How Sensitive Are Taxpayers to Marginal Tax Rates? Evidence from Income Bunching in the United States

Jacob A. Mortenson* and Andrew Whitten*

January 2016

Abstract

Understanding the way taxpayers respond to the tax code is critical for revenue and welfare analyses of taxation. One way taxpayers may respond is by bunching at kink points in the tax schedule to avoid high marginal tax rates. We study this phenomenon using over 400 million federal individual income tax returns in the United States from 1996 to 2014, analyzing state and federal statutory kinks as well as effective kinks created by tax credits and phase-outs of deductions and exemptions. Though most kinks do not cause statistically discernible bunching, we find strong responses at other kinks. Consistent with prior research, we see bunching patterns grow over time at the first kink in the Earned Income Tax Credit (EITC) schedule. In addition, we present new evidence documenting (i) the emergence and rapid rise of bunching at the second EITC kink and the Child Tax Credit refundability plateau, (ii) strong responses to the temporary Making Work Pay Tax Credit, and (iii) weak responses at three statutory kinks. In general, responsiveness is strongest at kinks that maximize tax credits, particularly at global refund-maximizing points in the schedule. Though the self-employed bunch more, we find wage earners also respond in recent years. Their responses appear to be driven by evasion, however, as there is no bunching in employer-reported W-2 wages. When translating bunching patterns to elasticities of taxable income, we find a range of values from zero to 1.50, with substantial variation by kink and household type.

Keywords: Bunching, Elasticity of Taxable Income, Public Finance.

JEL Code: H24.

---

*Both authors: Georgetown University and Congressional Joint Committee on Taxation. Contact email: raw65@georgetown.edu. We thank Jerry Auten, Jim Cilke, Tim Dowd, Nada Eissa, Rob Harvey, Patrick Langetieg, Arik Levinson, Rob McClelland, John Rust, Heidi Schramm, and numerous seminar participants for helpful comments. This research embodies work undertaken for the staff of the Joint Committee on Taxation, but as members of both parties and both houses of Congress comprise the Joint Committee on Taxation, this work should not be construed to represent the position of any member of the Committee.

1 Introduction

This paper estimates taxpayer responsiveness to marginal income tax rates by measuring the degree to which taxpayers bunch at kink points in income tax schedules. A kink point is an income amount for a given taxpayer at which marginal tax rates change discretely, marking the end of one tax bracket and the beginning of the next. Standard economic theory predicts that some taxpayers will avoid brackets with high tax rates by bunching at kinks where tax rates increase, resulting in extra mass in the distribution of income close to these kinks. We use measures of excess mass to estimate elasticities of taxable income (ETIs), building upon methods developed by Saez (2010) and Chetty et al. (2011).

ETIs are necessary statistics for welfare and revenue analyses of current and proposed income tax regimes (Chetty, 2009). They capture the degree to which taxpayers adjust taxable income in response to marginal tax rates, allowing for responses through labor supply, deduction, and evasion decisions. These elasticities are functions of both taxpayer preferences and policy choices such as the set of allowable deductions or the level of tax enforcement (Kopczuk, 2005). Estimating heterogeneous responses across household types is important, as the tax code accomplishes different objectives for different groups. For example, low earners receive income subsidies partly designed to increase labor supply, while high earners receive deductions and credits for activities like charitable giving and improving home energy efficiency.

The bunching approach yields ETI estimates because the amount of excess mass at kinks is a function of taxpayer sensitivity to marginal tax rates and kink size, which is given by the tax code. Measured this way, ETIs capture responsiveness to changes in tax rates that are within-year, freeing researchers from relying on tax reforms for identifying variation.[1] Saez (2010) develops this technique, analyzing bunching around kinks in the Earned Income Tax Credit (EITC) and federal income tax schedules using public-use tax data from 1960 to 2004. He finds substantial bunching only around the first EITC kink and at $0 of taxable income. These yield elasticities for the general population of approximately 0.10–0.33 and 0.11–0.26, respectively. All other kinks appear to generate no bunching and yield elasticities that are statistically indistinguishable from zero. Importantly, when looking closer at the elasticities around the EITC schedule, Saez finds the response is driven entirely by the self-employed. Once they are removed from the sample, estimates fall to 0.00–0.03 and are statistically insignificant. Isolating the self-employed yields elasticities of 0.75–1.10.

Following Saez, the bunching approach has been used to analyze responsiveness to the

[1] The conventional approach to estimating ETIs compares taxpayers' incomes before and after major tax reforms, exploiting the fact that reforms tend to affect different groups differently. Saez et al. (2012) review this literature, finding that the most convincing estimation specifications yield elasticities between 0.1 and 0.4. However, Weber (2014) argues that previous attempts have failed to eliminate the bias that arises from the mechanical endogeneity of tax rates due to progressive rate schedules. She proposes a set of approximately exogenous instruments for marginal tax rates and analyzes the Tax Reform Act of 1986, estimating an ETI of around 0.9.

Annual Earnings Test for Social Security income (Burtless and Moffitt, 1984; Friedberg, 2000; Gelber et al., 2013), to the Saver's Credit notches (Ramnath, 2013), and to discontinuities in tax schedules in Denmark (Le Maire and Schjerning, 2013; Chetty et al., 2011), Sweden (Bastani and Selin, 2014), Pakistan (Kleven and Waseem, 2013), Ireland (Hargarden, 2015), and the United Kingdom (Devereux et al., 2014). Kleven (2015) discusses this research as well as applications of the bunching approach outside of the tax literature.

Our bunching measures use detailed administrative data drawn from the universe of federal income tax returns in the United States from 1996 to 2014. With over 400 million observations in total, most of our estimators – including those for narrowly defined household types in a given year – use tens or hundreds of thousands of observations, resulting in smooth distributions over the intervals surrounding kinks. We build upon work by Saez (2010) and Chetty et al. (2011) in generating ETI estimates, and our bunching results associated with the EITC are broadly consistent with Saez's. However, Saez measures bunching using public-use data covering 0.1% of the population, while our dataset, which is significantly larger and more recent, allows us to make a series of improvements to bunching and income measures.

We make five contributions to the modern public finance literature estimating income responses to marginal tax rate changes. First, we measure annual variation in tax rate sensitivity over a nineteen-year period that includes two tax reforms and several business cycle fluctuations, including the Great Recession. Second, we derive the bunching estimator of Saez (2010) without imposing a functional form on utility. Third, we employ an estimation technique that fits more closely with the theory of discrete kinks and show how our technique compares with those of Saez (2010) and Chetty et al. (2011). Fourth, we provide new evidence of bunching behavior, documenting the emergence of bunching at five kinks where taxpayers were previously unresponsive, including bunching among wage earners. Finally, we advance a new explanation of bunching: taxpayers gravitate towards the refund-maximizing point of the schedule.

Our primary finding is the rapid rise of new bunching patterns at two low-income kinks. Statistically significant and economically meaningful bunching emerges in the mid 2000s at the second EITC kink and the refundability plateau of the Child Tax Credit (CTC). The second EITC kink marks the beginning of the credit's phaseout region and increases effective marginal tax rates by up to roughly 21 percentage points. The CTC refundability plateau marks the point at which the credit becomes fully refundable and increases effective marginal tax rates by 15 percentage points. Like the first EITC kink – another place we see strong bunching, confirming prior research – both of these kinks allow taxpayers to receive the maximum credit amounts, potentially elevating their salience. Moreover, for some groups they denote the global refund-maximizing point in the tax schedule, further increasing their salience.

We also find small, steady bunching at four other kinks for many groups. Consistent with Saez (2010), we see bunching at the beginning of the federal income tax schedule, where tax rates increase by ten percentage points. In addition, we see responsiveness at the second and third kinks in the statutory schedule, where marginal rates increase by ten and three percentage points, respectively. Finally, we find bunching at the kink that maximizes the Making Work Pay Tax Credit (MWPTC) – a small credit made available to low-income taxpayers in 2009 and 2010 that changes effective marginal tax rates by roughly six percentage points.

The patterns we observe, and their corresponding elasticity estimates, are not constant over time. At the first EITC kink, we see growing responsiveness, consistent with Chetty et al. (2013). At the second EITC kink and the CTC refundability plateau, bunching is nonexistent in the late 1990s and early 2000s, but emerges rapidly in the mid 2000s and then decreases near the end of our sample. This highlights the changing nature of tax rate sensitivity, and suggests the average taxpayer has more knowledge of the tax code in 2014 than in 1996. It also suggests ETI estimates based on tax reforms in the 1980s and 1990s may be weak proxies for contemporary responsiveness to marginal tax rates.

Many of the dynamics we observe are due to taxpayers following the refund-maximizing point as it alternates between kinks. Single, self-employed taxpayers with two children nicely illustrate this phenomenon. During 2004 to 2008, these taxpayers bunch at the second EITC kink, which maximizes their total refund, and ignore the nearby CTC refundability plateau. During 2009 and 2010, the two kinks are located at essentially the same place, and the group continues to bunch there. During 2011 to 2014, however, the CTC refundability plateau becomes the refund-maximizing point for these taxpayers, and they shift to bunch there, ignoring the nearby second EITC kink.

Responsiveness at tax kinks must be driven by labor supply, labor demand, deduction, or tax evasion decisions. Because most of the bunching we observe occurs at kinks defined with respect to earned income, deduction opportunities such as charitable giving cannot play a large role. While it is difficult to distinguish real economic activity from tax evasion for the self-employed, wage earnings are reported by employers to the IRS on Form W-2. The evidence we present suggests wage earner responsiveness is driven purely by evasion. We match reported W-2 earnings to primary and secondary tax filers, and though we observe substantial bunching among taxpayer-reported wage income, we do not observe any bunching at low-income kinks among employer-reported wage income. Moreover, misreporting here occurs predominantly in the form of over-reporting wage income when marginal tax rates are negative rather than under-reporting when marginal tax rates are positive.

Our results shed light on the features of the tax code that elicit the strongest behavioral responses. We observe the largest ETIs not at statutory kinks, but at effective kinks created by low-income tax credits. ETIs exceeding 0.1 occur only at kinks that maximize the EITC, CTC, and MWPTC. These credits substantially alter income reporting decisions, and our results indicate these distortions are growing over time.

2 Data and Institutional Background

Our analysis of taxpayer bunching uses data drawn from the Internal Revenue Service's Compliance Data Warehouse (CDW). The CDW contains the universe of tax returns (e.g. Form 1040 and its schedules) and information returns (e.g. Form W-2) of individuals in the United States. Each observation in our data is a tax unit that filed a tax return for a given year. This could be an individual or a married couple filing jointly. Most of the data consist of fields on the tax return and its schedules. These include ordinary and capital gains income, as well as deductions, credits, and taxes paid. Certain demographic information is also found on tax returns, such as marital status, number of children, years of birth of those in the tax unit, and state of residence. We also make use of wage and industry information from the Form W-2 as well as information on date of birth and sex at the time of birth from the Social Security Administration's Data Master File.

Our primary set of data is a sample from the CDW consisting of all tax returns in the seven states with no state income taxes: Alaska, Florida, Nevada, South Dakota, Texas, Washington, and Wyoming. We choose this subsample for two reasons. First, the combined population of these states accounts for roughly 20% of the U.S. population and is sufficiently large to identify heterogeneous responses in narrow subpopulations. Second, and most important, state income tax regimes interfere with our analysis by creating new kinks and amplifying existing federal kinks. When state kinks are near the federal kinks we study, they may affect the distribution of income we observe, confounding the identification of responses to federal kinks. When state kinks amplify federal kinks, as in the case of state-level EITCs, ignoring state-specific taxes and subsidies – as done in Saez (2010) – biases upwards elasticity estimates.

In addition, we draw several other samples for specific kinks. One is a sample from the CDW consisting of *all* tax returns in the neighborhood of high-income statutory kinks and kinks created by phaseouts of personal exemptions and itemized deductions. Another is a sample of *all* tax returns that claim the American Opportunity Tax Credit near the beginning of the phase-out of the credit. The mass from the seven states listed above is insufficient to analyze these kinks. Yet another is a sample from the CDW consisting of all tax returns from 2003 to 2014 in California, Connecticut, and New Jersey in the neighborhood of each state's largest kink.

2.1 Federal Tax Code

Despite its well-deserved reputation for complexity, the U.S. federal income tax code has a straightforward statutory schedule.[2] In 1996, the first year in our sample, the schedule

[2]The complexity comes from the many rules that govern the definition of *taxable income*, which is total income less deductions and exemptions, as well as the long list of tax credits available to certain taxpayers.

Table 1: Federal ordinary income tax: Statutory marginal tax rates (%)

	Bracket						
Year(s)	1st	2nd	3rd	4th	5th	6th	7th
1996-2000	—	15	28	31	36	39.6	—
2001	—	15	27.5	30.5	35.5	39.1	—
2002	10	15	27	30	35	38.6	—
2003-2012	10	15	25	28	33	35	—
2013-2014	10	15	25	28	33	35	39.6

The location of the kinks are adjusted for inflation annually. Taxpayers must update their knowledge of tax schedules annually in order to bunch. See Figure 1 for a graphical depiction of the schedule, including effective kinks created by income phase-outs associated with credits.

for ordinary income had five tax brackets whose marginal tax rates are detailed in Table 1.[3] This schedule remained stable on an inflation-adjusted basis until the Bush Tax Cuts of 2001-2003, which added a 10% bracket at the beginning of the schedule and generally lowered rates.[4] The Bush tax rates remained in place until the American Taxpayer Relief Act of 2012, which reinstated a top bracket of 39.6%.

The kink points separating the federal tax brackets vary by year and filing status. To keep terminology uniform, throughout the paper we take the "first" kink to be the divider between the first and second brackets according to the post-Bush Tax Cuts schedule. Similarly, we take the "second" kink to be the divider between the second and third brackets, and so on. Thus, in our terminology, the first kink did not exist in 1996-2001 and the sixth kink did not exist in our sample until 2013. The "zeroth" kink marks the beginning of the schedule in all years.

Unlike the EITC and CTC schedules detailed below, the statutory income tax schedule is progressive. All kinks see marginal tax rates increase and are therefore convex. Most of these kinks create small changes in the net-of-tax rate. In the presence of significant optimization frictions, we might not expect bunching at those kinks. Two of the kinks, however, create absolute changes in tax rates of 10 percentage points or more: the zeroth and second statutory kinks. All else equal, we expect to observe stronger responsiveness at these kinks. Estimating bunching at the zeroth kink requires care, however, as it also represents the 1040 filing threshold for most tax units. That is, most individuals with taxable

[3]The actual implementation of Table 1's marginal tax rates involves a large number of $50 micro-brackets, with discrete changes in tax liability only at the beginning of each bracket. Hence, the effective tax rate on marginal income is actually zero for most taxpayers for small enough marginal income increments. Like most other researchers, we ignore this, sticking with the simpler approximation of the tax code given by the table. This is valid if taxpayers' marginal decisions involve dollar increments larger than $50.

[4]Following convention, we refer to the Economic Growth and Tax Relief Reconciliation Act of 2001 and the subsequent Jobs and Growth Tax Relief Reconciliation Act of 2003 collectively as the "Bush Tax Cuts."

income below this kink are not required to file federal income tax returns, potentially creating a censoring problem. We avoid this issue by only examining self-employed taxpayers, as their filing threshold is $400 of self-employment income during our sample period. We further limit this sample to those taxpayers with no dependents to abstract away from tax credits related to children that effectively leave marginal incentives unchanged at this kink.

2.2 Earned Income Tax Credit

The EITC is one of the largest poverty alleviation policies (and tax expenditures) in the United States, with some 28.8 million low-income tax units receiving $68 billion dollars in 2013.[5] These figures have grown since 1996, when roughly 19.5 million tax units received $28.8 billion.[6] All low-income taxpayers between the ages of 25 and 64 are eligible, and the age restriction only applies to taxpayers with no qualifying children. The credit's schedule varies based on filing status (single or married), number of qualifying children, and tax year. Childless households may qualify for a small credit (a maximum of $496 in 2014), but the EITC is more generous for households with children. For example, a taxpayer with three qualifying children in 2014 could potentially receive a credit of $6,143.

The term "earned" in the credit's title refers to the definition of income to which the credit applies: labor and self-employment income. This definition of income does not allow for deductions, ensuring that behavioral responses to the EITC are in the form of (reported) earnings responses.

