Malas Notches*

Ben Lockwood†

First version: 4 July 2016
This version: 22 November 2019

Abstract

This paper shows that the sufficient statistic approach to the welfare properties of income (and other) taxes does not easily extend to tax systems with notches, because with notches, changes in bunching induced by changes in tax rates have a first-order effect on tax revenues. In an income tax setting, we show that the marginal excess burden (MEB) of a change in the top rate of tax is given by the Feldstein (1999) formula for the MEB of a proportional tax, plus a correction term. This formula applies even if there is tax evasion. These correction terms cannot be calculated just from knowledge of the elasticity of taxable income and quantitatively, they can be large. An application to VAT is discussed; with a calibration to UK data, the MEB of the VAT is roughly three times what is would be if VAT was simply a proportional tax.

JEL Classification: H20,H21,H31

Keywords: tax kink, tax notch, excess burden, sufficient statistic

*I would like to thank Omiros Kouvavas for outstanding research assistance, and Miguel Almunia, Steve Bond, Thomas Brosy, Michael Devereux, Michael Keen, and Joel Slemrod, participants at the 2017 ASSA conference, and an editor and two referees for helpful comments. I also acknowledge financial support from the ESRC under grant ES/L000016/1.

†CBT, CEPR and Department of Economics, University of Warwick, Coventry CV4 7AL, England; Email: B.Lockwood@warwick.ac.uk

1 Introduction

In a recent survey, Chetty (2009b) argues that an important new development in public economics is the so-called sufficient statistic approach, which "derives formulas for the welfare consequences of policies that are functions of high-level elasticities rather than deep primitives" (Chetty (2009b), p 451). In turn, this means that to assess the welfare properties of these policies, only these elasticities, rather than fully structural models, need to be estimated.[1]

The sufficient statistic approach originated in a seminal paper by Feldstein (1999), who showed that the marginal excess burden (MEB) of a proportional income tax only depends on the behavioral responses to the tax via a sufficient statistic, the elasticity of taxable income (ETI). The ETI summarises the response of a given household to changes in the tax rate, although these changes can be at several margins (hours, effort, etc.) Feldstein's paper has given rise to a large literature devoted to obtaining empirical estimates of the ETI (Gruber and Saez (2002), Saez et al. (2012), Kleven and Schultz (2014), Weber (2014)).

Subsequently, Saez (2001) and Saez et al. (2012) showed that the Feldstein formula for the MEB could be extended to the top rate of tax in a progressive piece-wise linear income tax system, and they also established formulae for the revenue and welfare-maximizing rate of tax. These formulae also have the sufficient statistic feature; specifically, they depend only on the ETI, a statistic of the income distribution, which is constant if the top tail of the income distribution is Pareto[2], and possibly a welfare weight.

In this paper, we ask the question as to whether these sufficient statistic properties of key formulae also extend to tax systems with notches. Generally, a tax notch occurs when there is a discontinuous change in the tax liability as the tax base varies (Slemrod (2013), Kleven (2016)).

In practice, we do see notches in several major kinds of taxes, and these are being increasingly studied in the empirical literature. For example, in Pakistan, there are notches of up to 5% in the personal income tax (Kleven and Waseem (2013)), and in Ireland, an emergency income levy after the financial crisis had a notch of up to 4% (Hargaden (2015))[3]. There are small notches in the federal income tax in the US, and larger notches induced by income-dependent entitlement to tax credits (Slemrod (2013)). In Germany, there is a large notch in income tax generated by the Mini-Job program (Tazhitdinova

[1] Chetty (2009a) also argues that this sufficient statistic approach is also valuable in several other contexts, such as evaluating the welfare gain from social insurance programs, and the welfare effects of changes in taxes with optimization frictions.

[2] The formula is that the marginal excess burden equals $\frac{tea}{1-t-te}$, where t is the rate of tax, e is the personal elasticity of taxable income with respect to the net of tax rate $1-t$, and a is the Pareto parameter.

[3] From Table 1 of Hargaden (2015), in 2010, earnings of above 26000 Euro incurred a charge of 1040 Euro.

(2018)). [4]

Notches also exist in other major taxes. For example, notches are, or were until recently, present in housing transactions taxes in the UK and the US (Best and Kleven (2013), Kopczuk and Munroe (2015)). They also arise in the corporate income tax in Costa Rica (Bachas and Soto (2015)). Slemrod (2013) notes that there are many examples of commodity tax notches, where a marginal change in some characteristic can change the product classification so as to produce a discrete change in the tax liability.[5] Finally, as argued by Liu and Lockwood (2015), a VAT threshold can be thought of as a tax notch; a firm's VAT liability changes discontinuously when its sales go over the registration threshold. Indeed, given the importance and near-ubiquity of VAT, this is in fact the most important example of a tax notch.

We first study notches in the income tax setting of Saez (2010) and others, where households differ in ability or taste so that the disutility of generating taxable income varies across households. For simplicity, we assume a two-bracket tax i.e. a tax with a lower rate below a threshold, and a higher rate above. In this setting, our first contribution is to derive an exact formula for the marginal excess burden (MEB) of the higher rate of tax, This formula similar to Feldstein (1999)'s formula for the MEB of a proportional income tax, but includes a correction factor that captures the effect of the bunching response to an increase in the top rate tax on tax revenue.

The bunching response measures the change in the *number* of households bunching at the threshold to avoid paying the top rate of tax, and is a property of the distribution of households. In what follows, to make it clear that this is a property of the distribution, we will call it the *aggregate bunching response.* It is thus distinct from the change in taxable income of a *particular* household induced by a change in the tax rate. The latter is measured by the elasticity of taxable income, and in what follows, we will call the second kind of response the *individual* response, as it pertains to a particular individual or household.[6]

Our main point is that with a notch, unlike the case of a kink, the aggregate bunching response affects tax revenue because with a notch, the tax schedule is discontinuous at the threshold. Specifically, an increase in the top rate of tax increases bunching, which - due to the notch - lowers tax revenue, and thus raises the MEB. Moreover, this correction factor to the Feldstein formula, denoted C, *cannot be expressed as a simple function of the usual sufficient statistics* i.e. the ETI and the Pareto parameter of the upper tail of

[4] This is aimed at increasing the labor supply of low-income individuals: earnings below the mini-job threshold, are exempt from income tax and the employee portion of social security taxes, while earnings above the threshold are not.

[5] For example, in the US, the Gas Guzzler Tax, under which high-performance cars are subject upon initial sale to a per-vehicle tax that is higher, the lower is the fuel economy of the car.

[6] The individual response could include responses in hours or intensity of work effort, usually known as intensive-margin responses, and the decision whether to work or not, usually known as the extensive-margin response. Thus, our distinction between the aggregate bunching and individual response is quite different to the intensive vs. extensive distinction.

the income distribution. It does depend on these variables, but it also depends on the lower rate of tax, the position of the notch, and a counterfactual, i.e. the earnings that the individual at the top of the interval (the top buncher) would choose if faced with the higher rate of tax. So, the sufficient statistic approach seems to break down with tax notches.

However, all is not lost. We show how the counterfactual earnings of the top buncher can be computed theoretically, using the indifference condition that the top buncher is indifferent between bunching and being above the notch. Alternatively, in any empirical study of bunching, it can be computed empirically, using the estimate of excess mass at the notch (the parameter B in Kleven and Waseem (2013)). Thus, this paper is the first to show how bunching estimates at notches can be used to make welfare calculations.

Of course, if the correction factor turns out to be small, the Feldstein formula still provides a good approximation to the MEB. Our third contribution is to investigate whether this is the case. Calibrations show that the percentage error from using the Feldstein formula for the MEB can be very large. At baseline values, the marginal excess burden is underestimated by a factor of six. So, the conclusion is that at least in the income tax setting, the sufficient statistic approach is not practical.

We then turn to apply our approach to the VAT, which is the most empirically important example of a tax notch. We present a simple model of small traders who differ in productivity, and are subject to VAT at rate t above a threshold level of sales. We show that this model is formally equivalent to our income tax model, in the sense that registered firms above the threshold face an effective rate of VAT t_R on value-added, and firms below the threshold face a lower but positive effective rate t_N.[7]

We then show that the MEB of an increase in the statutory rate of VAT is given by the Feldstein formula for a proportional tax plus a correction factor as in the income tax case. However, the details of the correction factor are more complex, because an increase in the statutory rate t increases *both* the effective rates t_R, t_N. A calibration of the model shows that the proportional tax formula for the MEB of the VAT underestimates the true MEB by a factor of up to three.

Finally, it should be noted that in this paper, we take all parameters of the tax system, including the notch, as given, and only vary the top rate of tax. A broader question, to be addressed in future work, is whether a notch can ever be part of an optimal tax system[8].

The remainder of the paper is arranged as follows. After the literature review in

[7]It may seem counter-intuitive that non-registered firms face a positive rate of effective VAT; this is because non-registered firms cannot claim back VAT on inputs, so-called "embedded" VAT.

[8]In the standard Mirrlees framework, where the tax is fully non-linear, this not the case; the optimal tax schedule is always continuous in income. However, where skills are continuously distributed, and the government is restricted to a finite number of tax rates, the answer to this question is less obvious. In fact, Blinder and Rosen (1985) note in the context of subsidies for charitable giving, with heterogeneous tastes, sometimes a notch can improve on a linear subsidy in the sense of having a lower total efficiency cost, defined as the sum of excess burden and the cost of raising the revenue. However, they do not undertake a full social-welfare-maximizing exercise.

Section 2, in Section 3, we set up the model. Section 4 has the main analytical results for the income tax, Section 4.3 has an extension to tax evasion, and Section 5 the simulations. Section 6 deals with the extension to the VAT, and Section 7 concludes.

