Does the Earned Income Tax Credit Reduce Saving by Low-Income Households?

*Caroline E. Weber**
University of Michigan

Job Market Paper
November 2011

Abstract

This paper analyzes the effect of the Earned Income Tax Credit (EITC) on non-labor income, particularly investment income. Policy-makers have devoted substantial time and resources toward increasing the saving rate of low-income households with programs like the Saver's Credit and Individual Development Accounts. Yet the EITC—the largest federal cash transfer program in the U.S.—provides a substantial disincentive for individuals to save and realize investment income because EITC benefits decline as investment income rises over certain income ranges. I find a significant response of investment income to these disincentives both when I semi-parametrically estimate the change in investment income in the region in which non-labor income begins being taxed, and also when I estimate the parameter parametrically using differences in tax rate changes between taxpayers with differing numbers of dependents for identification. A one percent increase in the after-tax return to saving causes a 3.05 percent increase in investment income. Nearly 40 percent of the decline over the last two decades in the fraction of EITC recipients with savings in income-bearing accounts can be explained by changing EITC incentives.

A special thanks to my dissertation committee for their invaluable encouragement and comments: Jim Hines, Joel Slemrod (Chair), Jeff Smith, and Kevin Stange. Also, many thanks to David Agrawal, Dave Cashin, Rob Garlick, Susan Godlonton, Emmanuel Saez, Daniel Wilson (discussant), International Institute of Public Finance 2011 participants, U.S. Department of Treasury seminar participants, and University of Michigan seminar participants for helpful comments. All remaining errors are my own.

*Ph.D. Candidate, Department of Economics, University of Michigan, 611 Tappan Street, Ann Arbor, MI 48109-1220: ceweber@umich.edu.

1 Introduction

Policy-makers have devoted substantial time and resources toward increasing the saving rate of low-income households with programs like the Saver's Credit and Individual Development Accounts. Yet the Earned Income Tax Credit (EITC)—the largest federal cash transfer program in the U.S.—provides a substantial disincentive for individuals to save and realize investment income because EITC benefits decline as investment income rises over certain income ranges. Over the last two decades, an average of 17.6 percent of low-income individuals that claim the EITC have some dividend and interest income, but strikingly, the fraction has declined by more than 50 percent over time, from 26.2 percent in 1988 to just 12.3 percent in 2006. In this paper, I determine the extent to which the disincentive to save created by the EITC has contributed to the decline in saving in income-bearing accounts. I use Individual Public Use Tax Files for 1988-2006, which contain data on a random stratified cross-sectional sample of individual income tax returns.

This paper uses two different sources of variation in the implicit EITC tax rate in the region in which the EITC benefit is declining (known as the phase-out tax rate) to examine the degree to which investment income falls as the implicit phase-out tax rate rises. The first identification strategy examines a narrow region around the second EITC tax kink. When an individual is below this tax kink there is no implicit tax on a marginal dollar of income whereas the phase-out tax rate applies on the margin to all types of adjusted gross income (AGI)—earned income, non-labor income, and deductions—above the second tax kink for most individuals. The most prevalent type of non-labor income is investment income. I find that AGI responds to the phase-out tax rate, and that this response is significant for both wage-earners and the self-employed. For wage-earners, this response is concentrated in non-labor income, particularly investment income. Investment income declines significantly above the tax kink, where the after-tax return to saving is lower. The implied elasticity of investment income with respect to the after-tax return to saving, which bounds the true elasticity from below, is 1.88.

The main estimates come from a difference-in-differences specification, which exploits differences in tax rate changes between taxpayers with differing numbers of dependents for identification. The extensive margin estimates imply that nearly 40 percent of the decline over the last two decades in the fraction of EITC recipients with savings in income-bearing accounts can be explained by changing EITC incentives. The elasticity of investment income is 3.05; that is, a one percent increase in the after-tax return to saving increases the amount of investment income by 3.05 percent. Given the increase in the phase-out tax rate over this period, the elasticity estimates imply a $1,776 decline at the mean in the annual amount in savings for those that held positive investment income in 1988-1990. This calculation assumes a five percent rate of return on saving, on average; if it is lower, then the decline will be even higher. A key advantage of exploiting both types of variation is that the assumptions required in order to obtain consistent estimates are roughly orthogonal in the two cases. The fact that the results from both are similar provides additional credibility to the estimates obtained.

The paper proceeds as follows. Section 2 provides relevant background information regarding the EITC as well as a brief literature review. Section 3 discusses the data and provides an overview of the empirical strategy. Section 4 conducts the estimation around the second tax kink in the EITC schedule. Section 5 conducts the difference-in-differences estimation, which provides the main estimates in this paper. Section 6 concludes.

2 Background

2.1 EITC Details

The EITC, which was introduced in 1975, has expanded over time and is now the largest federal cash transfer program in the U.S. The EITC is administered through the tax system, which is not true of other major cash transfer programs in the U.S. This means that the credit is received as a lump sum annually and is based on annual income, whereas other transfer

programs are administered at a monthly frequency.[1] The federal EITC is refundable, which allows individuals to receive the benefit even if it is larger than the taxes they owe. Some states also offer a credit, which usually takes the form of a "piggy-back" credit (meaning that it is calculated as a certain percentage of the federal credit). The state credit is usually refundable.

When determining the EITC amount, several types of income are relevant. Earned income is defined as wage and salary income plus self-employment income (minus one-half the self-employment tax in the years 1990-2006), where self-employment income includes income from a sole proprietorship or a farm.[2] Non-labor income includes categories such as investment income and unemployment insurance receipts; for a complete list, see Section 3.1. Deductions have varied over time; examples include alimony paid, moving expenses, and IRA contributions.[3] I will refer to non-labor income minus deductions as unearned income. AGI, which is a term used on the tax form, is the sum of earned and unearned income.

The EITC schedule consists of three regions—the phase-in region, the plateau region, and the phase-out region—and is depicted in Figure 1. The example shown in Figure 1 is for taxpayers with two children, and reflects the tax schedule in real dollars for unmarried taxpayers in 1996-2006 (it also applies for married filers in 1996-2001).[4] When an individual has no earned income, the EITC credit is zero. In the phase-in region, the implicit marginal tax rate is negative (the benefit is increasing); in the plateau region, the implicit marginal tax rate is zero (the benefit is constant); in the phase-out region, the implicit marginal tax rate is positive (the benefit is declining). When the end of the phase-out region is reached, the credit is again zero. The solid line is the EITC schedule as depicted in the existing EITC

[1]Individuals can receive their benefits at a monthly frequency using the Advance EITC, but very few people choose this option.

[2]There are a few minor adjustments to the measure of self-employment income for a small number of individuals, but these cannot be taken into account in this paper because the tax return data do not provide these details.

[3]The term deduction refers to "above-the-line" deductions; that is, it does not include those only available to individuals that itemize. Also, since one-half the self-employment tax is included in earned income, it will not be counted as a deduction in this paper (otherwise it will be counted twice, once in earned income and once in unearned income).

[4]The tax schedule is inflation-indexed in this period.

literature. It applies anytime an individual has no or negative unearned income, which is true for 75 percent of two-child EITC filers in 1996-2006.

When an individual has positive unearned income, it is included as income for the purposes of calculating the EITC anytime AGI is above the second kink (i.e. the individual is in the phase-out region). When AGI is above the second kink, a filer's EITC is the minimum of the benefit as determined by their earned income and the benefit given by their AGI. In this two-dimensional graph, this can be depicted by shifting the solid phase-out region line towards the dashed line (or even past it). The phase-out region is shifted in for 25 percent of two-child filers in 1996-2006; the median amount of unearned income for these taxpayers is $1,227. Approximately one-fifth of these individuals have unearned income that is sizable enough such that they face no plateau region at all, facing a spike instead. Note that the location of the second kink appears to be person-specific only because the horizontal axis in Figure 1 is earned income, while AGI is the variable determining the EITC benefit received for most individuals. Figure 2 further illuminates the interaction between earned and unearned income in determining an individual's EITC by plotting a three-dimensional graph in earned income, unearned income, and EITC benefit. As earned income rises, the EITC first rises and then falls. As unearned income rises, it first has no effect on the EITC and then causes it to decline at a certain threshold, where the threshold varies depending on the individual's earned income amount.

Table 1 lists historic federal EITC parameters for taxpayers with at least one qualifying child.[5] Households without children became eligible for the EITC after 1994, but the phase-out region occurs at a much lower income level ($6,710 in 2006 dollars), so I do not anticipate these individuals to have a substantial investment income response (or to have substantial

[5]The state "piggy-back" credits followed a similar expansionary trend. Prior to 1997, at most six states offered an EITC credit in any given year. The state EITC expanded in terms of generosity and the number of states offering the credit in the years 1997-2006. In 2006, 19 states offered a state EITC for families with children (for most states the same EITC applies for those without children) and the average rate across states was 15.5 percent. Fifteen of these states had a refundable credit. Conditional on having at least one child, only Wisconsin varies the percentage of the federal EITC benefit given depending on the number of dependents.

investment income). They also are a relatively poor comparison group for families with children. Therefore, these individuals are not part of the analysis that follows.

2.2 Related Literature

There is a substantial literature on low-income saving decisions, but the effect of the EITC on saving behavior has never been examined beyond the role played by the lump sum nature of its distribution. This section provides a brief overview of existing work on the EITC and saving decisions of low-income households as it relates to this paper.

The EITC literature has focused almost entirely on the effect of the EITC on labor supply, the source of true earned income.[6] The overall effect of the EITC on labor supply is theoretically ambiguous because the phase-in region encourages work and the phase-out region discourages work on the margin. The literature generally finds that the EITC increases work, mostly on the extensive margin, for single individuals.[7] True earned income may not equal earned income reported to the tax authorities. Studies that make use of tax return data distinguish between wage-earning and self-employed individuals. These studies find that, while the intensive margin response by wage-earning individuals is approximately zero, there is a substantial response among the self-employed and there is a larger response to the phase-in region (LaLumia, 2009; Saez, 2010). This response is a combination of labor supply, tax avoidance, and tax evasion responses.

The literature that uses audit data to examine the likelihood that qualifying individuals claim the EITC and the literature on tax evasion as it pertains to the EITC both have some relevance for interpreting the results in this paper given that tax return data are used. Blumenthal et al. (2005) use data from the Taxpayer Compliance Measurement Program (TCMP) on both filers and non-filers in 1988, which contain audit data for a stratified random sample of these individuals. The authors estimate that 89 percent of EITC individuals above

[6]There is, though, an emerging literature that has looked at other outcomes. Examples include the EITC's effect on fertility (Baughman and Dickert-Conlin, 2003), time use (Gelber and Mitchell, 2011), and children (Dahl and Lochner, 2010).

[7]See Hotz and Scholz (2003) for an extensive review of the EITC literature through 2002.

the filing threshold claimed the EITC credit if eligible.[8] Thus, most EITC-eligible individuals in the region I examine will be observed in the tax return data.

McCubbin (2000) finds that claiming an extra dependent to which the taxpayer is not entitled increases with EITC generosity (a one percent increase in EITC generosity increases the likelihood that a taxpayer reports an extra child by about four percent) and decreases with income using IRS Criminal Investigations Division random audit data for 1994. Given this, I will consider possible biases induced in the estimates in Section 5 if individuals manipulate the number of children they report.

There is not much evidence that wage-earners evade by adjusting their earned income, but there is evidence that self-employed individuals do. Joulfaian and Rider (1996) examine the responsiveness of income reported to the EITC marginal tax rate using 1988 audit data from the TCMP. They find that 94.6 percent of wage-earning individuals accurately report their earned income, and the EITC rate has no effect on the decision to report accurately. For sole proprietors, 19 percent report accurately and they find that a one percent higher EITC rate increases the probability of underreporting by 0.3 percent.

There is a literature on the saving decisions of low-income households, but not about the role played by the EITC.[9] One strand of this literature examines the effects of other need-based programs, such as AFDC/TANF and Medicaid, on savings. Hubbard et al. (1995) show that these programs affect savings decisions both because they decrease the need for precautionary savings in case of a negative income shock and the asset limits imposed by these programs create a high implicit tax rate when binding. Gruber and Yelowitz (1999) find empirical evidence supporting these theoretical results in the context of Medicaid. In contrast, recent papers examining the effect of AFDC/TANF asset limits (Hurst and

[8]The fraction was substantially lower if the taxpayer was below the filing threshold. The second EITC kink is near (within $1,000) or above the filing threshold for all years.

[9]The only evidence of the effect of the EITC on saving decisions is ethnographic and it is interested not in the effect of the EITC schedule, but rather in the lump sum nature of its distribution. Romich and Weisner (2000) conduct extensive interview-based analysis of 42 EITC recipient families in Wisconsin. They find that individuals prefer to receive the credit as one lump sum, rather than using the advance payment option. They find that 32 percent of the individuals in their sample had saved some of their EITC credit in a bank account and kept it in the account at least two months after the receiving their EITC credit.

Ziliak, 2006; Sullivan, 2006) on liquid assets, such as saving accounts, find no substantial effect. The EITC changes the after-tax return on savings for individuals in the phase-out region, providing a mechanism to decrease savings in addition to asset limits or precautionary savings. This mechanism will be the primary focus of this paper, but precautionary savings motives will be discussed when relevant in Section 5. The asset limits are much more likely to be binding in the case of TANF or Medicaid because the limits are effectively much lower for TANF and Medicaid than the limit introduced for the EITC in 1996.

A separate literature examines the effectiveness of programs designed to induce low-income individuals to save more. In addition to providing evidence on the degree to which individuals can be encouraged to save more, this literature also highlights a recent policy focus on increasing low-income saving. One popular program is Individual Development Accounts (IDA), which encourage individuals to save by providing matching funds (up to a limit each year) for income saved in these accounts. The individuals receive the matching funds when they withdraw their savings for qualified purposes (e.g. buying a house or paying for college). Over 400 IDA programs existed by 2006, which were privately funded or funded by states (Mills et al., 2008). Mills et al. (2008) examine a particular IDA program in Oklahoma with a randomized design; they find a large IDA take-up rate among the treated group (almost 90 percent). Those with IDA accounts are more likely to purchase a home, but this is somewhat offset by a decrease in other forms of asset accumulation (this IDA provided a larger match when the funds were withdrawn for the purpose of buying a home).

Another program, the Saver's Credit, was implemented through the tax system in 2002 to encourage saving for retirement among low-income individuals. The literature has found little impact on saving of this credit; for example, see Ramnath (2010). Duflo et al. (2006) conduct a randomized field experiment of a program similar in spirit to the Saver's Credit, and find that they can indeed induce individuals to save more. They cite differences between their experiment and the Saver's Credit, such as the complexity and non-refundability of the Saver's Credit, as likely to be driving the difference in findings between their experiment and

the studies that examine the Saver's Credit.

The existing EITC literature has found little effect of the EITC on labor supply decisions along the intensive margin (except perhaps for self-employed individuals). My paper examines the extent to which it has an effect on another margin—non-labor income, particularly investment income—which has not been previously examined. Given the existing empirical literature on low-income saving, the degree to which the EITC will affect low-income saving is not obvious, since the effects of other programs on low-income saving is mixed.

3 Empirical Strategy

This section provides an overview of the empirical strategy used in this paper. Because the methods used in each section are quite different, the details can be found in their respective sections. Subsection 3.1 provides a detailed description of the data. Subsection 3.2 explains the focus of this paper, the degree to which a change in saving in an income-bearing account reflects an actual change in savings, and the different forms of variation used for identification in each section. Subsection 3.3 discusses the central role of awareness in studying the impact of the EITC on saving, which is key to interpreting the estimates.

3.1 Data

This paper uses the Individual Public Use Tax Files for 1988-2006, which were created by the Statistics of Income Division of the Internal Revenue Service. The sample is stratified, so all estimates are population weighted. Some components of AGI—wage and salary income, alimony paid, and alimony received—are blurred in the tax return data for low-income individuals in the years 1996-2006.[10] Alimony income was blurred at the national level. Wage and salary income was blurred at the national level in 1996, at the state level for years 1997-1999, and at the state x married level for years 2000-2006. The blurring was done as

[10] Alimony paid and received were blurred beginning in 1995.

follows: "for every three records, in descending order, the average...[of a given income type to be blurred] has been determined and that value has been placed in the...field for each of the three records." (Weber, 2001). This means that earned income for all individuals that have positive wage earnings is measured with noise in the years in which blurring occurred. An alternative measure of earnings can be constructed by backing out earned income from AGI. This measure is also imperfect because most, but not all, types of unearned income are reported in the tax return data, and those that are also blurred must be excluded.[11] I use the latter measure for all years throughout the paper, so the accuracy of my imputation method can be precisely determined. In the years before blurring, the latter measure is within $15 of the actual wage for 99.5 percent of taxpayers for the sample used in Section 5. In the years after blurring, the blurred wage is within $15 of the latter measure 65 percent of the time, indicating that there is substantial noise in the blurred wage. Still, the estimates are robust to the use of either measure; for example, the baseline estimates in Section 5 decline by about 2 percent when the blurred wage is used instead.

