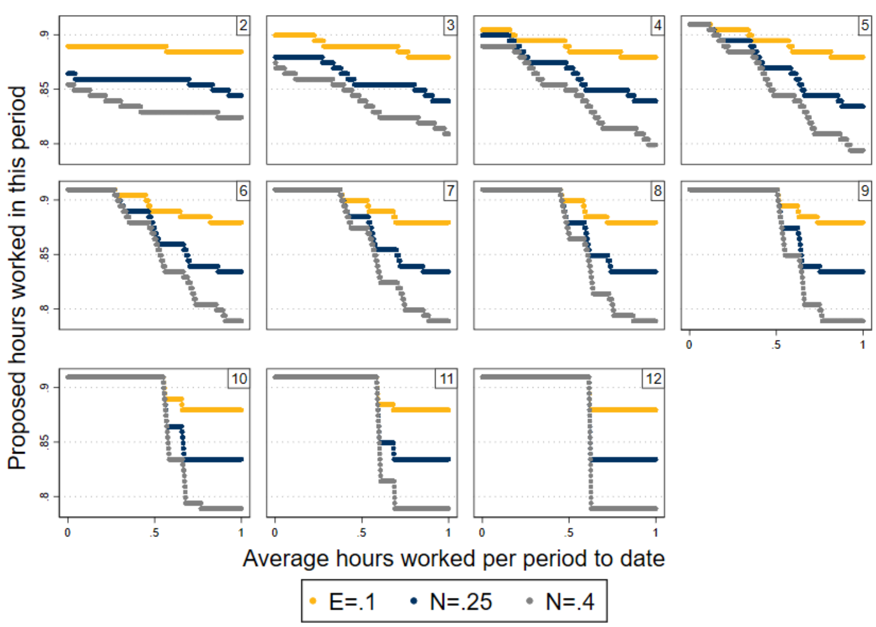


Notes. Figure 3 plots proposed hours in the subsequent period given average hours worked per period to date for a set of representative agents. In Panel A, I plot choices for three workers with different ability parameters n. These workers have the same elasticity $e = .3$ and face the same tax policy ($\tau_0 = 0, \tau_1 = .3$. In Panel B, I plot choices for three worker, all with $n = .9$, but with elasticities of .1, .25 and .4.

Figure 4: Agent's predicted hours choice based on average hours worked to date, by n and e

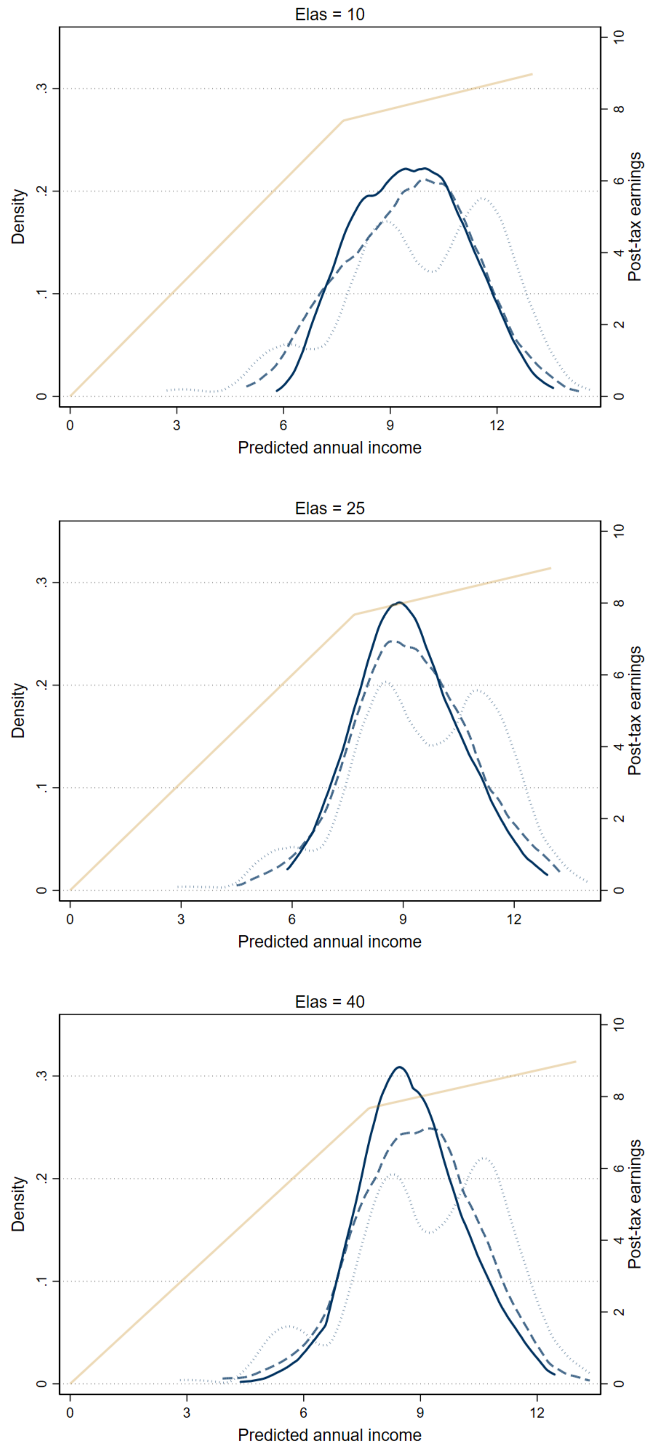
(a) By ability n

(b) By elasticity e



Notes. Figure 4 plots proposed hours in the subsequent period given average hours worked per period to date for a set of representative agents. In Panel A, I plot choices for three workers with different ability parameters n. These workers have the same elasticity $e = .3$ and face the same tax policy ($\tau_0 = 0, \tau_1 = .3$). In Panel B, I plot choices for three worker, all with $n = .9$, but with elasticities of .1, .25 and .4.