As earned income increases from zero, all households face a phase-in region, a plateau, and a phase-out region. In the phase-in region, additional earned income is subsidized at rates between 7.65% and 45% depending on the number of qualifying dependents in the household. In the plateau region, the taxpayer receives the maximum credit amount. Each additional dollar of qualifying income does not affect the credit amount. In the phase-out region, the subsidy is removed at rates between 7.65% and 21.06%, again depending on the number of qualifying dependents. This increases effective tax rates in this region, potentially discouraging earnings.[7]

We expect individuals to bunch around the first and second kinks, marking the beginning and end of the plateau region. Because it is non-convex, the kink at the end of the phase-out region should induce an absence of mass. Previous studies have not identified responses at non-convex kinks, but for many taxpayers this kink creates the largest percentage change in the net-of-tax rate. Thus, if we see a response to any non-convex kinks, we expect it here.

[5]See Eissa and Hoynes (2011) and Nichols and Rothstein (2015) for detailed discussions of the EITC.

[6]These figures are taken from the Statistic of Income's "Tax Stats" website, specifically the section on the EITC here: http://www.irs.gov/Individuals/Earned-Income-Tax-Credit-Statistics.

[7]Phaseout of the EITC occurs using the greater of earned income and adjusted gross income (AGI). In our empirical analysis we ignore this issue, assuming all taxpayers have weakly greater earned income than AGI. Our measures, therefore, likely understate responsiveness at the second EITC kink.

2.3 Child Tax Credit

The Child Tax Credit (CTC) is available to taxpayers on a per child basis, but phases out for those above certain income thresholds. As with the EITC, the CTC phases in as a function of earned income.[8] The credit amount and refundability parameters have varied since the credit's introduction in 1997. Initially the CTC was $400 per qualifying child. In 1999 it increased to $500; in 2001 and 2002 the credit was $600; and in 2003 the credit increased to its present value of $1,000 per qualifying child.

The credit creates two convex and two non-convex kinks. The first non-convex kink is the refundability threshold, which was introduced in 2001. At this threshold the portion of the credit exceeding the taxpayer's liability can be claimed by the individual, but only at a rate of 10% or 15% of earned income exceeding the threshold, depending on the year. This has no effect on households whose tax liability exceeds the credit amount. For the remaining households, however, the kink effectively decreases marginal tax rates by 10% or 15%. The threshold was $10,000 from 2001 to 2007 (indexed to inflation beginning in 2002), reduced in 2008 to $8,500, and reduced again in 2009 to $3,000 (not adjusted for inflation). The refundability rate was 10% from 2001 to 2003 and 15% after.

The first convex kink occurs at the point where the CTC has been fully refunded. After this point the credit is fully maximized, creating a plateau region. This "refundability plateau kink" is comparable to the first EITC kink, where the credit is maximized and creates a plateau. As other research has found bunching at the first EITC kink, the CTC refundability plateau is perhaps the most likely place to find a bunching response to the CTC.

The second convex kink is at the end of the credit's plateau and the beginning of the phase-out region, which is $75,000 for singles and $110,000 for married filing jointly (neither are indexed to inflation). The credit is reduced by $50 for every $1,000 in additional modified AGI, effectively increasing marginal tax rates by five percentage points. The end of the phase-out region – where the credit amount is completely eliminated – marks the second non-convex kink. At this point the taxpayer's marginal tax rate decreases ceteris paribus. In a frictionless world, we would expect an absence of mass at this point. However, given that marginal tax rates change by only five percentage points, we do not expect to find responsiveness at either the beginning or end of the CTC phaseout region.

2.4 Making Work Pay Tax Credit

The Making Work Pay Tax Credit was a refundable tax credit available to low-income and middle-income workers in 2009 and 2010. The credit was administered through a reduction in withholdings on Form W-2, and as a result many low-income individuals received the

[8] Technically it is the Additional Child Tax Credit – the refundable portion of the CTC – that phases in. See Crandall-Hollick (2013) for a detailed description of the credit and its legislative history.

credit even without filing a tax return. The credit effectively reduced the tax rate on earned income by 6.2% up to \$6,451 of earned income for singles and \$12,903 of earned income for married couples filing jointly. The maximum credit amounts were \$400 and \$800 for singles and married couples, respectively. The credit began to phase-out at a rate of roughly 2% at \$150,000 of modified AGI for married couples filing jointly and \$75,000 for all others, and was fully exhausted at \$190,000 and \$95,000, respectively.

The end of the phase-in region, beginning of the phase-out region, and end of the phase-out region all created kinks. However, we only expect responsiveness at the first kink, for two reasons. First, it is a relatively large, convex kink: 6.2 percentage points for all returns. Second, it maximizes a refundable tax credit and was a salient component of the American Recovery and Reinvestment Act of 2009. The other kinks are relatively small, and the incentives they create are easily overwhelmed by moderate optimization frictions.

3 Bunching Analysis

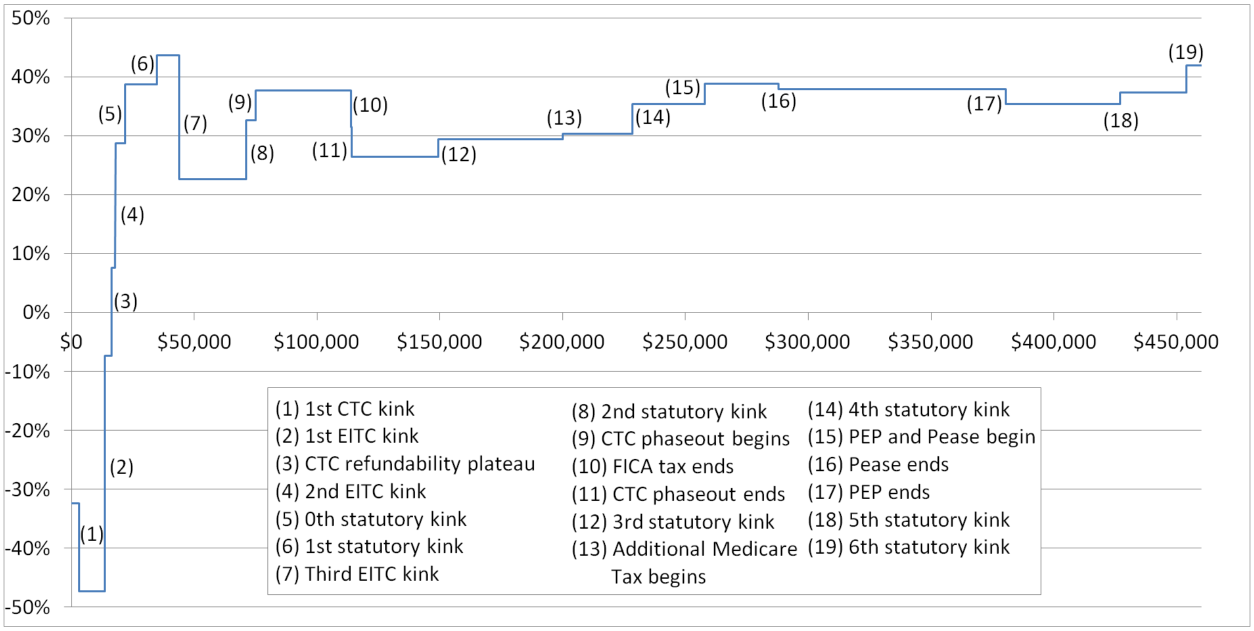
We now turn to documenting bunching patterns at the kinks described in the previous sections. Later, in Section 5, we translate the bunching measures described here into elasticities of taxable income. Figure 1 shows the location of the kinks we study for a single filer with two children in 2014. Kinks with increasing marginal tax rates are convex, while those with decreasing marginal tax rates are non-convex. The size of each kink is given in Table 2, where size is measured by the percentage change of the net-of-tax rate (NTR), which corresponds to the denominator in the definition of the ETI.[9] The five largest kinks occur at gross incomes below \$50,000, reflecting the strong distortions of the EITC and CTC. There are, however, some sizable kinks at high incomes as well. The sixth largest kink is the second statutory kink, occurring at \$71,250, where statutory rates rise from 15% to 25%. In addition, there are moderately-sized kinks at \$75,000 and \$113,700, at the beginning of the CTC phase-out and the threshold for FICA taxes, respectively.

In Section 5.1 we show that when the ETI is positive and there are no optimization frictions, all convex kinks will generate bunching. However, at most kinks there is no evidence of responsiveness in any of the years of our sample. This includes the largest kink in Figure 1, the non-convex kink at the end of the EITC phase-out region (the seventh kink in the schedule). Standard theory predicts a dip in the distribution of income near a non-convex kink. We do not observe this in the data, even for groups that are highly sensitive to other kinks.

We also see no response at most statutory kinks, including all high income kinks. In addition to exploring federal high-income kinks, we analyze the largest high-income kinks created by state tax regimes, which occur in California, Connecticut, and New Jersey. These

[9]Note that the sizes and locations of the kinks are different for taxpayers with different filing status, household size, or self-employment status.

Figure 1: Kinks faced by a single parent with two children in 2014



We assume that the taxpayer (i) only has wage income, (ii) pays no state income taxes, (iii) has $10,000 in itemized deductions, and (iv) claims the EITC and CTC. We ignore the Alternative Minimum Tax. To measure PEP kink sizes we take the most conservative approach, assuming the marginal increment to income is $2,500. See Section 6 for further discussion of this assumption. FICA refers to the Federal Insurance Contributions Act tax, which applies to earned income up to a year-specific cap. Kinks associated with the Making Work Pay Tax Credit – applicable in 2009 and 2010 – are not pictured here.

state kinks are small (less than 3 percentage point changes), but occur at income levels greatly exceeding any federal kink (up to $2 million). We find no evidence of bunching at any of these kinks during any of the years in our sample period. All of the bunching patterns we observe occur at incomes below $75,000, and the strongest patterns occur at kinks below $25,000.[10]

Our broad finding of zero responsiveness implies that taxable income is insensitive to marginal tax rates in the neighborhood of most kinks. This could be driven by several mutually compatible causes. First, gathering information about the tax schedule is costly and taxpayers may have imperfect knowledge of their local tax schedule, consistent with Chetty and Saez (2013). Second, taxpayers may not base their decisions on marginal incentives, as in Ito (2014). Third, taxpayers may know their local schedule and want to respond to marginal incentives, but may be constrained by optimization frictions such as adjustment costs or lumpy earnings opportunities. This explanation is consistent with Gelber et al. (2013), but is less convincing when deduction opportunities are present, e.g. at statutory

[10]In addition, in Appendix C we test for bunching at income eligibility thresholds for various transfer programs, including Medicaid, SNAP, and disability benefits. We find none, though annual tax return data – as opposed to data with finer temporal variation – are not well suited to analyze these programs, whose eligibility criteria are primarily monthly.

Table 2: Kinks faced by a single parent with two children in 2014, ranked by size

Kink	Gross Income	Percentage Point ΔNTR	Percent ΔNTR	Response for some group during our sample?
Third EITC kink	$47,756	+21.06	+37.41	No
First EITC kink	$13,650	−40.00	−27.15	Yes
Second EITC kink	$17,830	−21.06	−22.80	Yes
Zeroth statutory kink	$21,850	−10.00	−14.03	Yes
CTC refundability plateau	$16,333	−15.00	−13.97	Yes
Second statutory kink	$71,250	−10.00	−12.93	Yes
Beginning of CTC refundability	$3,000	+15.00	+11.33	No
Threshold for FICA taxes	$113,700	+06.20	+09.94	No
First statutory kink	$34,800	−05.00	−08.16	No
Beginning of CTC phase-out	$75,000	−05.00	−07.42	No
Sixth statutory kink	$454,050	−04.60	−07.34	No
End of CTC phase-out	$114,000	+05.00	+07.29	No
Fourth statutory kink	$228,450	−05.00	−07.18	No
Beginning of PEP and Pease	$257,800	−03.52	−05.45	No
End of PEP	$380,300	+02.53	+04.08	No
Third statutory kink	$149,400	−03.00	−04.08	No
Fifth statutory kink	$426,950	−02.00	−03.09	No
End of Pease	$287,800	+00.99	+01.62	No
Additional Medicare Tax threshold	$200,000	−00.90	−01.28	No

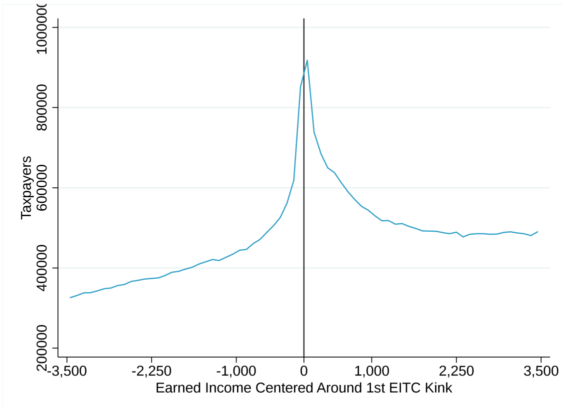
Kinks are ranked in descending size, measured by percent change in the net-of-tax rate. See the caption of Figure 1 for our assumptions.

kinks. Deductions, such as charitable giving, allow taxpayers to precisely manipulate their taxable income at the end of the year, after gross income is observed. Fourth, marginal tax rates are functions of annual income and deductions. Individuals respond to marginal tax rates throughout the year based on expectations of income and deduction activity. If income is sufficiently volatile or expectations are sufficiently imprecise, taxpayers may fail to respond to kink points. This problem is potentially compounded by the presence of multiple income earners and income types.

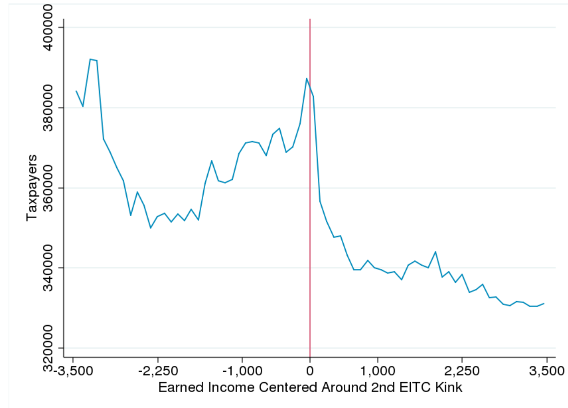
When kinks are small, the hypothesis that taxpayers ignore marginal incentives is particularly appealing. Chetty (2012) shows that ignoring many of the kinks in the tax schedule leads to utility losses of less than 1% compared to a utility-maximizing choice. In light of this, the lack of responsiveness at most middle-income and all high-income kinks is unsurprising.

Taxpayers are not universally unresponsive, however. We observe bunching at several low- and middle-income kinks. Similar to patterns documented in Saez (2010) and Chetty et al. (2013), we find sharp bunching at the first EITC kink in all years of our sample. This is where the strongest bunching occurs. We also document bunching at the zeroth statutory federal kink, consistent with Saez (2010). In addition, we provide new evidence of bunching at the second EITC kink, CTC refundability plateau kink, MWPTC kink, and the second and third statutory kinks. Responsiveness at the two EITC kinks, the CTC kink, and the

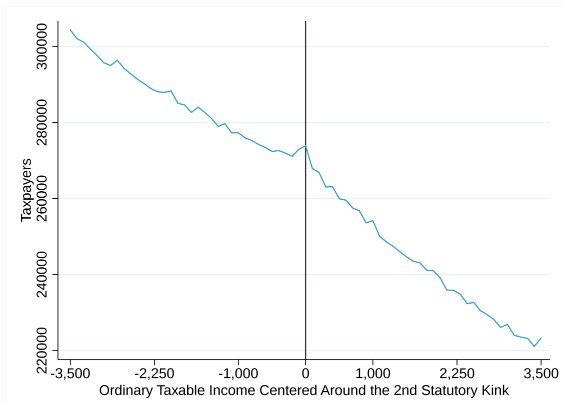
Figure 2: Bunching at four kinks



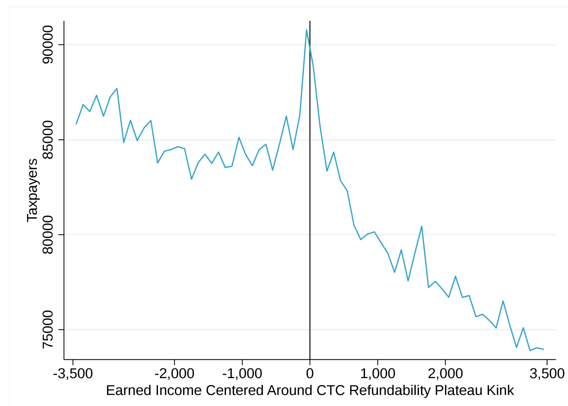
(a) 1st EITC kink



(b) 2nd EITC kink



(c) 2nd statutory kink



(d) CTC refundability plateau

Panels (a) and (b) feature all EITC-eligible filers in our sample, from 1996 to 2014 and 2002 to 2014, respectively, with the following exception. When the kinks are within \$2,000, we drop all taxpayers in (a) that respond to the second kink, and we drop all taxpayers in (b) that respond to the first kink. Panel (c) includes all taxpayers in all years of our sample. Panel (d) includes all taxpayers in our sample that have children, except those located within \$2,000 of the first or second EITC kinks, from 2004 to 2014.

second statutory kink are displayed in Figure 2 for selected years and household types. We explore these patterns in detail in the following sections.

