LIS
Working Paper Series

No. 711

The Laffer curve for high incomes

Jacob Lundberg

August 2017

Luxembourg Income Study (LIS), asbl

The Laffer curve for high incomes

Jacob Lundberg*

31 August 2017

Abstract

An expression for the Laffer curve for high incomes is derived, assuming a constant Pareto parameter and elasticity of taxable income. The peak of this Laffer curve is given by the well-known Saez (2001) expression. Microsimulations using Swedish population data show that the simulated curve matches the theoretically derived Laffer curve well, suggesting that the analytical expression is not too much of a simplification. Policy conclusions do not change much when income effects are taken into account. A country-level dataset of top effective marginal tax rates and Pareto parameters is assembled. This is used to draw Laffer curves for 27 OECD countries. Revenue-maximizing tax rates and degrees of self-financing for a small tax cut are also computed. The results indicate that degrees of self-financing range between 28 and 195 percent. Five countries have higher tax rates than the peak of the Laffer curve.

*Department of Economics and Uppsala Center for Fiscal Studies, Uppsala University. Email: lundberg.jacob@gmail.com. I am grateful to Spencer Bastani, Per Engström, Katarina Nordblom, Emmanuel Saez, Håkan Selin, Helena Svaleryd, Daniel Waldenström and seminar participants at the Department of Economics, Uppsala University, for their comments. Financial support from the Jan Wallander and Tom Hedelius Foundation is gratefully acknowledged.

1 Introduction

The appropriate tax rate on high incomes is intensely debated both in academia and in the political arena. The Laffer curve – the relationship between the tax rate and tax revenues – is a recurring topic in this debate. The curve became famous after a 1974 Washington dinner when conservative economist Arthur Laffer drew it on a napkin, although the insight that the tax rate may affect the tax base is much older. (Laffer, 2004) The napkin is currently on display at the Smithsonian Institution. Laffer went on to become an economic advisor to U.S. president Reagan, and since then the Laffer curve has been closely associated with Reaganomics and the tax reforms of the 1980s. It is perhaps the concept in public economics that is most well-known among the general public. The shape of the Laffer curve, and countries' positions on it, is important for policy because the purpose of taxation is to raise revenue.

In the past 20 years, new empirical and theoretical insights have allowed economists to be more concrete about the fiscal and welfare effects of top income taxation. Saez (2001), building on Diamond (1998), made a seminal contribution by showing that the revenue-maximizing top marginal tax rate, i.e., the peak of the Laffer curve, can be expressed as a function of only two parameters within the framework of Mirrleesian optimal taxation: $\tau^* = 1/(1+\alpha\varepsilon)$. These parameters are the elasticity of taxable income with respect to the net-of-tax rate (ε) and the Pareto parameter (α), a measure of the thinness of the right tail of the income distribution. The taxable income elasticity measures the strength of taxpayer responses to taxation and the inverse of the Pareto parameter is the percentage of the average high-income taxpayer's income that is subject to the top marginal tax rate. Intuitively, income taxation is more distortionary if the elasticity of taxable income is higher and if a lower proportion of average income is in the top tax bracket, because this means less tax revenue.

This paper is concerned with the Laffer curve for high incomes, i.e., tax revenues from the top tax bracket as a function of the top effective tax rate. The main contribution is the derivation of an analytical expression for the high-income Laffer curve and the evaluation of this by way of microsimulations on Swedish register data. The expression for the Laffer curve enables the researcher to approximate the fiscal impact of top tax rate changes – for example, how much would be gained by a move to the revenue-maximizing rate – without access to microdata on incomes. Previous research has simulated Laffer curves in various countries, but the explicit expression for the Laffer curve presented in this paper has not been derived before to my knowledge.

Two major assumptions are needed to derive the Laffer curve: First, the individual maximizes a quasilinear, isoelastic utility function, i.e., the taxable income elasticity is constant and there are no income effects. Second, potential incomes – the levels of taxable income that individuals would choose to supply if there were no taxation – follow a Pareto distribution. The Pareto distribution is a power-law type distribution and has been shown to be a good approximation for high incomes in many countries. These assumptions are the same as the ones needed to derive the well-known expression for the revenue-maximizing tax rate.

The logic behind the derivation is as follows. If potential incomes are Pareto-distributed, the Pareto parameter of the realized income distribution will be independent of the tax rate. Because the Pareto parameter is a function of average income in the top tax bracket, this implies that the average income of top-bracket taxpayers will be constant even when the tax rate changes. Instead, the number of taxpayers in the top tax bracket will vary depending on the tax rate. For example, if the top marginal tax rate is lowered, this will induce those who are already in the top bracket to increase their income. At the same time, taxpayers in lower tax brackets will increase their income to make their way into the top tax bracket, pushing down average income to its starting point. The number of high-income taxpayers will have increased, but their average income will be constant.

Finding an expression for tax revenues thus becomes a matter of calculating the number of people who are subject to the top marginal tax rate. Given Pareto distribution of potential incomes and an isoelastic utility function, this is relatively straightforward. I show that the high-income Laffer curve has the form $R = \tau(1-\tau)^{\alpha\varepsilon}$, where τ is the top marginal tax rate. The peak of this Laffer curve coincides with the top tax rate derived by Saez (2001). Another desirable property is that tax revenues from the top bracket are zero when the marginal tax rate is either 0 or 100 percent.

To simulate Laffer curves, the distribution of potential incomes is obtained from the observed labour income distribution in Sweden, assuming that the current income distribution is the result of individuals maximizing an isoelastic utility function subject to the current Swedish tax schedule, which is comprised of income tax, consumption taxes and the tax portion of social contributions. This allows me to evaluate the analytical Laffer curve and assess the importance of some of the assumptions needed to derive it – primarily the assumption that potential incomes are Pareto-distributed. I let individuals maximize utility given a counterfactual tax schedule where high incomes are subject to a tax rate that varies between 0 and 100 percent. The simulated Laffer curve is close to the theoretically derived Laffer curve: the analytical Laffer curve peaks at 61 percent while the simulated curve peaks at 64 percent. This indicates that the assumptions needed to derive the analytical expression are not overly restrictive. The main explanation for this is that the Swedish income distribution is remarkably close to an exact Pareto distribution above the threshold for central government income tax, which is the region for which the marginal tax rate is varied.

In the main analysis, income effects are ignored. This is in line with many other public finance papers and simplifies the derivations significantly because when utility is quasi-linear, there exists a closed form for the taxable income supply function. Extending the analysis to account for income effects of reasonable magnitude affects the results little, at least in countries where the Pareto parameter is relatively high. The intuition is that changing the top marginal tax rate may alter the incentive to earn income at the margin substantially, thereby inducing sizeable substitution effects, while net income may not increase as much, implying small income effects.

As an application of the analytical expression and illustration of its usefulness, I draw high-income Laffer curves for 27 OECD countries, using a specially assembled country-level dataset of Pareto parameters and top effective marginal tax rates, i.e., including payroll taxes, social contributions and consumption taxes. I also compute revenue-maximizing tax rates and degrees of self-financing for a small tax cut, given a taxable income elasticity of 0.2. The results, though they should be interpreted with some caution, suggest that five countries have surpassed the peak of the Laffer curve and would thus gain revenue by cutting the top tax rate.

2 Related literature

The theoretical derivations in this paper mainly build on Saez (2001), but a number of researchers have characterized the Laffer curve or analyzed top income taxation in various countries before. The only other explicit expression for the high-income Laffer curve that I have found in the literature is derived by Badel (2013). His Laffer curve is given by the expression $R = \tau(z_0(1-\tau)^{\varepsilon} - b)$, where z_0 is potential income and b is the top bracket threshold. However, this only holds for a representative individual and thus fails to take into account the fact that the number of high-income taxpayers in general will be a function of the tax rate. The peak of this curve does not coincide with the Saez revenue-maximizing rate and the curve predicts negative tax revenue for high tax rates.

The Laffer curve for a proportional tax can easily be obtained by setting $\alpha = 1$ in the expression for the high-income Laffer curve (see this by setting $b = 0$ in equation 4 below), so that $R = \tau(1-\tau)^{\varepsilon}$. This expression is known in the literature (e.g., Usher, 2014). Piketty & Saez (2013) also discuss the proportional-tax Laffer curve, but do not derive an explicit expression.

A few papers simulate high-income Laffer curves, but do not derive any explicit expressions. Giertz (2009) simulates Laffer curves for the United States using a few different elasticities, but provides little information on how the simulations are carried out. Badel & Huggett (2014) also simulate Laffer curves for top-income earners, taking human capital formation into account. Their model differs from the present paper in significant ways, for example in that it is dynamic and not parameterized to match estimates from the quasi-experimental literature. Badel and Huggett also characterize the income distribution quite crudely. They show that using the Saez (2001) revenue-maximizing tax rate with an econometrically identified taxable income elasticity results in substantially higher revenue-maximizing rates than the true (numerically simulated) value when endogenous human capital accumulation is accounted for.

Bastani & Seli (2014) take a methodologically similar approach to the present paper. The simulations in section 6 resemble the simulation exercise in Bastani and Selin with respect to, e.g., the utility function and the country of interest (Sweden). There is also a similarity in that numerical simulations are carried out in order to evaluate a simple analytical expression. However, Bastani and Selin analyze a bunching estimator of the taxable income elasticity rather than the Laffer curve.

Some authors modify the Saez (2001) formula for the peak of the Laffer curve without deriving the curve itself. Jacquet & Lehman (2016) account for individual heterogeneity in elasticities and show that the expression in this case is different from the one in Saez (2001). For this reason, I assume that there is no such heterogeneity when I derive the Laffer curve below. Badel & Huggett (2015) set up a dynamic model and derive a formula for the revenue-maximizing rate that depends on three different elasticities. They allow for responses of taxpayers below the top bracket and for impact on other tax bases such as capital income. Saez et al. (2012) derive the revenue-maximizing tax rate when there is income shifting, and Piketty & Saez (2013) consider rent-seeking and migration. While my Laffer curve only includes taxable income responses as in Saez (2001) – and income effects as an extension – in future work, it should be possible to derive Laffer curve expressions for a wider set of effects.

