The Sufficient Statistic Approach: Predicting the Top of the Laffer Curve*

Alejandro Badel
Bureau of Labor Statistics
Georgetown University
ale.badel@gmail.com

Mark Huggett
Georgetown University
mh5@georgetown.edu

THIS DRAFT: DECEMBER 7, 2016

Abstract
We provide a formula for the tax rate at the top of the Laffer curve as a function of three elasticities. Our formula applies to static models and to steady states of dynamic models and is relevant for the top tax rate on any component of income. We show how to apply it to many classic models. We also illustrate the application of the formula using a quantitative human capital model.

Keywords: Sufficient Statistic, Laffer Curve, Marginal Tax Rate, Elasticity

JEL Classification: D91, E21, H2, J24

*This work used the Extreme Science and Engineering Discovery Environment (XSEDE) Grant SES 130025, supported by the National Science Foundation, to access the Stampede cluster at the Texas Advanced Computing Center. This work also used the cluster at Georgetown University and the Federal Reserve Bank of Kansas City.

1 Introduction

Imagine that an important public policy issue, involving the functioning of the entire economy, could be settled by a simple formula together with only a few inputs estimated from data. Imagine further that the simple formula, connecting the public policy variable to empirical inputs, is general in that it holds within a wide class of theoretical models that are relevant to the issue at hand. The scenario just described is the goal of the sufficient statistic approach. Chetty (2009) states *"The central concept of the sufficient statistic approach ... is to derive formulas for the welfare* [revenue] *consequences of policies that are functions of high-level elasticities rather than deep primitives."*

An important application of the sufficient statistic approach is to predict the tax rate at the top of the Laffer curve (i.e. the revenue maximizing top tax rate). Thus, the public policy variable under consideration is a tax rate on some component of income or expenditure. This could be the tax rate on consumption, labor income, capital income or something more specific such as the top federal tax rate on ordinary income.

One may want to predict the top of the Laffer curve for several reasons. First, it may be widely agreed that setting a tax rate beyond the revenue maximizing rate is counterproductive. If so, an accurate prediction usefully narrows the tax policy debate. Second, one may argue, as do Diamond and Saez (2011), that the revenue maximizing tax rate on top earners closely approximates the welfare maximizing top tax rate for some welfare criteria. From this perspective, the revenue maximizing tax rate then becomes a quantitative policy guide.

From a quick look at the literature, one might conclude that the theoretical groundwork on this issue is complete. There is a widely-used sufficient statistic formula $\tau^* = 1/(1 + a\varepsilon)$ that characterizes the revenue maximizing top marginal tax rate τ^* that applies beyond a threshold. Moreover, there is also a closely related formula $\tau^* = (1 - g)/(1 - g + a\varepsilon)$ for the welfare maximizing marginal tax rate that applies beyond a threshold. This formula is stated in terms of the same two empirical inputs (a, ε) and a social welfare weight $g \geq 0$ put on the marginal consumption of top earners. See Diamond and Saez (2011) and Piketty and Saez (2013a), among many others, for a discussion of these formulae. The Mirrlees Review is an important public policy document that applies these formulae to offer quantitative policy advice.[1]

A more critical reading of the literature suggests that the widely-used formula does not actually apply to a wide class of relevant models. The widely-used formula $\tau^* = 1/(1 + a\varepsilon)$ is not valid in dynamic models. For example, it does not apply to steady states in either the infinitely-lived agent or the overlapping generations versions of the neoclassical growth model. These are the two workhorse models of macroeconomics. A large literature analyzes the taxation of

[1] See Brewer, Saez and Shephard (2010).

consumption, labor income and capital income using these models.[2] The widely-used formula also does not apply to the class of heterogeneous-agent models that currently dominate as positive models of the distribution of earnings, income, consumption and wealth.[3]

The sufficient-statistic formula in Theorem 1 of this paper applies to static models and to steady states of dynamic models. It applies regardless of whether a specific model allows for general equilibrium or only partial equilibrium responses. We apply the formula to the Mirrlees (1971) model, to the two workhorse models of macroeconomics and to a human capital model. The formula is stated below in terms of three elasticities, including the single elasticity ε_1 of the widely-used formula.

$$\tau^* = \frac{1 - a_2\varepsilon_2 - a_3\varepsilon_3}{1 + a_1\varepsilon_1}$$

One of the new elasticities ε_2 captures the possibility that tax revenue from agent types below the threshold will respond to changes in the top rate. For example, this can occur because such agents anticipate the possibility of passing the threshold later in life and change their decisions or because of factor price changes triggered by the decisions of agent types above the threshold. The other new elasticity ε_3 captures the possibility that agent types that are above the threshold have incomes or expenditures that are taxed separately and that the tax revenue from these sources changes when the top tax rate changes. For example, consumption and various types of capital income are commonly taxed separately from labor income.

This paper makes two contributions. First, Theorem 1 provides a tax rate formula with wide application and yet it is stated in terms of only three elasticities. It is a generalization of the widely-used formula. Thus, the formula in Theorem 1 should replace the widely-used formula in future work. It should also help guide future empirical work that estimates elasticities. Currently, there are no estimates in the literature for two of the three elasticities that enter the formula. Second, we bench test the formula using a human capital model. One message of the bench test is that the formula accurately predicts the top of the Laffer curve even when the formula inputs are calculated far from the top of the Laffer curve. Prediction is an important use of the formula in applied work.

The paper is organized as follows. Section 2 presents the tax rate formula. Section 3 shows that the formula applies in a straightforward way to several classic models. Section 4 bench tests the formula using a quantitative human capital model. Section 5 concludes.

[2] Auerbach and Kotlikoff (1987) is an early quantitative exploration of tax reforms within overlapping generations models.

[3] Heathcote, Storesletten and Violante (2009) review this literature. The models that they review feature agents that are ex-ante heterogeneous, face idiosyncratic risk and trade in incomplete financial markets.

Related Literature

Our paper is most closely related to three literatures. First, it relates to the so called sufficient statistics literature. Piketty and Saez (2013a) review the parts of this literature that focus on formulae for labor income taxation in static models. Our formula is a generalization of the revenue maximizing top rate formula from Saez (2001). Our formula applies to any component of income and not just labor income. Recently, sufficient statistic formulae have been developed that apply to steady states of dynamic models. For example, Piketty and Saez (2013b) present formulae for welfare maximizing tax rates on inheritance that apply to steady states of a specific dynamic model. We differ from this work in focus and in method. Our focus is revenue maximization rather than welfare maximization. In terms of method, we derive a formula based on an abstract modeling language rather than a specific model. A consequence is that our formula has very wide application as it is easy to map equilibria in many specific static or dynamic models into our abstract modeling language.

Second, it relates to the class of heterogeneous-agent models surveyed by Heathcote, Storesletten and Violante (2009). Our tax rate formula applies to many models in this large class. For example, Badel and Huggett (2014) apply our tax rate formula to a specific model. Our formula could also be applied to the models in Guner, Lopez-Daneri and Ventura (2014) and Kindermann and Krueger (2014). The model counterparts to the three formula coefficients and elasticities would be critical for understanding why these quantitative theoretical models have substantially different revenue maximizing top tax rates.

Third, it relates to the elasticity of taxable income literature surveyed by Saez, Slemrod and Giertz (2012). This literature provides elasticity estimates for sufficient statistic formulae.[4] We conjecture that our formula will be helpful in the development of methods to estimate the three relevant elasticities in dynamic models. This is because in specific economic models the elasticities can be calculated independently of the specific regression framework proposed. Badel and Huggett (2014) illustrate this point and assess existing regression methods for the elasticity ε_1 within a specific dynamic model.

2 Tax Rate Formula

The tax rate formula is based on three model elements: (i) a distribution of agent types $(X, \mathcal{X}, \mathrm{P})$, (ii) an income choice $y(x, \tau)$ that maps an agent type $x \in X$ and a parameter τ of the tax system into an income choice and (iii) a class of tax functions $T(y; \tau)$ mapping

[4] The elasticities relevant in sufficient statistic formulae are sometime termed "policy elasticities" (see Hendren (2015)) to emphasize that they are specific to changes in a specific policy variable.

income choice and a tax system parameter τ into the total tax paid. Total tax revenue is then $\int_X T(y(x,\tau);\tau)dP$. Our approach does not rely on specifying an explicit dynamic or static equilibrium model up front. Instead, our tax rate formula can be applied in a straightforward way by mapping equilibrium allocations of specific static or dynamic models into these three basic model elements.

2.1 Assumptions

Assumption $A1$ says that the distribution of agent types is represented by a probability space composed of a space of types X, a σ-field $\mathcal{X}$ on X and a probability measure P defined over sets in $\mathcal{X}$. Assumption $A2$ places structure on the class of tax functions. The tax functions differ in a single parameter τ, where τ is interpreted as the linear tax rate that applies to income beyond a threshold $\underline{y}$. Below this threshold the tax function can be nonlinear but all tax functions in the class are the same below the threshold. Assumption $A3$ says that key aggregates are differentiable in τ. The aggregates are based on integrals over the sets $X_1 = \{x \in X : y(x,\tau^*) > \underline{y}\}$ and $X_2 = X - X_1$ in $\mathcal{X}$, where $\tau^* \in (0,1)$ is a fixed value that serves to define and fix these sets.

