# Evaluating Tax Reforms without Elasticities: What Bunching *Can* Identify

Dylan T. Moore*

[**Job Market Paper**: Click here for latest version]

October 19, 2021

**Abstract**

Progressive, piecewise linear tax schedules are ubiquitous in modern economies. Evaluating proposed reforms to such schedules requires predicting how taxpayers will respond to changes in tax rates and brackets. This is typically done using estimates of the elasticity of taxable income (ETI), but challenges to both the internal and external validity of ETI estimates raises concerns about this approach. I present a new method for forecasting the revenue impact of proposed tax bracket changes using the pre-reform distribution of taxable income alone. The method relies on a novel nonparametric identification result: the "bunching mass" at a given bracket threshold (the share of taxpayers locating there) is a sufficient statistic for the revenue effect of taxpayer behavioral responses to small movements of the threshold. I explore implications for ex ante tax policy analysis, documenting what this method can tell us about revenue and welfare effects of proposed reforms, and how it compares to current practices. Extending my results to account for heterogeneous tax schedules and optimization frictions, I present an application to the Earned Income Tax Credit. Connecting my findings to the literature on bunching-based estimation of the ETI, I show that using bunching to analyze small tax bracket changes is far more robust than attempting to use it estimate a policy-relevant ETI. However, this robustness comes at the cost of narrowing the range of policy questions bunching methods can address. To address this concern, I close the paper by discussing how to use bunching to evaluate large bracket changes, and characterizing the conditions under which a bunching-based ETI estimate may be used to assess proposed tax rate changes.

*PhD Candidate. Department of Economics, University of Michigan, Ann Arbor, MI, 48109, United States. Email: dtmoore@umich.edu. I am grateful for helpful comments on earlier drafts received from Ashley C. Craig, James R. Hines Jr., Nirupama Rao, Nathan Seegert, Joel Slemrod, Ellen Stuart, and Tejaswi Velayudhan. I also thank seminar participants at the University of Michigan as well as conference participants at the Michigan Tax Invitational 2021 and the Young Economist Symposium 2021 for their valuable feedback.

Nearly all OECD economies feature progressive, piecewise linear income tax schedules in which the marginal tax rate varies across different brackets, increasing discontinuously at the thresholds separating brackets. Major tax reforms in these economies often feature both rate and bracket changes. Consider, for example, the Tax Cuts and Jobs Act of 2017 (TCJA), which substantially reformed the US income tax schedule. This reform reduced tax rates across five out of seven brackets of the US federal income tax schedule, but simultaneously changed the location of many tax bracket thresholds.[1] For example, single tax filers in 2017 entered the top tax bracket when their taxable income surpassed $418400. The TCJA moved this threshold to $500000 in 2018. The TCJA also featured a substantial increase in the standard deduction, which can be thought of as having changed the location of all bracket thresholds as function of pre-deduction income.[2]

Policymakers and other observers analyzing a proposed tax reform like the TCJA require information about how the prospective tax policy changes will alter taxpayer behavior in order to make accurate forecasts of the revenue impact of the reform. That is, they need to be able to estimate the *behavioral response to taxation*. In the modern tax analysis literature, the elasticity of taxable income (ETI) has become a centrally important measure of behavioral responses.[3] Estimates of the ETI are commonly employed by government agencies tasked with forecasting the fiscal impact of proposed tax reforms.[4] However, credibly estimating the ETI has proven challenging in practice. Estimates vary considerably according to the methodological approach adopted [Saez et al., 2012, Weber, 2014, Aronsson et al., 2018, Kumar and Liang, 2020, Neisser et al., 2021]. Furthermore, there is a strong theoretical and empirical rationale for expecting that the ETI can vary substantially with features of the tax system and other contextual factors, further complicating the question of how to select amongst ETI estimates for a given analysis [Slemrod and Kopczuk, 2002, Kopczuk, 2005, Jacquet and Lehmann, 2021, Neisser et al., 2021]. This uncertainty about the ETI can have important practical consequences, as estimated revenue effects of tax reforms is, at times, quite sensitive to the choice of ETI.[5]

This paper shows that for some important features of tax policy there may be no need to rely on ETI estimates in revenue forecasting, exploring a new approach to predicting the revenue and welfare effects of tax

[1] In many countries such as the US, tax bracket thresholds changed every year in a bid to account for inflation. When I talk about a reform that changes these thresholds, I mean that it substantially altered their location in terms of real income.

[2] While the tax code defines brackets in terms of *taxable income* after subtracting various deductions, from an economic theory perspective we care about what taxpayer budget constraints look like as a function of total income.

[3] Feldstein [1999] shows that the ETI can be viewed as a sufficient statistic for the efficiency costs (deadweight loss) of taxation.

[4] For example, the Canadian Parliamentary Budget Office (PBO) uses estimates of the ETI to assess the likely revenue effects of actual and hypothetical tax reforms at the request of Members of Parliament and political parties during election campaigns. See, for instance, https://www.pbo-dpb.gc.ca/web/default/files/Documents/Reports/2016/PIT/PIT_EN.pdf.

[5] For instance, Saez et al. [2012] note that whereas early estimates of the ETI in the US were so high that they implied the existence of income tax cuts which would actually raise revenue (a Pareto improvement), later estimates were much lower, suggesting that much more revenue could be raised by increasing tax rates.

reforms like the TCJA which include changes to brackets. The proposed method takes advantage of a quirk of taxpayer behavior under a progressive, piecewise linear tax schedule. Such schedules induce convex kinks in taxpayer budget sets at bracket thresholds, causing the marginal benefit of earning an additional dollar to drop discontinuously at the threshold. Consequently, for a large mass of taxpayers it is optimal to "bunch": to earn precisely the amount of income that lands them at the kink point (i.e. at the tax bracket threshold) is their optimal choice. The central finding of this paper is that the share of taxpayers who are bunching at a given kink (the "bunching mass") identifies the revenue impact of taxpayer responses to small changes in the location of the kink point. This implies that the revenue effect of moving a tax bracket threshold can be identified from the observed distribution of taxable income alone.[6] After documenting this result, the paper explores how it might be used to evaluate actual and potential reforms of real-world tax schedules. I also discuss the implications of these findings for the literature on the use of bunching to estimate the ETI.

Section 1 presents the main finding in the case of a simple two-bracket piecewise linear tax schedule. I assume only that taxpayers have continuous, convex preferences and that they make a frictionless choice of income, allowing for arbitrary heterogeneity in taxpayer preferences.[7] In this setting, the bunching mass is a sufficient statistic the first-order revenue effect of the behavioral response to changing the location of the threshold separating the two brackets. That is, bunching identifies the behavioral response effect of moving a convex kink point. The magnitude of the rate change at the kink is irrelevant to this result.

The derivation of this finding is straightforward, but the intuition behind it is somewhat subtle. Consider what happens following a small decrease in the location of the kink point. Such a reform raises the marginal tax rate on income within a window just below the original kink location. Absent income or labor force participation effects, the tax revenue effect of this reform has three parts. First, additional revenue is raised from taxpayers above the kink (the *mechanical effect*). The remaining two parts capture the *behavioral response effect* of the reform. Holding constant the size of the bunching mass, there is a loss in revenue due to the fact the bunching taxpayers have been moved to a lower level income. I label this the *relocation effect*. However, the bunching mass will not in fact remain constant: some individuals who bunched at the old kink will choose not to bunch at the new kink and will instead locate somewhere in the upper tax bracket. On the other hand, some individuals who used to locate in the lower bracket will begin bunching at the new kink point. I label these the *new* and *former buncher effects* respectively.

The central result of this paper follows from the fact that while the relocation effect is first-order, the new and former buncher effects are second-order. The bunching mass is then a sufficient statistic for the behavior

[6]Note, in this paper I say that a parameter is "identified" by the observed distribution of taxable income if—under the conditions specified—it can be obtained as a function of the distribution. This is consistent with the general definition of identification proposed by Lewbel [2019].

[7]That is, I assume that a taxpayer's observed choice of income reflects their optimal choice of income.

response effect of the reform because the relocation effect is proportional to the size of the bunching mass. Importantly, this result obtains in spite of the fact that moving a kink point has a first-order effect on the size of the bunching mass. It is a consequence of the fact that—to a first approximation—the decisions of taxpayers who are at the margin of entering or leaving the bunching mass cause no changes in tax revenue.

Section 2 of the paper explores the implications of this finding for analysis of progressive, piecewise linear tax schedules. Applying an optimal tax theory lens, novel tests for the Pareto efficiency of an observed tax schedule are presented. All prior such tests rely on strong functional form assumptions and account for the behavioral response to taxation using estimates of the ETI. By contrast, testing the efficiency of kink point location only requires knowledge of the observed distribution of taxable income as long as preferences are convex and there are no income or participation effects. Testable conditions for local Rawlsian (revenue-maximizing) optimality of kink point location are also identified. For non-Rawlsian welfare functions, welfare effects of moving kinks are not point identified because the behavioral response to such a reform generates a first-order impact on buncher utility and the magnitude of this effect is unobservable. However, it is possible to identify bounds on welfare effects given known welfare weights.

When tax reforms change both rates and brackets, estimates of the ETI are still needed to predict the part of the behavioral response effect caused by the rate changes. However, the bunching mass can still be used to predict the part of the behavioral response caused by the bracket changes. Indeed, I show that this is the correct way to account for the behavioral response to small bracket changes in tax reform evaluation. By contrast, commonly employed approaches to tax revenue forecasting which rely on using the ETI to estimate these behavioral responses are biased even if the ETI is correctly calibrated.

To illustrate the potential practical value of these findings, section 3 presents an application to analyzing the location of the first kink in the Earned Income Tax Credit (EITC). Application to real-world income data requires extending the model to incorporate two key features of real-world tax policy. First, in practical policy settings, different taxpayers often face different effective tax schedules at the same level of income. In the case of the EITC for example, variation in state-level EITC top-ups and welfare benefit phaseouts generates substantial differences in the fiscal impact of moving the EITC across different groups of bunching taxpayers. Second, in real-world taxable income data, taxpayers do not appear to bunch precisely at kink points. Rather, empirical evidence of bunching generally takes the form a a diffuse lump of taxpayers spread around the kink. Following the prior literature on bunching methods I assume that this departure from the prediction of the frictionless model is caused by optimization error. That is, I assume that taxpayer choices of taxable income are jointly produced by their optimal choice of income and an idiosyncratic random shock.

I show how both of these complexities of a practical tax policy setting can be incorporated without compromising the sufficient statistic interpretation of the bunching mass. Explicitly accounting for the role of frictions in can substantially change the magnitude of the behavioral response effect to moving the kink. I also discuss how income effects and extensive margin responses to the EITC could be incorporated into this analysis.

Following this empirical application, section 4 discusses the findings presented in this paper in the context of the existing literature on bunching methodology. The standard approach to bunching methodology first developed in Saez [2010] uses the bunching mass to generate estimates of the ETI. However, recent work on this approach demonstrates that nonparametric identification of the ETI using the bunching mass and other features of the observed distribution of taxable income is not possible [Blomquist et al., 2019, Bertanha et al., 2021]. While various remedies to this limitation have been proposed all these approaches rely on either strong functional form assumptions or only provide bounds on the ETI. Furthermore, even if the identification were possible, the ETI parameter which the standard bunching method seeks to identify may or may not provide policy-relevant information when agent preferences are heterogeneous [Blomquist et al., 2019].

Such work leaves open the question of whether or not the bunching mass has a generally valid policy-relevant interpretation. This paper shows that it does: the bunching mass is directly informative about the behavioral response effect of local movements of tax bracket thresholds. Nonetheless, as I have noted, estimates of the ETI are still required to evaluate the impact of tax rate changes. Thus, the contribution of this paper is to document a new use case for empirical bunching designs which yields more robust results than ETI-focused approaches at the expense of addressing a much more narrow set of policy questions.

To address this concern, section 4 also includes a discussion of the key factors that determine whether or not the ETI parameter which bunching methods seek to identify has a policy-relevant interpretation. The applicability of this parameter is compromised when the local average ETI is endogenous to the tax rate. However, the bias introduced into policy analysis by a reliance on this parameter may be small if the change in tax rates at the kink is small, the variance of taxpayer elasticities is low, and taxpayer preferences are close to isoelastic.

In addition, section 5 discusses another way bunching can be used to evaluate tax reforms which go beyond small bracket changes. I show that bunching can also provide some information about the impact of “large” (discrete) changes in the location of an existing kink point. Higher-order approximations of the behavioral response effect of such changes are identified by the observed distribution of taxable income. I also discuss a partial identification approach to estimating bounds on these revenue effects.

**Related Literature** As noted above, this paper is closely related to prior work on bunching methodology, and this connection is further discussed in section 4. Two prior papers in the bunching methodology literature that deserve particular mention will not appear in that later discussion. Marx [2018] discusses the welfare effect of moving a regulatory notch which generates bunching behavior in a study of charitable organization behavior. His results are not directly relevant to those discussed here but are of note because of their focus on a reform that changes the location of the threshold where bunching occurs. As well, Goff [2021] presents an identification result which is closely related to that presented in section 1 of this paper in the context of an application of bunching methods to estimating the impact of overtime pay regulations. Goff [2021] shows how bunching of hours at the overtime threshold can be used to identify the impact of changing the overtime hours threshold. The identification results presented here were developed independently.

# 1 The Bunching Mass as a Sufficient Statistic

In this section, I first introduce the model and key assumptions. Next, I present the basic finding and explore the intuition driving the result. This is followed by a discussion of applied welfare analysis using this finding. Finally, I show that that the central result is robust to allowing for income effects and extensive margin responses.

## 1.1 Model

Consider an economy populated by agents of different types $\theta \in \Theta$, where the type space $\Theta$ is convex and may be multidimensional. Suppose that taxpayer type is continuously distributed according to $F$, with a corresponding type density $f(\theta)$.

Each type of agent $\theta$ chooses taxable income $z$ to solve a type-specific utility maximization problem

$$\max_{z} \{u(c, z; \theta) : c = z - T(z)\}, \tag{1}$$

given some income tax schedule $T(\cdot)$. I derive the main result under very weak assumptions on agent preferences.

**Assumption 1** (Convex, Continuous Preferences)**.** *For all types $\theta \in \Theta$, the utility function $u(c, z; \theta)$ is:*

*(i) continuous in $(c, z)$;*

*(ii) strictly increasing and concave in consumption $(c)$, and;*

*(iii) and strictly decreasing and strictly convex in taxable income $(z)$.*

This assumption ensures that whenever the tax schedule $T$ is (weakly) convex, the maximization problem (1) admits a unique solution for each type $\theta$, which I label $z(\theta)$. Assumption 1 also ensures that this solution changes continuously in response to infinitesimal tax reforms (i.e. that small tax reforms don't induce "jumps" in the taxpayer's location on the budget constraint). Finally, I note that because taxpayer preferences are rational, and their choice of taxable income is always unique, their choices will satisfy the strong axiom of revealed preference (SARP).

It is worth highlighting the types of preferences that can be accommodated by assumption 1. For example, it allows for the possibility that agent's choice of taxable income results from adding together multiple income types as well as avoidance or evasion decisions. This assumption also allows for the possibility of a non-differentiable utility function, allowing for the possibility that agents to possess reference-dependent preferences.

### A Two-Bracket Tax Schedule

Let $T(\cdot)$ be a piecewise linear tax schedule with a marginal tax rate of $t_0$ on income below $z^\star$, a marginal tax rate of $t_1 > t_0$ on income above $z^\star$, and a demogrant of $G$ (i.e. $T(0) \equiv 0$). Such a schedule can be simply written as follows

$$T(z) \equiv \begin{cases} T_0(z) & \text{if } z \leq z^\star \\ T_1(z) & \text{if } z > z^\star \end{cases}, \tag{2}$$

where

$$T_0(z) \equiv t_0 z - G, \tag{3}$$

and

$$T_1(z) \equiv t_1(z - z^\star) + t_0 z^\star - G. \tag{4}$$

This schedule consists of two linear tax brackets separated by a bracket threshold at $z^\star$. The upper panel of the figure 1 depicts the budget set induced by this type of tax schedule. I will often refer to $z^\star$ as the "kink point" because, as the figure shows, this budget set features a convex kink at $z^\star$.

**Taxpayer Choices Under Linear Tax Schedules**

To characterize the choices that taxpayers make when facing tax schedule 2, we must first consider the counterfactual choices they would make when facing the linear tax schedules $T_0(z)$ and $T_1(z)$. These two tax schedules differ in both their marginal tax rates ($t_0$ vs. $t_1$) and their associated virtual income ($G$ vs. $(t_1 - t_0) z^\star + G$). Let $z(t, V; \theta)$ be the choice of taxable income a type $\theta$ taxpayer would make under a linear tax schedule with marginal tax rate $t < 1$ and virtual income $V \in \mathbb{R}$ and let $H(z; t, V)$ be the cumulative distribution function of taxable income induced by this linear tax schedule.[8] For simplicity of exposition, I require that this CDF is continuously differentiable in $z$.[9]

**Assumption 2** (Regularity Condition)**.** *For any linear tax schedule with marginal tax rate $t < 1$ and virtual income $V$, the cumulative distribution function of taxable income induced by this schedule ($H(z; t, V)$) is continuously differentiable for all $z$, with a corresponding density function $h(z; t, V)$.*

Additionally, in this section I will assume that differences in the virtual income amount do not affect taxpayer decisions: that is, there are no income effects. Later, I discuss the implications of income effects for my results.

**Assumption 3** (No Income Effects)**.** *Any pair of linear tax schedules with the same marginal tax rate $t$ induce the same distribution of taxable income, irrespective of differences in demogrant amounts. That is, for any $t < 1$ and $V, V' \in \mathbb{R}$,*

$$H(z; t, V) = H(z; t, V')$$

*for all $z$.*

Under this assumption, we can simplify notation somewhat, writing a type $\theta$ taxpayer's choice under a linear tax schedule with marginal rate $t$ as $z(t; \theta)$, the corresponding CDF of taxable income as $H(z; t)$, and the corresponding density function as $h(z; t)$. Importantly, note that the distribution $H(z; t_1)$ is invariant to changes in the parameter $z^\star$ that affect the virtual income associated with $T_1$ and that both $H(z; t_1)$ and $H(z; t_0)$ are invariant to changes in the demogrant amount $G$.

To further simplify notation, let $z_k(\theta) \equiv z(\theta, t_k)$, $H_k(z) \equiv H(z; t_k)$, and $h_k(z) \equiv h(z; t_k)$ for $k \in \{1, 2\}$.

[8] Formally, $H(z; t, V) \equiv \int \mathbf{1}\{z(\theta; t, V) \leq z\} \, \mathrm{d}F(\theta)$.

[9] This assumption could be weakened somewhat. Obtaining my key results requires that the CDF of taxable income under a linear tax schedule is continuously differentiable at a tax bracket threshold. However, these results are robust to the possibility that the CDF of taxable income under a linear tax features discontinuities or non-differentiability at income levels away from the bracket threshold. Thus, the result is robust to the possibility that, for example, some taxpayers have reference dependent preferences that induce bunching at a reference point, as long as this reference point is not located at the tax bracket threshold and is unaffected by changes to the threshold.

**Taxpayer Choices under a Two-Bracket Tax Schedule**

When facing the two-bracket tax schedule from equation (2), the choice of taxable income for any taxpayer type $\theta$ can be written as:[10]

$$z(\theta) = \begin{cases} z_0(\theta) & \text{if } z_0(\theta) < z^\star \\ z_1(\theta) & \text{if } z_1(\theta) > z^\star \\ z^\star & \text{if } z_1(\theta) \leq z^\star \leq z_0(\theta) \end{cases} . \tag{5}$$

The lower panel of figure 1 illustrates this visually, showing how taxpayers make decisions when facing the two bracket tax schedule 2. All taxpayers will choose to earn income up to the point where the marginal cost of earning exceeds the marginal benefit. For some, this means locating in the interior of the lower bracket and choosing to earn the same amount of income they would under $T_0$. For others, this means locating in the interior of the upper bracket and earning the same amount they would under $T_1$.

However, not all taxpayers choices can be characterized this way. The kink in the tax schedule induces a discontinuous drop in the marginal benefit of income at $z^\star$. Consequently, some taxpayer marginal cost curves pass through the discontinuity, never intersecting the marginal benefit. For such taxpayers, the marginal benefit of earning each dollar of income below $z^\star$ exceeds their marginal cost but for each dollar beyond $z^\star$ the opposite is true.[11] Thus, they will locate at exactly $z^\star$. These taxpayers are *bunchers.*

Figure 2 shows how the choices of a population of different types of taxpayers facing the tax schedule (2) generates a distribution of taxable income which features "bunching" at the kink point: that is, a mass point at $z^\star$. As in figure 1, the upper panel of figure 2 depicts how taxpayers make choices by plotting the marginal benefit of income against the marginal cost curves of different types of taxpayers. The lower panel of the figure shows the density of taxable income that is induced by the type distribution $F$ and the type-to-choice mapping $z(\cdot)$ (equation 5). Below the kink point, the observed distribution of taxable income coincides with the distribution under the linear tax schedule $T_0$ ($H_0$), and above this point it coincides with the distribution under the linear tax schedule $T_1$ ($H_1$):

$$H(z) = \begin{cases} H_0(z) & \text{if } z < z^\star \\ H_1(z) & \text{if } z \geq z^\star \end{cases} \tag{6}$$

Importantly, this distribution features a mass point at $z^\star$, *the bunching mass,* which is simply the fraction of

[10] All taxpayers must satisfy one, and only one of the three conditions in equation (5). The alternative possibility ($z(t_0;\theta) < z^\star < z(t_1;\theta)$) would violate SARP, since the $z(t_0;\theta)$ would be affordable under $T_1$ and $z(t_1;\theta)$ would be affordable under $T_0$.

[11] Put another way, if such taxpayers faced the linear tax schedule $T_0$ they would want to locate above the kink, but if they faced a linear tax schedule $T_1$ they would want to locate below the kink.