Diamond & Saez (2011) apply the Saez (2001) expression for the United States by setting the taxable income elasticity to 0.25 and the Pareto parameter to 1.5. In the present paper, a similar Pareto parameter of 1.61 and a slightly lower elasticity of 0.2 are used. For the case of Sweden, Pirttilä & Selin (2011) calculate revenue-maximizing rates for high incomes and Sørensen (2010) estimates the degree of self-financing of a cut to the top marginal tax rate, using the expressions derived below. Sørensen uses the exact same numbers for the taxable income elasticity and the Pareto parameter as I do, while Pirttilä and Selin use a lower Pareto parameter because they include both capital and labour income in the income definition. I argue that the Pareto parameter should only be calculated on labour income, as capital income is taxed separately in Sweden.

I am not aware of any papers that specifically analyze the high-income Laffer curve in Sweden, but Stuart (1981) constructs a representative-agent model of the Swedish economy where the household can allocate its labour in taxed or nontaxed sectors. The revenue-maximizing average marginal tax rate (keeping progressivity constant) is found to vary between 43 and 73 percent depending on assumptions about parameter values. Feige & McGee (1983) set up a very similar model but with some extensions, e.g., an endogenous capital stock. In their preferred parameterization they find a revenue-maximizing average tax rate (on both capital and labour) of 58 percent. Both of these papers conclude that Sweden was most likely on the wrong side of the Laffer curve.

3 Theoretical preliminaries

This section intuitively derives expressions for the marginal degree of self-financing and the revenue-maximizing tax rate, both of which are well-known in the literature (e.g., Saez, 2001; Saez et al., 2012; Sørensen, 2010) and of great policy interest. They will be estimated for 27 countries in section 5. A formal derivation is provided in the appendix.

I begin by denoting the top marginal tax rate by τ and the threshold where it starts to apply by b. Revenues from the top tax bracket are then given by

$$R = (\bar{z}_b - b)\tau N, \tag{1}$$

where N is the number of people earning more than b and $\bar{z}_b$ is their average income.

A tax reform will affect the incentive to earn taxable income. Taxpayers should be expected to respond by reducing hours worked or labour effort, increasing the amount of deductions or similarly changing taxable income. The standard measure of taxpayer responses to taxes is the elasticity of taxable income with respect to the net-of-tax rate, defined as

$$\varepsilon = \frac{dz/z}{d(1-\tau)/(1-\tau)} = -\frac{dz}{d\tau}\frac{1-\tau}{z}, \tag{2}$$

where z is taxable income and τ is the marginal tax rate.

A central policy issue is how tax revenues will be affected by a change in the tax rate. Due to behavioural responses, both the number and average income of top-bracket taxpayers will in general depend on the tax rate. However, when considering a small tax reform, changes in the number of high-income taxpayers will be of second-order importance for revenue. This is shown formally in the appendix. Therefore, the derivative with respect to the tax rate is

$$\frac{\partial R}{\partial \tau} = N\left([\bar{z}_b - b] + \tau\frac{d\bar{z}_b}{d\tau}\right) = N\left([\bar{z}_b - b] - \frac{\tau\varepsilon\bar{z}_b}{1-\tau}\right). \tag{3}$$

The first term shows the mechanical fiscal effect of the tax reform, i.e., the change in tax revenue when the tax base is kept constant. The second term captures behavioural responses. Note that the two terms have opposite signs. It is shown in the appendix that ε in this formula is the income-weighted average taxable income elasticity.

We can divide the second term by the first to obtain the marginal degree of self-financing (DSF). In the case of a tax cut, the degree of self-financing is the fraction of the mechanical revenue loss that is recouped through behavioural responses. In the case of a tax hike, it is the proportion of the mechanical revenue gain that is lost due to behavioural responses.

The mechanical change in revenue will depend on the difference between average income and the tax threshold (because this is the part of income where the top tax rate applies), while the behavioural change in revenue will depend solely on average income. The ratio of these two quantities is called the Pareto parameter, denoted α, and is crucial to discussions on high-income taxation:

$$\alpha = \frac{\bar{z}_b}{\bar{z}_b - b}. \tag{4}$$

For example, if the Pareto parameter is three, one-third of the average top-bracket taxpayer's income is subject to the top marginal tax rate. Using the definition above, the DSF can be expressed simply as

$$DSF = -\frac{\frac{dR}{d\tau} - \frac{dR}{d\tau}|_z}{\frac{dR}{d\tau}|_z} = -\frac{\tau\frac{d\bar{z}_b}{d\tau}}{\bar{z}_b - b} = \frac{\alpha\varepsilon\tau}{1-\tau}. \tag{5}$$

Intuitively, the marginal degree of self-financing is increasing in the Pareto parameter because a smaller fraction of average income is in the top tax bracket. It is increasing in the taxable income elasticity because this implies larger behavioural responses. The DSF is increasing in the current tax rate partly because the revenue impact is larger if the initial tax rate is larger (τ in the numerator) and partly because a higher initial tax rate means the net-of-tax rate will be affected more proportionately by a given tax change (τ in the denominator).

At the peak of the Laffer curve, behavioural responses will completely offset the mechanical revenue affect of a small tax cut or increase, i.e., the DSF is 100 percent. Setting equation 5 to 1 and assuming that α and ε are constant, we find the Saez (2001) expression $\tau^* = 1/(1+\alpha\varepsilon)$, as expected. This tax rate is socially optimal if the government does not care about the living standard of high-income earners. This is the case for infinitely high incomes if the government is utilitarian and marginal utility is decreasing, for example.

3.1 Income effects

The analysis can be extended to include income effects by noting from equation 1 that the average high-income individual's change in net income due to a tax reform will be equal to $-[\bar{z}_b - b]d\tau$, i.e., the distance between average income and the top tax bracket threshold, multiplied by the change in the tax rate. This will induce income effects on taxable income. The impact on taxable income can be obtained by multiplying this with the derivative of taxable income with respect to exogenous income, m. Thus the change in taxable income is given by

$$\frac{dz}{d\tau} = -\frac{\varepsilon_c z}{1-\tau} - [z-b]\frac{\partial z}{\partial m} = -\frac{\varepsilon_c z + \eta[z-b]}{1-\tau}, \tag{6}$$

where ε_c is the compensated taxable income elasticity and $\eta = (1-\tau)\partial z/\partial m < 0$ is the income effect parameter.[1] The first term comes directly from the definition of the elasticity and the second term captures income effects. Plugging this into equation 5 yields

$$DSF = \frac{(\alpha\varepsilon_c + \eta)\tau}{1-\tau}. \tag{7}$$

The taxable income elasticity and the income effect parameter are both population averages, but with different weighting; see the appendix for details. We see from the formula that income effects are less important for high-income taxation, as the compensated response is amplified by the Pareto parameter $\alpha > 1$ while the income effect is not. The intuition is that a tax cut for the top tax bracket may increase the incentive to earn taxable income at the margin considerably while net income does not increase much, implying that demand for leisure will not increase much either.

4 An expression for the high-income Laffer curve

In this section, I derive an expression for the Laffer curve for high incomes, i.e., tax revenues from the high-income segment as a function of the top marginal tax rate. In contrast to the expressions derived in the previous section, this formula does not appear in the literature. My derivation requires three assumptions. These assumptions are the same as the ones needed to derive Saez' revenue-maximizing top tax rate discussed above.

First, I assume that the right tail of the potential income distribution – potential incomes that are greater than b – is Pareto distributed. The Pareto distribution is defined such that the mass in the right tail is $1 - F(x) = (k/x)^\alpha$, where F is the cumulative distribution function, k is strictly positive and gives the minimum of the distribution and α is the Pareto parameter. Setting the minimum income to b, the cumulative distribution function of potential incomes F_0 is given by the following equation:

$$1 - F_0(z_0) = N_0 \left(\frac{b}{z_0}\right)^\alpha. \tag{8}$$

The density has been multiplied by N_0, which is the proportion of taxpayers whose potential income exceeds b, i.e., those who would be in the top tax bracket if the tax rate were zero. The population of taxpayers is normalized to one.

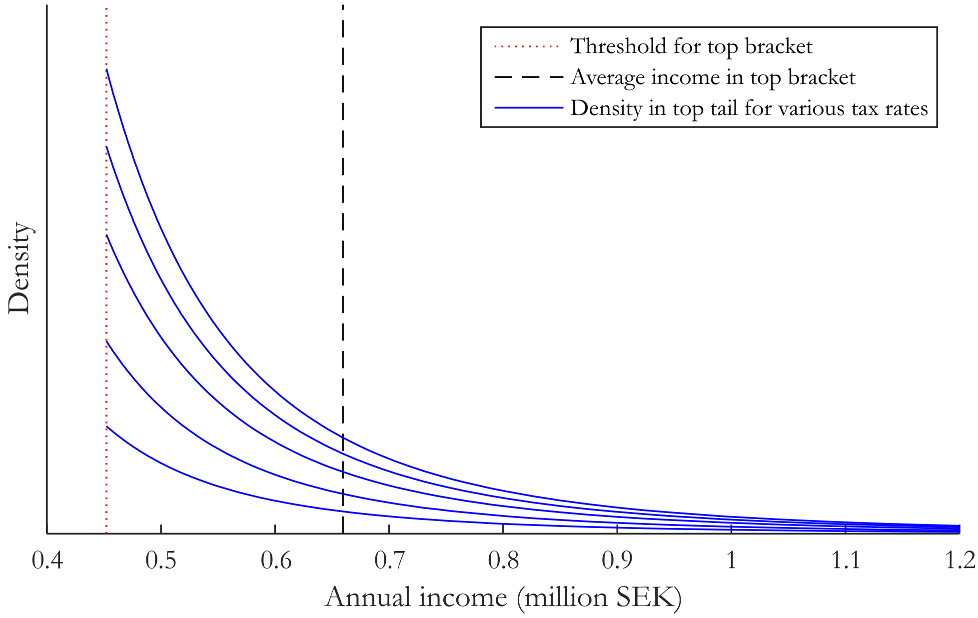
Second, the individual's budget set must be convex, requiring that the top marginal tax rate is also the highest. Given convex preferences, this is sufficient to rule out taxpayers jumping between (interiors of) segments of the tax schedule.[2]

The third assumption is that the taxable income elasticity is constant across individuals and tax rates and that there are no income effects and no extensive margin responses (no fixed costs of working). This implies a quasi-linear isoelastic utility function:

$$u(c,z) = c - \frac{z_0}{1+\frac{1}{\varepsilon}}\left(\frac{z}{z_0}\right)^{1+\frac{1}{\varepsilon}}. \tag{9}$$

[1] The income effect parameter is related to the compensated and uncompensated elasticities through the Slutsky equation: $\varepsilon_u = \varepsilon_c + \eta$.