A1. $(X, \mathcal{X}, \mathrm{P})$ *is a probability space.*

A2. There is a threshold $\underline{y} \geq 0$ *such that*

(i) $T(y;\tau) - T(\underline{y};\tau) = \tau[y - \underline{y}], \forall y > \underline{y}, \forall \tau \in (0,1)$ *and*

(ii) $T(y;\tau) = T(y;\tau'), \forall y \leq \underline{y}, \forall \tau, \tau' \in (0,1)$.

A3. $\int_{X_1} y(x,\tau)dP$ *and* $\int_{X_2} T(y(x,\tau);\tau)dP$ *are strictly positive and are differentiable in* τ.

We also consider a generalization where the tax system depends on multiple sources of income or expenditure. The three elements of the generalized model are (i) a distribution of agent types $(X, \mathcal{X}, \mathrm{P})$, (ii) an $n \geq 2$ dimensional income-expenditure choice $(y_1(x,\tau), ..., y_n(x,\tau))$ and (iii) a class of tax functions $T(y_1, ..., y_n; \tau)$ mapping choices and a tax system parameter τ into the total tax paid.

Assumptions $A1' - A3'$ restate assumptions A1 - A3 for the generalized model. $A2'$ assumes that the tax system is additively separable in that the first component of income y_1 determines a portion of the tax liability of an agent separately from the other components. For example, this structure captures a situation where labor income y_1 and capital income y_2 are taxed using separate tax schedules or where labor income y_1 and consumption y_2 are taxed separately. In Assumption $A3'$ the integrals are calculated over the sets $X_1 = \{x \in X : y_1(x,\tau^*) > \underline{y}\}$ and $X_2 = X - X_1$, where $\tau^* \in (0,1)$ is a fixed value.

A1′. $(X, \mathcal{X}, \mathrm{P})$ *is a probability space.*

A2′. T *is separable in that* $T(y_1, ..., y_n; \tau) = T_1(y_1; \tau) + T_2(y_2, ..., y_n), \forall (y_1, ..., y_n, \tau)$. *Moreover, there is a threshold* $\underline{y} \geq 0$ *such that*

(i) $T_1(y_1; \tau) - T_1(\underline{y}; \tau) = \tau[y_1 - \underline{y}], \forall y_1 > \underline{y}, \forall \tau \in (0,1)$ *and*

(ii) $T_1(y_1; \tau) = T_1(y_1; \tau'), \forall y_1 \leq \underline{y}, \forall \tau, \tau' \in (0,1)$.

A3′. $\int_{X_1} y_1 dP, \int_{X_1} T_2(y_2, ..., y_n) dP$ *and* $\int_{X_2} T(y_1, ..., y_n; \tau) dP$ *are strictly positive and are differentiable in* τ.

2.2 Formula

Before stating the formula in Theorem 1, we express total tax revenue as the sum of tax revenue from the set of agent types with incomes above a threshold $X_1 = \{x \in X : y(x, \tau^*) > \underline{y}\}$ and from all remaining types $X_2 = X - X_1$. Total tax revenue can be stated in the same manner when the tax system depends on $n \geq 2$ components of income or expenditure by again defining two sets $X_1 = \{x \in X : y_1(x, \tau^*) > \underline{y}\}$ and $X_2 = X - X_1$. This is done below.

$$\int_X T(y(x,\tau);\tau) dP = \int_{X_1} T(y(x,\tau);\tau) dP + \int_{X_2} T(y(x,\tau);\tau) dP$$

$$\int_X T(y_1, ..., y_n; \tau) dP = \int_{X_1} T(y_1, ..., y_n; \tau) dP + \int_{X_2} T(y_1, ..., y_n; \tau) dP$$

With these expressions in hand, we now state the theorem.

Theorem 1:

(i) Assume $A1 - A3$. *If* $\tau^* \in (0,1)$ *is revenue maximizing, then* $\tau^* = \frac{1 - a_2 \varepsilon_2}{1 + a_1 \varepsilon_1}$, *where*

$$(a_1, a_2) = \left(\frac{\int_{X_1} y dP}{\int_{X_1} \left[y - \underline{y}\right] dP}, \frac{\int_{X_2} T(y;\tau^*) dP}{\int_{X_1} \left[y - \underline{y}\right] dP}\right) \text{ and } (\varepsilon_1, \varepsilon_2) = \left(\frac{d \log \int_{X_1} y dP}{d \log(1-\tau)}, \frac{d \log \int_{X_2} T(y;\tau^*) dP}{d \log(1-\tau)}\right).$$

(ii) Assume $A1' - A3'$. *If* $\tau^* \in (0,1)$ *is revenue maximizing, then* $\tau^* = \frac{1 - a_2 \varepsilon_2 - a_3 \varepsilon_3}{1 + a_1 \varepsilon_1}$, *where*

$$(a_1, a_2, a_3) = \left(\frac{\int_{X_1} y_1 dP}{\int_{X_1} \left[y_1 - \underline{y}\right] dP}, \frac{\int_{X_2} T(y_1, ..., y_n; \tau^*) dP}{\int_{X_1} \left[y_1 - \underline{y}\right] dP}, \frac{\int_{X_1} T_2(y_2, ..., y_n) dP}{\int_{X_1} \left[y_1 - \underline{y}\right] dP}\right) \text{ and}$$

$$(\varepsilon_1, \varepsilon_2, \varepsilon_3) = (\frac{d\log \int_{X_1} y_1 dP}{d\log(1-\tau)}, \frac{d\log \int_{X_2} T(y_1, ..., y_n; \tau^*)dP}{d\log(1-\tau)}, \frac{d\log \int_{X_1} T_2(y_2, ..., y_n)dP}{d\log(1-\tau)}).$$

Proof:

(i) If $\tau^* \in (0,1)$ maximizes revenue then it also maximizes $\tau \int_{X_1} [y(x;\tau) - \underline{y}]dP + \int_{X_2} T(y(x;\tau), \tau)dP$. This holds by subtracting the constant term $\int_{X_1} T(\underline{y};\tau)dP$ from total revenue and using A2. The following necessary condition then holds:

$$\int_{X_1} \left[y(x,\tau^*) - \underline{y}\right] dP - \tau^* \frac{d\int_{X_1} y(x;\tau^*)dP}{d(1-\tau)} - \frac{d\int_{X_2} T(y(x,\tau^*);\tau^*)dP}{d(1-\tau)} = 0$$

Divide the necessary condition by $\int_{X_1} \left[y(x,\tau^*) - \underline{y}\right] dP$ and rearrange using the elasticities stated in the Theorem. This implies $1 - \frac{\tau^*}{1-\tau^*} a_1\varepsilon_1 - \frac{1}{1-\tau^*} a_2\varepsilon_2 = 0$ which in turn implies $\tau^* = \frac{1-a_2\varepsilon_2}{1+a_1\varepsilon_1}$.

(ii) If $\tau^* \in (0,1)$ maximizes revenue then it also maximizes $\tau \int_{X_1} (y_1 - \underline{y})dP + \int_{X_1} T_2(y_2, ..., y_n)dP + \int_{X_2} T(y_1, ..., y_n; \tau)dP$. This holds by subtracting the constant term $\int_{X_1} T_1(\underline{y};\tau)dP$ from total revenue and using $A2'$. The following necessary condition then holds:

$$\int_{X_1} \left[y_1 - \underline{y}\right] dP - \tau^* \frac{d\int_{X_1} y_1 dP}{d(1-\tau)} - \frac{d\int_{X_1} T_2(y_2, ..., y_n)dP}{d(1-\tau)} - \frac{d\int_{X_2} T(y_1, ..., y_n; \tau^*)dP}{d(1-\tau)} = 0$$

Divide all terms in the previous equation by $\int_{X_1} \left[y_1 - \underline{y}\right] dP$ and then rearrange using the elasticities stated in the Theorem. This implies $1 - \frac{\tau^*}{1-\tau^*} a_1\varepsilon_1 - \frac{1}{1-\tau^*} a_2\varepsilon_2 - \frac{1}{1-\tau^*} a_3\varepsilon_3 = 0$ which in turn implies $\tau^* = \frac{1-a_2\varepsilon_2-a_3\varepsilon_3}{1+a_1\varepsilon_1}$. $\|$

Comments:

1. The formula is appealing from the perspective of the sufficient statistic approach. It is stated in terms of at most three elasticities. Nevertheless, it applies to economies where taxes are determined based on many different income or expenditure types. It applies to non-parametric economic models analyzed in partial or in general equilibrium.