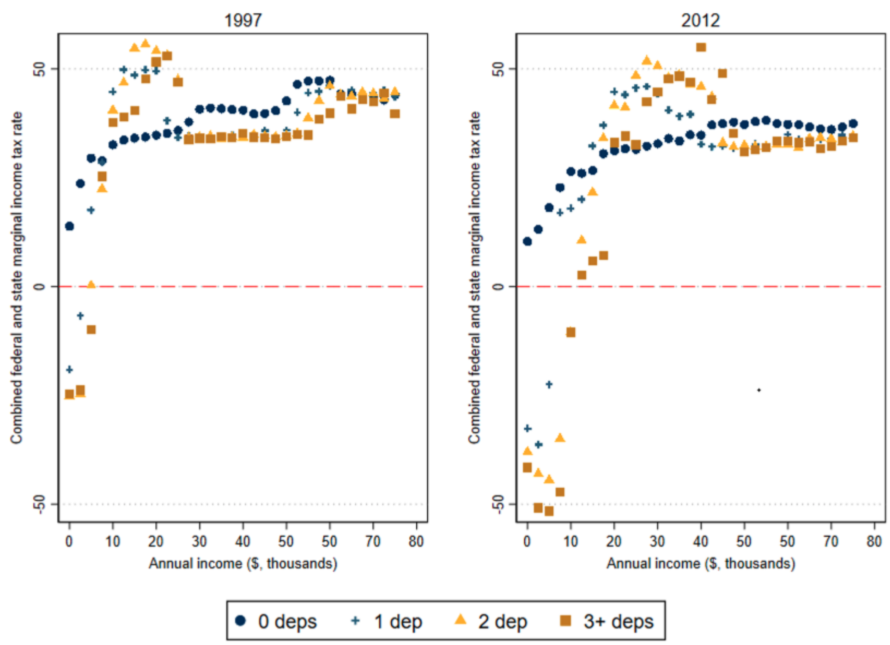
Figure 5: Distribution of proposed annual income as of three periods



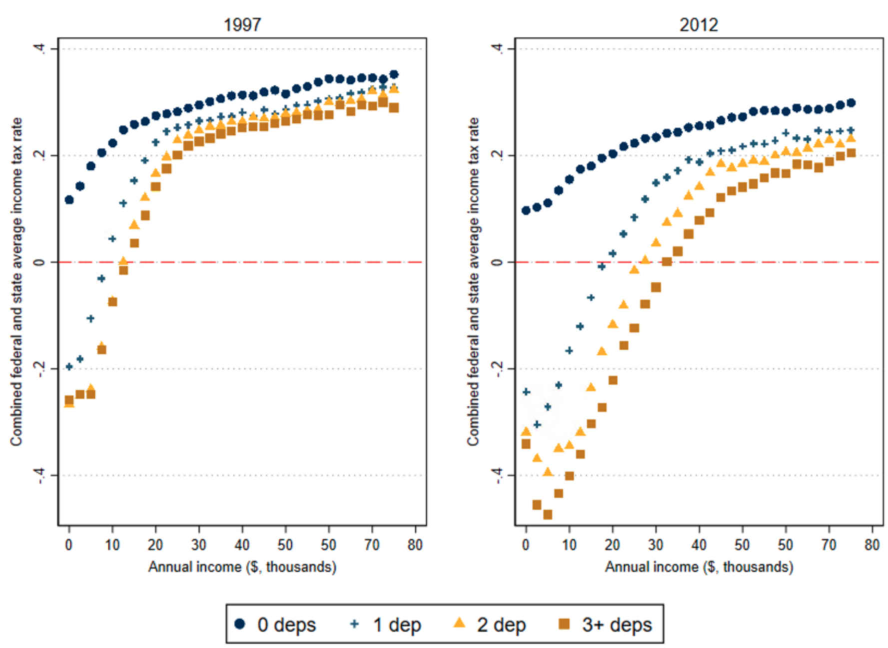
Notes. Figure 5 presents results from a three simulations of the model summarized in **??**. I plot the distributions of predicted annual income as of three periods for 500 agents with various levels of n. For each agent, I simulate responses using an elasticity e of .1, .25 and .4. The other parameter values are: $p = .8$, $\tau_0 = 0$, $\tau_1 = .3$, and $z^* = 7.6$. The light dotted line indicates predicted earnings as of $d = 4$, the dashed line is as of $d = 8$, and the solid line is the last period, $d = 12$.

Figure 6: Combined marginal and average tax rates for married households, 1997 vs 2012, SIPP sample