[2] See Saez (2001), p. 217.



Note: The Pareto parameter is 3.18 and the taxable income elasticity is 0.2. Densities are shown for the tax rates 0 (i.e., potential income), 25, 50, 75 and 90 percent. Higher tax rates have lower densities.

Figure 1: Example of how the right tail of the income distribution varies with the top marginal tax rate

Potential labour income, i.e., income in the absence of taxation, is denoted z_0, consumption c and virtual income y.[3] The individual maximizes utility subject to a budget constraint $c = (1-\tau)z + y$. This gives a taxable income supply function

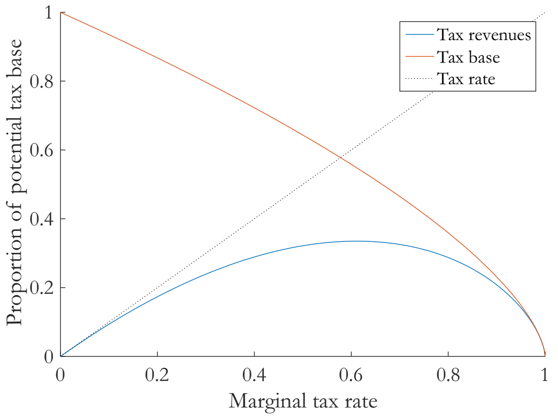
$$z = z_0(1-\tau)^{\varepsilon}. \tag{10}$$

This implies that there is a one-to-one mapping between potential income and realized income. For each marginal tax rate, i.e., within each segment of the tax schedule, there will be a multiplicative relationship between z_0 and z. Incomes will therefore be Pareto-distributed (within each segment of the tax schedule) if the potential income distribution is, and the Pareto parameter of the income distribution will be the same as the Pareto parameter of the potential income distribution. This in turn implies that the Pareto parameter will be independent of the tax rate.[4]

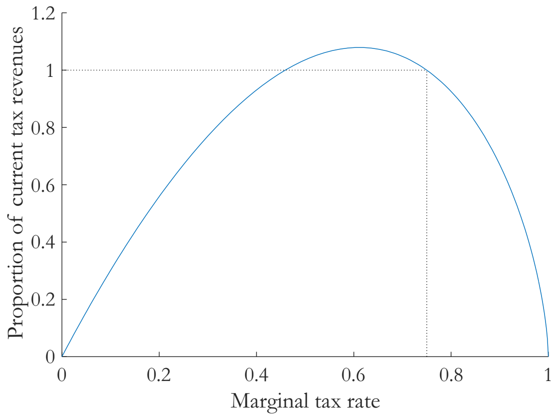
Next, it is crucial to note that the Pareto parameter (equation 4) is a function of the average income of top-bracket taxpayers ($\bar{z}_b$). Therefore, keeping the Pareto parameter constant requires that $\bar{z}_b$ is also constant. Instead, the number of people in the top tax bracket, N, must change after a tax reform. In section 3, it was noted that changes in N are only of second-order importance for revenue and could be ignored when analyzing small tax reforms. For non-marginal tax changes, however, changes in N must be considered. Figure 1 shows an example for the Swedish case ($b = 452{,}100$ and $\bar{z}_b = 659{,}000$, implying $\alpha = 3.18$), where total density in the top tail will vary with the tax rate while average income in the top tax bracket is constant. If the marginal tax schedule is piecewise linear

[3] Virtual income is given by $y = \tau z - T(z)$ and is a way of linearizing the budget constraint around a given segment on the tax schedule.

[4] See the discussion in Saez (2001), p. 212.



(a) The tax base and tax revenues as a function of the top marginal tax rate, expressed as a proportion of the potential tax base.



(b) Tax revenues from the top tax bracket as a function of the top marginal tax rate, expressed as a proportion of inital tax revenues, given an initial tax rate of 75 percent.

Figure 2: High-income Laffer curves for Sweden

and increasing, the model predicts bunching of taxpayers at the kink point b.[5] N will change as individuals move between the kink and the top income segment.

By inverting the taxable income supply function (equation 10), we find that everyone whose potential income is greater than $b_0 = b/(1-\tau)^{\varepsilon}$ will be in the top tax bracket if the tax rate is set to τ. Plugging this into equation 8, we see that $N(\tau) = 1 - F_0(b_0) = N_0(1-\tau)^{\alpha\varepsilon}$. Substituting this into equation 1, we conclude that the high-income Laffer curve is given by

$$R(\tau) = N_0(\bar{z}_b - b)\tau(1-\tau)^{\alpha\varepsilon}. \tag{11}$$

$N_0(\bar{z}_b - b)$ is the potential tax base, i.e., total income in the top tax bracket in the absence of taxation. Recall that $\bar{z}_b$ is independent of the tax rate. It can be verified that the maximum of the curve is given by the Saez top tax formula $\tau = 1/(1+\alpha\varepsilon)$. Tax revenues are zero at tax rates of 0 and 100 percent, as expected.[6]

In figure 2a, I plot tax revenues as a proportion of the potential tax base for the case of $\varepsilon = 0.2$ and $\alpha = 3.18$. I also plot the tax base for each tax level as a proportion of the potential tax base. The 45-degree line indicates what tax revenues would have been in the absence of behavioural responses.

Because the potential tax base is not observed, it may be more useful to express the Laffer curve as the ratio of post-reform to pre-reform tax revenue:

$$\frac{R(\tau_2)}{R(\tau_1)} = \frac{\tau_2}{\tau_1}\left(\frac{1-\tau_2}{1-\tau_1}\right)^{\alpha\varepsilon}. \tag{12}$$

In figure 2b, I use this formula to plot the Swedish high-income Laffer curve, using the same values as in section 5 ($\tau_1 = 0.75$, $\varepsilon = 0.2$, $\alpha = 3.18$).

[5]Observed bunching in Sweden, which also can be spotted in figure 5b, is very small, as shown by Bastani & Selin (2014). Chetty (2012) shows that quite small optimization frictions can reconcile the virtual absence of bunching with the elasticities estimated in the quasiexperimental literature. One way of thinking about it is that there are a number of latent bunchers in the vicinity of the kink point, but that frictions cause these taxpayers to miss the kink.

[6]See the appendix for an alternative derivation of equation 11.

4.1 Income effects

Through the use of differential equations, the Laffer curve can be derived in a more direct but less elegant way. In section 3, I derived an expression for the marginal degree of self-financing, which is related to the slope of the Laffer curve. If one knows the slope of the Laffer curve at each point, it is possible to trace out the curve itself. Note that equation 5 can be rewritten

$$DSF = 1 - \frac{\frac{dR}{d\tau}}{\frac{dR}{d\tau}|_z} = 1 - \frac{dR/R}{d\tau/\tau}. \tag{13}$$

The last step uses the fact that tax revenues can be expressed $R = \tau Z$, where Z is the tax base. In a mechanical calculation (holding Z constant), $dR/d\tau = Z = R/\tau$.

Without income effects, $DSF = \alpha\varepsilon\tau/(1-\tau)$. This together with equation 13 constitutes a differential equation in R and τ, which can be solved assuming that α and ε are constant even for large changes in τ. As noted above, the Pareto parameter is indeed independent of the tax rate if high potential incomes follow a Pareto distribution. The differential equation has the general solution $R(\tau) = C\tau\,(1-\tau)^{\alpha\varepsilon}$. Dividing both sides by τ and letting $\tau \to 0$, we see that the constant is the potential tax base, as expected. Thus we have derived equation 11.

This approach can be used to obtain an expression for the Laffer curve with income effects. The method in the previous section cannot be used to derive a Laffer curve with income effects because no explicit expression for the taxable income supply function exists for this class of utility functions unless the utility function is quasilinear, implying no income effects. Applying the differential-equation method to equation 7, the general solution is $R(\tau) = C\tau(1-\tau)^{\alpha\varepsilon_c+\eta}$. Plugging in the potential tax base, we find that the Laffer curve with income effects is given by

$$R(\tau) = N_0(\bar{z}_b - b)\tau(1-\tau)^{\alpha\varepsilon_c+\eta}. \tag{14}$$

This requires that ε_c and η are constant. The income effect parameter η will not in general be constant, however, so this is only an approximation. As η is negative, including income effects will shift the Laffer curve somewhat to the right. The maximum occurs at $\tau = 1/(1+\alpha\varepsilon_c+\eta)$, an expression that is also derived by Saez (2001).

5 Laffer curves in OECD countries

The Laffer curve expression, along with expressions for the revenue-maximizing rate and the degree of self-financing, is a powerful tool for analyzing top-income taxation with minimal data requirements. To illustrate this, I draw Laffer curves for 27 OECD countries. Three parameters are needed for each country: the elasticity of taxable income, the effective top marginal tax rate and the Pareto parameter.