2. The widely-used formula $\tau^* = 1/(1+a\varepsilon)$ is effectively a special case of the sufficient statistic formula in Theorem 1. What type of situations does the widely-used formula not address that the formula in Theorem 1 successfully addresses? There are two general categories. The first category includes situations where agent types below the threshold, in the set X_2, have their income and expenditures $(y_1, ..., y_n)$ and corresponding tax liabilities change as τ changes. This can happen, in static or dynamic models, when factor prices change due to the response from agent types above the threshold. In dynamic models this can also happen because agents transit through the income distribution. Thus, agents can be below the threshold at one age and above it at a later age. This implies that agent types below the threshold can have income or expenditure choices that vary with τ. In all these circumstances the tax revenue from agent types in X_2 changes as τ changes and thus the term $a_2\varepsilon_2$ is non-zero.

The second category covers scenarios in which many components of income or expenditure are taxed in practice. Consider a change in the parameter τ that governs the taxation of component y_1. Then agent types above the threshold, in the set X_1, will adjust other components $(y_2, ..., y_n)$ of income or expenditure. The revenue consequences of such adjustments need to be accounted for. The term $a_3\varepsilon_3$ will be non-zero when these revenues change. The next two sections give concrete examples of when the terms $a_2\varepsilon_2$ or $a_3\varepsilon_3$ are non-zero.

3. To state the formula using elasticities requires that each of the integrals (e.g. $\int_{X_2} T(y;\tau)dP$), over which the elasticity is taken, is non-zero. If any of the integral terms is zero, then the result can still be stated in some cases. For example, if the integral $\int_{X_2} T(y;\tau)dP$ is zero (i.e. total net taxes on agent types at or below the threshold are zero) and the integral does not vary on the margin as the tax rate τ varies, then the term $a_2\varepsilon_2$ in the formula in Theorem 1 can be replaced with a zero. Examples 1 and 3 in the next section illustrate this point.

3 Examples

We now consider three classic models: the Mirrlees model as well as the overlapping generations and infinitely-lived agent versions of the neoclassical growth model. We map equilibrium elements in each model into the language of Theorem 1. While the examples associate the tax rate parameter τ with a labor income tax rate, this is purely for convenience.

3.1 Mirrlees Model

Mirrlees (1971) considered a static model in which agents make a consumption and labor decision. In our version of this model, the government runs a balanced budget where taxes fund a lump-sum transfer $Tr(\tau)$. The model's primitives are a utility function $u(c, l)$, agent's productivity $x \in X$, a productivity distribution P and proportional tax rates on consumption τ_c and labor income τ.

Definition: *An equilibrium is* $(c(x;\tau), l(x;\tau), Tr(\tau))$ *such that given any* $\tau \in (0,1)$

1. *optimization:* $(c(x;\tau), l(x;\tau)) \in argmax\{u(c,l) \ : (1+\tau_c)c \leq wxl(1-\tau) + Tr(\tau), l \geq 0\}$
2. *government:* $Tr(\tau) = \tau \int_X wxl(x;\tau)dP + \tau_c \int_X c(x;\tau)dP$
3. *feasibility:* $\int_X c(x;\tau)dP = w \int_X xl(x;\tau)dP$

We state equilibria in closed form using the following functional forms and restrictions:

$u(c,l) = c - \alpha \frac{l^{1+\frac{1}{\nu}}}{1+\frac{1}{\nu}}$ and $\alpha, \nu > 0$

$X = R_+$ and $(X, \mathcal{X}, P)$ implies that $\int_X x^{1+\nu} dP$ is finite and $P(\{0\}) = 0$

Equilibrium allocations are straightforward to state:

$l(x;\tau) = [\frac{wx(1-\tau)}{\alpha(1+\tau_c)}]^\nu$

$c(x;\tau) = [wx[\frac{wx(1-\tau)}{\alpha(1+\tau_c)}]^\nu (1-\tau) + Tr(\tau)]/(1+\tau_c)$

$Tr(\tau) = (\tau + \tau_c) w^{1+\nu} [\frac{(1-\tau)}{\alpha(1+\tau_c)}]^\nu \int_X x^{1+\nu} dP$

We now map equilibrium allocations into the elements used to state Theorem 1.

Step 1: Set $(X, \mathcal{X}, P)$ to that governing productivity.
Step 2: Set $y_1(x,\tau) = wxl(x;\tau)$ and $y_2(x,\tau) = c(x;\tau)$
Step 3: Set $T(y_1, y_2; \tau) = \tau y_1 + \tau_c y_2$.

We now calculate the terms in the formula. It is clear that $\varepsilon_1 = \nu$ by a direct calculation of the elasticity and that $a_1 = 1$ as the threshold is $\underline{y} = 0$.[5] It is also clear that $\int_{X_2} T(y_1, y_2; \tau) dP = 0$ as effectively all agent types are above the threshold as $P(X_2) = P(\{0\}) = 0$. Thus, the term $a_2 \varepsilon_2$ in the formula can be replaced with a zero, consistent with Comment 3 to Theorem 1. It is easy to see that $(a_3, \varepsilon_3) = (\tau_c, \nu)$. Finally, it is also easy to calculate the revenue maximizing tax rate directly and see that it agrees with the rate implied by the formula.

$$\tau^* = \frac{1 - a_2\varepsilon_2 - a_3\varepsilon_3}{1 + a_1\varepsilon_1} = \frac{1 - 0 - \tau_c\nu}{1 + 1 \times \nu} = \frac{1 - \tau_c\nu}{1+\nu}$$

[5]Theorem 1 directs one to calculate the elasticities and the related coefficients when the sets (X_1, X_2) are defined at the revenue maximizing tax rate τ^*. We calculate $(a_1, \varepsilon_1) = (1, \nu)$ when these sets are specified for any fixed value of $\tau \in (0,1)$.

3.2 Overlapping Generations Growth Model

Diamond (1965) analyzes an overlapping generations model with two-period lived agents and a neoclassical production function $F(K, L)$ with constant returns. In our version of this model, age 1 and age 2 agents are equally numerous at any point in time and each age group has a mass of 1. Agents solve problem P1, where they choose labor, consumption and savings when young. They face proportional labor income and consumption taxes with rates τ and τ_c, respectively. The government collects taxes and makes a lump-sum transfer $Tr(\tau)$ to young agents.

(P1) $\max U(c_1, c_2, l)$ *s.t.*

$(1+\tau_c)c_1 + k \leq w(\tau)zl(1-\tau) + Tr(\tau), (1+\tau_c)c_2 \leq k(1+r(\tau))$ and $l \in [0, 1]$

Age 1 agents are heterogeneous in labor productivity $z \in Z \subset R_+$. The distribution of labor productivity is given by a probability space $(Z, \mathcal{Z}, \hat{P})$. Define two aggregates $K(\tau) = \int_Z k(z;\tau)d\hat{P}$ and $L(\tau) = \int_Z zl(z;\tau)d\hat{P}$.

Definition: *A steady-state equilibrium is* $(c_1(z;\tau), c_2(z;\tau), l(z;\tau), k(z;\tau))$, *a transfer* $Tr(\tau)$ *and factor prices* $(w(\tau), r(\tau))$ *such that for any* $\tau \in (0, 1)$

1. *optimization:* $(c_1(z;\tau), c_2(z;\tau), l(z;\tau), k(z;\tau))$ *solve P1.*
2. *prices:* $w(\tau) = F_2(K(\tau), L(\tau))$ *and* $1 + r(\tau) = F_1(K(\tau), L(\tau))$
3. *government:* $Tr(\tau) = \tau_c \int_Z (c_1(z;\tau) + c_2(z;\tau))d\hat{P} + \tau \int_Z w(\tau)zl(z;\tau)d\hat{P}$
4. *feasibility:* $\int_Z (c_1(z;\tau) + c_2(z;\tau))d\hat{P} + K(\tau) = F(K(\tau), L(\tau))$

We now map equilibrium allocations into the language used to state Theorem 1.

Step 1: Define the probability space of agent types. An agent type is $x = (z, j)$ consisting of the agent's productivity z when young and the agent's current age j.

$$x = (z, j) \in X = Z \times \{1, 2\} \textit{ and } P(A) = \int_Z [\frac{1}{2} 1_{\{(z,1) \in A\}} + \frac{1}{2} 1_{\{(z,2) \in A\}}] d\hat{P}, \forall A \in \mathcal{X}$$

Step 2: Define choices (y_1, y_2) as labor income and consumption, respectively.

$$(y_1(x;\tau), y_2(x;\tau)) = \begin{cases} (w(\tau)zl(z;\tau), c_1(z;\tau)) & for\ x = (z, 1), \forall z \in Z \\ (0, c_2(z;\tau)) & for\ x = (z, 2), \forall z \in Z \end{cases}$$

Step 3: Set $T(y_1, y_2; \tau) = \tau y_1 + \tau_c y_2$. Aggregate taxes $\int_X T(y_1, y_2; \tau) dP$ are proportional to the right-hand side of equilibrium condition 3.