Table 2 provides descriptive statistics for all EITC-eligible individuals for 1988-2006. The first two rows summarize the components of earned income. Almost 16 percent of individuals have some self-employment income, but it is the sole source of earned income for only 1.6 percent of taxpayers. The third row provides summary statistics for total unearned income, and the next nine rows break out unearned income by type.[12] Note that 33.6 percent of individuals have some unearned income, with a mean of $571.66. The most frequent type of unearned income is dividend and interest income (18.6 percent). The next three rows—capital gains or losses, partnership income, and other gains or losses—reflect other types of income (or loss), but less than five percent of individuals have these types of income. Taxable refunds only apply to those who itemize deductions and are reported by a small fraction of

[11]The latter affects a relatively small number of individuals in the sample used in this paper.

[12]Note that unearned income does not include blurred sources (alimony) and unreported sources (other income and a few deductions). Less than one percent of individuals in this sample have alimony income. In the years before the blurring of low-income wage data was instituted, I can back out unearned income directly by subtracting earned income from AGI. The likelihood that individuals have unearned income in these years (1988-1995) is 1.4 percentage points higher using this measure.

individuals (3.2 percent). Taxable pension, annuity and IRA distributions are only received by about five percent of the sample and, unlike the previous categories, are unlikely to respond unless individuals adjust the amount received each year, decreasing it if they are temporarily EITC-eligible. Almost no one has taxable Social Security benefits, which also are not expected to be adjusted in response to EITC schedule changes. A substantial number of individuals have unemployment insurance (10.7 percent). Individuals will not adjust this form of income directly in response to the EITC schedule (because it does not pay to give up a dollar of unemployment insurance to obtain 10-20 cents of EITC benefits), but it may be adjusted indirectly if individuals change the number of weeks worked in response to the EITC. About five percent of individuals have deductions and the mean is $103.94; deductions are likely responsive to the EITC schedule. About 30 percent of the sample is married, and over 60 percent use a tax preparer. Conditional on filing a tax form, approximately 92 percent of EITC-eligible individuals claim the EITC.[13] As noted in Subsection, 2.1, I exclude those with no dependents. For most of the analysis in Section 5, I restrict the sample to those with one and two dependents; for comparability, I impose the same restriction throughout the paper. The mean number of dependents is 1.41.[14]

The analysis of tax return data has several advantages. Most importantly, it provides a precise (or near precise) measure of the reported amount of each type of income, and is not subject to survey non-response, rounding, and other approximations. The main disadvantage of tax return data is that one cannot observe where individuals place the money that they would have held in an income-bearing account in the absence of a positive tax rate or those that do not file. For welfare, both moving the money out of the income-bearing account (assuming this is not costless to the individuals) and a real decline in saving are relevant;

[13]The actual number is likely slightly higher, because I cannot perfectly measure eligibility with the data available.

[14]Dependents do not always correspond exactly to the number of children claimed for EITC purposes; the latter is not in this data set for all years. For years in which both are available, the actual measure is about one-tenth of a child lower. In addition to data availability, the advantage of using dependents is that it avoids additional manipulation of claiming children only for the purpose of maximizing the EITC. For years in which I have both measures, the baseline estimates decline slightly when the actual measure is used, but the estimates are not statistically different.

this data can only address the former. Tax evasion on investment income (i.e. not reporting some or all investment income) is also a potential drawback, although I will be able to provide some evidence below that evasion is likely not a key mechanism behind the results found.

3.2 Research Design

The key research question examined in this paper is whether the design of the EITC has the unintended consequence of distorting non-labor-income decisions in the phase-out region, particularly the decision to save in an income-bearing account. This paper chooses to focus mostly on the decision to save in an interest- or dividend-bearing account. I also analyze a broader measure of investment income, which includes capital gains and losses, partnership income, other gains and losses, and IRA contributions (IRA contributions are a deduction).[15] I focus on a particular part of unearned income because unearned income contains a broad range of different types of income. As a result, the coefficient on unearned income does not have much policy relevance outside of choosing an optimal EITC design. In contrast, examining the effect of the EITC on saving of low-income households fits into the substantial literature on the effect of mean-tested programs on savings. Furthermore, unearned income contains both types of income one would expect to be responsive to the tax rate (e.g. dividend and interest income) and types that one would not (e.g. Social Security benefits).

The decision to save in an interest- or dividend-bearing account is related, but not identical, to the decision to save, as the EITC could induce individuals to stop saving in interest- or dividend-bearing accounts, instead placing their money in an interest-free account. This paper has nothing to say about this mechanism, but this response is distortionary and presumably not a behavior policymakers would like to encourage. There are several other

[15]Capital gains are not included in the primary measure of the return on saving in an income-bearing account because they are a very lumpy measure of asset accumulation, and there is a large capacity to retime the sale of an asset to a period in which the tax rate is lower. In reality, very few individuals in this range of the income distribution have capital gains, and I will show that the estimates are not affected by the inclusion of capital gains.

reasons why the measured response may not be be identical to the decision to save. First, if two taxpayers are cohabiting, one could save while the other claims the child and gets the EITC. This paper cannot provide evidence of the degree to which the measured response is due to this mechanism, because I cannot observe tax returns of two separate cohabitors. Second, individuals could save in an interest- or dividend-bearing account, but choose not to report it on their tax return—tax evasion. I argue that tax evasion is unlikely because there is information reporting by financial institutions on interest and dividend income above $10 per account, but it is possible, particularly for those with dividend or interest income less than $10. I will return to the issue of information reporting and tax evasion in Section 5, where I provide evidence that, under reasonable assumptions, any evasion of taxes on investment income has a minimal effect on the results. Going forward, much of the discussion will proceed as though individuals are altering their real savings decisions, but readers should keep in mind that the estimates in this paper are reduced-form.

To identify the effect of the EITC phase-out tax rate on the decision to save in an income-bearing account, I exploit two different forms of variation in the tax rate faced. Section 4 uses variation in the marginal EITC tax rate above the second EITC kink, relative to the marginal EITC rate below. The evidence in this section comes from a narrow region in the vicinity of the tax kink. It provides evidence on the overall response in earned and unearned income above the second kink, as well as which types of income are most responsive to the higher tax rate above the second kink, and thus driving the overall response. The second source of variation used in this paper exploits the feature that the phase-out tax rate has varied depending on the number of dependents an individual claims since 1991. Section 5 exploits this variation using a difference-in-differences estimator.

An advantage of these two different forms of identification is that they require roughly orthogonal identifying assumptions. The key identification assumption for estimating the response of investment income in Section 4 is that investment income as a fraction of AGI would increase at a constant rate, on average, in the region around the second EITC tax

kink if no tax kink existed in this region. Broadly speaking, the estimation strategy in Section 5 requires that the composition or secular behavior of those observed with one or two dependents cannot change differentially for one relative to two dependents in a way that is correlated with changes in the tax rate and outcomes. The fact that the results from both are similar provides additional credibility to the estimates obtained.

Individuals receiving the EITC are also subject to the regular tax schedule and may also be receiving subsidies from other cash transfer programs, such as AFDC/TANF, Medicaid and foodstamps/SNAP. These interactions will matter only if changes in these programs coincide with changes in the EITC tax rate and also discourage saving. Often, individuals in the phase-out region are not eligible to receive benefits from these other need-based programs. Moreover, the disincentives to save from these programs declined (while the disincentives from the EITC increased) over the period I examine and the decline in disincentives did not vary by number of children (Hurst and Ziliak, 2006; Gruber and Yelowitz, 1999), which means that these programs are unlikely to bias my results. Regarding the regular tax schedule, interactions between the EITC schedule and the regular tax schedule will be discussed below when they are important for the analysis.

3.3 Awareness

Awareness of the tax system is important in studies where the treatment is the marginal tax rate faced, because individuals must know and respond to the tax rate they will face in advance; this is in contrast to more standard treatment effect studies (e.g. the effect of a job training program) in which individuals know the treatment they receive before responding. When individuals are filing their tax return, it is too late to respond, except by engaging in tax evasion or making an IRA contribution.[16] A small literature finds that awareness of detailed features of the EITC schedule is limited (Liebman, 1998; Romich and Weisner, 2000). But, there is no reason to think that awareness of the EITC is different than that for

[16]IRA contributions can be made up to April 15 and be applied as a deduction to the previous year. However, very few EITC individuals make IRA contributions.

other features of the tax schedule or other cash transfer programs.

Awareness matters for several reasons. First, individuals must be aware that their unearned income is taxed if they are in the phase-out region. The fact that unearned income is taxed in the phase-out region has often been ignored in the EITC literature. However, this feature of the EITC is prominent when taxpayers fill out the form to calculate their EITC.[17] A picture of this form for 2006 is displayed in Figure 3. If (and only if) individuals' AGI places them in the phase-out region, they are asked to calculate their credit based on their earned income, and then again based on their AGI. They are then required to take the minimum of the two credit amounts. If individuals' AGI makes them ineligible for the credit, they are stopped before they start to fill in the form. Thus, individuals who are filing for the first time, or even are in the phase-out region for the first time, may not know about this feature until they fill out the tax form. This is too late for the individuals to respond, unless the individuals decide to engage in tax evasion or make IRA contributions.

Second, individuals must be aware of changes in the phase-out tax rate over time. Third, conditional on knowing that their income will be taxed at a higher rate in the phase-out region, individuals must also know whether they will be in the phase-out region, and thus face the phase-out tax rate, in advance. Before 1987, the EITC schedule was not inflation-indexed, which meant that the location of the tax kink remained the same in nominal dollars year after year, unless there was a tax reform. After 1987, the EITC schedule was inflation-indexed with a lag. For example, the inflation parameter which generated the 2006 tax schedule was the CPI for September 2004-August 2005. Rounding is then used (usually the values are rounded down to the nearest $50). For this reason, it is somewhat unlikely that individuals' income will naturally move in lock-step with the tax kink after inflation-indexing was introduced. An individual can know at the beginning of the tax year what the tax parameters will be and respond accordingly, but this requires them to look up the

[17]These features are less prevalent when individuals use TurboTax, but this type of electronic filing was not too common in the years examined. Only 13 percent of taxpayers filed their returns from home using tax preparation software in 2005 and this was likely not concentrated among low-income households. Free tax preparation software for low-income households did not exist until 2006.

location of the tax kink each year. As individuals' income gets further away from the tax kink, the amount of awareness necessary for them to realize the treatment they face and respond accordingly decreases, because they can more easily guess whether or not they will remain in the phase-out region in the following year.

The baseline estimates in Sections 4 and 5 ignore the issue of awareness. The estimates in Section 4 are expected to be lower than those in Section 5 precisely because it requires a greater degree of awareness to respond each year in a narrow region around the tax kink, than somewhere in the phase-out region. However, several additional specifications in Section 5 will analyze ways in which individuals become more or less aware of EITC-related incentives, including their use of a paid preparer and having unemployment insurance (where unemployment insurance is considered a rough proxy for whether the individual experiences a transitory shock in earnings).

4 Estimating the Response at the Second Tax Kink

This section examines the behavior of people whose income places them in the income region around the second tax kink, exploiting the change in the EITC marginal tax rate at the kink. The section first estimates the overall response in AGI to the change in the tax rate using a modification of the method proposed by Saez (2010), which relies on estimates of bunching in the region around the tax kink. I find small, but statistically significant elasticities at the second EITC kink for *both* wage-earning and self-employed individuals. I then implement a procedure that provides evidence regarding which types of income are most responsive to the change in the EITC rate at the kink; that is, which types decline significantly above the kink. The response for wage-earners is driven by changes in unearned income, particularly investment income.

4.1 Bunching

A classical model predicts bunching at tax kinks when the budget set is convex, because individuals above the tax kink will wish to decrease their income as the tax rate above the kink rises given their preferences (labor-leisure preferences in the context of earned income, saving-consumption preferences in the context of dividend and interest income, and so forth). However, once they reach the tax kink, they have no incentive to decrease their income further. This is because the tax rate below the kink did not change and they were already willing to earn income at this tax rate before the tax rate increased above the kink. If individuals were able to perfectly adjust their AGI, all these individuals would be located at a point mass at the kink; however, adjustment costs and income volatility generate imperfect bunching around the tax kink. Saez (2010) shows that the amount of bunching in the region near a tax kink can be used to calculate an elasticity of AGI with respect to the marginal net-of tax rate using the following formula:

$$\hat{\varepsilon} \simeq \frac{\hat{B}/\hat{h}_0(z^*)}{z^* ln\left(\frac{1-\tau_l}{1-\tau_h}\right)}, \tag{1}$$

where z^* is the point in the income distribution at which the tax kink is located, τ_l and τ_h are the marginal tax rates on either side of z^*, such that $\tau_l < \tau_h$. Let there be a region $[z^* - \delta_b, z^* + \delta_b]$ in which all imperfect bunching occurs. Then, $\hat{B}$ is the estimated excess mass and $\hat{h}_0(z^*)$ is the estimated counterfactual density of taxpayers in this region; that is, the density of taxpayers if there was no tax kink. These parameters are depicted graphically in Figure 4. Saez (2010) implements this method to estimate EITC elasticities at each kink; however, his analysis ignored that individuals' location relative to the second kink is usually determined by AGI, not earned income, as stressed in this paper.

To estimate $\hat{B}$ and $\hat{h}_0(z^*)$, I first estimate the smoothed density $\hat{h}$ using a local linear regression within \$12,000 of the tax kink.[18] I use the smoothed density estimates to con-

[18]The figures are trimmed from -\$9,000-\$9,000 to exclude the edges which are noisy because the estimates

struct the elasticities, because this method provides substantial efficiency gains, relative to constructing the estimates using unsmoothed estimates, which is the method used by Saez (2010) and Kleven and Waseem (2011). I use an automatic bin and bandwidth selection criteria[19] to select the bin and bandwidth size within $[z^* - \delta_b, z^* + \delta_b]$ and apply this bin and bandwidth to the entire income range.[20] As long as the bin size chosen is small relative to the selected bandwidth, the exact choice of bin size is usually not important. However, Saez (2010) effectively over-smooths the data by using a bin size ($500) that is an order of magnitude larger than the bin size used here for his smoothed density figures (it is equivalent to using a uniform kernel at all points within the bin size). Over-smoothing the data would induce a downward bias in my elasticity estimates; for example, using the baseline bandwidth reported in this paper for self-employed individuals for all years, but replacing the bin size with $500, biases the estimates downward by more than 10 percent, with a minimal decrease in variance.

Let $\hat{h}^*_-$ be the estimated mean density in the region $[z^* - \delta_b - \delta_c, z^* - \delta_b]$, $\hat{h}^*$ be the estimated mean density in the region $[z^* - \delta_b, z^* + \delta_b]$, and $\hat{h}^*_+$ be the estimated mean density in the region $[z^* + \delta_b, z^* + \delta_b + \delta_c]$, where δ_c is the width of the counterfactual region on either side of the imperfect bunching region. Let $\hat{H}^*_- = \delta_c \hat{h}^*_-$, $\hat{H}^* = 2\delta_b \hat{h}^*$, $\hat{H}^*_+ = \delta_c \hat{h}^*_+$, where these variables denote the cumulative density in their respective regions. Ideally, $\hat{h}^*_+$ would be the counterfactual density, absent a tax kink, in the region $[z^* + \delta_b, z^* + \delta_b + \delta_c]$. However, when τ_h is the tax rate above the tax kink, it overestimates the counterfactual if individuals above z^* decrease their income by a certain percentage in response to the higher tax rate (as opposed to decreasing their income by a fixed amount); this is consistent with what our

are smoothed over a small number of observation at the edges.

[19]The bandwidth is based on the rule-of-thumb approach given in Fan and Gijbels (1996). This is the method used in McCrary (2008), whose estimation procedure is the same as is used here, except I do not allow for a break at z^*. Since I am examining tax kinks, not tax notches, imperfect bunching should take place on *both* sides of the tax kink. Therefore, to gain efficiency in the region of interest—just on either side of the tax kink—I do not allow for a break in the density at zero because one is not expected. The results when I do and do not allow a break are qualitatively similar, but the latter provides substantial efficiency gains.

[20]A bin size and bandwidth chosen over a larger region would not be appropriate to efficiently measure the density in the imperfect bunching region.

models predict and the parameter our estimation strategies are designed to uncover.[21] It overestimates the counterfactual in the region of the tax kink because the counterfactual distribution gets compressed as each individual decreases their income by a percent of their total income. Therefore, the elasticity estimates of a particular income type with respect to the marginal net-of-tax rate $(1-\tau)$ presented in this section can be interpreted as lower bounds on the truth.[22]

The counterfactual density is the mean of the densities above and below the region of imperfect bunching: $\hat{h}_0(z^*) = \frac{1}{2}(\hat{h}^*_- + \hat{h}^*_+)$. This method assumes that the densities on either side of the region of imperfect bunching are a good approximation for the density in the imperfect bunching region.[23] The excess mass is constructed by subtracting the counterfactual mass from the total mass in the bunching region: $\hat{B} = \hat{H}^* - \frac{\delta_b}{\delta_c}(\hat{H}^*_- + \hat{H}^*_+)$.

The other important consideration in constructing these estimates is the choice of the size of the imperfect bunching region, $2\delta_b$, and the size of the counterfactual region, δ_c. I set $\delta_b = \delta_c = \$1,000$ for the baseline estimates, which are the same as those used by Saez (2010), to analyze the second tax kink. The choices for δ_b and δ_c are based on the following considerations. First, the gap between the first and second kink for two-child families since 1996 is \$3,470.[24] Therefore, it must be that $2\delta_b + \delta_c \leq 3,470$, so that the counterfactual

[21] If τ_l is the tax rate above the tax kink, this discussion applies to $\hat{h}^*_-$ instead of $\hat{h}^*_+$.