A large literature in public economics uses tax reforms as identifying variation to estimate the taxable income elasticity. Piketty & Saez (2013) write that "most estimates of aggregate elasticities of taxable income are between 0.1 and 0.4 with 0.25 perhaps being

Table 1: Taxation of high incomes in 27 OECD countries

Country	**Pareto parameter**	**Current top tax rate**	**Laffer curve peak**	**Degree of self-financing**
Australia	1.86	55%	73%	46%
Austria	3.14	63%	61%	108%
Belgium	2.03	74%	71%	115%
Canada	1.83	58%	73%	50%
Czech Republic	2.95	46%	63%	50%
Denmark	3.04	66%	62%	119%
Finland	2.40	72%	68%	123%
France	2.20	69%	69%	96%
Germany	1.66	57%	75%	44%
Greece	2.30	53%	68%	51%
Ireland	1.98	64%	72%	71%
Israel	2.97	58%	63%	83%
Italy	2.18	55%	70%	54%
Japan	2.37	60%	68%	70%
Luxembourg	3.39	59%	60%	96%
Mexico	2.23	39%	69%	28%
Netherlands	3.35	59%	60%	97%
New Zealand	2.10	44%	70%	32%
Norway	2.02	63%	71%	70%
Poland	3.25	47%	61%	58%
Slovakia	2.71	36%	65%	30%
South Korea	1.81	49%	73%	35%
Spain	2.08	52%	71%	45%
Sweden	3.18	75%	61%	195%
Switzerland	1.73	51%	74%	36%
United Kingdom	1.79	59%	74%	52%
United States	1.61	48%	76%	30%

Note: A taxable income elasticity of 0.2 is assumed. Source: See appendix.

a reasonable estimate". Chetty (2012) calculates that an elasticity of 0.33 is consistent with several central papers. It is conceivable that the elasticity varies over the income distribution, and some studies report higher elasticities for top incomes (e.g., Gruber & Saez, 2002). However, in the absence of strong evidence of heterogeneous elasticities, I will follow the convention in the literature and assume a constant elasticity. It is also possible for the elasticity to vary across countries, due to institutional or cultural differences, but the literature is not rich enough to provide credible estimates for all the countries studied. It is important to note that this literature only captures the response in the first few years after a tax reform, thus potentially more important long-term responses like human capital accumulation and career choices are missed. When considering fiscal effects, it is sensible to use a somewhat lower elasticity to account for the fact that some taxable income responses may be due to, e.g., converting labour income into capital income (see the discussion in Lundberg, 2017). For this reason, I assume the elasticity of taxable income to be 0.2.

Country-level data on effective marginal tax rates is not readily available. For this reason, effective marginal tax rates are calculated from data on income tax rates, social contributions and consumption taxes.[7] When applicable, the deductibility of employees' social contributions is accounted for. The consumption tax rate is calculated by dividing VAT, sales tax and excise tax receipts by total consumption, excluding wage outlays by the public sector. It is thus assumed that high-income earners face the same effective consumption tax rate as the population at large. Because of this procedure and because all tax rates are not from the same year, the calculated effective tax rates should be interpreted with some caution. All details are given in the appendix. The highest top effective tax rate is found in Sweden, at 75 percent[8], while Slovakia has the lowest at 36 percent.

A Pareto parameter is needed for each country in order to draw the Laffer curve. If the assumption of Pareto-distributed potential incomes is true for all countries, the Pareto parameter will not be endogenous to the tax rate, but will reflect intrinsic inequality in earnings potential, due to, e.g., differences in human capital. It is well known in the optimal taxation literature that the shape of the skill distribution is of central importance for optimal tax policy. The two main sources for the Pareto parameter are the World Wealth and Income Database (WID) and the Luxembourg Income Study (LIS). In a few cases, credible country-specific sources – that take into account which incomes are included in the tax base for the country concerned – are used. Pareto parameters are estimated on individual-level data on labour income from the LIS. In the WID, Pareto parameters are calculated from the income share of the top one percent using the formula given by Atkinson et al. (2010, p. 753); when estimating on LIS data, equation 4 is applied to the top five percent.[9] To the extent that top incomes are not Pareto-distributed, the size of the Pareto parameter will depend on the method and cut-off used for its estimation; this is a source of uncertainty.

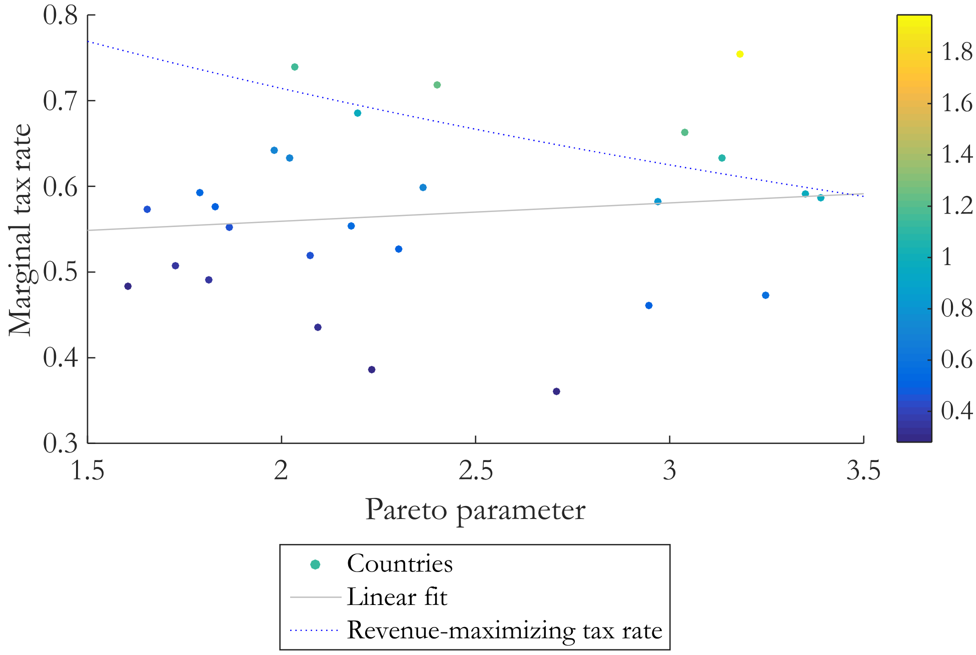
A full table of Pareto parameters can be found in the appendix. Note that WID parameters are always lower than the LIS estimates and that the discrepancy is quite large in some cases. The difference in cut-off points may explain some of this. Another possible explanation is that while the WID estimates pertain to all income, the LIS parameters are estimated on labour income only. One can argue that it is more appropriate to use labour income Pareto parameters, especially for countries – such as the Nordic countries – where capital income is taxed separately from labour income. It should also be noted that the more reliable country-specific sources are closer to the LIS numbers. However, in the interest of conservatism, the WID is used whenever data is available. The order of priority is thus (1) country-specific sources, (2) the WID and (3) the LIS. Because of the uncertainty surrounding the Pareto parameters, the specific source for this parameter should be considered before drawing policy conclusions about a particular country.

Pareto parameters and effective marginal tax rates thus estimated are shown in table 1.

[7] I am grateful to Alexander Fritz Englund for valuable research assistance in computing effective tax rates, and to Timbro for financial support. The compilation of marginal tax rates has been published separately as Fritz Englund & Lundberg (2017).

[8] This includes the effects of the EITC phase-out, which does not raise the marginal tax rates of those with very high incomes. Because most high-income taxpayers are in the EITC phase-out region, I include it in the effective marginal tax rate.

[9] The 95th percentile of strictly positive incomes for each country was used as income threshold (b in equation 4). The data was examined for signs of top-coding. Only the Norwegian data showed clear evidence of top-coding.



Note: The revenue-maximizing rate is drawn given a taxable income elasticity of 0.2. The colour of the dot indicates the degree of self-financing (legend on the right).

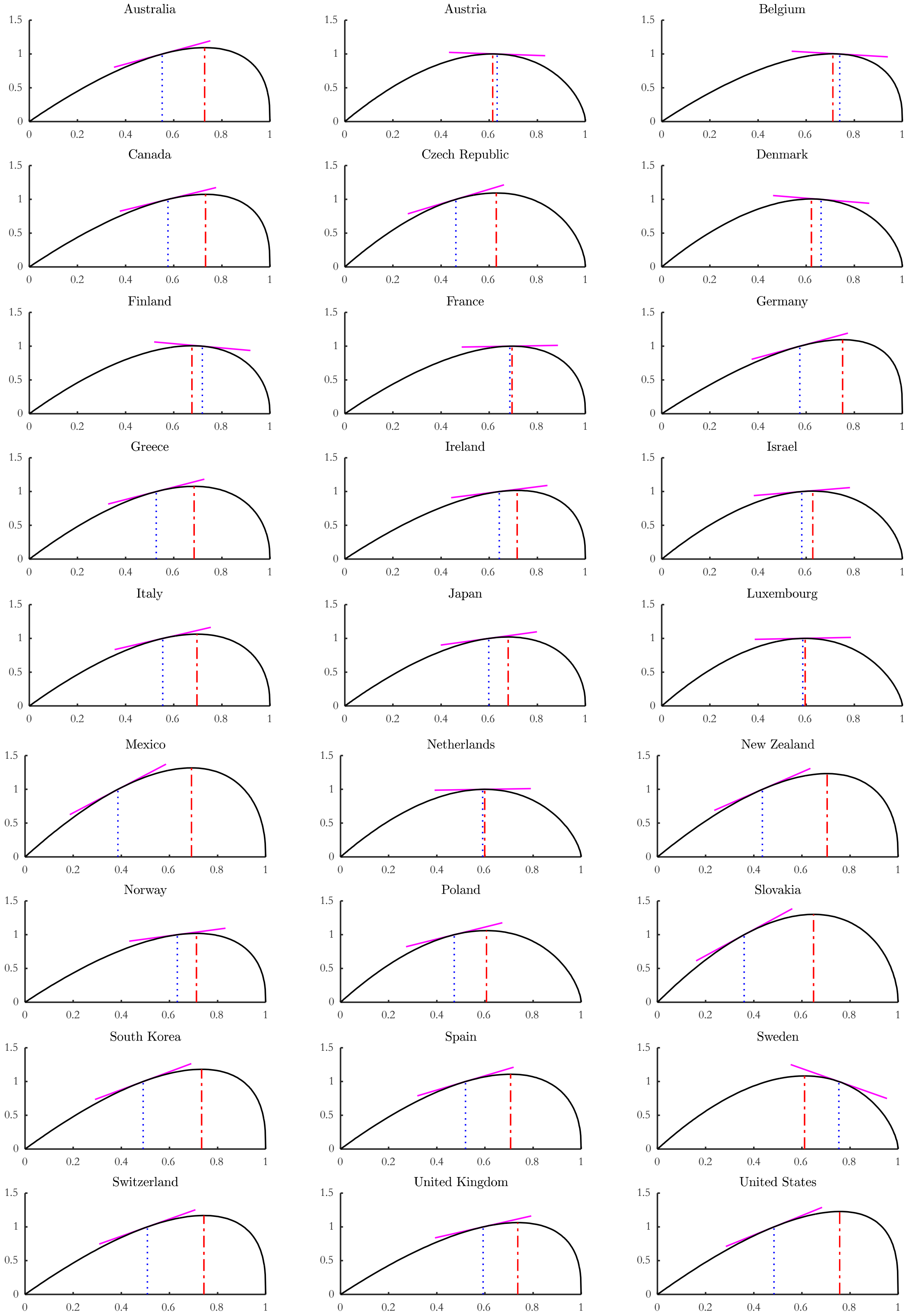
Figure 3: The high-income marginal tax rate and Pareto parameter in 27 OECD countries

Also shown are revenue-maximizing top marginal tax rates and degrees of self-financing (see section 3). The Pareto parameters range from 1.61 (the United States) to 3.39 (Luxembourg). Revenue-maximizing rates accordingly range between 60 and 76 percent. In five cases, the country is estimated to be on the wrong side of the Laffer curve, implying a degree of self-financing exceeding 100 percent. The average effective marginal tax rate is 57 percent, while the average estimated revenue-maximizing tax rate is 68 percent. The average degree of self-financing is 70 percent. These results hold if the Pareto parameters used are accurate and if the true taxable income elasticity is indeed 0.2. As the degree of self-financing is directly proportional to both the Pareto parameter and the elasticity, it is easy for the reader to perform robustness checks.[10]

A scatterplot of Pareto parameters and effective marginal tax rates is shown in figure 3. Instead of the expected negative relationship, there is a slight positive correlation between the top tax rate and the Pareto parameter. Again, the data issues surrounding the Pareto parameters should be considered before drawing conclusions from this.