3.1 Estimation Technique

When measuring bunching, the key issue is how taxpayers would behave in the absence of a kink. In particular, we must specify an alternative local tax schedule as well as the local distribution of income under the alternative tax schedule. We estimate this counterfactual behavior separately for two scenarios, corresponding to the two marginal tax rates (MTRs) that hold above and below the kink. Let us call t_0 the MTR that applies below the kink, and t_1 the MTR that applies above it. First, for those bunchers located below (left of) the kink, we estimate their behavior under a locally constant MTR equal to t_0. In other words, we assume their MTR continues unchanged throughout the kink region. Second, for those bunchers located above (right of) the kink, we estimate their behavior under a locally constant MTR equal to t_1, assuming their MTR also held below the kink.

In estimating these counterfactual scenarios separately, we break from the bunching analysis developed by Chetty et al. (2011). To our knowledge, all extant research that reports bunching coefficients uses their style of estimating one counterfactual distribution for bunchers on both sides of the kink. Though not always made explicit, the underlying assumption for the counterfactual tax schedule is a constant MTR equal to t_0. We estimate the two scenarios separately to strengthen the link with the theory developed in Section 5.1, which allows us to translate bunching patterns into elasticity estimates. Researchers that use the Chetty et al. (2011) bunching estimator translate their coefficients into elasticities using an infinitesimal formula that is valid only for small kinks. In contrast, our estimation technique allows us to translate bunching coefficients into elasticities using Equation 5, which is appropriate for discrete jumps in MTRs.[11] We compare our methods with those of Chetty et al. (2011) and Saez (2010) in Section 5.4. In general, our methods produce bunching coefficients (and elasticities) smaller than Saez's and larger than those of Chetty et al.

For each counterfactual scenario, we estimate the income distribution using observed data near the kink but not so close as to be affected by bunching behavior. Specifically, we group households into bins and estimate distinct linear projections on both sides of the kink. For the counterfactual scenario where the MTR is t_0, we use bins $-R, ..., -1, 0$, where bin 0 contains the kink. For the counterfactual scenario where the MTR is t_1, we use bins $0, 1, ..., R$. We call the union of these sets of bins the "bunching region."

For the counterfactual scenario where the MTR is t_0, we estimate the following equation by ordinary least squares:

$$y_j = \alpha^0 + \beta^0 z_j + \sum_{k=-W}^{0} \gamma_k^0 \cdot \mathbf{1}[j = k] + \varepsilon_j^0, \tag{1}$$

where y_j denotes the number of taxpayers in bin j, z_j denotes the income level of bin j, W denotes the number of bins in the bunching window near the kink, and ε_j^0 denotes the residual.[12] Parameters γ_k^0 capture the number of taxpayers in the bunching window unexplained by the linear prediction $(\alpha^0 + \beta^0 z_k)$. In other words, γ_k^0 measures the amount of excess mass in bin k relative to the counterfactual expectation.

[11] See Section 5.1 for details. An additional concern with the approach of Chetty et al. is that it uses the existing income distribution above the kink (i.e. under MTR t_1) to inform the counterfactual distribution under MTR t_0. Saez (2010) shows these distributions are generally not equal, nor are they directly proportional. Thus it is unclear what, if any, information the actual distribution above the kink offers when estimating the counterfactual distribution under MTR t_0.

[12] We tried including higher-order polynomial terms of z_j, but this would often over-fit the data, producing unrealistic counterfactual projections inside the bunching window.

For the counterfactual scenario where the MTR is t_1, we estimate a similar equation:

$$y_j = \alpha^1 + \beta^1 z_j + \sum_{k=0}^{W} \gamma_k^1 \cdot \mathbf{1}[j = k] + \varepsilon_j^1. \quad (2)$$

Our default parameter values, which we select by visual inspection, are a binwidth of \$100 (δ = \$100), a bunching region of 71 bins ($R = 35$), and a bunching window of 21 bins ($W = 10$).[13] Our default counterfactuals are therefore derived from the actual distribution of income between \$1,000 and \$3,500 away from each kink. Letting circumflexes denote estimated coefficients, we calculate the total number of bunchers as $\hat{B} = \sum_{k=-W}^{-1} \hat{\gamma}_i^0 + \sum_{k=1}^{W} \hat{\gamma}_i^1 + (1/2)(\hat{\gamma}_0^0 + \hat{\gamma}_0^1)$. Figure 3 graphically depicts this estimation technique for married filers near the second statutory kink in 2002. The estimated number of bunchers is simply the difference between the observed and counterfactual distributions of income inside the bunching window.

In a few instances, two kinks are too close together to perform the analysis as described. Suppose we wish to analyze kink K, but kink L lies somewhere inside K's bunching region. If taxpayers bunch at L, this can lead to unreasonable estimates for the counterfactual distributions needed for K's analysis. For this reason we do not report bunching coefficients or ETI estimates for kink K whenever (i) taxpayers bunch at some kink L, and (ii) the distance between kinks K and L is between \$1,000 and \$2,000. When the distance between the kinks is less than \$1,000, so that kink L lies within kink K's bunching window, the problem is not the estimates for K's counterfactual distributions. Instead, the difficulty is that it is hard to tell which kink bunchers are responding to. In this case, we estimate the total number of bunchers in the usual way, except we divide them into two groups. If kink K sees marginal tax rates change by Δt_K, and kink L sees marginal tax rates change by Δt_L, then we assign fraction $\Delta t_K / (\Delta t_K + \Delta t_L)$ to K, and one minus this fraction to L.

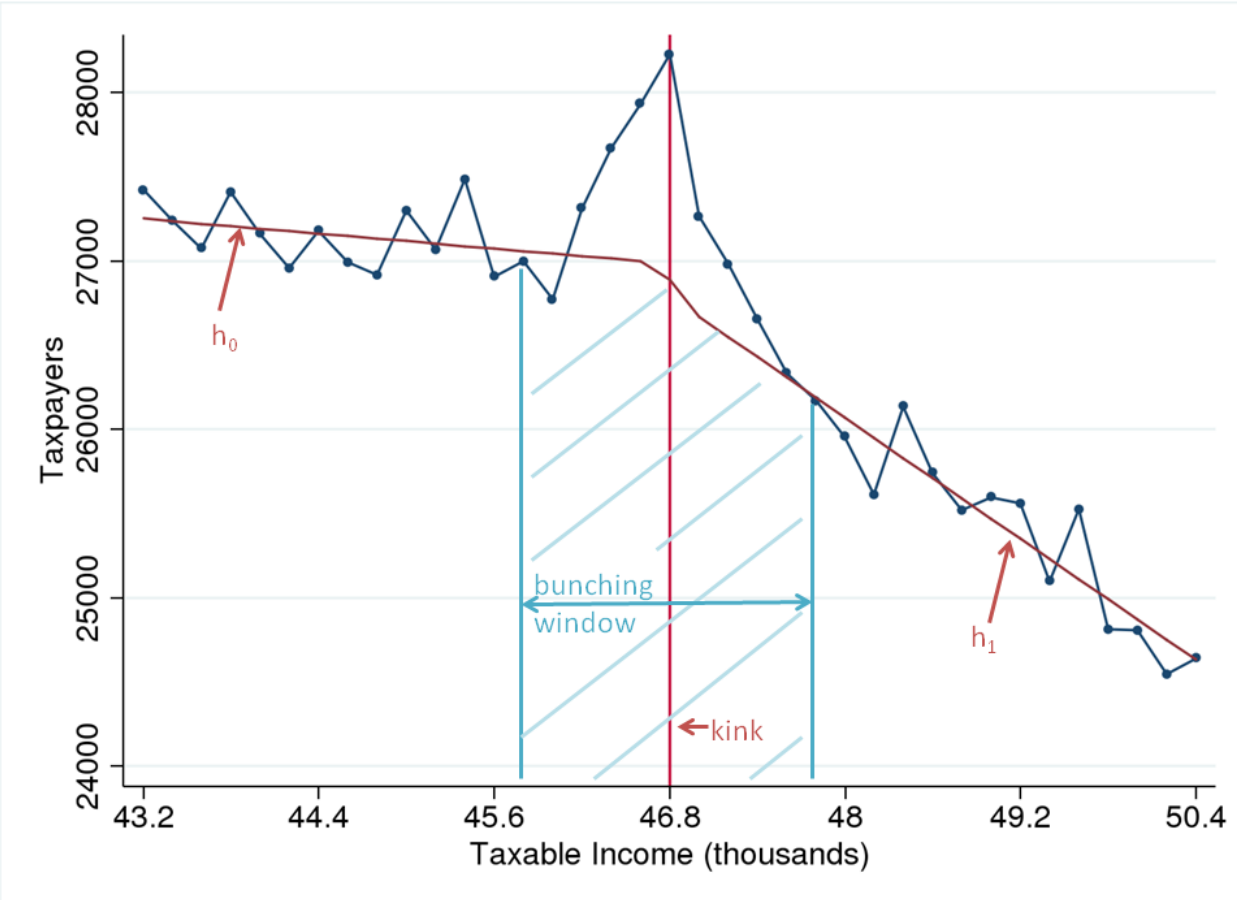
Regardless of whether other kinks are nearby, the total number of bunchers is a flawed metric for taxpayer responsiveness. All else equal, the number of bunchers will be larger when analyzing kinks affecting a larger mass of taxpayers. Hence, we report a unitless bunching coefficient $\hat{b}$ equal to $\hat{B}$ (or the fraction thereof assigned to the kink being analyzed) divided by the average number of non-bunchers in \$100 bins inside the bunching window. In other words, letting P_k denote the observed population in bin k, we define

$$\hat{b} \equiv \hat{B} / \left[\frac{\sum_{k=-W}^{W} P_k - \hat{B}}{2W+1} \cdot \frac{\$100}{\delta} \right].$$

We use a bootstrap procedure to obtain standard errors for $\hat{B}$ and $\hat{b}$ by adding randomly sampled estimated residuals (from the original regressions) to the predicted values of the

[13] We show that our results are robust to parameter choice in Appendix A.

Figure 3: Actual and estimated counterfactual distributions of income



The distribution of income is displayed for married couples filing jointly in 2002 who have no capital gains realizations. The estimation parameters are $R = 18$, $W = 5$, and $\delta = \$200$.

original regressions, repeatedly estimating $\hat{B}$ and $\hat{b}$ from the new, simulated data.[14]

3.2 Bunching Estimation Results

Though all taxpayers face incentives to bunch at the convex kinks of Figure 2, some taxpayers are more responsive to these incentives than others. To compare bunching patterns across groups, Table 3 presents estimated bunching coefficients at four kinks where we find a response, using 2014 data. In general, the first EITC kink elicits the largest response. It sees the largest bunching coefficient, 19.97, corresponding to single, self-employed individuals. The bunching coefficient indicates the mass of bunchers is approximately 20 times the average number of non-bunchers in \$100 bins inside the bunching window. This implies 49%

[14]We thank Raj Chetty, John Friedman, Tore Olsen, and Luigi Pistaferri for public provision of a Stata program designed specifically to implement their estimation technique. Our code builds directly on theirs, and we plan to make our code publicly available in the near future.

Table 3: Bunching coefficients calculated at four kinks in 2014

	1st EITC kink	2nd EITC kink	2nd statutory kink	CTC refundability plateau
Single, wage earners	3.45 (0.43) [N=535,000]	0.28 (0.15) [N=602,900]	0.34 (0.08) [N=556,800]	0.93 (0.08) [N=339,400]
Single, self-employed	19.97 (0.49) [N=481,100]	1.35 (0.66) [N=266,500]	0.91 (0.43) [N=26,900]	13.65 (0.44) [N=182,900]
Married filing jointly, wage earners	0.58 (0.36) [N=104,000]	0.17 (0.20) [N=184,500]	0.32 (0.14) [N=218,800]	0.26 (0.10) [N=132,500]
Married filing jointly, self-employed	11.37 (0.46) [N=112,200]	1.53 (0.34) [N=88,300]	0.97 (0.34) [N=30,900]	1.76 (0.33) [N=112,200]
Married filing separately, wage earners	—	—	0.11 (0.36) [N=27,100]	0.61 (1.06) [N=4,200]
Married filing separately, self-employed	—	—	-2.30 (1.15) [N=1,400]	6.72 (4.11) [N=700]

Bunching coefficients are reported for various household types, with standard errors in parentheses. The number of taxpayers in the bunching region (rounded to the nearest hundred) is presented in brackets. Wage earners are those with positive wage income and zero self-employment income. The self-employed are those with positive self-employment income. Single status includes "head of household" filers. Estimates are omitted when the kink is between $1,000 and $2,000 away from another kink where taxpayers bunch, as discussed in Section 3.1. Married filers who file separately are ineligible for the EITC and thus are excluded from its analysis. The income definition is earned income for the EITC and CTC kinks, and taxable income for the statutory kinks. The first EITC kink and CTC refundability kink for households with one child in 2014 are near $10,000, which induces round-number bunching for married households.

of single, self-employed taxpayers in the bunching window (i.e. within $1,000 of the kink) are there because of the changing marginal incentives at the kink.[15] According to the theory developed in Section 5.1, these taxpayers desire income greater than the kink when facing the low tax rate, and income less than the kink when facing the high tax rate.[16]

Contrasting with prior research, we observe wage earners (i.e. those without self-employment income) bunching at many kinks in recent years. Single wage earners, in particular, ex-

[15] To see this, let X denote the average number of non-bunchers in $100 bins inside the bunching window. Then $\hat{B} = \hat{b} \cdot X$, and the fraction of bunchers among the population within $1,000 of the kink is $\hat{b} \cdot X/(21X + \hat{b} \cdot X) = \hat{b}/(21 + \hat{b}) \approx 0.49$ for $\hat{b} = 19.97$.

[16] Other groups and other kinks exhibit smaller, statistically insignificant bunching coefficients that are occasionally negative. Negative numbers imply the kink causes *less* mass to locate near the kink. This is plausible only if the ETI is negative, which has little empirical support. As none of the negative coefficients are statistically distinguishable from zero, we interpret them as evidence that taxpayers are not responding to the kink.

hibit statistically significant bunching coefficients at the first EITC kink, second EITC kink, CTC refundability plateau, and second statutory kink beginning in the early to mid-2000s. Married-filing-separately wage earners bunch at the third statutory kink, but married-filing-jointly wage earners do not exhibit statistically significant bunching (absent round-number bunching, as is the case in 2014) at any kink in any year. However, as shown in Table 3, the magnitude of responses by wage-earners is much smaller than those with self employment income. All of the statistically significant bunching coefficients for the self-employed are larger than those of their wage-earning counterparts in 2014, and most of these differences are themselves statistically significant.

It is unclear, though, whether the greater responsiveness of the self-employed can be attributed to greater real labor responses or higher rates of tax evasion. One the one hand, manipulating earned income is inherently easier when one is both the employer and employee, so the self-employed likely exhibit larger real labor responses. Moreover, self-employment income is easily adjusted at the end of the tax year by incurring business expenses. On the other hand, unlike wage income, self-employment income is not subject to third-party reporting and is therefore easier to hide from tax authorities (or inflate when income is subsidized). Thus, the self-employed likely exhibit larger tax evasion responses as well.