The coefficients premultiplying each of the elasticities are easy to determine. The threshold is $\underline{y} = 0$ as the tax τ is a proportional labor income tax. The coefficients are $(a_1, a_2, a_3) = (1, \frac{\tau_c \int_Z c_2(z;\tau^*) d\hat{P}}{w(\tau^*)L(\tau^*)}, \frac{\tau_c \int_Z c_1(z;\tau^*) d\hat{P}}{w(\tau^*)L(\tau^*)})$. The coefficient a_2 from Theorem 1 is the ratio of the total tax revenue from agent types in X_2 to the total incomes y_1 above the threshold $\underline{y}$ for agent types in X_1. This equals the ratio of total consumption taxes paid by age 2 agents to total labor income. The coefficient a_3 is the ratio of total tax revenue from agents in X_1 on types of income or expenditure other than type y_1 to total incomes y_1 above the threshold.

$$\tau^* = \frac{1 - a_2\varepsilon_2 - a_3\varepsilon_3}{1 + a_1\varepsilon_1} = \frac{1 - \frac{\tau_c \int_Z c_2(z;\tau^*) d\hat{P}}{w(\tau^*)L(\tau^*)} \times \varepsilon_2 - \frac{\tau_c \int_Z c_1(z;\tau^*) d\hat{P}}{w(\tau^*)L(\tau^*)} \times \varepsilon_3}{1 + 1 \times \varepsilon_1}$$

The take-away point from Example 2 is that the formula in Theorem 1 applies to steady states of dynamic models once an agent type is viewed in the right way.[6]

3.3 Infinitely-Lived Agent Growth Model

Trabandt and Uhlig (2011) analyze Laffer curves using the neoclassical growth model. The model features a production function $F(k, l)$ and an infinitely-lived agent. They calibrate some model parameters so that steady states of their model match aggregate features of the US economy and 14 European economies and preset other model parameters. They calculate Laffer curves by varying either the tax rate on labor income, capital income or consumption to determine how the steady-state equilibrium lump-sum transfer responds to the tax rate. We show that our sufficient statistic formula applies to Laffer curves in their model. We focus on the Laffer curve related to varying the labor income tax rate, but our formula also applies to the Laffer curves arising from varying the other tax rates in their model.

The equilibrium concept given below differs from that in Trabandt and Uhlig in that it abstracts from steady-state growth and imports in order to simplify the exposition.

$$(P1) \quad \max \textstyle\sum_{t=0}^{\infty} \beta^t u(c_t, l_t) \quad s.t.$$

$$(1 + \tau_c)c_t + (k_{t+1} - k_t(1 - \delta)) + b_{t+1} \leq (1 - \tau_l)w_t l_t + k_t(1 + r_t(1 - \tau_k)) + b_t R_t^b + Tr_t$$

[6]For transparency we abstract from long-run growth. It is straightforward to add a constant growth rate of labor augmenting technological change and analyze equilibria displaying balanced growth. The tax system parameter τ will alter the "level" but not the growth rate of balanced-growth equilibrium variables. The formula can then be applied to these "level variables". The tax rate produced by the formula will maximize revenue each period among all the balanced-growth equilibria indexed by the tax rate τ.

$l_t \in [0,1]$ and k_0 is given

Definition: *A steady-state equilibrium is an allocation* $\{c_t, l_t, k_t\}_{t=0}^{\infty}$, *prices* $\{w_t, r_t, R_t^b\}_{t=0}^{\infty}$ *fiscal policy* $\{g_t, b_t, Tr_t\}_{t=0}^{\infty}$ *such that*

1. *optimization:* $(c_t, l_t, k_t, b_t) = (\bar{c}, \bar{l}, \bar{k}, \bar{b}), \forall t \geq 0$ *solves P1.*
2. *prices:* $w_t = \bar{w} = F_2(\bar{k}, \bar{l}), r_t = \bar{r} = F_1(\bar{k}, \bar{l}) - \delta$ *and* $R_t^b = \bar{R}^b, \forall t \geq 0$
3. *government:* $(g_t, b_t, Tr_t) = (\bar{g}, \bar{b}, \bar{Tr}), \forall t \geq 0$ *and* $\bar{g} + \bar{b}(\bar{R}^b - 1) + \bar{Tr} = \tau_l \bar{w}\bar{l} + \tau_c \bar{c} + \tau_k \bar{k}\bar{r}$
4. *feasibility:* $\bar{c} + \bar{k}\delta + \bar{g} = F(\bar{k}, \bar{l})$

We map equilibrium allocations into the language of Theorem 1. Denote the labor income tax rate $\tau_l = \tau$. Bars over variables denote steady-state quantities so that $\bar{c}(\tau)$ denotes the steady-state equilibrium consumption associated with labor income tax rate $\tau_l = \tau$, fixing the other tax rates (τ_c, τ_k) and government spending and debt $(\bar{g}, \bar{b})$. Transfers $\bar{Tr}(\tau)$ adjust to changes in revenue when τ is varied.

Step 1: $(X, \mathcal{X}, \mathrm{P})$ is $X = 1, \mathcal{X} = \{\{1\}, \emptyset\}, \mathrm{P}(\{1\}) = 1, \mathrm{P}(\emptyset) = 0$
Step 2: $y_1(x, \tau) = \bar{w}(\tau)\bar{l}(\tau), y_2(x, \tau) = \bar{c}(\tau)$ and $y_3(x, \tau) = \bar{r}(\tau)\bar{k}(\tau)$
Step 3: $T(y_1, y_2, y_3; \tau) = \tau y_1 + \tau_c y_2 + \tau_k y_3$

With the tax system in step 3, total taxes $\int_X T(y_1, y_2, y_3; \tau) dP$ equal the right-hand side of equilibrium condition 3. Therefore, the Laffer curve for transfers $Tr(\tau)$ equals total taxes less government spending and interest payments on the debt. Since $\bar{R}^b(\tau) = 1/\beta, \forall \tau \in [0,1)$ follows directly from the Euler equation in steady state, transfers are a monotone function of total taxes in this model.

Given the mapping, the coefficients pre-multiplying the elasticities are easy to calculate. The two relevant coefficients are $(a_1, a_3) = (1, \frac{\tau_c \bar{c}(\tau^*) + \tau_k \bar{k}(\tau^*)\bar{r}(\tau^*)}{\bar{w}(\tau^*)\bar{l}(\tau^*)})$. The representative-agent structure implies that there is just one agent type. Thus, the term $a_2\varepsilon_2$ is zero in this model as there are no agent types in the set X_2 having y_1 at or below the threshold $\underline{y} = 0$. Thus, the revenue from these types is zero at all tax rates. The coefficient a_3 is non-zero as there are other sources of taxes besides the labor tax on agent types in the set X_1. When the labor tax moves these other sources of tax revenue can move as well.

$$\tau^* = \frac{1 - a_2\varepsilon_2 - a_3\varepsilon_3}{1 + a_1\varepsilon_1} = \frac{1 - \frac{\tau_c \bar{c}(\tau^*) + \tau_k \bar{k}(\tau^*)\bar{r}(\tau^*)}{\bar{w}(\tau^*)\bar{l}(\tau^*)} \times \varepsilon_3}{1 + 1 \times \varepsilon_1}$$

The take-away point is that for each of the model economies considered by Trabandt and Uhlig (2011), there are two high-level elasticities $(\varepsilon_1, \varepsilon_3)$ and one coefficient a_3 that determine the top of the model Laffer curve. Thus, for the purpose of determining the top of the Laffer curve with respect to a specific tax rate, the empirical strategy could be quite different. Instead of calibrating the many parameters of a specific parametric version of the Trabandt-Uhlig model, one could focus on estimating the two high-level elasticities that are directly relevant for determining the top of the model Laffer curve.

4 Bench Testing the Formula

We now bench test the formula. One use of the formula in applied work is to predict the tax rate at the top of a Laffer curve. Prediction is based on using values of the three elasticities and the related coefficients determined away from the maximum.[7] The bench test determines the accuracy properties of the formula when the relevant inputs are determined away from the maximum.

4.1 A Human Capital Model

We bench test the formula using a version of the Ben-Porath (1967) model.[8] Agents maximize lifetime utility by choosing time allocation decisions (n_j, l_j, s_j) and by choosing consumption and asset choices (c_j, a_{j+1}). Leisure n_j, work time l_j and learning time s_j are distinct activities. Labor market earnings wh_jl_j are the product of a wage rate w, worker skill h_j and work time l_j before retirement. Worker skill evolves according to a function H which depends on current skill h_j, learning time s_j and learning ability a. Taxes are determined by a tax function T_j and by a lump-sum transfer Tr.

Problem P1: $\max \sum_{j=1}^{J} \beta^{j-1} u(c_j, n_j)$ subject to

$$c_j + k_{j+1} = e_j + k_j(1+r) - T_j(e_j, c_j, k_jr; \tau) + Tr \text{ and } c_j, k_{j+1}, n_j, s_j, l_j \geq 0$$

$$e_j = wh_jl_j \; for \; j < Retire \text{ and } e_j = 0 \; for \; j \geq Retire$$

$$h_{j+1} = H(h_j, s_j, a) \text{ and } n_j + s_j + l_j = 1, \text{ given } (h_1, a) \text{ and } k_1 = 0.$$

[7]Clearly, tax rate formulae are designed to hold at the maximum. Diamond and Saez (2011), and others, use such formulae to predict the revenue maximizing top tax rate.