[22] Previous literature has either ignored this issue (Saez, 2010; Kleven and Waseem, 2011) or assumed that the density is biased downwards because the bunching individuals came from the region above the tax kink (Chetty et al., 2011). The latter assumption is incorrect as long as individuals further above the tax kink are responding in the same way as those near the tax kink and the counterfactual region is not near the next tax kink (in the region near the maximum income taxed at a given rate, there is a decline because these individuals are no longer being replaced by individuals further up in the distribution).

[23] This is the most common way of constructing the counterfactual density in the literature (Saez, 2010; Kleven and Waseem, 2011). The other approach that has been used estimates a global polynomial to construct the counterfactual (Chetty et al., 2011). This method is ideal when the distribution has a shape that is well approximated by a low-order polynomial (even if ultimately a higher-order polynomial is used). This does not apply for this distribution, because there are two kinks that are relatively close together and are not defined by the same type of income. In the early years of the EITC when the kinks were further apart, it is possible, although perhaps not completely persuasive, to get estimates using this method. The results obtained are similar. In later years, it becomes impossible, particularly for those with two children, to use a global polynomial to control for one kink, and accurately estimate the counterfactual for the other. The global polynomial does a good job of estimating the counterfactual when it is approximately linear, which is the same assumption imposed when the method in this paper is used.

[24] The gap increases for married families by \$1,000 in the years 2002-2004 and \$2,000 in the years 2005-2006.

region for each kink does not include part of the bunching region for the other tax kink. If δ_b is smaller than the imperfect bunching region, the estimates will be biased downwards because part of the bunching will be excluded from $\hat{H}^*$ and will instead be included in $\hat{H}^*_+$. Assuming the counterfactual density is flat, choosing a δ_b that is larger than the imperfect bunching region does not bias the estimates, but will otherwise. In practice, choosing a large δ_b induces a bias in the estimates, because the extent to which $\hat{h}^*_-$ and $\hat{h}^*_+$ provide an accurate estimate of $h_0(z^*)$ usually declines further away from z^*. There is not much that can be done in the way of sensitivity analysis over the whole sample period given the constraint that the first kink imposes. I pursue this issue below, when I look exclusively at the years 1988-1993, when the first kink was about twice as far away as it was in 1994-2006.

I examine self-employed individuals and wage-earners separately because people in these two employment states arguably differ in their capacity to respond to tax kinks on several dimensions. An individual is defined as self-employed if they have any self-employment income. One way in which individuals can bunch at the tax kink is to simply not report income above the kink—tax evasion. This is quite difficult to do with wage and salary income and not get caught, as this income is generally subject to withholding and information reporting.[25] Self-employment income faces neither withholding nor information reporting.

Tax avoidance is another way in which individuals can respond. There are few tax avoidance possibilities on earnings for wages because all wages are reported on a W-2 and thus face withholding and information reporting requirements as discussed above.[26] Still, there is some flexibility if workers are able to substitute towards non-monetary forms of compensation. Self-employed individuals have a greater opportunity for tax avoidance. These individuals are allowed to deduct their expenses from their gross income, which allows them

[25]Note that there is a minimum requirement on withholding for wage-earning individuals. It is $221 for single individuals and $667 for married individuals in 2006 dollars. This should, in general, not have an effect on the estimates, unless individuals hold two jobs, one of which is below the threshold and they claim to only have one job on the Form W-4 for their primary job (Form W-4 is completed by individuals so that their employer knows how much to withhold). If individuals decide to report the income from their second job if they face the lower tax rate, but not if they face the higher rate, a small part of the response by these individuals could be tax evasion.

[26]It is possible that such individuals do receive cash payments for work done on the side for an employer.

to have an extra expense, say, which would place them at the tax kink without engaging in anything illegal.

A labor supply response could also move these individuals towards the tax kink. Self-employed individuals face lower adjustment costs, on average, associated with altering their labor supply decisions, so it is more feasible for them to make minor adjustments.

Each of these three possibilities—evasion, avoidance, and a labor supply response—all suggest that the income adjustments will be larger and more precise for the self-employed compared to wage-earners. The top panels of Figures 5 and 6 display the bunching at the second EITC kink for self-employed and wage-earning individuals for all years, respectively.[27] The densities are normalized so that they integrate to one. The variables $\hat{z}^*$, $\hat{\tau}_l$, and $\hat{\tau}_h$ refer to the population weighted averages for each sample. The solid yellow line below the density is the estimated counterfactual density of individuals in the bunching region. The bunching depicted in the figures is more sharply defined for the self-employed. The corresponding elasticities are given in Table 3 Columns (1) and (2). Standard errors for the elasticities were obtained by nonparametric bootstrap (500 replications). The elasticity estimate of adjusted gross income with respect to the marginal net-of-tax rate for self-employed individuals is 0.063 and is significant at the one percent level, while the estimate for wage-earners is 0.010 and insignificant.

Table 3 Columns (3)-(5) and the second panel of Figures 5 and 6 examine years 1988-1993 for self-employed and wage-earners, respectively. The key advantage of examining the early years of the tax credit is that the first kink was about twice as far away for filers with two or more taxpayers, so a broader range of δ_b and δ_c are feasible.[28] I fix $\delta_c = \$1,000$. For the

[27]The estimates are population weighted.

[28]Also, in the later years, the second kink interacted with several other tax schedule features. First, the filing threshold for married individuals was within $1,000 of the second kink in 1994-2006. It is unlikely that this had a substantial effect on filing, as it was clearly in the best interest of individuals on both sides of the tax kink to file to claim the EITC, but it is possible. Second, in 2002, the Saver's Credit was introduced and the first notch for the single filers is near the second EITC kink. Ramnath (2010) documents bunching at these kinks, but the number of individuals applying for the Saver's Credit was quite small, so it is unlikely that this would generate a substantial bias in these estimates. If either of these other programs were playing a significant role in the size of the overall estimates, those estimates would be biased upwards.

self-employed, I report estimates for $\delta_b = \$2,000$ and $\delta_b = \$1,000$ in Columns (3) and (4); the counterfactual density is drawn in Figure 5 for the case of $\delta_b = \$2,000$. The estimates are similar for both and visually $\delta_b = \$2,000$ appears correct, but there are efficiency gains when using $\delta_b = \$1,000$. The elasticity estimate for $\delta_b = \$1,000$ is 0.095 and is statistically significant at the 10 percent level. One could even make the case that δ_b should be \$3,000. In this case, the estimate roughly doubles and this rise is driven by the fall in density between -\$4,000 and -\$3,000, which is now part of the counterfactual density. For wage-earners, the plotted density strongly supports $\delta_b = \$3,000$.

The fact that the bunching is more spread out for wage-earning individuals is consistent with the fact that they have less flexibility in adjusting their earned income. The elasticity estimate for wage-earners is given in Column (5). It is 0.104 and is statistically significant at the five percent level. The wage-earner and self-employed elasticities are about the same (once the difference in the precision of bunching is taken into account) for these years. There are several possible explanations for this finding. First, when self-employed individuals engage in tax evasion, they may move all the way to the first tax kink (in the region between the first and second kink, they receive no additional EITC benefit, but do have to pay additional payroll tax on their self-employment earnings). Alternatively, while it may be easier for the self-employed to alter their earned income, wage-earners make up the difference by adjusting their unearned income more. Given the discussion in Subsection 3.3, these parameters can be interpreted as a lower bound relative to what would be found further in the phase-out region, where the amount of awareness needed to respond precisely is lower.

4.2 Changes in Average Income

So far, I have presented evidence that individuals respond to the second EITC kink in a small, but statistically significant, way. However, this analysis cannot be informative regarding which types of income are most responsive, nor does the literature have any precedent for such an examination beyond splitting the sample by those with and without a particular type

of income (e.g. self-employed income) and examining whether the elasticity is different across the two groups. When the response is on the intensive margin, this approach will uncover the relevant response heterogeneity. However, when the response is on the extensive margin, the response of these individuals will be captured in the elasticity of those categorized as not having this type of income, because the researchers will only observe these individuals after they have given up all of this type of income. This is especially problematic for types of unearned income where the likelihood of an extensive margin response is higher. The rest of this section proposes a new alternative method of examining behavior around the tax kink for different types of income to provide some evidence regarding which types of income are the most responsive and are thus contributing to the elasticities found above.

To explain this method, I use the example of investment income, and assume that a tax kink is introduced as part of a tax reform, such that the reform discretely increases the marginal tax rate above the tax kink. Suppose individuals are able to bunch perfectly at the tax kink, there is no growth in investment income, and investment income is the only type of income that responds to the increase in the marginal tax rate above the kink. Then, investment income will decline above the tax kink as depicted in the top panel of Figure 7. Note that investment income is a part of AGI, and thus a change in investment income moves an individual's location on *both* the horizontal and vertical axes. Below the tax kink, investment income does not respond. At the tax kink, average income is comprised of all those that were at the kink before the reform, plus all those who move to the kink in response to the tax reform (these are all the individuals whose desired percent decrease in investment income at the new tax rate is at least as large as the gap between their pre-reform AGI and the tax kink). This point in Figure 7 is located half way between the old and new investment income levels, which occurs when there is a uniform distribution of investment income, everyone holds the same amount of investment income pre-reform, and everyone responds in the same way. In general, this point can be anywhere within the old and new

levels, or even above the old level.[29] Above the kink, investment income is at a new lower level. Under these assumptions, one could drop the point at the tax kink and then implement a regression discontinuity design to estimate the effect of the higher tax rate above the kink. Dropping the point at the tax kink effectively eliminates all manipulation of the running variable, which would otherwise bias the estimates.

With imperfect bunching and investment income growth, the picture changes to look like that in the bottom panel of Figure 7; investment income growth, even without imperfect bunching, mitigates the discrete drop in investment income above the kink if individuals are responding along the intensive margin. This occurs because individuals above the kink respond by decreasing their investment income by a certain percentage, but also started with more. This figure allows other types of income to be growing as well; their rate of growth is a determinant of the slope of investment income. But, it assumes investment income is the only type of income that responds to the tax rate. I discuss what happens when this assumption is relaxed below. Imperfect bunching causes investment income to decline once individuals start responding to the tax kink at $-\delta_b$; that is, for a given level of AGI, investment income will be lower (it can also increase for the same reason as described in the perfect bunching case when investment income increases at the point of the tax kink). By the time AGI has exceeded the imperfect bunching region, investment income will be at a new, lower growth rate.[30]

By calculating the percent change in average investment income in a region above δ_b relative to the counterfactual investment income that would have existed absent the tax change, backing out the elasticity of investment income with respect to the marginal net-of-tax rate is straightforward. The counterfactual is empirically constructed from the region

[29]The latter is more likely to occur as the response by those with investment income decreases and the number of individuals with no investment income increases.

[30]Because imperfect bunching exists and the outcome (investment income) and assignment variable (AGI) are simultaneously realized, calculating the change in slope by taking the limit as approach the tax kink (using a method similar to that in Card et al. (2009)), would necessarily yield an estimate of approximately zero (i.e. with imperfect bunching, the individuals are approximately the same just on either side of the tax kink).

$[z^* - \delta_b - \delta_c, z^* - \delta_b]$; for the counterfactual to be valid, investment income must continue to grow at the same rate in the region in which the elasticity is calculated. However, the standard errors on the counterfactual increase as AGI increases above $-\delta_b$, because the accuracy with which I can predict the counterfactual using investment income in the region $[-\delta_b - \delta_c, -\delta_b]$ decreases as AGI increases. For this reason, I instead calculate the elasticity in the region of the tax kink $[-\delta_b, \delta_b]$, and acknowledge that this is a lower bound on the true elasticity.[31] I will now define some notation, which is depicted in Figure 7.

Let X^*_{i+} be average investment income for bin i and let X^*_{i-} be the counterfactual investment income in the same bin. The elasticity can then be calculated using the following formula:

$$\hat{\varepsilon} = \frac{\sum_{i=1}^{n_b} ln\left(\frac{X^*_{i-}}{X^*_{i+}}\right)/n_b}{\sum_{j=1}^{n_c} ln\left(\frac{1-\tau_{jl}}{1-\tau_{jh}}\right)/n_c}, \tag{2}$$

where n_b is the number of bins in the region $[-\delta_b, \delta_b]$ and n_c is the number of individuals in the region $[-\delta_b, \delta_b]$.

Note that this elasticity estimate is also a lower bound on the underlying average elasticity (assuming investment income is increasing, on average, in AGI) because investment income is endogenous with respect to AGI; that is, when investment income declines, so does AGI, pushing down the estimated response at a given AGI level. However, a positive response will be found as long as investment income is not the sole source of growth in AGI. Moreover, like the density method used above, this method implicitly assumes that there are no other factors that need to be held fixed in order to obtain a consistent estimate, which makes it possible to obtain a non-causal relationship. For example, suppose earned income decreases in response to the tax rate, but only for those with low investment income. This would lead to a spurious response in investment income. This will be partly ruled out in the discussion that follows; still, these are informative results, meant to provide evidence of the apparent response before proceeding to the next section, which analyzes the investment

[31] When the former method is used, the results are qualitatively similar, but the standard errors are substantially larger.

income response in a more precise, causal way.

The elasticity estimates from equation (2) are given in Table 3 Columns (6)-(9). The standard errors are obtained by a nonparametric bootstrap method (500 replications). I exclude those with unemployment insurance income from the analysis because a decline in unemployment insurance should be accompanied by an increase in earned income; this method is not designed to address this type of mechanical correlation. The first three columns examine the wage-earning sample. Columns (6) and (7) report the results for the two components of AGI—earned and unearned income—respectively. Figure 8 is the corresponding figure. Both the figure and the elasticity estimate for earned income (-0.015) point to the following fact: earned income grew at a faster rate (although not statistically significant) above the tax kink relative to below. This occurs mechanically if other types of income included in AGI are growing slower in response to the higher tax rate and earned income is not responding, or at least responding in a very small way, relative to the other types of income that do respond (the mechanical response is small because earned income is a much larger share of AGI than unearned income). The magnitude of the earned income response is consistent with it being entirely mechanical. In contrast, Figure 8 shows a decline in unearned income above the tax kink. The corresponding elasticity is 0.516, but is statistically insignificant, likely reflecting the overall volatility of unearned income and also that some types of unearned income are expected to be responsive, while others are not expected to be responsive.

Table 3 Column (8) and the top panel of Figure 9 examine one type of unearned income—investment income.[32] Although capital gains or losses are a lumpy realization of savings, I include these in this analysis of investment income, because most individuals that have capital gains also have dividend or interest income. If capital gains are excluded, and individuals respond to the higher marginal tax rate above the tax kink by decreasing capital gains, this

[32]The sample is restricted to those for whom the sum of dividend, interest, positive capital gains, and tax exempt interest income is under $2,300 in all years, which is $500 below the investment income asset limit in 1996-2006; this corresponds to my sample restriction in the next section, and is imposed here for comparability. The results are similar when it is not imposed.

will likely bias the estimates downwards. Both the figure and the elasticity provide evidence that investment income declined substantially above the tax kink—the elasticity estimate is 1.879 and is statistically significant at the 5 percent level.

In Table 3 Column (9) and the second panel of Figure 9, I consider whether self-employment income for self-employed individuals declined significantly in the region around the kink. Both the figure and the elasticity estimate suggest that it did; however, the estimates are not significant.

This section has found a significant response to the phase-out region tax rate among both wage-earners and self-employed individuals, and for the former, it is concentrated among non-labor income, particularly investment income. The elasticity of investment income with respect to the marginal net-of-tax rate is 1.88 and can be interpreted as a lower bound on the true elasticity.

5 Difference-in-Differences Estimates

This section focuses specifically on the responsiveness of investment income to the EITC phase-out marginal tax rate, exploiting differential variation in the phase-out tax rate between one- and two-dependent families. The evidence is consistent with the descriptive evidence from the last section, pointing to a significant decline in investment income in response to changes in the EITC schedule. There is no evidence that bias-inducing selection or tax evasion are driving the results, although both are considered.

The results from the last section inform the method used here. In particular, I do not rely on a comparison of individuals above and below the tax kink because bunching occurs at the tax kink, which generates selection into a particular observed tax rate and thus prevents a causal parameter from being obtained. Weber (2011) shows that picking an instrument to resolve this type of selection is difficult, and no obvious instrument exists in this context. To the extent that imperfect bunching is due to income volatility, selection into a particular tax

bracket may be mitigated, but this creates a new problem—the parameters of the treatment are not known with certainty in advance, making it difficult for individuals to respond.

I proceed by defining a region based on earned income within the phase-out region and identify the parameters of interest by relying on differential variation in the marginal tax rates between families with one and two children. Defining the region in this way will not necessarily yield consistent estimates if earned income responds to the tax rate. The evidence in the previous section is consistent with the assumption that earned income does not respond to the tax rate for wage-earners, but not for self-employed individuals. Therefore, I restrict the sample to wage-earning individuals. Given the results in Section 4, a response in earned income would necessarily be small, but I cannot completely rule it out; therefore, later in this section, I will discuss the conditions under which such a response would induce a bias in the estimates, and provide evidence that this form of selection does not appear to be having a significant impact on the estimates.