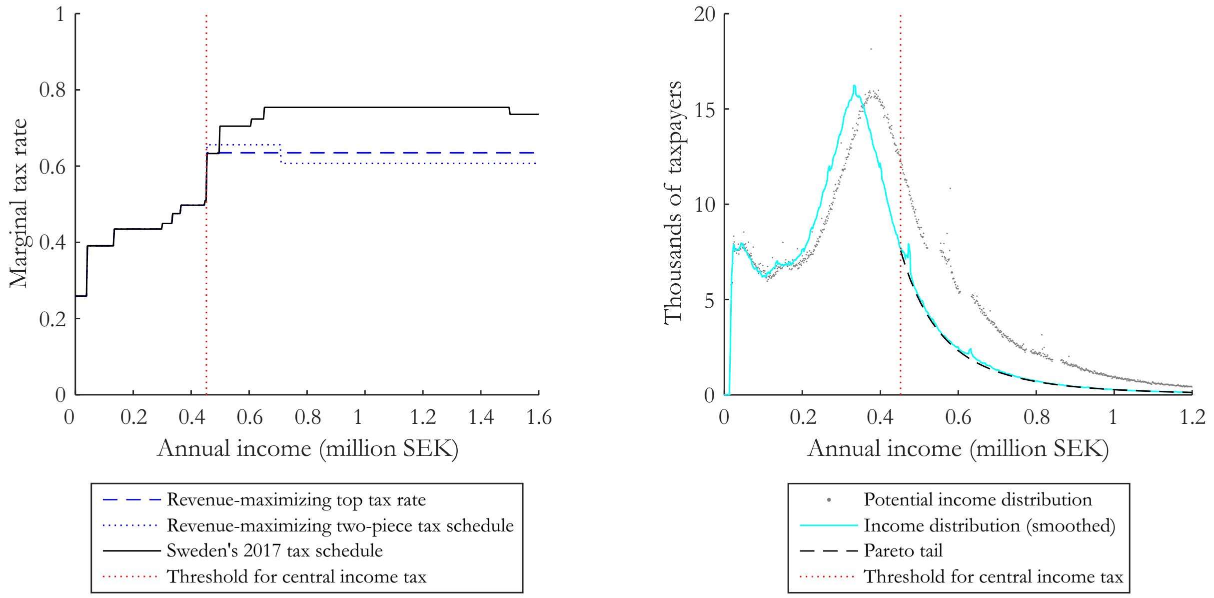
Laffer curves for 27 OECD countries are shown in figure 4. Tax revenues are expressed as a multiple of current tax revenues (i.e., equation 12 is used). The parameters in table 1 are used to draw the curves. We see that countries with low Pareto parameters, such as the United States, are skewed to the right, the peak occurring at higher tax rates.

[10]For example, if the elasticity is 0.1 all countries are to the left of the Laffer curve peak (Sweden being very close to the top). If the elasticity is 0.3, however, 12 countries are on the downward-sloping part of the curve and 17 countries are if the elasticity is 0.4. An elasticity of 0.75 is required for all countries to have surpassed the revenue-maximizing rate.



Note: The horizontal axis shows the marginal tax rate and the vertical axis shows tax revenues as a multiple of current tax revenues (equation 12). The current (dotted) and revenue-maximizing (dotted–dashed) tax rates are indicated. I also plot the slope at the current tax rate (which is related to the degree of self-financing).

Figure 4: Laffer curves in 27 OECD countries



(a) Effective marginal tax rates by income. (b) Actual and potential income distributions.

Figure 5: The Swedish marginal tax schedule and labour income distribution

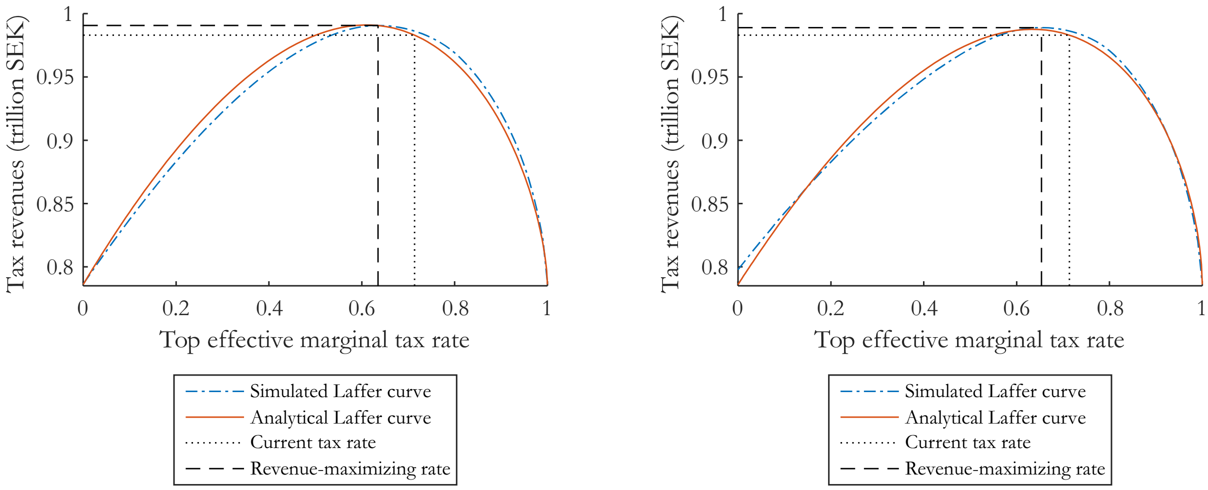
6 Simulations

In this section, I draw high-income Laffer curves for Sweden by running microsimulations on full-population register data. This allows me to test some of the assumptions required to derive the expression for the Laffer curve presented above. The model is equivalent to the Swedish Labour Income Microsimulation Model (SLIMM) described by Lundberg (2017), but without income effects (except in section 6.2) and participation responses.

It is well known that individuals' preferences are crucial to the size of behavioural responses and therefore the shape of the Laffer curve. I continue to use a quasilinear and isoelastic utility function (equation 9). Consequently, this is not an assumption that is tested. It was noted in the previous section that a taxable income elasticity of 0.2 is a reasonable or somewhat conservative midpoint of the international literature. This also seems to be the case for the literature on the Swedish taxable income elasticity; see Sørensen (2010), Pirttilä & Selin (2011) and Ericson et al. (2015) for surveys and Lundberg (2017) for a discussion on optimization frictions and fiscal externalities. Hence I continue to use 0.2 as elasticity. The two assumptions that I can assess the importance of are the assumption that high potential incomes are exactly Pareto-distributed and the assumption that marginal tax rates are increasing.

The income data used is the 2013 distribution of labour incomes in Sweden, constructed from Statistics Sweden's population-wide register data. This is scaled up by employment and nominal wage growth between 2013 and 2017. It is interesting to note from figure 5b that high incomes – starting around the threshold for central government income tax (452,100 SEK per year[11]) – are very well approximated by a Pareto distribution with a Pareto parameter of 3.18. The same conclusion is reached by Bastani & Lundberg (2016). They show that a quantity termed the local Pareto parameter (see their figure 12) is

[11]The current exchange rate is 9 SEK/USD.



(a) Without income effects.

(b) With income effects.

Figure 6: High-income Laffer curves in Sweden in 2017, i.e., total tax revenues from labour income as a function of the effective marginal tax rate on high incomes

remarkably stable from about SEK 400,000–500,000 per year, implying that incomes are close to being exactly Pareto-distributed.

The effective marginal tax rate in 2017 for each income level is then calculated – see figure 5a. This is made up of central and municipal income tax, the tax part of social security contributions and VAT and excise taxes (assumed to be 19 percent of income for all income levels).[12] The highest effective marginal tax rate is 75 percent.

Using the marginal tax rates, the distribution of potential incomes is obtained by inverting the taxable income supply function (equation 10): $z_0 = z/(1-\tau)^{\varepsilon}$. Because the income distribution is mostly smooth while the marginal tax schedule has discontinuities, it follows from this functional form that the potential income distribution has holes where the marginal tax rate jumps. This can be seen in figure 5b. The flipside is that if the potential income distribution were smooth, the observed income distribution would feature spikes at the kink points of the tax schedule. In reality, very little such bunching is observed. (Bastani & Selin, 2014) This is usually explained by the presence of optimization frictions, i.e., adjustment costs or the like that prevent individuals from attaining the full optimum. The present model does not feature optimization frictions as it would be difficult to unscramble optimization errors and identify potential income when the taxable income supply function contains a random element.

In the simulations as described above, total labour income (including social security contributions) is SEK 2.2 trillion and total tax revenue is 1 trillion. About five million people earned some labour income during the year. One million of these paid central government income tax. Total potential labour income is SEK 2.6 trillion. This means that total income would increase by 19 percent if all labour taxation were abolished.