2 Related Literature

This paper speaks to a number of related literatures. First, it is already known that due to externalities of one kind or another, the sufficient statistic approach has its limitations. Saez et al. (2012) give the examples of deductibility from income tax of charitable giving and mortgage interest payments for residential housing. In these cases, an increase in the marginal rate of tax will boost charity income and home ownership respectively, which may be valuable objectives in themselves. Saez et al. (2012) call these classical externalities[9].

Fiscal externalities, where the actions of the household generate additional revenue for the government and thus benefits other households, can also cause the sufficient statistic approach to fail, or at least require adjustment, but in these cases a simple change to the formula is sometimes possible. The analysis of income tax evasion of Chetty (2009b) is a case in point[10]. As Gillitzer and Slemrod (2016) show, in this case the standard formula for the marginal efficiency cost of funds can be adjusted in the same way it must be adjusted for any fiscal externality, i.e. whenever a change in tax rates induces taxpayers to shift income to another tax. Our results are rather different to these cases of both classical and fiscal externalities. In our setting, there is no fiscal or other externality- rather, the sufficient statistic approach fails because the aggregate bunching response has a first-order effect on tax revenue. Indeed, in Section 4.3 below, we show our main qualitative results continue to apply in the presence of evasion, which makes the point that our argument is distinct from an externality one.

A second related literature is on VAT. Here, there are two distinct sets of related papers. First, there is a growing literature on the effect of VAT thresholds on firm behavior. Theoretical contributions include Keen and Mintz (2004), Kanbur and Keen (2014) and Liu and Lockwood (2015), and empirical studies include Liu and Lockwood (2015) and Harju et al. (2016). The theoretical work of Kanbur, Keen and Mintz focuses on the optimal threshold of the VAT, holding the rate of tax fixed, and is thus complementary to this paper, which characterizes the MEB of an increase in the rate, holding the threshold fixed. In fact, we effectively ask the question of whether it is legitimate to ignore the

[9]See Doerrenberg et al. (2015) for a more formal statement of this argument, and estimates of how deductions respond to tax rate changes for the case of Germany.

[10]Chetty shows that when the household can evade the personal income tax at a cost, if that cost is a pure transfer payment i.e. a fine times a probability of detection, there is effectively a positive fiscal externality of evasion - it generates additional revenue for the government and thus benefit for all households. In this case, as we might expect, we see that the elasticity of taxable income over-estimates the excess burden of the tax.

threshold altogether when calculating the MEB of the VAT.

Therefore, our paper relates to a literature on the marginal excess burden of indirect taxes, including VAT (e.g. Ballard et al. (1985), Rutherford and Paltsev (1999)). In these papers, when the marginal excess burden of VAT is calculated, it is always assumed that the VAT is a proportional tax i.e. the VAT threshold is ignored. This paper shows that this simplifying assumption yields seriously biased estimates.

A third related literature is that on the MEB and welfare-maximizing taxes with kinks in the tax schedule. Here, we make a small contribution as a by-product of our main focus, which is on notches. In the case of kinks, it generally understood that the marginal excess burden of the top rate of income tax, and the welfare-maximizing top rate depends via simple formulae, only on the elasticity of the ETI, and the Pareto statistic of the income distribution. However, there seems to be some confusion about the conditions required for this result. Saez et al. (2012) suggest that what is required is that assumption that "behavioral responses take place only along the intensive margin", or more precisely that the aggregate bunching response of an increase in the top rate of tax is of second order relative to the extensive margin response.[11] This assumption is very strong, as even with a kink, there is always a bunching response. Our Proposition 1 below shows that this assumption is *not* necessary, because no matter what the size of the bunching response, the response has no effect on tax revenue, to first order, as the tax schedule is continuous. All that is required is that the distribution of taxpayer types is continuous, a standard assumption.

3 The Model and Preliminary Results

3.1 Set-Up

We follow Saez (2010) in our set-up. There are individual taxpayers indexed by a skill or taste parameter $n \in [\underline{n}, \overline{n}]$, assumed continuously distributed in the population with distribution $H(n)$ and density $h(n)$. A type n individual has preferences over consumption c and taxable income z of the form

$$u(c, z; n) = c - \psi(z; n) \tag{1}$$

where $\psi(z; n)$ is the disutility of earning income z. So, as utility is linear in c, we are assuming away income effects. We also assume:

A1. $\psi_z > 0, \psi_{zz} > 0,\ \psi_n, \psi_{nz} < 0.$

[11] Specifically, they say the following. "The change dt could induce a small fraction dN of the N taxpayers to leave (or join if $dt < 0$) the top bracket. As long as behavioral responses take place only along the intensive margin, each individual response is proportional to dt so that the total revenue effect of such responses is second order ($dN.dt$) and hence can be ignored in our derivation."

A1 says that the cost of generating taxable income is strictly increasing and strictly concave in z. It also allows us to interpret a higher n as a higher skill level (i.e. higher wage), or a lower taste for leisure. In particular, the higher n, the lower the total and marginal disutility of generating a given amount of taxable income. Assumption A1 is satisfied for example, by the iso-elastic specification of Saez (2010):

$$\psi(z;n) = \frac{n}{1+\frac{1}{e}} \left(\frac{z}{n}\right)^{1+\frac{1}{e}} \tag{2}$$

The budget constraint is $c = z - T(z)$, where $T(.)$ is the tax function. So, a household's utility over z is $u(z;n) = z - T(z) - \psi(z;n)$.

Finally, for future reference, define the optimal taxable income at tax rate t for a type n taxpayer to be;

$$z(1-t,n) \equiv \arg\max_{z\geq 0}\{(1-t)z - \psi(z;n)\}$$

Generally, Assumption A1 does not imply that $z(1-t,n) > 0$, so we allow for corner solutions with zero earnings i.e. where the household does not work. However, in the iso-elastic case (2), there will always be an interior solution, as the marginal cost of z goes to zero with z. Note from A1 that if there is an interior solution $z(1-t,n) > 0$, then $z_{1-t}, z_n > 0$, where subscripts denote derivatives. So, z_{1-t} is the response of taxable income to the net-of-tax rate. Following the terminology introduced in the introduction, we call this the *individual* response to the tax.

3.2 Kinks and Notches

For simplicity, we focus on a two-bracket tax, although our arguments apply straightforwardly to the case of the highest tax in a piecewise-linear tax system with any number of brackets. We will assume that the tax system is progressive; that is, the tax rate on incomes in the higher income bracket is strictly greater than the tax on incomes in the lower income bracket.

So, with a two-bracket tax, for a kink, the tax function is

$$T_K(z) = \begin{cases} t_L z, & z \leq z_0 \\ t_L z_0 + t_H(z - z_0), & z > z_0 \end{cases} \tag{3}$$

for $z_0 > 0$, $t_H > t_L \geq 0$; that is, all income below the kink point z_0 is taxed at the lower rate t_L, and all income in excess of the kink is taxed at the higher rate. For a notch, the tax function is

$$T_N(z) = \begin{cases} t_L z, & z \leq z_0 \\ t_H z, & z > z_0 \end{cases} \tag{4}$$

with $t_H > t_L \geq 0$. That is, when taxable income is below z_0, a tax at rate t_L is paid on all income, but when z is above z_0, a tax at rate t_H is paid on *all* income.

Note here that we are studying what Kleven and Waseem (2013) call a proportional tax notch. The more general case is where there is also a pure notch, where a lump-sum tax or subsidy is also paid when earnings exceed z_0. We choose to focus on the proportional notch partly for simplicity, and partly because most of the empirical cases of notches discussed in the introduction are of this type.

3.3 Bunching

With either a kink or a notch, all types in an interval $n \in [n_L, n_H]$ will bunch at taxable income z_0. In both cases, the lowest type who bunches is the one who is just willing to earn taxable income z_0 at the lower tax rate. So, n_L is defined by the condition

$$z(1 - t_L, n_L) = z_0 \tag{5}$$

With a kink, the highest type who bunches, n_H, is defined by the condition that the optimal choice of taxable income at tax t_H is just z_0 i.e.

$$z(1 - t_H; n_H) = z_0 \tag{6}$$

With a notch, n_H is defined by the condition that the n_H type must be indifferent between staying at the notch and paying tax t_L, and choosing z optimally, and paying t_H on *all income*. To write this indifference condition, we first define the indirect utility function

$$v(1 - t; n) \equiv \max_{z \geq 0} \{(1 - t)z - \psi(z; n)\}$$

Then, the condition defining n_H can be written:

$$(1 - t_L)z_0 - \psi(z_0; n_H) = v(1 - t_H; n_H) \tag{7}$$

The left-hand side of (7) is utility when taxable income is constrained to be at the notch value z_0. Note that this indifference condition implies $z(1 - t_H, n_H) > z_0$, because if $z(1 - t_H, n_H) < z_0$, the n_H-type could choose z optimally *and* stay below the notch.

3.4 The Aggregate Bunching Response

Here, we study the effect of a change in t_H on the mass of individuals who bunch i.e. on the size of the interval $[n_L, n_H]$. Note first from (5) that n_L is unaffected by t_H for both a kink and a notch. Next, in the kink case, we can calculate from (6) that

$$\frac{\partial n_H}{\partial t_H} = \frac{z_{1-t_H}}{z_n} > 0 \tag{8}$$

So, we have a aggregate bunching response to an increase in t_H: i.e. an increase in the tax rate above the kink makes going above the kink less attractive, and so more people bunch below the kink.