The identifying variation in this section comes from the differential variation in the phase-out EITC rate between one- and two-child families over time. This variation is plotted in Figure 10. There, I plot the marginal net-of-tax rate, which is one minus the marginal tax rate; this is the variable that will be used in the main estimating equation. Given the form of identifying variation, the specifications in this section assume that individuals do not make endogenous fertility decisions or decisions to claim an extra dependent in a way that is correlated with their investment income response. These types of exogeneity condition are assumed in much of the EITC literature, as most of it relies on differential variation in tax rates across the number of children for identification. I consider this story later in this section and find no evidence that the estimates are biased by individuals selecting the number of children they have in a way that is correlated with their outcomes.

Before proceeding to the main estimation, I provide graphical evidence of the response in investment income over time by number of children, and estimate a binary treatment (i.e. did an individual have *any* interest or dividend income) difference-in-differences equation.

For this analysis, I split the years into two groups, pre and post. Pre includes all years through the end of the EITC expansion (1988-1995), and post includes all the years after the expansion (1996-2006). I estimate the following difference-in-differences intent-to-treat (ITT) specification for wage-earning individuals with earned income in the range \$18,000-\$30,000:

$$INV_{it} = \alpha + \gamma post * twokids + \beta_1 twokids_{it} + \beta_2 X_{it} + \eta_t + \nu_{it}, \tag{3}$$

where γ is the parameter of interest, INV_{it} is an indicator for whether an individual has positive interest or dividend income, *twokids* is an indicator for whether the individual has more than one child, *post* is an indicator for the observation being in 1996-2006, X_{it} are additional covariates, and η_t are year fixed effects.[33] The vector X_{it} includes earnings, earnings squared, married, earnings x *twokids*, earnings squared x *twokids*, and married x *twokids*. I pool data over all years in the sample. All income values are in 2006 dollars.[34] I choose \$18,000 as the lower cutoff because it is at least \$1,000 above the second kink point in all years; thus, the estimates will likely not capture bunching decisions around the kink point. It is above the filing threshold for each type of marital status and above the positive tax rate threshold for all single and head of household taxpayers, but not for married individuals. For them, the positive tax rate is implemented part way through the window examined in this section. However, given that families with one dependent face this positive tax rate first, this will bias the estimates downward, because it will make one-dependent families appear more responsive relative to two-child households. Earned income of \$30,000 was chosen as the baseline upper bound, because it is at least \$1,000 below the end of the phase-out region for all years. The baseline estimates are robust to the choice of upper and lower bound; I will analyze the sensitivity of the main estimates to the bounds in Table 6.

[33] A separate indicator for *post* is not included. It is not needed because year-fixed effects are included.

[34] I use the deflator used by the IRS, which is the average CPI-U from September in year $t-2$ through August in year $t-1$.

I drop individuals with more than $2,300 of investment income (the EITC definition of investment income is dividend, interest, tax exempt interest and positive capital gain income) in any year. This decision is driven by the fact that there was an asset limit imposed in 1996, excluding those with more than $2,800 dollars of investment income in 2006 dollars from claiming the EITC.[35] Before 1996, 1.1 percent of two-child families held investment income over this limit and were 10 percent more likely to have investment income over the limit than one-child families. Two-child families are also more apt to respond to the limit by decreasing their investment income below the threshold because they will lose more in benefits than an equivalent one-child household. Consequently, estimates that include individuals with investment income over the limit or impose the limit only after 1996 will be biased upward. I impose a limit of $500 less than the asset limit to avoid differential rates of selection by number of children into having investment income just under the cutoff.

Table 4 provides descriptive statistics for the baseline specification. Average interest and dividend income reported, including all zeros, is $48.60. About 20 percent of individuals have positive interest and dividend income. Almost one-third of the sample is married, and the average number of children is 1.4. Fifty-nine percent use a paid preparer.

The estimated γ from equation (3) is -0.036 and the p-value is 0.000. The main estimation below defines treatment as $log(1-\tau)$. Rescaling the estimated γ using two-stage least squares so that it is directly comparable to the main estimates, yields an estimate of 0.789 with a p-value of 0.000. This estimate implies that, for a one percent decrease in the marginal net-of-tax rate, the likelihood an individual opens an interest- or dividend-bearing account declines by 0.008. For example, if the saving rate was 0.4 before a 10 percent decline in the marginal net-of-tax rate, it would be 0.32 afterwards.

Figure 11 provides graphical evidence of the behavioral response to changes in the EITC phase-out tax rate of the choice to hold income in interest- or dividend-bearing accounts over time using the same sample restrictions. The estimated line comes from a cubic in

[35]There are a few special circumstances when this limit does not hold.

year interacted with post and number of children.[36] In the early years, when there was no (or a small) difference between the tax rates for one- and two-child households, their investment income was not statistically distinguishable, suggesting that one-child families provide a good comparison group for two-child families. It takes a year or two for individuals to adjust to the new tax rates, which is consistent with other EITC literature (Saez, 2010). In 1996, there is a statistically significant drop in the likelihood that two-child families have investment income and this difference remains statistically different from that for one-child families through 2006. This graphical evidence is consistent with the estimates obtained from the simple difference-in-differences estimates; the gap between the likelihood that taxpayers with one and two children have investment income is about 0.03 in the post period, and this increase in the gap is stable over time. This suggests that the additional covariates included in equation (3) do not have a substantial effect on the estimates.

The main specification in this section takes advantage of all the variation in $log(1-\tau)$ by including it directly and estimating the following intent-to-treat (ITT) specification:

$$INV_{it} = \alpha + \gamma log(1-\tau_{it}) + \beta_1 twokids_{it} + \beta_2 X_{it} + \eta_t + \nu_{it}, \tag{4}$$

where the covariates and sample restrictions are the same as for equation (3). Note that one may wish to estimate the response of investment income to the after-tax rate of return $log((1-\tau_{it})r_t)$, where r_t is the average annual real interest rate. Because $log((1-\tau_{it})r_t) = log((1-\tau_{it})) + log(r_t)$ and there are year fixed-effects in this model, the estimates obtained from this specification are equivalent. The same argument applies for including the state EITC rates when state-year fixed effects are included (except for Wisconsin).[37]

[36]Note that another possible type of selection could occur if a substantial number of new families filed for the EITC after the phase-out rate increased (which coincided with an increase in EITC generosity) and these new filers were much less likely to have investment income. I cannot completely rule this out, but it is worth noting that while EITC claims were rising over this period, there was no discrete jump at 1996, nor is the pattern consistent with the pattern in Figure 11. Additionally, adding a control for the total number of EITC claimants by year in equation (3) has a minimal effect on the estimates.

[37]It does not apply to Wisconsin because a larger percentage of the federal credit is given to families with two children.

The estimates given by equation (4) produce ITT estimates because all individuals are included as long as they have earned income in the specified range, regardless of whether they actually decide to claim the EITC, or are even eligible based on their overall AGI level. It also ignores other tax rates and credits faced. I choose not to report treatment on the treated (TOT) estimates to address these issues, because the value of such estimates is limited. The estimates are quite similar in practice (only four percent of my baseline sample was not EITC-eligible) and, in this context, given the nature of the selection and its correlation with the ITT measure, Weber (2011) shows that the estimate will be biased.

In Tables 5, 6, and 7, the top rows of each column estimate the extensive margin effect—the dependent variable is an indicator for whether the individual receives any interest or dividend income. These estimates are the mean marginal effects from a probit specification. The bottom rows estimate the elasticity of dividend and interest income with respect to the marginal net-of-tax rate. These estimates come from a two-tiered model proposed by Cragg (1971).[38] The first tier estimates the extensive margin decision and is estimated using the same probit model as is used in the top rows. The second tier estimates the intensive margin effect (i.e. conditional on choosing to save a positive amount, how much does the individual save) using a truncated normal regression, where the dependent variable is the log of interest and dividend income. Taking the log of interest and dividend income is attractive because the specification directly produces an elasticity, and it is necessary because the unscaled dependent variable has a large left skew that, if left untransformed, would violate the assumptions of the estimation method. The reported coefficients are mean marginal effects from this two-tier model. Standard errors were calculated using the delta method. I do not report the second tier; however it is worth noting that the second tier coefficient on $log(1-\tau_{it})$ is never statistically significant. This suggests that the significant response occurs along the extensive margin and much of what the elasticities are capturing is the average amount by which dividend and interest income changes when individuals choose to open, or

[38]This type of model is also referred to as a hurdle model.

close, an interest- or dividend-bearing account.

Table 5 presents some results. Baseline estimates are given in Column (1). The extensive margin estimate is 0.612, which means that a one percent increase in the marginal net-of-tax rate increases the probability of having interest or dividend income by 0.006. The phase-out marginal net-of-tax rate decreased by 9.6 percent over the last two decades, which implies a 0.058 decrease in the likelihood that individuals have interest or dividend income. The overall decrease in the likelihood that individuals hold interest or dividend income for EITC recipients in this population is 0.134; hence, among those examined in this section, 43.5 percent of the decline in the fraction of EITC recipients with savings in income-bearing accounts over the last two decades in the can be explained by changing EITC incentives. Overall, the extensive margin estimate implies that 39.4 percent of the decline in the fraction of EITC recipients with savings in income-bearing accounts over the last two decades can be explained by changing EITC incentives.[39]

The elasticity estimate is 3.047, which implies that a one percent increase in the marginal net-of-tax rate increases interest and dividend income by 3.047 percent. The elasticity estimate implies that for each one percent increase in the marginal net-of-tax rate, dividend and interest income increases by $1.48 at the mean. Suppose that the interest rate is 5 percent. Then, if individuals spend their income (as opposed to saving it, but not placing it in an interest- or dividend-bearing account), this implies an increase of $29.62 at the mean in the annual amount in savings. Among those that have positive savings in 1988-1990, this implies a $185.82 increase in savings. Because the marginal net-of-tax rate decreased by 9.6 percent over this period, these estimates imply that the annual amount in savings, among those with positive saving in 1988-1990, would have been $1,775.82 higher in 2006 if the disincentives to save had remained at their 1988 level.

Table 5 Column (2) examines a broader measure of investment income, including capital

[39]These estimates still apply only to those with one or two dependents, who were wage-earners with investment income less than $2,300.

gains or losses, partnership income, other gains or losses, and IRA contributions;[40] this has a minimal effect on the estimates. I exclude capital gains in the baseline estimates because it is a lumpy realization of savings, rather than the annual amount. Column (3) includes the controls X in a more flexible way, by interacting each control with an indicator variable for the observation occurring before 1992 and another indicator variable for the observation occurring after 1998. The year 1992 is two years before the largest expansions in the gap between one- and two-child families, and 1998 is two years after the last change in the tax rate. This specification allows for different trends in investment income in the years before 1992 and after 1998. The estimates increase slightly and the additional controls have a minimal effect on the standard errors. Column (4) adds state fixed-effects to control for differences in state EITC generosity.

Table 5 Columns (5)-(7) examine heterogeneity in the baseline estimates by taxpayer type. Column (5) estimates the response separately for unmarried (single and head of household) individuals and married individuals; this estimate is perhaps not too interesting on its own, but the lack of difference between these two types lends credibility to the identification used in Subsection 5.1. Married individuals are slightly less responsive than unmarried individuals and the difference in responsiveness is highly insignificant.

One might expect that those who temporarily face the EITC schedule, may be less responsive, since savings is often a longer-term decision, and awareness of detailed features of the schedule may be low. On the other hand, if an individual is a temporary EITC recipient because of an unemployment spell, their response could be more elastic, as it may be an appealing time to draw down their savings. I use an indicator for having unemployment insurance as a proxy for temporarily facing the EITC schedule due to a negative shock in earned income; it is of course not a perfect proxy as an individual could be in the EITC range both with and without becoming unemployed for part of the year. The results are reported

[40]Note that all negative values are truncated at zero and estimated using the same two-tier model above to take this truncation into account.

in Table 5 Column (6).[41] The results suggest that, on average, those with unemployment insurance receipts do not respond differently; the estimates are not statistically different.

Table 5 Column (7) considers how the response differs among those that do and do not use a paid preparer.[42] Those who choose to have a paid preparer may be more responsive to begin with, and the paid preparer may educate them about the costs and benefits associated with having investment income when in the phase-out region. With regards to the latter, using a paid preparer in the current year can be interpreted as a proxy for having a paid preparer in the past, because learning in the current year is not helpful unless the tax preparer encourages individuals to engage in tax evasion. Both those with and without a paid preparer respond, but the response of those who use a paid preparer is twice as large. This supports the hypothesis that paid preparers increase individuals' understanding of the EITC schedule.

Table 6 considers two possible forms of selection discussed above, which could be biasing the estimates. Column (1) repeats Column (4) in Table 5. The first type of selection occurs when individuals select into or out of the sample by adjusting their earned income in response to the EITC phase-out tax rate. If individuals are moving in and out of the sample and this selection is correlated with the likelihood they respond along the investment income margin (or have investment income), then the estimates may be biased. The estimates remain unbiased only if this selection is occurring at a constant rate—that is, entry and exit are equal. Assuming a positive correlation among earned and investment income responses, the estimates will be biased downwards (upwards) if exit (entry) is larger. This occurs because those who are entering or exiting will adjust their investment income more to a given change in the marginal tax rate than individuals whose earned income does not respond to the tax change. When exit (entry) is larger, my estimates will exclude (include) more individuals

[41] Unemployment insurance indicators and a complete set of interactions with the other covariates are included in the estimating equation to control for underlying differences in investment income among those with unemployment insurance.

[42] Paid preparer usage and a complete set of interactions with paid preparer are included in the estimating equation to control for underlying differences in investment income among those with a paid preparer.

that are highly responsive. The last section found no evidence of a response in earned income; however, a small response among certain taxpayers could exist. If responsiveness changes monotonically with earned income, the entry and exit rates will converge as the estimation window narrows, making the bias decrease. Moreover, if I divide the sample in half and estimate the response over \$18,000-\$24,000 and \$24,000-\$30,000, both estimates should experience a reduction in bias, thus moving the estimates in the same direction. Table 6 Columns (2) and (3) split the data in this way and find that the estimates are not statistically different from one another and that the coefficients move in opposite directions relative to the baseline, providing no support for this type of bias-inducing selection.

Another potential obstacle to obtaining consistent estimates in this section is that individuals can choose the number of children they wish to have (or to claim on their tax return). Empirically, the literature suggests that these responses are relatively small (Baughman and Dickert-Conlin, 2003; McCubbin, 2000). This will bias the estimates upwards if the correlation between choosing to have a second child to obtain a higher EITC amount and choosing to have lower dividend and interest income is greater than the same correlation for choosing to have a first child. The incentives for having a first child are higher than those for having a second for all years in the sample, but the incentives for claiming a second child have increased over time more than those for having a first child, and these changes are correlated with changes in the phase-out tax rate, which would bias the estimates upwards. While there is a strong incentive to claim a second child, there is no incentive to claim a third child, except in the state of Wisconsin.[43] Table 6 Column (4) drops two-child families and instead includes three-child families, effectively removing the selection problem for two children, but leaving the selection for one child; now, any selection for one-child families that is correlated with the phase-out tax rate will bias the estimates downwards. There are about four times fewer three-child than two-child families, which increases the standard errors, all else equal. While the estimates are no longer statistically significant, the point estimates are

[43]Wisconsin provides more EITC to those with three children.

close to the baseline estimates. These results provide no evidence that selection into having a second child plays a significant role in the baseline estimates.

While the existence of tax evasion could not bias the estimates, it would substantially change the interpretation and welfare implications of the results. In general, it is appropriate to assume that the elasticity of the tax evasion with respect to the marginal net-of-tax rate is a function of the costs associated with engaging in tax evasion.[44] In this context, this implies that there is more evasion below the information reporting threshold at $10 and evasion will increase as the marginal tax rate increases. Figure 12 provides evidence on the amount of interest income and dividend income evasion below the information reporting line. The estimates are based on the density test proposed by McCrary (2008), which relies on the same bin size and bandwidth selection as described in Section 4, except that the bandwidth is chosen as the average of the bandwidth chosen below the information reporting notch and above the notch. In order to make a fair comparison, I choose the bandwidth when the data is restricted to $10 on either side of the tax notch (otherwise, the bandwidth would be far too large to obtain a reasonable estimate below the information reporting notch). The analysis is done in nominal dollars.[45]

The top panel of Figure 12 examines interest income over years 1988-2006. There is a 120 percent increase in the density just above the information reporting threshold and this estimate is statistically significant at the 1 percent level. This suggests that there is substantial evasion in interest income below $10, or alternatively individuals often forget to report their interest income unless they receive a tax form (1099-INT) at the end of the year with the amount earned, which only occurs if there was information reporting. If the elasticity with respect to the tax rate is positive and increasing with a decrease in the cost,

[44] This point comes up frequently in the elasticity of taxable income literature (e.g., Saez et al., 2010).

[45] The reason is that interest and dividend income is rounded to the nearest dollar on the tax form, so if I were to renormalize the variable to 0 at the information reporting notch and then convert the variable to real dollars, there are a large number at 0, which is just above the notch, and then a gap until plus or minus one. Moreover, the number now in minus 2 to minus 1 is lower because the inflation rate scales many of them out, but does not have the same effect at 0. Therefore, the estimates when using real dollars substantially overstate the true effect.

the jump should increase after 1996, relative to the small EITC years (1988-1993); but neither of these separate estimates is statistically different from the estimate for all years, and the estimate for the years after 1996 is lower, not higher. This suggests that evasion occurs, but that it does not vary systematically with the tax rate. The bottom panel considers dividend income; however, the amount of dividend income in this range is about five times less, so the results are not statistically significant; in fact, there does not appear to be any evidence of evasion on this margin.