In order to draw Laffer curves, I let individuals maximize utility (equation 9) given a counterfactual tax schedule where incomes over the threshold for central government income tax are subject to a constant effective marginal tax rate ranging from 0 to 100

[12] See Lundberg (2017) for details on how the tax component of social contributions and the consumption tax rate are calculated.

percent. This income region is suitable for testing the Laffer curve expression because the income distribution is well approximated by a Pareto distribution here. For each tax rate, I calculate total tax revenue from labour income. The result is shown in figure 6a, along with an analytical Laffer curve drawn using equation 12.[13] Overall, the two curves are very similar. This indicates that the assumptions made in section 4 are not far from reality, given the utility function used. The analytical Laffer curve peaks at the Saez top rate $\tau = 1/(1+\alpha\varepsilon) = 1/(1+3.18\times 0.2) = 61.1\%$, while the simulated Laffer curve peaks at 63.5 percent (also shown in figure 5a). The difference is explained by the fact that potential incomes are not exactly Pareto-distributed. The simulations imply that lowering the top tax rate to 63.5 percent would increase tax revenue by SEK 7.6 billion. A mechanical calculation yields a revenue shortfall of SEK 22 billion, implying that the reform would have a degree of self-financing of 135 percent.

6.1 Two-piece top tax bracket

The simulations so far have been restricted to consider a single tax rate for high incomes. If potential incomes are exactly Pareto distributed (and the elasticity is constant), the revenue-maximizing tax schedule will indeed be linear for high incomes, because the Pareto parameter is unchanged regardless of the value of b. As a test of this, I allow for two tax rates for incomes over SEK 452,100, the current threshold for central government income tax. Denoting this threshold by b_1, I introduce a second threshold $b_2 > b_1$. A tax rate of τ_1 applies between b_1 and b_2, and τ_2 applies above b_2. Maximizing over τ_1, b_2 and τ_2, I find the two-piece tax schedule for top incomes that would maximize tax revenue. As illustrated in figure 5a, this has a tax rate at 66 percent from the central tax threshold up to SEK 710,000 per year, and after that 61 percent. Adopting this tax schedule is estimated to raise SEK 7.9 billion in additional tax revenue – only 300 million more compared to the case with only one tax bracket for high incomes. As incomes are approximately Pareto-distributed, it is not surprising that not much is gained by allowing for a second tax bracket.

6.2 Income effects

Proceeding to add income effects, I consider a utility function of the following form:

$$u(c,z) = \frac{c^{1-\gamma}}{1-\gamma} - \frac{z_0}{1+\frac{1}{e}}\left(\frac{z}{z_0}\right)^{1+\frac{1}{e}}, \tag{15}$$

where e is the Frisch elasticity of labour supply and γ is approximately the ratio of income effects to the compensated response. The Frisch elasticity is the elasticity of labour supply holding the marginal utility of consumption constant. It is approximately equal to the compensated elasticity of taxable income. The parameter z_0 will no longer have the interpretation of potential income, but will be related to earnings capacity (see the appendix). In order to target a taxable income elasticity of 0.2 and an income effect

[13] Current tax revenue from the top tax bracket (the part of incomes that exceeds SEK 451,200) is SEK 197 billion. The current tax rate (τ_1 in equation 12) is set to 71 percent, which is the average marginal tax rate for those who pay central government income tax. Tax revenue from lower tax brackets (assumed constant at the current SEK 786 billion) is then added to obtain total tax revenue.

parameter (defined by $\eta = (1-\tau)\partial z/\partial m$) of around or slightly below –0.1, I set $e = 0.23$ and $\gamma = 0.5$. Such income effects are approximately in line with the results of Cesarini et al. (2015), who use Swedish lotteries to estimate a marginal propensity to earn out of unearned income (which is the same as the income effect parameter) of −0.11 in a calibrated model; note that this includes income effects on the extensive margin as well, which are not applicable in this setting because disposable income out of work is unaffected by labour tax reforms. This is discussed in detail by Lundberg (2017). In the simulations it is assumed that there is no non-labour income. This is of no practical importance because it is individuals' behaviour, as measured by the taxable income elasticity and the income effect parameter, that matters – not the exact parameterization of the utility function. By increasing everyone's non-labour income from zero to one percent of their potential income, the average income effect parameter is numerically calculated to be −0.083. Similarly, I increase every taxpayer's net-of-tax rate by one percent and find an average uncompensated elasticity of 0.124, implying a compensated elasticity of 0.207.

As no analytical expression for the taxable income supply function exists, I calculate and invert it numerically in order to map observed incomes into a distribution of z_0, i.e., this distribution is calibrated such that individual optimization returns exactly the observed income distribution.[14] I then proceed as above by letting individuals maximize utility while facing counterfactual tax schedules with the top tax rate varying between 0 and 100 percent. The resulting Laffer curve is shown in figure 6b. As expected, the curve is not very different from the case without income effects. The peak of the simulated curve occurs at 65 percent, while the analytical curve (see section 4.1) peaks at $1/(1+\alpha\varepsilon_c+\eta) = 63\%$, for the numerically calculated parameter values discussed above.

7 Conclusion

The main contribution of this paper is the derivation of an expression for the Laffer curve for high labour incomes of the form $R = \tau(1-\tau)^{\alpha\varepsilon}$ and the testing of this expression by way of microsimulations. The derivation requires a constant Pareto parameter α and taxable income elasticity ε. This analytical expression allows the calculation of the fiscal impact of tax reforms with minimal data requirements. Its peak is given by $\tau = 1/(1+\alpha\varepsilon)$ and the degree of self-financing of a small tax cut is $\alpha\varepsilon\tau/(1-\tau)$, both of which are well-known expressions in the literature.

A simulation exercise using Swedish population data yields Laffer curves that are very similar to the ones drawn using the analytical expression. This is done by hypothetically altering the effective marginal tax rate for the richest fifth of working Swedes, i.e., those that are subject to central government income tax, and letting individuals maximize utility given these counterfactual tax schedules. The simulated high-income Laffer curve peaks at 64 percent while the analytical Laffer curve peaks at 61 percent. This implies that the assumptions behind the theoretically derived Laffer curve are not too restrictive, for a given elasticity of taxable income.

Swedish top incomes are shown to follow a Pareto distribution closely. Thus the revenue-maximizing tax schedule for high incomes is well approximated by a single tax rate. Allowing for two high-income tax brackets increases potential tax revenue by only 0.03 percent.

[14]The Matlab functions `fzero` and `fminbnd` are used for the inversion.

Theoretically, income effects should affect these conclusions little, as a lower marginal tax rate does not raise net income by much – and thus does not increase the demand for leisure by much – in a high-Pareto-parameter country like Sweden. This prediction is supported by the simulations, where the revenue-maximizing top tax rate increases only slightly, to 65 percent, when accounting for income effects of reasonable magnitude.

By assembling a large country-level dataset on Pareto parameters and effective marginal tax rates, I am able to draw top-income Laffer curves for 27 OECD countries. More work is needed to explain the discrepancy between data sources on the magnitude of the Pareto parameter. Therefore one should be careful with drawing policy conclusions about specific countries. Having this in mind, the average top effective marginal tax rate in the dataset is 57 percent. This can be contrasted with an estimated revenue-maximizing tax rate of 68 percent if the elasticity of taxable income is 0.2. The average degree of self-financing is 70 percent, but the range is wide: from a low of 28 percent (Mexico) to a high of 195 percent (Sweden).

8 References

Alvaredo, Facundo, Atkinson, Anthony B., Piketty, Thomas, Saez, Emmanuel & Zucman, Gabriel (2016), "The World Wealth and Income Database". <http://www.wid.world>. Accessed 2016-06-28.

Atkinson, Anthony B., Piketty, Thomas & Saez, Emmanuel (2010), "Top incomes in the long run of history", in Atkinson, Anthony B. & Piketty, Thomas (eds.), *Top Incomes: A Global Perspective*. Oxford: Oxford University Press.

Badel, Alejandro (2013), "Higher taxes for top earners: Can they really increase revenue?", *The Regional Economist*, 21 (4).

Badel, Alejandro & Huggett, Mark (2014), "Taxing top earners: A human capital perspective", Federal Reserve Bank of St. Louis Working Paper 2014-017B.

Badel, Alejandro & Huggett, Mark (2015), "The sufficient statistic approach: Predicting the top of the laffer curve", Federal Reserve Bank of St. Louis Working Paper 2015-038A.

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

Brøns-Petersen, Otto (2016), "Derfor er topskatten så selvfinansierende at afskaffe", Punditokraterne. <http://punditokraterne.dk/2016/07/28/derfor-er-topskatten-saa-selvfinansierende-at-afskaffe>.

Cesarini, David, Lindqvist, Erik, Notowidigdo, Matthew J., & Östling, Robert (2015), "The effect of wealth on individual and household labor supply: Evidence from swedish lotteries", NBER Working Paper 21,762.

Chetty, Raj (2012), "Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply", *Econometrica*, 80 (3).

Diamond, Peter (1998), "Optimal income taxation: An example with a u-shaped pattern of optimal marginal tax rates", *American Economic Review*, 88 (1).

Diamond, Peter & Saez, Emmanuel (2011), "The case for a progressive tax: From basic research to policy recommendations", *Journal of Economic Perspectives*, 25 (4).

Ericson, Peter, Flood, Lennart & Islam, Nizamul (2015), "Taxes, wages and working hours", *Empirical Economics*, 49 (2).

European Union (2015), "Taxation trends in the European Union: Data for the EU member states, Iceland and Norway". Luxembourg: Publications Office of the EU.

Feige, Edgar L. & McGee, Robert T. (1983), "Sweden's Laffer curve: Taxation and the unobserved economy", *Scandinavian Journal of Economics*, 85 (4).

Fritz Englund, Alexander & Lundberg, Jacob (2017), "Effective marginal tax rates: An international comparison", report, Epicenter.

Giertz, Seth H. (2009), "The elasticity of taxable income: Influences on economic efficiency and tax revenues, and implications for tax policy", in Viard, Alan D. (ed.), *Tax Policy Lessons from the 2000s*. Washington, D.C.: AEI Press.

Gruber, Jon & Saez, Emmanuel (2002), "The elasticity of taxable income: Evidence and implications", *Journal of Public Economics*, 84.