[8]This is a central model in the analysis of the distribution of earnings (see Weiss (1986), Neal and Rosen (2000) and Rubinstein and Weiss (2006)). See Huggett, Ventura and Yaron (2006) and Huggett, Ventura and Yaron (2011) for a quantitative analysis of such models.

The economy has an overlapping generations structure. The fraction μ_j of age j agents in the population satisfies $\mu_{j+1} = \mu_j/(1+n)$, where n is a population growth rate. There is an aggregate production function $F(K, L)$. Physical capital depreciates at rate δ. Aggregate capital, labor and consumption $(K(\tau), L(\tau), C(\tau))$ are straightforward functions of the decisions of agents, population fractions and the distribution of initial conditions. Decisions are functions of age and initial conditions. Initial conditions are initial skill and learning ability level (h_1, a). The distribution of initial conditions is given by a probability measure P.

$$K(\tau) = \sum_{j=1}^{J} \mu_j \int k_j(h_1, a; \tau)dP \text{ and } L(\tau) = \sum_{j=1}^{J} \mu_j \int h_j(h_1, a; \tau)l_j(h_1, a; \tau)dP$$

Definition: *An equilibrium consists of decisions* $(c_j, k_j, n_j, l_j, s_j, h_j)$*, factor prices* $(w(\tau), r(\tau))$ *and government transfers* $Tr(\tau)$ *such that (1)-(4) hold, given government spending* G*: (1) optimization:* $(c_j, k_j, n_j, l_j, s_j, h_j)$ *solve P1, (2) prices:* $w(\tau) = F_2(K(\tau), L(\tau))$ *and* $r(\tau) = F_1(K(\tau), L(\tau)) - \delta$*, (3) government:* $Tr(\tau) + G = \sum_{j=1}^{J} \mu_j \int T_j(e_j, c_j, k_j r; \tau)dP$ *and (4) feasibility:* $C(\tau) + K(\tau)(n+\delta) + G = F(K(\tau), L(\tau))$.

4.2 Setting Model Parameters

Table 1 presents functional forms and model parameter values. Parameters are set in three main steps. First, demographic, technology and tax system parameters are set without solving the model. Demographic parameters are set so that agents begin life at a real-life age of 23, retire at age 65 and live up to a real-life age of 85. The population growth rate is $n = 0.01$. Technology parameters (A, γ, δ) are set so that (i) γ matches the US value for capital's share, (ii) δ is consistent with the US investment-output ratio and the capital-output ratio, given n and (iii) A is normalized so that the equilibrium wage rate is $w = 1.0$. US values for capital's share, the investment-output ratio and the capital-output ratio equal $(.352, .174, 3.22)$ based on averages from national accounts data over the period 1960-2015. The tax system consists of a progressive tax on earnings T, proportional tax rates (τ_c, τ_k) on consumption and capital income and a lump-sum retirement transfer. Tax system parameters are based on the US evidence presented below. In the benchmark model the common lump-sum transfer Tr in the definition of equilibrium is zero.

Second, the utility function parameter is set to $\nu = 0.35$ for the benchmark case. This implies a Frisch elasticity of leisure with respect to human capital of -0.35 and a Frisch elasticity of total labor time (i.e. $s_j + l_j$) with respect to human capital equal to $\nu(n_j/(l_j + s_j))$. Thus, the Frisch elasticity of total labor time is 0.525 when total labor time is 40 percent of the time

Table 1 - Benchmark Model Parameter Values

Category	Functional Forms	Parameter Values
Demographics	$\mu_{j+1} = \mu_j/(1+n)$	$n = 0.01, J = 63, Retire = 43$ $j = 1, ..., 63$ (ages 23-85)
Technology	$F(K, L) = AK^{\gamma}L^{1-\gamma}$	$(A, \gamma, \delta) = (0.877, 0.352, 0.044)$
Tax System	$T(e; \tau) + \tau_c c + \tau_k kr$ for $j < Retire$ $\tau_c c + \tau_k kr - transfer$ for $j \geq Retire$	$T(e; \tau)$ is based on Figure 1 $\tau_c = 0.10$ and $\tau_k = 0.20$ $transfer = 18115$
Preferences	$u(c, n) = \log c + \phi \frac{n^{1-1/\nu}}{1-1/\nu}$	$\beta = 0.967, \phi = 0.618, \nu = 0.35$
Human Capital	$H(h, s, a) = h(1-\delta_h) + a(hs)^{\alpha}$	$(\alpha, \delta_h) = (0.677, 0.0043)$
Initial Conditions	$a \sim PLN(\mu_a, \sigma_a^2, \lambda_a)$ and $\epsilon \sim LN(0, \sigma^2)$ $\log h_1 = \beta_0 + \beta_1 \log a + \log \epsilon$	$(\mu_a, \sigma_a, \lambda_a) = (-0.977, 0.167, 3.45)$ $(\beta_0, \beta_1, \sigma) = (4.68, 0.939, 0.711)$

Note: PLN and LN denote the Pareto-Lognormal and the Lognormal distributions respectively. The joint distribution of initial conditions is constructed by assuming that ϵ is independent of learning ability, which follows a PLN distribution.

endowment. Third, all remaining model parameters ($(\delta, \beta, \phi, \alpha, \delta_h)$ and initial conditions) are chosen so that a model equilibrium best matches the US targets that we document below.

The remainder of this section discusses the model tax system, US targets and the model fit.

Tax System

We input earnings in thousand dollar increments into TAXSIM for a couple filing jointly that is living in a specific state in the 2010 tax year. TAXSIM calculates total taxes, which include federal and state income taxes and the employee and employer parts of social security and medicare taxes.[9] We calculate a marginal tax rate as the change in total taxes divided by the change in total earnings, where total earnings also include the employer part of social security and medicare taxes.

Figure 1 plots the relationship between earnings and marginal tax rates when averaged across states.[10] The model marginal tax rate function is a piecewise-linear function that passes through the empirical rate schedule. The top tax rate in the model is 0.422 which is the empirical marginal tax rate at an income of 319.5 thousand dollars. Appendix A.2 calculates that this is the 99-th percentile threshold for income in the US in 2010 based on our income measure.

The consumption tax rate $\tau_c = 0.10$ is the ratio of taxes on production in 2010 to total consumption expenditures.[11] These taxes include sales, excise and property taxes. The capital income tax is set to $\tau_k = 0.20$. We enter long-term capital gains in thousand dollar increments

[9] See Feenberg and Coutts (1993).

[10] Averages are computed using state employment as weights. Source: http://www.bls.gov/lau/rdscnp16.htm

[11] See Bureau of Economic Analysis Table 3.1 and Table 1.1.5.

into TAXSIM for a couple filing jointly with earnings of 160 thousand dollars in 2010 that lives in a specific state. We compute a marginal tax rate as the change in taxes divided by the change in income. The resulting schedule, when averaged across states, is flat beyond 280 thousand dollars of capital gains and the marginal rate at this threshold equals $\tau_k = 0.20$. The federal rate on long-term capital gains was 15 percent in 2010. We set the common social security benefit to 18.115 thousand dollars. This is the yearly old-age benefit for a worker retiring in 2010 based on an earnings history equal to average earnings.[12]

US Targets

We target several features of US data. First, we target earnings and hours profiles. Second, we target the 99th percentile of income in 2010 and the Pareto statistic at the 99th percentile of income in 2010. We describe next how these targets are constructed from US data. Earnings and hours profiles based on US data are displayed in Figure 2.

Earnings and hours profiles are based on the estimated age polynomials from a regression of the earnings statistic of interest on a third-order polynomial in age and a time dummy variable for each year. We run this regression on tabulated US Social Security Administration (SSA) male earnings data from Guvenen, Ozkan and Song (2014). The earnings facts are age profiles for (i) median earnings, (ii) the 99-50, 90-50 and 10-50 earnings percentile ratios and (iii) the Pareto statistic at the 99th percentile. The facts on the average fraction of time spent working are based on Panel Study of Income Dynamics (PSID) male hours data from Heathcote, Perri and Violante (2010).[13] Appendix A.1 describes the SSA and PSID data sets.[14]

We document that the 99th percentile of income is 319.5 thousand dollars in 2010 and that the Pareto statistic at this percentile is 1.70 in 2010. The Pareto statistic in cross-section data at a given threshold is the data counterpart to the formula coefficient a_1 in Theorem 1. The Pareto statistic of a distribution is $\bar{y}/(\bar{y} - \underline{y})$, where $\bar{y}$ is the mean for observations beyond a threshold $\underline{y}$. We use Statistics of Income (SOI) data and a definition of income to determine these two income facts. Appendix A.2 states the income measure that we use, the categories of income in SOI data and our methods for calculating $(\underline{y}, \bar{y})$.