(a) Marginal income tax rate



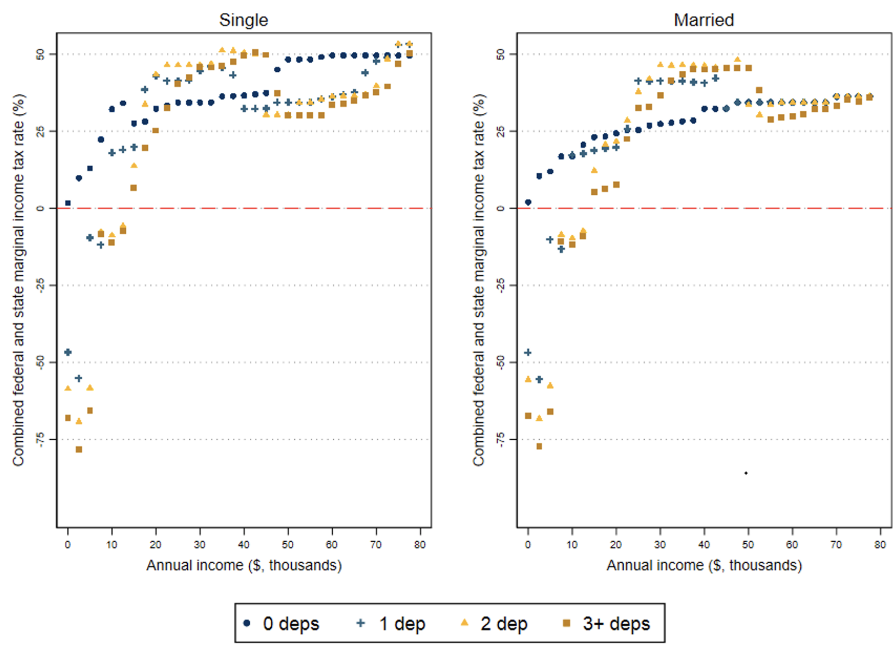
(b) Average income tax rate



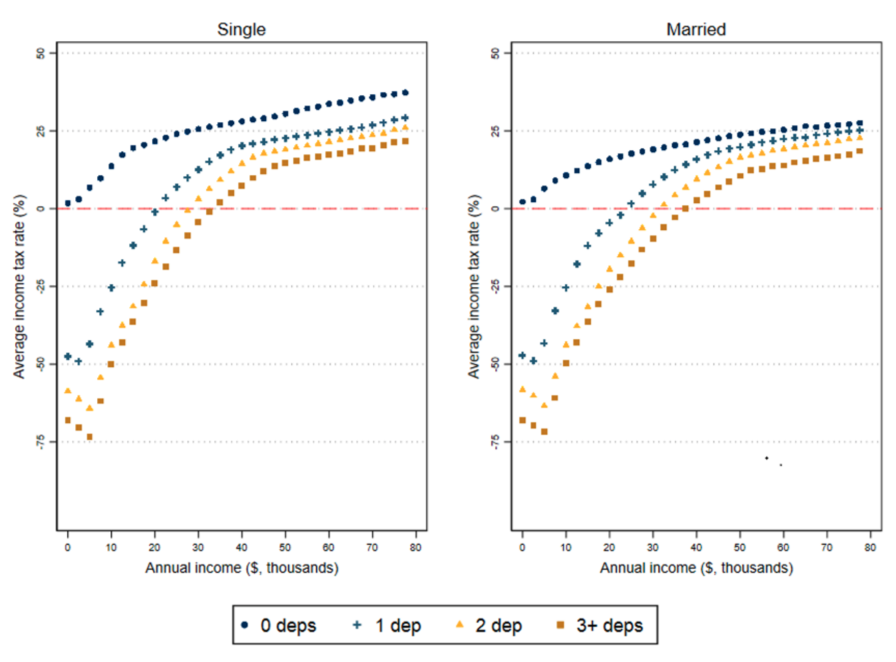
Notes. Figure 6 illustrates how marginal and average tax rates vary for married households with 0, 1, 2 or 3+ dependents in 1997 versus 2012. By year and for each number of dependents, I group households into annual income bins of $2,500. In each bin, I calculate the average and average marginal tax rate faced by households at year's end.

Figure 7: Combined marginal and average tax rates for married households, by number of dependents and annual income, SNAP households, 2017

(a) Marginal income tax rate



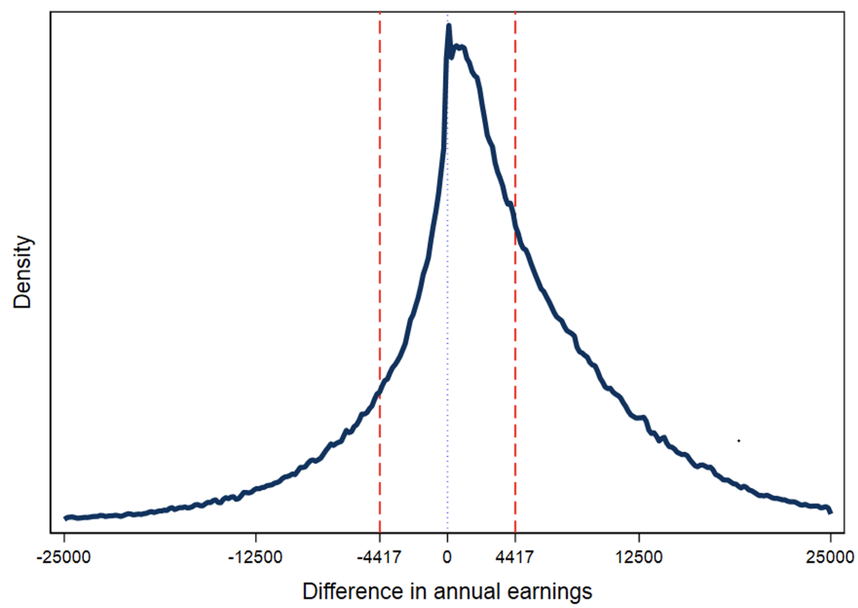
(b) Average income tax rate



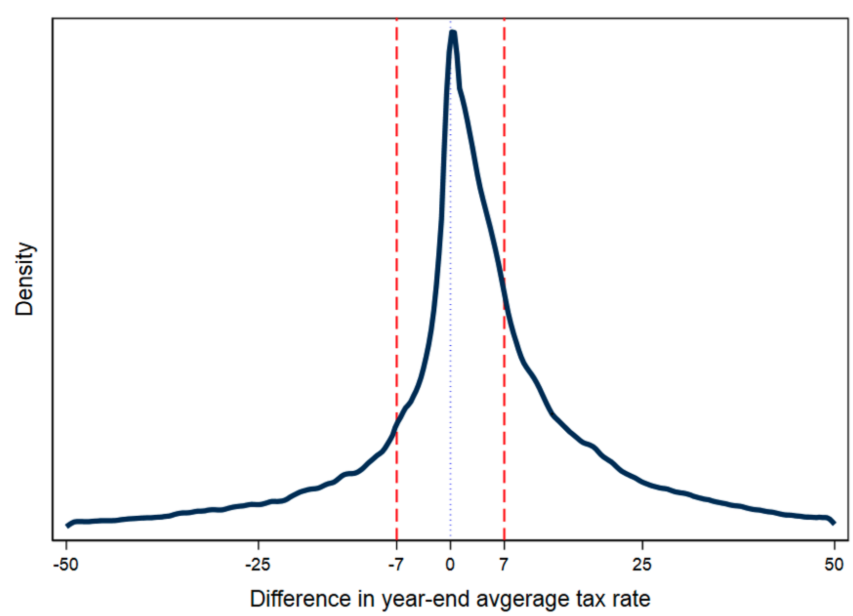
Notes. Figure 7 summarizes how average and marginal tax rates vary by household income and number of dependents for single and married filers enrolled in SNAP in California in 2016. I group tax units into bins of $2,000 in annual income by filing status and number of dependents, and within each bin, identify the mean marginal and average tax rate.