In the notch case, note that $v_t = -z$, where v_t is the derivative of v with respect to t. Then, using this fact and the implicit function rule, we can calculate from (7) that

$$\frac{\partial n_H}{\partial t_H} = \frac{z(1-t_H, n_H)}{\psi_n(z_0; n_H) - \psi_n(z(1-t_H, n_H); n_H)} \tag{9}$$

Also, as $\psi_{nz}(z;n) < 0$ and $z(1-t_H, n_H) > z_0$, we see that the denominator of (9) is positive, and consequently from (9):

$$\frac{\partial n_H}{\partial t_H} > 0 \tag{10}$$

So, again we see that there is a aggregate bunching response to a change in t_H; an increase in the tax rate above the notch makes going above the notch less attractive, and so more people bunch at the notch.

4 Main Results

4.1 The Effect of the Aggregate Bunching Response on Tax Revenue

Here, we establish a key result that the effects of the aggregate bunching response on tax revenue with a kink and a notch are qualitatively different, being zero and negative respectively. With a kink, revenue can be written

$$R = t_L \int_{\underline{n}}^{n_L} z(1-t_L; n)h(n)dn + t_L(1-H(n_L))z_0 + t_H \int_{n_H}^{\overline{n}} (z(1-t_H; n) - z_0)h(n)dn \tag{11}$$

Note from the second term in (11) that all households with $n \geq n_L$ pay tax at the lower rate on the first z_0 of earnings.

So, in the kink case, the aggregate bunching effect on tax revenue i.e. the effect of a change in t_H on R via a change in n_H is, from (6) and (11):

$$\frac{\partial R}{\partial n_H} = -t_H(z(1-t_H; n_H) - z_0)h(n_H) = 0 \tag{12}$$

So, overall, with a kink, the effect of the aggregate bunching response on tax revenue is zero. This is simply due to the fact that a kinked tax schedule is continuous in z.

With a notch, revenue is

$$R = t_L \int_{\underline{n}}^{n_L} z(1-t_L; n)h(n)dn + t_L(H(n_H) - H(n_L))z_0 + t_H \int_{n_H}^{\overline{n}} z(1-t_H; n)h(n)dn \tag{13}$$

Comparing this to (11), we see a key difference. Because the higher rate applies to *all* income for those earning above z_0, the threshold z_0 no longer enters into the the tax base for t_H, and so the size of the term on z_0 in the tax base for falls from $1 - H(n_L)$ to $H(n_H) - H(n_L)$, reflecting the fact that now only individuals below n_H pay any tax at the lower rate.

Note from (13) that;

$$\frac{\partial R}{\partial n_H} = (t_L z_0 - t_H z(1 - t_H; n_H))h(n_H) < 0 \tag{14}$$

This is strictly negative as $t_H > t_L$, $z(1 - t_H; n_H) > z_0$. So, in contrast to the kink case, the aggregate bunching effect on tax revenue R from an increase in t_H is negative, as $\frac{\partial n_H}{\partial t_H} > 0$ from (10). This is because a small increase in n_H has two effects on revenue that are both negative. First, there is a discontinuity in the tax *base*; the earnings of these who now locate at the notch fall discontinuously from $z(1 - t_H; n_H)$ to z_0. Second, there is a discontinuity in the tax *rate* applying to that base; all these earnings are taxed at a lower rate, t_L rather than t_H.

So, we conclude:

Proposition 1. *The effect of the bunching response on tax revenue is zero for a kink, but strictly negative for a notch.*

This result is the key one that drives the rest of the paper. Proposition 1 also helps to clarify some confusion in the literature. As already noted, Saez et al. (2012) argue that for sufficient statistic formulae to apply in the kink case, what is required is that assumption that "behavioral responses take place only along the intensive margin", or more precisely that the aggregate bunching response of an increase in the top rate of tax is of second order relative to the individual response. Proposition 1 shows that this assumption is not required, because no matter how large is $\frac{\partial n_H}{\partial t_H}$, $\frac{\partial R}{\partial n_H} = 0$ in the kink case.

4.2 The Marginal Excess Burden

Here, we derive a formula for the marginal excess burden (MEB) of t_H when there is a notch and show that it can be written as the MEB of a proportional tax plus a correction factor. To define the MEB, note that due to quasi-linearity, the natural measure of welfare is the integral of indirect utilities, say W, plus revenue R, which is assumed to be redistributed as a lump-sum back to households when calculating the MEB. So,

$$MEB = -\frac{d(W + R)/dt_H}{dR/dt_H} \tag{15}$$

The minus sign ensures that the marginal excess burden is measured as a positive number.

Generally, whether there is a kink or a notch, a simple envelope argument tells us that a change dt_H only has a direct effect on W; all indirect effects, via individual or aggregate bunching responses are zero, as households are optimising. In turn, due to the assumption of a quasi-linear utility function, this direct effect is simply the total increase in tax paid at the higher rate i.e. dt_H times the base of the higher rate of tax. That is, mathematically:

$$\frac{dW}{dt_H} = -\int_{n_H}^{\overline{n}} z(1-t_H;n)h(n)dn = -B_H \tag{16}$$

where B_H is the base of the higher rate of tax. Plugging (16) back into the MEB formula (15), dividing through by B_H, and rearranging, we get

$$MEB = \frac{1-E/F}{E/F}, \quad E = \frac{t_H}{R}\frac{dR}{dt_H}, \quad F = \frac{t_H B_H}{R} \tag{17}$$

So, we see that we can *always* write the MEB in terms of an observable, F, the share of revenue raised by the top rate of tax, and E, the aggregate elasticity of revenue with respect to the top rate of tax. The problem with this characterization of the MEB is twofold.

First, it is not easy to credibly estimate E, as one must typically rely on cross-country data, and in that case, exogenous variation in the tax t is hard to find. For example, if the UK raised its top rate of tax t_H from 40% to 50% - as actually happened 2010 - and revenue R rose by 5%, we cannot infer that the elasticity is 0.5 as other things are not equal. Moreover, the only plausible control group would be are other similar countries, which are small in number, have their own changes in taxes, and so on.

Second, and more fundamentally, E will depend on both individual household responses to the top rate of tax t_H, and the distribution of income, and and we wish to know how both these factors determine E. For the case of a kink, such a formula has been provided by Saez (2001), and is given in (23) below. It is the main objective of this paper to develop a similar formula for the case of a notch, and explore its implications.

The first step in this exercise is to calculate the overall effect of an increase in t_H on tax revenue R via the different channels. From (13), we have:

$$\frac{dR}{dt_H} = B_H + \underbrace{t_H \left.\frac{\partial B_H}{\partial t_H}\right|_{n_H \text{ const}}}_{\text{individual}} + \underbrace{\frac{\partial R}{\partial n_H}\frac{\partial n_H}{\partial t_H}}_{\text{aggregate bunching}} \tag{18}$$

As before, B_H is the base in which the higher rate of tax is levied.

So, (18) is composed of three terms, the mechanical effect B_H, and two behavioral effects on tax revenue, the individual and aggregate bunching effects. The individual effect on tax revenue is standard; it describes how the tax base changes because of changes in earnings, conditional on the taxpayer staying the same tax bracket.

So, plugging (16), (18) back into the MEB formula (15), dividing through by B_H, multiplying by $1-t_H$, and noting that holding n_H constant, $\frac{\partial B_H}{\partial(1-t_H)} = -\frac{\partial B_H}{\partial t_H}$, we can establish the following result.

Proposition 2. *with a tax notch, the marginal excess burden of the top rate of income tax is*

$$MEB = \frac{t_H \bar{e} + C}{1 - t_H(1+\bar{e}) - C}, \; C = -\frac{1-t_H}{B_H}\frac{\partial R}{\partial n_H}\frac{\partial n_H}{\partial t_H} \tag{19}$$

where

$$\bar{e} = \frac{1-t_H}{B_H}\frac{\partial B_H}{\partial(1-t_H)}\bigg|_{n_H \text{ const}} = \frac{1-t_H}{B_H}\int_{n_H}^{\overline{n}} \frac{\partial z(1-t_H;n)}{\partial(1-t_H)} h(n)dn \tag{20}$$

Here, $\bar{e}$ is the elasticity of the tax base B_H with respect to the net of tax rate $1-t_H$, holding n_H constant, and so is just the average ETI. Also, C is a correction factor, which captures the effect of a changing n_H, the aggregate bunching response, on the MEB, via its effect on revenue.

Some comments are appropriate at this point. First, (19) is the formula for the marginal excess burden of a *proportional* income tax, as shown by Feldstein (1999), plus a correction factor C. This is intuitive; all households above n_H are paying tax at rate t_H on all their income, so for these households, t_H is indeed a proportional tax. So, as already remarked, the correction factor C just captures the effect of a changing n_H, the aggregate bunching response, on the MEB, via its effect on revenue.

As a next step, we would like to be able to investigate in more detail to what extent the correction factor C is quantitatively important. To do this, we make two standard assumptions. The first is that the disutility of income is iso-elastic i.e. as in (2). In that case, all individuals have the same ETI, namely e, and so $\bar{e} = e$, a constant independent of n_H. The second is that the distribution of n is Pareto above n_H. We can then prove: [12]

Proposition 3. *Assume iso-elastic utility (2), and that the distribution of n is Pareto, with shape and scale parameters $a, \underline{n}$. Then, the MEB with a notch is*

$$MEB = \frac{t_H e + C}{1 - t_H(1+e) - C}, \tag{21}$$

where

$$C = \frac{(t_H - t_L z_0/\tilde{z}_H)(a-1)(1+e)}{1 - \left(\frac{z_0}{\tilde{z}_H}\right)^{(1+e)/e}} > 0. \tag{22}$$

Moreover, in (22), $\tilde{z}_H = n_H(1-t_H)^e$ and n_H *is defined by (7).*

This result enables us to compare precisely how the MEB compares to the MEB in a kinked tax system. As shown for example, by Saez (2001), under our assumptions, the

[12]This and subsequent Propositions are proved in the Appendix.

latter is

$$MEB_K = \frac{t_H ea}{1 - t_H(1 + ea)} \tag{23}$$

Clearly, MEB_K depends only on simple sufficient statistics; other than the tax rate t_H, it depends only on e, the individual elasticity of taxable income, and a, the shape parameter of the income distribution.