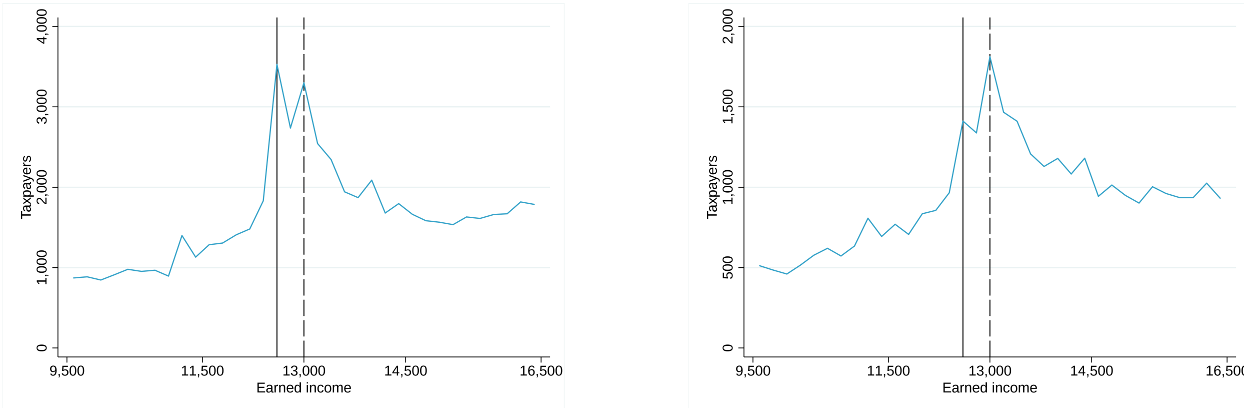
Because wage earnings are reported by third parties, they are thought to be more reliable indicators of real economic activity (Slemrod, 2007). Further, because EITC kinks are determined by gross income, adjustments or itemized deductions are irrelevant. Thus, at first glance, our results seem to indicate wage-earner bunching is generated by real labor supply (or demand) responses. However, in Section 4.1 we show that bunching in taxpayer-reported wages does not manifest in employer-reported wages. This suggests that evasion is the primary channel for responsiveness among this group.

In addition to the four kinks of Table 3, we document statistically significant but economically insignificant bunching at the zeroth and third kinks in the statutory schedule.[17] The former coincides with the filing threshold for wage earners, and therefore suffers from censoring in the data. In addition, taxpayers with children do not see a change in marginal incentives at this kink because the non-refundable portion of the Child Tax Credit immediately eliminates their liability. Thus, we can cleanly measure bunching here only for childless, self-employed taxpayers. The bunching coefficients we observe are more stable over time for married taxpayers than singles, but both groups produce small (0.2 to 0.9), statistically significant bunching in roughly half of the years in the sample.

At the third statutory kink we face no inherent sample restrictions, yet we find statisti-

[17] We primarily observe bunching through earned income, as opposed to adjusted gross income or taxable income. The amount of information needed to calculate adjusted gross income (total income net of adjustments) and taxable income (adjusted gross income net of deductions) is larger than that needed to calculate earned income. For this reason, annual taxable income forecasts necessary to bunch at statutory kinks are strictly noisier than analogous earned income forecasts necessary to bunch at EITC kinks or the CTC refundability plateau.

Figure 4: Responses to the Making Work Pay Tax Credit (2010)



(a) Married, self-employed, 2 children (b) Married, self-employed, 3+ children

Income distributions are displayed for married, self-employed taxpayers with two or more children. The solid vertical line denotes the first EITC kink, while the dashed line denotes the MWPTC kink. The self-employed are those with positive self-employment income. Married taxpayers who file separately from their spouse are excluded. The figures look very similar in 2009.

cally significant bunching only for married taxpayers who file separately from their spouse. Other groups, including married couples filing jointly, fail to respond during any years of our sample. At this kink and all kinks above it, we analyze the universe of tax returns because the mass in the seven states comprising our main sample is insufficient to distinguish bunching from noise. Among married taxpayers who file separately, we find stronger responsiveness among wage earners than the self-employed. The former bunch in most years of our sample and see bunching coefficients up to 1.33 (with a standard error of 0.22), while the latter bunch in only a handful of years.

Finally, we see responsiveness to the temporary Making Work Pay Tax Credit by married, self-employed taxpayers with two or more children. In 2009 and 2010, this credit created a convex kink of roughly six percentage points at the end of its phase-in. When fully phased in, the credit delivered \$800 to married taxpayers filing jointly. Unfortunately, given its proximity to the first EITC kink for the responsive households, our bunching and elasticity estimates require assumptions about which bunchers near the kink are assigned to the EITC and MWPTC kinks. Nonetheless, Figure 4 clearly shows separate responses to each kink, indicating some taxpayers were responding specifically to the MWPTC.

3.2.1 Evolution of bunching patterns

One of the most striking features of the bunching patterns we observe is their evolution over time. In 1996, substantial bunching occurs only at the first EITC kink and only for those with self-employment income. Bunching coefficients in 1996 for the four groups included in Table 2 do not exceed six. By 2014, there is substantial bunching at three kinks, wage earners are bunching at several kinks, and bunching coefficients at the first EITC kink reach as high as twenty.

Figure 5 displays this temporal variation. Several findings emerge from studying the figure. First, bunching coefficients are generally increasing over time. Second, in the aggregate, substantial wage-earner bunching (with coefficients exceeding unity) is nonexistent until emerging in 2010 at the first EITC kink. This is why we find wage-earner bunching where Saez (2010) and Chetty et al. (2013) found none, as their samples end in 2004 and 2009, respectively. Third, bunching by the self-employed in response to the CTC emerges in 2009, and the CTC's coefficients for singles rise rapidly to levels previously achieved only at the first EITC kink. This is especially noteworthy given that the CTC kink changes effective marginal tax rates by 15 percentage points, whereas the first EITC kink sees effective marginal tax rates rise by up to 45 percentage points.

Though responsiveness is generally increasing, in a few instances bunching coefficients fall quite dramatically. For example, the second EITC kink sees a drop in bunching following a peak in 2011. The CTC refundability plateau also sees bunching coefficients fall in recent years. In addition, the first EITC kink sees less bunching by all groups in 2008 at the onset of the Great Recession. An intriguing hypothesis is that economic downturns may influence taxpayers' ability to bunch. Indeed, in Ireland bunching measures have been shown to be correlated with the business cycle (Hargarden, 2015). However, in our data there is no year in which *all* coefficients fall. When bunching falls at one kink, it always rises at another. In Section 4.2, we argue that these patterns are driven not by macroeconomic fluctuations, but rather by small adjustments to the tax code that shift the overall refund-maximizing point from one kink to another.

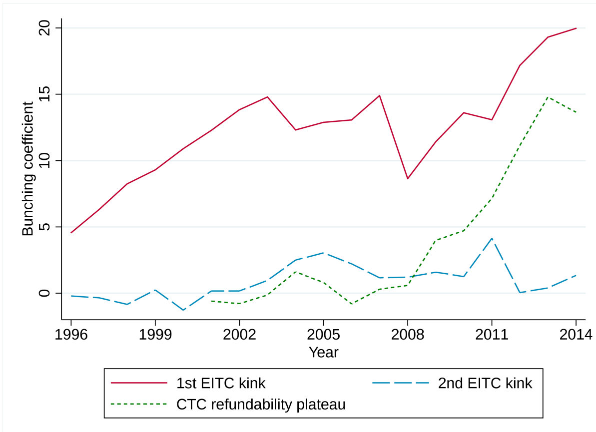
4 Explanations for Bunching Variation

We have documented substantial bunching at many large kinks and some small kinks in the effective marginal tax rate schedule. These responses indicate taxpayers have sophisticated knowledge of the tax code as well as the ability and willingness to precisely control their reported incomes. However, we have said little about what form responsiveness takes. There are four general avenues of response to labor income taxation: labor supply, labor demand, deductions, and tax evasion. In this section, we investigate these channels. In addition, we show that bunching responses are significantly more likely to occur at the overall refund-maximizing point in the schedule.

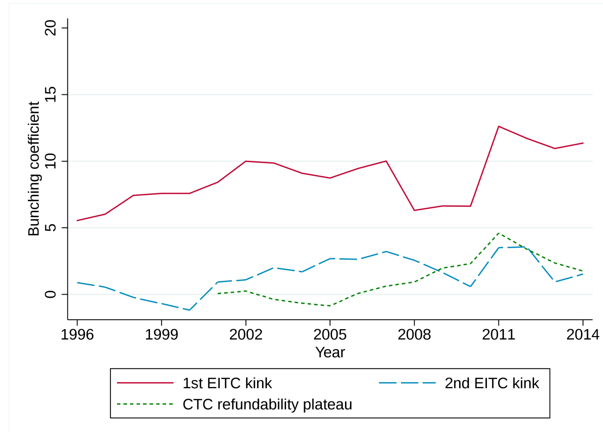
4.1 Wage-Earner Bunching: Tax Evasion?

It is difficult to disentangle the various mechanisms available for bunching when taxpayers are self-employed. As these taxpayers are both employer and employee, labor supply and

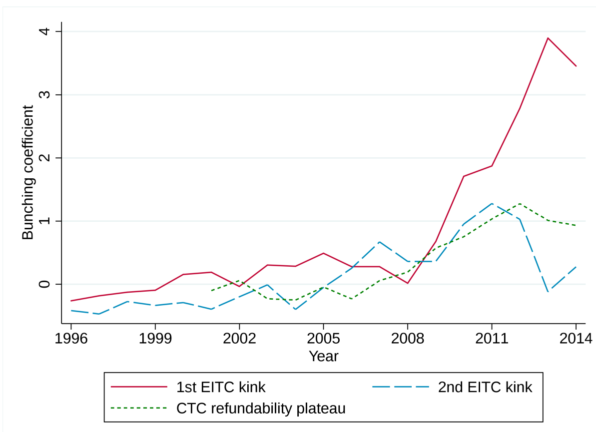
Figure 5: Bunching over time at three kinks



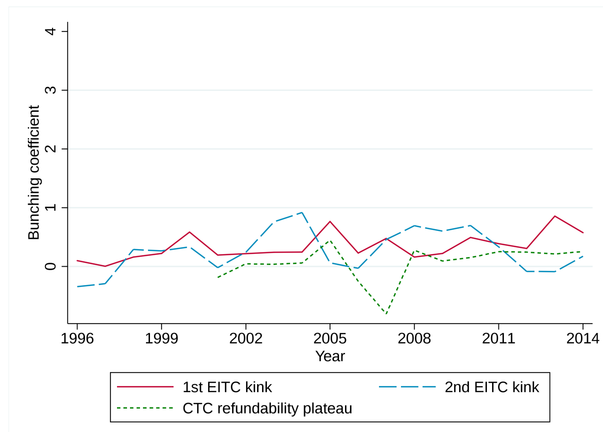
(a) Single, self-employed



(b) Married, self-employed



(c) Single, wage earners



(d) Married, wage earners

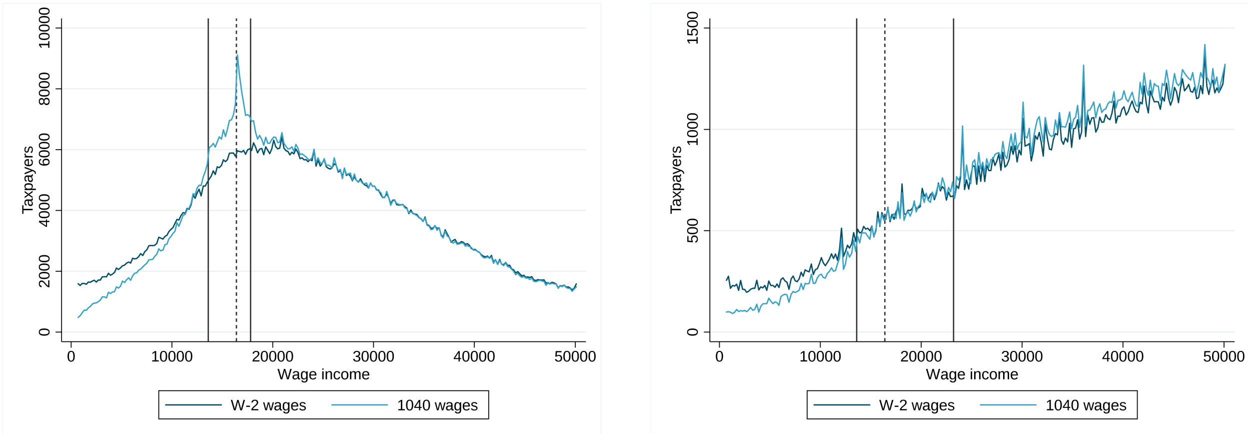
Trends in bunching coefficients at three kinks are displayed. Wage earners are those with positive wage income and zero self-employment income. The self-employed are those with positive self-employment income. Single status includes "head of household" filers. For the Child Tax Credit coefficients, childless taxpayers are excluded.

demand responses are conflated. And because the tax schedule is defined with respect to *net* self-employment income, it is tough to distinguish income responses from deduction responses. If a taxpayer reports \$15,000 in income and \$5,000 in expenses in order to bunch at a \$10,000 kink, we cannot know whether she first chose income and then adjusted expenses to land near the kink, or vice versa. Further, because net self-employment income is not subject to third-party reporting, it is difficult to detect tax evasion responses.[18]

We have much more traction, however, with analysis of pure wage earners. For this group, the taxable-income definition at the kinks where we observe large responses (created

[18] Audit data can help to identify evasion responses. We are unaware of any studies of bunching in the U.S. that use audit data, however, in a tax audit experiment in Denmark, Kleven et al. (2011) find that around half of the bunching response of the self-employed is eliminated post-audit.

Figure 6: Taxpayer- and employer-reported wage income (2014)



(a) Single, two children

(b) Married, two children

Distributions of taxpayer-reported and employer-reported wages from $600 to $50,000 are displayed for those taxpayers for which we observe a Form W-2. The solid vertical bars mark the first and second EITC kinks, and the dashed vertical bar denotes the CTC refundability plateau. Wage earners are those with positive wage income and zero self-employment income. Single status includes "head of household" filers. Taxpayer-reported wages are derived from Form 1040, whereas employer-reported wages are dervied from Form W-2.

by the EITC and CTC) does not allow for deductions, immediately ruling out this channel. In addition, all employers that pay an employee more than more than $600 in a given year are required to file form W-2 with the IRS.[19] Because of this third-party reporting, wage income is thought to be subject to significantly less evasion than self-employment income (Slemrod, 2007). Moreover, we can test for systematic mismatches between taxpayer-reported and employer-reported wage income.

Figure 6 displays the distributions of taxpayer- and employer-reported wage income for those taxpayers for which we observe W-2 wages. We highlight taxpayers with two children in 2014, although the patterns are broadly similar for singles with varying number of kids in other recent years. Looking first at the distributions for singles in panel (a), we see that above roughly $25,000 in income, the two distributions are virtually identical. This is consistent with prior research that shows taxpayer-reported wages are highly reliable. However, below $25,000 a different picture emerges. There is a significant amount of extra mass in the distribution of taxpayer-reported wages in the EITC plateau region (between the two solid vertical bars). This extra mass exhibits sharp bunching precisely at the CTC refundability plateau (marked by the dashed bar), however, employer-reported wages show no indication of bunching here. This rules out the possibility of a labor supply response to kinks and indicates this group's bunching is purely a reporting phenomenon.

The evidence is different for the married taxpayers of panel (b). There is no excess mass in

[19] If any taxes are withheld, even employees earning less than $600 must have a W-2 filed. Household employees have a looser threshold: $1,900 in 2014.

the EITC plateau region, and no bunching in either taxpayer-reported or employer-reported wages. We do see more mass in taxpayer-reported wages at incomes above $25,000, but this is unlikely to be driven by taxpayer reporting behavior. Marginal tax rates are positive in this region, so taxpayers have no strategic reason to fabricate income in this region. A more mundane explanation is that we are unable to perfectly match primary filers' Forms 1040 with secondary filers' Forms W-2. In other words, for a number of dual-earner households we may be observing only one spouse's W-2 wage income. Unfortunately, we cannot test this directly as taxpayers do not report whether one or two earners' wages are included on the Form 1040.

We caution, however, that a mismatch between taxpayer- and employer-reported earnings does not necessarily reflect tax evasion. We do not observe taxpayers reporting *less* income than their W-2s indicate, which is transparently illegal (provided the W-2s are accurate). Relative to the distribution of taxpayer-reported incomes, we see significantly more mass in employer-reported incomes below $10,000. Thus, bunching among wage earners is driven exclusively by taxpayers reporting *more* than their employer-reported wages. This is legal if taxpayers have unreported tip income or additional earnings below the W-2 filing threshold (generally $600, but $1,900 for household employees) from other employers. However, it strains credibility to suppose that *only* taxpayers in these situations are able to legally manipulate their incomes in response to kinks. Instead, we view the evidence as suggestive of tax evasion in the form of fabricated earnings, under-reported additional income, or self-employment income mis-characterized as wage income.[20]

4.2 Bunching for Maximum Refunds

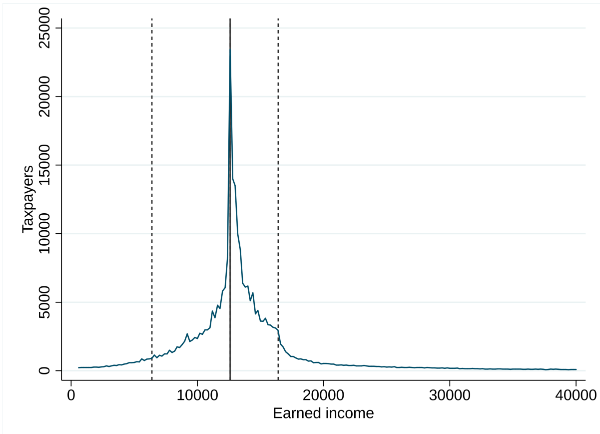
In Section 3.2.1, we demonstrated that bunching patterns are not constant over time. Here we delve deeper into variation in bunching patterns, seeking to explain changes in bunching coefficients.