[12]The old-age benefit is from Table 2.A26 in the 2010 Annual Statistical Supplement of the Social Security Bulletin. See https://www.ssa.gov/policy/docs/statcomps/supplement/2010/2a20-2a28.pdf

[13]The average fraction of time spent working is total work hours per year in PSID data divided by discretionary time (i.e. 14 hours per day times 365 days per year).

[14]The age polynomials from the regression on earnings data are normalized to pass through the data statistics at age 45 in 2010, with the exception of median earnings. We regress the log of median earnings as indicated above, exponentiate the estimated age polynomial and then scale the result so that median earnings equal 1 at age 25. The age coefficients from the PSID hours regression are normalized to pass through the average value across years at age 45 in the way specified in Appendix A.1.

Model Fit

Figure 2 plots US targets and the corresponding model implications. The model produces a fairly flat total hours profile (i.e. $s_j + l_j$) as in the data and at the same time produces an increasing median earnings profile. Two model forces contribute to the upward sloping median earnings profile: (1) agents accumulate human capital with age because they invest time in learning and (2) agents shift time to market work and away from learning as they age. The model produces a strong increase in the 99-50 earnings ratio with age and the decrease in the Pareto statistic with age. Both model features are due to the fact that agents with higher learning ability have steeper age-earnings profiles than agents with lower learning ability, other things equal.

4.3 Model Laffer Curves

Figure 3 presents Laffer curves for the benchmark model with $\nu = 0.35$ and for a model with an alternative value for the preference parameter $\nu = 0.25$.[15] We vary the parameter ν so that the top of the Laffer curve potentially occurs at a different top tax rate or produces a different magnitude for the lump-sum transfer.[16]

The peak of the Laffer curve in both models occurs at a top tax rate of approximately 57 percent. This top tax rate raises extra revenue of roughly a tenth of one percent of output in the original steady state in the benchmark model and of slightly more in percentage terms in the model with $\nu = 0.25$. The extra revenue funds a lump-sum transfer.

We now relate the top of the Laffer curves in Figure 3 to the top as predicted by the tax rate formula. To do this, we map equilibrium variables from the model into the three elements used in Theorem 1. This is done in three steps.

Step 1: $x = (h_1, a, j) \in X = R_+ \times R_+ \times \{1, ..., J\}$

Step 2: $y_1(x, \tau) = w(\tau)h_j(h_1, a; \tau)l_j(h_1, a; \tau), y_2(x, \tau) = c_j(h_1, a; \tau), y_3(x, \tau) = r(\tau)k_j(h_1, a; \tau)$ and $y_4(x, \tau) = transfer$ when $j \geq Retire$ and zero otherwise.

Step 3: $T(y_1, y_2, y_3, y_4; \tau) = T(y_1; \tau) + \tau_c y_2 + \tau_k y_3 - y_4$

Based on Steps 1-3, we calculate the coefficients (a_1, a_2, a_3) and the elasticities $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ for each model.[17] Table 2 presents the results. The coefficient a_1 in both models is near the

[15] The model Laffer curve is calculated (see Appendix A.3) by varying the top tax rate τ, computing the model equilibrium for each value of τ and plotting the resulting equilibrium total tax revenue.

[16] After varying ν, we adjust the model parameters (β, ϕ) and (β_0, β_1) to best match US targets.

[17] First, calculate $X_1 = \{x \in X : y_1(x; \tau) > \underline{y}\}$ using the benchmark value $\tau = .422$ and the threshold $\underline{y}$

Table 2: Revenue Maximizing Top Tax Rate Formula

Terms	Benchmark Model $\nu = 0.35$	Alternative Model $\nu = 0.25$
$a_1 \times \varepsilon_1$	$1.70 \times .267 = .454$	$1.75 \times .265 = .465$
$a_2 \times \varepsilon_2$	$2.04 \times .057 = .116$	$2.39 \times .032 = .076$
$a_3 \times \varepsilon_3$	$.076 \times .403 = .031$	$.080 \times .493 = .040$
$\tau^* = \frac{1-a_2\varepsilon_2-a_3\varepsilon_3}{1+a_1\varepsilon_1}$	0.59	0.60
τ at peak of Laffer curve	0.57	0.57

Note: The coefficients (a_1, a_2, a_3) are calculated at the equilibrium with top tax rate $\tau = 0.422$. The elasticities $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ are calculated as a difference quotient. Individual terms are rounded to three digits. The tax rate τ at the peak of the Laffer curve is taken from Figure 3.

value $a_1^{US} = 1.70$ at the 99th percentile of income in 2010 that was calculated in the previous subsection. This value was a target of the calibration exercise. The coefficients (a_2, a_3) are positive in both models. The coefficient a_2 is the ratio of all the (net) tax revenue from agents below the threshold to the aggregate earnings that is taxed at the top rate, whereas a_3 is the ratio of the consumption and capital income taxes paid by agents above the threshold to the aggregate earnings that is taxed at the top rate.

Table 2 also presents the value of the model elasticities $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$. All three elasticities are positive. Thus, the two extra terms in the numerator of the tax rate formula are positive and act to reduce the revenue maximizing top tax rate. The fact that the two extra terms are positive tells one that as the top tax rate τ increases, and $(1 - \tau)$ decreases, less tax revenue is collected from the bottom 99 percent of agent types and that less tax revenue is collected from the top 1 percent in the form of consumption and capital income taxes. This is intuitive as higher top tax rates coincide with smaller aggregate capital and labor inputs. Thus, there is less aggregate consumption and capital income subject to taxation.

Table 2 also shows that the predicted top of the model Laffer curve is not far from the actual top in both models. Thus, one conclusion of this bench test is that the formula accurately predicts the top of the Laffer curve within these specific models. A separate conclusion is that the two extra terms, that are absent from the widely-used formula, are not negligible. For example, the numerator term in the formula for the benchmark model is approximately 0.85. The additional terms depress the revenue maximizing tax rate by an extra 15 percent from what the widely-used formula would suggest. Thus, abstracting from these forces would lead one to overstate the revenue maximizing top tax rate in a quantitatively relevant way.

from Figure 1. X_1 is determined by the grid on (h_1, a) used to compute equilibria - see Appendix A.3. Second, calculate elasticities as a difference quotient based on $\tau = 0.422$ and 0.447. Third, calculate (a_1, a_2, a_3) by computing the ratios in Theorem 1, using the set X_1 and X_2.

5 Discussion

This paper presents a formula for the revenue maximizing top tax rate that applies broadly to static models and to steady states of dynamic models. We raise two issues related to this sufficient statistic formula:

Issue 1: The formula applies to many model frameworks that have been used within macroeconomics, labor economics and public economics. Within a given framework, the formula applies without making parametric assumptions and applies to partial or general equilibrium models. Thus, the use of this sufficient statistic formula would not appear to be very restrictive. This raises the methodological issue of whether estimating three high-level elasticities may be a superior research strategy, compared to estimating the many primitive parameters of a specific parametric economic model and then computing policy counterfactuals, for the purpose of predicting the top of the Laffer curve and providing quantitative guidance for setting a top tax rate.

Issue 2: Suppose that we adopt the view that the long-term response to a permanent tax reform is of most interest for policy making. This view seems to be widely shared. The policy focus is then on the steady-state effects of a permanent change in the top tax rate.

Given this view, are existing elasticity estimates, from the elasticity of taxable income literature, ready to be used as direct inputs into our sufficient statistic formula? We provide three reasons to be cautious in doing so. First, the literature has focused on estimating short-term responses. Saez et al. (2012, p. 13) state *"The long-term response is of most interest for policy making ... The empirical literature has primarily focused on short-term (one year) and medium-term (up to five year) responses ..."*. Second, Badel and Huggett (2014) find that, within their dynamic human capital model, commonly employed regression methods underestimate the elasticity ε_1. Thus, from the perspective of that model framework, there are reasons to believe that existing reduced-form estimates are biased downwards. Third, marginal tax rates applying to US top earners display strong mean reversion. For example, Mertens (2015) uses proxies for exogenous variation in top tax rates and concludes that shocks to US average marginal tax rates for the top 1 percent lead to transitory movements in top tax rates in practice. This raises the important issue of how to estimate long-run elasticities corresponding to a permanent change in the top tax rate when exogenous innovations in top marginal rates lead to movements in top tax rates that are transitory.

References

1. Auerbach, A. and L. Kotlikoff (1987), Dynamic Fiscal Policy, Cambridge University Press, Cambridge.

2. Badel, A. and M. Huggett (2014), Taxing Top Earners: A Human Capital Perspective, manuscript.

3. Ben-Porath, Y. (1967), The Production of Human Capital and the Life Cycle of Earnings, Journal of Political Economy, 75, 352-65.

4. Brewer, M., Saez, E. and A. Shephard (2010), Means Testing and Tax Rates on Earnings, in Dimensions of Tax Design: The Mirrlees Review, Editors S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, M, Gammie, P. Johnson, G. Myles and J. Poterba, Oxford University Press.