Figure 8: Distribution of cross-year differences in household income, average tax rate, and predicted ω, SNAP sample

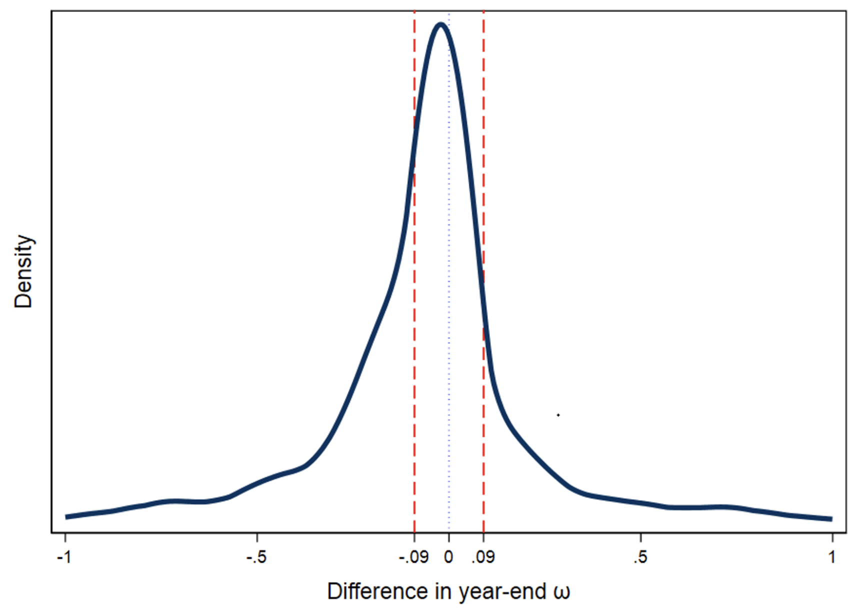
(a) Annual earnings



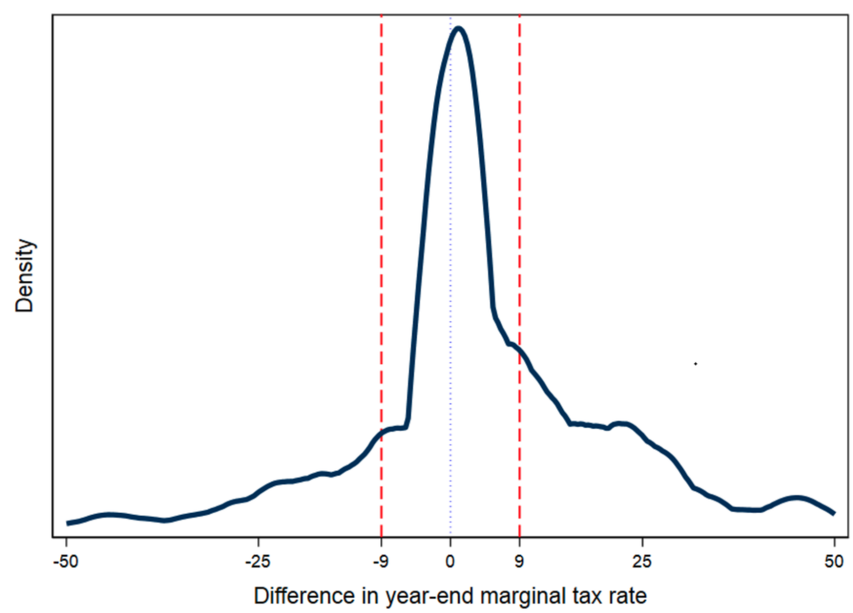
(b) Average tax rate



(c) Predicted final quarter net of tax wage rate

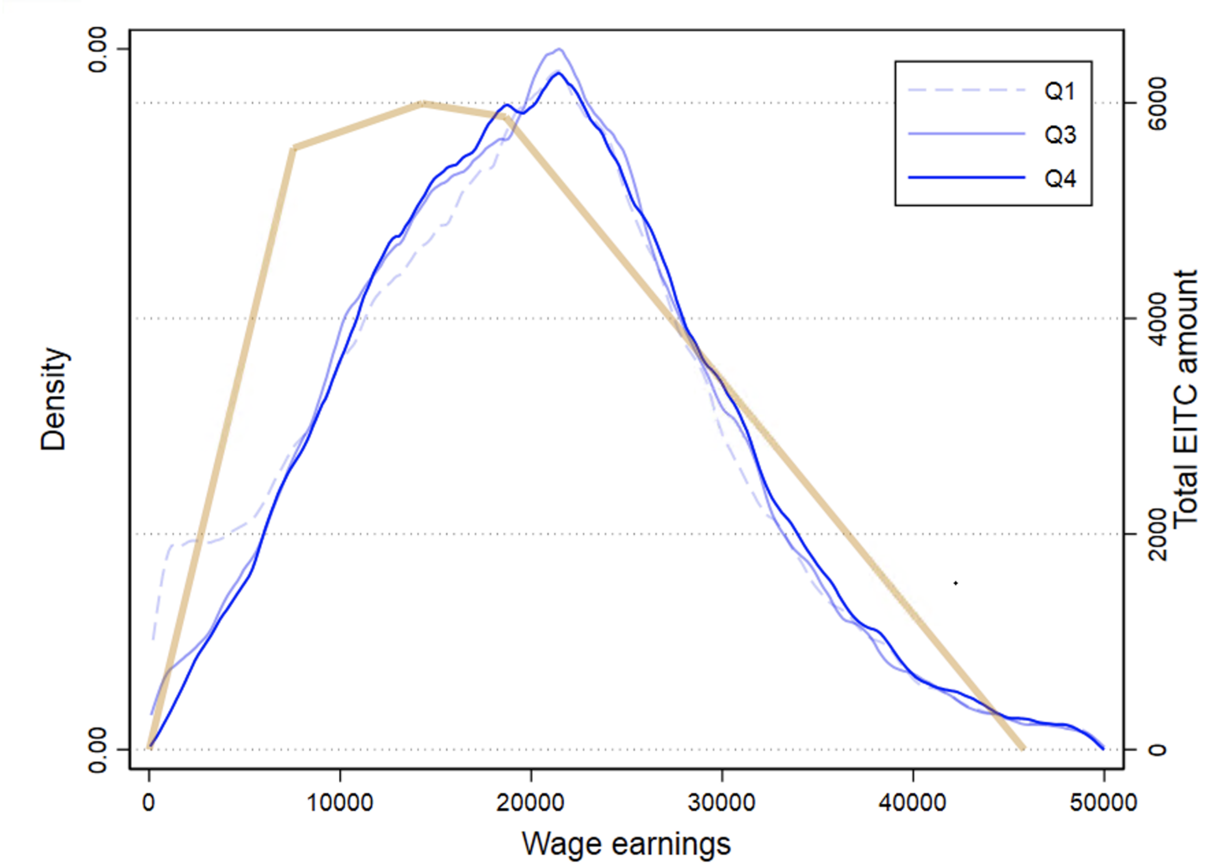


(d) Marginal tax rate



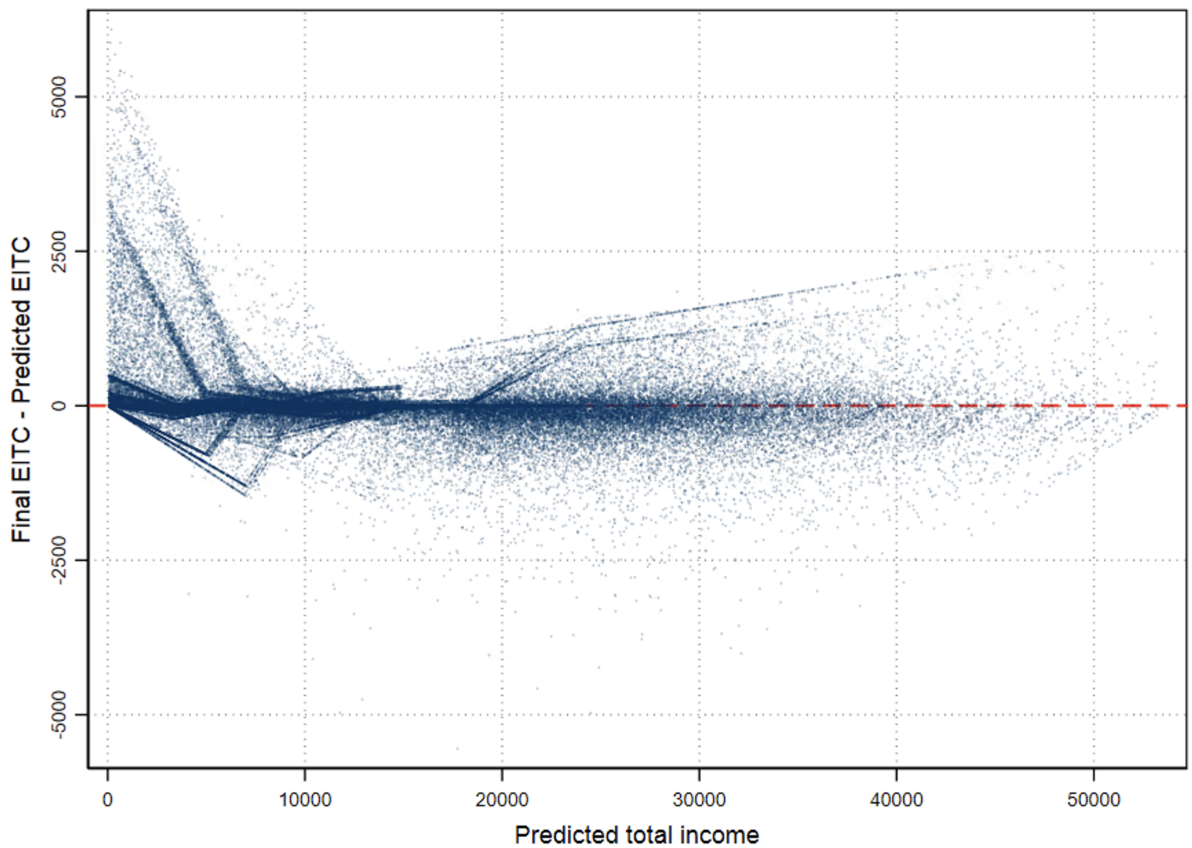
Notes. Figure 8 illustrates the differences in household income, average tax rate and ω across tax years within the same SNAP households. For all households, I subtract the value from the value from the previous tax year. I plot the kernel density of all these differences. For household income, I use a bandwidth of $100 and limit to differences within $25,000. For average tax rate, I use a bandwidth of half a percentage point and limit to differences within 50 percentage points. For predicted net of tax wage rate on fourth quarter earnings, I use a bandwidth of 5 percentage points and limit to differences within 75 percentage points. For the marginal tax rate, I use a bandwidth of 2.5 percentage points and limit to differences within 50 percentage points. The red dotted lines indicate the median absolute value difference, meaning half of households exhibit a difference between those bounds and the other half exhibit a difference outside those bounds.