Figure 1: Taxpayer Choices

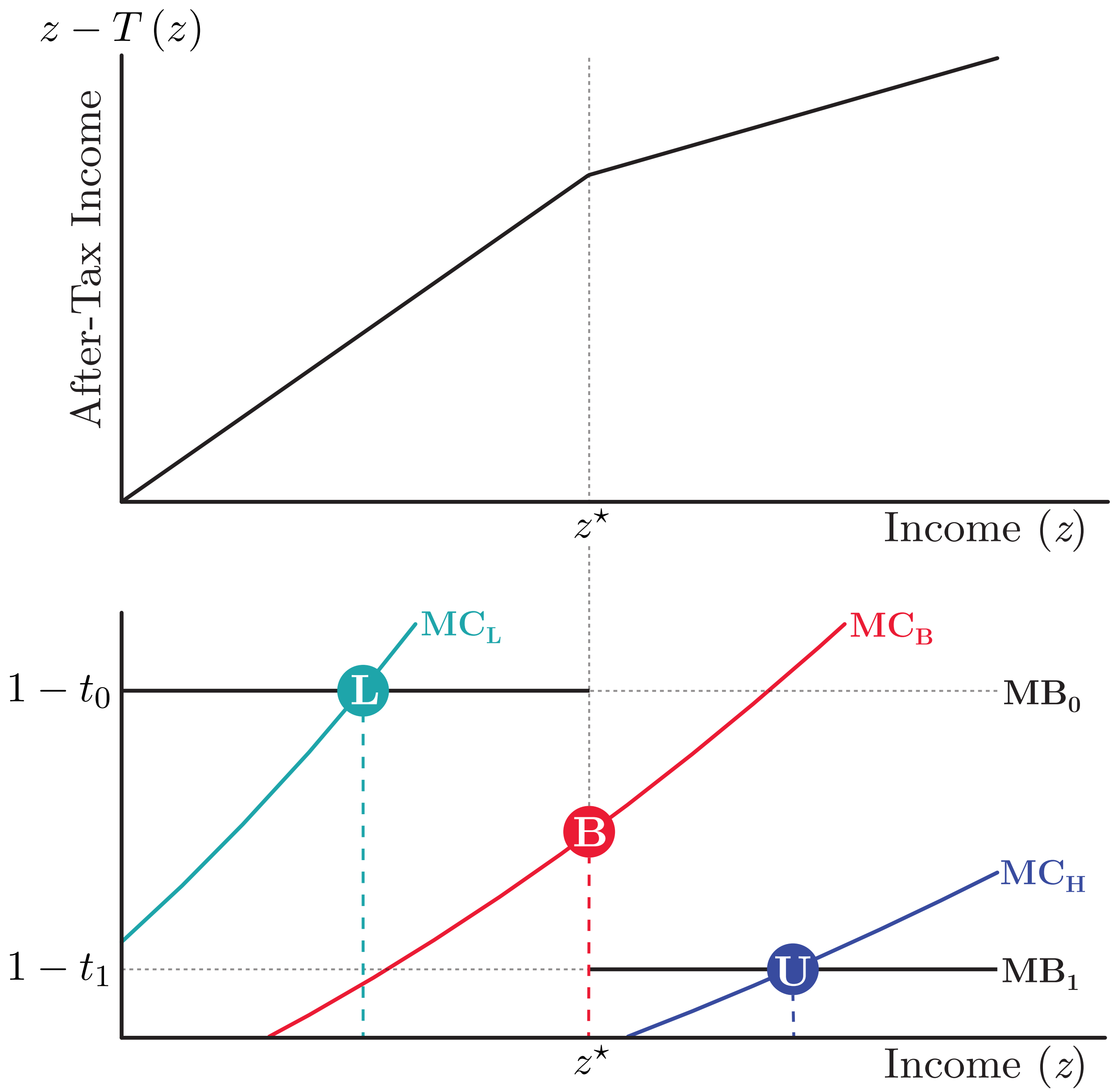



**Figure Notes:** The top of this figure depicts the budget set of taxpayers facing the tax schedule described in equation 2. The bottom of the figure shows several examples of taxpayer choices. The horizontal black lines on the figure show the marginal benefit taxpayers derive from earning an additional dollar of income. The marginal benefit drops discontinuously at $z^\star$ due to the kink in the taxpayers' budget constraint. The three upward sloping colored lines are marginal cost curves for three different types of taxpayers. Taxpayer L has a marginal cost curve which intersects the marginal benefit at a point below $z^\star$, making them a lower bracket taxpayer. Taxpayer U has a marginal cost curve which intersects the marginal benefit at a point above $z^\star$, making them an upper bracket taxpayer. Taxpayer B has a marginal cost curve which does not intersects the marginal benefit function at any point, instead running through the discontinuity of the function. Because this taxpayer's marginal benefit of earning exceeds their marginal cost at every point below $z^\star$ but falls below their marginal cost at every point above $z^\star$ they will choose to locate at precisely the kink point $z^\star$.

agents who would prefer to locate above the kink point under $T_0$, but below the kink point under $T_1$:

$$\Pr\left\{z_1(\theta) < z^\star < z_0(\theta)\right\} = H_1(z^\star) - H_0(z^\star). \tag{7}$$

Figure 2 provides some intuition for why this occurs, depicting a set of taxpayers who would all make distinct choices under a linear tax schedule, but many of whom lump together at $z^\star$ when facing a kinked tax schedule.

It is important to note that this presentation ignores important complications that arise in real-world bunching applications due to frictions in taxpayer income choices. Taxable income data are not usually consistent with the existence of a mass point at a kink point: rather, evidence of bunching comes in the form of a diffuse clustering of taxpayers around a kink. Consistent with the existing literature on bunching methodology, I derive my main results by abstracting from such concerns. I revisit them in the discussion of my empirical application in section (3).

## 1.2 The Bunching Mass as a Sufficient Statistic

Without loss of generality, suppose that $G = 0$. Under assumptions 1–3 tax revenue under the kinked tax schedule (2) is simply:

$$R(z^\star) \equiv \underbrace{t_0 \int_0^{z^\star} z h_0(z)\,\mathrm{d}z}_{\text{revenue below the kink}} + \underbrace{t_0 z^\star\left(H_1(z^\star) - H_0(z^\star)\right)}_{\text{revenue from bunchers}} + \underbrace{\int_{z^\star}^{\infty} \left[t_1(z - z^\star) + t_0 z^\star\right] h_1(z)\,\mathrm{d}z}_{\text{revenue above the kink}}. \tag{8}$$

The lower panel of figure 2 connects each of these three components of tax revenue to the relevant part of the observed distribution of taxable income.

Now, consider a tax reform displayed in figure 3, which infinitesimally lowers the kink point by $\mathrm{d}z^\star$ (moving it from $z^\star$ to $z^\star - \mathrm{d}z^\star$). This reform discretely increases the marginal tax rate within a small window of taxable income immediately below $z^\star$ while simultaneously reducing the demogrant (virtual income) associated with the second tax bracket. Figure 4 shows how such a reform would change taxpayer choices and the resulting the distribution of taxable income presented relative to the pre-reform case depicted in figure 2. Under assumptions 1–3, the revenue effect of this reform can be decomposed into four parts.

The *mechanical effect* of the reform reflects the fact that any agent with initial taxable income above the kink point now pays a higher marginal rate on their earnings in the interval $(z^\star - \mathrm{d}z^\star, z^\star]$, so that each such agent's tax liability increases by $(t_1 - t_0)\,\mathrm{d}z^\star$. The remaining three effects constitute the behavioral response to the reform.

Figure 2: The Observed (Pre-Reform) Distribution of Taxable Income

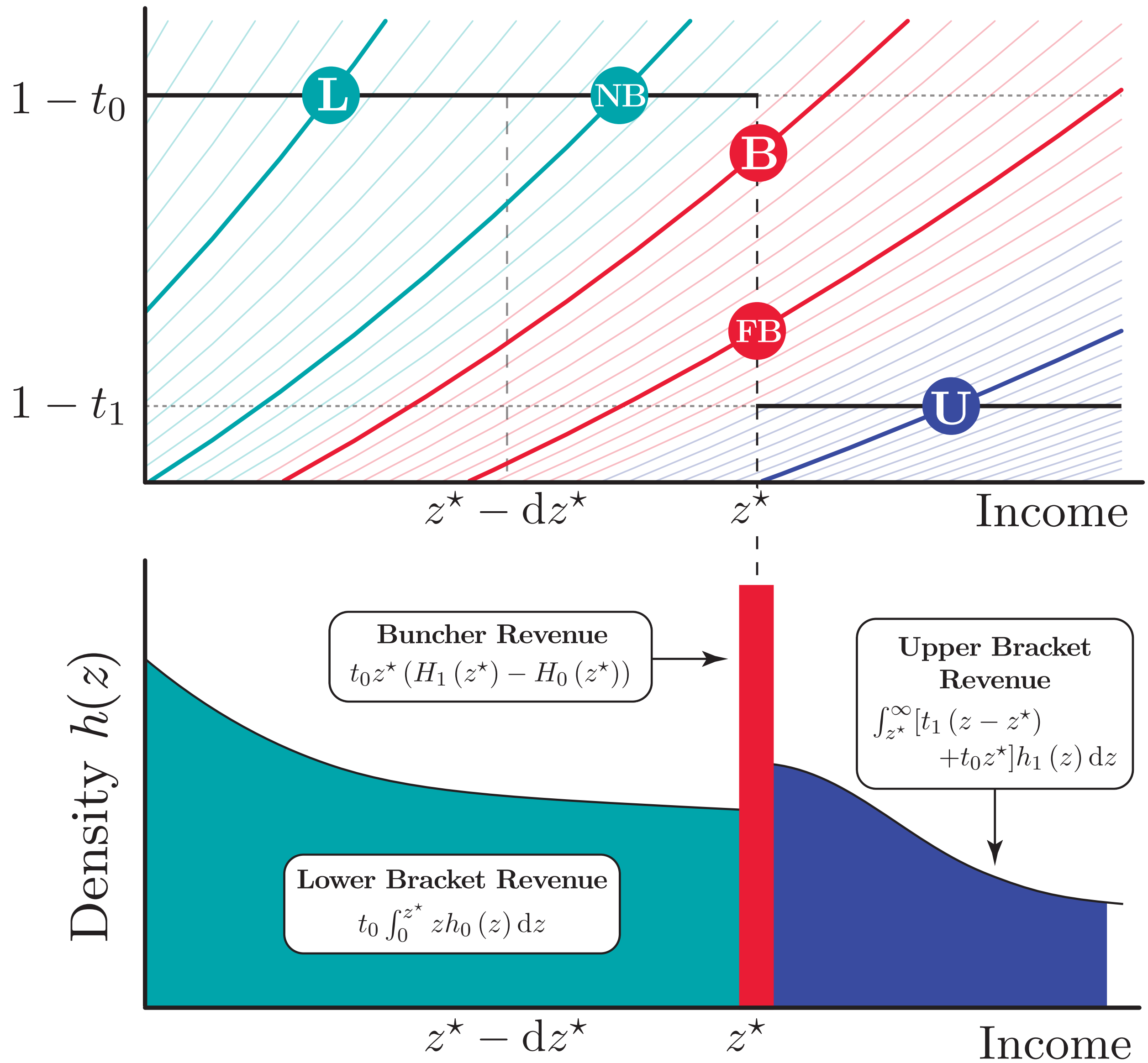
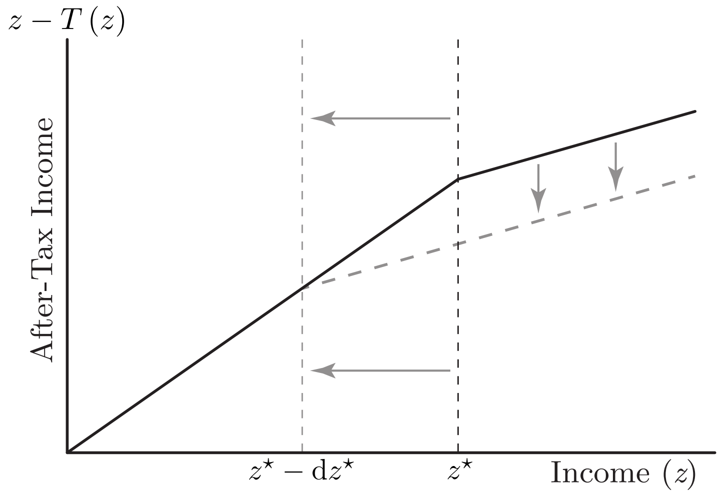



**Figure Notes:** This figure shows how a population of taxpayers with differing marginal costs of earning income who each face the same kinked budget constraint will induce an observed distribution of taxable income which features a mass point (the bunching mass) at the kink $z^\star$. The upper panel shows the marginal benefit of earning income for taxpayers who are facing the tax schedule described in equation 2 alongside a number of representative marginal cost curves. As depicted in figure 1, lower bracket taxpayers are those whose marginal cost curves intersect the marginal benefit below $z^\star$, upper bracket taxpayers are those whose marginal cost curves intersect the marginal benefit above $z^\star$, and bunchers are those marginal cost curves run through the discontinuity in the marginal benefit function at $z^\star$. The lower panel shows the distribution of taxable income that results. At the kink point, it features a mass point (the bunching mass) due to fact that the set of taxpayers with cost curves running through the discontinuity has a non-zero measure. Below the kink point the observed density of taxable income is identical to $h_0$: the density that would be observed under the linear tax schedule $T_0$. Above the kink, it is identical to $h_1$: the density under the linear tax schedule $T_1$. The bottom panel is labeled to show the tax revenue derived from different parts of the distribution of taxable income. Note, certain marginal costs curves in the top panel have been rendered darker colors and their corresponding taxpayer choices have been given labels. These labels are here for reference and will be discussed in the notes for figure 4.

Figure 3: Infinitesimally Lowering a Kink Point




The *relocation effect* captures the reduction in tax revenue that would occur if we simply moved the bunchers from $z^\star$ to $z^\star - \mathrm{d}z^\star$, holding constant the size of the bunching mass. The bunching agents have a pre-reform mass of $H_1(z^\star) - H_0(z^\star)$, with all bunchers paying tax liability of $t_0 z^\star$. If all these agents choose to bunch at the new kink point after the reform and no new agents join the bunching mass, then moving the kink point simply causes a loss of $-t_0 \mathrm{d}z^\star$ dollars of tax revenue from a group of agent with a mass of $H_1(z^\star) - H_0(z^\star)$.

However, the size of the bunching mass will not in fact remain constant, because some agents are at at the margin between bunching and not bunching pre-reform. Consider the group of agents initially in the lower tax bracket with $z_0(\theta) \in (z^\star - \mathrm{d}z^\star, z^\star)$. After the tax reform, these agents will become bunchers and their taxable income will fall from $z_0(\theta)$ to $z^\star - \mathrm{d}z^\star$, inducing a reduction in tax revenue by $t_0(z_0(\theta) - z^\star + \mathrm{d}z^\star)$. For the case of an infinitesimal kink point shift, the mass of these agents is $h_0(z^\star)\mathrm{d}z^\star$ and their pre-reform taxable income is $z_0(\theta) = z^\star$. Thus, the *new buncher effect*—which accounts for this additional loss in revenue—can be written as

$$-t_0 h_0(z^\star)(\mathrm{d}z^\star)^2.$$

There will also be a group of agents with $z_1(\theta) \in (z^\star - \mathrm{d}z^\star, z^\star)$ who will leave the bunching mass after the reform and choose a value of taxable income $z_1(\theta)$ in the higher bracket, generating additional tax revenue of $t_1(z_1(\theta) - z^\star + \mathrm{d}z^\star)$ relative the. For the case of an infinitesimal kink point shift, the mass of these agents is $h_1(z^\star)\mathrm{d}z^\star$ and their post-reform taxable income is $z_1(\theta) = z^\star$. Thus, the *former buncher effect*—which accounts for this additional gain in revenue—can be written as

$$t_1 h_1(z^\star)(\mathrm{d}z^\star)^2.$$

Figure 4: Effect of Lowering a Kink Point on the Distribution of Taxable Income

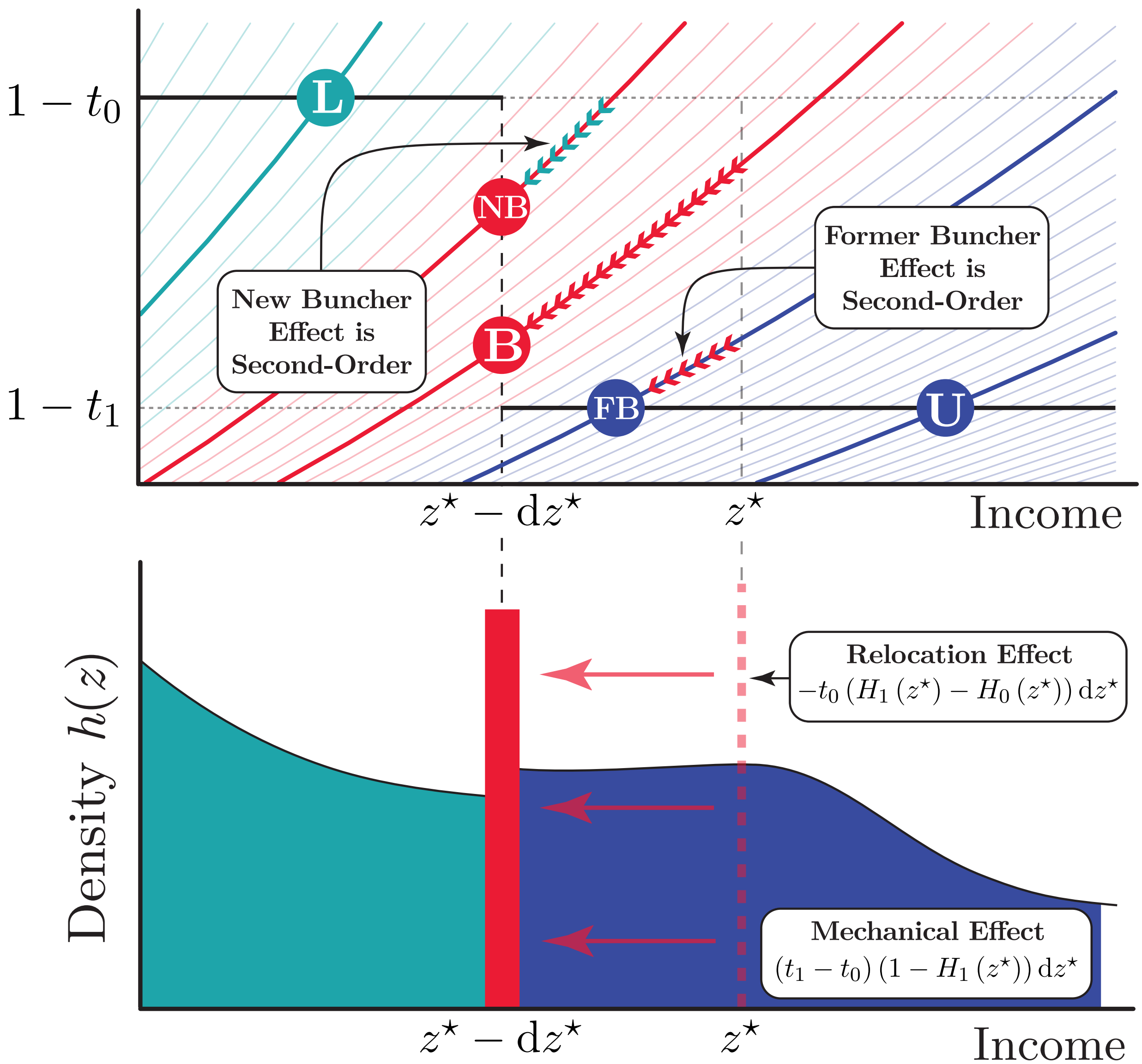



**Figure Notes:** This figure depicts the evolution of taxpayer choices presented in figure 2 after an infinitesimal lowering of the kink point from $z^\star$ to $z^\star - \mathrm{d}z^\star$, as well as the consequent changes in tax revenue. In the top panel, five marginal cost curves have been rendered in bolder colors and corresponding choices of taxable income have been given labels. Each of these curves provides a representative of one of a particular type of taxpayer. These same five labeled curves are also presented in figure 2 for reference. Some taxpayers are like L and U: their choice of taxable income is unaffected by the reform. However, the reform does extract additional revenue from taxpayers like U via the mechanical effect. Some taxpayers are like B: they are bunchers both before and after the reform. Unlike L and U, B's choice of taxable income changes, falling from $z^\star$ to $z^\star - \mathrm{d}z^\star$. Finally, some taxpayers are new bunchers (like NB) and former bunchers (like FB). NB is a taxpayer who is induced to bunch as a result of the reform, resulting in some reduction of their taxable income from some value in the interval $(z^\star, z^\star - \mathrm{d}z^\star)$ to the new kink point $z^\star - \mathrm{d}z^\star$. FB is a taxpayer who is induced to stop bunching as a result of the reform, with a reduction of taxable income from some value in the old kink point $z^\star$ to some new value in the interval $(z^\star, z^\star - \mathrm{d}z^\star)$. The relocation effect is the change in revenue that would result if we simply moved the bunching mass from $z^\star$ to $z^\star - \mathrm{d}z^\star$, holding constant its size. The central result of the paper is that the relocation effect provides a first-order approximation of the behavioral response effect. That is, the former and new buncher effects are second order.

The central result of this paper follows immediately from the fact that, whereas the relocation effect is first-order, both the new and former buncher effects are second-order. As noted above, this is true in spite of the fact that moving a kink point has a first-order effect on the size of the bunching mass.[12] However, this change has no first-order revenue effect. The informal derivation above provides some intuition for why: the new and former buncher effects consist of an infinitesimal change in revenue for each member of an infinitesimally small group. By contrast, the relocation effect consists of an infinitesimally small change in revenue for a large group of taxpayers. Therefore, the relocation effect alone accounts for the first-order revenue impact of the behavioral response to moving a kink point. But notice, the relocation effect is directly proportional to the bunching mass, leading to the following striking result.

