RESEARCH DIVISION

Working Paper Series

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

Alejandro Badel
and
Mark Huggett

Working Paper 2015-038A
https://doi.org/10.20955/wp.2015.038

November 2015

FEDERAL RESERVE BANK OF ST. LOUIS
Research Division
P.O. Box 442
St. Louis, MO 63166

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

Alejandro Badel
Federal Reserve Bank of St. Louis
ale.badel@gmail.com

Mark Huggett
Georgetown University
mh5@georgetown.edu

THIS DRAFT: NOVEMBER 10, 2015

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. One of the elasticities that enters our formula has been estimated in the elasticity of taxable income literature. We apply standard empirical methods from this literature to data produced by reforming the tax system in a model economy. We find that these standard methods underestimate the relevant elasticity in models with endogenous human capital accumulation.

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 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, was general in that it held within a wide class of theoretical models that were 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 tax rate). Thus, the public policy variable under consideration is a tax rate on some specified 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, following 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 $\tau^* = 1/(1 + a\epsilon)$ that characterizes the revenue maximizing marginal tax rate τ^* that applies beyond a threshold. Moreover, there is also a closely related formula $\tau^* = (1 - g)/(1 - g + a\epsilon)$ for the welfare maximizing marginal tax rate that 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), Piketty and Saez (2013) among many others for a discussion of these formulae. The Mirrlees Review uses them to offer quantitative advice for setting the top income tax rate - see Adam et al. (2010, Chapter 2).

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 $\tau^* = 1/(1 + a\epsilon)$ 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 modern macroeconomics. A large literature analyzes the taxation of consumption, labor income and capital income using these models.[1] The widely-used formula 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.[2]

The sufficient-statistic formula in Theorem 1 of this paper applies to static models and to steady states of dynamic models. Specifically, we show that it applies to the Mirrlees (1971) model, to the two workhorse models of modern macroeconomics and to human capital models. Theorem 1 is stated in terms of three elasticities including the single elasticity of the widely-used formula. One of the new elasticities captures the possibility, within dynamic models, that 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. The other new elasticity captures the possibility that agent types that are above the threshold have incomes or expenditures that are taxed separately. For example, consumption and various types of capital income are commonly taxed separately from labor income. The formula in Theorem 1 allows both of these new elasticities to be non-zero.

This paper makes two contributions. First, Theorem 1 provides a tax rate formula with wide application yet it is stated in terms of only three elasticities. 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 from the bench test is that the formula accurately predicts the top of the Laffer curve even when the top tax rate in the model is initially quite far from the revenue maximizing top tax rate. This is important because prediction is precisely the intended use of the formula in applied work. A second message of the bench test is that standard methods, from the elasticity of taxable income literature, underestimate one key elasticity in the human capital model. An underestimated elasticity would lead one to overstate the revenue maximizing top tax rate. Future work should develop methods to accurately estimate the three key long-run elasticities in different modeling frameworks and then apply these methods to data.

Our paper is most closely related to three literatures. First, it relates to the optimal labor income taxation literature. See Piketty and Saez (2013) for a review that highlights sufficient statistic formulae. Second, it relates to the class of heterogeneous-agent models surveyed by

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

[2] 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. Badel and Huggett (2014), Guner, Lopez-Daneri and Ventura (2014) and Kindermann and Krueger (2014) compute Laffer curves produced by changing top tax rates within such models and examine the resulting steady states. Altig and Carlstrom (1999) analyze implications of changing tax progressivity in a model without idiosyncratic risk.

Heathcote, Storesletten and Violante (2009) because our tax rate formula applies to many models in this large class. Badel and Huggett (2014) apply our tax rate formula to a specific heterogeneous-agent model and find it accurately predicts the top of the model Laffer curve. Our formula could also be applied to the models in Guner, Lopez-Daneri and Ventura (2014) and Kindermann and Krueger (2014) to better understand why these quantitative models have substantially different revenue maximizing top tax rates. Third, it relates to the elasticity of taxable income literature - a key source of elasticity estimates for sufficient statistic formulae. Saez, Slemrod and Giertz (2012) survey this literature. The work of Golosov, Tsyvinski and Werquin (2014) is close in some respects to our work. They analyze local perturbations of tax systems of a more general type than the elementary perturbation that we analyze. Our formula applies to a much wider class of models.

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 static and dynamic models. Section 4 bench tests the formula using a quantitative human capital model. Section 5 concludes.