Bunching at kinks changes at both the extensive and intensive margins. For example, bunching at the CTC refundability plateau and the second EITC kink is nonexistent until emerging late in the 2000s. Similarly, in many instances bunching coefficients rise at one kink at the same time another kink sees coefficients fall. We hypothesize this is due to refund-maximizing behavior.

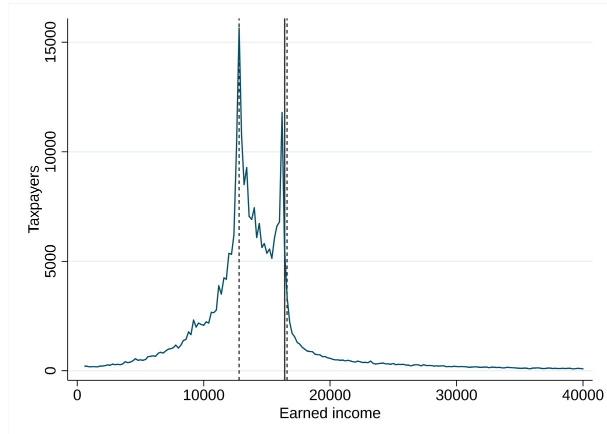
Fortunately, we can test this hypothesis, as different groups see different kinks mark the refund-maximizing point in the schedule. Moreover, within the same groups the refund-maximizing point sometimes changes from one year to another due to relatively small changes in the tax code. To see this, consider Figure 7, which displays income distributions for single, self-employed taxpayers with two children from 2010 to 2013. In all panels, the refund-

[20]Reporting self-employment income as wage income evades the Social Security and Medicare taxes that mirror payroll taxes on wage income.

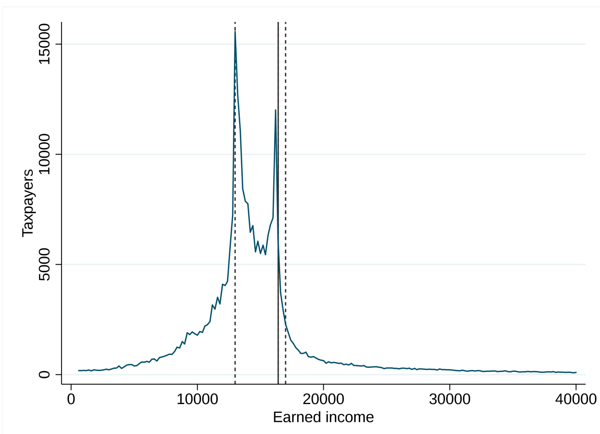
Figure 7: Tracking the refund-maximizing kink: singles



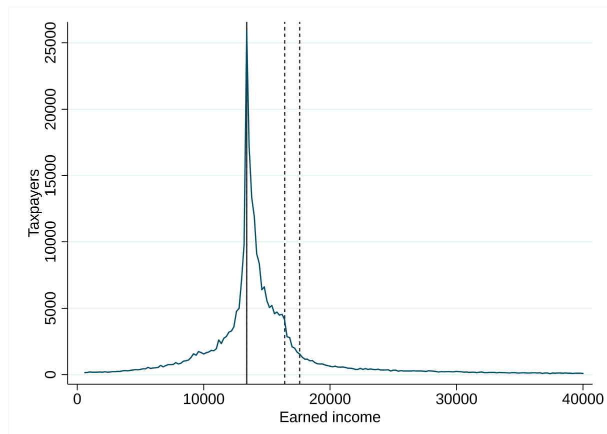
(a) Single, self-employed, two children (2010)



(b) Single, self-employed, two children (2011)



(c) Single, self-employed, two children (2012)



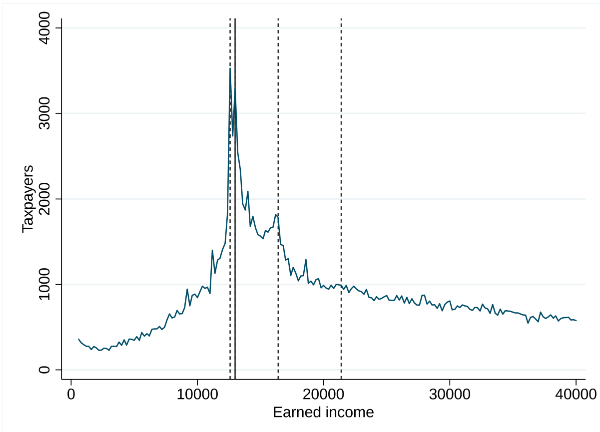
(d) Single, self-employed, two children (2013)

Income distributions are displayed for single, self-employed taxpayers with two children from 2010 to 2013. The solid vertical line denotes the refund-maximizing kink, while dashed vertical lines denote other kinks where taxpayers respond. Single status includes "head-of-household" filers. The self-employed are those with positive self-employment income.

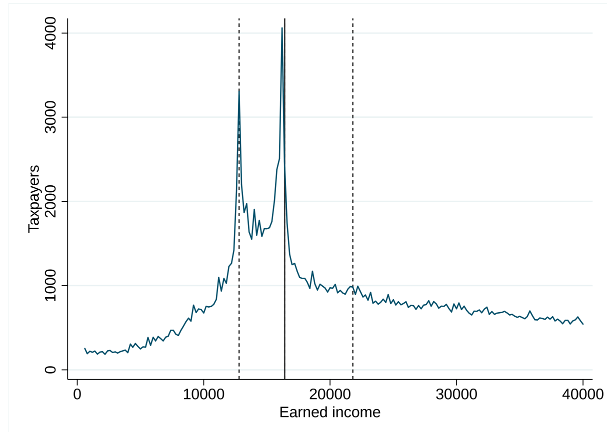
maximizing kink is marked by a solid vertical line. In 2010, the first EITC kink maximizes this group's refund. However, in 2011 the CTC refundability plateau moves to just below the second EITC kink and becomes the refund maximizer due to a temporary 2 percentage point reduction in payroll taxes. Revealing their sophistication, taxpayers immediately respond. A large mass shifts from the first EITC kink to the CTC refundability plateau. Bunching coefficients at the first EITC kink fall from 15.8 to 9.6, while at the CTC kink they rise from 2.1 to 4.8, and at the nearby second EITC kink they rise from 4.5 to 7.1.[21] The distribution of income looks similar in 2012, as the CTC kink remains the refund-maximizer. However, in 2013, the payroll tax holiday ends and the first EITC kink reclaims its status as the refund-

[21] Recall that when two kinks are this close, we assign the total mass to each kink in proportion to the size of the kink. See Section 3.1 for further detail.

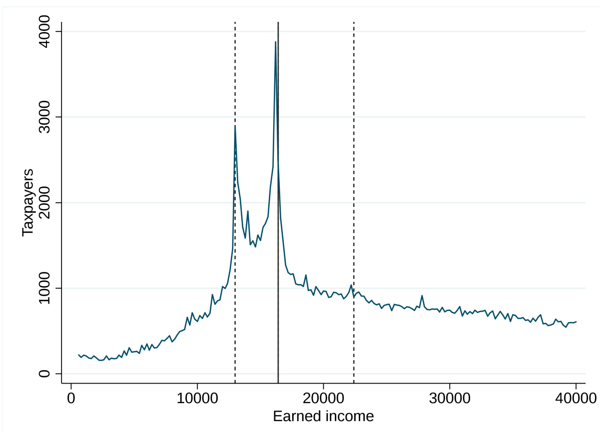
Figure 8: Tracking the refund-maximizing kink: married couples



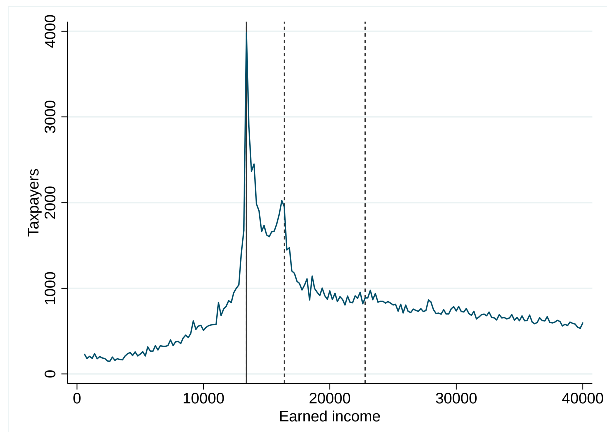
(a) Married, self-employed, two children (2010)



(b) Married, self-employed, two children (2011)



(c) Married, self-employed, two children (2012)



(d) Married, self-employed, two children (2013)

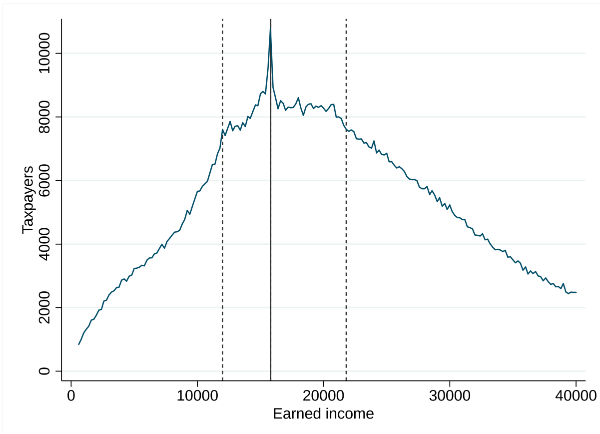
Income distributions are displayed for married, self-employed taxpayers with two children from 2010 to 2013. The solid vertical line denotes the refund-maximizing kink, while dashed vertical lines denote other kinks where taxpayers respond. Married couples who file separately are excluded. The self-employed are those with positive self-employment income.

maximizer. Once again, taxpayers immediately respond by shifting to the new optimum; bunching plummets at the CTC kink while it soars at the first EITC kink.

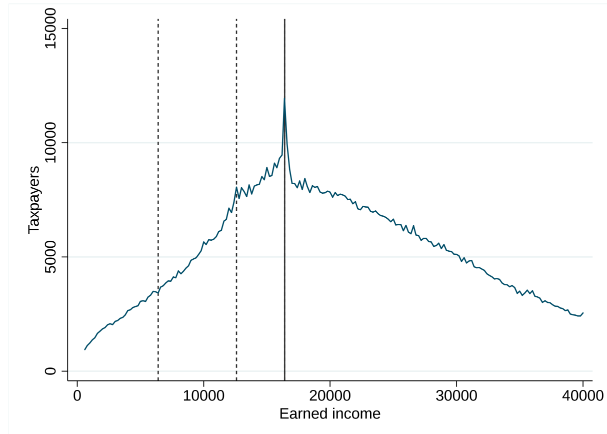
We see refund-maximum tracking among married couples as well. Figure 8 displays income distributions for the married counterparts to Figure 7. They, too, see the refund-maximizing kink shift from the first EITC kink to the CTC refundability plateau in 2011 and then back to the first EITC kink in 2013. Like their single counterparts, they respond immediately to each of these shifts.

We also observe sophisticated refund-maximizing behavior among single wage earners. Figure 9 shows this for single wage-earners with two children. In 2008, this group's refund-maximizing kink is the second EITC kink, and that is where the only significant bunching occurs. Then, from 2009 to 2011 the second EITC kink is in essentially the same location as

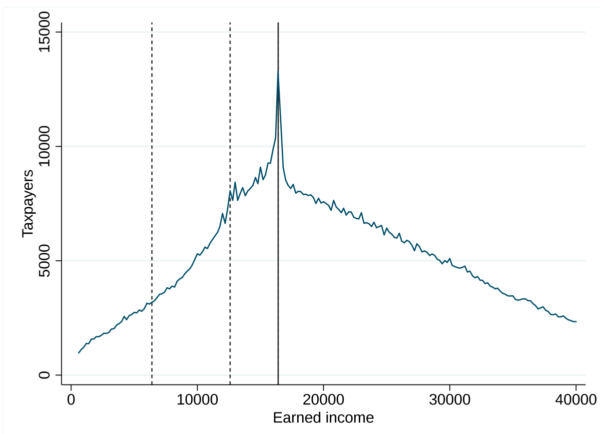
Figure 9: Tracking the refund-maximizing kink: single wage-earners



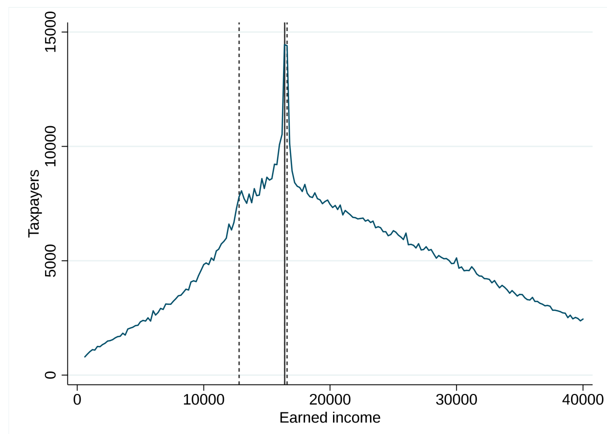
(a) Single, wage-earners, two children (2008)



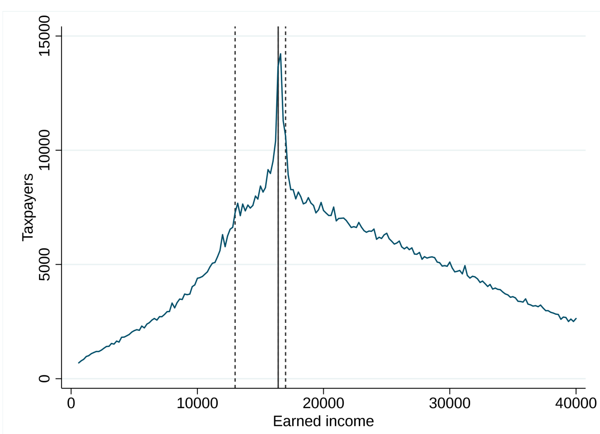
(b) Single, wage-earners, two children (2009)



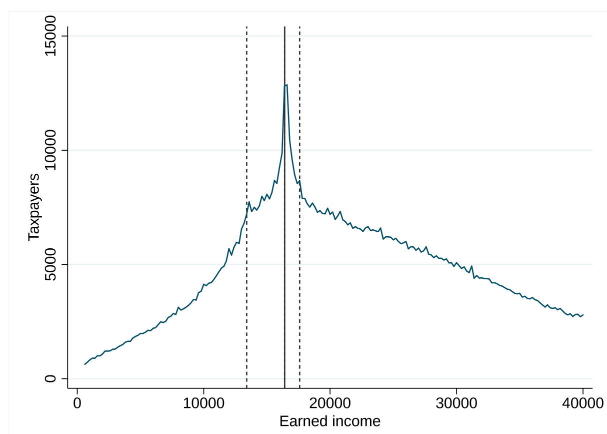
(c) Single, wage-earners, two children (2010)



(d) Single, wage-earners, two children (2011)



(e) Single, wage-earners, two children (2012)



(f) Single, wage-earners, two children (2013)

Income distributions are displayed for single, wage-earning taxpayers with two children from 2008 to 2013. The solid vertical line denotes the refund-maximizing kink, while dashed vertical lines denote other kinks where taxpayers respond. Single status includes “head-of-household” filers. Wage earners are those with zero self-employment income.

the CTC refundability plateau, and taxpayers continue to bunch there. However, in 2012 the two kinks separate, as the second EITC kink rises due to inflation indexing, leaving behind the CTC kink as the new, unique refund-maximizer. In 2013 and 2014 (not pictured), the kinks continue to separate, and in all three years bunching clusters sharply around with CTC kink, with taxpayers essentially ignoring the second EITC kink.

5 Observed Elasticities of Taxable Income

In this section we translate bunching patterns into estimates of elasticities of taxable income (ETIs). ETIs capture the degree to which taxable income responds to its marginal tax rate, and incorporate individuals' choices of work hours, deductions, and levels of tax evasion. They are necessary statistics for revenue and welfare analyses of income taxation (Chetty, 2009), and the mix of possible responses they capture – from reporting phenomena to "real" economic activities – make them distinct from labor supply elasticities.