**Theorem 1** (Sufficiency of the Bunching Mass)**.** *Under assumptions 1, 2, and 3, if the tax schedule is piecewise linear with two tax brackets (as defined in equation 2) then the first-order revenue effect of decreasing the kink point $z^\star$ is*

$$-R'\left(z^\star\right) = \underbrace{\left(t_1 - t_0\right)\left(1 - H_1\left(z^\star\right)\right)}_{\textit{mechanical effect}} \underbrace{-t_0\left(H_1\left(z^\star\right) - H_0\left(z^\star\right)\right)}_{\textit{behavioral response effect}}. \tag{9}$$

*The bunching mass is a sufficient statistic for the revenue impact of the behavioral response to the reform. The probability of locating above the kink is a sufficient statistic for the mechanical revenue effects of the reform.*

# 2 Evaluation without Elasticities

In this section, I build on Theorem 1 and its corollary to demonstrate how bunching designs can be used to evaluate proposed reforms to progressive, piecewise linear tax schedules. I first present a novel test for the Pareto efficiency of an observed tax schedule. This amounts to a test of whether a given tax schedule is locally on the wrong side of the Laffer curve, so that a local tax increase (implemented by decreasing a tax bracket threshold) would actually decrease revenue. Next, I characterize the welfare effects of this reform according to standard welfarist objective functions. Finally, I discuss existing techniques for predicting the revenue effects of tax reforms, showing that they do not correctly account for the effect of changing tax brackets.

[12] Indeed, differentiating equation (7) with respect to $z^\star$ reveals that the reform changes the size of the bunching mass by $h_1\left(z^\star\right) - h_0\left(z^\star\right)$.

## 2.1 An Empirical Test of Pareto Efficiency

Corollary 1 immediately suggests a simple test for Pareto efficiency. If a marginal reduction in the location of a kink point generates revenue losses due to the behavioral response effect which exceed the revenue generated by the mechanical effect, then the tax schedule is inefficient.

**Theorem 2** (Test for Pareto Efficiency)**.** *Under assumptions 1, 2, and 3, then a piecewise linear tax schedule with two tax brackets (as defined in equation 2) which is Pareto efficient must satisfy*

$$R'\left(z^{\star}\right) \leq 0. \tag{10}$$

*If $R'\left(z^{\star}\right) = 0$, then such a tax schedule must also satisfy the second-order condition*

$$R''\left(z^{\star}\right) = t_1 h_1\left(z^{\star}\right) - t_0 h_0\left(z^{\star}\right) < 0. \tag{11}$$

*These conditions are empirically testable using the observed (pre-reform) distribution of taxable income.*

*Proof.* Consider the tax reform depicted in figure 5 which infinitesimally increases $z^{\star}$. If $R'\left(z^{\star}\right) > 0$, then such a reform would increase tax revenue while also expanding the budget set faced by taxpayers. By revealed preference, all taxpayers weakly prefer this new tax schedule. Thus, whenever equation (10) does not hold there exists some Pareto-improving increase in the location of the kink point.

The second order-condition (11) provides a test for whether or not current the tax schedule is at a local minimum of tax revenue. If $R'\left(z^{\star}\right) = 0$ but $R''\left(z^{\star}\right) \geq 0$, then there once again exists a Pareto-improving tax reform that increases the location of the kink point, since all taxpayers prefer such a schedule and it generates (weakly) greater tax revenue. This condition can be verified empirically because the density terms that appear on the right-hand side of equation (11)—while not directly observable—are identified as limit points of the observed income distribution.[13] □

Theorem 2's test of Pareto efficiency is an analogue of the test for local Laffer effects initially proposed by Werning [2007]. However, no prior implementations of such tests require can account for the behavioral

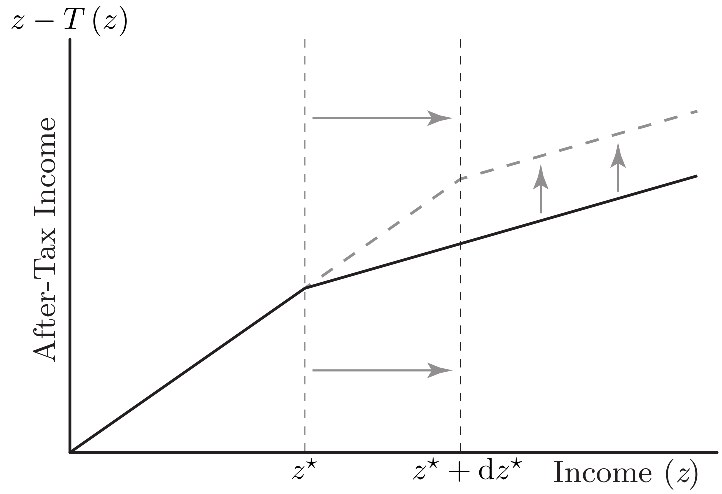
[13] Specifically, taking the limit of the observed density from below the kink gets us one of these densities

$$\lim_{z \to z^{\star}-} h\left(z\right) = \lim_{z \to z^{\star}} h_0\left(z\right) = h_0\left(z^{\star}\right)$$

and from above the kink gets us the other

$$\lim_{z \to z^{\star}+} h\left(z\right) = \lim_{z \to z^{\star}} h_1\left(z\right) = h_1\left(z^{\star}\right).$$

Figure 5: Infinitesimally Raising a Kink Point




response to taxation without committing to a specific model of taxpayer behavior and calibrating model parameters.By contrast, this test is accounts for the behavioral response to a changing tax bracket thresholds without using any information beyond the observed distribution of taxable income. This robustness comes at the cost of a narrow scope of inquiry however, as the test is uninformative about other tax reforms of interest.

In section B, I expand on this Pareto efficiency test, constructing additional testable conditions for efficiency using an analogue of the two-bracket reforms discussed by Bierbrauer et al. [2020].

## 2.2 Welfare Effects of Moving Kinks

The efficiency results above demonstrate that bunching can be used to provide a nonparametric test of the Pareto principle, but this is a very weak criterion. It is also of interest to consider welfare effects under a stronger definition of welfare.

The sharpest welfare effect results are obtained by assuming a Rawlsian (revenue-maximizing) objective. A Rawlsian planner has no reason to forgo any revenue gains, irrespective of whether these come at the cost of some agents private welfare.[14]

[14] Here, I am adopting the standard optimal tax theory convention of labeling the maximin objective

$$\max\left\{\min_{\theta\in\Theta} u\left(c\left(\theta\right), z\left(\theta\right); \theta\right)\right\}$$

as a "Rawlsian" welfare maximization problem. I also adopt the standard assumption that this maximization problem is equivalent to

$$\max\left\{\min_{\theta\in\Theta} c\left(\theta\right)\right\},$$

so that the agent with the lowest level of utility are those with the lowest levels of consumption. Further assuming that the marginal dollar of tax revenue finances increases in the demogrant $G$, the social planner's problem reduces to one of revenue maximization.

**Theorem 3** (Test for Rawlsian Optimality)**.** *Under assumptions 1, 2, and 3, a piecewise linear tax schedule with two tax brackets (as defined in equation 2) which is Rawlsian-optimal (revenue-maximizing) must satisfy*

$$R'(z^\star) = 0, \tag{12}$$

*and*

$$R''(z^\star) = t_1 h_1(z^\star) - t_0 h_0(z^\star) < 0. \tag{13}$$

This theorem simply states that the location of a kink point in a revenue-maximizing tax schedule must satisfy standard necessary and sufficient conditions for a local maximum.[15] As in theorem 2, the second-order condition (13) ensures that the kink is not located at a local minimum.

### 2.2.1 Bergsonian Welfare Functions

Next, we the case where the planner may have some reason to forgo revenue-increasing tax reforms. For this discussion, it will be helpful to first reframe the finding presented in Theorem 1 in the language of fiscal externalities. The fiscal externality of a given tax reform is the ratio of the behavior response effect to the mechanical effect of the reform. This metric provides an intuitive summary of the efficiency costs of a reform, reflecting the revenue loss due to the behavioral response per dollar raised through the mechanical effect of the reform.

**Corollary 1** (Fiscal Externality is Identified)**.** *Under assumptions 1, 2, and 3, if the tax schedule is piecewise linear with two tax brackets (as defined in equation 2) then the fiscal externality (FE) of marginally decreasing the kink point $z^\star$ is*

$$FE(z^\star) \equiv -\frac{t_0\left(H_1(z^\star) - H_0(z^\star)\right)}{(t_1 - t_0)\left(1 - H_1(z^\star)\right)}, \tag{14}$$

*and is identified by the observed (pre-reform) distribution of taxable income.*

Suppose we adopt a standard, additively separable Bergsonian welfare function as our measure of social welfare in a society with the two bracket tax schedule defined in equation (2). That is, suppose that social welfare when the kink point is located at $z^\star$ is

$$W(z^\star) \equiv \frac{\int \omega\left(v(\theta; z^\star)\right) \mathrm{d}F(\theta)}{\lambda} + R(z^\star), \tag{15}$$

[15]Note, theorem 3 does not provide a test for whether the kink point is at a globally revenue-maximizing location, as $R(z^\star)$ need not be a strictly concave function in general.

where $v(\theta; z^{\star})$ is the indirect utility function of a type $\theta$ agent, $\lambda$ is the marginal social value of government revenue, and $\omega(\cdot)$ is some weakly increasing, weakly convex function of utility.

I define the marginal rate of substitution for a type $\theta$ agent earning income $z$ as

$$MRS(z,\theta) \equiv \frac{\partial u(c,z;\theta)/\partial z}{\partial u(c,z;\theta)/\partial c},$$

where consumption $c \equiv z - T(z)$. Further, I shall denote the marginal social welfare weight of a type $\theta$ agent as

$$g(\theta) \equiv \frac{\omega'(v(\theta; z^{\star}))}{\lambda} \frac{\partial u(c,z;\theta)}{\partial c}.$$

The welfare weight reflects the marginal social value of a dollar of private consumption for a type $\theta$ agent.

I assume that the marginal dollar of government revenue is used to finance increases in the demogrant (or something with equivalent marginal social value). Thus, given the above definition, the average marginal social welfare weight is

$$\int g(\theta)\,\mathrm{d}F(\theta) = 1.$$

That is to say the marginal social value of a dollar of government revenue is one.[16]

The first-order welfare effect of decreasing the location of the kink point can then be written as

$$-W'(z^{\star}) = \underbrace{(t_1 - t_0)\int_{z^{\star}}^{\infty}(1-\bar{g}(z))\,h_1(z)\,\mathrm{d}z}_{\text{mechanical effect}} \quad \underbrace{-t_0\left(H_1(z^{\star}) - H_0(z^{\star})\right)}_{\text{revenue effect of behavioral response}}$$
$$\underbrace{-\int_{\{\theta: z_1(\theta) < z^{\star} < z_0(\theta)\}} g(\theta)\left(1 - t_0 - MRS(z^{\star},\theta)\right)\mathrm{d}F(\theta)}_{\text{utility effect of behavioral response}} \qquad (16)$$

where $\bar{g}(z) \equiv \mathbb{E}[g(\theta)\,|z(\theta) = z]$ is the average marginal social welfare weight of agents with an income of $z$ under the current tax schedule.

Notice, the first-order welfare effect of decreasing the location of a kink point differs from most similar expressions found in the optimal tax literature, because the behavioral response to this reform has a first-order effect on welfare for individuals in the bunching mass. To understand this, recall that the absence of such first-order welfare effects depends critically on the envelope theorem. But for bunchers, the envelope condition

[16] The marginal cost of increasing the demogrant is 1 and the marginal social benefit is $\int g(\theta)\,\mathrm{d}F(\theta)$. They must be equalized in the optimal tax system.

need not be satisfied.[17] In particular, for at least some types of agents who choose to bunch we have

$$1 - t_0 - MRS\left(z^{\star}, \theta\right) > 0. \tag{17}$$

Thus, reducing the location of the kink point has a first-order welfare effect because it has a first-order effect on the private welfare of bunchers, decreasing utility by

$$\left(1 - t_0 - MRS\left(z^{\star}, \theta\right)\right) \frac{\partial u}{\partial c}$$

for any type $\theta$ who is currently bunching and continues to bunch following the reform. The third term of equation (16) multiplies by this utility effect by $G'\left(V\left(\theta; z^{\star}\right)\right)$ and integrates over all agents who are bunching to obtain the total first-order welfare effect.[18]

The exception to the general failure of the envelope theorem are the so-called *marginal bunchers* that Saez [2010] used to motivate his original bunching method. These are taxpayers who choose to locate at $z^{\star}$ under both the observed tax schedule $\left(z\left(\theta\right) = z^{\star}\right)$ and the counterfactual linear tax schedule with tax rate $t_0$ $\left(z_0\left(\theta\right) = z^{\star}\right)$. For such individuals, the envelope condition holds

$$1 - t_0 - MRS\left(z^{\star}, \theta\right) = 0.$$

Notice, in my framework, these marginal bunchers can also be thought of as representing the *new bunchers* who will enter the bunching mass following an infinitesimal decrease in the location of the kink point. This explains why equation (16) does not include any terms accounting for the welfare implications the *new buncher effect*: for an infinitesimal change, there are no such effects to account for.

On the other hand equation (16) does account for a direct welfare impact of the *former buncher effect*. Former bunchers represent a second type of marginal buncher. For an infinitesimal reform, the former bunchers are taxpayers who choose to locate at $z^{\star}$ under both the observed tax schedule $\left(z\left(\theta\right) = z^{\star}\right)$ and the counterfactual linear tax schedule with tax rate $t_1$ $\left(z_1\left(\theta\right) = z^{\star}\right)$. This implies that the indifference curves of the former bunchers are tangent to the slope of the higher tax bracket before the reform,

$$1 - t_1 - MRS\left(z^{\star}, \theta\right) = 0.$$

[17] Figure 1 shows this visually. The marginal cost curve of a typical buncher does not intersect the marginal benefit function rather it runs through the discontinuity in this function at $z^{\star}$. Thus, their marginal benefit of earning exceeds their marginal cost at $z^{\star}$: the envelope condition fails.

[18] In the optimal tax theory literature, this type of first-order welfare effect usually only appears in the presence of some kind of market failure. For example, externalities, labor market frictions, and behavioral biases can induce first-order welfare effects to behavioral responses to tax reforms.

Thus, the first-order effect of the reform on their utility is the same as for taxpayers located above $z^\star$:

$$\left(1-t_0-MRS\left(z^\star,\theta\right)\right)\frac{\partial u}{\partial c}=\left(t_1-t_0\right)\frac{\partial u}{\partial c}.$$

Building on these results, I propose an empirical test of welfarist optimality of the observed tax schedule given a known function describing the average marginal social welfare weight at each income level $\bar{g}\left(\cdot\right)$. First, I introduce one additional convenient piece of notation. Let

$$\hat{g}_+\left(z\right)\equiv\mathbb{E}\left[g\left(\theta\right)|z\left(\theta\right)>z\right]$$

be the average marginal social welfare weight of taxpayers with income above the kink point.

**Theorem 4** (Infeasible Test for Welfarist-Optimality)**.** *Under assumptions 1, 2, and 3, a piecewise linear tax schedule with two tax brackets (as defined in equation 2) which is welfare-maximizing according to (15) must satisfy*

$$\underbrace{1-\hat{g}_+\left(z^\star\right)}_{\text{mechanical welfare effect}}+\underbrace{\left(1+\frac{t_1-t_0}{t_0}k\bar{g}\left(z^\star\right)\right)FE\left(z^\star\right)}_{\text{behavioral response welfare effect}}=0 \tag{18}$$

*where*

$$k\equiv\frac{\mathbb{E}\left[g\left(\theta\right)\left(1-t_0-MRS\left(z^\star,\theta\right)\right)|z\left(\theta\right)=z^\star\right]}{\bar{g}\left(z^\star\right)\left(t_1-t_0\right)} \tag{19}$$

*is the ratio of the true first-order welfare effect of the reform caused by the behavioral response of bunchers relative to the upper bound of this effect: $\bar{g}\left(z^\star\right)\left(t_1-t_0\right)$, and is not identified by observed data.*

*If the left-hand side of equation (4) is positive (negative), then there exists a welfare-improving decrease (increase) of the kink point.*

To gain some intuition about Theorem 4, consider what condition (18) would look like if the welfare of the bunchers was irrelevant: $\bar{g}\left(z^\star\right)=0$. In this case, the condition simply states that at the optimum the fiscal externality of the reform must be equal to the average gain in social welfare caused by the mechanical effect of the reform. This average gain is the difference between the marginal social value of a dollar of government revenue (recall, this is equal to 1) and the average marginal social value of a dollar of private consumption for the agents who are impacted by the mechanical effects (which is $\hat{g}_+\left(z^\star\right)$). Thus, unlike in the Rawlsian case, at the optimum the social planner will sometimes opt to forgo feasible revenue-increasing tax reforms because the redistributive benefits of the reform may be insufficient to justify the welfare loss causes by the behavioral response to the reform (as measured by the fiscal externality).

Returning to the general case, where $\bar{g}(z^\star) \neq 0$, the fact that moving a kink point causes behavioral responses which have a first-order welfare effect simply inflates the welfare implications of the fiscal externality. This reflects the fact that the welfare loss caused by the behavioral response includes both lost revenue *and* these first-order welfare impacts.

Without committing to strong functional form assumptions, the value of $k$ is unidentified, so theorem 4 cannot be used as the basis for an empirical test of the optimality of a tax schedule. However, given known values for the average welfare weight parameters $\hat{g}_+(z^\star)$ and $\bar{g}(z^\star)$, we can obtain an empirical test motivated by a partial identification approach to the problem.

**Corollary 2** (Feasible Test for Welfarist-Optimality)**.** *Under assumptions 1, 2, and 3, a piecewise linear tax schedule with two tax brackets (as defined in equation 2) which is welfare-maximizing according to (15) must satisfy*

$$1+\left(1+\left(\frac{t_1-t_0}{t_0}\right)\bar{g}(z^\star)\right)FE(z^\star)<\hat{g}_+(z^\star)<1+FE(z^\star). \tag{20}$$

*On the other hand, if*

$$1+\left(1+\left(\frac{t_1-t_0}{t_0}\right)\bar{g}(z^\star)\right)FE(z^\star)\geq\hat{g}_+(z^\star)$$

*then there exists a welfare-improving decrease of the kink point. If*

$$1+FE(z^\star)\leq\hat{g}_+(z^\star)$$

*then there exists a welfare-improving increase of the kink point.*

This test simply checks to make sure that condition (18) from theorem 4 is satisfied for at least one $k\in(0,1)$. If there is no such value of $k$, then we can conclude that a welfare-improving movement of the kink point exists. Importantly, a condition (18) is a necessary but insufficient condition for welfare maximization.

# 3 Empirical Application

This section applies the ideas discussed above to revisit an empirical application from Saez [2010]. Saez applies his bunching methodology to estimate elasticities using a sample of US personal income tax returns: the IRS Individual Public Use Tax Files. In particular, I revisit his investigation of bunching at the first kink point of the Earned Income Tax Credit (EITC) schedule between 1995 and 2004. While the dataset has changed somewhat since Saez's original exercise, my work closely follows another recent replication of Saez [2010] conducted by Bertanha et al. [2021].[19]

Between 1995 and 2004, the EITC provided a 34% subsidy on the marginal dollar of family earnings in a single child household below \$8580 (in 2008 \$). For both groups, the marginal subsidy shrank to zero the earnings levels above these thresholds, introducing a convex kink in the household budget constraint. These households also faced a second convex kink point at \$15740 (in 2008 \$) where the marginal subsidy became a marginal tax of 16%, as the subsidy was phased out.

Recall, the results presented in the preceding sections apply to a scenario where labor supply choices are frictionless. However, real-world taxable income distributions do not appear to support the assumption of a frictionless model. As Saez [2010] noted, in the IRS data some groups of EITC recipients do exhibit evidence of bunching at the first kink point, but this is in the form of a diffuse lump in the density of taxable income at the kink. This can be rationalized by a model where taxpayers' choice of taxable income is subject to frictions. For example, perhaps it includes some noise which is outside their control.

To address this, I employ the filtering method proposed by Bertanha et al. [2021]. Their method relies on assuming that the observed taxable income of agents whose optimal choice would be at the kink point includes some optimization error but the observed choices of other agents are undistorted. In particular, the observed taxable income of a type $\theta$ agent be $z^o(\theta)$ and suppose that this function satisfies

$$\log\left(z^o(\theta)\right) = \begin{cases} \log\left(z^\star\right) + \epsilon & \text{if } z(\theta) = z^\star \\ \log\left(z(\theta)\right) & \text{if } z(\theta) \neq z^\star \end{cases}$$

where $\epsilon$ is an optimization error independently drawn from the distribution $G$. Further, let the support of $G$ be the bounded interval $[\underline{\epsilon}, \overline{\epsilon}]$ where $\underline{\epsilon} < 0 < \overline{\epsilon}$. The filtering method of Bertanha et al. [2021] proposes estimating the underlying frictionless distribution of taxable income by borrowing techniques from the regression discontinuity literature. In particular, this method involves fitting a 7-th order polynomial to the empirical cumulative distribution function of taxable income while excluding any points with logged taxable income in the interval$[\log(z^\star) - \underline{\epsilon},\ \log(z^\star) + \overline{\epsilon}]$ and allowing for a discontinuity in the function at $z^\star$. The predicted values from this polynomial provide estimates of the frictionless cumulative distribution function within the excluded region. The size of the discontinuity in the fitted function is the size of the bunching mass.

[19] Using my version of the dataset (provided by the NBER), I have been able to reproduce the results of Bertanha et al. [2021] extremely closely, but not perfectly. The small discrepancy may be accounted for by some minor changes in the dataset or in the labeling of certain variables.

## 3.1 Accounting for Heterogeneous Tax Schedules

Note, for a complete analysis of the fiscal externality of moving the EITC kink, we need to account for all other ways that the taxable income choices of EITC taxpayers influences government revenues/expenditure. Thus, the statutory marginal tax rates from the EITC schedule are not the relevant marginal rates for our analysis. Instead, I will employ measures of marginal rates which incorporate state income taxes, payroll taxes, and the phaseout of two major welfare programs (AFDC/TANF and SNAP).

I use the NBER's Taxsim tax calculator to determine for each taxpayer in the dataset their the total marginal tax rate on the wage income. This includes the impact of things like other federal tax credits, the effect of state tax policies, and FICA. The state income tax calculations include state-level EITC top-up programs which are usually specified as a percentage of federal EITC payments, and therefore have kink points which almost always line up with the federal kink points. I calculate AFDC/TANF payments using a state-level benefit calculator created by Kroft et al. [2020] and employ a crude approximation to the impact of SNAP (food stamps), assuming a constant phase out rate of 20% applies throughout the relevant range of income, following Kleven (2020). Again following Kleven (2020), I assume a 54% take-up rate for both AFDC/TANF and SNAP.