2 Tax Rate Formula

The tax rate formula is based on three basic 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 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 $n \geq 2$ components of income or expenditure. The three basic 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 the vector of 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. Alternatively, one might view y_1 as being ordinary income and y_2 as being the sum of long-term capital gains and qualified dividends as defined by the Internal Revenue Service in the US. 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

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

(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

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

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

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. Thus, it does not require assumptions on functional forms for the primitives of specific models. However, it does require that certain aggregates are differentiable in the tax parameter.

2. The widely-used formula $\tau^* = 1/(1 + a\epsilon)$ 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 $a_2\epsilon_2$ 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 $a_3\epsilon_3$ will be non-zero when these revenues change. The next two sections give concrete examples of when the terms $a_2\epsilon_2$ or $a_3\epsilon_3$ 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 but without using an elasticity for that integral. For example, if the integral $\int_{X_2} T(y;\tau)dP$ is zero (i.e. total net taxes on agent types below the threshold are zero) and the integral does not vary on the margin as the tax rate τ varies, then the term $a_2\epsilon_2$ 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 Example 1: Mirrlees Model

Mirrlees (1971) considered a static model in which agents make a consumption and labor decision in the presence of a tax and transfer system. 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 a class of tax functions $T(y;\tau)$.

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. *market clearing:* $\int_X c(x;\tau)dP = w \int_X xl(x;\tau)dP$

3. *government budget:* $Tr(\tau) = \tau \int_X wxl(x;\tau)dP + \tau_c \int_X c(x;\tau)dP$

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 the probability space used in Example 1.

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 $a_1 = 1$ as the threshold is $\underline{y} = 0$ and that $\epsilon_1 = \nu$ by a direct calculation of the elasticity.[3] 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 $a_2\epsilon_2$ in the formula can be replaced with a zero, consistent with Comment 3 to Theorem 1. It is easy to see that $(a_3, \epsilon_3) = (\tau_c, \nu)$. 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.

$$\tau^* = \frac{1 - a_2\epsilon_2 - a_3\epsilon_3}{1 + a_1\epsilon_1} = \frac{1 - 0 - \tau_c\nu}{1 + 1 \times \nu} = \frac{1 - \tau_c\nu}{1 + \nu}$$

[3] 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 $(a_1, \epsilon_1) = (1, \nu)$ when these sets are specified for any fixed value of $\tau \in (0, 1)$.

3.2 Example 2: Diamond 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 the 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.

$$\text{(P1)} \quad \max U(c_1, c_2, l) \ s.t.$$

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

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. *factor prices:* $w(\tau) = F_2(K(\tau), L(\tau))$ *and* $1 + r(\tau) = F_1(K(\tau), L(\tau))$
3. *market clearing:* $\int_Z (c_1(z;\tau) + c_2(z;\tau))d\hat{P} + K(\tau) = F(K(\tau), L(\tau))$
4. *government budget:* $Tr(\tau) = \tau_c \int_Z (c_1(z;\tau) + c_2(z;\tau))d\hat{P} + \tau \int_Z w(\tau)zl(z;\tau)d\hat{P}$

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.

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

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 4.

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.[4]

3.3 Example 3: Trabandt and Uhlig

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.

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

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

[4]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 along a balanced-growth path.

$l_t \in [0,1]$ 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. *market clearing:* $\bar{c} + \bar{k}\delta + \bar{g} = F(\bar{k}, \bar{l})$

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 $a_2\epsilon_2$ 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. The intended use of the formula in applied work is to predict 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. Thus, a first step in our bench test is to determine the accuracy properties of using the formula when the relevant inputs are determined away from the maximum. A second step is to estimate a key model elasticity using standard methods from a large empirical literature on tax reforms. This allows us to determine the performance of existing empirical methods, in the context of specific models, because the theoretically-relevant, high-level elasticities are computed separately.

4.1 A Human Capital Model

We bench test the formula using a version of the Ben-Porath (1967) model. This human capital model is a central model in the analysis of the distribution of earnings (see Weiss (1986), Neal and Rosen (2000) and Rubinstein and Weiss (2006)). In our version of this model, agents maximize lifetime utility by choosing time allocation decisions (n_j, l_j, s_j) and by choosing a consumption c_j and asset choice 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. Worker skill evolves according to a function H which is increasing in current skill h_j, learning time s_j and learning ability a.

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

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

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

Taxes are determined by a tax function $T(wh_jl_j;\tau)$ that taxes labor market earnings and by a lump-sum transfer $Tr(\tau)$ received at each age. The tax function and transfers are indexed by a one-dimensional parameter τ. Agents differ in initial conditions which are initial skill and learning ability level (h_1, a). The distribution of these initial conditions is given by a probability measure P. 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.