Figure 9: Distribution of annual income and predicted annual income as of the end of the first quarter, SNAP households with two dependents, 2017



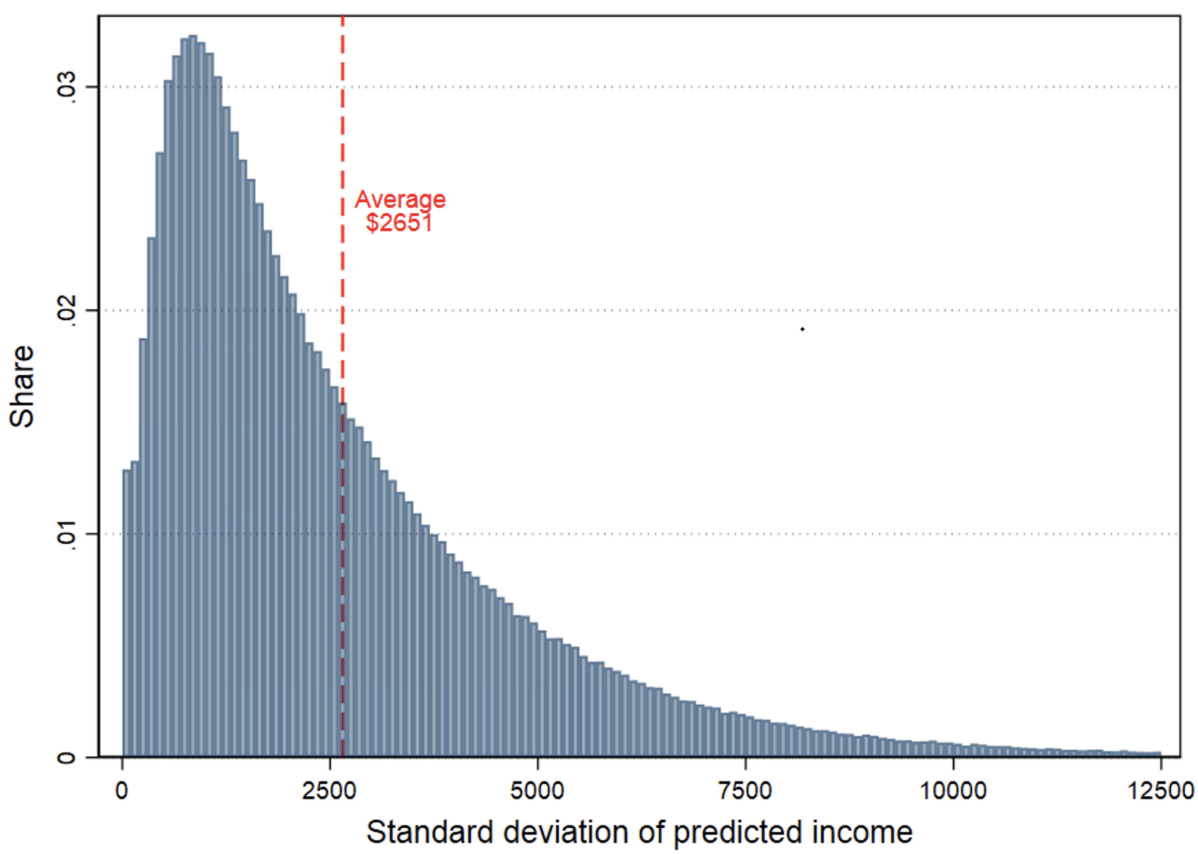
Notes. Figure 9 plots kernel density distributions of predicted annual wage earnings as of the end of the first quarter and third quarter, as well as the distribution of actual wage earnings, for SNAP households with two dependents in 2017. I use a bandwidth of $500 and limit to households with a maximum of $50,000. I overlay these distributions on the combined federal and state EITC schedule for a single filer with two dependents in 2017. From the first quarter to the last, the distribution of earnings shifts such that fewer households are located towards the beginning of the phase-in region and more are clustered at the top of the EITC range.

Figure 10: Predicted total EITC amounts versus final total EITC amounts, SNAP sample



Notes. Figure 10 illustrates how those differences vary by predicted annual income. Each dot represents one SNAP households' predicted and actual EITC amount. I use a five percent sample of SNAP households in tax years 2015 to 2017

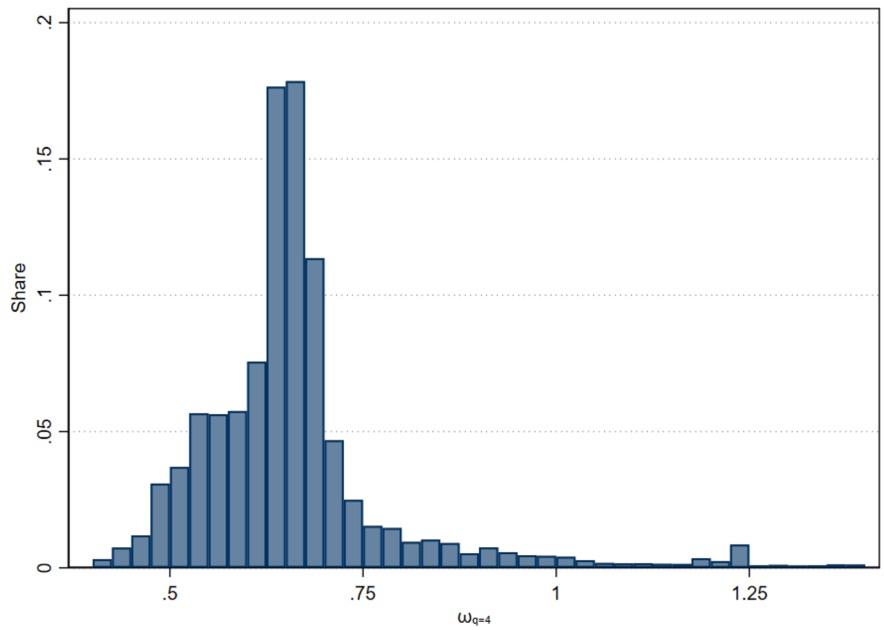
Figure 11: Distribution of standard deviations in predicted earnings, SNAP households, 2015-2017