By contrast, from (22), it is clear that C is a more complex object. It depends not only on sufficient statistics e, a, and the top rate of tax, t_H, but also on other parameters of the tax system t_L, z_0, and on $\tilde{z}_H$, which is the unconstrained earnings of the type n_H, given that they face the higher rate of tax.

So, there are two ways of solving for C. One is simply to compute C using the formulae (22), (7), choosing calibrated values for e, a, z_0, and that is what we do in this paper. Alternatively, as shown by Kleven and Waseem (2013), in any empirical study of a notch, the earnings $n_H(1 - t_L)^e$ can be estimated. Specifically, $n_H(1 - t_L)^e$ is simply $z^* + \Delta z^*$ in the notation of their paper, where z^* is the earnings notch and as explained there, $\Delta z^*/z^*$ can be estimated from excess bunching at the notch. Given this, $\tilde{z}_H$ can be recovered simply by multiplying $z^* + \Delta z^*$ by $(1 - t_H)^e/(1 - t_L)^e$, using the empirical estimate of e.

4.3 Tax Evasion

Before turning to simulations with a calibrated version of our model, we consider how our results extend to the case where the taxpayer can evade, or shelter, some of her income at a resource cost. In this section, we briefly sketch the argument; the details are given in the Online Appendix.

We generalize our framework using Chetty (2009b). We now interpret z as *reported* income, and we denote by s income that is sheltered from the government. A type n individual now has preferences

$$u(c, s; n) = c - g(s) - \psi(z + s; n) \tag{24}$$

Note two changes from (1). First, there is a cost of sheltering income from the tax authorities, captured by g; we assume that $g', g'' > 0$. As Chetty (2009b) says, this could reflect the loss in profits from transacting in cash instead of electronic payments or the cost of choosing a distorted consumption bundle to avoid taxes. Second, the disutility of income depends on the sum of reported and sheltered income i.e. $z + s$.

The budget constraint is

$$c = z + s - T(z) - a(s) \tag{25}$$

where as in Chetty (2009b), $a(s)$ is the expected cost to the household of audit, which is assumed to be increasing and weakly convex in s. This captures any fines paid if s is detected by the tax authorities, times the probability of detection.[13] Note that the tax paid depends only on reported income. The, household maximizes (24) with respect to z, s subject to (25), giving rise to choice of reported income $z(1-t)$.

Then, the behavior of the household faced with a kink or a notch is qualitatively the same as before. That is, under either type of tax schedule, households in the bunching interval $[n_L, n_H]$ keep z just at the threshold. In the case of a kink, n_L, n_H are characterized by (5), (6) as before. In the case of a notch, (7) is modified to allow for the endogenous choice of sheltered income s. Given this, it is still the case that the effect on revenue R of a change in n_H is zero in the kink case and negative in the notch case, as this simply follows from the (dis-)continuity of R in the kink (notch) case. So, Proposition 1 continues to hold.

Moreover, as shown in the Online Appendix, in the special case where there is no audit cost of evasion, i.e. $a \equiv 0$, Proposition 2 continues to hold. In the more realistic case where there is an audit cost, the MEB is equal to the MEB of a proportional tax plus *two* correction factors, one for the notch C as before, and one offsetting negative term capturing the fact that the audit cost is a transfer and thus lowers the MEB of the tax. As the first is positive and the second is negative, they have offsetting effects on the MEB.

5 Simulations

We have seen that the MEB of an increase in t_H is given by the corresponding formula for a proportional tax t_H plus a correction factor, C. Moreover, the MEB formula for a proportional tax is very simple, depending only on the intensive-margin elasticity e, and thus can easily be calculated.

So, a key question is whether we can get a good approximation to MEB by setting $C = 0$ i.e. treating t_H as a proportional tax. In this section, we investigate whether the MEB, calculated assuming that t_H is a proportional tax, is a good approximation to the true MEB.

To do this, we need to calibrate the model. In particular, we require values for e, a, t_H, t_L, and z_0. Our baseline parameter values are chosen as follows. Following Piketty and Saez (2013), we set $a = 1.5$, and following Saez et al. (2012) and Kleven and Schultz (2014), we set $e = 0.25$. Regarding the tax rates, we first set $t_L = 0.2$, which is broadly in line with the average income and payroll tax paid by US households[14]. It is also the basic rate of income tax in the UK. For the notch, we use the fact that notches in personal

[13] We simplify slightly by making a independent of t. In practice, audit fines are often proportional to taxes owed, and this generalization is simple to make; there will be an additional term a_t in the integral in (N.4).

[14] "Overview Of The Federal Tax System As In Effect For 2015", Joint Committee on Taxation, Congress of the United States.

income tax, where they exist, are small. For example, Kleven and Waseem (2013) show that in the Pakistani income tax, the notch ranges between 2 and 5 percentage points. So, we will take our baseline notch $t_H - t_L = \Delta t = 0.03$.

To choose $\underline{n}, z_0$ we assume that only the top 20% of the population pay a higher rate of income tax, roughly the proportion in the UK. Define n_0 to be the skill level corresponding to taxable income just at the notch i.e. $n_0(1-t_L)^e = z_0$. This requires that 80% of the population have skills below n_0 i.e. $H(n_0) = 1 - \left(\frac{\underline{n}}{n_0}\right)^{\alpha} = 0.8$, or $\frac{\underline{n}}{n_0} = (0.2)^{1/1.5} = 0.342$. Given that only the ratio $\frac{\underline{n}}{n_0}$ is determined, we set $\underline{n} = 1$, so $n_0 = 2.924$. But then $z_0 = 2.924(0.8)^{0.25} = 2.168$.

Finally, from (22), we need a value for n_H. Under the assumption (2), the indifference condition (7) reduces to

$$e(n_H)^{-1/e}(z_0)^{1+\frac{1}{e}} + n_H(1-t_H)^{1+e} - (1-t_L)z_0(1+e) = 0 \tag{26}$$

Equation (26) has two roots, and we take the larger root to ensure that $n_H(1-t_L)^e > z_0$. Finally, parameter values are chosen so that the denominator in (21) is positive, which is equivalent to $dR/dt_H > 0$ i.e. that the tax rate is on the left side of the Laffer curve. This requires simply that the notch is greater than 0.0015.[15]

Figures 1, 2 show both the true MEB, as given by (21), and the approximation, treating t_H as a proportional tax i.e. setting $C = 0$ in (21). The former is denoted by MEB in the Figures, and the latter by MEB_A.

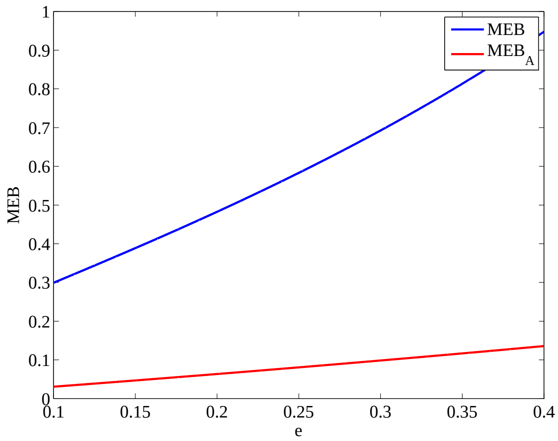

Figure 1: MEB as e varies

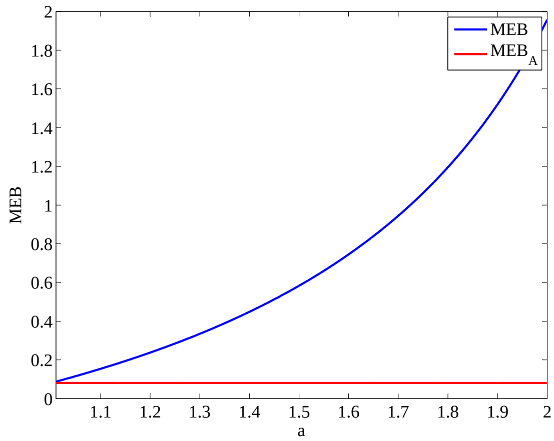

Figure 2: MEB as a varies

The error in using MEB_A at the baseline values can be read off from Figure 1, setting $e = 0.25$. It can be seen that true MEB is about 0.6, whereas the approximation is about 0.1. So, the error in using the proportional formula is about a factor of six. Figure 1 also shows that MEB is increasing in e, at a faster rate than MEB_A, so when $e = 0.4$ for

[15]For the denominator in (21) to be positive, we require $1 - t_H(1+e) > C$, which is satisfied for $t_H - t_L > 0.0015$.

example, the error in using MEB_A is almost an order of magnitude.

Figure 2 shows that MEB is also increasing in a, the Pareto parameter which measures (inversely) the size of the tail of the income distribution. As MEB_A is independent of a, this means that the the error in using MEB_A is increasing in a.

6 An Application to VAT

As remarked in the introduction, perhaps the most important example of a tax notch is the value-added tax. In this section, we present a simple model of the value-added tax, which is mathematically equivalent to the model developed above. We then calibrate the model using UK data from Liu and Lockwood (2015), to estimate the MEB from the VAT, taking into account bunching at the threshold.

6.1 The Set-Up

Here, we briefly outline the set-up of the model. A detailed exposition is in Appendix A.2. We consider a single industry with a fixed, large number of small traders producing a homogeneous good. Each small trader combines his own labor input with an intermediate input to produce output via a fixed coefficients technology. An implication of this technology is that value-added is proportional to output. As in the income tax model, individual traders are indexed by a skill parameter, and have a disutility of supplying labour of the same iso-elastic form as in (2).