Definition: *An equilibrium consists of decisions* $(c_j, k_j, n_j, l_j, s_j, h_j)$ *and a government transfers* $Tr(\tau)$ *such that conditions 1 and 2 hold, given an exogenous value for the wage and interest rate* (w, r) *and the tax function parameter* τ*:*

1. *Decisions:* $(c_j(h_1, a; \tau), k_j(h_1, a; \tau), n_j(h_1, a; \tau), l_j(h_1, a; \tau), s_j(h_1, a; \tau), h_j(h_1, a; \tau))$ *solve Problem P1, given* $(w, r, Tr(\tau))$.
2. *Government Budget:* $Tr(\tau) = \sum_{j=1}^{J} \mu_j \int_{H\times A} T(wh_j(h_1, a; \tau)l_j(h_1, a; \tau); \tau)dP$

4.2 Model Parameters

We document properties of the US federal income tax schedule and the US age-earnings distribution. Figure 1 graphs the US federal marginal income tax rates in 2010. The income thresholds in Figure 1 for different tax brackets are stated as multiples of average US income in 2010. The model tax function T is determined by a polynomial.[5] Figure 1 shows that the marginal rates implied by the model tax function closely approximate US federal marginal rates. We associate the tax function parameter τ with the top tax rate in Figure 1. When the parameter τ is increased from the benchmark value of $\tau = 0.35$, the tax function $T(y; \tau)$ implied by this new value of τ is unchanged below the top threshold but differs above the threshold. Thus, the class T of tax functions satisfies assumption A2 from section 2.

Figure 2 documents earnings and hours facts. The earnings facts are based on the age coefficients from a regression of a third-order polynomial in age and a time dummy variable for each year run on tabulated US Social Security Administration (SSA) male earnings data from Guvenen, Ozkan and Song (2014). The earnings facts are (i) median earnings, (ii) the 99-50, 90-50 and 10-50 earnings percentile ratios and (iii) the Pareto statistic. The facts on the average fraction of time spent working are from Panel Study of Income Dynamics (PSID) male hours data from Heathcote, Perri and Violante (2010).[6] Appendix A.1 describes what is measured in the SSA and PSID data sets.[7]

[5] Appendix A.2 describes data sources and approximation methods.

[6] 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).

[7] The age coefficients 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 which is normalized to 100 at age 55. The age coefficients from the PSID hours regression are normalized to pass through the average value across years at age 45 as discussed in Appendix A.1.

Table 1 - Benchmark Model Parameter Values

Category	Functional Forms	Parameter Values
Demographics	$\mu_{j+1} = \mu_j/(1+n)$	$n = 0.01, J = 40$ $j = 1, ..., 40$ (ages 23-62)
Tax System	T	Figure 1 and Appendix
Preferences	$u(c,n) = \log c + \phi \frac{n^{1-1/\nu}}{1-1/\nu}$	$\beta = 0.952, \phi = .231, \nu = 0.25$
Human Capital	$H(h,s,a) = h(1-\delta) + a(hs)^\alpha$	$(\alpha, \delta) = (0.833, 0.00001)$
Initial Conditions	$(\log h_1, \log a) \sim N((\mu_h, \mu_a), \Sigma)$	$(\mu_h, \mu_a) = (5.03, -1.07)$ $(\sigma_h, \sigma_a, \rho_{h,a}) = (.789, .407, .242)$

Note: Parameters for Demographics, the Tax System and that utility function parameter ν are preset without solving for equilibrium. The remaining parameters are set so that the model equilibrium best matches the facts in Figure 2. Parameters are rounded to 3 significant digits. The parameters $(\sigma_h, \sigma_a, \rho_{h,a})$ refer to the standard deviation of log human capital and log learning ability and the correlation between these log variables.

Figure 2 calculates how the Pareto statistic at the 99th percentile for earnings varies with age. The Pareto statistic is the mean for observations above a threshold divided by the difference between this mean and the threshold. We set the threshold to be the 99th percentile for each age and year in the data set. We highlight the Pareto statistic because the cross-sectional value of the Pareto statistic is the coefficient a_1 in the tax rate formula in Theorem 1. We focus on the Pareto statistic at the 99th percentile because the top federal tax rate iin 2010 begins at approximately the 99th percentile of income as displayed in Figure 1.