Importantly, this has the implication that the marginal tax rates on either side of the first EITC kink may differ substantially across the taxpayers in my analysis, as tax/benefit policy is not constant across states nor within states across years. A valid empirical analysis therefore requires extending the baseline results presented in sections 1 and 2. For simplicity, I will present a two bracket extension, though this can be easily generalized to the multibracket case.

Suppose that type $\theta$ agents face the following two bracket tax schedule

$$T\left(z, z^{\star}; \theta\right) \equiv \begin{cases} t_0\left(\theta\right) z & \text{if } z \leq z^{\star} \\ t_1\left(\theta\right) z + \left[t_0\left(\theta\right) - t_1\left(\theta\right)\right] z^{\star} & \text{if } z > z^{\star} \end{cases},$$

where $t_0\left(\theta\right)$, and $t_1\left(\theta\right)$ are type-$\theta$-specific tax schedule parameters.[20] The first-order revenue effect a tax reform which infinitesimally reduces $z^{\star}$ is[21]

[20] This result can also be extended to cases where the kink point is a type-specific parameter.

[21] The fiscal externality of this reform can be written as

$$FE \equiv -\frac{\mathbb{E}\left[t_0\left(\theta\right) | z_1\left(\theta\right) < z^{\star} < z_0\left(\theta\right)\right]}{\mathbb{E}\left[t_1\left(\theta\right) - t_0\left(\theta\right) | z\left(\theta\right) > z^{\star}\right]} \frac{H_1\left(z^{\star}\right) - H_0\left(z^{\star}\right)}{1 - H_1\left(z^{\star}\right)}. \tag{21}$$

$$-R'(z^\star) = \mathbb{E}\left[t_1(\theta) - t_0(\theta) \,|z(\theta) > z^\star\right](1 - H_1(z^\star)) - \mathbb{E}\left[t_0(\theta)\,|z_1(\theta) < z^\star < z_0(\theta)\right](H_1(z^\star) - H_0(z^\star)). \tag{22}$$

**Results**

Equation 21 can be empirically estimated using my data and tax/benefit calculators. I estimate the expected tax rates on either side of the kink point $\mathbb{E}\left[t_0(\theta)\,|z_0(\theta) = z^\star\right]$ and $\mathbb{E}\left[t_1(\theta)\,|z_1(\theta) = z^\star\right]$ via local linear regression. I then assume that the average marginal rate just below the kink for bunchers is the same as the observed average marginal rate just below the kink:

$$\mathbb{E}\left[t_0(\theta)\,|z_1(\theta) < z^\star < z_0(\theta)\right] = \mathbb{E}\left[t_0(\theta)\,|z_0(\theta) = z^\star\right].$$

I further assume that

$$\mathbb{E}\left[t_0(\theta)\,|z(\theta) = z^\star\right] = \mathbb{E}\left[t_0(\theta)\,|z(\theta) > z^\star\right]$$

and

$$\mathbb{E}\left[t_1(\theta)\,|z(\theta) = z^\star\right] = \mathbb{E}\left[t_1(\theta)\,|z(\theta) > z^\star\right]$$

so that the marginal rate in each bracket for taxpayers who locate near the kink are, on average, the same as those for taxpayers who locate away from the kink point.[22]

Table 2 presents estimates of each of the components that make up the right hand side of equation (22): the size of the bunching mass, the fraction of taxpayers above the kink point, and the estimated average tax rates on either side of the kink. I show the results for all taxpayers, as well as for several subgroups of taxpayers. The size of the estimated bunching mass varies substantially across groups, with 1.4% of all taxpayers in the sample bunching compared to 8.1% of unmarried, self-employed taxpayers. Contrary to what statutory rates would imply, the marginal rates on either side of the EITC kink are positive for the average taxpayer in my sample. This is consistent with the observation made by Bierbrauer et al. [2020], who note that the negative marginal rates of the EITC are offset by the large positive marginal rates induced by the phaseout of welfare benefits throughout most the relevant range of income. Only in the case of unmarried, self-employed taxpayers is the marginal tax rate just below the kink negative. The estimated size of the kink is also slightly smaller than what statutory EITC rates would imply.

[22] The ideal version of this exercise would use Taxsim to determine the exact marginal tax rates on either side of the kink for every taxpayer in the sample. I defer this more comprehensive analysis to future work.

Table 1: Components of Equation (22)

| | All Taxpayers | Unmarried | Self-Employed | Unmarried & Self-Employed |
|---|---|---|---|---|
| $H_1(z^\star) - H_0(z^\star)$ | 0.014 | 0.022 | 0.034 | 0.081 |
| $1 - H_1(z^\star)$ | 0.902 | 0.838 | 0.885 | 0.741 |
| $\mathbb{E}[t_0(\theta)\,\lvert z(\theta) = z^\star]$ | 0.021 | 0.002 | 0 | $-0.01$ |
| $\mathbb{E}[t_1(\theta)\,\lvert z(\theta) = z^\star]$ | 0.304 | 0.309 | 0.321 | 0.301 |

Table 2 presents a first-order approximation to the change in average annual revenue that would have resulted from reducing the first EITC kink by \$100 over the sample period.[23] This effect is broken down into the part due to the mechanical effect of the reform and the part due to the behavioral response effect. The last row of the table also reports the estimated fiscal externality of reducing the location of the kink.

In every case the behavioral response effect is quite small relative to the mechanical effect. This can be seen most clearly in the small fiscal externality estimates, which suggest that the for each dollar raised through the mechanical effect of reducing the first EITC kink, about a tenth of a cent would be lost to this behavioral response. Note as well that the behavioral response effect is actually positive for self-employed taxpayers because for these individuals the average marginal tax rate in the lower bracket is negative. Consequently, reducing the location of the first EITC kink causes both a mechanical increase in revenue from all taxpayers above the kink who now receive a lower total subsidy amount and an increase in revenue due to the behavioral response effect because moving the bunching taxpayers to a lower taxable income reduces the subsidy they receive.

Theorem 2 provided inequalities which must be satisfied in order for the observed tax schedule to be plausibly welfare-optimal (ignoring income effects and extensive margin responses). One of these conditions states that $1 + FE(z^\star)$ provides an upper bound on the average welfare weight above the kink point ($\hat{g}_+(z^\star)$). Thus, these results imply that the $\hat{g}_+(z^\star) < 1$.

## 3.2 Frictions

The purpose of this paper is not to address the important outstanding question in bunching methodology of how best to deal with the fact that the observed distribution of taxable income does not exhibit sharp bunching at kink points and is therefore inconsistent with the frictionless labor supply model used to justify bunching methods. As noted above, I employ the filtering method proposed by Bertanha et al. [2021] to

[23] I obtain the first-order approximation by plugging the estimates from Table (1) into the right hand side of equation (22) and multiplying the result by \$100 to obtain an estimate of the change in average tax revenue per taxpayer as a result of the reform. To get the total revenue effect numbers estimates I simply scale this change by the average number of taxpayers in a given year in the sample.

Table 2: Revenue Effect of Moving the First EITC Kink (Single Child Households, 1995-2004)

| | All Taxpayers | Unmarried | Self-Employed | Unmarried & Self-Employed |
|---|---|---|---|---|
| Mechanical Effect (2008 \$) | $48,293,930$ | $23,938,589$ | $9,543,895$ | $2,195,578$ |
| Behav. Response Effect (2008 \$) | $-55,122$ | $-4,051$ | $123$ | $7,403$ |
| Total Revenue Effect (2008 \$) | $48,238,808$ | $23,934,538$ | $9,544,016$ | $2,202,982$ |
| Fiscal Externality | $-0.0011$ | $-0.0002$ | $0$ | $0.0034$ |

*Notes: The revenue effects reported in the first three rows of this table are first-order approximations to the average additional revenue per year that would result from reducing the first EITC kink by \$100.*

obtain an estimate of the underlying frictionless distribution of taxable income in my empirical application. However, in this section I will address another important aspect of a model with frictions: the manner in which they alter the fiscal externality of moving a kink point.

Suppose that the observed labor supply of taxpayers included optimization errors. That is

$$z = y\xi$$

where $y$ is the choice of labor supply the taxpayer would make according to a frictionless labor supply model and $\xi$ is a multiplicative optimization error. Let these errors be distribution according to $G$ for all taxpayers.

These frictions imply that bunching taxpayers are not only located at the kink point but rather scattered around at many different levels of taxable income, with some locating in the lower bracket and others in the top bracket. The first-order revenue effect of moving a kink point thus affects different bunching taxpayers in different ways. In appendix can be shown that the fiscal externality of reducing a kink point in this case becomes

$$FE^{f}\left(z^{\star}\right) = -\left(\frac{t_0 G\left(1\right)\frac{\mathbb{E}\left[z|y=z^{\star},z\leq z^{\star}\right]}{z^{\star}} + t_1\left(1-G\left(1\right)\right)\frac{\mathbb{E}\left[z|y=z^{\star},z>z^{\star}\right]}{z^{\star}}}{t_1 - t_0}\right)\left(\frac{H_y\left(z^{\star};t_1\right)-H_y\left(z^{\star};t_0\right)}{1-H_z\left(z^{\star};z^{\star}\right)}\right) \tag{23}$$

where $H_y\left(z;t\right)$ is the frictionless distribution of taxable income under tax rate $t$, $H_z\left(z;z^{\star}\right)$ is the observed distribution of taxable income when the kink point is located at $z^{\star}$, and $G\left(1\right)$ is the fraction of taxpayers induced to work below their optimum point by optimization frictions.

There are two important differences between this version of the fiscal externality and the frictionless version. On the one hand, the share of taxpayers above the kink in the frictionless distribution is higher than in the observed distribution

$$1-H_y\left(z^{\star};t_1\right) < 1-H_z\left(z^{\star};z^{\star}\right)$$

so the denominator is larger, attenuating the fiscal externality. On the other hand, the numerator has also changed because the effective marginal tax rate for bunchers is now

$$t_0 G(1) \underbrace{\left[\frac{\mathbb{E}[z|y = z^\star, z \leq z^\star]}{z^\star}\right]}_{\in(0,1)} + t_1 (1 - G(1)) \underbrace{\left[\frac{\mathbb{E}[z|y = z^\star, z > z^\star]}{z^\star}\right]}_{>1}$$

This reflects the variation in the marginal tax rate of bunchers due to the variation in their location. Notice, if $\mathbb{E}[z|y = z^\star] = z^\star$ (i.e. with mean 1 frictions) then this is just a weighted average of $t_0$ and $t_1$ so it will be greater than $t_0$. However, this is not guaranteed in general.

**Empirical Results**

The filtering method of Bertanha et al. [2021] that I applied to estimate a frictionless distribution of taxable income also generates an estimated error distribution. I assume that this serves as a good estimate for the error distribution $G$.[24] I use the empirically estimated average marginal tax rates from the prior section can be used in place of the fixed tax rates in equation (23). This allows me to estimate the effective marginal tax rate for bunchers.

Table 3 presents the estimated fiscal externality of decreasing the kink point explicitly accounting for optimization errors. The frictionless fiscal externality estimates are presented for comparison.[25] Accounting for frictions substantially changes the estimated fiscal externalities. For the set of all taxpayers, the fiscal externality implies that for each dollar of mechanical revenue raised by decreasing the kink point, 1 cent is lost to the compensated behavior response. While this number remains relatively small, it has increased in size nearly tenfold. For other groups, we also observe larger magnitude estimates than in the frictionless case. And, most notably, the fiscal externality for unmarried, self-employed individuals has changed sign: it is now negative. This group also has the largest magnitude fiscal externality: the estimates suggests that for each dollar of mechanical revenue raised from this group through a reduction in the first EITC kink, their behavioral response to the reform would result in a 5.6 cent reduction in revenue.

The reason for these differences from the frictionless results is that when there are optimization frictions the effective marginal tax rate for bunchers is a combination of both the marginal rate above the kink and the marginal rate below the kink. Because the marginal rate above the kink is much higher than the rate

[24] It is important to note that while my results are derived under the assumption that all taxpayers face optimization errors, the filtering method of Bertanha et al. [2021] requires the assumption that only bunchers are affected by these errors. Thus, the use of the estimated error distribution in this exercise is not not ideal. However, as Bertanha et al. [2021] "a proper deconvolution theory must be developed to tackle this problem". Cattaneo et al. [2018] have proposed one approach to doing this.

[25] These are simply the same estimates previously presented in Table **??**.

Table 3: Fiscal Externalities with Optimization Frictions

| | **All Taxpayers** | **Unmarried** | **Self-Employed** | **Unmarried & Self-Employed** |
|---|---|---|---|---|
| **Friction-Inclusive FE** | $-0.0104$ | $-0.0147$ | $-0.0228$ | $-0.0568$ |
| **Frictionless FE** | $-0.0011$ | $-0.0002$ | 0 | 0.0034 |

below, accounting for frictions drives up the fiscal externality. This effect is especially large because the estimated rate below the kink is so small. These effects were large enough to offset any attenuation of the fiscal externality resulting from the fact that the mechanical effect accounting frictions is also larger. While the fiscal externality estimates presented here may still be viewed as relatively small, this application illustrates the potential importance of explicitly accounting for optimization friction in policy analysis of kink point location.

## 3.3 What Doesn't Bunching Identify? Income and Participation Effects

Excluded from the analysis above are two potentially important components of the behavioral response to tax reforms: income effects and labor force participation (extensive margin) responses. The latter are of particular interest, as empirical work on the EITC suggests it may have substantial labor force participation effects. Here, I summarize results presented in appendix A.3 which show how these effects alter the revenue impact and fiscal externality of moving a kink point. Importantly, the bunching mass remains a sufficient statistic for the part of the fiscal externality of the tax reform which is attributable to substitution effects.

### 3.3.1 Income Effects

Suppose that assumption 3 does not hold because some agents have non-zero income effects. Consider once again the tax reform presented in figure (3) which infinitesimally reduces the location of the kink point $z^{\star}$ and notice that upper bracket taxpayers experience a reduction in their virtual income as a result of the reform. If leisure is a normal good, these taxpayers should be expected to increase their taxable income to partially offset this loss in income. Thus, moving the kink point changes the mapping from the distribution of types into the distribution of taxable income under tax schedule $T_1$ and, consequently, changes the observed distribution of taxable income above the kink point.

### 3.3.2 Labor Force Participation Effects

Assumption 1 together with convexity of the tax schedule rules out extensive margin responses by excluding the possibility that an agent's indifference curve may be tangent to the budget constraint at more than one point. One way to relax this assumption is to assume that agents face a fixed cost of labor force participation and they will only enter the labor force and earn positive income if the gains from doing so exceed this cost. In this scenario, changes in the location of the kink point can induce workers to enter or exit the labor force by changing the utility they receive conditional on participating. Returning to the tax reform presented in figure (3), there are two groups of taxpayers who experience such participation effects.

First, as noted above, a reduction in $z^\star$ lowers the virtual income associated with the upper tax bracket, decreasing the consumption of taxpayers who would choose to do so conditional on entering the labor force. For taxpayers at the margin of participating—those whose fixed cost of work is exactly offset by the benefits of work—this reduced consumption will induce them to exit the labor force. Second, recall that changing the location of a kink point has a first-order effect on the utility of agents located at the kink (as I discussed in section 2.2). Again, for taxpayers at the margin of participating, these first-order utility effects will induce them to exit. In both cases, the participation effects causes a reduction tax revenue.

### 3.3.3 Bringing It All Together

Although income and participation effects alters the revenue impact of moving a kink point, the bunching mass still identifies a key component of this impact: that attributable to the substitution effect. In particular, the fiscal externality of moving a kink point becomes

$$FE\left(z^\star\right) = FE_{sub}\left(z^\star\right) + FE_{inc}\left(z^\star\right) + FE_{part}\left(z^\star\right),$$

where

$$FE_{sub}\left(z^\star\right) \equiv -\frac{t_0\left(H_1\left(z^\star\right) - H_0\left(z^\star\right)\right)}{\left(t_1 - t_0\right)\left(1 - H_1\left(z^\star\right)\right)} \leq 0,$$

is the part of fiscal externality caused by the substitution effects of the reform, and is identified by the observed (pre-reform) distribution of taxable income. $FE_{inc}\left(z^\star\right)$ and $FE_{part}\left(z^\star\right)$ are the parts of the fiscal externality caused by income and participation effects (respectively). These terms are defined in detail in Appendix A.3. Importantly, neither is identified by the observed distribution of taxable income. Accounting for these effects in an ex ante evaluation of a proposed tax reform requires relying on external evidence about the probable magnitude of these effects.

**Other Extensive Margin Responses** Participation effects are not the only extensive margin responses that might be relevant to my analysis of the EITC. Small tax reforms can also cause taxpayers can also "jump" from one level of taxable income to another whenever the tax schedule induces a non-convex budget set. Even under assumption 1, such non-convexity introduces the possibility that taxpayer indifference curves can be tangent to the budget constraint at more than one point; that a taxpayer may be at the margin between two distinct choices of positive income. We might therefore be concerned that moving a convex kink in a tax schedule which also includes an non-convex kink can potentially induce such taxpayers to jump from one income level to another. In principle, these kinds of effects can be incorporated into the fiscal externality of the reform as I have done above with income and participation effects, but I do not discuss this here. Bergstrom and Dodds [2021] provide an extensive discussion of how these types of responses can affect optimal nonlinear income taxation.

### 3.3.4 Application to the EITC

The discussion above shows that the size of the bunching mass can tell us about one component of the behavioral response effects of moving the first EITC kink, but it may be important to account for the other aspects of the response. Reducing the kink point might generate a substantial fiscal externality due to taxpayers exiting the labor force at or above the kink point. Income effects—if present—would generate an increase in revenue from taxpayers increasing their incomes. Using income and participation elasticity estimates derived from the literature it would be possible to supplement my empirical analysis to account for these responses. However, correctly accounting for the fiscal externality of participation effects requires more detailed information about the change in total net government payments that occurs when an individual above the kink moves to zero income. Such an analysis is complex and beyond the scope of this paper. Bierbrauer et al. [2020] conduct an analysis of the efficiency consequences of the introduction of the EITC which accounts for these effects but confine their analysis to California due to the difficulty associated with conducting a full analysis.

# 4 What is the Bunching Elasticity?

This section revisits standard bunching methodology in light of the results presented above. Bunching-based elasticity estimates are closely connected to the behavioral response to moving a kink point but their application to other tax reforms of interest faces external validity challenges. The conditions that give rise to these challenges are characterized analytically and additionally explored in a simulation exercise.

## 4.1 Relating the Bunching Mass to Elasticities

Discussion of the standard bunching method requires a special case of the modeling framework that I introduced in section 1.1. In particular, suppose that each agent's type consists of a wage rate $w \in \mathbb{R}_{++}$ and a vector of other characteristics $\psi \in \Psi$ where $\Psi$ is a convex (possibly multidimensional) parameter space. That is to say, the agent's type is $\theta = (w, \psi)$ and the type space is $\Theta = \mathbb{R}_{++} \times \Psi$. For convenience, I will refer to a given value of $\psi$ as a *group*.

Let $F(w, \psi)$ represent the joint distribution of wages and group membership. For simplicity, I maintain the assumption that there are no income effects (assumption 3) but the results below can be easily extended to incorporate income effects. Thus, I can define $z(t; w, \psi)$ as the taxable income a type $(w, \psi)$ agent chooses when facing a linear income tax rate $t < 1$ (irrespective of the intercept of their budget constraint). Let $H(z; t)$ be the distribution of taxable income under a tax rate of $t$, as induced by the function $z(t; w, \psi)$ and the parameter distribution $F$.

In addition to assumptions 1, 2, and 3, in this section I assume that within each group $\psi$ agent preferences satisfy the single-crossing condition with respect to the wage rate.

**Assumption 4** (Single Crossing Condition)**.** *For all types $(w, \psi) \in \mathbb{R}_{++} \times \Psi$ and any tax rate $t < 1$, taxable income is strictly increasing in the wage rate $\frac{\partial z(t;w,\psi)}{\partial w} > 0$.*

With this additional assumption, it is possible to write the bunching mass as a function of compensated elasticities of taxable income. Let

$$\varepsilon^c(t; w, \psi) \equiv -\frac{1-t}{z^\star}\frac{\partial z(t; w, \psi)}{\partial t} \tag{24}$$

denote the compensated elasticity of taxable income for type $(w, \psi)$ agents facing a linear tax rate of $t$.

**Theorem 5** (Bunching Mass in Terms of Elasticities)**.** *Under assumptions 1, 2, 3, and 4, and given a piecewise linear tax schedule with two tax brackets (as defined in equation 2), the bunching mass at the kink point can be written as:*

$$H\left(z^{\star};t_1\right)-H\left(z^{\star};t_0\right)=\int_{t_0}^{t_1}\bar{\varepsilon}^c\left(z^{\star};t\right)\cdot\frac{z^{\star}h\left(z^{\star};t\right)}{1-t}dt, \tag{25}$$

*where $\bar{\varepsilon}^c\left(z^{\star};t\right)$ is the average (compensated) elasticity of taxable income among agents who would choose to locate at $z^{\star}$ when facing a counterfactual linear tax schedule with a tax rate of $t$,*

$$\bar{\varepsilon}^c\left(z^{\star};t\right)\equiv\mathbb{E}\left[\varepsilon^c\left(t;w,\psi\right)\left|z\left(t;w,\psi\right)=z^{\star}\right.\right]. \tag{26}$$

*Proof.* The bunching mass at $z^{\star}$ can be written as

$$\begin{aligned}H\left(z^{\star};t_1\right)-H\left(z^{\star};t_0\right)&=\int_{\psi}\left[H\left(z^{\star}|\psi;t_1\right)-\tilde{H}\left(z^{\star}|\psi;t_0\right)\right]f_{\psi}\left(\psi\right)\mathrm{d}\psi\\&=\int_{\psi}\int_{t_0}^{t_1}\frac{\partial H\left(z^{\star}|\psi;t\right)}{\partial t}f_{\psi}\left(\psi\right)\mathrm{d}\psi\end{aligned}$$

where the second equality follows from the fundamental theorem of calculus.