5. Chetty, R. (2009), Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods, Annual Review of Economics, 1, 451- 88.

6. Diamond, P. (1965), National Debt in a Neoclassical Growth Model, American Economic Review, 55, 1126- 50.

7. Diamond, P. and E. Saez (2011), The Case for a Progressive Tax: From Basic Research to Policy Recommendations, Journal of Economic Perspectives, 25, 165-90.

8. Feenberg, D. and E. Coutts (1993), An Introduction to the TAXSIM Model, Journal of Policy Analysis and Management, 12, 189-94.

9. Guner, N., Lopez-Daneri, M. and G. Ventura (2014), Heterogeneity and Government Revenues: Higher Taxes at the Top?, manuscript.

10. Guvenen, F., Ozkan, S. and J. Song (2014), The Nature of Countercyclical Income Risk, Journal of Political Economy, 122, 621-60.

11. Heathcote, J., Storesletten, K. and G. Violante (2009), Quantitative Macroeconomics with Heterogeneous Households, Annual Review of Economics, 1, 319-54.

12. Heathcote, J., Perri, F. and G. Violante (2110), Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States: 1967-2006, Review of Economic Dynamics, 13, 15-51.

13. Hendren, N. (2015), The Policy Elasticity, manuscript.

14. Huggett, M., Ventura, G. and A. Yaron (2006), Human Capital and Earnings Distribution Dynamics, Journal of Monetary Economics, 53, 265- 90.

15. Huggett, M., Ventura, G. and A. Yaron (2011), Sources of Lifetime Inequality, American Economic Review, 101, 2923-2954.

16. Kindermann, F. and D. Krueger (2014), High Marginal Tax Rates on the Top 1% ? Lessons from a Life Cycle Model with Idiosyncratic Income Risk, manuscript.

17. Mertens, K. (2015), Marginal Tax Rates and Income: New Time Series Evidence, manuscript.

18. Mirrlees, J. (1971), An Exploration into the Theory of Optimum Income Taxation, Review of Economic Studies, 38, 175- 208.

19. Neal, D. and S. Rosen (2000), Theories of the Distribution of Earnings, Handbook of Income Distribution, (eds) Atkinson and Bourguignon.

20. Piketty, T and E. Saez (2003), Income Inequality in the United States, 1913-1998, Quarterly Journal of Economics, 118, 1-39.

21. Piketty, T and E. Saez (2013a), Optimal Labor Income Taxation, Handbook of Public Economics, Volume 5, editors A. Auerbach, R. Chetty, M. Feldstein and E. Saez, Elsevier.

22. Piketty, T and E. Saez (2013b), A Theory of Optimal Inheritance Taxation, Econometrica, 81, 1851-86.

23. Rubinstein, Y. and Y. Weiss (2006), Post Schooling Wage Growth: Investment, Search and Learning, Handbook of the Economics of Education, (eds) Hanushek and Welch.

24. Saez, E. (2001), Using Elasticities to Derive Optimal Income Tax Rates, Review of Economic Studies, 68, 205-29.

25. Saez, E., Slemrod, J. and S. Giertz (2012), The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review, Journal of Economic Literature, 50:1, 3-50.

26. Trabandt, M. and H. Uhlig (2011), The Laffer Curve Revisited, Journal of Monetary Economics, 58, 305- 27.

27. Weiss, Y. (1986), The Determination of Life Cycle Earnings: A Survey, Handbook of Labor Economics, (eds) Ashenfelter and Layard, Elsevier.

A Appendix

A.1 Data

SSA Data We use tabulated Social Security Administration (SSA) earnings data from Guvenen, Ozkan and Song (2014). We use age-year tabulations of the 10, 25, 50, 75, 90, 95 and 99th earnings percentile for males age $j \in \{25, 35, 45, 55\}$ in year $t \in \{1978, 1979, ..., 2011\}$. These tabulations are based on a 10 percent random sample of males from the Master Earnings File (MEF). The MEF contains all earnings data collected by SSA based on W-2 forms. Earnings data are not top coded and include wages and salaries, bonuses and exercised stock options as reported on the W-2 form (Box 1). The earnings data is converted into real units using the 2005 Personal Consumption Expenditure deflator.

We construct the Pareto statistic at the 99th earnings percentile for age j and year t as follows. We assume that the earnings distribution follows a Type-1 Pareto distribution beyond the 99th percentile for age j and year t. We construct the parameters describing this distribution via the method of moments and the data values for the 95th and 99th earnings percentiles (e_{95}, e_{99}) for a given age and year. The c.d.f. of a Pareto distribution is $F(e; \alpha, \lambda) = 1 - (\frac{e}{\alpha})^{-\lambda}$. We solve the system $.95 = F(e_{95}; \alpha, \lambda)$ and $.99 = F(e_{99}; \alpha, \lambda)$. This implies $\lambda = \frac{\log .05 - \log .01}{\log e_{99} - \log e_{95}}$. To construct the Pareto statistic at the 99th percentile for age j and year t, it remains to calculate the mean earnings for earnings beyond the 99th percentile that is implied by the Pareto distribution for that age and year. The mean follows the formula $E[e|e \geq e_{99}] = \frac{\lambda e_{99}}{\lambda - 1}$.

PSID Data We use Panel Study of Income Dynamics (PSID) data provided by Heathcote, Perri and Violante (2010), HPV hereafter. The data comes from the PSID 1967 to 1996 annual surveys and from the 1999 to 2003 biennial surveys.

Sample Selection We keep only data on male heads of household between the ages of 23 and 62 reporting to have worked at least 260 hours during the last year with non-missing records for labor earnings. In order to minimize measurement error, we delete records with positive labor income and zero hours of work or an hourly wage less than half of the federal minimum in the reporting year.

Variable Definitions The annual earnings variable provided by HPV includes all income from wages, salaries, commissions, bonuses, overtime and the labor part of self-employment income. Annual hours of work is defined as the sum total of hours worked during the previous year on the main job, on extra jobs and overtime hours. This variable is computed using information on usual hours worked per week times the number of actual weeks worked in the last year.

Age-Year Cells We split the dataset into age-year cells and compute the relevant moment within each cell. We put a PSID observation in the (a, y) cell if the interview was conducted during year y and the reported head of household's age was in the interval $[a, a + 4]$ in year y. The life-cycle profiles we calculate correspond to $(\beta_{23} + d, \beta_{24} + d, \beta_{25} + d, ..., \beta_{63} + d)$, where the β_a are the estimated age coefficients and d is a vertical displacement selected in the manner described in section 4.

A.2 Income Threshold and Pareto Statistic

We calculate $(\underline{y}, \bar{y})$ - the 99-th percentile of income in the US in 2010 and the mean income above this threshold in four steps. We use tax units as the unit of measurement and an income concept that is the sum of the following income sources from tabular Statistics of Income (SOI) data: (i) wages and salaries, (ii) interest, (iii) non-qualified dividends, (iv) business income, (v) IRA distributions, (vi) pensions and annuities, (vii) total rent and royalty, (viii) partnership and S-corporation income and (ix) estate and trust income.[18] We exclude capital gains and qualified dividends. Qualified dividends and long-term capital gains are taxed at preferential rates in 2010.

1. SOI Table 1.4 tabulates income, by type of income, for tax units sorted by adjusted gross income (AGI) bins. SOI Table 1.4 also reports the number of tax units in various AGI bins. In 2010, 250 and 500 thousand dollars are the $p_1 = .9823$ and $p_2 = .9947$ percentiles of AGI based on the number of potential tax units in 2010 reported in Piketty and Saez (2003 update).

2. Assume that tax units in the $[250, 500)$ and $[500, \infty)$ AGI bins (in thousands of dollars) are also the tax units that would fall in the two corresponding bins based on the same percentiles for our definition of income. Assume that income is distributed Pareto beyond the p_1 percentile of income. The Pareto cdf is $F(y) = 1 - (\alpha/y)^{\lambda}$. Conditional means satisfy $E[y|y > y_i] = y_i(\lambda/(\lambda - 1))$.

3. Denote (y_1, y_2) the (p_1, p_2) percentiles of our income measure and $(\bar{y}_1, \bar{y}_2)$ the respective means, conditional on income exceeding (y_1, y_2). Calculate $(\bar{y}_1, \bar{y}_2)$ based on tabulated income types in the $[250, \infty)$ and $[500, \infty)$ AGI bins. Solve the equations below to get $(\alpha, \lambda, y_1, y_2)$:

$$1 - p_1 = (\alpha/y_1)^{\lambda},\ 1 - p_2 = (\alpha/y_2)^{\lambda},\ \bar{y}_1 = y_1(\lambda/(\lambda - 1)),\ \bar{y}_2 = y_2(\lambda/(\lambda - 1))$$

4. Solve $(\underline{y}, \bar{y}) = (319.5, 775.8)$ using step 3 and the equations: $.99 = 1 - (\alpha/\underline{y})^{\lambda}$ *and* $\bar{y} = \underline{y}(\lambda/(\lambda - 1))$.

A.3 Computation

The algorithm to compute the model Laffer curve is given below. This computation takes all model parameters and the structure of the tax system below the threshold as given.