As an additional check on evasion below $10 in particular, I replace all reports less than or equal to $10 with zeros for all years. This forces the response of all with income less than $10 to be zero, which eliminates any evasion response of these individuals as well as any other response. These estimates are displayed in Column (5) of Table 6. They are slightly lower, but not statistically different from the baseline estimates. The decrease could be due to eliminating this form of tax evasion, but may also be due to the fact that I have eliminated the possibility of response for all those with dividend or interest income under $10.

5.1 Separating the Income and Substitution Effects

The estimation conducted up to this point has estimated a single parameter capturing the response to the marginal tax rate. This parameter encompasses both a substitution and an income effect. In particular, an increase in the phase-out tax rate has two effects: 1) the marginal cost of holding a dollar in an interest- or dividend-bearing account increases (substitution effect), and 2) the individuals' wealth changes because the EITC amount changes (income effect). Wealth increases if the location of the end of the phase-out region is held fixed. Wealth decreases if the location of the beginning of the phase-out region is held fixed. This section adds a control for the income effect, so that the substitution effect—the main parameter of interest in this paper—is not biased. I find that the estimates in previous section were biased downwards, although the difference is not statistically significant.

The sign of the income effect on saving is ambiguous. As the EITC increases, individuals

have more income, which may lead them to save more. As discussed in Subsection 2.2, a key mechanism considered in the context of other need-based programs is that these programs decrease the need for precautionary savings. This is true in the context of the EITC as well because, for individuals in the phase-out region, a negative shock in income is met with an increase in the amount of EITC received. This suggests that the sign should be negative.

To estimate the two effects separately, I add a log function of after-tax earned income ($earnings + \tau(end - earnings)$), where end is the end of the phase-out region and τ is the EITC phase-out marginal tax rate.[46] Note that actual after-tax income is defined as ($agi+\tau(end-agi)$) (assuming $agi > earnings$) and τ will change if individuals are not located in the phase-out region given their AGI. In this specification, individuals that decrease their investment income by a larger percentage in response to the tax rate change would receive a smaller income treatment by definition because AGI includes investment income. This would create a nontrivial bias in the estimates. Note that the existing literature that separately estimates the substitution and income effects (e.g. Gruber and Saez, 2002) estimates the income effect in this way, and therefore the results from this literature are biased. Even using earnings in the definition of the proxy for the EITC benefit amount requires a stronger assumption on earnings than before—earnings cannot respond to the marginal tax rate. If earnings respond and the response is positively correlated with the investment income response, this will create a downwards bias in the income effect estimates, but I find no evidence that earnings responds in this paper. Even if earnings responds, this bias is, by definition, smaller than the bias induced by using AGI, because earnings are one component of AGI. The proxy precisely captures the change in the EITC benefit received for a given level of earnings, as long as AGI is not outside the phase-out region. The income effect term can therefore be interpreted as an ITT estimate.

Empirically, the identification of the income effect separately from the substitution effect comes from two sources. One source is variation in end, which is the end of the phase-out

[46] Gruber and Saez (2002) show that the income effect measure should take this form.

region. The variable *end* increased for two-child taxpayers in 1996 and married taxpayers in 2002 and 2005. For this analysis to be valid, I must assume that marital status does not change endogenously with the changes in *end*. If more responsive individuals are more likely to get married once *end* increases, this will bias the income effect estimates upwards. The personal income tax schedule also changed in 2002 and the change affected married, head of household, and single individuals differently. Prior to 2002, all of the individuals who faced a positive personal income tax rate in the EITC sample examined here faced a 15 percent income tax. In 2002, a new 10 percent bracket was created. To ensure that all individuals experienced the same tax rate change in 2002 (so there is no differential effect of this tax change which might be attributed to a change in *end*), I drop all single individuals.[47] The income effect will also be partly identified from variation in τ; changes in τ matter for both the substitution and income effect terms, but it enters each in a different way. Note that without including flexible controls in earnings, identification of the income effect would also come from variation in earnings.[48] However, investment income likely increases with earnings, regardless of the tax rate, making this variation invalid.

The results are presented in Table 7. Columns (1) and (2) replicate the baseline specification given in Columns (1) and (4) in Table 5 with the restricted sample used in this subsection. Columns (3) and (4) add the income effect term to this specification. There is not enough variation to separately identify the income effect at a statistically significant level. The substitution effect gets slightly larger when controlling for the income effect, but the difference is not statistically significant. When the income effect is positive and it is ignored in the estimation, the estimates should be biased downwards because an increase in the phase-out tax rate usually coincided with an increase in EITC benefits in the period examined.

[47]Most unmarried individuals with children file as a head of household, so this does not drop many individuals from the sample.

[48]A quadratic in earnings is used in the paper; however the results are not sensitive to including a cubic or quartic polynomial in earnings instead.

6 Conclusion

When designing the EITC, non-labor income was likely included as a determinant of the amount of EITC received to ensure the credit was going to those that were low-wealth. This paper finds that an unintended consequence of this provision is to substantially distort non-labor income, particularly investment income; in fact, along the intensive margin, the non-labor income distortions are far larger than the earned income distortions for wage-earning individuals. Such behavior induces deadweight loss because engaging in such behavior is not costless for these individuals regardless of whether a decline in investment income translates into a decline in savings. To the extent that the behavioral response estimates in this paper reflect an actual decline in savings, this creates additional concerns, particularly since the government would like to encourage low-income households to save more, as evidenced by programs like the Saver's Credit and Individual Development Accounts.

This paper provides evidence that the response in investment income is substantial. Indeed, nearly 40 percent of the decline in saving in income-bearing accounts by EITC recipients over the last two decades can be explained by the changing incentives for saving caused by the EITC schedule. The response is twice as large among those that use a paid preparer, consistent with an increased awareness of the relevant incentives among this group.

The policy implications for EITC design depend on how policymakers weight the deadweight loss induced by this provision, relative to its original intent—to exclude individuals that were not actually poor from receiving the EITC. Making the EITC amount exclusively a function of earned income would eliminate the distortion found in this paper, but would increase the number of claimants by 4 percent; these additional claimants would receive an average of $1,014 in benefits, even though their average AGI is $52,179.

References

Baughman, R. and Dickert-Conlin, S. (2003). Did expanding the EITC promote motherhood? *The American Economic Review, Papers and Proceedings*, 93(2):247–251.

Blumenthal, M., Erard, B., and Ho, C.-C. (2005). Participation and compliance with the Earned Income Tax Credit. *National Tax Journal*, 58(2):189–213.

Card, D., Lee, D. S., and Pei, Z. (2009). Quasi-experimental identification and estimation in the regression kink design. Princeton University, Industrial Relations Section Working Paper No. 553.

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. *Quarterly Journal of Economics*, 126(2):749–804.

Cragg, J. G. (1971). Some statistical models for limited dependent variables with applications to the demand for durable goods. *Econometrica*, 39(5):829–844.

Dahl, G. B. and Lochner, L. (2010). The impact of family income on child achievement: Evidence from the Earned Income Tax Credit. NBER Working Paper No. 14599.

Duflo, E., Gale, W., Liebman, J., Orzag, P., and Saez, E. (2006). Saving incentives for low- and middle-income families: Evidence from a field experiment with H&R Block. *Quarterly Journal of Economics*, 121(4):1311–1346.

Fan, J. and Gijbels, I. (1996). *Local Polynomial Modelling and its Applications*. Chapman & Hall, New York.

Gelber, A. M. and Mitchell, J. W. (2011). Taxes and time allocation: Evidence from single women and men. *Review of Economic Studies*, forthcoming.

Gruber, J. and Saez, E. (2002). The elasticity of taxable income: Evidence and implications. *Journal of Public Economics*, 84(1):1–32.

Gruber, J. and Yelowitz, A. (1999). Public health insurance and private savings. *Journal of Political Economy*, 107(6, pt. 1):1249–1274.

Hotz, V. J. and Scholz, J. K. (2003). The Earned Income Tax Credit. In Moffitt, R., editor, *Means-Tested Transfer Programs in the United States*. University of Chicago Press and NBER, Chicago.

Hubbard, R. G., Skinner, J., and Zeldes, S. P. (1995). Precautionary saving and social insurance. *Journal of Political Economy*, 103(2):360–399.

Hurst, E. and Ziliak, J. P. (2006). Do welfare asset limits affect household saving? *The Journal of Human Resources*, 41(1):46–71.

Joulfaian, D. and Rider, M. (1996). Tax evasion in the presence of negative income tax rates. *National Tax Journal*, 49(4):553–572.

Kleven, H. J. and Waseem, M. (2011). Tax notches in Pakistan: Tax evasion, real responses, and income shifting. http://personal.lse.ac.uk/kleven/Downloads/MyPapers/workingPapers/kleven-waseem_may2011.pdf.

LaLumia, S. (2009). The earned income tax credit and reported self-employment income. *National Tax Journal*, 62(2):191–207.

Liebman, J. B. (1998). The impact of the earned income tax credit on incentives and income distribution. In Poterba, J. M., editor, *Tax policy and the economy, Volume 12*, pages 83–120. MIT Press, Massachusetts.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714.

McCubbin, J. (2000). EITC noncompliance: The misreporting of children and the size of the EITC. *National Tax Journal*, 53(4, part 2):1135–1164.

Mills, G., Gale, W. G., Patterson, R., Engelhardt, G. V., Eriksen, M. D., and Apostolov, E. (2008). Effects of individual development accounts on asset purchases and saving behavior: Evidence from a controlled experiment. *Journal of Public Economics*, 92(5-6):1509–1530.

Ramnath, S. P. (2010). Taxpayers' response to notches: Evidence from the Saver's Credit. http://sitemaker.umich.edu/ramnath/files/saversresponse.pdf.

Romich, J. L. and Weisner, T. (2000). How familes view and use the EITC: Advance payments versus lump sum delivery. *National Tax Journal*, 53(4):1245–1265.

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

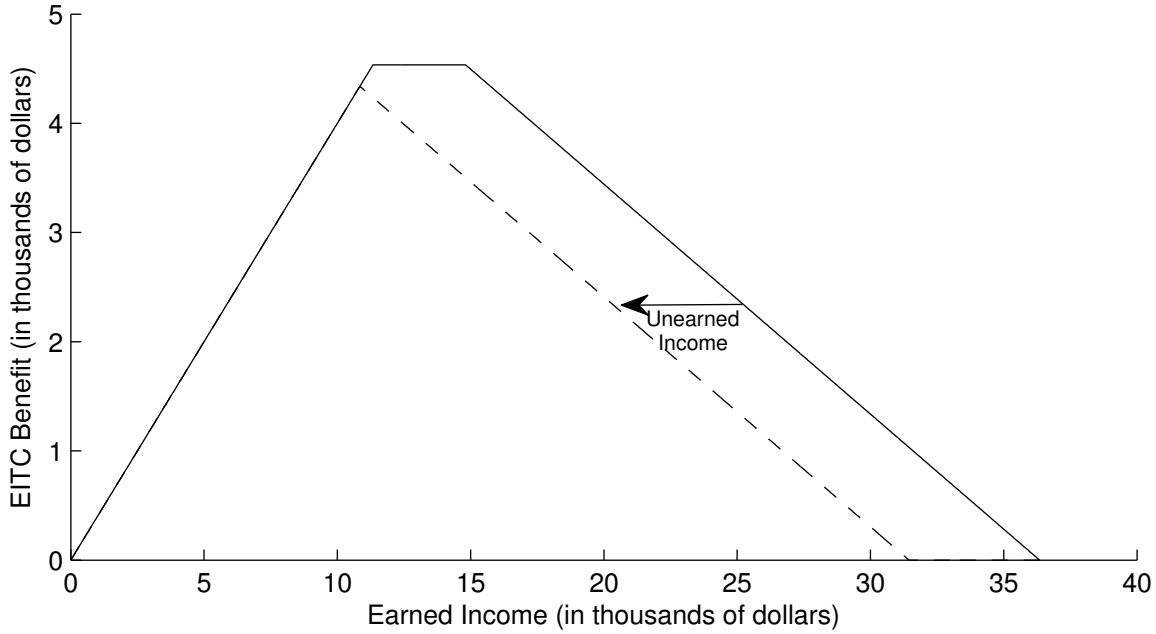
Saez, E., Slemrod, J., and Giertz, S. (2010). The elasticity of taxable incomes with respect to marginal tax rates: A critical review. *Journal of Economic Literature*, forthcoming.

Sullivan, J. X. (2006). Welfare reform, saving, and vehicle ownership: Do asset limits and vehicle exemptions matter? *The Journal of Human Resources*, 41(1):72–105.

Weber, C. (2011). Identifying the causal effect of a marginal tax rate change when there are multiple tax brackets. In process.

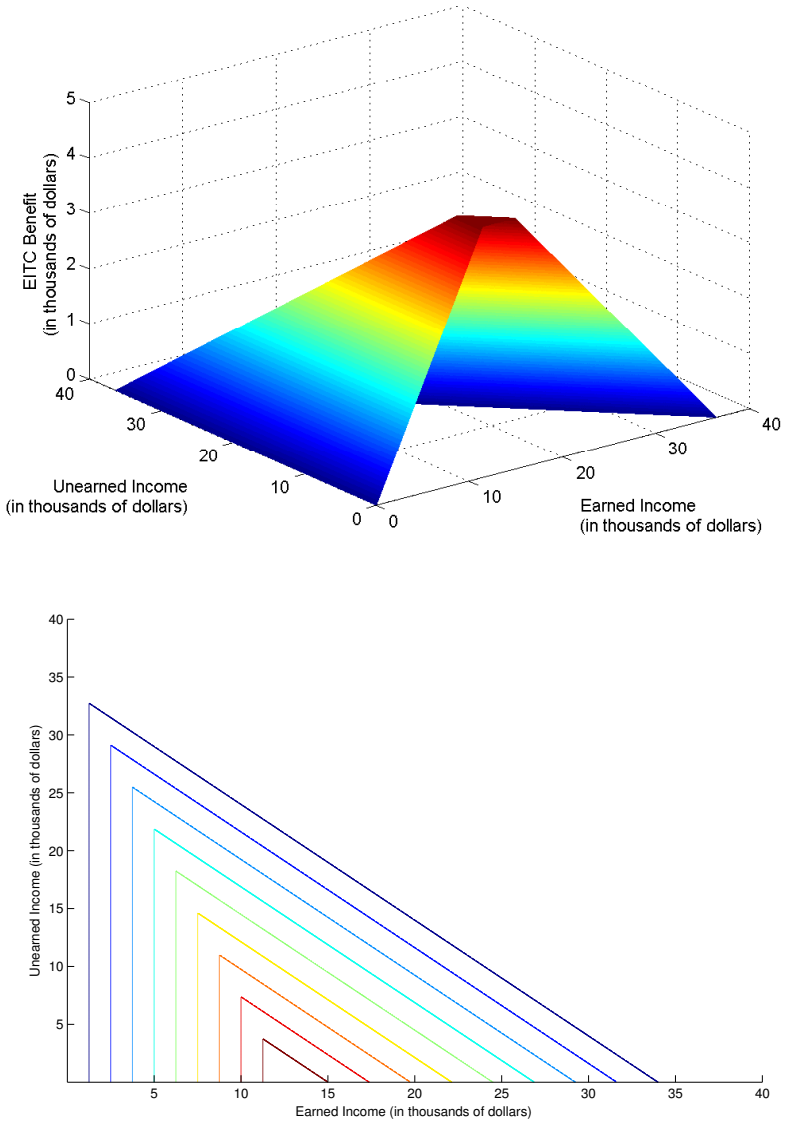
Weber, M. (2001). General description booklet for the 1996 public use tax file. Written under the direction of Peter Sailer, Chief, Special Projects Section, Individual Statistics Branch, Statistics of Income Division, Internal Revenue Service.