Traders sell to final consumers, who have perfectly elastic demand for the good. This is analogous to the assumption made in the taxable income literature that the wage is fixed, i.e. labor demand is perfectly elastic at a fixed wage. The traders face a VAT system. If the trader is registered, he must charge VAT on sales at rate t, but can claim back VAT paid on the input. The trader must register for VAT if the value of sales exceeds the threshold but can register voluntarily even if this is not the case.

6.2 Trader Payoffs, Effective VAT Rates, and Bunching

Let n measure the skill of the trader. It is shown in the Appendix that the payoff of trader n can be written as a function of value-added z and the VAT system as follows;

$$u(z;n) = z - T(z) - \frac{n}{1+\frac{1}{e}}\left(\frac{z}{n}\right)^{1+\frac{1}{e}} \tag{27}$$

Here, $T(z)$ is the amount of VAT paid by the trader. Moreover, $T(z)$ can be written in terms of *effective* VAT rates:

$$T(z) = \begin{cases} t_N z, & z \leq z_0 \\ t_R z, & z > z_0 \end{cases}, \; t_R = \frac{t}{(1+t)(1-\gamma)}, t_N = \frac{\gamma t}{1-\gamma}. \tag{28}$$

Here, t_N, t_R are the effective VAT rates faced by non-registered and registered traders respectively on the value-added they generate. These depend on the statutory rate of VAT, t, the VAT threshold z_0, expressed as a level of value-added, above which the firm will register, and γ which measures the intensity of the intermediate input in production.[16]

The idea is the following. First, if any intermediate input is used i.e. $\gamma > 0$, the trader is effectively taxed at rate t_N even if his turnover is below the threshold and he does not register, because his input is subject to VAT. This effective rate is increasing in γ and t. Second, if the trader's value-added is above the threshold, he pays a rate t_R, which is also increasing in γ and t. Finally, to rule out voluntary registration, we will assume that registration incurs a higher effective tax rate i.e. $t_R > t_N$ which requires $1 > (1+t)\gamma$.

Then, (27), (28) describe a utility function and a tax schedule as function of value-added z that are mathematically equivalent to the income tax model although, obviously, the economic interpretation of z is different. From this equivalence, we can infer the following. Faced with the tax schedule (28), all traders in the interval $n \in [n_L, n_R]$ will bunch at the VAT threshold z_0. Moreover, $n_L = z_0/(1-t_N)^e$, and n_R solves (7) with t_H, t_L replaced by t_R, t_N.

6.3 The Marginal Excess Burden of the VAT

Here, we use the mathematical equivalence of the VAT and income tax models to move swiftly to a formula for the MEB of the VAT. First, let $z(1-t;n) = (1-t)^e n$ be the value-added chosen by an unconstrained firm facing tax t. Then, it is shown in A.2 that the revenue from the VAT is as in (13), with t_H, t_L replaced by t_R, t_N. Then, the revenue from the VAT can be written compactly as

$$R = t_N B_N + t_R B_R \tag{29}$$

In (29), the bases on which t_N, t_R are levied are the value-added of non-registered and registered traders respectively i.e.

$$B_N = \int_{\underline{n}}^{n_N} (1-t_N)^e n h(n) dn + z_0 (H(n_R) - H(n_N)), \;\; B_R = \int_{n_R}^{\overline{n}} (1-t_R)^e n h(n) dn \tag{30}$$

Now note that a change in the statutory rate t of VAT will change both effective tax rates t_N, t_R unless $\gamma = 0$ i.e. no intermediate inputs are used. This is of course, analogous to a reform that changes both t_H and t_L in the income tax model. So, for the VAT, the formula for the MEB becomes somewhat more complex. To present the formula for the MEB in this case, we need a few more definitions. First, from (30), the intensive-margin

[16] It is shown in the Appendix that $z_0 = (1-\gamma)y_0$, where y_0 is the threshold expressed in the usual way as a value of sales.

elasticities of B_R, B_N with respect to the net-of-tax rate are

$$\frac{1-t_R}{B_R}\frac{\partial B_R}{\partial t_R}\bigg|_{n_R \text{ const}} = e, \quad \frac{1-t_N}{B_N}\frac{\partial B_N}{\partial(1-t_N)}\bigg|_{n_N \text{ const}} = e\phi, \tag{31}$$

where

$$\phi = \frac{\int_{\underline{n}}^{n_N} z(1-t_N;n)h(n)dn}{B_N} < 1 \tag{32}$$

The term ϕ captures a new effect of bunching; with bunching, a mass $H(n_R) - H(n_N)$ of the non-registered firms that are bunching are unresponsive to a change in the rate of VAT, which lowers the aggregate intensive-margin elasticity of the tax base B_N with respect to t_N.[17]

Moreover, recall that an increase in t causes both t_N and t_R to increase, so

$$\theta = \frac{\frac{B_R}{1-t_R}\frac{\partial t_R}{\partial t}}{\frac{B_R}{1-t_R}\frac{\partial t_R}{\partial t} + \frac{B_N}{1-t_N}\frac{\partial t_N}{\partial t}} \tag{33}$$

measures the importance of a change in t_R on revenue relative to a change in t_N. Armed with these new definitions, we can state our result, which is proved in the Appendix.

Proposition 4. *Assume that the distribution of sales is Pareto, with shape and scale parameters $a, \underline{n}$. Then, the MEB of the VAT is*

$$MEB = \frac{\tau\varepsilon + C}{1 - \tau(1+\varepsilon) - C} \tag{34}$$

where

$$\tau = (1-\theta)t_N + \theta t_R, \quad \varepsilon = \frac{(1-\theta)t_N\phi + \theta t_R}{(1-\theta)t_N + \theta t_R}e \tag{35}$$

and finally the correction factor is

$$C = -\frac{\frac{\partial R}{\partial n_R}\left(\frac{\partial n_R}{\partial t_N}\frac{\partial t_N}{\partial t} + \frac{\partial n_R}{\partial t_R}\frac{\partial t_R}{\partial t}\right)}{\frac{B_R}{1-t_R}\frac{\partial t_R}{\partial t} + \frac{B_N}{1-t_N}\frac{\partial t_N}{\partial t}} \tag{36}$$

So, we note now that bunching impacts the calculation of the MEB in two ways. First, as before, there is a correction factor C in (34). The correction factor is more complex than in the income tax case. The reason for the additional complexity is clear from (36); an increase in t now increases both t_R, t_N and in turn, both of these effective taxes affect n_R, the top of the bunching interval, and thus revenue. An explicit formula for C in terms of parameters can be derived as in (22) above; this is done in the Online Appendix.

[17] A similar point has been noted before by Slemrod et al. (1994) and Apps et al. (2014) who consider the design of a two-bracket income tax. Because the tax system they studied was kinked, not notched, the formula for the optimal lower rate of tax depends only on the intensive-margin elasticity, but this elasticity is dampened by the fact that taxpayers at the kink do not adjust their behavior in response to the tax.

In addition, there is a second, new effect of bunching in (35). Bunching dampens the intensive-margin response to a change in t, because at a fixed n_N, n_R, firms in this interval will not adjust their sales in response to a change in t. This is captured by the term ϕ which lowers the intensive margin response from e to ε.

An interesting special case is where the small traders do not use any intermediate input, so i.e. $\gamma = 0$. Then from (28), $t_N = 0$, $t_R = \frac{t}{1+t}$, so (34) simplifies to

$$MEB = \frac{\frac{t}{1+t}e + C}{1 - \frac{t}{1+t}(1+e) - C} \tag{37}$$

It can be checked that in this case, C is given by the explicit formula (22), replacing t_H, t_L by $t_R, 0$ respectively.

6.4 Simulations

Here we calibrate the VAT model, and plot the true MEB in (34) and an approximation to the MEB as parameters vary. [18] The approximation is the one treating VAT as a proportional tax i.e. setting $C = 0$ in (37), which gives

$$MEB_A = \frac{\frac{t}{1+t}e}{1 - \frac{t}{1+t}(1+e)}$$

The parameters are calibrated as follows. In the UK, the statutory rate of VAT is 20%, so $t = 0.2$. Liu and Lockwood (2016) calculate that for the universe of firms in the UK that file a corporate tax return, $\gamma = 0.45$. This gives $t_N = 0.16$, $t_R = 0.30$.

Next, define n_0 to be the productivity level corresponding to turnover just at the threshold i.e. $n_0(1-t_N)^e = z_0$. From Liu and Lockwood (2015), 62.5% of firms are below the threshold. So, $\frac{\underline{n}}{n_0}$ must satisfy $H(n_0) = 1 - \left(\frac{\underline{n}}{n_0}\right)^{1.2} = 0.625$, or $\frac{\underline{n}}{n_0} = (0.375)^{1/1.2} = 0.442$. Given that only the ratio $\frac{\underline{n}}{n_0}$ is determined, we set $\underline{n} = 1$, so $n_0 = 2.26$. But then $z_0 = 2.53(0.84)^{0.25} = 2.164$.

Finally, we need a value for a. A prior question is whether the "upper tail" of the distribution of firm sales y is well-described by a Pareto distribution. In the case of personal incomes, a Pareto distribution of the upper tail is widely accepted, but less is known about firms. In the US, there is evidence that the size distribution of firms as measured by sales is Pareto (Luttmer (2007)), and Luttmer estimates a value for the US of $a = 1.06$. In the Online Appendix, we provide evidence that this is also the case for the UK, using firm sales from administrative data on corporate tax returns. We show that for firms above the VAT threshold, the estimate a is about 1.2. So, this is the figure we will use in the simulations.

[18] The details of the calibration are described in the Online Appendix.