5.1 Bunching Theory

Here we show that bunching at convex kink points is a straightforward implication of positive ETIs. We assume throughout that taxpayers (i) are fully aware of the tax schedule, and (ii) can manipulate earnings with arbitrary precision.[22] These assumptions are unlikely to hold in practice, but are useful for calculating lower bounds on the "structural" elasticities that govern frictionless responses. Intuitively, the bunching patterns of Section 3 are attenuated by optimization frictions because some taxpayers that would bunch in a frictionless world do not bunch in reality. In general, this means observed bunching patterns will be weaker than structural elasticities would imply.

Under the strong assumptions above, we map observed bunching patterns to elasticities that we call "observed" ETIs, borrowing nomenclature from Chetty et al. (2011). In the analysis below, the mapping between bunching patterns and elasticities is strictly monotone, such that the diluted bunching patterns of Section 3 translate to observed elasticities strictly less than their structural counterparts. Thus the ETI estimates we report represent lower bounds on the structural elasticities that inform welfare analysis of taxation.

Let $e \equiv \frac{dz}{d(1-t)}\frac{1-t}{z}$ denote the elasticity of taxable income (z) with respect to the net-of-tax rate $(1-t)$, where t denotes the marginal tax rate.[23] Suppose this elasticity is constant in a neighborhood around kink point z^*. Under this assumption, and using the definition of

[22]Given a positive ETI, bunching can arise under imperfect optimization as well. We merely need some fraction of taxpayers to be aware of the tax schedule, with sufficiently small optimization frictions (e.g. small adjustment costs).

[23]Technically, e is the uncompensated elasticity of taxable income. In the absence of income effects, it can also be interpreted as a compensated elasticity.

the ETI, we solve the differential equation $e \cdot \frac{d(1-t)}{1-t} = \frac{dz}{z}$. The result is that income takes the functional form

$$z = n \cdot (1-t)^e \tag{3}$$

locally around kink z^* for some n, which we call potential income.[24] This solution is identical to that which emerges from the structural assumptions of Saez (2010). In that setting, agents also choose taxable income according to equation (3), but this comes from maximizing a quasi-linear, iso-elastic utility function under a standard budget constraint. Here we have shown that the only assumption embedded in the structural approach necessary for income to take this functional form is a constant elasticity in the neighborhood of the kink. The rest of the trappings of the structural approach are superfluous (and therefore benign).[25]

With our empirical approach in mind, we now assume earners face the same local tax schedule and have the same elasticity, $e > 0$.[26] In this case, the only parameter that creates dispersion in income is potential income (n), which we take to be distributed according to a smooth, atomless density $f(\cdot)$, with distribution function $F(\cdot)$. Potential income is assumed to be independent of the tax schedule. Thus, we rule out general equilibrium effects that might arise if the tax schedule indirectly affected the wage schedule available to taxpayers through changes in aggregate labor supply.

Suppose that the marginal tax rate is t_0 up to kink z^*, and t_1 thereafter. Here we analyze convex kinks, such that $t_0 < t_1$. In this case, the space of potential income can be divided into three regions. Those with low potential income choose income below the kink; those with high potential income choose income above the kink; and those with moderate potential income locate precisely at the kink. Specifically, define $n_j \equiv z^*/(1-t_j)^e$ for $j = 0, 1$. Then $z = n \cdot (1-t_0)^e$ for $n \leq n_0$ and $z = n \cdot (1-t_1)^e$ for $n \geq n_1$. The remaining mass of earners with $n \in (n_0, n_1)$ are unable to achieve their desired solution according to (3). When facing rate t_0, they desire income greater than z^*, but when facing rate t_1, they want less income than z^*. Consequently, they bunch precisely at the kink (i.e. $z = z^*$). Their mass is given by

$$B \equiv \int_{n_0}^{n_1} f(n)dn \approx (n_1 - n_0)\frac{f(n_0) + f(n_1)}{2}. \tag{4}$$

This expression represents the mass of bunchers assuming all taxpayers share the same elasticity. This will be useful later when we translate bunching patterns to elasticity estimates.

[24]Potential income is the level of income that would obtain absent taxes (i.e. when $t = 0$).

[25]In addition to yielding solution (3), the structural approach assumes away income effects and therefore assigns e the dual roles of the compensated and uncompensated ETI. As there is some evidence of trivially small income effects (Gruber and Saez, 2002; Bastani and Selin, 2014), we follow the structural approach in interpreting our estimates for ETIs as compensated elasticities.

[26]Though the constant elasticity assumption is ubiquitous in the literature estimating ETIs, our evidence indicates it is unlikely to hold when estimating population parameters. It is more plausible in our case, as we merely assume the elasticity is locally constant around each kink for a given household type in a given year.

However, homogeneity in elasticities is not needed to imply the existence of a bunching response. Bunching would arise under heterogeneous elasticities provided they are positive and locally constant, assuming taxpayers respond frictionlessly to fully salient tax schedules.

As it stands, equation (4) is not estimable because we do not observe the distribution of potential income. However, we do observe the distributions of income under marginal tax rates t_0 and t_1, respectively, and these distributions are related to the distribution of potential income. Let H_j denote the distribution function and h_j the density over income that would arise if the tax rate were t_j both above and below the kink. From equation (3), $H_j(z) = F(z/(1-t_j)^e)$, so that $h_j(z) = f(z/(1-t_j)^e)/(1-t_j)^e$. This implies $f(n_j) = h_j(z^*) \cdot (1-t_j)^e$. Applying this result to equation (4), with a bit of algebraic manipulation, gives

$$B \approx \frac{z^*}{2}\left(h_0(z^*)\left[\left(\frac{1-t_0}{1-t_1}\right)^e - 1\right] + h_1(z^*)\left[1 - \left(\frac{1-t_1}{1-t_0}\right)^e\right]\right), \tag{5}$$

which establishes the relationship between the mass of bunchers and the counterfactual distributions of income under marginal tax rates t_0 and t_1.[27] This equation is mathematically equivalent to Saez's Equation (5), but here we have derived it directly from the definition of the elasticity of taxable income, without imposing structural assumptions within a framework of utility maximization.

Our Equation (5) allows us to translate bunching patterns into observed elasticities. Tax rates t_0, t_1, and kink point z^* are directly observable in the tax code. Thus, we can identify an estimate for e given estimates for parameters B, $h_0(z^*)$, and $h_1(z^*)$. In Section 3.1, we estimated these using predicted counterfactual distributions. With these estimates in hand, we calculate observed elasticities for various groups, using the delta method to obtain standard errors. These estimates are discussed in the next section.

5.2 Elasticity Estimates

Tables 4 and 5 present observed elasticities with standard errors in parentheses. As in Section 3, we present selected results for the year 2014. We estimate elasticities separately for groups with distinct kink locations and/or marginal tax rates because these parameters enter directly into Equation (5). We calculate marginal tax rates taking into account the federal tax schedule, payroll taxes (half of which are assumed to fall on the taxpayer), and the EITC, CTC, and MWPTC schedules. When calculating effective marginal tax rates for the CTC, we assume all taxpayers qualify for the EITC.

We obtain a range of values from statistical zeros to elasticities as large as 1.50.[28] The

[27] Note that if potential income has a linear distribution function (e.g. if it is distributed uniformly) then Equations (4) and (5) hold with precise equality.

[28] The observed elasticity of 1.50 corresponds to single, self-employed taxpayers with one child bunching around the first EITC kink in 2007.

Table 4: Observed elasticities at the first two EITC kinks in 2014

	First EITC kink			
	0 children	1 child	2 children	3+ children
Single, wage earners	—	0.10 (0.01) [N=308,400]	0.08 (0.03) [N=177,900]	0.04 (0.01) [N=56,900]
Single, self-employed	—	0.95 (0.03) [N=221,800]	0.35 (0.02) [N=188,100]	0.29 (0.02) [N=67,200]
Married, wage earners	0.31 (0.13) [N=30,600]	0.03 (0.02) [N=28,800]	-0.01 (0.02) [N=28,800]	-0.00 (0.01) [N=15,100]
Married, self-employed	0.06 (0.10) [N=25,000]	0.53 (0.03) [N=33,700]	0.22 (0.02) [N=34,100]	0.21 (0.02) [N=18,400]

	Second EITC kink			
	0 children	1 child	2 children	3+ children
Single, wage earners	—	0.00 (0.01) [N=325,400]	—	0.05 (0.01) [N=72,700]
Single, self-employed	—	0.02 (0.02) [N=36,100]	—	0.19 (0.05) [N=36,200]
Married, wage earners	-0.04 (0.05) [N=49,200]	-0.01 (0.01) [N=54,000]	-0.00 (0.01) [N=47,100]	0.01 (0.01) [N=31,600]
Married, self-employed	-0.05 (0.06) [N=29,200]	0.03 (0.01) [N=18,900]	0.03 (0.01) [N=21,200]	0.05 (0.01) [N=17,800]

Observed ETIs are reported for various household types, with standard errors in parentheses. The number of taxpayers in the bunching region (rounded to the nearest hundred) is presented in brackets. Estimates are omitted when the kink is between $1,000 and $2,000 away from another kink where taxpayers bunch, as discussed in Section 3.1. Married taxpayers are excluded if filing separately. Wage earners are those with positive wage income and zero self-employment income. The self-employed are those with positive self-employment income. Single status includes "head of household" filers. The income definition is earned income for the EITC and CTC kinks, and taxable income for the statutory kinks. Marginal tax rates are estimated with labor income as the operative channel of response and are inclusive of payroll taxes.

largest elasticities consistently come from single, self-employed taxpayers with one child bunching around the first EITC kink. This comes as no surprise, as this group has the largest bunching coefficients and faces a smaller kink than their counterparts with more children. Recall that kink size is controlled for in equation (5), thus observed elasticities are increasing in bunching coefficients but decreasing in kink size.

Table 5: Observed elasticities at two kinks in 2014

	Second statutory kink	CTC refundability plateau			
		1 child	2 children	3 children	4 children
Single, wage earners	0.01 (0.00) [N=481,400]	0.11 (0.01) [N=252,200]	—	0.01 (0.01) [N=72,200]	0.03 (0.01) [N=12,700]
Single, self-employed	0.02 (0.01) [N=23,200]	1.13 (0.04) [N=175,800]	—	0.12 (0.03) [N=6,300]	0.06 (0.05) [N=700]
Married, wage earners	0.00 (0.00) [N=218,800]	0.04 (0.03) [N=38,200]	0.01 (0.03) [N=45,300]	0.01 (0.00) [N=32,600]	0.01 (0.01) [N=12,800]
Married, self-employed	0.01 (0.00) [N=30,900]	0.55 (0.03) [N=39,900]	-0.04 (0.09) [N=48,300]	0.03 (0.00) [N=18,500]	-0.00 (0.01) [N=4,600]

Observed ETIs are reported for various household types, with standard errors in parentheses. The number of taxpayers in the bunching region (rounded to the nearest hundred) is presented in brackets. Estimates are omitted when the kink is between $1,000 and $2,000 away from another kink where taxpayers bunch, as discussed in Section 3.1. Married taxpayers are excluded if filing separately. Wage earners are those with positive wage income and zero self-employment income. The self-employed are those with positive self-employment income. Single status includes "head of household" filers when analyzing the CTC but excludes them from the second statutory kink analysis. The income definition is earned income for the EITC and CTC kinks, and taxable income for the statutory kinks. Marginal tax rates are determined with labor income as the operative channel of response and are inclusive of payroll taxes.

Outside of self-employed taxpayers with one child at the first EITC kink, observed elasticities are bounded above by 1.13 and typically fall between zero and 0.5. Observed elasticities at the CTC refundability plateau are also largest for the self-employed with one child. Around the second EITC kink, we see the largest observed elasticities among the self-employed with three or more children. At the second statutory kink, observed elasticities are precisely estimated and are very close to zero.[29]

5.3 Elasticities Over Time

One of the most striking results of our research is the changing nature of responsiveness across time. Bunching at the second EITC kink and the CTC refundability plateau kink is nonexistent for the first half of our sample, yet both kinks see bunching patterns emerge in the late 2000s. The first EITC kink sees bunching from the beginning of our sample, but for some groups this response grows stronger over time.

Table 6 presents average elasticities (weighted by sample size) for various groups at four

[29]Because observed elasticities act as lower bounds on structural elasticities, the variability we see in observed elasticities need not imply structural elasticities are changing. Given a fixed structural elasticity, observed elasticities will vary if optimization frictions or kink salience change over time. Further research is needed to tease apart these effects.

Table 6: Average observed elasticities at various kinks

		Single		Married filing jointly	
		Wage earners	Self-employed	Wage earners	Self-employed
First EITC kink	1996-2000:	-0.05	0.19	0.01	0.19
	2001-2007:	-0.01	0.43	0.02	0.28
	2008-2014:	0.06	0.54	0.02	0.28
Second EITC kink	1996-2000:	-0.04	-0.12	0.02	-0.04
	2001-2007:	-0.01	0.03	0.01	0.06
	2008-2014:	0.01	-0.00	0.00	0.07
CTC refundability plateau	1996-2000:	—	—	—	—
	2001-2007:	0.00	0.01	-0.00	0.01
	2008-2014:	0.03	0.22	0.06	0.57
Second statutory kink	1996-2000:	0.01	0.01	0.00	0.00
	2001-2007:	0.01	0.01	0.00	0.01
	2008-2014:	0.00	0.01	0.00	0.01

Weighted-average observed ETIs are reported for various household types, with weights equal to sample size. Married taxpayers are excluded if filing separately. Wage earners are those with positive wage income and zero self-employment income. The self-employed are those with positive self-employment income. Single status includes "head of household" filers. The income definition is earned income for the EITC and CTC kinks, and taxable income for the statutory kinks. Marginal tax rates are estimated with labor income as the operative channel of response and are inclusive of payroll taxes. Note that the CTC refundability plateau kink did not exist until 2001. Elasticity estimates at the zeroth kink are all indistinguishable from zero and are not reported here.

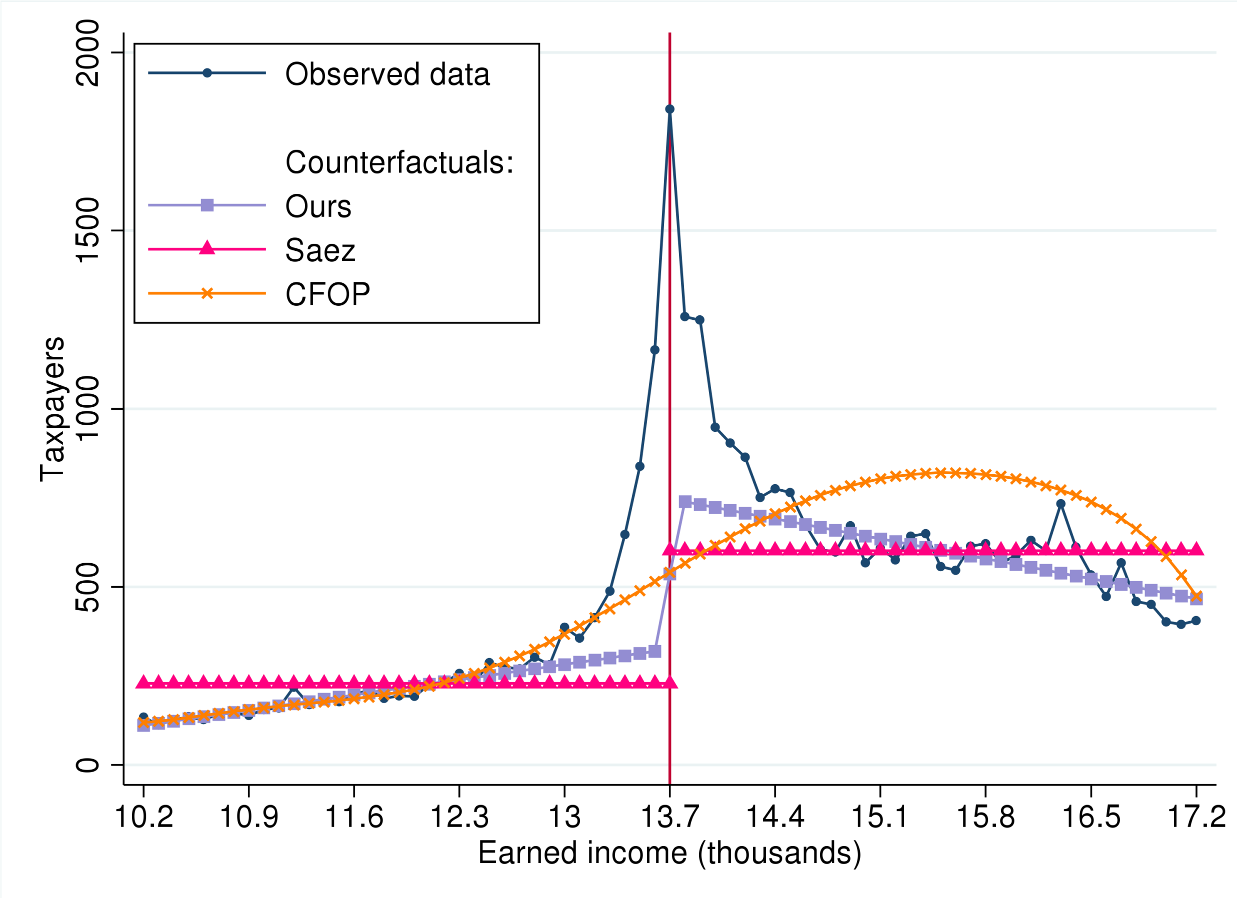
kinks where we observe a bunching response. Observed elasticities remain confined to the range $[0.00, 0.01]$ at the second statutory kink, and are relatively stable at the second EITC kink as well. Wage earners at the other two kinks also see little variation in ETIs. In contrast, the response of the self-employed changes dramatically at the first EITC kink and the CTC refundability plateau.