Figure 1: EITC Schedule



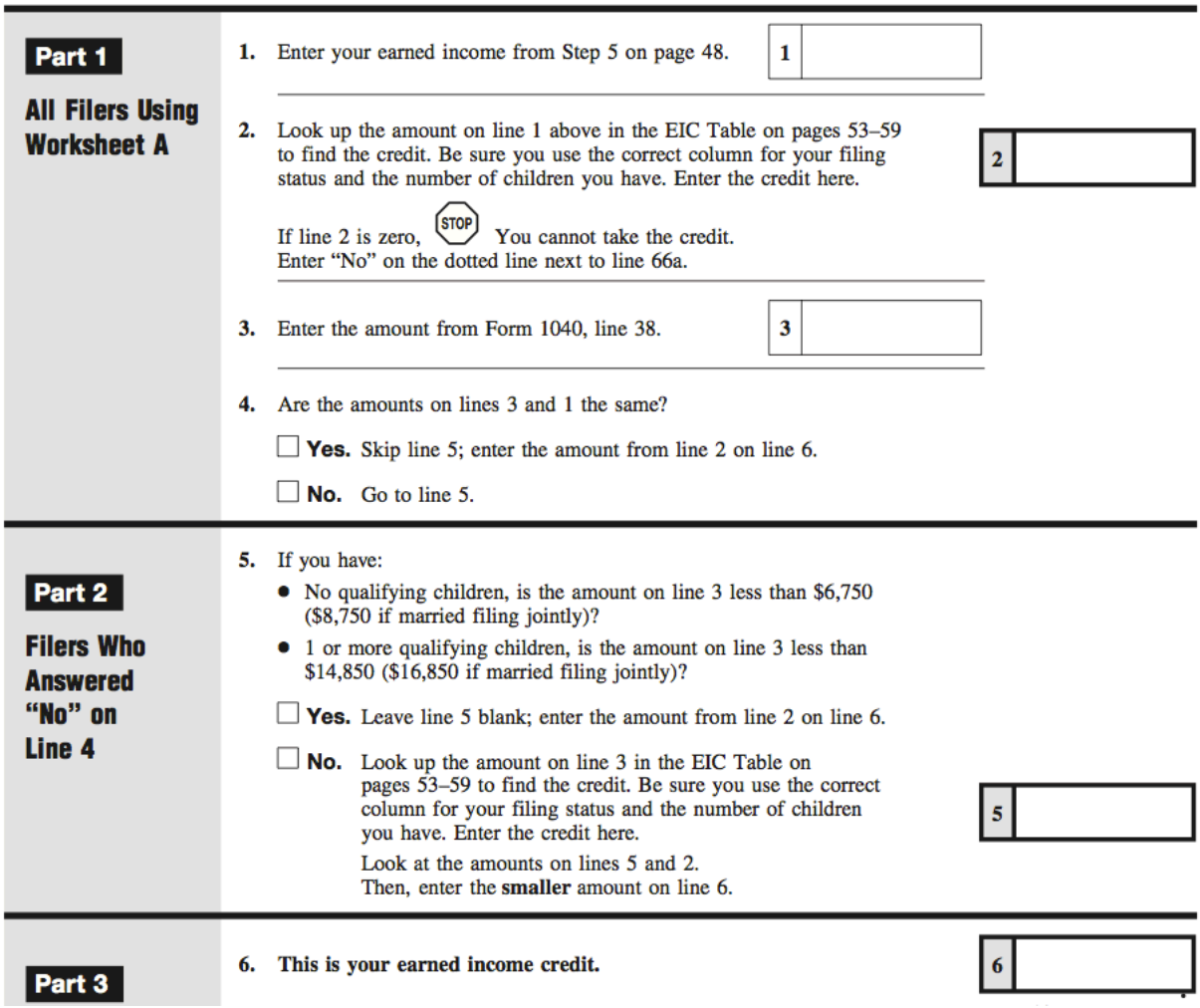
This figure displays the EITC schedule in earned income for taxpayers with two dependents in 1996-2006. 75 percent of these taxpayers have unearned income less than or equal to zero and thus face the EITC schedule given by the solid line. 25 percent have positive unearned income and face an EITC schedule, such that the phase-out line is shifted in towards the dashed line. For about a quarter of these individuals, the line is shifted far enough that the plateau region is eliminated; they face a spike in their EITC schedule.

Figure 2: EITC Schedule with Unearned Income



The top figure plots the entire 3-dimensional surface in earned income, unearned income, and the EITC amount for taxpayers with two dependents in 1996-2006. The bottom figure provides contour lines of the top figure. When earned income is low, EITC benefits remain low, regardless of unearned income levels. This is because, when earned plus unearned income (AGI) is large enough to reach the second kink, individuals are still forced to take the benefit based on their earned income because it is lower than their AGI benefit. As earned income rises, the EITC first rises and then falls, either with or without a plateau region depending on the amount of unearned income, as was depicted in FIgure 1. As unearned income rises, it first has no effect on the EITC and then causes it to decline at a certain threshold, where the threshold varies depending on the individual's earned income amount.

Figure 3: EITC Calculation Form in 2006



Part 1

All Filers Using Worksheet A

1. Enter your earned income from Step 5 on page 48. **1**

2. Look up the amount on line 1 above in the EIC Table on pages 53–59 to find the credit. Be sure you use the correct column for your filing status and the number of children you have. Enter the credit here. **2**

   If line 2 is zero, STOP You cannot take the credit. Enter "No" on the dotted line next to line 66a.

3. Enter the amount from Form 1040, line 38. **3**

4. Are the amounts on lines 3 and 1 the same?

   ☐ **Yes.** Skip line 5; enter the amount from line 2 on line 6.

   ☐ **No.** Go to line 5.

Part 2

Filers Who Answered "No" on Line 4

5. If you have:
   - No qualifying children, is the amount on line 3 less than $6,750 ($8,750 if married filing jointly)?
   - 1 or more qualifying children, is the amount on line 3 less than $14,850 ($16,850 if married filing jointly)?

   ☐ **Yes.** Leave line 5 blank; enter the amount from line 2 on line 6.

   ☐ **No.** Look up the amount on line 3 in the EIC Table on pages 53–59 to find the credit. Be sure you use the correct column for your filing status and the number of children you have. Enter the credit here. **5**

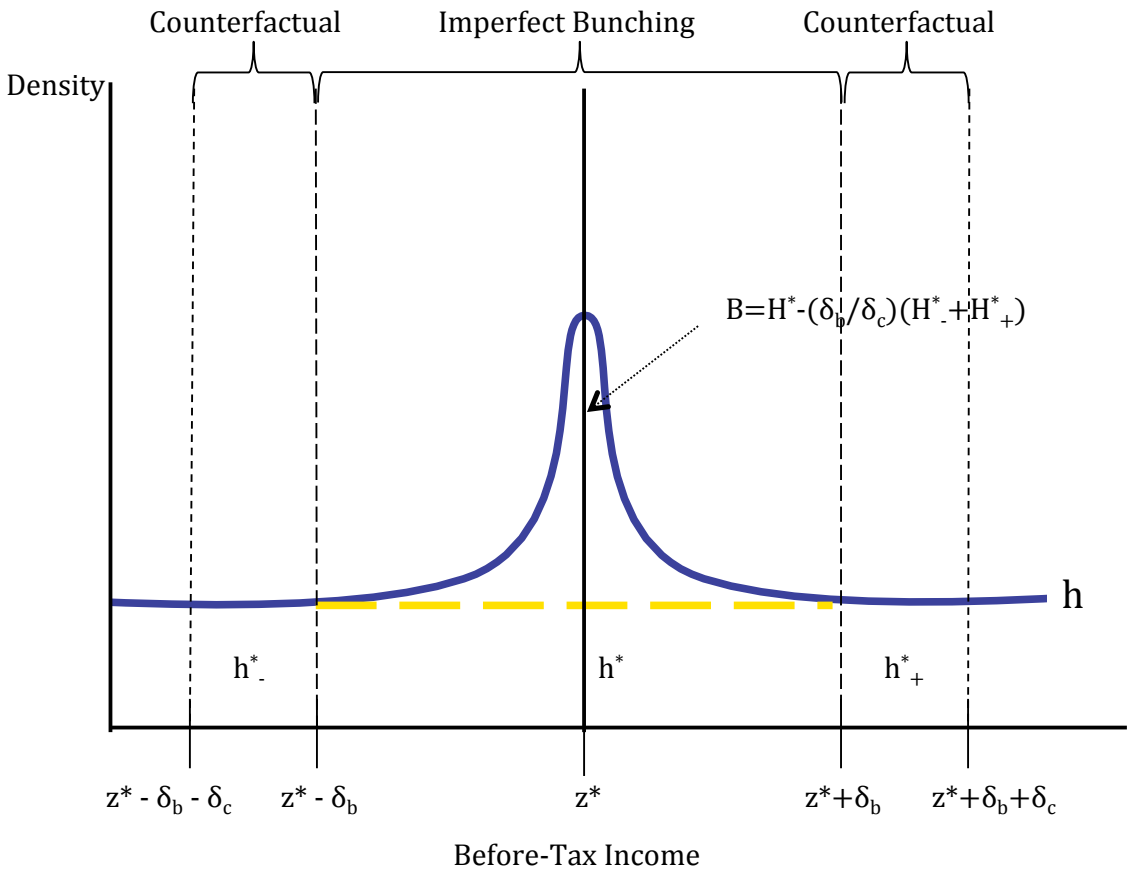
   Look at the amounts on lines 5 and 2. Then, enter the **smaller** amount on line 6.

Part 3

6. **This is your earned income credit.** **6**

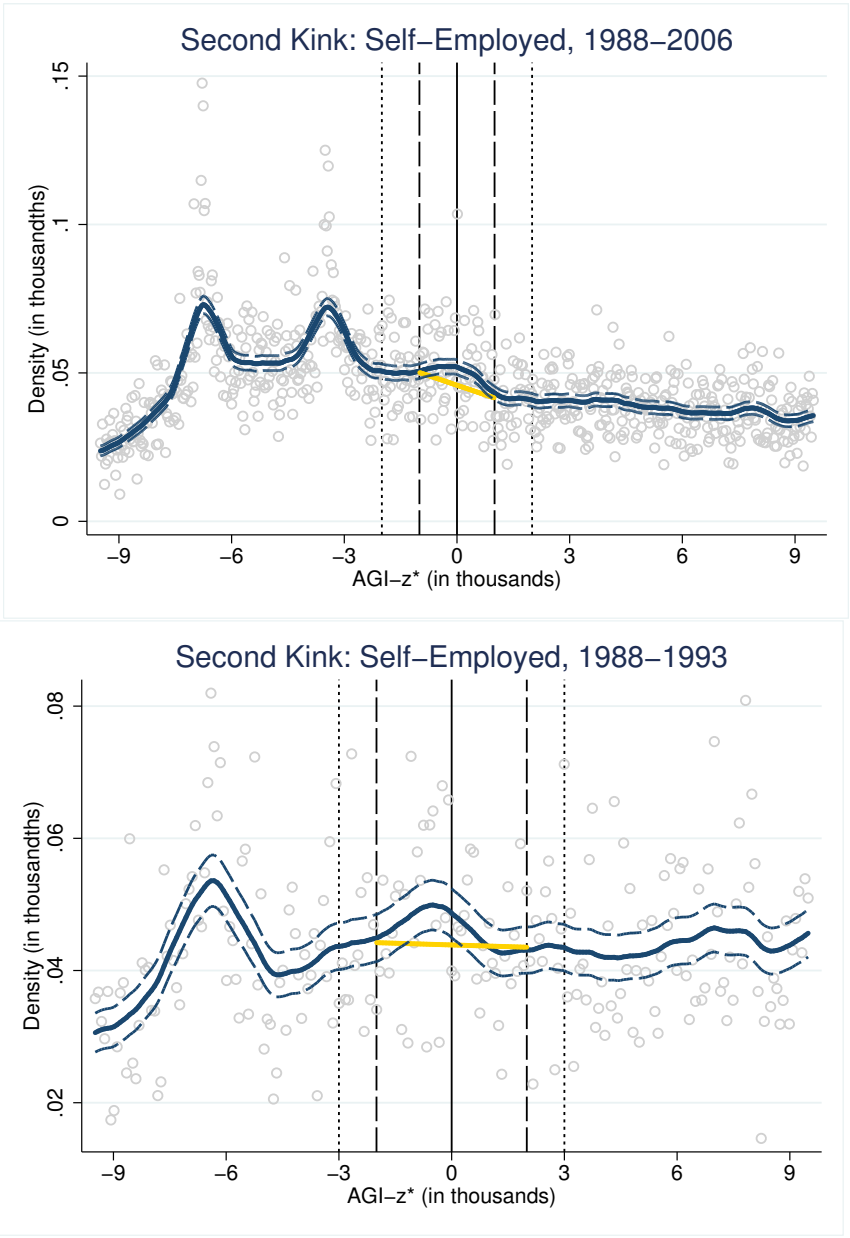
This figure displays the main EITC calculation form for 2006. On line 1, individuals are asked to write down their earned income. On line 2, they calculate their credit using a tax table, and are asked to stop if they are ineligible. On line 3, they enter their AGI. On line 4, they are asked whether their earned income and AGI are equal (i.e., they have no unearned income). If they are, they can skip the rest of the form and enter the credit amount from line 2 on line 6, where line 6 is the amount of EITC claimed. If earned income and AGI are not equal, they proceed to the next question before line 5, which asks them whether their AGI places them in the phase-out region. If it does not, they again skip the rest of the form and enter the credit amount from line 2 on line 6. If their AGI does place them in the phase-out region, line 5 asks them to calculate their EITC based on their AGI. They are then required to take the minimum of lines 2 and 5 and put this amount on line 6—this is the amount of EITC they will receive.

Figure 4: Estimating Elasticity from Bunching Around a Tax Kink



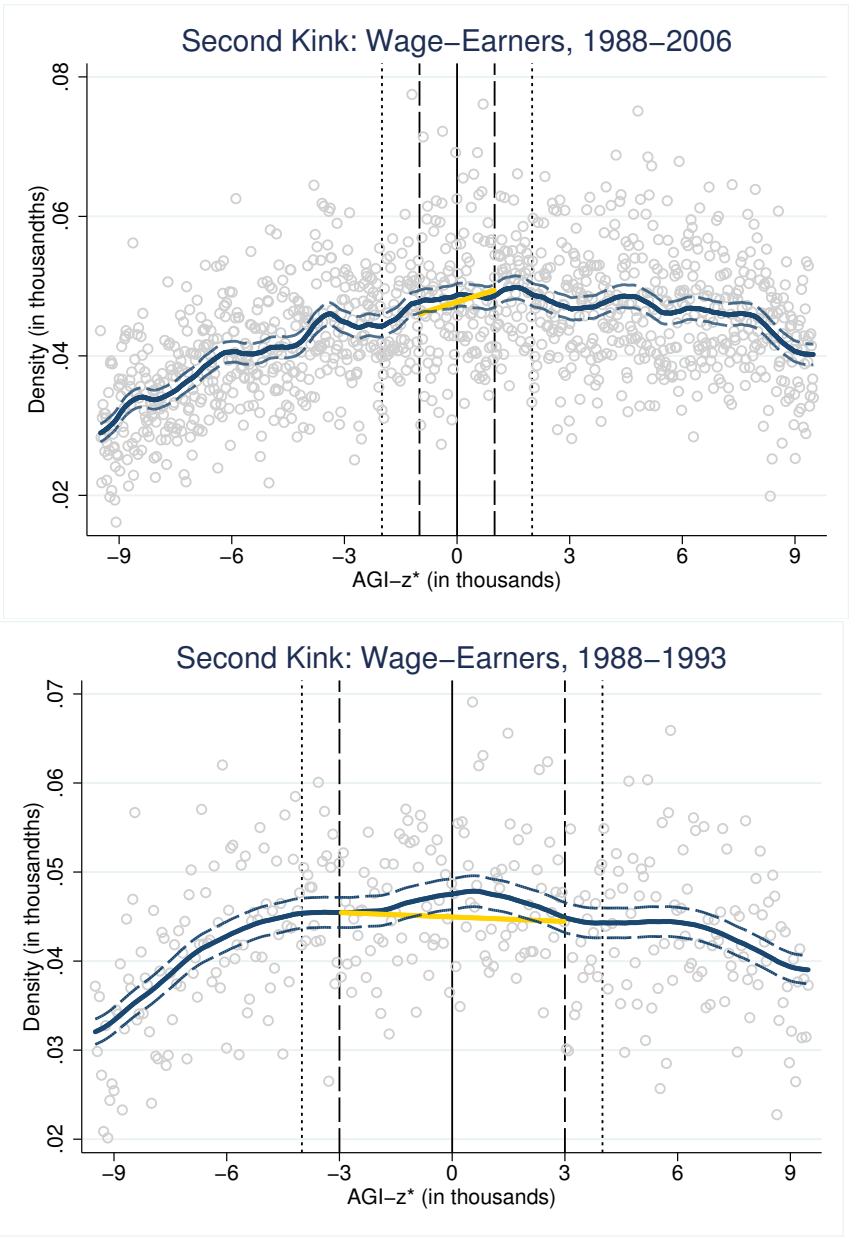
This figure illustrates how bunching at a tax kink can be used to identify the response to the change in marginal tax rate at the tax kink. z^* is the tax kink, $2\delta_b$ is the width of the imperfect bunching region, δ_c is the width of the counterfactual region. h^*_- and h^*_+ are the densities in the counterfactual regions, h^* is the density in the bunching region, each of which is depicted by the blue line in its respective region. $H^* = 2\delta_b h^*$ is the cumulative density in the imperfect bunching region. $H^*_- = \delta_c h^*_-$ and $H^*_+ = \delta_c h^*_+$ are the cumulative densities in the counterfactual region. B is the amount of imperfect bunching in the bunching region, which is given by the area between the solid blue and dashed yellow lines.