We specify functional forms for the utility function u, human capital H, the tax function T and for the distribution of initial conditions. Table 1 presents functional forms and parameter values. We preset some model parameters. Specifically, we set $(w, r, n, J) = (1.0, .04, .01, 40)$ so that the wage is normalized to 1, the real interest rate is 4 percent, the population growth rate is 1 percent and the working lifetime is 40 model periods covering a real-life age of 23 to 62. We preset the tax function T as described above. Finally, we also preset the utility function parameter to $\nu = 0.25$. This implies that the model has a constant Frisch elasticity of leisure with respect to human capital equal to $-\nu$ and a Frisch elasticity of total labor time $(s_j + l_j)$ equal to $\nu(n_j/(l_j + s_j))$.[8] Thus, the model has a Frisch elasticity of total labor time equal to .375 when $\nu = .25$ and $n_j/(l_j + s_j) = 1.5$

We set all remaining model parameters (governing preferences, human capital and initial conditions) to minimize the weighted squared deviation of the log of equilibrium model moments from the log of the data moments in Figure 2. Figure 2 plots the corresponding model econ-

[8]The standard necessary conditions for a solution to Problem P1 with constant marginal tax rates imply that over the life cycle $\Delta \log n_j = \nu[\log \beta(1+r) - \Delta \log h_j]$. Thus, after correcting for a trend term, leisure decreases by ν percent after a 1 percent evolutionary increase in human capital across periods.

omy moments that result from this minimization problem.[9] Table 1 lists the value of model parameters.

We stress that the goal of the bench test is not to analyze a rich model that accounts for many aspects of the US tax system and many aspects of the earnings and hours distribution in US data. Instead, the bench test highlights whether or not the formula is useful in predicting the top of the Laffer curve in interesting applications for which the answer is not known in advance. We analyze a dynamic model because our work claims that the formula applies to both static and dynamic models. We choose to analyze a dynamic human capital model for two reasons. First, skill accumulation is a natural explanation for why top earners are disproportionately age 50 and above - see Figure 2. Second, we hypothesize that current methods for estimating earnings elasticities that are in wide use in the literature will systematically underestimate the elasticities that are theoretically relevant for determining the top of the Laffer curve. If true, then bench testing the formula on a human capital model will be especially valuable as it will highlight a potential weakness in the current status of a large empirical literature.

4.3 Theorem 1 and Model Laffer Curves

Figure 3 presents Laffer curves. The Laffer curve for the benchmark model is based on the parameters in Table 1.[10] The Laffer curve for the other model in Figure 3 is based on setting $\nu = 0.10$ and resetting the remaining model parameters to best match targets following the procedures for the benchmark model. The goal is to construct model economies where the top of the Laffer curve occurs at top tax rates that are at varying distances from the initial top rate in Figure 1. These models are then used to bench test the accuracy of the formula when the inputs to the formula are calculated away from the maximum.

We now relate the top of the model Laffer curve to the top predicted by the tax rate formula. To do this, we map equilibrium variables into the language used in Theorem 1. Agent types have to be defined in terms of exogenous variables to apply Theorem 1. A natural formulation is that an agent type is $x = (h_1, a, j)$ and is determined by the initial condition (h_1, a) and age j.

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

[9] There is a tension in choosing model parameters governing initial conditions. For example, increasing the variance of log learning ability σ_a^2 tends to reduce the model Pareto statistic at the 99th percentile but increase the 99-50 ratio in Figure 2. We conjecture that a more flexible class of distributions over initial conditions is key to a better fit to the upper tail of the earnings distribution.

[10] The model Laffer curve is calculated by varying the top tax rate τ, computing the model equilibrium for each value of τ and plotting the resulting equilibrium total tax revenue. Transfers per agent equal total taxes per agent as implied by the government budget constraint.

Table 2: Revenue Maximizing Top Tax Rate Formula

Terms	Benchmark Model $\nu = 0.25$	Alternative Model $\nu = 0.10$
$a_1 \times \epsilon_1$	$2.29 \times .550 = 1.26$	$1.68 \times .258 = .433$
$a_2 \times \epsilon_2$	$3.26 \times .010 = .031$	$1.83 \times .001 = .002$
$a_3 \times \epsilon_3$	0	0
$\tau^* = \frac{1-a_2\epsilon_2-a_3\epsilon_3}{1+a_1\epsilon_1}$	0.43	0.70
τ at peak of Laffer curve	0.425	0.675

Note: The coefficients (a_1, a_2, a_3) are calculated at the equilibrium with top tax rate $\tau = 0.35$. The elasticities $(\epsilon_1, \epsilon_2, \epsilon_3)$ are calculated as a difference quotient. Individual terms in the formula are rounded to at most three digits to provide a simple numerical display. The tax rate τ at the peak of the Laffer curve are taken from Figure 3.