Note, the single crossing condition (assumption 4) ensures that for any linear tax rate $t$ and within any group $\psi$ there exists a wage rate $w^{\star}\left(t,\psi\right)$ at which group members would choose to locate at the kink point. That is to say, there exists a function $w^{\star}\left(t,\psi\right)$ satisfying $z\left(t;w^{\star}\left(t,\psi\right),\psi\right)=z^{\star}$ for all $t<1$ and $\psi\in\Psi$. Furthermore, the single crossing condition ensures that there is a straightforward connection between the conditional distribution of taxable income and the conditional distribution of wages:

$$H\left(z\left(t;w,\psi\right)|\psi;t\right)=F\left(w|\psi\right).$$

Differentiating both sides of this expression with respect to $t$ we obtain

$$\frac{\partial H\left(z\left(t;w,\psi\right)|\psi;t\right)}{\partial t}=-\frac{\partial z\left(t;w,\psi\right)}{\partial t}h\left(z\left(t;w,\psi\right)|\psi;t\right)$$

and therefore the bunching mass can be expressed as

$$
\begin{aligned}
H\left(z^{\star};t_{1}\right)-H\left(z^{\star};t_{0}\right) &= -\int_{\psi}\int_{t_{0}}^{t_{1}}\frac{\partial z\left(t;w^{\star}\left(t,\psi\right),\psi\right)}{\partial t}h\left(z^{\star}|\psi;t\right)f_{\psi}\left(\psi\right)\mathrm{d}t\mathrm{d}\psi \\
&= -\int_{t_{0}}^{t_{1}}\left\{\int_{\psi}\frac{\partial z\left(t;w^{\star}\left(t,\psi\right),\psi\right)}{\partial t}\frac{h\left(z^{\star}|\psi;t\right)f_{\psi}\left(\psi\right)}{h\left(z^{\star};t\right)}\mathrm{d}\psi\right\}h\left(z^{\star};t\right)\mathrm{d}t
\end{aligned}
$$

The result then follows from using equation (24) to simplify the interior integral above. □

Theorem 5 is simply a restatement of a result derived by Blomquist et al. [2019] and represents the most general formulation of the relationship between the bunching mass and taxpayer elasticities. Importantly, equation 25 shows that the bunching mass is a function of both agents' elasticities and unobserved densities of taxable income. Thus, the bunching mass alone is not informative about taxpayer elasticities. In order to obtain a weighted average of elasticities from an estimate of the bunching mass, it is necessary to additionally an have an estimate of the *normalizing constant*

$$
\mathcal{H}\equiv z^{\star}\int_{t_{0}}^{t_{1}}\frac{h\left(z^{\star};t\right)}{1-t}\mathrm{d}t. \tag{27}
$$

By dividing the bunching mass by $\mathcal{H}$ it is possible to identify the *bunching elasticity*

$$
\bar{\varepsilon}_{B}^{c}\equiv\int_{t_{0}}^{t_{1}}\bar{\varepsilon}^{c}\left(z^{\star};t\right)\rho\left(t\right)\mathrm{d}t. \tag{28}
$$

which is a weighted average elasticity which places some positive weight

$$
\rho\left(t\right)\equiv\frac{\frac{h\left(z^{\star};t\right)}{1-t}}{\int_{t_{0}}^{t_{1}}\frac{h\left(z^{\star};t'\right)}{1-t'}\mathrm{d}t'}, \tag{29}
$$

on the local average elasticity at the kink point under a linear tax rate $t$ ($\bar{\varepsilon}^{c}\left(z^{\star};t\right)$) for each tax rate $t\in\left[t_{0},t_{1}\right]$.

**Corollary 3.** *Under assumptions 1, 2, 3 and 4, and given a piecewise linear tax schedule with two tax brackets (as defined in equation 2), the bunching mass at the kink point together with the normalizing constant* $\mathcal{H}$ *identifies the bunching elasticity:*

$$
\bar{\varepsilon}_{B}^{c}=\frac{H\left(z^{\star};t_{1}\right)-H\left(z^{\star};t_{0}\right)}{\mathcal{H}} \tag{30}
$$

The central focus of Blomquist et al. [2019] and Bertanha et al. [2021] is the challenge to identification of the bunching elasticity caused by the fact that $\mathcal{H}$ is not identified by the observed distribution of taxable income. Saez [2010] and Chetty et al. [2011] propose methods for addressing this issue, but Blomquist et al. [2019] and Bertanha et al. [2021] highlight potential problems with these approaches. Although these papers propose alternative approaches to obtaining identification or partial identification of a weighted average of elasticities, these proposals require either imposing strong restrictions on how the density of taxable income at $z^\star$ $(h(z^\star; t))$ varies across different linear tax schedules, or assuming that agents have quasi-linear, isoelastic preferences with no variation in their elasticity.

Blomquist et al. [2019] present a version of Corollary 3, and note that the bunching elasticity $(\bar{\varepsilon}^c_B)$ is subject to "issues of external validity" because it provides information about a very specific convex combination of elasticities.[26] In the next subsection, I build on their work by more precisely characterizing these issues.

## 4.2 Policy Analysis with the Bunching Elasticity

Corollary 1 already provides one unambiguously valid interpretation of the bunching elasticity. Using equation (30) the fiscal externality of moving a kink point can be rewritten in terms of the bunching elasticity:[27]

$$FE(z^\star) \equiv -\frac{t_0}{t_1 - t_0} \cdot \frac{\mathcal{H}}{1 - H_1(z^\star)} \cdot \bar{\varepsilon}^c_B$$

However, Corollary 1 also makes it clear that decomposing the bunching mass into the bunching elasticity $(\bar{\varepsilon}^c_B)$ and the normalizing constant $(\mathcal{H})$ is unnecessary for the purpose of identifying the fiscal externality of moving a kink point. Thus, the value of such a decomposition depends on whether the bunching elasticity is for analyzing other tax reforms of interest.

A natural candidate for a tax reform of interest is the effect of an infinitesimal change in the marginal income tax rate either immediately below or immediately above the kink. In particular, consider a tax schedule perturbation in the style of Saez [2001] which infinitesimally increases the marginal tax rate just below $z^\star$ and leaves it unchanged everywhere else. Applying results from the literature on the perturbation approach to optimal nonlinear taxation, the fiscal externality of such a reform is[28]

$$FE^-_t(z^\star) \equiv -\frac{t_0}{1 - t_0} \cdot \frac{z^\star h(z^\star; t_0)}{1 - H(z^\star; t_0)} \cdot \bar{\varepsilon}^c(z^\star; t_0). \tag{31}$$

[26] It is important to note that this weighted average is not, in general, interpretable as the average elasticity of bunchers. In appendix ??, I show that in the special case where taxpayers have quasi-linear, isoelastic preferences but exhibit heterogeneous elasticities the bunching elasticity is in fact strictly smaller than the average elasticity of bunching taxpayers.

[27] Or, equivalently, the fiscal externality of increasing the tax rate just below the kink point from $t_0$ to $t_1$.

[28] These results are usually framed in terms of welfare effects rather than fiscal externalities. The perturbation approach to optimal taxation began with Saez [2001] and has subsequently been formalized and extended to the case of multidimensional heterogeneity by Gerritsen [2016], Golosov et al. [2014], and Jacquet and Lehmann [2021]. Bergstrom and Dodds [2021] extend these results to reforms of tax schedules which are almost everywhere continuously differentiable (i.e. tax schedules that may contain kink points) so the expressions below follow from their results.

Similarly, the fiscal externality of a tax reform which infinitesimally increases the marginal tax rate just above $z^{\star}$ is

$$FE_t^{+}(z^{\star}) \equiv -\frac{t_1}{1-t_1} \cdot \frac{z^{\star} h(z^{\star}; t_1)}{1-H(z^{\star}; t_1)} \cdot \bar{\varepsilon}^c(z^{\star}; t_1). \tag{32}$$

The fiscal externalities of these two local tax reforms can be estimated using the average elasticity of two different types of "marginal bunchers". In particular, $FE_t^{-}(z^{\star})$ depends on the average elasticity of agents who would locate at the kink $z^{\star}$ under the linear tax rate $t_0$ ($\bar{\varepsilon}^c(z^{\star}; t_0)$), and $FE_t^{+}(z^{\star})$ depends on the average elasticity of agents who would locate at the kink $z^{\star}$ under the linear tax rate $t_1$ ($\bar{\varepsilon}^c(z^{\star}; t_1)$).

Having precisely defined these two local tax reforms of interest, and clarified the elasticities needed to identify their revenue consequences helps us shed light on two key issues: the external validity of the bunching elasticity, and the differences between the revenue effects of moving a kink point and of infinitesimally changing marginal rates near the kink point.

#### 4.2.1 Determinants of External Validity

The definition of the bunching elasticity (equation (28)) makes it clear that neither of the two policy-relevant local average elasticities contained in equations (31) and (32) will generally be identified by the standard bunching method. Although the bunching elasticity does assign positive weight to both $\bar{\varepsilon}^c(z^{\star}; t_0)$ and $\bar{\varepsilon}^c(z^{\star}; t_1)$, in general, a given value of $\bar{\varepsilon}^c_B$ is consistent with any finite, positive values of $\bar{\varepsilon}^c(z^{\star}; t_0)$ and $\bar{\varepsilon}^c(z^{\star}; t_1)$. It is thus important to clarify under what circumstances we should expect $\bar{\varepsilon}^c_B$ to differ substantially from $\bar{\varepsilon}^c(z^{\star}; t_0)$ and $\bar{\varepsilon}^c(z^{\star}; t_1)$.

Equation (28) also makes it clear what the source of any such differences would be: endogeneity of the local average ETI at the kink point ($\bar{\varepsilon}^c(z^{\star}; t)$) to the tax rate. It can be shown that such endogeneity results from two sources:[29]

$$\begin{aligned}\frac{\mathrm{d}\bar{\varepsilon}^c(z^{\star}; t)}{\mathrm{d}t} = & \underbrace{-\frac{1}{1-t}\mathrm{Cov}\left(\varepsilon^c, (1+\alpha)\,\varepsilon^c | z = z^{\star}\right)}_{\text{composition effect}} \\ & \underbrace{+\mathbb{E}\left[\frac{\partial \varepsilon^c}{\partial t} | z = z^{\star}\right] + \mathbb{E}\left[\left(2 - \frac{\bar{\varepsilon}^c(z^{\star}; t)}{\varepsilon^c}\right)\frac{\partial \varepsilon^c}{\partial w}\frac{\partial w^{\star}(t;\psi)}{\partial t} | z = z^{\star}\right]}_{\text{misspecification effect}}, \end{aligned} \tag{33}$$

where $\alpha(z; t, \psi) \equiv -\left(1 + \frac{z h'(z|\psi; t)}{h(z|\psi; t)}\right)$ is the (second-order) Pareto parameter of the group-$\psi$-specific distribu-

[29] See appendix A.4 for the derivation.

tion of taxable income at a given level of income $z$ under tax rate $t$.[30] Note, for simplicity, I have suppressed the arguments of various functions in equation (33) and will continue to do so below.[31]

To understand the first term in equation (33), note that when agents exhibit multidimensional heterogeneity (i.e. when there is more than one group $\psi$), changing the tax rate influences the relative size of different groups among agents located at a given value of taxable income (in this case, $z^\star$). This generates what Jacquet and Lehmann [2021] label a *composition effect*.[32] We can break composition effects down further:

$$\mathrm{Cov}\left(\varepsilon^c, (1+\alpha)\,\varepsilon^c | z = z^\star\right) = \underbrace{\left(1 + \mathbb{E}\left[\alpha | z = z^\star\right]\right) \mathrm{Var}\left(\varepsilon^c | z = z^\star\right)}_{\text{elasticity heterogeneity across groups}} + \underbrace{\mathrm{Cov}\left(\varepsilon^c\left(\varepsilon^c - \bar{\varepsilon}^c\left(z^\star; t\right)\right), \alpha | z = z^\star\right)}_{\text{heterogeneity of Pareto parameter across groups}} \tag{34}$$

The first term in this expression captures the fact that, all else equal, greater variance of elasticities among agents at $z^\star$ implies a larger magnitude composition effect. This variance effect is proportional to the average group-specific Pareto parameter at $z^\star$. The second term is non-zero whenever group-specific Pareto parameters are correlated with a particular function of group-specific elasticities. Note that both of these types of composition effects are present if agents have isoelastic preferences with constant elasticities within groups; it is variation in elasticities across groups that generate composition effects.

The second term in equation (33) reflects another setting where the tax rate can influence the local average ETI $\bar{\varepsilon}^c\left(z^\star; t\right)$: if agent preferences are not isoelastic. I call this the *misspecification effect*, in reference to the ubiquity of isoelastic function form assumptions in the bunching literature (and in empirical public finance generally). As with the composition effect, the misspecification effect can be broken down into two parts:

$$\underbrace{\mathbb{E}\left[\frac{\partial \varepsilon^c}{\partial t} | z = z^\star\right]}_{\text{direct effect of tax rate}} + \underbrace{\mathbb{E}\left[\left(2 - \frac{\bar{\varepsilon}^c\left(z^\star; t\right)}{\varepsilon^c}\right) \frac{\partial \varepsilon^c}{\partial w} \frac{\partial w^\star\left(t; \psi\right)}{\partial t} | z = z^\star\right]}_{\text{elasticity heterogeneity within groups}}. \tag{35}$$

The first term in this expression reflects the fact that with non-isoelastic preferences, the tax rate can directly influence the elasticity of taxable income. The interpretation of the second term is somewhat complex, but

[30]Note, for a continuous random variable $X$ with distribution $F_X$ the function

$$\frac{x f_X\left(x\right)}{1 - F_X\left(x\right)}$$

is sometimes described as the Pareto parameter at $x$ in the public finance literature, because if $F_X$ has a Pareto tail, this measure converges to the Pareto parameter as $x \to \infty$. This claim is also true for the function

$$-\left(1 + \frac{x f'_X\left(x\right)}{f_X\left(x\right)}\right),$$

and this measure has also sometimes been described in the public finance literature as the "Pareto parameter" (for example, see Hendren [2020]). Because this paper includes references to both these measures, I distinguish the second measure by labeling it as the *second-order Pareto parameter at x*.

[31]In particular, I have employed the convention that $\varepsilon^c \equiv \varepsilon^c\left(t; w, \psi\right)$ and $\alpha \equiv \alpha\left(z; t, \psi\right)$, and $z \equiv z\left(t; w, \psi\right)$.

[32]They discuss the implications of such effects for attempts to use sufficient statistics pertaining to an observed tax system for the purpose of inferring the optimal tax system.

this term will be present whenever agents within the same group can have different elasticities (i.e. the group-specific elasticity varies with the wage rate). Intuitively, it captures the fact that changes in the tax rate will generally change which members of a given group choose to locate at a particular level of taxable income, and so if elasticities vary within groups such changes can generate an indirect effect on the local average ETI. Note, both of these effects would be present even in a model with unidimensional heterogeneity (i.e. if there were only one group $\psi$).

To further clarify the nature of composition and misspecification effects, let us consider some motivating examples.

**Example 1** (Isoelastic Preferences with Constant Pareto Parameter)**.** Suppose that all taxpayers have quasi-linear, isoelastic preferences of the form

$$u\left(z;w,\psi\right) \equiv z - T\left(z\right) - \frac{w}{1+\frac{1}{\psi}}\left(\frac{z}{w}\right)^{1+\frac{1}{\psi}},$$

so that

$$z\left(t;w,\psi\right) = w\left(1-t\right)^{\psi}.$$

With this specification, agents within a given group all have the same elasticity: $\psi$. Thus, there are no misspecification effects. Further suppose that all groups have wages are distributed according to $F$, a distribution which is Pareto Type I above some threshold wage $w_0$. That is to say, for any $w \geq w_0$

$$F_w\left(w\right) = F_w\left(w_0\right) + \left(1 - F_w\left(w_0\right)\right)\left(1 - \left(\frac{w}{w_0}\right)^{-a}\right)$$

for some Pareto slope parameter $a > 1$. Given these assumptions, it can be shown that for any taxable income $z > w_0$, all groups $\psi$ will have the same group-specific Pareto parameter: $\alpha\left(z;t,\psi\right) = a$ for all $\psi$ and $t < 1$. This implies that the second term in equation (34) is zero. The effect of the tax rate on the local average ETI can thus simply be written as

$$\frac{\mathrm{d}\bar{\varepsilon}^{c}\left(z^{\star};t\right)}{\mathrm{d}t} = -\frac{1}{1-t}\cdot\left(1+a\right)\cdot \mathrm{Var}\left(\psi | z = z^{\star}\right).$$

**Example 2** (Isoelastic Preferences with Non-Constant Pareto Parameter)**.** Suppose that all taxpayers have quasi-linear, isoelastic preferences but of a slightly different form than that discussed above:

$$u\left(z;w,\psi\right) \equiv z - T\left(z\right) - \frac{1}{1+\frac{1}{\psi}}\left(\frac{z}{w}\right)^{1+\frac{1}{\psi}}.$$

This specification implies a different taxable income function

$$z\left(t;w,\psi\right)=w^{1+\psi}\left(1-t\right)^{\psi}.$$

As in the previous example, within a given group all agents have the same ETI ($\psi$) and so there are no misspecification effects. Again, suppose that wages are identically distributed across groups according to $F$, a distribution which is Pareto Type I with slope $a$ at any wage above $w_0$. It can be shown that for any taxable income $z>w_0$, different groups have different local Pareto parameters. In particular, $\alpha\left(z;t,\psi\right)=\frac{a}{1+\psi}$ for all $\psi$ and $t<1$. This implies that the second term in equation (34) is non-zero. The effect of the tax rate on the local average ETI can thus simply be written as

$$\frac{\mathrm{d}\bar{\varepsilon}^{c}\left(z^{\star};t\right)}{\mathrm{d}t}=-\frac{1}{1-t}\left[\mathrm{Var}\left(\psi|z=z^{\star}\right)+a\mathrm{Cov}\left(\psi,\frac{\psi}{1+\psi}|z=z^{\star}\right)\right].$$

**Example 3** (Unidimensional Heterogeneity with Multiple Incomes)**.** Suppose that all taxpayer have two source of income ($z_1$ and $z_2$) and that their total taxable income is simply the sum of these two incomes. For simplicity, I assume that both labor activities have the same wage rate. Additionally, suppose that there is only one group $\psi$, so that agents differ only in their wage rate $w$. Under this assumption, there are no composition effects but there is a misspecification effect. Finally suppose that agents have a separable utility function, with the disutility of income for each source taking a quasi-linear, isoelastic form.

In particular, suppose that an agent with wage rate $w$ chooses values of $z_1$ and $z_2$ to solve the problem

$$\max_{z_1,z_2>0}\left\{z_1+z_2-T\left(z_1+z_2\right)-\frac{w}{1+\frac{1}{e_1}}\left(\frac{z_1}{w}\right)^{1+\frac{1}{e_1}}-\frac{w}{1+\frac{1}{e_2}}\left(\frac{z_2}{w}\right)^{1+\frac{1}{e_2}}\right\}.$$

Such an agent facing a linear tax rate $t$ will choose to generate $z_1\left(t;w\right)\equiv w\left(1-t\right)^{e_1}$ dollars of income from the first source and $z_2\left(t;w\right)\equiv w\left(1-t\right)^{e_2}$ dollars from the second source. Total taxable income for such an agent is

$$\begin{aligned}z\left(t;w\right)&\equiv z_1\left(t;w\right)+z_2\left(t;w\right)\\&=w\left[\left(1-t\right)^{e_1}+\left(1-t\right)^{e_2}\right].\end{aligned}$$

Notice, the elasticities of $z_1$ and $z_2$ with respect to the net-of-tax rate are $e_1$ and $e_2$. These elasticities are constant across all agents and different tax rate. However the elasticity of taxable income is a convex combination of the elasticities of each of the two income sources

$$\begin{aligned}\varepsilon^c(t) &\equiv \frac{z_1(t;w)}{z(t;w)} \cdot e_1 + \frac{z_2(t;w)}{z(t;w)} \cdot e_2 \\ &= \left(\frac{(1-t)^{e_1}}{(1-t)^{e_1} + (1-t)^{e_2}}\right) e_1 + \left(\frac{(1-t)^{e_2}}{(1-t)^{e_1} + (1-t)^{e_2}}\right) e_2,\end{aligned}$$

where the weights on the elasticity of each income source correspond to the share of the income source in the agent's total income. As shown above, these income shares are independent of taxpayer wage rate so the misspecification effect in this example is entirely a direct effect of taxation on this elasticity. This direct effect stems from the fact that, as long as $e_1 \neq e_2$, a change in the tax rate changes the share of each type of income in total taxable income. Specifically, the misspecification effect is

$$\frac{\mathrm{d}\varepsilon^c(t)}{\mathrm{d}t} = -\frac{(e_1 - e_2)^2}{1-t} \cdot \frac{z_1(t;w)}{z(t;w)} \cdot \frac{z_2(t;w)}{z(t;w)}.$$

#### 4.2.2 Infinitesimal Rate Changes vs. Moving Kinks

Superficially, the external validity concerns associated with use of the bunching elasticity might seem unimportant in light of the earlier results of this paper. As was noted above, moving a kink point provides a mechanism for implementing a specific local rate change. It is therefore important to ask how the impact of moving a kink point differs from our other local tax reforms of interest. By comparing (31) and (32) with the fiscal externality of moving a kink point, we can gain insight into this question.

Inserting equation (25) into equation (14) we obtain an expression for the fiscal externality of moving a kink point as a function of the average elasticity of taxable income at $z^\star$ under various counterfactual linear tax rates between $t_0$ and $t_1$:

$$FE(z^\star) = -\frac{t_0 \int_{t_0}^{t_1} \frac{z^\star h(z^\star;t)}{1-t} \bar{\varepsilon}^c(z^\star;t)\,\mathrm{d}t}{(t_1 - t_0)(1 - H(z^\star;t_1))} \tag{36}$$

This expression bears some superficial similarities in structure to the fiscal externalities of the local tax reforms introduced above (equations 31 and 32). To make the connection more precise, we can rewrite $FE(z^\star)$ in terms of the fiscal externalities of these other local reforms. In each case, the fiscal externality of moving a kink point can be written as the sum of a re-scaled version of the fiscal externality of the local tax reform and an adjustment term:

$$FE(z^\star) = \left(\frac{1 - H(z^\star;t_0)}{1 - H(z^\star;t_1)}\right) FE_t^-(z^\star) - \frac{t_0 z^\star \int_{t_0}^{t_1} \int_{t_0}^{t} \left\{\frac{\mathrm{d}}{\mathrm{d}t'}\left[\frac{h(z^\star;t')}{1-t'} \bar{\varepsilon}^c(z^\star;t')\right]\right\} \mathrm{d}t'\mathrm{d}t}{(t_1 - t_0)(1 - H(z^\star;t_1))}, \tag{37}$$

and

$$FE\left(z^{\star}\right)=\left(\frac{t_0}{t_1}\right)FE_t^{+}\left(z^{\star}\right)+\frac{t_0 z^{\star}\int_{t_0}^{t_1}\int_{t}^{t_1}\left\{\frac{\mathrm{d}}{\mathrm{d}t'}\left[\frac{h\left(z^{\star};t'\right)}{1-t'}\bar{\varepsilon}^c\left(z^{\star};t'\right)\right]\right\}\mathrm{d}t'\mathrm{d}t}{\left(t_1-t_0\right)\left(1-H\left(z^{\star};t_1\right)\right)}. \tag{38}$$

Note, the adjustment terms depend on how $\frac{h(z^{\star};t)}{1-t}\bar{\varepsilon}^c\left(z^{\star};t\right)$ varies over the interval $[t_0,t_1]$. Evaluating the limit of equations (37) and (38) as $t_1\to t_0$, we can see that the fiscal externalities of all three reforms converge for an arbitrarily small kink size.[33] However, it is clear that, in general, these tax reforms may have quite different revenue effects from one another.