Our results are given in Figures 3-4 below. Here, we see that the true MEB is about 3 times higher than the approximation. Also, the true MEB is increasing in both e and a. This difference is much smaller than in the income tax case, which is due partly by the lower value of a in the VAT case. Indeed, we can see in Figure 4 that the accuracy of the approximation MEB_A falls rapidly as a rises, because MEB is increasing in a whereas MEB_A is independent of a.

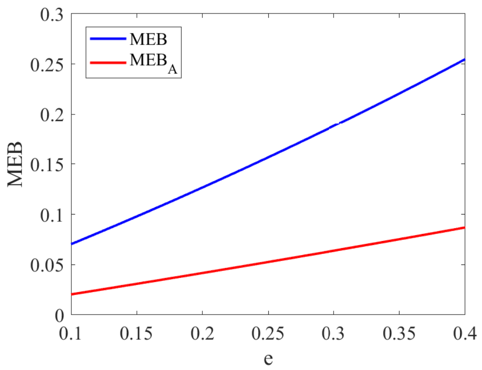

Figure 3: MEB of VAT as e varies, $\gamma = 0.45$

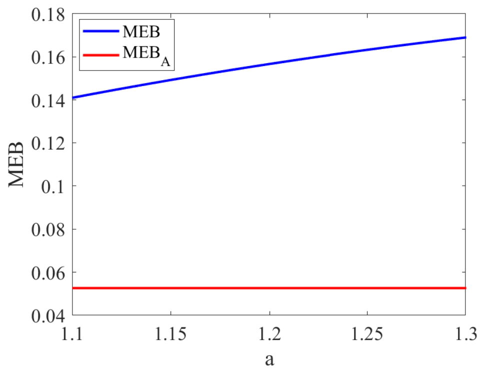

Figure 4: MEB of VAT as a varies, $\gamma = 0.45$

7 Conclusions

This paper shows that the sufficient statistic approach to the welfare properties of income (and other) taxes does not easily extend to tax systems with notches, because with notches, changes in aggregate bunching induced by changes in tax rates have a first-order effect on tax revenues. In an income tax setting, we showed that the MEB of a change in the top rate of tax is given by the Feldstein (1999) formula for the MEB of a *proportional* tax, plus a correction term. This formula also applies when the model is extended to allow for tax evasion. Also, under certain conditions, the optimal top rate of tax is given by the formula for the optimal proportional tax, minus a correction term. These correction terms can be computed empirically, using an estimate of excess mass at the notch. Quantitatively, these correction terms can be very large.

An application to VAT was also discussed. A simple model of small traders who differ in productivity, and are subject to VAT at rate t above a threshold level of sales was shown to be formally equivalent to the income tax model. We showed that the MEB of an increase in the statutory rate of VAT is given by the Feldstein formula for a proportional tax plus a correction factor as in the income tax case. With a calibration to UK data, the MEB of the VAT is roughly three times what is would be if VAT was simply a proportional tax.

References

Apps, Patricia, Ngo Long, and Ray Rees, "Optimal piecewise linear income taxation," *Journal of Public Economic Theory*, 2014, *16* (4), 523–545.

Bachas, Pierre and Mauricio Soto, "Not (ch) your average tax system: corporate taxation under weak enforcement," *Unpublished manuscript*, 2015.

Ballard, Charles L, John B Shoven, and John Whalley, "The total welfare cost of the United States tax system: a general equilibrium approach," *National Tax Journal*, 1985, pp. 125–140.

Best, Michael and Henrik Jacobsen Kleven, "Housing Market Responses to Transaction Taxes: Evidence from Notches and Stimulus in the UK," Memeo, London School of Economics 2013.

Blinder, Alan S and Harvey S Rosen, "Notches," *The American Economic Review*, 1985, *75* (4), 736–747.

Chetty, Raj, "Is the Taxable Income Elasticity Sufficient to Calculate Deadweight Loss? The Implications of Evasion and Avoidance," *American Economic Journal: Economic Policy*, August 2009, *1* (2), 31–52.

_ , "Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods," *Annu. Rev. Econ.*, 2009, *1* (1), 451–488.

Doerrenberg, Philipp, Andreas Peichl, and Sebastian Siegloch, "The elasticity of taxable income in the presence of deduction possibilities," *Journal of Public Economics*, 2015.

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

Gillitzer, Christian and Joel Slemrod, "Does Evasion Invalidate the Welfare Sufficiency of the ETI?," *The BE Journal of Economic Analysis & Policy*, 2016, *16* (4).

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

Hargaden, Enda Patrick, "Taxpayer Responses over the Cycle: Evidence from Irish Notches," 2015.

Harju, Jarkko, Tuomas Matikka, and Timo Rauhanen, "The Effects of Size-Based Regulation on Small Firms: Evidence from Vat Threshold," Working Paper 75, VATT Institute for Economic Research June 2016.

Kanbur, Ravi and Michael Keen, "Thresholds, Informality, and Partitions of Compliance," *International Tax and Public Finance*, 2014, *21* (4), 536–559.

Keen, Michael and Jack Mintz, "The Optimal Threshold for a Value-Added Tax," *Journal of Public Economics*, 2004, *88* (3-4), 559–576.

Kleven, Henrik, "Bunching," *Annual Review of Economics*, 2016, *8*.

Kleven, Henrik Jacobsen and Esben Anton Schultz, "Estimating taxable income responses using Danish tax reforms," *American Economic Journal: Economic Policy*, 2014, *6* (4), 271–301.

_ **and Mazhar Waseem**, "Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan," *The Quarterly Journal of Economics*, 2013, *128* (2), 669–723.

Kopczuk, Wojciech and David Munroe, "Mansion Tax: The Effect of Transfer Taxes on the Residential Real Estate Market," *American Economic Journal: Economic Policy*, May 2015, *7* (2), 214–57.

Liu, Li and Ben Lockwood, "VAT Notches," Discussion Papers 10606, CEPR May 2015.

Luttmer, Erzo GJ, "Selection, growth, and the size distribution of firms," *The Quarterly Journal of Economics*, 2007, *122* (3), 1103–1144.

Piketty, Thomas and Emmanuel Saez, "Optimal Labor Income Taxation," *Handbook of Public Economics*, 2013, *5*, 391.

Rutherford, Thomas and Sergey Paltsev, "From an input-output table to a general equilibrium model: assessing the excess burden of indirect taxes in Russia," *Draft, University of Colorado*, 1999.

Saez, Emmanuel, "Using Elasticities to Derive Optimal Income Tax Rates," *The Review of Economic Studies*, 2001, *68* (1), 205–229.

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

_ **, Joel Slemrod, and Seth H. Giertz**, "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review," *Journal of Economic Literature*, 2012, *50* (1), 3–50.

Slemrod, Joel, "Buenas notches: lines and notches in tax system design," *eJournal of Tax Research*, 2013, *11* (3), 259.

_ **, Shlomo Yitzhaki, Joram Mayshar, and Michael Lundholm**, "The optimal two-bracket linear income tax," *Journal of Public Economics*, 1994, *53* (2), 269–290.

Tazhitdinova, Alisa, "Does the Magnitude of Incentives Matter? Labor Supply and Demand Responses to an Unusually Large Tax Break," 2018.

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

A Appendix

A.1 Proofs of Propositions

Proof of Proposition 3. Under the assumptions made in this Proposition, $\bar{e} = e$. So, the MEB formula (21) follows from (19). It remains to derive the formula (22) for C. From (9), noting that

$$\psi_n = -\frac{1}{1+e}\left(\frac{z}{n}\right)^{1+1/e}. \tag{A.1}$$

and $z(1-t;n) = (1-t)^e n$, we have

$$\frac{\partial n_H}{\partial t_H} = \frac{(1-t_H)^e n_H(1+e)}{(1-t_H)^{1+e} - \left(\frac{z_0}{n_H}\right)^{1+1/e}} \tag{A.2}$$

Next, from (14) and (16), using the fact that $z(1-t;n) = (1-t)^e n$, we have

$$\frac{1}{B_H}\frac{\partial R}{\partial n_H} = \frac{(t_L z_0 - t_H(1-t_H)^e n_H)h(n_H)}{(1-t_H)^e \int_{n_H}^{\overline{n}} nh(n)dn} \tag{A.3}$$

So, plugging (A.2),(A.3) into the formula for C in (19), we have:

$$C = \frac{(1-t_H)(t_H(1-t_H)^e n_H - t_L z_0)}{(1-t_H)^e E[n\,|n \geq n_H]}\frac{h(n_H)}{(1-H(n_H))}\frac{n_H(1+e)}{(1-t_H)^{1+e} - \left(\frac{z_0}{n_H}\right)^{1+1/e}}$$

where we have used $\int_{n_H}^{\overline{n}} nh(n)dn = E[n\,|n \geq n_H]\,(1-H(n_H))$ in (A.4).