Algorithm:

1. Guess $(Tr(\tau), K(\tau)/L(\tau))$ for any top tax rate $\tau \geq 0.422$, keeping all other features of the tax system unchanged. $Tr(\tau)$ denotes the lump-sum transfer given to all agents.

2. Compute the agent's decision rules, given $Tr(\tau)$ and the factor prices (w, r) implied by $K(\tau)/L(\tau)$ from step 1.

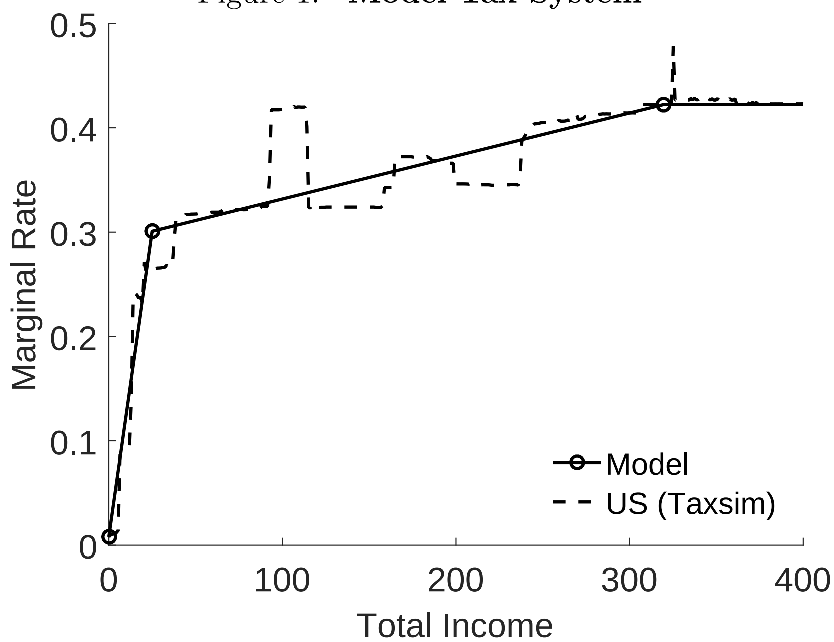
3. Compute the implied values of $(Tr, K/L)$ using the decision rules from step 2, the distribution of initial conditions, the government budget constraint and the fixed initial value for G. If the guessed and implied values are within tolerance, then stop. Otherwise, revise the guess and repeat steps 1-3.

4. The model Laffer curve is the relationship between τ and $Tr(\tau)$.

The distribution of initial conditions employed in Table 1 is discretized. We employ six values for learning ability and fifty values for initial human capital that are conditional on the value of learning ability. Human capital values are equispaced on a log grid following $\log h_1 = \beta_0 + \beta_1 \log a + \log \epsilon$ and vary from -2.5 to 2.5 standard deviations of the log shock innovation. Conditional probabilities are set using a Tauchen procedure based on the parameters in Table 1.

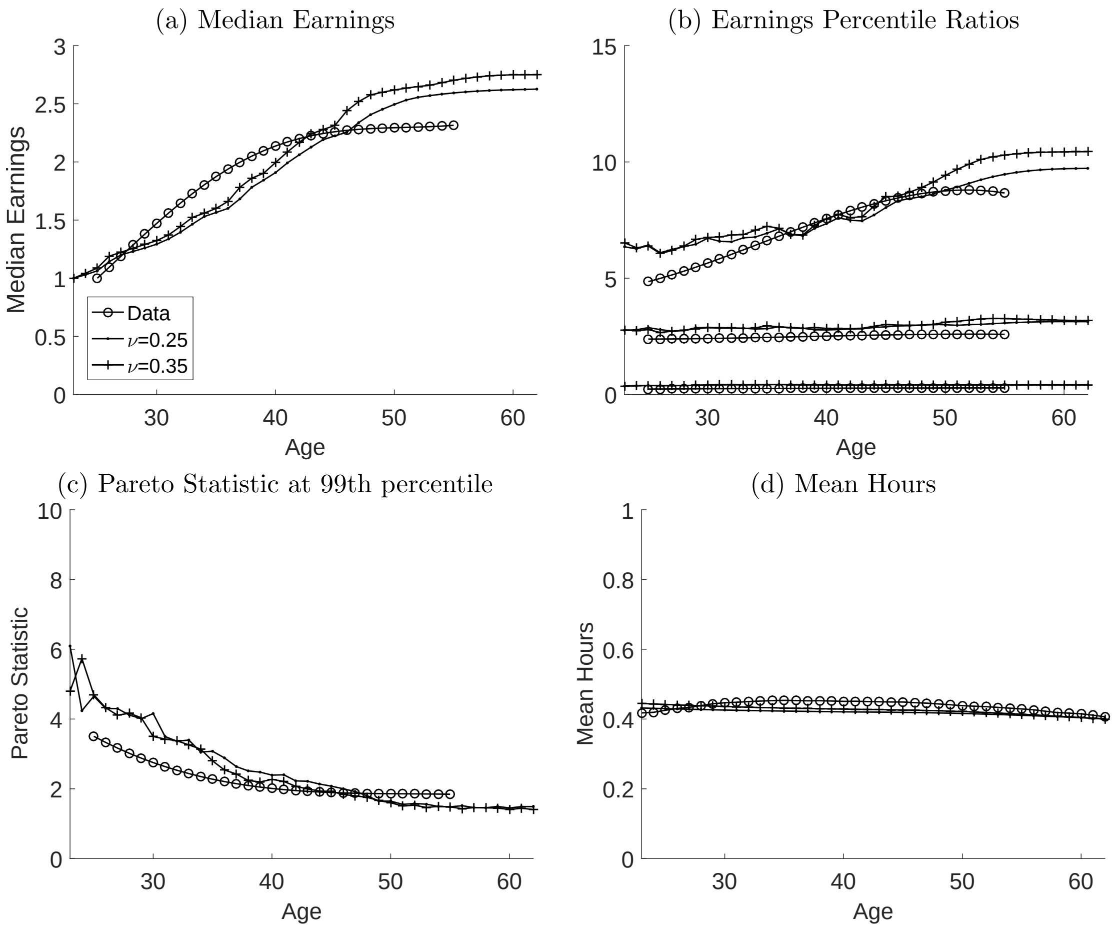
[18] https://www.irs.gov/uac/soi-tax-stats-individual-income-tax-returns-publication-1304-complete-report

Figure 1: **Model Tax System**



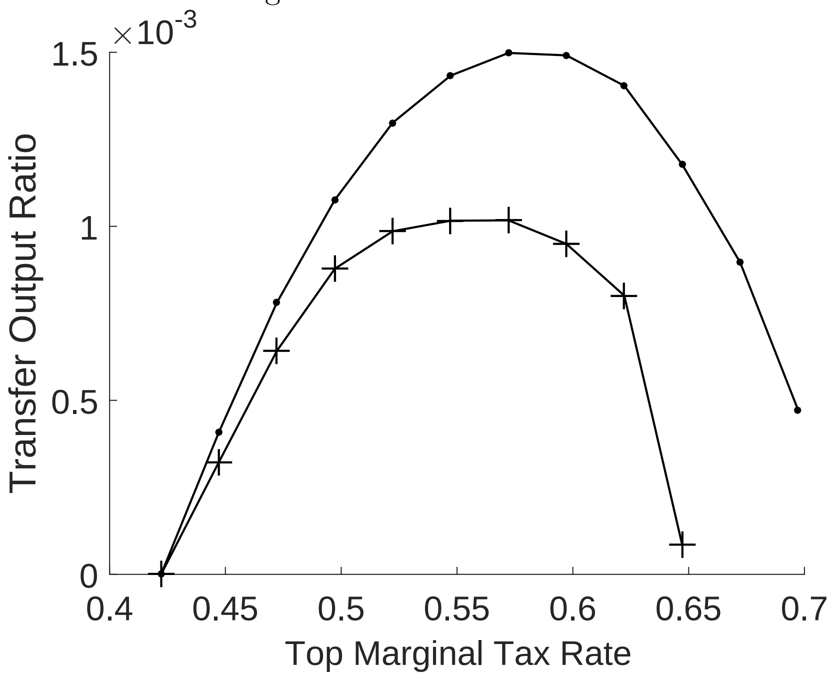
Note: The dashed line is the US schedule produced by TAXSIM. Circles describe the marginal tax rate schedule in the model. Income is in thousand dollar units.

Figure 2: **Life-Cycle Profiles: Data and Model**



Note: Open circles describe profiles for the US economy. The small crosses and the solid line describe profiles for the two model economies with $\nu = 0.35$ and $\nu = 0.25$, respectively. The ratios in Figure 2(b) correspond to the 99-50, 90-50 and the 10-50 earnings percentile ratios.

Figure 3: **Laffer Curves**



Note: Crosses describe the Laffer curve in the benchmark model with $\nu = 0.35$. Circles describe the Laffer curve in the alternative model with $\nu = 0.25$. The vertical axis plots the aggregate transfer-output ratio, where output is the level under the tax system with top rate equal to $\tau = 0.422$.