Figure 5: Estimated Density Around the Second EITC Kink for Self-Employed Individuals



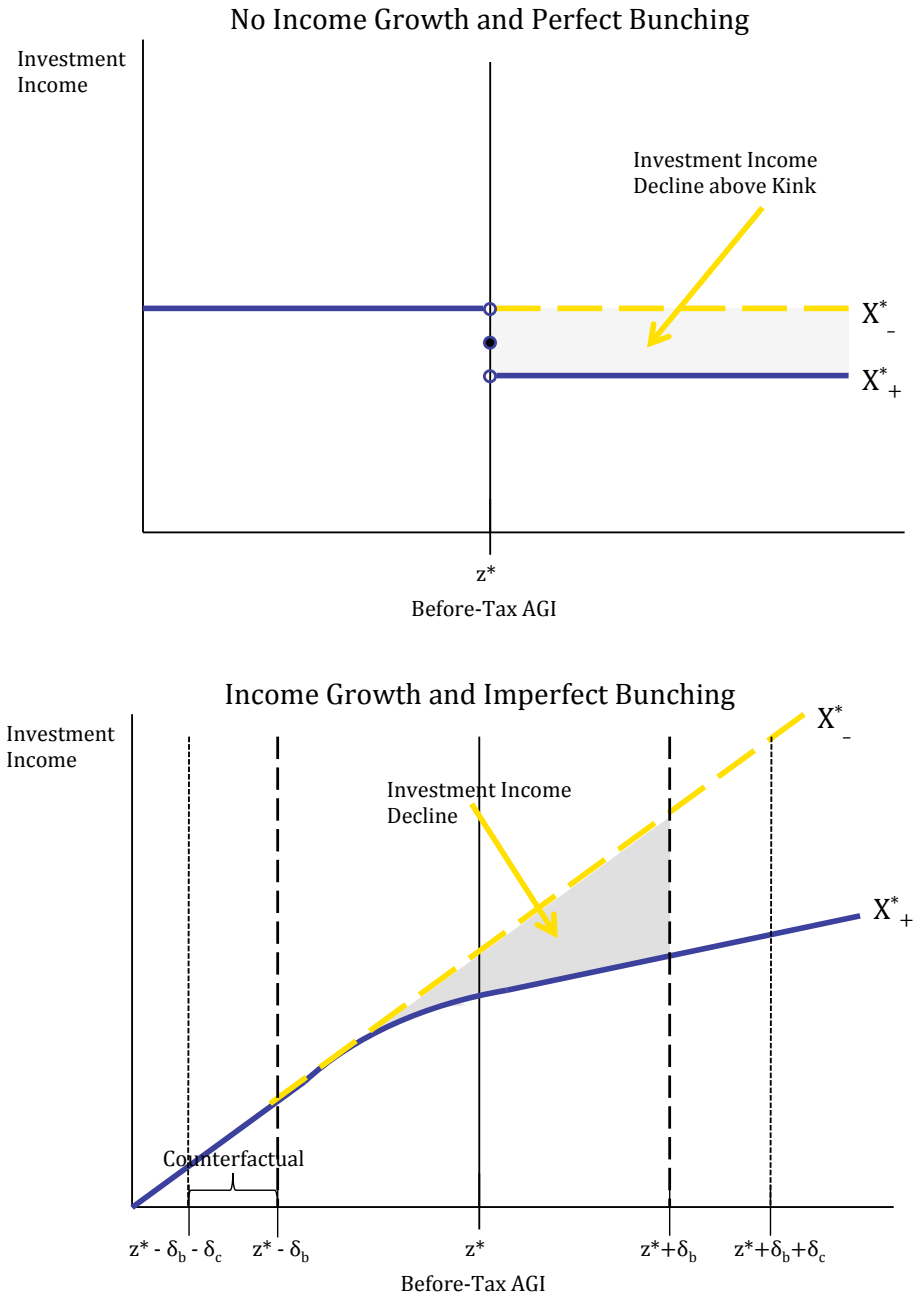
All adjusted gross income (AGI) values are in 2006 dollars, renormalized so that the kink (z^*) is at zero for all years. The estimates are population weighted. The density estimates are normalized to integrate to one. 95 percent confidence intervals are given by the dashed lines. The scatterplot provides the unsmoothed density estimates. The region inside the long-dash vertical lines is used to calculate the bunching at the tax kink, and the region between the short-dash and long-dash lines is used to calculate the counterfactual density. The implied counterfactual is given by the solid yellow line. The bin size and bandwidth are 27 and 797 for the top panel and are 83 and 1,469 for the bottom panel.

Figure 6: Estimated Density Around the Second EITC Kink for Wage-Earning Individuals



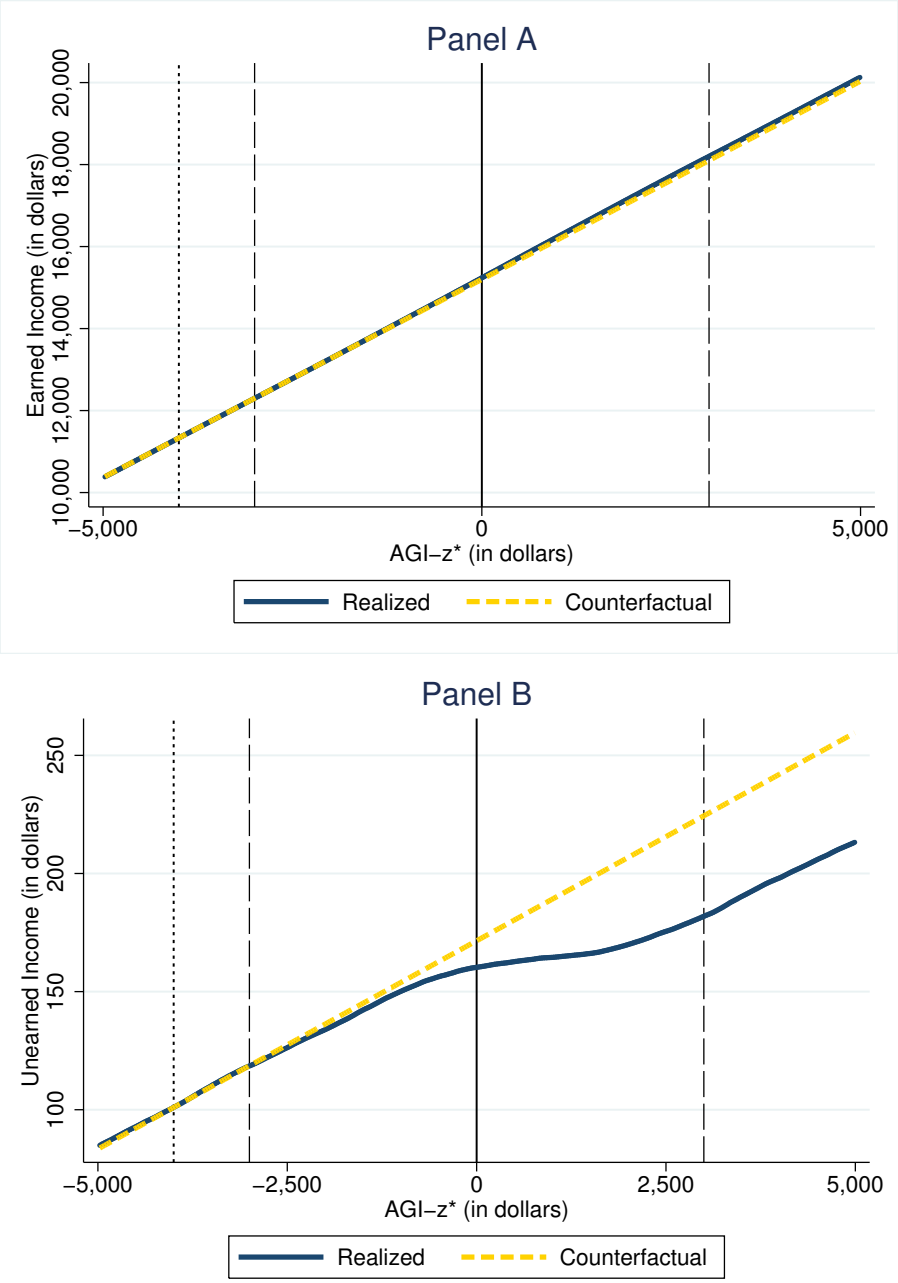
All adjusted gross income (AGI) values are in 2006 dollars, renormalized so that the kink (z^*) is at zero for all years. The estimates are population weighted. The density estimates are normalized to integrate to one. 95 percent confidence intervals are given by the dashed lines. The scatterplot provides the unsmoothed density estimates. The region inside the long-dash vertical lines is used to calculate the bunching at the tax kink, and the region between the short-dash and long-dash lines is used to calculate the counterfactual density. The implied counterfactual is given by the solid yellow line. The bin size and bandwidth are 17 and 745 for the top panel and are 55 and 2,310 for the bottom panel.

Figure 7: Estimating Elasticity from Change in Average Income Around the Tax Kink



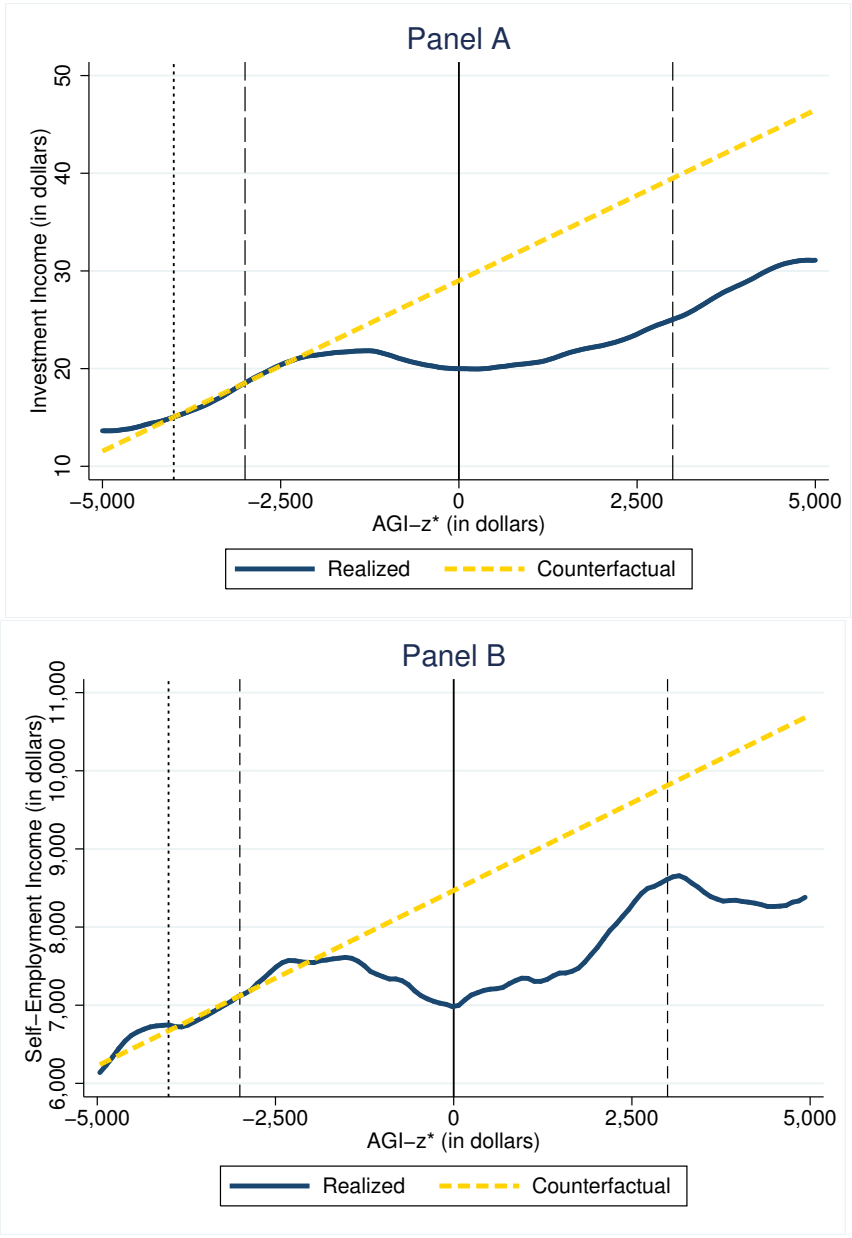
This figure shows how a lower bound on the elasticity of response of a given income type X can be calculated in the region around the tax kink. X^*_+ (the solid blue line) gives the realized amount of investment income. X^*_- (the yellow dashed line) gives the counterfactual amount of investment income. The shaded grey region is the implied decline in investment in the region around the tax kink that is due to the increase in marginal tax rate above the kink.

Figure 8: Estimated Average Earned and Unearned Income Around the Second EITC Kink



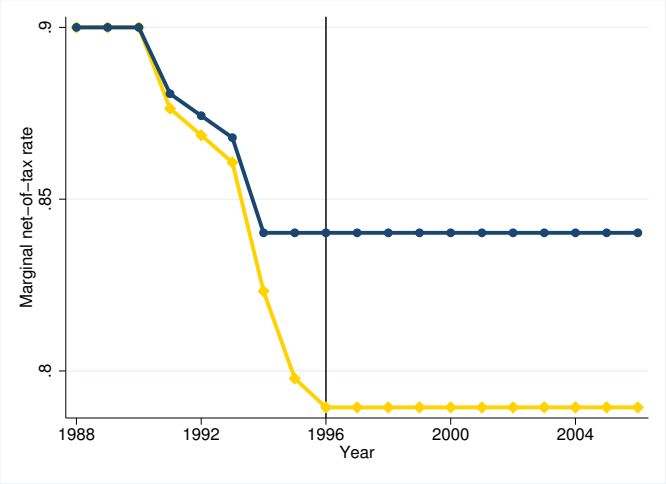
Wage-earners for years 1988-2006 with no unemployment insurance income are included. All adjusted gross income (AGI) values are in 2006 dollars, renormalized so that the kink (z^*) is at zero for all years. The estimates are population weighted. The region inside the dashed vertical lines is used to calculate the elasticity. Realized average income is given by the solid blue line. The implied counterfactual is given by the dashed yellow line and is constructed using investment income in the region between the short-dash and long-dash vertical lines (-4,000 to -3,000). The bin size and bandwidth for earned income are 31 and 2,518 and are 31 and 3,787 for unearned income.

Figure 9: Estimated Average Income Around the Second EITC Kink



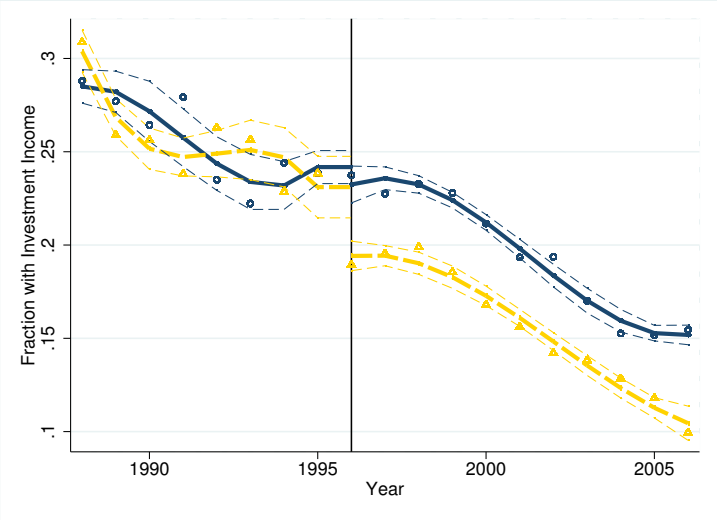
Panel A includes wage-earners for years 1988-2006 with no unemployment insurance income whose EITC investment income is less than $2,300. Panel B includes all individuals with positive self-employment income and no unemployment insurance income for years 1988-1993. The estimates are population weighted. All adjusted gross income (AGI) values are in 2006 dollars, renormalized so that the kink (z^*) is at zero for all years. The region inside the long-dash vertical lines is used to calculate the elasticity. Realized average income is given by the solid blue line. The implied counterfactual is given by the dashed yellow line and is constructed using investment income in the region between the short-dash and long-dash vertical lines (-4,000 to -3,000). The bin size and bandwidth are 31 and 2,417 for Panel A and are 88 and 1,384 for Panel B.