To better understand the adjustment terms in equations (37) and (38), note that the integrand in these terms can be decomposed into three effects:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{h\left(z^{\star};t\right)}{1-t}\bar{\varepsilon}^c\left(z^{\star};t\right)\right]=&\underbrace{\frac{h\left(z^{\star};t\right)}{\left(1-t\right)^2}\left(\bar{\varepsilon}^c\left(z^{\star};t\right)-\left(1+\mathbb{E}\left[\alpha|z=z^{\star}\right]\right)\bar{\varepsilon}^c\left(z^{\star};t\right)^2\right)}_{\text{baseline effect}}\\&\underbrace{-\frac{h\left(z^{\star};t\right)}{\left(1-t\right)^2}\left(\left(1+\mathbb{E}\left[\alpha|z=z^{\star}\right]\right)\mathrm{Var}\left(\varepsilon^c|z=z^{\star}\right)+\mathrm{Cov}\left(\left(\varepsilon^c\right)^2,\alpha|z=z^{\star}\right)\right)}_{\text{composition effect}}\\&+\underbrace{\frac{h\left(z^{\star};t\right)}{1-t}\left(\mathbb{E}\left[\frac{\partial\varepsilon^c}{\partial t}|z=z^{\star}\right]+2\mathbb{E}\left[\frac{\partial\varepsilon^c}{\partial w}\frac{\partial w^{\star}\left(t;\psi\right)}{\partial t}|z=z^{\star}\right]\right)}_{\text{misspecification effect}}. \end{aligned}\tag{39}$$

As with equation (33) the *composition effect* in equation (39) is non-zero if there is heterogeneity in agent elasticities across groups and the *misspecification effect* is non-zero if agent preferences are non-isoelastic, though the precise nature of these effects differs slightly from their counterparts equation (33). In addition, equation (39) includes a *baseline effect*, which reflects that fact that even if taxpayers have isoelastic preferences and homogeneous elasticities, $\frac{h(z^{\star};t)}{1-t}$ remains endogenous to the tax rate.

If $\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{h(z^{\star};t)}{1-t}\bar{\varepsilon}^c\left(z^{\star};t\right)\right]>0$ for all $t\in[t_0,t_1]$, then equations (37) and (38) imply that the fiscal externality of moving a kink point always falls between the fiscal externalities of the two local rate changes:

$$0>FE_t^{-}\left(z^{\star}\right)>FE\left(z^{\star}\right)>FE_t^{+}\left(z^{\star}\right).$$

However, the sign of $\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{h(z^{\star};t)}{1-t}\bar{\varepsilon}^c\left(z^{\star};t\right)\right]$ is ambiguous, so in principle possible that the fiscal externality of

[33] This provides a more rigorous derivation of the result I presented in equation (**??**).

moving a kink point $FE(z^\star)$ could be above or below both $FE_t^-(z^\star)$ and $FE_t^+(z^\star)$.[34,35] This implies that there can exist Pareto- or welfare-improving movements of a kink point even if such improvements cannot be obtained via infinitesimal local rate changes near the kink point. On the other hand, the converse also holds true: in some cases, an infinitesimal local rate change might be efficiency- or welfare-enhancing even when moving a kink point is not. Indeed, this latter points holds irrespective of the sign of $\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{h(z^\star;t)}{1-t}\bar{\varepsilon}^c(z^\star;t)\right]$.

### 4.3 Discussion

In comparing my proposed bunching methodology to the standard approach, two salient considerations emerge. First, my method remains valid in the presence of multidimensional heterogeneity and departures from isoelastic functional form assumptions, whereas the standard bunching method—even in an ideal application where $\mathcal{H}$ is known—is not. Standard bunching estimates of the ETI can thus generate misleading policy implications in circumstances where my method will generate accurate policy implications. Second, policy analysis which focuses only on moving kink points to the exclusion of other local tax reforms can result in missing out on beneficial reform opportunities.

One perspective on the ambiguity in current knowledge about the factors that determine how close a bunching-based elasticity ($\bar{\varepsilon}_B^c$) will be to the policy-relevant local average elasticities $\bar{\varepsilon}^c(z^\star;t_0)$ and $\bar{\varepsilon}^c(z^\star;t_1)$ is that a strength of the new bunching method proposed in this paper is its robustness to any realization of these unknown factors. On the other hand, this ambiguity also implies that it is impossible to know to what extent other beneficial local tax reforms might be overlooked by a policy analysis that focuses exclusively on moving kink points.

## 5 Discrete Kink Point Movements

The headline result of this paper pertains to infinitesimal movements of kink points. Since any real-world tax reform is necessarily discrete in nature, it may be insightful to consider what can (and cannot) be learned about the revenue effect of a discrete movement in a kink point using the bunching mass. In particular, consider decrease in the location of the kink point from some initial $z^\star$ to some new point $z^{\star\star} < z^\star$.

[34]To see that the sign of $\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{h(z^\star;t)}{1-t}\bar{\varepsilon}^c(z^\star;t)\right]$ is ambiguous note because all three effects in equation (39) can be positive or negative. The baseline effect is positive if and only if

$$\bar{\varepsilon}^c(z^\star;t) < \frac{1}{1+\mathbb{E}[\alpha|z=z^\star]}.$$

The composition effect has an ambiguous sign because $\mathbb{E}[\alpha|z=z^\star]$ can be above or below $-1$ and $\mathrm{Cov}\left((\varepsilon^c)^2, \alpha|z=z^\star\right)$ can be positive or negative. Finally, both components of the misspecification effect have ambiguous signs.

[35]In section **??** I present simulation results that include an example where this occurs.

This reform has mechanical effects, which stem from the fact that any taxpayer initially located above the kink point ($z_1(\theta) > z^\star$) faces an increase in their tax liability of $(t_1 - t_0)(z^\star - z^{\star\star})$. Assuming there are no income effects, these agents will not change their choice of taxable income in response to this reform, and therefore the mechanical effect of the reform on tax revenue is

$$(t_1 - t_0)(1 - H_1(z^\star))(z^\star - z^{\star\star}). \tag{40}$$

Turning to the behavioral response, as in the infinitesimal case we can split it into three components.

First, we have the relocation effect. Holding constant the size of the bunching mass, bunchers used to pay $t_0 z^\star$ in taxes, but now instead pay $t_0 z^{\star\star}$, changing revenue by

$$-t_0(z^\star - z^{\star\star})(H_1(z^\star) - H_0(z^\star)), \tag{41}$$

Adding (41) and (40) together we obtain the first-order approximation of the revenue effect of a discrete change in the kink point (equation 9).

This approximation ignores the two additional components of the behavioral response: the new buncher effect, and the former buncher effect. some taxpayers who previously located below the kink will become bunchers. Specifically, those taxpayers whose choices under $T_0$ and $T_1$ satisfy,

$$z_1(\theta) < z^{\star\star} < z_0(\theta) < z^\star.$$

The first-order approximation implicitly assumes these taxpayers continue to pay $t_0 z_0(\theta)$ in taxes, but in reality, they now pay the lower amount $t_0 z^{\star\star}$. Thus, correcting the first-order approximation to account for the new buncher effect reduces tax revenue by

$$-t_0 \int_{z^{\star\star}}^{z^\star} (z - z^{\star\star}) h_0(z)\, \mathrm{d}z. \tag{42}$$

On the other hand, some taxpayers who previously chose to bunch will stop bunching and locate somewhere above the new kink point. These taxpayers are those whose choices under $T_0$ and $T_1$ satisfy

$$z^{\star\star} < z_1(\theta) < z^\star < z_0(\theta).$$

Approximating the behavioral response effect using the relocation effect implicitly assumes that these taxpayers remain as bunchers but move to the new kink point, paying $t_0 z^{\star\star}$ in taxes. But in fact, they pay the larger amount of $t_1 (z_1(\theta) - z^{\star\star}) + t_0 z^{\star\star}$. Thus, correcting the first-order approximation for the former buncher effect increases tax revenue by

$$t_1 \int_{z^{\star\star}}^{z^{\star}} (z - z^{\star\star})\, h_1(z)\, \mathrm{d}z. \tag{43}$$

Adding (40), (41), (42), and (43) together we obtain the total effect of a discrete reform on tax revenue:[36]

$$\begin{aligned} R(z^{\star\star}) - R(z^{\star}) = & \underbrace{(t_1 - t_0)(z^{\star} - z^{\star\star})(1 - H_1(z^{\star}))}_{\text{mechanical effect}} \quad \underbrace{-t_0 (z^{\star} - z^{\star\star})(H_1(z^{\star}) - H_0(z^{\star}))}_{\text{relocation effect}} \\ & - \underbrace{t_0 \int_{z^{\star\star}}^{z^{\star}} (z - z^{\star\star})\, h_0(z)\, \mathrm{d}z}_{\text{new buncher correction}} + \underbrace{t_1 \int_{z^{\star\star}}^{z^{\star}} (z - z^{\star\star})\, h_1(z)\, \mathrm{d}z}_{\text{former buncher correction}} \end{aligned} \tag{44}$$

Equation (44) provides some intuition for my main result. The bunching mass is a sufficient statistic for the behavioral response to a small change in the location of the kink point because, for sufficiently small reforms, the net impact of the new and former buncher effects is negligible.

Equation (44) presents the revenue effect of a discrete decrease in the location of a kink point. The same equation can be used to describe the effect of increasing the location of a kink point from $z^{\star}$ to $z^{\star\star} > z^{\star}$, but it is helpful to provide an explicit reinterpretation. The revenue effect of this change can be written as

$$\begin{aligned} R(z^{\star\star}) - R(z^{\star}) = & \underbrace{-(t_1 - t_0)(z^{\star\star} - z^{\star})(1 - H_1(z^{\star}))}_{\text{first-order approximation of mechanical effect}} + \underbrace{t_0 (z^{\star\star} - z^{\star})(H_1(z^{\star}) - H_0(z^{\star}))}_{\text{first-order approximation of behavioral response}} \\ & \cdots - \underbrace{t_0 \int_{z^{\star}}^{z^{\star\star}} (z^{\star\star} - z)\, h_0(z)\, \mathrm{d}z}_{\text{former buncher correction}} + \underbrace{t_1 \int_{z^{\star}}^{z^{\star\star}} (z^{\star\star} - z)\, h_1(z)\, \mathrm{d}z}_{\text{new buncher correction}} \end{aligned} \tag{45}$$

Recall, the first two terms in this expression provide a first-order approximation to the effect of this reform. This approximation implicitly assumes that the revenue effect of the reform consists of a reduction in the tax liability of all taxpayers initially located above the original kink point $z^{\star}$, and an increase in revenue resulting from moving a fixed bunching mass to a higher locations. The interpretation of errors this approximation introduces differs slightly from the case of a kink point decrease however, where the approximation only introduced error into the behavioral response to the reform. Here, it includes errors in both the behavioral response and the mechanical effect.

To see this, note that the first-order approximation implicit assumes that taxpayers initially located in the

[36] The observant reader will notice that the derivation of equation (44) presented above abstracts from an additional form of behavioral response to discretely decreasing a kink point: that some taxpayers will move from an interior solution in the lower tax bracket to an interior solution in the higher tax bracket. However, this is done only for expositional simplicity; equation (44) can also be derived using the revenue function (8). The heuristic derivation works because the new buncher effect accounts for the removal of these individuals from the lower bracket and the former buncher effect accounts for their appearance in the higher bracket.

upper tax bracket between the original kink point and the new kink point will continue to choose $z_1(\theta)$ and will pay taxes of

$$t_1(z_1(\theta) - z^\star) + t_0 z^\star - \underbrace{(t_1 - t_0)(z^{\star\star} - z^\star)}_{\text{mechanical effect}}.$$

Instead, these individuals choose to become bunchers at the new kink point $z^{\star\star}$, paying taxes of $t_0 z^{\star\star}$. The correction for this error take a simple form: the new buncher effect in equation (45).

There is also once again a former buncher effect, which corrects the error arising in the first-order approximation due to the fact that some taxpayers exit the bunching mass after this reform. Following an increase in the kink point, these taxpayers now locate somewhere on the lower tax bracket, paying taxes of $t_0 z_0(\theta)$ rather than the $t_0 z^{\star\star}$ implicitly assumed when using the first-approximation to the behavioral response to a reform.

**Partial Identification and Approximations**

Together, equations (44) and (45) provide a complete guide to the revenue effects of moving the location of a kink point.

If the pre-reform distribution of taxable income is observed, it is easy to see that three out of the four terms contained in equations (44) and (45) are identified. In both cases, the new buncher effect is identified by the observed income distribution (in addition to two first-order approximation terms). This is because the new buncher effect corrects for the first-order approximation's implicit assumption that they will continue to make the same choices that we currently observed them making, when in fact they will end up at a known counterfactual bunching location $z^{\star\star}$.

By contrast, the former buncher effect is not identified by observed data, because it serves to correct the first-order approximation's implicit assumption that some taxpayers will locate at the new bunching location $z^{\star\star}$ when in fact, they will end up choosing to locate some point on the new budget constraint which is infeasible on the pre-reform budget constraint. As we don't observe these choices in pre-reform data, we cannot identify this effect. Notice, this non-identification result stems from the same fundamental problem that underlies recent non-identification results regarding the standard bunching method: counterfactual decisions of bunchers are not observed.[37]

While this is a valuable result in itself, it is possible to expand on the practical value of the result by adopting a partial identification lens. For example, without any additional assumptions, we can identify a lower bound

[37]As I discuss in the previous section, this is what causes non-identification in the standard bunching method when preferences are quasi-linear and isoelastic, and the ETI is homogeneous. Outside this special case, knowledge of counterfactual decisions of bunchers is insufficient.

on the revenue effect of a discrete decrease in the kink point by evaluating equation (44) with the former buncher effect set to zero. This works because, in the case of a decrease in the kink point, the former buncher effect is always positive, so ignoring it will lead to an underestimate of the effect. A similar result holds for the case of a discrete increase in the kink point, but in this case evaluating equation (45) with the former buncher effect set to zero yields an upper bound, because for such reforms the former buncher effect is always negative.

Rather than simply assuming the former buncher effect is zero, we could use a trapezoidal approximation to the former buncher effect. It turns out that such an approximation is identified by the pre-reform distribution of taxable. In the case of a discrete decrease in a kink point we have

$$t_1 \int_{z^{\star\star}}^{z^\star} \left(z - z^{\star\star}\right) h_1\left(z\right) \mathrm{d}z \approx \frac{\left(z^\star - z^{\star\star}\right)^2}{2} t_1 h_1\left(z^\star\right) \tag{46}$$

and in the case of a discrete increase we have

$$- t_0 \int_{z^\star}^{z^{\star\star}} \left(z^{\star\star} - z\right) h_0\left(z\right) \mathrm{d}z \approx -\frac{\left(z^\star - z^{\star\star}\right)^2}{2} t_0 h_0\left(z^\star\right). \tag{47}$$

Note, similar strategies have been employed in the traditional bunching literature from its inception. Saez [2010] proposes the use of trapezoidal approximation for counterfactual distribution of income for bunchers. However, in the standard bunching literature, identification of these approximations themselves requires much stronger assumptions about agent behavior than those presented here, because they include a counterfactual unobserved density. Equations (46) and (47) do not include any such counterfactual densities because the error the former buncher effect is largest in the region of taxable income close to the original kink point $z^\star$, but shrinks to zero close to the new kink point $z^{\star\star}$.[38]

Imposing weak additional restrictions on the density of taxable income allows for a reinterpretation of approximations (46) and (47) as bounds on the former buncher effect. In the case of a kink point decrease, if $h_1\left(z\right)$ is weakly decreasing for all $z \in [z^{\star\star}, z^\star]$, then (46) provides a lower bound on the former bunching effect. Conversely, if $h_0\left(z\right)$ is weakly decreasing over this interval, then we obtain an upper bound. Similarly, in the case of a kink point increase, if $h_0\left(z\right)$ is weakly decreasing (increasing) for all $z \in [z^\star, z^{\star\star}]$, then (47) provides an upper (lower) bound on the former bunching effect. Thus, when density functions are assumed to be monotone, sharper bounds on the revenue effect of discrete changes are possible.

[38]Using higher derivatives of of the distribution of taxable income, it is possible to extend these approximations even further. In fact, for any $n \in \mathbb{N}$ an $n$th-order Taylor approximation to the former buncher effect is identified by the observed distribution, though the practical value of this extension is limited by the feasibility of estimating such derivatives.

# 6 Conclusion

Saez [2010] presents a tantalizing idea: that bunching in the observed distribution of taxable income can teach us about the response to taxation even absent any variation in the tax schedule over time or across different groups of taxpayers. At first blush, the approach seems to promise that the imposition of minor restrictions—motivated by economic theory—permits for ex ante evaluation of tax policy changes. Recent work on bunching methodology has called this promise into question, highlighting the fact that identification in the standard bunching method does not rest purely on the restrictions of economic theory but rather relies upon specific functional form assumptions, restricted forms of agent heterogeneity, and requires knowledge of unobserved counterfactual densities of taxable income.

In this paper, I have shown that by shifting our focus away from the ETI, it is possible to deliver on the promise of bunching methodology. With minimal economic theory restrictions, the bunching mass becomes a sufficient statistic for the compensated behavioral response to moving a kink point. This use of the bunching mass is fully nonparametric and requires no information beyond the observed distribution of taxable income. Building on this basic result, I developed new empirical tests for Pareto efficiency and welfarist-optimality of piecewise linear tax schedules. An empirical application to the EITC showed that these results can be extended to incorporate important features of real-world policy settings: heterogeneous tax schedules and optimization frictions. Correctly incorporating optimization frictions into the analysis is shown to have a meaningful impact on estimated revenue effects of moving a kink point.

By providing a precise, nonparametric interpretation of the bunching mass in terms of the effect of moving a kink point, the paper also sheds light on the nature of the bunching elasticity. Bunching elasticities can provide a good guide to the behavioral response to other tax reforms if that response can be expected to be sufficiently similar to the response to moving a kink point. Whether this is a reasonable expectation depends on factors which are not currently well-documented empirically, such as the variance of the ETI. Gathering evidence about these factors could therefore be of great value for providing bunching practitioners with guidance on the external validity of their elasticity estimates.

This paper's central finding also demonstrates that the identification challenges raised in the recent literature on standard bunching methodology stem from a narrow focus on a particular parameter of interest: the ETI. The concept of the elasticity of taxable income as a sufficient statistic is a powerful, but as I have shown, it is not always necessary to extract this parameter in order to learn about the welfare effects of some tax reform. Bunching methodology may therefore offer a case study in the value of ensuring that any policy-motivated empirical work is tightly integrated with a formal welfare analysis of policy reforms of interest.

# References


Thomas Aronsson, Katharina Jenderny, and Gauthier Lanot. The quality of the estimators of the eti. In *Proceedings. Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association*, volume 111, pages 1–59. JSTOR, 2018.

Katy Bergstrom and William Dodds. Optimal taxation with multiple dimensions of heterogeneity. 2021.

Marinho Bertanha, Andrew H McCallum, and Nathan Seegert. Better bunching, nicer notching. *arXiv preprint arXiv:2101.01170*, 2021.

Felix Bierbrauer, Pierre Boyer, and Emanuel Hansen. Pareto-improving tax reforms and the earned income tax credit. 2020.

Soren Blomquist, Whitney K Newey, Anil Kumar, and Che-Yuan Liang. The bunching estimator cannot identify the taxable income elasticity. 2019.

M Cattaneo, Michael Jansson, Xinwei Ma, and Joel Slemrod. Bunching designs: Estimation and inference. Working paper, UCSD Bunching Workshop, 2018.

Raj Chetty, John N Friedman, Tore Olsen, and Luigi Pistaferri. Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from danish tax records. *The quarterly journal of economics*, 126(2):749–804, 2011.

Martin Feldstein. Tax avoidance and the deadweight loss of the income tax. *Review of Economics and Statistics*, 81(4):674–680, 1999.

Aart Gerritsen. Optimal nonlinear taxation: the dual approach. 2016.

Leonard Goff. Treatment effects in bunching designs: The impact of the federal overtime rule on hours. Technical report, Working Paper, 2021.

Mikhail Golosov, Aleh Tsyvinski, and Nicolas Werquin. A variational approach to the analysis of tax systems. Technical report, National Bureau of Economic Research, 2014.

Nathaniel Hendren. Measuring economic efficiency using inverse-optimum weights. *Journal of public Economics*, 187:104198, 2020.

Laurence Jacquet and Etienne Lehmann. Optimal income taxation with composition effects. *Journal of the European Economic Association*, 19(2):1299–1341, 2021.

Wojciech Kopczuk. Tax bases, tax rates and the elasticity of reported income. *Journal of Public Economics*, 89(11-12):2093–2119, 2005.

Kory Kroft, Kavan Kucko, Etienne Lehmann, and Johannes Schmieder. Optimal income taxation with unemployment and wage responses: a sufficient statistics approach. *American Economic Journal: Economic Policy*, 12(1):254–92, 2020.

Anil Kumar and Che-Yuan Liang. Estimating taxable income responses with elasticity heterogeneity. *Journal of Public Economics*, 188:104209, 2020.

Arthur Lewbel. The identification zoo: Meanings of identification in econometrics. *Journal of Economic Literature*, 57(4):835–903, 2019.