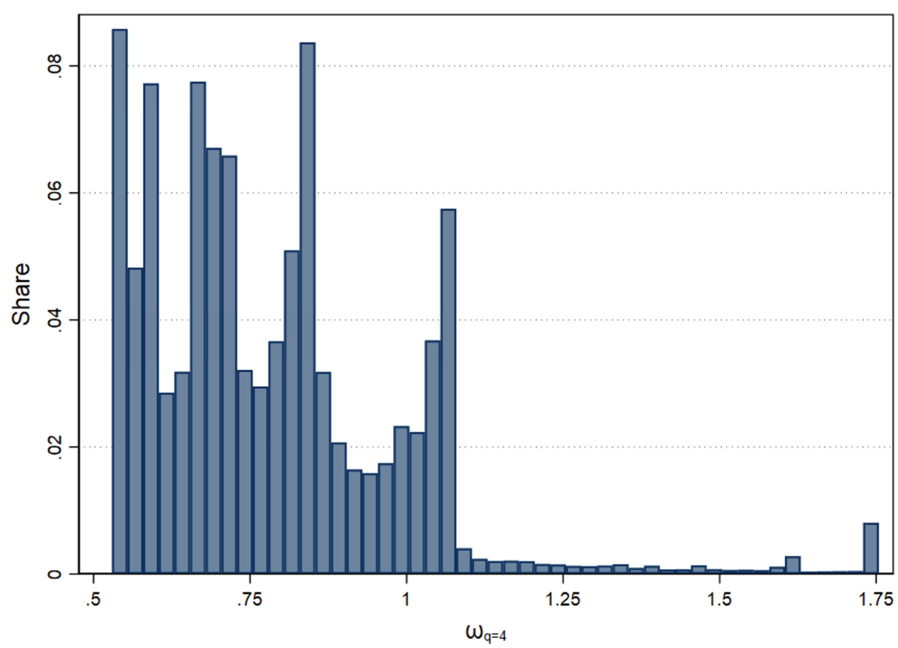
Notes. Figure 11 illustrates the distribution of standard deviations across households in their predicted earnings over four quarters of each tax year. I implement the same restrictions as in the rest of my analysis. I predict annual earnings as of each quarter by extrapolating from to-date earnings (i.e., predicted earnings as of June equal double the income earned in the first two quarters). For each household in each tax year, I calculate the standard deviation over the four predicted values, where the final value is equal to the household's actual earnings.

Figure 12: Distribution of ω in Q4 in the SIPP and SNAP samples

(a) SIPP sample



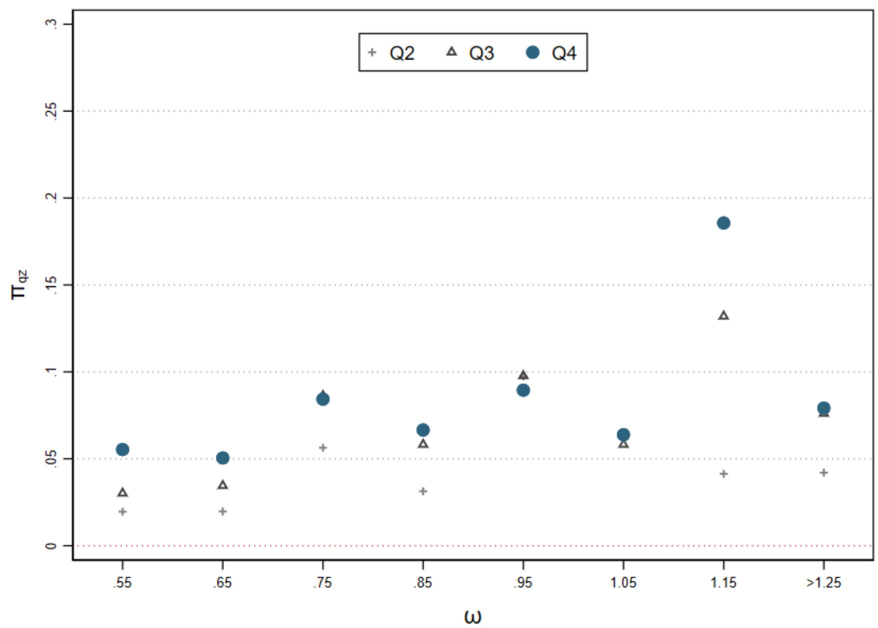
(b) SNAP sample



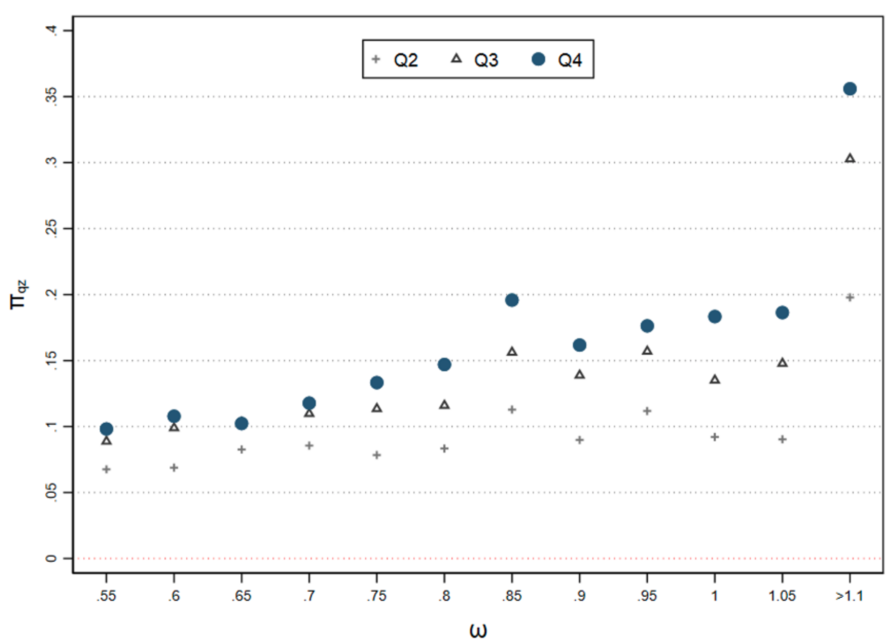
Notes. Figure 12 illustrates the distribution of values of ω, the predicted net of tax wage rate for subsequent quarter's earnings, for the fourth quarter in both the SIPP and SNAP sample.