Figure 10: Variation in the Marginal Net-of-Tax Rate by Number of Dependents



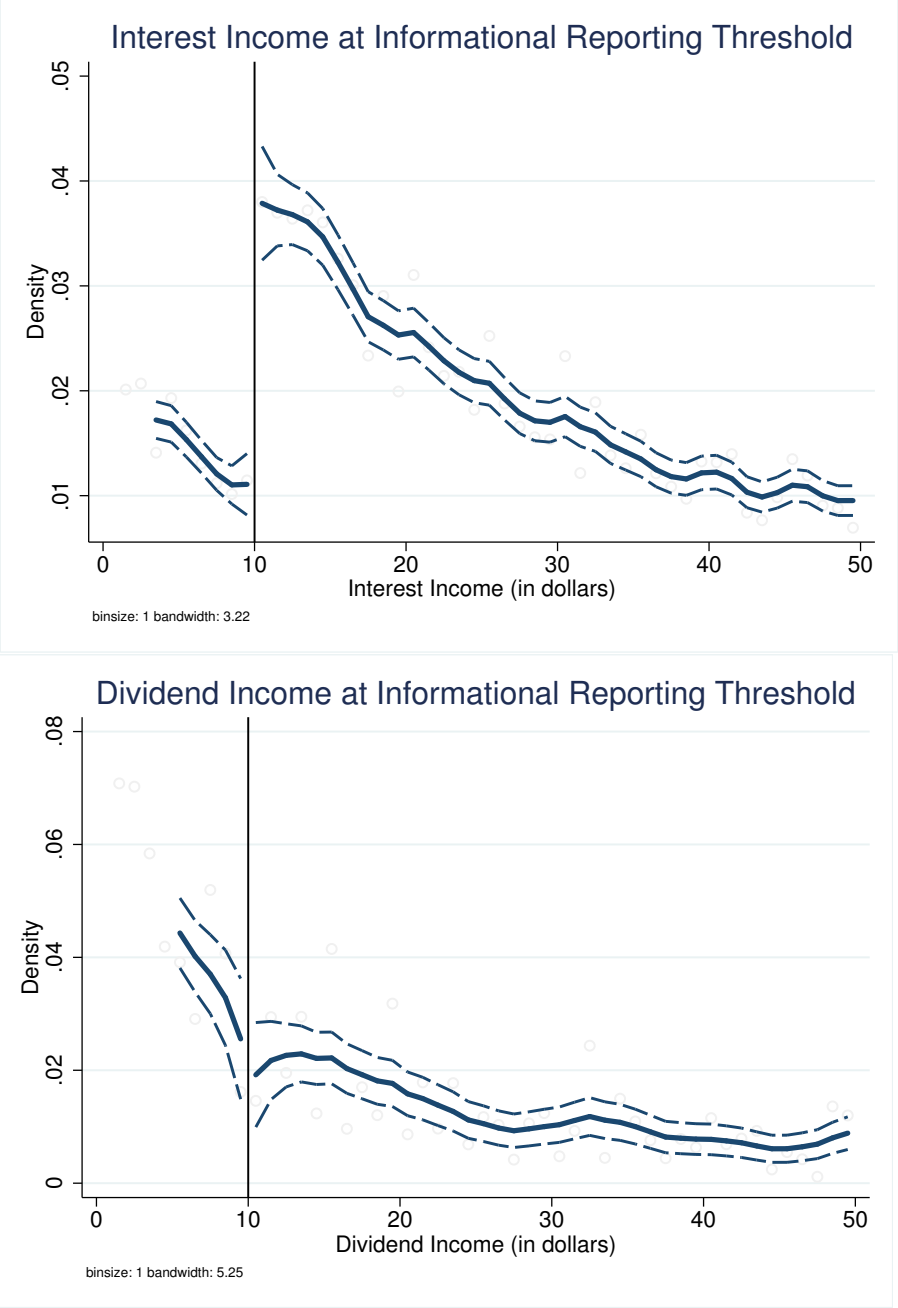
This figure plots the variation in the marginal net-of-tax phase-out rate by year and number of dependents. The marginal net-of-tax rate for one dependent is depicted by the blue line with circles and the marginal net-of-tax rate for two or more dependents is depicted as the yellow line with diamonds. The marginal net-of-tax rate is one minus the marginal tax rate.

Figure 11: Likelihood Individuals have Investment Income by Year and Number of Dependents



This figure plots the likelihood an individual holds investment income by year for one (blue line) and two dependents (yellow line) based on the restrictions used in the baseline estimates, with a break at 1996 (marked by the vertical black line). The thin dashed lines are 95 percent confidence intervals. The scatter plot gives the value for each dependent-year pair. The lines are from a cubic polynomial regression in year interacted with a pre/post indicator and number of dependents.

Figure 12: Individuals' Responses to the Informational Reporting Threshold



These estimates include any individual with positive interest or dividend income, subject to the baseline sample restrictions in Section 5. The implied log difference in density at the information reporting threshold for interest income is 1.21 (0.21) and for dividend income is 0.18 (0.33). Years 1996-2006 are included. Interest and dividend income is in nominal dollars. The estimates are population weighted. The density estimates are normalized to integrate to one. The smoothed density estimates do not start at zero because they cannot be estimated accurately (or could, but not with the same bandwidth) until interest or dividend income is greater than or equal to the bandwidth.

Table 1: EITC Schedule Parameters

EITC Range	Bracket	Marginal Tax Rate
1979-1984 (in 2006 $), Qualifying Children ≥ 1 :[1]		
Phase-in	0-$9,779	-10%
Plateau	$9,779-$11,735	0%
Phase-out	$11,735[2]-$19,559	12.5%
1985-1986 (in 2006 $), Qualifying Children ≥ 1 :[1]		
Phase-in	0-$9,061	-11%
Plateau	$9,061-$11,780	0%
Phase-out	$11,780-$19,935	12.22%
1987 (in 2006 $), Qualifying Children ≥ 1 :		
Phase-in	0-$10,743	-14%
Plateau	$10,743-$12,227	0%
Phase-out	$12,227-$27,267	10%
1988-1990 (in 2006 $), Qualifying Children ≥ 1 :[3]		
Phase-in	0-$10,747	-14%
Plateau	$10,747-$16,933	0%
Phase-out	$16,933-$31,979	10%
1991 (in 2006 $), Qualifying Children $= 1$:[3]		
Phase-in	0-$10,748	-16.7%
Plateau	$10,748-$16,935	0%
Phase-out	$16,935-$31,988	11.93%
1991 (in 2006 $), Qualifying Children ≥ 2 :[3]		
Phase-in	0-$10,608	-17.3%
Plateau	$10,608-$16,935	0%
Phase-out	$16,935-$31,988	12.36%
1992 (in 2006 $), Qualifying Children $= 1$:[3]		
Phase-in	0-$10,752	-17.6%
Plateau	$10,752-$16,929	0%
Phase-out	$16,929-$31,986	12.57%
1992 (in 2006 $), Qualifying Children ≥ 2 :[3]		
Phase-in	0-$10,752	-18.4%
Plateau	$10,752-$16,929	0%
Phase-out	$16,929-$31,986	13.14%
1993 (in 2006 $), Qualifying Children $= 1$:[3]		
Phase-in	0-$10,754	-18.5%
Plateau	$10,754-$16,928	0%
Phase-out	$16,928-$31,983	13.21%
1993 (in 2006 $), Qualifying Children ≥ 2 :[3]		
Phase-in	0-$10,754	-19.5%
Plateau	$10,754-$16,928	0%

Phase-out	$16,928-$31,983	13.93%
1994 (in 2006 $), Qualifying Children = 1 :[3]		
Phase-in	0-$10,434	-26.3%
Plateau	$10,434-$14,810	0%
Phase-out	$14,810-$31,983	15.98%
1994 (in 2006 $), Qualifying Children ≥ 2 :[3]		
Phase-in	0-$11,343	-30%
Plateau	$11,343-$14,810	0%
Phase-out	$14,810-$34,058	17.68%
1995 (in 2006 $), Qualifying Children = 1 :[3]		
Phase-in	0-$8,083	-34%
Plateau	$8,083-$14,815	0%
Phase-out	$14,815-$32,013	15.98%
1995 (in 2006 $), Qualifying Children ≥ 2 :[3]		
Phase-in	0-$11,338	-36%
Plateau	$11,338-$14,815	0%
Phase-out	$14,815-$35,001	20.22%
1996-2006 (in 2006 $), Qualifying Children = 1 :[3,4]		
Phase-in	0-$8,080	-34%
Plateau	$8,080-$14,810	0%
Phase-out	$14,810-$32,001	15.98%
1996-2006 (in 2006 $), Qualifying Children ≥ 2 :[3,4]		
Phase-in	0-$11,340	-34%
Plateau	$11,340-$14,810	0%
Phase-out	$14,810-$36,348	21.06%

[1]The EITC schedule was fixed in nominal dollars in these years. Therefore, in real terms, the bracket cutoffs declined substantially over this time. In each case, the real values in the last year of the range are listed.

[2]The beginning of the phase-out range is underlined in each year because it marks the location of the second tax kink.

[3]The parameters are inflation-indexed, but are then rounded. In each case, the real values based on the actual values in the last year in the range are listed.

[4]For married taxpayers, the phase-out region was shifted right by $1,000 in 2002-2004 and $2,000 in 2005-2006.

Table 2: Descriptive Statistics for EITC Eligible Taxpayers

Variable	Observations	Mean	Std. Dev.	Fraction $\neq 0$	Std. Dev.
Wage Income	105,848	16,017.45	9,672.61	0.984	0.124
Self-Employment Income	105,848	1,042.46	4,517.18	0.157	0.364
Unearned Income	105,848	571.66	10,649.28	0.336	0.472
Dividend & Interest Income	105,848	104.69	1,677.13	0.186	0.389
Capital Gains or Losses	105,848	8.57	1,828.51	0.022	0.147
Partnership Income	105,848	-66.17	10,386.36	0.031	0.174
Other Gains or Losses	105,848	-14.42	3,178.40	0.005	0.067
Taxable Refunds	105,848	16.46	296.59	0.032	0.176
Taxable Pension, Annuity & IRA Distributions	105,848	16.38	296.67	0.051	0.219
Social Security Income	105,848	5.11	118.08	0.003	0.055
Unemployment Insurance	105,848	335.56	1,372.48	0.107	0.309
Deductions	105,848	103.94	855.12	0.050	0.218
Married?	105,848	0.28	0.45	-	-
Number of Dependents	105,848	1.41	0.49	-	-
Tax Preparer?	102,161	0.62	0.49	-	-
Claim EITC?	105,848	0.92	0.28	-	-

This table reports summary statistics for all EITC eligible individuals with one or two dependents. Tax preparation variables are missing for 1990 and are occasionally missing in other years. All income values are in 2006 dollars.

Table 3: Elasticities at Second EITC Kink

	Bunching Elasticity Estimates				
	(1)	(2)	(3)	(4)	(5)
Elasticity	0.063***	0.010	0.104	0.095*	0.104**
	(0.020)	(0.011)	(0.073)	(0.052)	(0.047)
Type	Self-employed	Wage-Earner	Self-employed	Self-employed	Wage-earner
Income Type	AGI	AGI	AGI	AGI	AGI
Years	1988-2006	1988-2006	1988-1993	1988-1993	1988-1993
δ_b	1,000	1,000	2,000	1,000	3,000
δ_c	1,000	1,000	1,000	1,000	1,000

	Income Elasticity Estimates			
	(6)	(7)	(8)	(9)
Elasticity	−0.015	0.516	1.879**	1.125
	(0.013)	(0.451)	(0.832)	(1.060)
Type	Wage-earner	Wage-earner	Wage-earner	Self-employed
Income Type	Earned	Unearned	Investment	Self-employment
Years	1988-2006	1988-2006	1988-2006	1988-1993
δ_b	3,000	3,000	3,000	2,000
δ_c	1,000	1,000	1,000	1,000

The bunching estimates are calculated by first calculating the amount of imperfect bunching around the tax kink using a local linear regression as described in the text and the corresponding figures. Then, equation (1) is used to construct elasticities based on this estimate. The income elasticity estimates are calculated by first calculating average income of the given income type in the region around the tax kink using a local linear regression as described in the text and the corresponding figures. Then, equation (2) is used to calculate elasticities based on this estimate. The standard errors are calculated via non-parametric bootstrap (500 replications). These estimates exclude individuals that did not file for the EITC, had no dependents, or had more than two dependents. The wage-earner estimates include only wage-earners and the self-employed include only self-employed individuals. $2\delta_b$ is the width of the imperfect bunching region and δ_c is the width of the counterfactual region.

Table 4: Repeated-Cross-Section Descriptive Statistics

Variable	Observations	Mean	Std. Dev.	Fraction $\neq 0$	Std. Dev.
Earned Income	31,480	23,679.15	3,444.38	-	-
Unearned Income	31,480	1,160.53	7,695.71	0.351	0.477
Dividend & Interest Income	31,480	48.60	205.94	0.206	0.404
Capital Gains or Losses	31,480	-10.35	227.29	0.016	0.125
Partnership Income	31,480	245.71	5,971.41	0.028	0.164
Other Gains or Losses	31,480	-7.32	1,777.94	0.001	0.029
Unemployment Insurance	31,480	316.48	1,389.29	0.100	0.300
Married?	31,480	0.28	0.45	-	-
Number of Dependents	31,480	1.40	0.49	-	-
Tax Preparer?	30,217	0.59	0.49	-	-
EITC Eligible?	31,480	0.96	0.19	-	-

This table reports summary statistics for individuals in the baseline estimates. Tax preparation variables are missing for 1990 and are occasionally missing in other years. All income values are in 2006 dollars.

Table 5: Baseline Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Extensive Margin:							
$log(1-\tau)$	0.612***	0.613***	0.868***	0.816***	0.867***	0.824***	0.353*
	(0.174)	(0.156)	(0.213)	(0.221)	(0.267)	(0.217)	(0.188)
$log(1-\tau)$ x *married*					−0.098		
					(0.306)		
$log(1-\tau)$ x $\mathbf{1}[ui>0]$						0.114	
						(0.240)	
$log(1-\tau)$ x *pdprep*							0.7366***
							(0.209)
P-value of sum					0.005	0.000	0.000
Elasticty:							
$log(1-\tau)$	3.047***	3.000***	3.682***	3.435***	3.897***	3.505***	1.292*
	(0.810)	(0.724)	(0.835)	(0.936)	(1.061)	(0.917)	(0.748)
$log(1-\tau)$ x *married*					−0.806		
					(1.605)		
$log(1-\tau)$ x $\mathbf{1}[ui>0]$						0.171	
						(1.306)	
$log(1-\tau)$ x *pdprep*							3.3626***
							(1.042)
P-value of sum					0.021	0.000	0.000
Broad Measure of Investment Income?	No	Yes	No	No	No	No	No
Controls x 1992, 1998 included?	No	No	Yes	Yes	Yes	Yes	Yes
State FE included?	No	No	No	Yes	Yes	Yes	Yes
Bandwidth:	18K-30K	18K-30K	18K-30K	18K-30K	18K-30K	18K-30K	18K-30K
Observations	30.480	30.480	30.480	30.480	30.480	30.480	30,217

The extensive margin estimates are mean marginal effects from a probit specification, where the dependent variable is an indicator variable for whether the taxpayer had positive interest or dividend income. The elasticity estimates are mean marginal elasticities from a two-tier model, where the dependent variable in the first tier is an indicator variable for whether the taxpayer had positive interest or dividend income and the dependent variable in the second tier is the log of interest and dividend income. The dependent variables in the second column are constructed from the sum of dividend, interest, partnership, capital gains and losses, other gains and losses. and IRA contributions. All columns include covariates X (as specified in the paper) and year fixed-effects. All income values are in 2006 dollars. Estimates are population weighted. Estimates are for years 1988-2006. Standard errors clustered by year are in parentheses.

Table 6: Robustness Checks

	(1)	(2)	(3)	(4)	(5)
Extensive Margin:					
$log(1-\tau)$	0.816***	0.912***	0.711***	0.596	0.538**
	(0.221)	(0.305)	(0.197)	(0.579)	(0.253)
Elasticty:					
$log(1-\tau)$	3.435***	2.539**	4.554***	3.063	2.981***
	(0.936)	(1.141)	(0.920)	(2.496)	(0.966)
Controls x 1992, 1998 included?	Yes	Yes	Yes	Yes	Yes
State FE included?	Yes	Yes	Yes	Yes	Yes
Bandwidth:	18K-30K	18K-24K	24K-30K	18K-30K	18K-30K
Two dependent taxpayers included?	Yes	Yes	Yes	No	Yes
Three dependent taxpayers included?	No	No	No	Yes	No
Observations	31,480	16,821	14,652	22,996	31,480

The extensive margin estimates are mean marginal effects from a probit specification, where the dependent variable is an indicator variable for whether the taxpayer had positive interest or dividend income. The elasticity estimates are mean marginal elasticities from a two-tier model, where the dependent variable in the first tier is an indicator variable for whether the taxpayer had positive interest or dividend income and the dependent variable in the second tier is the log of interest and dividend income. The last column replaces all values for interest and dividend income with zero whenever these income types are less than $10 before constructing the dependent variables. All columns include covariates X (as specified in the paper) and year fixed-effects. All income values are in 2006 dollars. Estimates are population weighted. Estimates are for years 1988-2006. Standard errors clustered by year are in parentheses.

Table 7: Substitution and Income Effects

	(1)	(2)	(3)	(4)
Extensive Margin:				
$log(1-\tau)$	0.638***	0.781***	0.810*	0.952
	(0.190)	(0.268)	(0.445)	(0.617)
$log(earnings + \tau(end - earnings))$			0.202	0.202
			(0.389)	(0.611)
Elasticity:				
$log(1-\tau)$	3.144***	3.347***	4.456**	4.501
	(0.889)	(1.135)	(2.066)	(3.250)
$log(earnings + \tau(end - earnings))$			1.543	1.377
			(1.813)	(3.204)
Controls x 1992, 1998 included?	No	Yes	No	Yes
State FE included?	No	Yes	No	Yes
Bandwidth:	18K-30K	18K-30K	18K-30K	18K-30K
Single Filers Included	No	No	No	No
Observations	30,266	30,266	30,266	30,266

The extensive margin estimates are mean marginal effects from a probit specification, where the dependent variable is an indicator variable for whether the taxpayer had positive interest or dividend income. The elasticity estimates are mean marginal elasticities from a two-tier model, where the dependent variable in the first tier is an indicator variable for whether the taxpayer had positive interest or dividend income and the dependent variable in the second tier is the log of interest and dividend income. All columns include covariates X (as specified in the paper) and year fixed-effects. All income values are in 2006 dollars. Estimates are population weighted. Estimates are for years 1988-2006. Standard errors clustered by year are in parentheses. Note that single filers is referring to the taxpayer category; unmarried head of household filers are still included.