Benjamin M Marx. The cost of requiring charities to report financial information. 2018.

Carina Neisser et al. The elasticity of taxable income: A meta-regression analysis. Technical report, University of Bonn and University of Cologne, Germany, 2021.

Emmanuel Saez. Using elasticities to derive optimal income tax rates. *The review of economic studies*, 68(1): 205–229, 2001.

Emmanuel Saez. Do taxpayers bunch at kink points? *American economic Journal: economic policy*, 2(3): 180–212, 2010.

Emmanuel Saez, Joel Slemrod, and Seth H Giertz. The elasticity of taxable income with respect to marginal tax rates: A critical review. *Journal of economic literature*, 50(1):3–50, 2012.

Joel Slemrod and Wojciech Kopczuk. The optimal elasticity of taxable income. *Journal of Public Economics*, 84(1):91–112, 2002.

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

Iván Werning. Pareto efficient income taxation. 2007.

# A Derivations

## A.1 Heterogeneous Tax Rates

Suppose that type $\theta$ agents face the following two bracket tax schedule

$$T(z;\theta) \equiv \begin{cases} t_0(\theta) z & \text{if } z \leq z^\star(\theta) \\ t_1(\theta) z + [t_0(\theta) - t_1(\theta)] z^\star(\theta) & \text{if } z > z^\star(\theta) \end{cases},$$

where $z^\star(\theta)$, $t_0(\theta)$, and $t_1(\theta)$ are type-$\theta$-specific tax schedule parameters. Let tax revenue from a type $\theta$ taxpayer be denoted as

$$\mathcal{T}(\theta) \equiv T(z(\theta);\theta).$$

Total tax revenue can then be written as

$$R \equiv \int \mathcal{T}(\theta) f(\theta) \mathrm{d}\theta.$$

Now consider a tax reform indexed by $\kappa$ which infinitesimally changes the location of the kink point for type $\theta$ taxpayers by $\frac{\mathrm{d}z^\star(\theta)}{\mathrm{d}\kappa}$. The revenue effect of this reform is

$$\frac{\mathrm{d}R}{\mathrm{d}\kappa} = \int \frac{\mathrm{d}\mathcal{T}(\theta)}{\mathrm{d}z^\star(\theta)} \frac{\mathrm{d}z^\star(\theta)}{\mathrm{d}\kappa} f(\theta) \mathrm{d}\theta$$

where

$$\frac{\mathrm{d}\mathcal{T}(\theta)}{\mathrm{d}z^\star(\theta)} \equiv \begin{cases} 0 & \text{if } z_0(\theta) \leq z^\star(\theta) \\ t_0(\theta) & \text{if } z_1(\theta) < z^\star(\theta) < z_0(\theta) \\ t_0(\theta) - t_1(\theta) & \text{if } z_1(\theta) \geq z^\star(\theta) \end{cases}.$$

As in the standard case, some taxpayers are below the kink, and face no change in tax payments, some are above the kink and face a mechanical change in tax payments, and some are (non-marginal) bunchers and face a change in tax payments due to their behavioral response.The difference is that now the kink point locations and marginal rates are variable across individuals. For a general reform we have

$$\frac{\mathrm{d}R}{\mathrm{d}\kappa} = \int_{\theta\in\Theta_1} [t_1(\theta) - t_0(\theta)] \frac{\mathrm{d}z^\star(\theta)}{\mathrm{d}\kappa} f(\theta) \mathrm{d}\theta - \int_{\theta\in\Theta_B} t_0(\theta) \frac{\mathrm{d}z^\star(\theta)}{\mathrm{d}\kappa} f(\theta) \mathrm{d}\theta \tag{48}$$

where $\Pr\{z_1(\theta) < z^\star(\theta) < z_0(\theta)\}$ is the generalized bunching mass—the fraction of taxpayers who are bunching at their own kink point—and $\Pr\{z_1(\theta) > z^\star(\theta)\}$ is the fraction of taxpayers who are located above their own kink point.

In the main text, I assume that all taxpayers have the same kink point location so that $z(\theta) = z^\star$ for all $\theta$ so that

$$\Pr\{z_1(\theta) < z^\star(\theta) < z_0(\theta)\} = H_1(z^\star) - H_0(z^\star)$$

and $\Pr\{z_1(\theta) > z^\star(\theta)\} = 1 - H_1(z^\star)$. Letting $\frac{\mathrm{d}z^\star(\theta)}{\mathrm{d}\kappa} = -1$, equation (48) simplifies to equation (22).

## A.2 Revenue Impact with Frictions

alternative approach:

$$\begin{aligned}
R(z^\star) \equiv \int_0^{z^\star} & \left[ t_0 \int_0^{z^\star/y} y\xi g(\xi)\,\mathrm{d}\xi + \int_{z^\star/y}^{\infty} [t_1 y\xi - (t_1 - t_0) z^\star] g(\xi)\,\mathrm{d}\xi \right] h_y(y; t_0)\,\mathrm{d}y \\
& + \left[ t_0 \int_0^1 z^\star \xi g(\xi)\,\mathrm{d}\xi + \int_1^\infty [t_1 z^\star \xi - (t_1 - t_0) z^\star] g(\xi)\,\mathrm{d}\xi \right] (H_y(z^\star; t_1) - H_y(z^\star; t_0)) \\
& \qquad + \int_{z^\star}^{\infty} \left[ t_0 \int_0^{z^\star/y} y\xi g(\xi)\,\mathrm{d}\xi + \int_{z^\star/y}^{\infty} [t_1 y\xi - (t_1 - t_0) z^\star] g(\xi)\,\mathrm{d}\xi \right] h_y(y; t_1)\,\mathrm{d}y
\end{aligned}$$

$$R(z^\star) \;\; \equiv \;\; \int_0^{z^\star} R(y; z^\star) h_y(y; t_0)\,\mathrm{d}y \; + \; R(z^\star; z^\star)(H_y(z^\star; t_1) - H_y(z^\star; t_0)) \; + \; \int_{z^\star}^{\infty} R(y; z^\star) h_y(y; t_1)\,\mathrm{d}y$$

where

$$R(y; z^\star) = t_0 \int_0^{z^\star/y} y\xi g(\xi)\,\mathrm{d}\xi + \int_{z^\star/y}^{\infty} [t_1 y\xi - (t_1 - t_0) z^\star] g(\xi)\,\mathrm{d}\xi$$

thus

$$\begin{aligned}
R'(z^\star) = \underbrace{R(z^\star; z^\star) h_y(z^\star; t_0) + R(z^\star; z^\star)(h_y(z^\star; t_1) - h_y(z^\star; t_0)) - R(z^\star; z^\star) h_y(z^\star; t_1)}_{=0} \\
+ \int_0^{z^\star} \frac{\partial R(y; z^\star)}{\partial z^\star} h_y(y; t_0)\,\mathrm{d}y + \left[ \left. \frac{\partial R(y; z^\star)}{\partial y} \right|_{y=z^\star} + \frac{\partial R(y; z^\star)}{\partial z^\star} \right] (H_y(z^\star; t_1) - H_y(z^\star; t_0)) + \int_{z^\star}^{\infty} \frac{\partial R(y; z^\star)}{\partial z^\star} h_y(y; t_1)\,\mathrm{d}y
\end{aligned}$$

$$\begin{aligned}
\frac{\partial R(y; z^\star)}{\partial z^\star} &= \underbrace{t_0 \frac{z^\star}{y} g\left(\frac{z^\star}{y}\right) - \left[ t_1 \frac{z^\star}{y} - t_1 \frac{z^\star}{y} + t_0 \frac{z^\star}{y} \right] g\left(\frac{z^\star}{y}\right)}_{=0} \underbrace{- \int_{z^\star/y}^{\infty} (t_1 - t_0) g(\xi)\,\mathrm{d}\xi}_{\text{mechanical effect}} \\
&= -(t_1 - t_0)\left(1 - G\left(\frac{z^\star}{y}\right)\right)
\end{aligned}$$

$$\begin{aligned}\frac{\partial R\left(y;z^{\star}\right)}{\partial y} &= t_0\int_0^{z^{\star}/y}\xi g\left(\xi\right)\mathrm{d}\xi + t_1\int_{z^{\star}/y}^{\infty}\xi g\left(\xi\right)\mathrm{d}\xi - t_0\frac{z^{\star}}{y^2}y\frac{z^{\star}}{y}g\left(\frac{z^{\star}}{y}\right) + \frac{z^{\star}}{y^2}\left[t_1 y\frac{z^{\star}}{y} - \left(t_1-t_0\right)z^{\star}\right]g\left(\frac{z^{\star}}{y}\right)\\ &= t_0\int_0^{z^{\star}/y}\xi g\left(\xi\right)\mathrm{d}\xi + t_1\int_{z^{\star}/y}^{\infty}\xi g\left(\xi\right)\mathrm{d}\xi \underbrace{-t_0\left(\frac{z^{\star}}{y}\right)^2 g\left(\frac{z^{\star}}{y}\right) + t_0\left(\frac{z^{\star}}{y}\right)^2 g\left(\frac{z^{\star}}{y}\right)}_{=0}\end{aligned}$$

$$\begin{aligned}\left.\frac{\partial R\left(y;z^{\star}\right)}{\partial y}\right|_{y=z^{\star}} &= t_0\int_0^1\xi g\left(\xi\right)\mathrm{d}\xi + t_1\int_1^{\infty}\xi g\left(\xi\right)\mathrm{d}\xi\\ &= t_0 G\left(1\right)\int_0^1\xi\frac{g\left(\xi\right)}{G\left(1\right)}\mathrm{d}\xi + t_1\left(1-G\left(1\right)\right)\int_1^{\infty}\xi\frac{g\left(\xi\right)}{1-G\left(1\right)}\mathrm{d}\xi\\ &= t_0 G\left(1\right)\int_0^1\xi\frac{g\left(\xi\right)}{G\left(1\right)}\mathrm{d}\xi + t_1\left(1-G\left(1\right)\right)\int_1^{\infty}\xi\frac{g\left(\xi\right)}{1-G\left(1\right)}\mathrm{d}\xi\end{aligned}$$

$$\begin{aligned}R'\left(z^{\star}\right) &= \underbrace{\int_0^{z^{\star}}\frac{\partial R\left(y;z^{\star}\right)}{\partial z^{\star}}h_y\left(y;t_0\right)\mathrm{d}y + \frac{\partial R\left(y;z^{\star}\right)}{\partial z^{\star}}\left(H_y\left(z^{\star};t_1\right)-H_y\left(z^{\star};t_0\right)\right) + \int_{z^{\star}}^{\infty}\frac{\partial R\left(y;z^{\star}\right)}{\partial z^{\star}}h_y\left(y;t_1\right)\mathrm{d}y}_{\text{mechanical effect}} + \underbrace{\left.\frac{\partial R\left(y;z^{\star}\right)}{\partial y}\right|_{y=z^{\star}}\left(I\right.}_{\text{behaviora}}\\ &= \underbrace{-\left(t_1-t_0\right)\left[\int_0^{z^{\star}}\int_{z^{\star}/y}^{\infty}g\left(\xi\right)\mathrm{d}\xi h_y\left(y;t_0\right)\mathrm{d}y + \int_1^{\infty}g\left(\xi\right)\mathrm{d}\xi\left(H_y\left(z^{\star};t_1\right)-H_y\left(z^{\star};t_0\right)\right) + \int_{z^{\star}}^{\infty}\int_{z^{\star}/y}^{\infty}g\left(\xi\right)\mathrm{d}\xi h_y\left(y;t_1\right)\mathrm{d}y\right]}_{\text{mechanical effect}} +\\ &= \underbrace{-\left(t_1-t_0\right)\left(1-H_z\left(z^{\star};t_0\right)\right)}_{\text{mechanical effect}} + \underbrace{\left(t_0 G\left(1\right)\frac{\mathbb{E}\left[z|y=z^{\star},z\leq z^{\star}\right]}{z^{\star}} + t_1\left(1-G\left(1\right)\right)\frac{\mathbb{E}\left[z|y=z^{\star},z>z^{\star}\right]}{z^{\star}}\right)\left(H_y\left(z^{\star};t_1\right)-H_y\left(z\right.\right.}_{\text{behavioral response effect}}\end{aligned}$$

$$\begin{aligned}R'\left(z^{\star}\right) = -\left(t_1-t_0\right)&\left[\int_0^{z^{\star}}\left(1-G\left(\frac{z^{\star}}{y}\right)\right)h_y\left(y;t_0\right)\mathrm{d}y + \left(1-G\left(1\right)\right)\left(H_y\left(z^{\star};t_1\right)-H_y\left(z^{\star};t_0\right)\right) + \int_{z^{\star}}^{\infty}\left(1-G\left(\frac{z^{\star}}{y}\right)\right)h_y\left(y\right.\right.\\ &+\left[t_0\int_0^1\xi g\left(\xi\right)\mathrm{d}\xi + t_1\int_1^{\infty}\xi g\left(\xi\right)\mathrm{d}\xi\right]\left(H_y\left(z^{\star};t_1\right)-H_y\left(z^{\star};t_0\right)\right)\end{aligned}$$

$$\begin{aligned}R'\left(z^{\star}\right) = -\left(t_1-t_0\right)&\left[\int_0^{z^{\star}}\left(1-G\left(\frac{z^{\star}}{y}\right)\right)h_y\left(y;t_0\right)\mathrm{d}y - \left(1-G\left(1\right)\right)\left(H_y\left(z^{\star};t_1\right)-H_y\left(z^{\star};t_0\right)\right) + \int_{z^{\star}}^{\infty}\left(1-G\left(\frac{z^{\star}}{y}\right)\right)h_y\left(y\right.\right.\\ &+\left[t_0\int_0^1\xi g\left(\xi\right)\mathrm{d}\xi + t_1\int_1^{\infty}\xi g\left(\xi\right)\mathrm{d}\xi\right]\left(H_y\left(z^{\star};t_1\right)-H_y\left(z^{\star};t_0\right)\right)\end{aligned}$$

$$\begin{aligned}\int_1^{\infty}\xi g\left(\xi\right)\mathrm{d}\xi &= \left[-\xi\left(1-G\left(\xi\right)\right)\right]_1^{\infty} + \int_1^{\infty}\left(1-G\left(\xi\right)\right)\mathrm{d}\xi\\ &= \left(1-G\left(1\right)\right) + \int_1^{\infty}\left(1-G\left(\xi\right)\right)\mathrm{d}\xi\end{aligned}$$

$$\int_0^1 \xi g(\xi)\,\mathrm{d}\xi = G(1) - \int_0^1 G(\xi)\,\mathrm{d}\xi$$

$$t_0 \int_0^1 \xi g(\xi)\,\mathrm{d}\xi + t_1 \int_1^\infty \xi g(\xi)\,\mathrm{d}\xi = t_0 G(1) + t_1(1 - G(1)) - t_0 \int_0^1 G(\xi)\,\mathrm{d}\xi + t_1 \int_1^\infty (1 - G(\xi))\,\mathrm{d}\xi$$

$$R'(z^\star) = -(t_1 - t_0)\left[\int_0^{z^\star} \left(1 - G\left(\frac{z^\star}{y}\right)\right) h_y(y;t_0)\,\mathrm{d}y - (1 - G(1))(H_y(z^\star;t_1) - H_y(z^\star;t_0)) + \int_{z^\star}^\infty \left(1 - G\left(\frac{z^\star}{y}\right)\right) h_y(y\right.$$
$$+\left[t_0 G(1) + t_1(1 - G(1)) - t_0 \int_0^1 G(\xi)\,\mathrm{d}\xi + t_1 \int_1^\infty (1 - G(\xi))\,\mathrm{d}\xi\right](H_y(z^\star;t_1) - H_y(z^\star;t_0))$$

## A.3 Income and Participation Effects

Suppose that taxpayers face the two-bracket tax schedule (2). For each type of taxpayer $\theta$, I assume that they make two choices:

1. Whether or not to enter the labor force.
2. Conditional on entering the labor force, how much income to earn.

Allowing for the possibility of income effects, a type $\theta$ taxpayer who enters the labor force will choose a value of taxable income that solves the problem utility maximization (1) where their utility function is assumed to satisfy Assumption 1. Recall, $z(t, V; \theta)$ is the choice of taxable income a type $\theta$ agent would make if facing a linear tax schedule with marginal tax rate $t$ and virtual income $V$. Conditional on entering the labor force, the choice of taxable income a type $\theta$ agent makes when facing the kinked tax schedule (2) can be written as

$$z(\theta) = \begin{cases} z(t_0, V_0; \theta) & \text{if } z(t_0, V_0; \theta) < z^\star \\ z(t_1, V_1; \theta) & \text{if } z(t_1, V_1; \theta) > z^\star \\ z^\star & \text{if } z(t_1, V_1; \theta) \leq z^\star \leq z(t_0, V_0; \theta) \end{cases},$$

where $V_0 \equiv G$ and $V_1 \equiv (t_1 - t_0) z^\star + G$. As in section (1.1), it remains that case that all taxpayers must satisfy one, and only one of the three conditions in equation listed above. The alternative possibility $(z(t_0;\theta) < z^\star < z(t_1;\theta))$ would violate SARP, since the $z(t_0;\theta)$ would be affordable under $T_1$ and $z(t_1;\theta)$ would be affordable under $T_0$.

Let $v_E(\theta)$ be the indirect utility associated with a type $\theta$ agent's utility maximization problem:

$$v_E(\theta) \equiv u(z(\theta) - T(z(\theta)), z(\theta); \theta).$$

On the other hand, let $v_U(\theta)$ be the utility a type $\theta$ agent would receive if unemployed:

$$v_U(\theta) \equiv u(T(0), 0; \theta).$$

I assume that a type $\theta$ taxpayer only chooses to enter the labor force if the private benefit of doing so exceeds some fixed cost of work $\xi$. Assuming that $\xi \sim Q(\cdot|\theta)$ for all type $\theta$ taxpayers, the probability that such a taxpayer enters the labor force is

$$Q(\theta) \equiv Q(v_E(\theta) - v_U(\theta)|\theta).$$

Total tax revenue in this economy can constructed by integrating over the type distribution:

$$R(z^\star) \equiv t_0 \int_{\theta \in \Theta_0(z^\star)} z(t_0, V_0; \theta) Q(\theta) f(\theta) \mathrm{d}\theta + t_0 z^\star \int_{\theta \in \Theta_B(z^\star)} Q(\theta) f(\theta) \mathrm{d}\theta + \int_{\theta \in \Theta_1(z^\star)} [t_1(z(t_1, V_1; \theta) - z^\star) + t_0 z^\star] Q(\theta) f(\theta) \tag{49}$$

where

$$\Theta_0(z^\star) = \{\theta \in \Theta : z(t_0, V_0; \theta) < z^\star\}$$

is the set of types who choose to locate at an interior point the lower tax bracket if they enter the labor force,

$$\Theta_1(z^\star) = \{\theta \in \Theta : z(t_1, V_1; \theta) > z^\star\}$$

is the set of types who choose to locate in the upper tax bracket if they enter the labor force, and

$$\Theta_B(z^\star) = \{\theta \in \Theta : z(t_1, V_1; \theta) \leq z^\star \leq z(t_0, V_0; \theta)\}$$

is the set of types who choose to bunch if they enter the labor force.

Differentiating equation (49), and applying the multidimensional version of Leibniz rule, we obtain

$$
\begin{aligned}
-R'(z^\star) = & \underbrace{(t_1 - t_0)\int_{\theta\in\Theta_1} Q(\theta) f(\theta)\,\mathrm{d}\theta}_{\text{mechanical effect}} \underbrace{-t_0 \int_{\theta\in\Theta_B} Q(\theta) f(\theta)\,\mathrm{d}\theta}_{\text{substitution effect}} \\
& \underbrace{-t_1 (t_1 - t_0)\int_{\theta\in\Theta_1} \frac{\partial z(t_1, V_1;\theta)}{\partial V_1} Q(\theta) f(\theta)\,\mathrm{d}\theta}_{\text{income effect}} \\
& \underbrace{-t_0 z^\star \int_{\theta\in\Theta_B} (1 - t_0 - MRS(z^\star,\theta)) \frac{\partial u(\theta)}{\partial c} q(\theta) f(\theta)\,\mathrm{d}\theta}_{\text{participation effect (bunchers)}} \\
& \underbrace{-\int_{\theta\in\Theta_1} [t_1 (z(t_1, V_1;\theta) - z^\star) + t_0 z^\star](t_1 - t_0) \frac{\partial u(\theta)}{\partial c} q(\theta) f(\theta)\,\mathrm{d}\theta}_{\text{participation effect (upper bracket taxpayers)}} \qquad (50)
\end{aligned}
$$

where $q(\theta) \equiv Q'(v_E(\theta) - v_U(\theta)|\theta)$ is the density of the fixed cost of working for type $\theta$ agents evaluated at $v_E(\theta) - v_U(\theta)$.

The first two terms in equation (50) are familiar from the discussion in section (1.2): the mechanical effect of the reform and the relocation effect. The *income effect* captures the revenue consequences of the behavioral responses of taxpayers located above the kink point who may choose to respond to their loss of income by increasing their earnings. For an infinitesimal reduction in the location of the kink point, these taxpayers experienced an income loss of $t_1 - t_0$, and their response is proportional to the size of this loss. If leisure is a normal good then $\frac{\partial z(t_1,V_1;\theta)}{\partial V_1} \leq 0$ for all types $\theta$, so this effect will be positive as long as $t_1 > 0$.

The *participation effects* capture the fact that some taxpayers will choose to exit the labor force after the reform. Specifically, these are taxpayers who were at the margin between participating and not participating before the reform. Because the utility these agents derive from working declines following the reform, they will no longer choose to do so. For bunchers, the resulting change in the participation rate is proportional to the first-order reduction in utility caused by their behavioral response to the reform (as discussed in section (2.2)). For upper bracket taxpayers, the resulting change in the participation rate is proportional to the utility loss they experience due to their reduction in virtual income.