Now, given that n follows a Pareto distribution with shape and scale parameters $a, \underline{n}$, we also know that

$$E[n\,|n \geq n_H] = \frac{an_H}{a-1}, \frac{h(n)}{1-H(n)} = \frac{a}{n} \tag{A.4}$$

Plugging (A.4) into (A.4), we get:

$$C = \frac{(1-t_H)(t_H(1-t_H)^e - t_L z_0/n_H)(a-1)(1+e)}{(1-t_H)^{1+e} - \left(\frac{z_0}{n_H}\right)^{1+1/e}} \tag{A.5}$$

Then, using the definition $\tilde{z}_H = n_H(1-t_H)^e$ to eliminate n_H in (A.5), and rearranging, we get (22) as required. □

Proof of Proposition 4. Let B_N, B_R be the bases of the effective taxes t_N, t_R defined in (30). Then from (29),(30) and remembering that a change in the statutory rate of VAT t changes t_N, t_R via (28), we have:

$$\frac{dW}{dt} = -\left(\frac{\partial t_N}{\partial t}B_N + \frac{\partial t_R}{\partial t}B_R\right) \tag{A.6}$$

$$\frac{dR}{dt} = \frac{\partial t_N}{\partial t}\left(B_N + t_N\left.\frac{\partial B_N}{\partial t_N}\right|_{n_R\text{ const}}\right) + \frac{\partial t_R}{\partial t}\left(B_R + t_R\left.\frac{\partial B_R}{\partial t_R}\right|_{n_R\text{ const}}\right) - C' \tag{A.7}$$

where

$$C' = -\frac{\partial R}{\partial n_R}\left(\frac{\partial t_N}{\partial t}\frac{\partial n_R}{\partial t_N} + \frac{\partial t_R}{\partial t}\frac{\partial n_R}{\partial t_R}\right) \tag{A.8}$$

So, plugging (A.6),(A.7) into (15), we have, after rearrangement

$$\begin{aligned} MEB &= -\frac{d(W+R)/dt}{dR/dt} \qquad \text{(A.9)} \\ &= \frac{\frac{B_N}{1-t_N}\frac{\partial t_N}{\partial t}t_N\left(\frac{1-t_N}{B_N}\frac{\partial B_N}{\partial(1-t_N)}\Big|_{n_R\text{ const}}\right) + \frac{B_R}{1-t_R}\frac{\partial t_R}{\partial t}t_R\left(\frac{1-t_R}{B_R}\frac{\partial B_R}{\partial(1-t_R)}\Big|_{n_R\text{ const}}\right) + C'}{\frac{B_N}{1-t_N}\frac{\partial t_N}{\partial t}\left(1-t_N-t_N\frac{1-t_N}{B_N}\frac{\partial B_N}{\partial t_N}\Big|_{n_R\text{ const}}\right) + \frac{B_R}{1-t_R}\frac{\partial t_R}{\partial t}\left(1-t_R-t_R\frac{1-t_R}{B_R}\frac{\partial B_R}{\partial t_R}\Big|_{n_R\text{ const}}\right) - C'} \\ &= \frac{\frac{B_N}{1-t_N}\frac{\partial t_N}{\partial t}e\phi + \frac{B_R}{1-t_R}\frac{\partial t_R}{\partial t}t_R e + C'}{\frac{B_N}{1-t_N}\frac{\partial t_N}{\partial t}(1-t_N(1+e\phi)) + \frac{B_R}{1-t_R}\frac{\partial t_R}{\partial t}(1-t_R(1+e)) - C'} \end{aligned}$$

where in the last line, we have used (31).So, dividing top and bottom of (A.9) by $\frac{B_R}{1-t_R}\frac{\partial t_R}{\partial t} + \frac{B_N}{1-t_N}\frac{\partial t_N}{\partial t}$ and using the definition of θ from (33), and the definition of C from (36), we get

$$MEB = \frac{(1-\theta)t_N e\phi + \theta t_R e + C}{1-(1-\theta)t_N(1+e\phi) - \theta t_R(1+e) - C} \tag{A.10}$$

Finally, using the definitions of $\tau = (1-\theta)t_N + \theta t_R$, $\varepsilon = \frac{(1-\theta)t_N\phi + \theta t_R}{(1-\theta)t_N + \theta t_R}e$, (A.10) can be rearranged to (35), as required. □

A.2 The VAT Model

Model Set-Up. Consider a single industry with a fixed, large number of small traders producing a homogeneous good. Each small trader combines his own labor input l with an intermediate input x to produce output y via a fixed coefficients technology

$$y = \min\left\{l, \frac{x}{\gamma}\right\}, \tag{A.11}$$

where γ measures the the input requirement per unit of output. In particular, for all traders, to produce one unit of output requires γ units of input.

Individual traders are indexed by a skill or taste parameter $m \in [\underline{m}, \overline{m}]$, assumed continuously distributed in the population with distribution $H(m(1-\gamma))$ and density $h(m(1-\gamma))$. A trader of type m has an overall payoff of

$$u(l;m) = \pi - \psi(l;m), \;\; \psi(l;m) = \frac{Am}{1+\frac{1}{e}}\left(\frac{l}{m}\right)^{1+\frac{1}{e}} \tag{A.12}$$

where π is profit and $\psi(l;m)$ is the disutility of labour. So, traders are differentiated by disutility of labor.

For simplicity, it is assumed that traders only sell to final consumers, who have perfectly elastic demand for the good at price $p = 1$. This is analogous to the assumption made in the taxable income literature that the wage is fixed, i.e. labor demand is perfectly elastic at a fixed wage. Finally, the intermediate input is produced only from labor supplied by non-trader households via a fixed-coefficients technology where one unit of labor is needed to produce one unit of the intermediate input. So, the tax-exclusive price of the output is w, the wage, which we also assume to be 1. This implies that the tax-exclusive value of sales is just y.

The traders and the producer of the intermediate input face a VAT system. It is assumed that the producer is VAT-registered. If the trader is registered, he must charge VAT on sales y at rate t, but can claim back any VAT paid on inputs. The trader must register for VAT if the value of sales y exceeds the threshold y_0, but can register voluntarily even if $y < y_0$.

We can now compute trader profit as follows. When not registered, the price of the input is $1+t$. So, the profit for the non-registered trader is

$$\pi_N = (1-\gamma(1+t))y. \tag{A.13}$$

where γ is the cost of inputs relative to revenue per unit sold. For the registered trader, we reason as follows. This trader must charge VAT on his output. None of the output VAT can be passed on to the buyer, as he has perfectly elastic demand. So, revenue per unit sold is $p/(1+t)$. But, if the trader is registered, he can claim back VAT on the input use x, so the price of the input is γ. So, overall, the profit for the registered trader is

$$\pi_R = \left(\frac{1}{1+t} - \gamma\right) y. \tag{A.14}$$

Trader Payoffs and Effective Tax Rates. Trader utility is profit minus the disutility of labour. So, combining (A.1), (A.13), (A.14) and defining $n \equiv m(1-\gamma)$, $l = y$, we get:

$$\begin{aligned} u_N &= (1-\gamma(1+t))y - \frac{A}{1-\gamma}\frac{n}{1+\frac{1}{e}}\left(\frac{y(1-\gamma)}{n}\right)^{1+\frac{1}{e}} \\ u_R &= \left(\frac{1}{1+t} - \gamma\right) y - \frac{A}{1-\gamma}\frac{n}{1+\frac{1}{e}}\left(\frac{y(1-\gamma)}{n}\right)^{1+\frac{1}{e}} \end{aligned} \tag{A.15}$$

Now, using $z = y(1-\gamma)$ in (A.15), and setting $A = 1-\gamma$, we get

$$\begin{aligned} u_N &= \frac{1-\gamma(1+t)}{1-\gamma} z - \frac{n}{1+\frac{1}{e}}\left(\frac{z}{n}\right)^{1+\frac{1}{e}} \\ u_R &= \left(\frac{1}{(1+t)(1-\gamma)} - \frac{\gamma}{1-\gamma}\right) z - \frac{n}{1+\frac{1}{e}}\left(\frac{z}{n}\right)^{1+\frac{1}{e}} \end{aligned} \tag{A.16}$$

Finally, we note from (28) that

$$1-t_N = \frac{1-\gamma(1+t)}{1-\gamma}, \ 1-t_R = \frac{1}{(1+t)(1-\gamma)} - \frac{\gamma}{1-\gamma} \tag{A.17}$$

Then, combining (A.16), (A.17), we get (27), (28) as required.

Tax Revenue. Now we derive (29) in the text. Let $y(n)$ be the sales of an n-type trader. Note that as m has distribution function $H(m(1-\gamma))$, n has has distribution function

$H(n)$. Then, revenue from the from the VAT is

$$R = \frac{t}{1+t}\int_{n_R}^{\overline{n}} y(n)h(n)dn + t\int_{\underline{n}}^{n_R} \gamma y(n)h(n)dn \tag{A.18}$$

The first term is revenue from VAT levied on the value of sales of registered firms, because the sale price is $1/(1+t)$, and the second term is revenue from inputs sold by the intermediate input producer to firms that do not register for VAT. Using $z(n) = y(n)(1-\gamma)$, we can write this as

$$R = \frac{t}{(1+t)(1-\gamma)}\int_{n_R}^{\overline{n}} z(n)h(n)dn + \frac{t\gamma}{1-\gamma}\int_{\underline{n}}^{n_R} z(n)h(n)dn \tag{A.19}$$

Finally, replacing $z(n)$ by $z(1-t_N;n), z_0$, or $z(1-t_R;n)$ where appropriate, and using (A.17) for the definitions of t_N, t_R, we get (29) as required.

Online Appendix

The Extended Model with Tax Avasion. Maximizing (24) in the paper with respect to z, s subject to (25) gives first-order conditions for the choice of z, s respectively;

$$1 - T'(z) - \psi_z(z + s; n) = 0 \tag{N.1}$$
$$1 - g'(s) - a'(s) - \psi_z(z + s; n) = 0 \tag{N.2}$$

Here, for convenience, we denote the derivative of ψ with respect to $z + s$ as ψ_z. Note from (N.1),(N.2) that the household equates the marginal cost of evasion $g' + a'$, to the marginal tax rate, T' as in Chetty (2009b). Now let $T' = t$. From (N.1),(N.2) we can then solve for reported and sheltered income as functions of the net of tax rate $1 - t$, which we write as $z(1-t; n)$ and $s(1-t; n)$ respectively. However, note that with evasion, even assuming the iso-elastic specification for disutility of income, (2), no longer gives a closed-form solution for z in terms of $1 - t$, as z, s are simultaneously determined via (N.1), (N.2).