Figure 13: Log earnings response to levels of predicted net of tax wage rate in subsequent quarter, SNAP households and tax units, 2015-2017

(a) SIPP sample



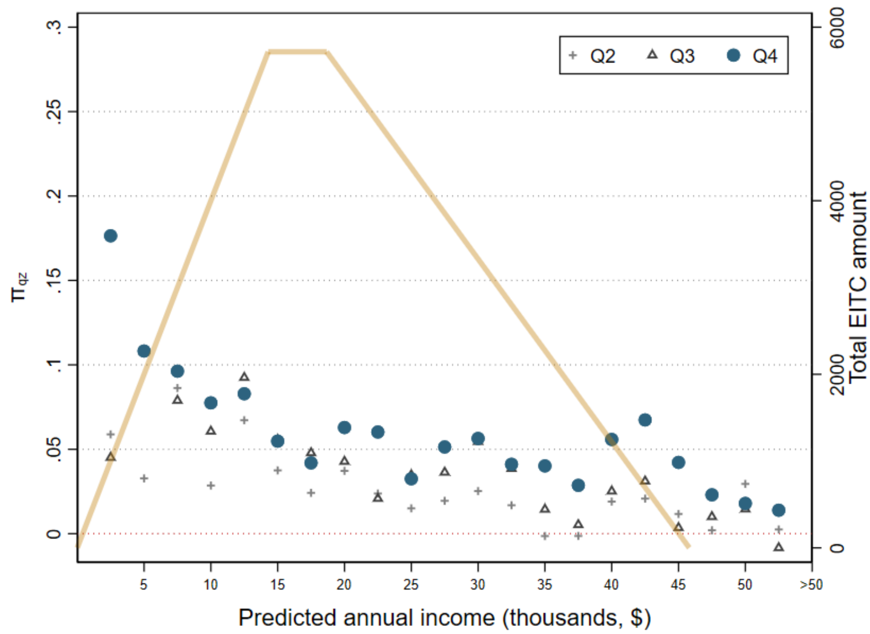
(b) SNAP sample



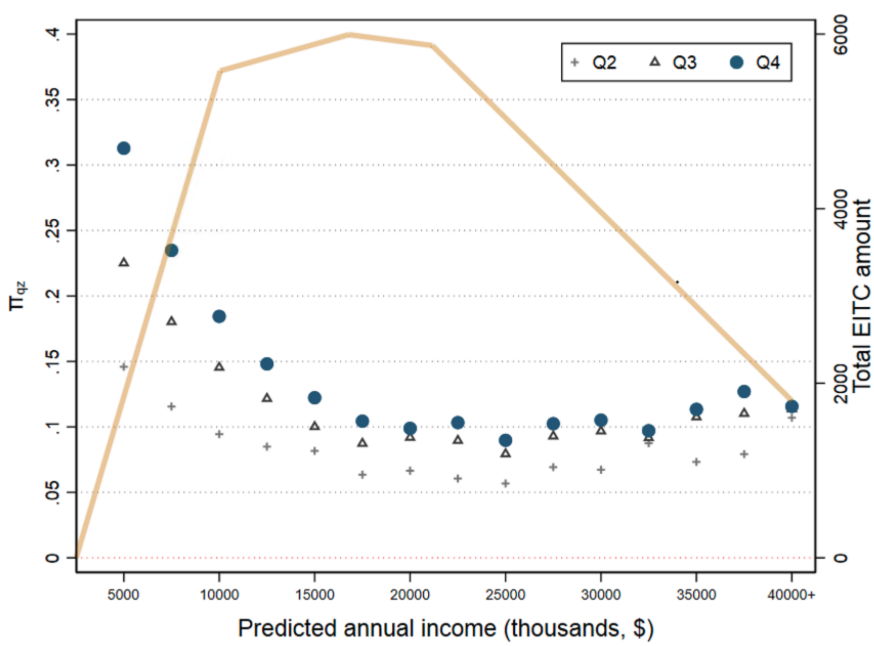
Notes. Figure 13 plots estimates of $\pi_{q\omega}$ from Equation 4. I bin households into twelve levels of ω: below .55, above 1.1, and in between, groups of .05. I implement the same restrictions described above. I limit to households whose earnings in three previous quarters are between $2,000 and $75,000, and with positive earnings in all quarters in a tax year. I also restrict to households whose total EDD wages equal their state AGI. Standard errors are clustered at the household-level.

Figure 14: Difference in earnings adjustment between quarters 2, 3 and 4 relative to quarter 1 over predicted annual income

(a) SIPP sample



(b) SNAP households



Notes. Figure 14 plots estimates of π_{qz} from Equation 5. The coefficients identify the difference in log earnings by predicted annual income between each calendar quarter relative to the first quarter. I overlay the estimates from the SIPP sample on the EITC schedule for a family with two dependents in 2008. I overlay the estimates among the SNAP households on the combined EITC schedule in California for a household with two dependents in 2017. Standard errors are clustered at the household-level.