Let

$$\pi(\theta) \equiv \frac{1}{Q(\theta)} \frac{\partial u(\theta)}{\partial c} q(\theta)$$

be the semi-elasticity of labor force participation rate of type $\theta$ agents with respect to their consumption $c(\theta)$ and let

$$\eta(\theta) \equiv -\frac{\partial z_1(\theta z(t_1, V_1;\theta))}{\partial V_1} \geq 0$$

be the magnitude of the income effect for type $\theta$ agents (conditional on labor force participation). We can rewrite equation (50) as

$$
\begin{aligned}
-R'(z^\star) = \underbrace{(t_1 - t_0)(1 - H_1(z^\star))}_{\text{mechanical effect}} \underbrace{-t_0 (H_1(z^\star) - H_0(z^\star))}_{\text{substitution effect}} \\
+ \underbrace{t_1 (t_1 - t_0)(1 - H_1(z^\star)) \hat{\eta}^+(z^\star)}_{\text{income effect}} \\
\underbrace{-t_0 z^\star (H_1(z^\star) - H_0(z^\star)) \mathbb{E}[k(\theta)\pi(\theta) | z(\theta) = z^\star]}_{\text{participation effect (bunchers)}} \\
\underbrace{-t_1 (t_1 - t_0)(1 - H_1(z^\star)) \mathbb{E}[(t_1(z(\theta) - z^\star) + t_0 z^\star)\pi(\theta) | z(\theta) > z^\star]}_{\text{participation effect (upper bracket taxpayers)}} \quad (51)
\end{aligned}
$$

where

$$\hat{\eta}^+(z^\star) \equiv \mathbb{E}[\eta(\theta) | z(\theta) > z^\star]$$

is the average income effect of the kink point and

$$k(\theta) \equiv 1 - t_0 - MRS(z^\star, \theta).$$

The total fiscal externality of moving the kink point is

$$FE(z^\star) = FE_{sub}(z^\star) + FE_{inc}(z^\star) + FE_{part}(z^\star),$$

where the portion attributable to the substitution effect is

$$FE_{sub}(z^\star) \equiv -\frac{t_0 (H_1(z^\star) - H_0(z^\star))}{(t_1 - t_0)(1 - H_1(z^\star))} \leq 0.$$

The parts of $FE(z^\star)$ associated with the income and participation effects can be obtain by normalizing these effects (as they appear in equation (51)) by the mechanical effect:

$$FE_{inc}(z^\star) \equiv t_1 \hat{\eta}^+(z^\star)$$

and

$$FE_{part}(z^\star) = \underbrace{z^\star \mathbb{E}[k(\theta)\pi(\theta) | z(\theta) = z^\star] FE_{sub}(z^\star)}_{\text{buncher participation effect}} \underbrace{-t_1 \mathbb{E}[(t_1(z(\theta) - z^\star) + t_0 z^\star)\pi(\theta) | z(\theta) > z^\star]}_{\text{upper bracket participation effect}}$$

Interestingly, the fiscal externality of the buncher participation effect is proportional to the fiscal externality caused by the substitution effect. Nonetheless, because $\mathbb{E}\left[k\left(\theta\right)\pi\left(\theta\right)|z\left(\theta\right)=z^{\star}\right]$ is unknown, this fiscal externality is not identified by the observed distribution of taxable income.

## A.4 Effect of Tax Rate on Local Average ETI

Note that the local average ETI at $z^{\star}$ can be written as

$$\bar{\varepsilon}^{c}\left(z^{\star};t\right)=\frac{\int_{\psi}\varepsilon^{c}\left(t;w^{\star}\left(t;\psi\right),\psi\right)h\left(z^{\star}|\psi;t\right)f_{\psi}\left(\psi\right)\mathrm{d}\psi}{\int_{\psi}h\left(z^{\star}|\psi;t\right)f_{\psi}\left(\psi\right)\mathrm{d}\psi}.$$

Differentiating this expression, we obtain

$$\begin{aligned}\frac{\mathrm{d}\bar{\varepsilon}^{c}\left(z^{\star};t\right)}{\mathrm{d}t}=&\frac{\int_{\psi}\varepsilon^{c}\left(t;w^{\star}\left(t;\psi\right),\psi\right)\frac{\partial h\left(z^{\star}|\psi;t\right)}{\partial t}f_{\psi}\left(\psi\right)\mathrm{d}\psi}{\int_{\psi}h\left(z^{\star}|\psi;t\right)f_{\psi}\left(\psi\right)\mathrm{d}\psi}-\bar{\varepsilon}^{c}\left(z^{\star};t\right)\frac{\int_{\psi}\frac{\partial h\left(z^{\star}|\psi;t\right)}{\partial t}f_{\psi}\left(\psi\right)\mathrm{d}\psi}{\int_{\psi}h\left(z^{\star}|\psi;t\right)f_{\psi}\left(\psi\right)\mathrm{d}\psi}\\&+\int_{\psi}\left[\frac{\partial\varepsilon^{c}\left(t;w^{\star}\left(t;\psi\right),\psi\right)}{\partial t}+\frac{\partial\varepsilon^{c}\left(t;w^{\star}\left(t;\psi\right),\psi\right)}{\partial w}\frac{\partial w^{\star}\left(t;\psi\right)}{\partial t}\right]\frac{h\left(z^{\star}|\psi;t\right)f_{\psi}\left(\psi\right)}{h\left(z^{\star};t\right)}\mathrm{d}\psi.\quad(52)\end{aligned}$$

To further simplify this expression, note that by differentiating the equality $H\left(z\left(t;w,\psi\right)|\psi;t\right)=F\left(w|\psi\right)$ we can obtain

$$\frac{\partial z}{\partial w}h\left(z\left(t;w,\psi\right)|\psi;t\right)=f\left(w|\psi\right),$$

and by further differentiating this expression we get

$$\frac{\partial h\left(z\left(t;w,\psi\right)|\psi;t\right)}{\partial t}=-\left(\frac{\partial^{2}z}{\partial w\partial t}/\frac{\partial z}{\partial w}\right)h\left(z\left(t;w,\psi\right)|\psi;t\right)-\frac{\partial z}{\partial t}h'\left(z\left(t;w,\psi\right)|\psi;t\right).\quad(53)$$

Further note that by differentiating both sides of the equality $z\left(t;w^{\star}\left(t;\psi\right),\psi\right)=z^{\star}$ we obtain

$$\frac{\partial w^{\star}\left(t;\psi\right)}{\partial t}=-\frac{\partial z\left(t;w^{\star}\left(t;\psi\right),\psi\right)/\partial t}{\partial z\left(t;w^{\star}\left(t;\psi\right),\psi\right)/\partial w}.$$

Combining this expression with equation (53), together with the definitions of $\varepsilon^{c}\left(t;w^{\star}\left(t;\psi\right),\psi\right)$ and $\alpha\left(z^{\star}|\psi;t\right)$ we get

$$\frac{\partial h\left(z^{\star}|\psi;t\right)}{\partial t}=\frac{\frac{\partial\varepsilon^{c}\left(t;w^{\star}\left(t;\psi\right),\psi\right)}{\partial w}\frac{\partial w^{\star}\left(t;\psi\right)}{\partial t}}{\varepsilon^{c}\left(t;w^{\star}\left(t;\psi\right),\psi\right)}h\left(z^{\star}|\psi;t\right)-\frac{1}{1-t}\varepsilon^{c}\left(t;w^{\star}\left(t;\psi\right),\psi\right)\left(1+\alpha\left(z^{\star}|\psi;t\right)\right)h\left(z^{\star}|\psi;t\right),\quad(54)$$

and substituting this into equation (52), we obtain equation (33).

$$\frac{\mathrm{d}\bar{\varepsilon}^c\left(z^\star;t\right)}{\mathrm{d}t} = -\frac{1}{1-t}\mathrm{Cov}\left(\varepsilon^c,\left(1+\alpha\right)\varepsilon^c|z=z^\star\right)+\mathbb{E}\left[\frac{\partial\varepsilon^c}{\partial t}|z=z^\star\right]+\mathbb{E}\left[\left(2-\frac{\bar{\varepsilon}^c\left(z^\star;t\right)}{\varepsilon^c}\right)\frac{\partial\varepsilon^c}{\partial w}\frac{\partial w^\star\left(t;\psi\right)}{\partial t}|z=z^\star\right]$$

### A.5 Connecting Fiscal Externalities of Different Reforms

First, note that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{h\left(z^\star;t\right)}{1-t}\bar{\varepsilon}^c\left(z^\star;t\right)\right] = \frac{h\left(z^\star;t\right)}{1-t}\frac{\partial\bar{\varepsilon}^c\left(z^\star;t\right)}{\partial t}+\left(\frac{h\left(z^\star;t\right)}{\left(1-t\right)^2}+\frac{1}{1-t}\frac{\partial h\left(z^\star;t\right)}{\partial t}\right)\bar{\varepsilon}^c\left(z^\star;t\right). \tag{55}$$

Integrating over equation (54) we obtain $\frac{\partial h(z^\star;t)}{\partial t}$ and by combining this with equation (33) we can re-write equation (55) as

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{h\left(z^\star;t\right)}{1-t}\bar{\varepsilon}^c\left(z^\star;t\right)\right] = \frac{h\left(z^\star;t\right)}{1-t}\left[\frac{1}{1-t}\left(\bar{\varepsilon}^c\left(z^\star;t\right)-\mathbb{E}\left[\left(1+\alpha\right)\left(\varepsilon^c\right)^2|z=z^\star\right]\right)+\mathbb{E}\left[\frac{\partial\varepsilon^c}{\partial t}|z=z^\star\right]+2\mathbb{E}\left[\frac{\partial\varepsilon^c}{\partial w}\frac{\partial w^\star\left(t;\psi\right)}{\partial t}|z=z^\star\right.\right.$$

Rearranging this expression gives equation (39).

# B Multiple Tax Brackets and Joint Reforms

My method can be easily extended to incorporate convex, piecewise linear tax schedules with more than two brackets.[39]

Suppose the tax schedule is

$$T\left(z\right) \equiv \max_{i\in\{0,1,\ldots,n\}} T_i\left(z\right) \tag{56}$$

where for each $i \in \{0,1,\ldots,n\}$,

$$T_i\left(z\right) \equiv t_i z - V_i. \tag{57}$$

The tax schedule is thus fully characterized by a vector of bracket-specific marginal tax rates $\mathbf{t} = (t_0, t_1, \ldots, t_n)$ and an accompanying vector of bracket-specific virtual incomes $\mathbf{V} = (V_0, V_1, \ldots, V_n)$. Suppose that $(\mathbf{V}, \mathbf{t})$ satisfies $V_i > V_{i-1}$ and $t_i > t_{i-1}$ for all $i \in \{1, 2, \ldots, n\}$. Then equations (56) and (57) describe a convex, piecewise linear tax schedule with $n+1$ different tax brackets. The tax schedule features $n$ different convex kink

[39] Indeed, it can actually be extended to include any budget set formed by the intersection of some finite set of smooth, convex budget sets. That is to say, the tax schedule within each tax bracket can be nonlinear, as long as it is convex and continuously differentiable. Thus, for example, my method could be readily be applied to analyze any bunching observed at the first kink point in the German income tax schedule, where the marginal tax rate rises discontinuously from 0 to 14% and thereafter increases continuously until reaching 42% (https://taxsummaries.pwc.com/germany/individual/taxes-on-personal-income). Such an extension is straightforward, and for simplicity I will focus here on only the piecewise linear case.

points whose locations are implicitly defined by the vectors of tax rates and virtual incomes $(\mathbf{V}, \mathbf{t})$. A taxpayer in the $i$th bracket who is not bunching at a kink point has income some income $z \in \left(\frac{V_i - V_{i-1}}{t_i - t_{i-1}}, \frac{V_{i+1} - V_i}{t_{i+1} - t_i}\right)$ with the kink point $z_i^\star \equiv \frac{V_{i+1} - V_i}{t_{i+1} - t_i}$. serving as the boundary between this bracket and the next.

As in section 1, the actual tax schedule is thus composed of many constituent linear tax schedules of the form given by (57). Let $z_i(\theta)$ be the choice of taxable income a type $\theta$ taxpayer would make if they faced a particular linear tax $T_i(z)$ and let $H(z; t_i)$ be the resulting distribution of taxable income. Under assumptions 1 and 2, the corresponding density of taxable income under this linear schedule $(h_i(z))$ is continuous. Furthermore, we can describe the observed choices of taxpayers under the piecewise linear tax schedule (56) as

$$z(\theta) = \begin{cases} z_i(\theta) & \text{if } z_i(\theta) \in \left(z_{i-1}^\star, z_i^\star\right) \\ z_i^\star & \text{if } z_{i+1}(\theta) < z_i^\star < z_i(\theta) \end{cases} \tag{58}$$

That is to say, if a taxpayer's choice on a particular counterfactual linear tax schedule $T_i(z)$ is located in the tax bracket where the actual tax schedule (56) coincides with $T_i(z)$, then the taxpayer will choose an interior solution. On the other hand, the taxpayer will locate at a given kink point $z_i^\star$ which divides the $i$-th and $(i+1)$-th tax brackets if their choice under the counterfactual linear tax schedule $T_i(z)$ is located above the kink point and their choice under $T_{i+1}(z)$ is located below the kink point. By the strong axiom of revealed preference, all taxpayers will satisfy one and only one of these possible conditions.[40]

It follows from equation (5), that the observed distribution of taxable income will be

$$H(z) = \begin{cases} H(z; t_i) & \text{if } z \in \left(z_{i-1}^\star, z_i^\star\right) \\ H(z_i^\star; t_{i+1}) - H(z_i^\star; t_i) & \text{if } z = z_i^\star \end{cases} \tag{59}$$

Tax revenue under the tax schedule (56) can be written as a

$$R(\mathbf{z}^\star) \equiv \sum_{i=0}^{n} \left( \int_{z_{i-1}^\star}^{z_i^\star} T_i(z)\, h(z; t_i)\, \mathrm{d}z \right) + \sum_{i=1}^{n} t_i z_i^\star \left(H(z_i^\star; t_{i+1}) - H(z_i^\star; t_i)\right) \tag{60}$$

where $T_i(z; \mathbf{z}^\star)$ refers to the counterfactual linear tax schedules of equation (57). I have adjusted the notation here to make clear that, in general, the virtual income for each of these counterfactual tax schedules depends

[40] That at least one such condition must hold follows from the fact that the agent must choose some location on the budget set. To see why only one condition can hold, note that the budget set induced by the actual tax schedule (56) is simply the intersection of the budget sets induced by our $n+1$ counterfactual linear tax schedules. This means that if for some type $\theta$ we have $z_i(\theta) \in \left(z_i^\star, z_{i+1}^\star\right)$ and $z_j(\theta) \in \left(z_j^\star, z_{j+1}^\star\right)$ for $i \neq j$, then $z_i(\theta) \succ z_j(\theta)$ by SARP because $z_j(\theta)$ is affordable under the linear schedule $T_i$. Similarly, $z_j(\theta) \succ z_i(\theta)$ because $z_i(\theta)$ is affordable under the linear schedule $T_j$, a contradiction. Similarly, suppose that for type $\theta$, $z_{i+1}(\theta) < z_{i+1}^\star < z_i(\theta)$ and $z_{j+1}(\theta) < z_{j+1}^\star < z_j(\theta)$ for some $i \neq j$. Suppose, without loss of generality, that $i > j$. This implies that $z_i(\theta)$ is affordable under $T_{j+1}$ and that $z_{j+1}(\theta)$ is affordable under $T_i$. Thus, by SARP, we have $z_i(\theta) \succ z_{j+1}(\theta)$ and $z_{j+1}(\theta) \succ z_i(\theta)$. Similar reasoning completes the proof.

on multiple kink point locations. Note as well, that for equation 60 to provide a full description of tax revenue requires adopting the convention that $V_{-1} = V_0$ so that $z^\star_{-1} = 0$.

If we continue to rule out income effects (assumption 3), the first-order revenue effect of equally decreasing the virtual income level for tax brackets with an index $j > i$ is

$$-\frac{\partial R(\mathbf{z}^\star)}{\partial \kappa} = 1 - H(z^\star_i; t_{i+1}) - \frac{t_i}{t_{i+1} - t_i}\left(H(z^\star_i; t_{i+1}) - H(z^\star_i; t_i)\right). \tag{61}$$

But notice, this expression is proportional to the first revenue effect of decreasing a kink because that is exactly what such a reform does: it moves the kink point by adjust virtual income levels/Consequently the first-order welfare effects are also the same. Thus, Theorems 1–4 continue to apply in the case of multiple tax brackets. The fiscal externality of decreasing all $V_j$ for $j > i$ is

$$FE(z^\star_i) \equiv -\frac{t_i\left(H(z^\star_i; t_{i+1}) - H(z^\star_i; t_i)\right)}{(t_{i+1} - t_i)\left(1 - H(z^\star_i; t_{i+1})\right)}. \tag{62}$$

## B.1 Moving Two Kinks

Suppose instead of moving a full set of virtual incomes we only alter decrease the virtual income $V_i$ in the $i$th bracket. Such a reform has the effect of implicitly changing the location of both kink points that border the bracket, decreasing the lower kink point ($z^\star_{i-1}$) and increasing the upper kink point ($z^\star_i$). By considering this reform, we can develop tests of Pareto efficiency and welfarist optimality that go beyond the tests developed in section 1.

Consider a tax reform, indexed by $\kappa \in \mathbb{R}$, that simultaneously moves the locations of two kinks points $z^\star_i$ and $z^\star_j$, satisfying $z^\star_j > z^\star_i$. In particular, the reform decreases the location of $z^\star_i$ and increases the location of $z^\star_j$ as follows

$$\frac{\partial z^\star_i}{\partial \kappa} = -1 \qquad \frac{\partial z^\star_j}{\partial \kappa} = \frac{t_i - t_{i-1}}{t_j - t_{j-1}}, \tag{63}$$

while leaving the locations of all other kink points constant. Notice, independently decreasing the location of the lower kink point $z^\star_i$ would result in a mechanical effect that collects additional revenue from all taxpayers above this kink. But by increasing the location of the higher kink point $z^\star_j$ in just the right proportion, this joint reform offsets this effect for any taxpayers above $z^\star_j$, keeping the mechanical effect confined to only those taxpayers between the two kink points under consideration ($z^\star_i$ and $z^\star_j$).

Thus, absent income effects, the first-order effect of this reform on tax revenue is

$$-\frac{\mathrm{d}R\left(\mathbf{z}^{\star}\right)}{\mathrm{d}V_i} = \underbrace{H\left(z_i^{\star};t_i\right) - H\left(z_{i-1}^{\star};t_i\right)}_{\text{mechanical effect}} \tag{64}$$

$$\underbrace{-\left(\frac{t_{i-1}}{t_i - t_{i-1}}\right)\overbrace{\left(H\left(z_{i-1}^{\star};t_i\right) - H\left(z_{i-1}^{\star};t_{i-1}\right)\right)}^{\text{bunching mass at } z_j^{\star}} + \frac{t_i}{t_{i+1}-t_i}\overbrace{\left(H\left(z_i^{\star};t_{i+1}\right) - H\left(z_i^{\star};t_i\right)\right)}^{\text{bunching mass at } z_i^{\star}}}_{\text{revenue effect of behavioral response}}.$$

As with single kink point reforms, this effect is identified from the observed distribution of taxable income, though it now requires using the size of two bunching masses (one for each kink point). However, this joint reform differs from single kink reforms in a key respect: the behavioral response consists of both a positive and negative component. On the one hand, the reform decreases taxable income for the bunchers at the lower kink point, a behavioral response that decreases tax revenue. On the other hand, the reform increases taxable income for bunchers at the higher kink point, a behavioral response that *increases tax revenue.*

The fiscal externality of this tax reform is

$$FE^{V_i} \equiv -\frac{\frac{t_{i-1}}{t_i - t_{i-1}}\left(H\left(z_{i-1}^{\star};t_i\right) - H\left(z_{i-1}^{\star};t_{i-1}\right)\right) - \frac{t_i}{t_{i+1}-t_i}\left(H\left(z_i^{\star};t_{i+1}\right) - H\left(z_i^{\star};t_i\right)\right)}{\left(H\left(z_i^{\star};t_i\right) - H\left(z_{i-1}^{\star};t_i\right)\right)}. \tag{65}$$

We can re-write this as a function of the fiscal externalities of decreasing the two kink points by themselves (using equation 62):

$$FE^{V_i} = \left(\frac{1 - H\left(z_{i-1}^{\star};t_i\right)}{H\left(z_i^{\star};t_i\right) - H\left(z_{i-1}^{\star};t_i\right)}\right) FE\left(z_i^{\star}\right) - \left(\frac{1 - H\left(z_i^{\star};t_{i+1}\right)}{H\left(z_i^{\star};t_i\right) - H\left(z_{i-1}^{\star};t_i\right)}\right) FE\left(z_{i+1}^{\star}\right). \tag{66}$$

This expression makes it clear that, in principle, it is possible for the fiscal externality of this type of joint reform to be positive.

**Theorem 6** (Test for Pareto Efficiency (Multibracket))**.** *Under assumptions 1, 2, and 3, then a piecewise linear tax schedule which is Pareto efficient must satisfy*

$$FE\left(z_i^{\star}\right) \geq -1 \tag{67}$$

*for every kink point $z_i^{\star}$, and*

$$FE^{V_i} \geq -1 \tag{68}$$

*for every virtual income level $V_i$.*

*Moreover, when a tax schedule satisfies both of these conditions (and the relevant second-order conditions in edge cases where $FE\left(z_i^{\star}\right) = -1$ or $FE^{V_i} = -1$) no Pareto-improvement can be obtained through any infinitesimal tax reform that jointly moves any set of kink points.*

The second part of this theorem is analogous to the result presented in Bierbrauer et al. [2020], which shows that if there are no Pareto-improving "single-bracket" or "two-bracket" reforms to marginal tax rates, then there are no Pareto-improving reforms.[41] Theorem 6 states that if the are no Pareto-improving moves of a single kink point, and no Pareto-improving joint reforms of the kind described above, then there are no other Pareto-improvements which can be obtained through jointly moving kink points (infinitesimally).

*Sketch of a proof:* This result can be derived by first demonstrating that the fiscal externality of any joint movement of a collection of kink points can be written as a linear combination of the fiscal externalities of moving individual kink points (as equation 66 does for the particular joint reform discussed above). It is not difficult to then show that any the existence of a Pareto-improving joint movement of three kink points would imply that either condition 67 or condition 68 must fail.

[41] Note, in Bierbrauer et al. [2020], "single-bracket" and "two-bracket" reforms do not refer to changes in tax rates within different tax brackets in a piecewise linear tax schedule but rather describe local reforms at a single point in the tax schedule which can be intuitively described as "single-bracket" and "two-bracket".