5.4 Comparing techniques

Here we compare our technique with those of Saez (2010) and Chetty, Friedman, Olsen, and Pistaferri (2011, hereafter CFOP). The two steps necessary to derive elasticity estimates are (i) estimating the number of bunchers, and (ii) translating bunching coefficients to observed elasticities. We borrow from both Saez and Chetty for the first step, and closely follow Saez for the second step. Because we use aspects of both methods, it is unsurprising that our results generally fall in between those of Saez and CFOP.

Figure 10 illustrates the different methods for estimating the number of bunchers, using data on married, self-employed taxpayers with two children around the first EITC kink in 2014. For this group no other kinks were nearby the first EITC kink, allowing us to isolate a pure response to one kink. Saez assumes uniformly distributed counterfactuals, simply taking the average values of the densities outside the bunching window and projecting these averages

Figure 10: Different techniques for estimating counterfactuals



The distribution of income is displayed for married couples filing jointly with two children in 2014. Estimates for the counterfactual distributions of income if there were no kink are also displayed.

inward. This is done for both sides of the kink separately, informing two counterfactual densities. Our technique is similar, except we take into account the slope of the observed densities, forming linear projections inward. The CFOP technique, in contrast, uses the observed densities on both sides of the kink to inform *one* counterfactual density. Because this counterfactual assumes the lower tax rate holds everywhere, the implication is that the mass of bunchers will relocate above the kink. Thus, CFOP adjust their counterfactual density upward to the right of the kink, enforcing that the total mass of taxpayers in the bunching region (i.e. the graph area) remains fixed. This "integration constraint" pushes up the counterfactual density, causing the measure of bunching to fall.[30]

In translating bunching patterns to elasticity estimates, Saez uses a bunching equation

[30]The integration constraint is not consistent, however, as some taxpayers near the top of the region will be pushed out by the new, lower tax rate. Thus their counterfactual density should have strictly less mass than the actual density. The implication is that CFOP have overinflated the counterfactual density, and therefore underestimate bunching.

that is mathematically equivalent to our Equation (5). This equation was derived for kinks with large discrete changes in marginal tax rates. In contrast, CFOP translate their bunching estimates to elasticities using a formula appropriate for small kinks. Specifically, for infinitesimal kinks they establish the following:

$$e \approx \frac{B}{z^* \cdot h_0(z^*) \cdot \log\left(\frac{1-t_0}{1-t_1}\right)}. \tag{6}$$

Table 7 compares the elasticities produced by the three techniques. In comparing our technique to Saez, the difference is straightforward. We use equivalent bunching formulae, and we both estimate two distinct counterfactuals. The only difference is that we relax Saez's assumption of uniformity for the counterfactual densities. The result is that most of our elasticities are smaller than Saez's, as can be seen by comparing columns four and five of the table. This is consistent with Figure 10, which shows the difference in bunching estimates with and without uniformity. Because densities tend to slope upwards toward the kink, Saez's counterfactuals are generally below ours in the bunching window around the kink. His technique therefore counts more bunchers than our method, translating into greater elasticities.[31] Whether these upward slopes would remain absent a kink is ultimately unknowable. We take the conservative stance (in terms of lower elasticities) and assume that they would.

In comparing our technique to CFOP, the situation is more complicated. We use a different technique for both estimating the number of bunchers and for translating bunching coefficients to elasticities. In addition, CFOP impose an integration constraint that is unnecessary for our method. To help isolate these effects, we calculate the CFOP observed elasticities both with and without the integration constraint in columns one and two. We also calculate our observed elasticities using the infinitesimal formula, given by Equation (6), in column three. We move from the CFOP technique in column one to our technique in column four by changing one aspect with each column in between. Column two relaxes the integration constraint. This generally leads to large increases among the statistically significant elasticity estimates. In one case the observed elasticity more than doubles. Next, column three changes the number of counterfactuals from one to two. This sometimes increases and sometimes decreases estimates, but the effects are generally small. Finally, column four replaces the infinitesimal bunching formula with our discrete formula. The effects here are also small, though they tend to decrease the observed elasticities. Thus, the most important difference between our technique and CFOP's is the integration constraint.

[31] Saez's estimates are not universally greater, however, because for some groups the distribution of income slopes downwards towards the kink.

Table 7: Observed elasticities calculated at the first EITC kink in 2014

	(1) CFOP	(2) CFOP	(3) Ours	(4) Ours	(5) Saez
			One child		
Single, wage earners	0.10 (0.01) [N=308,400]	0.10 (0.01) [N=308,400]	0.10 (0.01) [N=308,400]	0.10 (0.01) [N=308,400]	0.19 — [N=308,400]
Single, self-employed	0.40 (0.03) [N=221,800]	0.94 (0.04) [N=221,800]	1.03 (0.03) [N=221,800]	0.95 (0.03) [N=221,800]	2.12 — [N=221,800]
Married, wage earners	0.06 (0.03) [N=28,800]	0.06 (0.04) [N=28,800]	0.03 (0.02) [N=28,800]	0.03 (0.02) [N=28,800]	0.07 — [N=28,800]
Married, self-employed	0.46 (0.04) [N=33,700]	0.56 (0.05) [N=33,700]	0.55 (0.03) [N=33,700]	0.53 (0.03) [N=33,700]	1.05 — [N=33,700]
			Two children		
Single, wage earners	-0.05 (0.02) [N=177,900]	-0.05 (0.02) [N=177,900]	0.07 (0.02) [N=177,900]	0.08 (0.03) [N=177,900]	0.03 — [N=177,900]
Single, self-employed	0.22 (0.03) [N=188,100]	0.41 (0.03) [N=188,100]	0.39 (0.02) [N=188,100]	0.35 (0.02) [N=188,100]	0.67 — [N=188,100]
Married, wage earners	0.01 (0.04) [N=28,000]	0.01 (0.04) [N=28,000]	-0.01 (0.02) [N=28,000]	-0.01 (0.02) [N=28,000]	-0.01 — [N=28,000]
Married, self-employed	0.28 (0.04) [N=34,100]	0.32 (0.04) [N=34,100]	0.23 (0.02) [N=34,100]	0.22 (0.02) [N=34,100]	0.37 — [N=34,100]
Uniform densities?	No	No	No	No	Yes
Integration constraint?	Yes	No	No	No	No
Bunching equation	Infinitesimal	Infinitesimal	Infinitesimal	Discrete	Discrete

Elasticity estimates are reported, comparing our method with those of Saez and CFOP. For all columns, standard errors are in parentheses and sample sizes (rounded to the nearest hundred) are in brackets. Married taxpayers are excluded if filing separately. Wage earners are those with positive wage income and zero self-employment income. The self-employed are those with positive self-employment income. Single status includes "head of household" filers. The discrete and infinitesimal bunching formulae are given by Equations (5) and (6), respectively.

6 Discussion

We have estimated the responsiveness of income to changes in marginal tax rates by measuring the degree to which American taxpayers bunch at kink points in income tax schedules. Our exploration of federal and state statutory kinks, deduction and exemption phase-outs, EITC kinks, and CTC kinks indicates most kinks do not generate observable behavioral responses. In particular, all non-convex kinks fail to induce an observable response, and we do not see any responsiveness among high-income taxpayers at the federal at state income tax kinks we study.

However, we discover economically meaningful bunching at several large, convex kinks in the tax schedule, and statistically significant bunching at a total of seven kinks. The strongest response occurs at the first EITC kink. These patterns are consistent with the findings of Saez (2010) and Chetty et al. (2013). The response at the first kink exists during all years of our sample, 1996 to 2014, and increases in intensity over the course of the sample.

We also find new evidence of bunching responses, most notably at the CTC refundability plateau and the second EITC kink. These responses do not exist at the beginning of our sample, but emerge during the mid 2000s. In addition, we find small but visually compelling, statistically significant responses at the largest statutory kink, the beginning of the statutory schedule, the third statutory kink, and the MWPTC kink.

Contrary to prior research, we document bunching by wage earners at several kinks. However, we also demonstrate that wage earner bunching is almost certainly the product of tax evasion. Wage earner bunching in income reported on Form 1040 vanishes when examining Form W-2, which is reported by employers. Further, it appears most bunchers in wage earnings are systematically over-reporting their wage income, as opposed to hiding wage income.

Finally, we develop a new explanation of bunching by self-employed individuals: targeting the refund-maximizing point in the tax schedule. Large shifts in mass of bunchers over time across kinks are consistent with this behavior. This is strongly suggestive that self-employed individuals are fraudulently reporting income in order to maximize their tax refunds.

References

Bastani, Spencer and Håkan Selin (2014), "Bunching and non-bunching at kink points of the Swedish tax schedule." *Journal of Public Economics*, 109, 36–49.

Burtless, Gary and Robert Moffitt (1984), "The effect of social security benefits on the labor supply of the aged." *Retirement and economic behavior*, 135–171.

Chetty, Raj (2009), "Is the taxable income elasticity sufficient to calculate deadweight loss? The implications of evasion and avoidance." *American Economic Journal: Economic Policy*, 1, 31–52.

Chetty, Raj (2012), "Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply." *Econometrica*, 80, 969–1018.

Chetty, Raj, John N. Friedman, Tore Olsen, and Luigi Pistaferri (2011), "Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from Danish tax records." *The quarterly journal of economics*, 126, 749–804.

Chetty, Raj, John N. Friedman, and Emmanuel Saez (2013), "Using differences in knowledge across neighborhoods to uncover the impacts of the EITC on earnings." *American Economic Review*, 103, 2683–2721.

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

Crandall-Hollick, Margot (2013), "The child tax credit: Current law and legislative history." *Congressional Research Service Report*, 41873.

Devereux, Michael P., Li Liu, and Simon Loretz (2014), "The elasticity of corporate taxable income: New evidence from UK tax records." *American Economic Journal: Economic Policy*, 6, 19–53.

Eissa, Nada and Hilary Hoynes (2011), "Redistribution and tax expenditures: The Earned Income Tax Credit." *National Tax Journal*, 64, 689–730.

Foundation, Kaiser Family (2015), "Annual updates on eligibility rules, enrollment and renewal procedures, and cost-sharing practices in medicaid and chip." URL `kff.org/medicaid/report/annual-updates-on-eligibility-rules-enrollment-and/`. Accessed: 2015-09-04.

Friedberg, Leora (2000), "The labor supply effects of the social security earnings test." *Review of Economics and Statistics*, 82, 48–63.

Gelber, Alexander M., Damon Jones, and Daniel W. Sacks (2013), "Earnings adjustment frictions: Evidence from the social security earnings test." Technical report, National Bureau of Economic Research.

Gruber, Jon and Emmanuel Saez (2002), "The elasticity of taxable income: evidence and implications." *Journal of Public Economics*, 84, 1–32.

Hargarden, Enda Patrick (2015), "Taxpayer responses over the cycle: Evidence from irish notches."

Ito, Koichiro (2014), "Do consumers respond to marginal or average price? evidence from nonlinear electricity pricing." *American Economic Review*, 104, 537–563.

Kleven, Henrik J (2015), "Bunching." *Annual Review of Economics*, 8.

Kleven, Henrik J. and Mazhar Waseem (2013), "Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from Pakistan." *The Quarterly Journal of Economics*, 128, 669–723.

Kleven, Henrik Jacobsen, Martin B. Knudsen, Claus Thustrup Kreiner, Søren Pedersen, and Emmanuel Saez (2011), "Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark." *Econometrica*, 79, 651–692.

Kopczuk, Wojciech (2005), "Tax bases, tax rates and the elasticity of reported income." *Journal of Public Economics*, 89, 2093–2119.

Le Maire, Daniel and Bertel Schjerning (2013), "Tax bunching, income shifting and self-employment." *Journal of Public Economics*, 107, 1–18.

McClelland, Robert and Shannon Mok (2012), "A review of recent research on labor supply elasticities."

Nichols, Austin and Jesse Rothstein (2015), "The earned income tax credit." Technical report, National Bureau of Economic Research.

Ramnath, Shanthi (2013), "Taxpayers' responses to tax-based incentives for retirement savings: Evidence from the saver's credit notch." *Journal of Public Economics*, 101, 77–93.

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

Saez, Emmanuel, Joel Slemrod, and Seth H. Giertz (2012), "The elasticity of taxable income with respect to marginal tax rates: A critical review." *Journal of Economic Literature*, 50, 3–50.

Slemrod, Joel (2007), "Cheating ourselves: The economics of tax evasion." *The journal of economic perspectives*, 25–48.

Weber, Caroline E. (2014), "Toward obtaining a consistent estimate of the elasticity of taxable income using difference-in-differences." *Journal of Public Economics*, 117, 90–103.

Appendix A: Robustness Checks

Here we test our estimation technique for sensitivity to parameter choice. The three key parameters are binwidth and the sizes of the bunching window and bunching region. In Section 3.1, we label these parameters δ, W, and R, respectively. Binwidth simply measures how finely the data are collapsed when performing the analysis. The bunching window defines the area within which we count the total number of bunchers. We assume bunching does not occur outside the bunching window. Finally, the bunching region defines the area outside the bunching window that we use when constructing the counterfactual distribution of income if there were no kink.

As an example, our default parameter values are $\delta = \$100$, $W = 10$, and $R = 35$. This implies a bunching window of $W \cdot \delta = \$1,000$ and a bunching region of $R \cdot \delta = \$3,500$ around the kink. In other words, we assume the distribution of income within $1,000 of the kink is affected by bunching, but that outside this threshold the distribution is unaffected by bunching. Moreover, we use the observed distribution of income between $1,000 and $3,500 away from the kink to estimate the counterfactual distribution of income if the kink did not exist.

Tables 8 and 9 test how these parameters affect our bunching coefficients and observed elasticities for the four most responsive groups at the first EITC kink, using 2003 data. We choose this year as it is the most recent year in which self-employed, low-income taxpayers bunch *only* at the first EITC kink. Starting in 2004, we are constrained when choosing the size of the bunching region, as self-employed taxpayers bunch at the nearby second EITC kink. We discuss this constraint in further detail in Section 3.1. By presenting 2003 estimates here, we avoid this issue and thus are able to test a wide range of parameter choices.

The results indicate our findings are generally robust to parameter choice. For example, our preferred estimate for the observed elasticity of our most responsive group – single, self-employed individuals with one child – is 1.17. Binwidth choices of $50 or $250 lead to estimates of 1.17 and 1.22, with small standard errors. Changing the bunching region by $500 in either direction also has small effects, with alternative estimates of 1.14 and 1.25. The one parameter that has a large effect on the estimates is the choice of bunching window. In particular, choosing a bunching window of just $500 causes the observed elasticity to fall to 0.85. However, this is not a sensible choice for the bunching window, as visual inspection of Figure 11 makes clear. The income distribution is clearly affected by bunching behavior beyond $500 on either side of the kink. Cutting the bunching window short thus has two effects. First, it does not count those bunchers who are more than $500 beyond the kink. Second, it includes those bunchers when estimating the counterfactual distributions of income, artificially inflating these distributions. Both effects decrease the estimated number of bunchers and the corresponding observed elasticity.