Jacquet, Laurence & Lehmann, Etienne (2016), "Optimal taxation with heterogeneous skills and elasticities: Structural and sufficient statistics approaches", THEMA Working Paper No 2016-04.

Keane, Michael P. (2011), "Labor supply and taxes: A survey", *Journal of Economic Literature*, 49 (4).

KPMG (2016), "Individual income tax rates table". <https://home.kpmg.com/xx/en/home/services/tax/tax-tools-and-resources/tax-rates-online/individual-income-tax-rates-table.html>. Accessed 2016-08-17.

Laffer, Arthur B. (2004), "The Laffer curve: Past, present, and future", Backgrounder no. 1765, June 1. The Heritage Foundation. <http://www.heritage.org/research/reports/2004/06/the-laffer-curve-past-present-and-future>.

LIS Cross-National Data Center (2016), "Luxembourg Income Study Database" (multiple countries). Luxembourg: LIS Cross-National Data Center. <www.lisdatacenter.org>. Accessed 2016-06-30.

Lundberg, Jacob (2017), "Analyzing tax reforms using the Swedish Labour Income Microsimulation Model", mimeo, Uppsala University.

Mendoza, Enrique G., Razin, Assaf & Tesar, Linda L. (1994), "Effective tax rates in macroeconomics: Cross-country estimates of tax rates on factor incomes and consumption", *Journal of Monetary Economics*, 34 (3).

Norwegian Tax Administration (2016), "Calculation of Norwegian income taxes – income year 2016". <https://www.skatteetaten.no/globalassets/taxnorway/2016/temabrev/2016-09-utregning-av-norsk-inntektsskatt-eng.pdf>.

Piketty, Thomas & Saez, Emmanuel (2013), "Optimal labor income taxation", *Handbook of Public Economics, Volume 5*. Amsterdam: Elsevier.

Pirttilä, Jukka & Selin, Håkan (2011), "Tax policy and employment: How does the Swedish system fare?", Uppsala Center for Fiscal Studies Working Paper 2011:2.

Portuguese Tax and Customs Authority (2010), "The Tax Code – Income and Gains of Individuals, Decree-law No-442-A/88, of 30 November". Translated by William Cunningham. <http://info.portaldasfinancas.gov.pt/NR/rdonlyres/FA120865-C4A3-4383-AA12-63C97A94DB2A/0/IRS_code_with_OE2010_PEC_changes.pdf>.

PWC (2016), “Worldwide Tax Summaries Online” <http://taxsummaries.pwc.com/>. Accessed 2016-08-17.

Riihelä, Marja, Sullström, Risto & Tuomala, Matti (2014), “Top incomes and top tax rates: Implications for optimal taxation of top incomes in Finland”, Tampere Economic Working Paper 88.

Saez, Emmanuel (2001), “Using elasticities to derive optimal income tax rates”, *Review of Economic Studies*, 68 (1).

Saez, Emmanuel, Slemrod, Joel & Giertz, Seth H. (2012), “The elasticity of taxable income with respect to marginal tax rates: A critical review”, *Journal of Economic Literature*, 50 (1).

Stuart, Charles E. (1981), “Swedish tax rates, labor supply, and tax revenues”, *Journal of Political Economy*, 89 (5).

Sørensen, Peter Birch (2010), *Swedish Tax Policy: Recent Trends and Future Challenges.* Report to the Expert Group on Public Economics (ESO), 2010:4.

Usher, Dan (2014), “How High Might the Revenue-maximizing Tax Rate Be?”, Queen’s Economics Department Working Paper No. 1334.

Zoutman, Floris, Jacobs, Bas & Jongen, L.W. Egbert (2014), “Optimal redistributive taxes and redistributive preferences in the netherlands”, mimeo. <https://dl.dropboxusercontent.com/u/55134859/Website%20files/Zoutman2014Optimal2.pdf>.

Appendix

Formal derivations

This section shows how the theoretical results in section 3 can be derived more formally.

We consider a two-piece tax schedule with a kink point b. Incomes over this threshold are taxed at a rate τ. Without loss of generality, the tax rate in the first bracket is set to zero. The main assumption is that the marginal tax schedule is increasing.[15] Thus, the tax function is given by

$$T(z) = \max\{0, \tau(z-b)\},$$

where z is taxable income.

Individuals are heterogeneous in earnings capacity, z_0. Income increases monotonically with earnings capacity and all individuals with the same earnings capacity have the same income. Individuals maximize a utility function $u(c, z; z_0)$ subject to a budget constraint $c = z - T(z) + m$, where c is consumption and m is exogenous income.

Individuals will locate in the first segment of the tax schedule, on the kink (bunching) or in the second segment. For those in the second segment – those whose incomes strictly exceed b – virtual income is given by $y(\tau) = \tau z - T(z) + m = \tau b + m$ so that the budget constraint is $c = (1-\tau)z + y$. Introducing virtual income is a method of linearizing a piecewise linear budget constraint. The taxable income supply function for this group is denoted $z(1-\tau, y(\tau); z_0)$.

We are interested in how the individual's optimal taxable income will change when the tax rate is increased. There will be both income and substitution effects. When the tax rate changes, virtual income will also change. Because virtual income is the intercept of the linearized budget constraint, changing the tax rate will shift the budget constraint and this will induce income effects. For a taxpayer in the top tax bracket, $\partial y/\partial\tau = b$. Applying the definitions of the uncompensated taxable income elasticity ($\varepsilon_u = \partial z/\partial(1-\tau) \times (1-\tau)/z\ |_y$) and the income effect parameter ($\eta = (1-\tau)\partial z/\partial y$), the change in taxable income brought about by a tax reform can be expressed

$$\frac{dz(1-\tau, y(\tau))}{d\tau} = -\frac{\partial z}{\partial(1-\tau)}\ |_y + \frac{\partial z}{\partial y}\frac{\partial y}{\partial \tau} = -\frac{\varepsilon_u z - \eta b}{1-\tau} = -\frac{\varepsilon_c z + \eta(z-b)}{1-\tau}, \tag{16}$$

where the elasticity version of the Slutsky equation ($\varepsilon_u = \varepsilon_c + \eta$) is used in the last step.

Turning to considering aggregate effects, the density function of the earnings capacity distribution is denoted $f_0(z_0)$. The population of taxpayers is normalized to one for simplicity. The proportion of taxpayers in the top tax bracket (those who earn strictly more than b) is given by

$$N(\tau) = \int_{b_0(\tau)}^{\infty} f_0(z_0)dz_0,$$

where $b_0(\tau)$ is such that $z(1-\tau, y(\tau); b_0(\tau)) = b$, i.e., earnings capacity for the taxpayer who is at the margin of entering the top tax bracket.

[15] When both the budget set and individuals' preferences are convex, jumping between the interiors of the two segments of the tax schedule is ruled out and thus the tax rate in the second segment will not affect revenues from the first segment.

The average income of top-bracket taxpayers is

$$\bar{z}_b(\tau) = \int_{b_0(\tau)}^{\infty} z(1-\tau, y(\tau); z_0) f_0(z_0) dz_0 / N(\tau).$$

Integrating over taxpayers, tax revenues from the top tax bracket can be expressed

$$R(\tau) = \tau \int_{b_0(\tau)}^{\infty} \left[z(1-\tau, y(\tau); z_0) - b \right] f_0(z_0) dz_0 = \tau \left[\bar{z}_b(\tau) - b \right] N(\tau). \tag{17}$$

Using the Leibniz rule for the derivative of integrals, the revenue impact of a small tax increase is

$$\frac{dR}{d\tau} = \left[\bar{z}_b - b \right] N + \tau \left[\int_{b_0(\tau)}^{\infty} \frac{dz(1-\tau, y(\tau); z_0)}{d\tau} f_0(z_0) dz_0 - (z(1-\tau, y(\tau); b_0) - b) \frac{db_0(\tau)}{d\tau} \right] =$$
$$= \left[\bar{z}_b - b \right] N + \tau \int_{b_0(\tau)}^{\infty} \frac{dz(1-\tau, y(\tau); z_0)}{d\tau} f_0(z_0) dz_0. \tag{18}$$

The first term is the mechanical effect, the second term is the revenue impact of a change in average income and the third term is revenue effect of a changing number of high-income taxpayers. The third term is equal to zero because $z(1-\tau, y; b_0) = b$ by definition. Thus N can be regarded as constant for small changes in τ. This is also discussed by Saez et al. (2012, footnote 7).

The marginal degree of self-financing is the behavioural effect divided by the mechanical effect (which has the opposite sign):

$$DSF(\tau) = -\frac{\tau \int_{b_0}^{\infty} \frac{dz(1-\tau, y(\tau); z_0)}{d\tau} f_0(z_0) dz_0}{(\bar{z}_b - b)N}. \tag{19}$$

Next, we define the income-weighted average compensated taxable income elasticity to be

$$\bar{\varepsilon_c}(\tau) = \frac{\int_{b_0}^{\infty} \varepsilon_c(\tau; z_0) z(1-\tau, y; z_0) f_0(z_0) dz_0}{\int_{b_0}^{\infty} z(1-\tau, y; z_0) f_0(z_0) dz_0} = \frac{\int_{b_0}^{\infty} \varepsilon_c(\tau; z_0) z(1-\tau, y; z_0) f_0(z_0) dz_0}{\bar{z}_b N}.$$

The tax-base-weighted average income effect parameter is

$$\tilde{\eta}(\tau) = \frac{\int_{b_0}^{\infty} \eta(\tau; z_0) [z(1-\tau, y; z_0) - b] f_0(z_0) dz_0}{\int_{b_0}^{\infty} [z(1-\tau, y; z_0) - b] f_0(z_0) dz_0} = \frac{\int_{b_0}^{\infty} \eta(\tau; z_0) [z(1-\tau, y; z_0) - b] f_0(z_0) dz_0}{(\bar{z}_b - b)N}.$$

The tax base is the part of income that is in the top tax bracket. The reason for the different weightings will become clear in the derivations that follow. The existence of these averages is guaranteed by the second mean value theorem of integrals.