Then, the behavior of the household faced with a kink or a notch is qualitatively the same as before. That is, under either type of tax schedule, there is a bunching interval $[n_L, n_H]$. In the case of a kink, n_L, n_H are characterized by (5), (6) as before. In the case of a notch, (7) is modified to

$$(1 - t_L)z_0 + \hat{s} - g(\hat{s}) - a(\hat{s}) - \psi(z_0 + \hat{s}; n_H) = v(1 - t_H; n_H) \tag{N.3}$$

Here, $\hat{s}$ is the unique solution to (N.2) for type n_H, where also $z = z_0, T' = t_L$ and so $\hat{s}$ is the optimal amount of evaded income for this type at the threshold. Also,

$$v(1 - t; n) = \max_{z,s} \{(1 - t)z + s - \psi(z + s; n) - g(s) - a(s)\}$$

So, (N.3) says that the marginal buncher at the top is the one for whom the utility from keeping reported income at z_0, while optimizing sheltered income, is equal to the unconstrained indirect utility, given the higher rate of tax on all reported income.

Given this characterization of household behavior, it is then easily seen that much the analysis above generalizes to this case, with the mathematics being unchanged, while bearing in mind the new interpretation of $z(1-t; n)$. In particular, it is still the case that the effect on revenue R of a change in n_H is zero in the kink case and negative in the notch case, as this simply follows from the (dis-)continuity of R in the kink (notch) case. So, Proposition 1 continues to hold.

As for Proposition 2, we can establish, using the envelope theorem, that

$$\frac{dW}{dt_H} = -B_H + A_H,\ A_H = \int_{n_H}^{\overline{n}} a'(s(1 - t_H; n))s_t((1 - t_H; n))h(n)dn > 0 \tag{N.4}$$

The intuition for $A_H > 0$ is as in Chetty (2009b); the audit cost is a transfer to government, which is recycled back to the households, and so reduces the welfare cost of the tax increase.

Using (N.4), and the formula (18) above, simple manipulation gives a formula for the excess burden

$$MEB = \frac{t_H \bar{e} + C - \frac{(1-t_H)A_H}{B_H}}{1 - t_H(1+\bar{e}) - C} \tag{N.5}$$

where C is given by the same formula as in Proposition 2. So, we can make two observations.

First, in the special case where there is no audit cost of evasion, i.e. $a \equiv 0$, Proposition 2 continues to hold. So, our basic insight that the MEB of a notch is equal to the MEB of a proportional tax plus a correction factor generalizes to the case of evasion. However, Proposition 3, and in particular, the explicit formula for C, depends heavily on the simple functional form $z(1-t;n) = (1-t)^e n$, and so does not generalize, even if ψ g, and a are all assumed iso-elastic.

Second, in the more realistic case where there is an audit cost, the MEB is equal to the MEB of a proportional tax plus two correction factors, one for the notch, C, and the other $\frac{(1-t_H)A_H}{B_H}$, capturing the fact that the audit cost is a transfer. As the first is positive and the second is negative, they have offsetting effects on the MEB.

Details of MEB Simulation for the VAT Case. We need to express all the relevant elements of the MEB in terms of the parameters, t, γ, z_0,and n_R, n_N. In turn, we know that $n_N = z_0/(1-t_N)^e$ and that n_R is determined by

$$(1-t_N)z_0(1+e) - e(n_R)^{-1/e}(z_0)^{1+\frac{1}{e}} - n_R(1-t_R)^{1+e} = 0 \tag{N.6}$$

Assume that the distribution of firms is Pareto with shape and scale parameters $a, \underline{n}$. Without loss of generality, we assume $\underline{n} = 1$; so, the distribution and density of n is $H(n) = 1 - n^{-a}$, $h(n) = \frac{a}{n^{a+1}}$. So, using these formulae and $z(1-t;n) = (1-t)^e n$, we have by routine calculation;

$$\begin{aligned}
B_R &= (1-t_R)^e \int_{n_R}^{\overline{n}} nh(n)dn = (1-t_R)^e \frac{a}{a-1}(n_R)^{1-a} \qquad \text{(N.7)}\\
B_N &= (1-t_N)^e \int_{1}^{n_N} nh(n)dn + z_0(H(n_R) - H(n_N))\\
&= (1-t_N)^e \frac{a}{a-1}(1-(n_N)^{1-a}) + z_0\left((n_N)^{-a} - (n_R)^{-a}\right)
\end{aligned}$$

Moreover, from the formulae for t_N, t_R in the paper, we have:

$$\frac{\partial t_R}{\partial t} = \frac{1}{(1-\gamma)(1+t)^2}, \ \frac{\partial t_R}{\partial t} = \frac{\gamma}{(1-\gamma)} \tag{N.8}$$

So, plugging (N.8) into the formula for θ in the paper, we can write

$$\theta = \frac{\frac{B_R}{1-t_R}}{\frac{B_R}{1-t_R} + \frac{B_N}{1-t_N}\gamma(1+t)^2} \tag{N.9}$$

Plugging (N.7) into (N.10) allows us to compute θ as a function of t, γ, z_0,and n_N, n_R.

Next, using $z(1-t;n) = (1-t)^e n$, and the properties of the Pareto distribution, we have;

$$\phi = \frac{\int_{\underline{n}}^{n_N} z(1-t_N;n)h(n)dn}{B_N} = \frac{(1-t_N)^e \frac{a}{a-1}(1-(n_N)^{1-a})}{B_N} \tag{N.10}$$

So, using (N.7), (N.10), ϕ can be computed as a function of t, γ, z_0,and n_N, n_R.

Finally, recalling the definition of C in the paper, we have:

$$C = -\frac{\frac{\partial R}{\partial n_R}\left(\frac{\partial t_N}{\partial t}\frac{\partial n_R}{\partial t_N} + \frac{\partial t_R}{\partial t}\frac{\partial n_R}{\partial t_R}\right)}{\frac{B_N}{1-t_N}\frac{\partial t_N}{\partial t} + \frac{B_R}{1-t_R}\frac{\partial t_R}{\partial t}} = -\frac{\frac{\partial R}{\partial n_R}\left(\gamma\frac{\partial n_R}{\partial t_N} + \frac{1}{(1+t)^2}\frac{\partial n_R}{\partial t_R}\right)}{\frac{B_N}{1-t_N}\gamma + \frac{B_R}{1-t_R}\frac{1}{(1+t)^2}} \tag{N.11}$$

where in the second line, we use (N.8).

It remains to calculate $\frac{\partial n_R}{\partial t_N}, \frac{\partial n_R}{\partial t_R}, \frac{\partial R}{\partial n_R}$. From (N.6), we have:

$$\frac{\partial n_R}{\partial t_R} = \frac{(1-t_R)^e n_R(1+e)}{(1-t_R)^{1+e} - \left(\frac{z_0}{n_R}\right)^{1+1/e}} \tag{N.12}$$

$$\frac{\partial n_R}{\partial t_N} = \frac{-z_0(1+e)}{(1-t_R)^{1+e} - \left(\frac{z_0}{n_R}\right)^{1+1/e}}$$

Moreover, from the formula for $\frac{\partial R}{\partial n_R}$ in the paper, and the iso-elastic form of $z(1-t, n)$, we get

$$\frac{\partial R}{\partial n_R} = (t_N z_0 - t_R(1-t_R)^e n_R)h(n_R) \tag{N.13}$$

Plugging (N.12),(N.13) into (N.11), and using the formula for the density of the Pareto density to substitute out $h(n_R)$, we eventually get:

$$C = \frac{(t_R(1-t_R)^e n_R - t_N z_0)\left((1-t_R)^e n_R \frac{1}{(1+t)^2} - z_0\gamma\right)(1+e)}{\left((1-t_R)^{1+e} - \left(\frac{z_0}{n_R}\right)^{1+1/e}\right)\left(\frac{B_R}{1-t_R}\frac{1}{(1+t)^2} + \frac{B_N}{1-t_N}\gamma\right)}\frac{a}{(n_R)^{a+1}}$$

This expression for C is computable knowing t, γ, z_0,and n_R, n_L. Thus, all the components of MEB in the paper can be calculated. □

Calculation of the Pareto Parameter for UK firms. We use the method of Luttmer (2007) and others to estimate the distribution of of firm size for the UK using corporate tax return data. Firm size y is measured by sales. If the distribution of firm size is Pareto, the log of the size of the upper tail of the distribution of firm size is linear in y, with the coefficient on y being a.

We briefly describe the the data here: a fuller description is given in Liu and Lockwood (2015). We have annual sales of firms, taken from the universe of corporation tax records (CT600) in financial years 2004/5 to 2009/10. The data is then refined by eliminating companies which are part of a larger VAT group i.e. using only standard-alone independent companies. We also drop all observations with partial-year corporation tax records. In addition, we eliminate companies that mainly engage in overseas activities. This yields a data-set with 731,706 observations for 435,688 companies between April 1, 2004 and March 30, 2010.

To analyze the data, we group firm sales into bins of size £10,000. A visual inspection of the data (available on request) indicates that log of the size of the upper tail of the distribution of firm size is near to linear in y. We then regress (by year) the size of the upper tail of the distribution, denoted $1-F$, on firm size as measured by sales. We define the upper tail as starting at the VAT threshold in any year. The coefficient on sales in this regression gives a value for a. Inspection of the results in Table 1 below indicates that for our population of UK firms, a is approximately 1.2.

Table 1: Regression of Upper Tail of the Size Distribution on Firm Size

	Dependent Variable: log of 1-F					
	2004-05	2005-06	2006-07	2007-08	2008-09	2009-10
Sales	-1.202***	-1.182***	-1.180***	-1.202***	-1.193***	-1.182***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Constant	3.353***	3.192***	3.285***	3.413***	3.344***	3.133***
	(0.029)	(0.027)	(0.028)	(0.031)	(0.029)	(0.026)
R-squared	0.975	0.977	0.975	0.971	0.974	0.979
N	2119	2185	2336	2349	2421	2334