Other groups show similar patterns. Our preferred elasticity estimate for single, self-

Table 8: Bunching coefficients and observed elasticities calculated at the first EITC kink in 2003

	Bunching coefficient	Observed elasticity	Sample size	Binwidth	Bunching window	Bunching region
Single, self-employed, one child	32.51 (1.15)	1.17 (0.04)	116,000	$100	$1,000	$3,500
	22.68 (0.85)	0.85 (0.03)	116,000	$100	$500	$3,500
	37.51 (1.98)	1.30 (0.08)	116,000	$100	$1,500	$3,500
	31.82 (1.38)	1.14 (0.05)	109,900	$100	$1,000	$3,000
	34.29 (1.08)	1.25 (0.04)	121,500	$100	$1,000	$4,000
	32.51 (1.25)	1.17 (0.05)	115,700	$50	$1,000	$3,500
	33.90 (1.16)	1.22 (0.05)	116,900	$250	$1,000	$3,500
Single, self-employed, two children	23.34 (1.33)	0.56 (0.03)	106,500	$100	$1,000	$3,500
	13.30 (0.93)	0.34 (0.02)	106,500	$100	$500	$3,500
	32.72 (1.72)	0.75 (0.04)	106,500	$100	$1,500	$3,500
	20.58 (1.31)	0.49 (0.03)	100,900	$100	$1,000	$3,000
	25.35 (1.21)	0.61 (0.03)	111,100	$100	$1,000	$4,000
	23.34 (1.17)	0.55 (0.03)	106,200	$50	$1,000	$3,500
	26.22 (1.42)	0.63 (0.04)	107,200	$250	$1,000	$3,500

Bunching coefficients and observed elasticities are reported for married, self-employed filers with one or two children. Standard errors are in parentheses. The table shows the sensitivity of our estimates to variation in the estimation parameters, reported in the final three columns. Sample size reports the number of taxpayers within the bunching region and is rounded to the nearest hundred. The self-employed are those with positive self-employment income. Single status includes "head of household" filers. Taxpayers under 25 or over 65 years of age who do not have children are ineligible for the EITC and are therefore excluded.

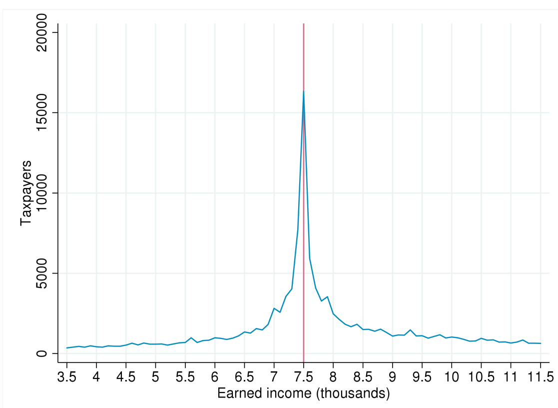
Table 9: Bunching coefficients and observed elasticities calculated at the first EITC kink in 2003

	Bunching coefficient	Observed elasticity	Sample size	Binwidth	Bunching window	Bunching region
Married filing jointly, self-employed, one child	15.84 (0.73)	0.60 (0.03)	34,700	$100	$1,000	$3,500
	12.75 (0.46)	0.49 (0.02)	34,700	$100	$500	$3,500
	14.72 (1.12)	0.53 (0.04)	34,700	$100	$1,500	$3,500
	16.13 (0.88)	0.61 (0.04)	31,800	$100	$1,000	$3,000
	17.02 (0.85)	0.65 (0.03)	37,600	$100	$1,000	$4,000
	15.84 (0.73)	0.59 (0.03)	34,500	$50	$1,000	$3,500
	16.10 (0.69)	0.60 (0.03)	35,100	$250	$1,000	$3,500
Married filing jointly, self-employed, two children	13.98 (0.87)	0.36 (0.02)	44,100	$100	$1,000	$3,500
	8.76 (0.62)	0.23 (0.02)	44,100	$100	$500	$3,500
	19.42 (1.33)	0.49 (0.04)	44,100	$100	$1,500	$3,500
	11.93 (0.89)	0.30 (0.02)	40,000	$100	$1,000	$3,000
	14.82 (0.89)	0.38 (0.02)	47,800	$100	$1,000	$4,000
	13.98 (0.87)	0.35 (0.02)	43,900	$50	$1,000	$3,500
	15.51 (0.99)	0.40 (0.03)	44,600	$250	$1,000	$3,500

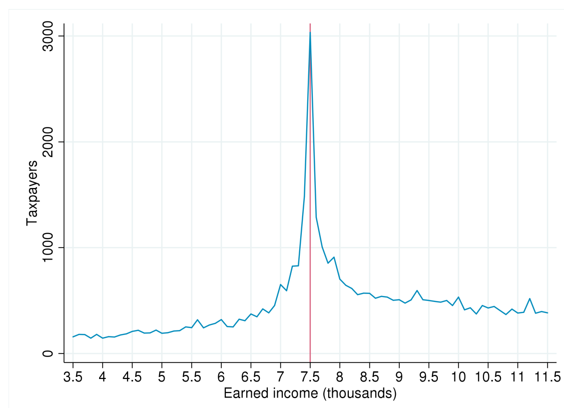
Bunching coefficients and observed elasticities are reported for married, self-employed filers with one or two children. Standard errors are in parentheses. The table shows the sensitivity of our estimates to variation in the estimation parameters, reported in the final three columns. Sample size reports the number of taxpayers within the bunching region and is rounded to the nearest hundred. The self-employed are those with positive self-employment income. Taxpayers under 25 or over 65 years of age who do not have children are ineligible for the EITC and are therefore excluded.

employed taxpayers with two children is 0.56. Except for the choice of bunching window, all other permutations leave the observed elasticity in the range $[0.49, 0.63]$. Unsurprisingly, a smaller bunching window of \$500 reduces the estimate to 0.34. Again, visual inspection of the distribution makes clear that it is still affected by bunching between \$500 and \$1,000 of the kink. For married, self-employed taxpayers, our preferred elasticity estimates are 0.60 and 0.36, respectively, for those with one or two children. Except for the choice of bunching window, alternative estimates for these parameters lie in the ranges $[0.59, 0.65]$ and $[0.30, 0.40]$, respectively. For all cases, the bunching window could arguably be expanded from \$1,000 to \$1,500. This would generally increase our elasticity estimates. We take the conservative approach in choosing \$1,000.

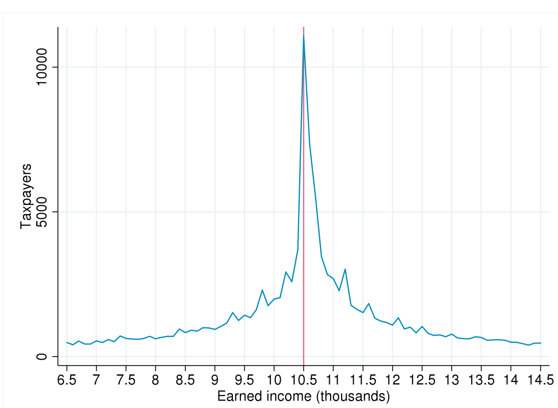
Figure 11: Income distribution of self-employed taxpayers near the first EITC kink in 2003



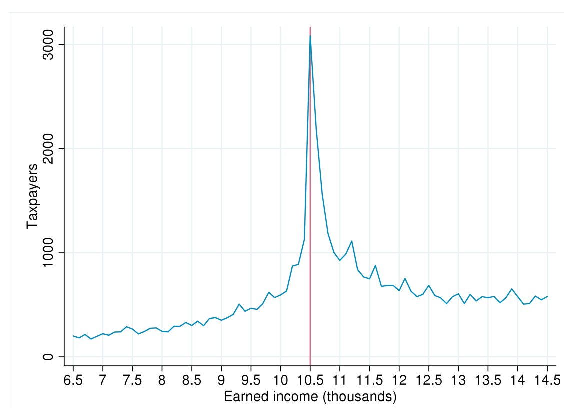
(a) Single, one child



(b) Married filing jointly, one child



(c) Single, two children



(d) Married filing jointly, two children

The distribution of income is displayed for various household types in 2003. Single status includes "head of household" filers. The self-employed are those with positive self-employment income.

Appendix B: Personal Exemption and Itemized Deduction Phase-outs

When determining taxable income, both personal exemptions and itemized deductions phase-out at high incomes, creating discontinuities in the budget constraints of high-income taxpayers. Our evidence suggests taxpayers do not respond to these incentives, but we describe them here for completeness. The phase-outs discussed in this section were in effect during our sample from 1996 to 2005, but were gradually removed beginning in 2006, with full removal from 2010 to 2012. They have since been reinstated.

The personal exemption phase-out (PEP) is a step function of AGI, generating notches in the budget constraint. Personal exemptions are reduced by 2% for each $2,500 of income exceeding the phase-out threshold until exemptions are exhausted. The beginning (and end) of the phase-out varies by filing status: $145,950 for singles, $182,450 for head of household, and $218,950 for married couples filing jointly in 2005.[32]

The itemized deduction phase-out (often referred to as "Pease" after former Ohio Congressman Donald Pease) reduces certain itemized deductions at a rate of 3 cents per dollar of AGI exceeding the threshold. However, Pease does not apply against itemized deductions generated from casualty and theft losses, investment interest, gambling losses, or medical expenses. The total percentage of itemized deductions eliminated by Pease is capped at 80% per taxpayer. Throughout the time period we study (1996-2005) this threshold is the same for all filing statuses except married couples filing separately, for whom the threshold is halved. In 2005 the threshold was $145,950 ($72,975), identical to the PEP threshold for singles.

Pease creates relatively small changes in marginal tax rates at its introduction and conclusion. For example, suppose a head of household with three children claims $20,000 of itemized deductions and earns exactly the Pease threshold of $145,950 in 2005. For a marginal increase of $1,000 above the Pease threshold, qualified itemized deductions are reduced by 3%, meaning the individual has 30 additional dollars of taxable income. If the taxpayer faces an initial marginal tax rate of 31%, she would see her marginal rate increase by around 1 percentage point ($\$30 \times 31\%/\$1000$) as a result of Pease, creating a small convex kink. Similarly, once Pease is phased out the change is also around 1 percentage point, which creates a small non-convex kink. In the presence of moderate optimization frictions, these kinks are unlikely to induce a behavioral response.

PEP generates larger marginal tax rate increases than Pease. However, because the discontinuities PEP generates are notches, not kinks, assumptions are needed to calculate the magnitude of the discontinuity relative to a kink. For example, the "size" of a kink is calculated by dividing the difference between the net-of-tax rates (one minus the marginal

[32]For all filers, the end of the phase-out region is $122,500 above the beginning.

tax rate) on either side of a kink, and dividing by the net-of-tax rate to the right of the kink. Calculating the size of a notch requires an assumption about the size of a marginal response by the taxpayer: do taxpayers adjust their income in \$1, \$50, or \$1,000 increments?

Suppose a taxpayer earns income at the PEP threshold. She has four personal exemptions, which reduce taxable income by $4 \times \$3,200 = \$12,800$. If this taxpayer earns at least 1 additional dollar but less than 2,500 additional dollars, her personal exemptions will be reduced by \$256 ($2\% \times \$12,800$). Assuming her marginal tax rate is 31% initially, this increases her tax liability by $31\% \times \$256 = \$79.36$. If we assume a marginal response constitutes a \$1 change in income, the implicit change in marginal tax rates is 7,936 percentage points. If instead we assume the marginal response is \$1,000, the implicit change in marginal tax rates is 7.936 percentage points. We take the most conservative measure, assuming the income increment is the full \$2,500. Thus we take this taxpayer's kink size to be 3.17 percentage points.[33]

[33] The formula for kink size, given an income increment of $X \leq 2500$, is $((79.36/X) \times 100\%)$.

Appendix C: Bunching Analysis of Medicaid, SNAP, and Federal Disability Insurance

In this appendix we analyze whether federal income tax data provide evidence that individuals adjust their incomes to remain eligible for Medicaid, Supplemental Nutrition Assistance Program (SNAP), or federal disability benefits. These are large, economically important programs that serve millions of people. In 2011, there were approximately 57 million Medicaid participants, 45 million SNAP beneficiaries, and 8.5 million workers receiving federal disability benefits. We analyze these programs as a robustness check to ensure perceived bunching at kink points in the tax schedule was not caused by other incentives. We describe each program in broad strokes, with an emphasis on program-specific income definitions and eligibility thresholds. We then discuss our findings and conclude with a short discussion of the strengths and weaknesses of our approach. In short, we see no evidence of bunching associated with any of these programs, but this may be due to the limitations of the tax data in the context of these programs.

C.1 A Brief Description of Each Program

Medicaid provides health insurance at subsidized rates to low-income individuals, primarily parents, pregnant mothers, and children. For many, the program is free. Income eligibility criteria are a function of the federal poverty line, which is an increasing function of the number of adults and children in the household. Medicaid is administered at the state level, and the definitions of qualifying income and eligibility thresholds (as a percentage of the federal poverty line) vary by state. Importantly, during the period we analyze (2002 to 2011) Medicaid introduces a "notch" in the budget set, as earning income above the threshold results in a complete loss of benefits. This is in contrast to having benefits phase-out, which would produce a kink. Given that the threshold is a notch and that Medicaid is a large benefit, we expect substantial responsiveness to this threshold.

SNAP provides funds to low-income individuals that can only be spent on certain types of food at participating retailers. Similar to Medicaid, it is a function of the federal poverty line, with larger households receiving greater benefits. Unlike Medicaid, the thresholds and income definitions do not vary across states, with the exceptions of Alaska and Hawai'i. Eligibility is limited to those with less than 130% of the federal poverty line in gross monthly income and less than 100% of the poverty line in net monthly income. Benefits begin phasing out with the first dollar of earned income. Individuals receive the difference between the maximum allotment per household (around $650 for a family of 4) and 30% of their monthly income. This produces a relatively modest non-convex kink at the end of the phase-out region.[34]

[34] See www.fns.usda.gov/snap/eligbility for a thorough description of SNAP, including eligibility criteria and definitions of income.

Federal disability insurance (DI) is administered by the Social Security Administration, providing monthly payments to individuals with a disability. Individuals must apply for DI and once approved are unable to earn income above a certain threshold without triggering a review of their claim or outright termination of benefits. Thus, the threshold (called the Substantial Gainful Activity threshold) is a notch. The Substantial Gainful Activity threshold ranged from $500 each month in 1996 to $1,000 each month in 2011. The size of the monthly benefit for DI recipients is a function of prior earnings and can range from a few hundred dollars to a few thousand dollars per month. As a result, we believe the potential for bunching to the left of this notch is substantial.[35]

C.2 Qualifying Income and Kink Construction

SNAP and DI income eligibility thresholds are measured monthly, and in some states Medicaid income is measured monthly as well. Given that tax data record income on an annual basis, we construct the applicable kinks or notches by simply multiplying the monthly thresholds by twelve.[36]

All three programs define income in an analogous manner to "earned income" qualifying for the EITC, but can include income items that are not recorded by the tax system, such as child support payments, housing subsidies, or Supplemental Security Income. In addition, various deductions are allowed and SNAP beneficiaries must satisfy the gross and net income tests described above. Medicaid income varies by state, but we use the same income definition for all states and all years.

We use the EITC sample as the basis for studying all three programs. These data are drawn with the following restrictions: all observations are from the seven states with no income taxes, all filed federal income tax returns, and all were between 25 and 65 years of age. We analyze Medicaid bunching from 2002 to 2011, SNAP bunching from 1996 to 2011, and DI bunching from 1999 to 2011.

C.3 Bunching Results for Medicaid, SNAP, and Federal Disability Insurance

We find an absence of bunching at each eligibility threshold in every state and every year. This is surprising given the value of these programs to participants – especially in the case of Medicaid and DI. The lack of bunching suggests either individuals are not adjusting their income in response to these programs, or our estimates are biased by substantial measurement error. Individuals might not respond if they have imperfect knowledge of income eligibility criteria, or because adjusting income is costly.

[35]See www.ssa.gov/disability for a thorough description of federal disability insurance.

[36]We retrieved annual state specific income eligibility thresholds for Medicaid from Foundation (2015).

On the other hand, individuals may be responding to these programs in ways that are undetectable with tax data. Federal income tax data do not contain all income or deduction items that comprise qualifying income for these programs. This is particularly relevant for the analysis of Medicaid, where the income definition varies across states and potentially over time. A second problem is that tax data are recorded annually, while eligibility for two (and in some states all three) programs are evaluated on a monthly basis. Thus, the only bunchers we can identify in our data for SNAP and DI are those that bunch in *every* month.