Applying these definitions and plugging equation 16 into equation 19, we can derive

$$DSF(\tau) = \frac{\tau \int_{b_0}^{\infty} \frac{\varepsilon_c(\tau; z_0) z(1-\tau, y(\tau); z_0) - \eta(\tau; z_0)[z(1-\tau, y(\tau); z_0) - b]}{1-\tau} f_0(z_0) dz_0}{(\bar{z}_b - b)N} =$$
$$= \frac{\tau \left[\bar{\varepsilon}_c \bar{z}_b N - \tilde{\eta}(\bar{z}_b - b)N \right]}{(1-\tau)(\bar{z}_b - b)N} = \frac{\tau \left[\alpha \bar{\varepsilon}_c + \tilde{\eta} \right]}{1-\tau}. \tag{20}$$

In the last step, we use the definition of the Pareto parameter $\alpha = \bar{z}_b/(\bar{z}_b - b)$ (equation 4). Thus we have arrived at equation 6. The substitution effect is proportional to all of taxable income (because the elasticity is defined in terms of the proportional change in taxable income) while the income effect depends only on the part that is in the top tax bracket (the tax base), because this determines how disposable income will change. Therefore it is natural that the compensated elasticity is income-weighted while the income effect parameter is tax-base-weighted.[16]

Note that α, $\bar{\varepsilon_c}$ and $\tilde{\eta}$ in general are functions of τ. Equation 20 is thus valid for the evaluation of small tax changes. When solving for τ to obtain an explicit expression, e.g., in order to set $DSF = 1$ to find the revenue-maximizing rate, it must be assumed that α, $\bar{\varepsilon_c}$ and $\tilde{\eta}$ are independent of τ. Unfortunately, it is difficult to find a utility specification where the behavioural parameters are constant on the individual level. Keane (2011) analyzes a utility function equivalent to equation 15. His derivations imply that ε_c and η will change after a tax reform.[17] Saez (2001) considers the limiting case $b \to \infty$, in which $\bar{\varepsilon_c}$ and $\tilde{\eta}$ can indeed be constant. For finite b, equation 14 – whose derivation requires constant $\bar{\varepsilon_c}$ and $\tilde{\eta}$ – is therefore only approximately correct.

In the case without income effects, i.e., $\eta = 0$, assuming constant ε_c is no longer so restrictive, as it is equivalent to assuming isoelastic and quasilinear utility (see equation 9). This also allows a connection with the derivation of the Laffer curve. With such a utility function, the taxable income supply function is $z(1-\tau; z_0) = z_0(1-\tau)^{\varepsilon}$, where z_0 now has the interpretation of potential income. Assuming constant α is equivalent to assuming Pareto distribution of high potential incomes, implying $f_0(z_0) = N_0 \alpha b^{\alpha}/z_0^{\alpha+1}$. Plugging these into equation 17, we can derive the high-income Laffer curve without income effects (equation 11). The assumptions needed to derive the well-known explicit expression for the revenue-maximizing rate ($\tau^* = 1/(1+\alpha\varepsilon)$) are therefore the same as the ones needed to derive the entire Laffer curve.

[16] In contrast, Saez (2001, p. 210) considers an income-weighted *uncompensated* elasticity and an *unweighted* income effect parameter.

[17] Keane's variable S, which he shows determines the magnitude of the behavioural parameters, is a function of virtual income in a piecewise linear tax system. Virtual income depends on the tax rate.

Effective top marginal tax rates

Country	τ_{IC}	τ_{IL}	τ_{IS}	τ_{SN}	τ_{SD}	τ_P	τ_C	τ_E
Australia	45%			4%		5%	8%	55%
Austria	55%					3%	16%	63%
Belgium	50%	4%			13%	32%	14%	74%
Canada	33%	13%	7%				9%	58%
Czech Republic	16%		8%	5%		9%	17%	46%
Denmark	52%				8%		24%	66%
Finland	32%	20%		2%	6%	24%	20%	72%
France	45%		4%	3%	6%	22%	15%	69%
Germany	45%		6%				14%	57%
Greece	45%						14%	53%
Hungary	15%			19%		27%	25%	60%
Ireland	40%		8%	4%		11%	17%	64%
Israel	50%						16%	58%
Italy	43%	4%	2%				13%	55%
Japan	45%	10%	1%		1%	1%	7%	60%
Lithuania	15%			9%		31%	15%	51%
Luxembourg	40%		4%				27%	59%
Mexico	35%						5%	39%
Netherlands	52%						15%	59%
New Zealand	33%						16%	44%
Norway	14%	25%		8%		14%	21%	63%
Poland	32%			2%		4%	16%	47%
Portugal	48%		9%		11%	24%	15%	74%
Slovakia	25%					1%	14%	36%
Slovenia	50%				22%	16%	21%	73%
South Korea	38%	4%		1%		1%	10%	49%
Spain	23%	23%					13%	52%
Sweden	25%	32%	3%			31%	19%	75%
Switzerland	12%	28%			6%	6%	8%	51%
United Kingdom	45%			2%		14%	12%	59%
United States	40%	4%		2%		1%	4%	48%

Sources: European Union (2015), KPMG (2016), PWC (2016), national sources (see country notes).

Explanation of column headers. τ_{IC}: central income tax

τ_{IL}: local, provincial, state etc. income tax

τ_{IS}: surtaxes, solidarity contributions etc.

τ_{SN}: non-deductible employee social contributions

τ_{SD}: deductible employee social contributions

τ_P: payroll taxes (employers' social contributions)

τ_C: average tax on consumption

τ_E: effective marginal tax rate. It is computed in this way:

$$\tau_E = \frac{\tau_I + \tau_{SN} + \tau_{SD}(1 - \tau_I) + \tau_C(1 - \tau_I - \tau_{SN} - \tau_{SD}(1 - \tau_I)) + \tau_P}{1 + \tau_P},$$

where $\tau_I = \tau_{IC} + \tau_{IL} + \tau_{IS}$ is national and local income tax and any surtaxes.

General notes. Tax rates for the very highest tax bracket in each country are shown. Payroll taxes and social contributions are only included if they are uncapped and apply to all incomes. Local tax rates are the national average unless stated otherwise. Income tax and social contribution rates are for 2015 or 2016.

The consumption tax rate is obtained from OECD data using the formula proposed by Mendoza et al. (1994): (general sales taxes + excise duties) / (private consumption expenditure + government consumption expenditure − government employee compensation). This takes into account the fact that some consumption taxes are paid by the government to itself. The data is from 2014, or 2013 in a few cases. No data for employee compensation was available for Canada, Mexico or New Zealand. In these cases, the government's compensation of employees was assumed to make up half of government consumption expenditures, which is the average of all countries.

Country notes. Australia: Payroll tax is the simple average of state tax rates.

Belgium: The average local tax rate is 7.54 percent of the national income tax.

Canada: Provincial tax rate for Ontario, including a 56 percent surtax on the provincial tax.

Czech Republic: The income tax base includes employer's social contributions. Income tax rates are therefore multiplied by 1.09.

Denmark: The sum of central and local tax rates is capped at 52 percent.

Italy: Local tax rate for Rome. The solidarity contribution of 3 percent is adjusted for the fact that it is deductible from central income tax.

Japan: The surtax is 2.1 percent of the central tax liability.

Luxembourg: The solidarity surcharge is 9 percent of the income tax liability.

Norway: Income tax rates from the Norwegian Tax Administration (2016).

Portugal: Employees' social contributions are deductible according to the Portuguese Tax and Customs Authority (2010). An extraordinary surtax of 3.5 percent and a solidarity surcharge of 5 percent apply.

Slovenia: Employees' social contributions are deductible according to email communication with the Slovenian Ministry of Finance.

South Korea: The local tax is 10 percent of the national tax.

Spain: Regional tax rate for Madrid.

Sweden: The surtax refers to the phase-out of the earned income tax credit for incomes between 600,000 and 1,500,000 SEK. More than 90 percent of top-bracket taxpayers are in this interval.

Switzerland: Local tax rate for Zurich. The cantonal tax rate is 13 percent and the municipal tax rate is 1.19 times the cantonal tax rate.

United States: Average state income tax from Diamond & Saez (2011). It is adjusted for the fact that it is deductible from federal income tax.

Pareto parameters

Country	LIS		WID		Other source	
	α	Year	α	Year	α	Reference
Australia	2.84	2010	1.86	2010		
Austria	3.14	2004				
Belgium	2.03	2000				
Brazil	1.98	2013				
Canada	2.92	2010	1.83	2010		
China	3.60	2002				
Colombia	2.47	2013	1.80	2010		
Czech Republic	2.95	2010				
Denmark	2.66	2010	2.17	2010	3.04	Brøns-Petersen (2016)
Dominican Rep	2.11	2007				
Egypt	2.16	2012				
Estonia	3.53	2010				
Finland	3.26	2013			2.4	Riihelä et al. (2014)
France	2.59	2010	2.20	2012		
Germany	2.95	2010	1.66	2010		
Georgia	3.45	2013				
Greece	2.30	2010				
Guatemala	1.84	2006				
Hungary	3.87*	2012				
Iceland	3.28	2010				
India	2.68	2011				
Ireland	2.88	2010	1.98	2009		
Israel	2.97	2012				
Italy	2.88	2010	2.18	2009		
Japan	3.60	2008	2.37	2010		
Luxembourg	3.39	2013				
Malaysia			1.75	2010		
Mexico	2.23	2012				
Netherlands	2.95	2010			3.35	Zoutman et al. (2014)
New Zealand			2.10	2012		
Norway	2.88	2010	2.02	2011		
Panama	2.28	2013				
Paraguay	1.79	2013				
Peru	2.32	2013				
Poland	3.25	2013				

Country	LIS		WID		Other source	
Russia	3.48	2013				
Serbia	4.55	2013				
Singapore			2.10	2012		
Slovakia	2.71	2010				
Slovenia	3.68*	2012				
South Africa	2.25	2012	2.18	2011		
South Korea	4.79	2006	1.81	2012		
Spain	3.32	2013	2.08	2012		
Sweden	2.55	2005	1.88	2013	3.18	own calculations
Switzerland			1.73	2010		
Taiwan	3.37	2013	1.79	2013		
United Kingdom	2.49	2013	1.79	2012		
United States	2.40	2013	1.61	2014		
Uruguay	2.70	2013	1.94	2012		

* Few observations; not used in calculations.

Sources: Alvaredo et al. (2016) (WID), own calculations based on LIS Cross-National Data Center (2016) (LIS), other sources.