Step 2: $y_1(x, \tau) = wh_j(h_1, a; \tau)l_j(h_1, a; \tau)$

Step 3: $T(y_1; \tau) = T(y_1; \tau)$

Based on Steps 1-3, we calculate the coefficients (a_1, a_2, a_3) and the elasticities $(\epsilon_1, \epsilon_2, \epsilon_3)$.[11] We analyze each of the model economies underlying the Laffer curves in Figure 3. Table 2 presents the results. The elasticity ϵ_1 is quite different across models. The lower value of the leisure elasticity parameter $\nu = 0.1$ is associated with an earnings elasticity ϵ_1 that is roughly half of the earnings elasticity in the benchmark economy. The coefficient a_1 is also quite different across models. This occurs despite the fact that the calibration procedure keeps the targets fixed. In practice the model Pareto statistic at the threshold for the top tax rate is close to the corresponding model Pareto statistic among agents age 50.

One conclusion from the bench test is that the predicted top of the model Laffer curve is not far from the actual top. This holds in both of the models analyzed. This occurs even though the top of one of the Laffer curves occurs at a top tax rate quite far from the initial top tax rate that was used to calculate the coefficients and elasticities that enter the formula. Thus, using the Badel-Huggett formula to predict the top of the Laffer curve within these specific dynamic models is not especially problematic when inputs are calculated far away from the top of the Laffer curve.

[11] To calculate elasticities, set $X_1 = \{x \in X : y_1(x; \tau) = wh_j(h_1, a; \tau)l_j(h_1, a; \tau) > \underline{y}\}$ using the benchmark value $\tau = .35$ and the threshold $\underline{y}$ from Figure 1. In practice, the set X_1 is determined by a grid on (h_1, a) used to compute equilibria as described in Appendix A.3. Elasticities are calculated as a difference quotient based on $\tau = 0.35$ and 0.40.

4.4 Comparing Model Elasticities to Estimated Elasticities

Saez, Slemrod and Giertz (2012) review the literature that estimates earnings or income elasticities with respect to the net-of-tax rate. Much of the literature applies the regression framework below. The parameter ϵ is the elasticity with respect to the net-of-tax rate, z_{it} is income or earnings of individual i at time t, $\tau_t(z_{it})$ is the marginal tax rate at time t that corresponds to income z_{it}, $f(z_{it}) = \log z_{it}$ is an income control and α_t are time dummy variables. The variable $(1 - \tau_t(z_{it}))$ is refered to as the net-of-tax rate. This regression framework was used by Gruber and Saez (2002) among many others.

$$\log\left(\frac{z_{it+1}}{z_{it}}\right) = \epsilon \log\left(\frac{1 - \tau_{t+1}(z_{it+1})}{1 - \tau_t(z_{it})}\right) + \beta f(z_{it}) + \alpha_t + \nu_{it+1}$$

We follow the literature and apply this regression framework. Unlike the literature, we apply it to a tax reform within the benchmark human capital model. In model period 1 and 2 agents live in the economy with top tax rate $\tau = 0.35$ set to the value in the benchmark model. In model period 3 agents are surprised to discover that the top tax rate is permanently changed to $\tau = 0.45$. We analyze the tax reform assuming that transfers are constant at the initial steady-state level. We draw a sample of 30,000 agents between the ages of 23-55 that have earnings within the top 10 percent of the earnings distribution. We follow these agents for 7 model periods. Model period 3 is the year of the tax reform.

We run the regression above on this sample. We estimate the elasticity using the three different instrumental variables procedures described and employed in Saez, Slemrod and Giertz (2012, Table 2). The precise instruments employed are described in Table 3. Table 3 presents the mean and the standard deviation of the estimated elasticity after applying this framework to 100 different randomly drawn balanced panels of 30,000 agents. The main finding is that the mean of the estimated elasticities are consistently far below the true long-run model elasticity. Moreover, standard errors are quite small so that sampling variability is not the explanation for why estimated elasticities are below long-run elasticities. The reader should keep in mind that the long-run model elasticity enters the sufficient statistic formula and accurately predicts the tax rate at the top of the Laffer curve.

5 Discussion

This paper has two main contributions. First, the formula in Theorem 1 applies broadly to static models and to steady states of dynamic models and yet depends on only three elasticities. Thus, this sufficient-statistic formula should replace the widely-used formula in future work.

Table 3 - Elasticity Estimates

Mean Elasticity ϵ	-0.067	0.112	0.091	0.131	0.090	0.131
S.D.	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Income Control $f(z)$	No	Yes	No	Yes	No	Yes
Time Effects α_t	No	No	Yes	Yes	Yes	Yes
Instrument 1: $1_{\{i \in T_2\}}$	Yes	Yes				
Instrument 2: $1_{\{i \in T_2 \ and \ t=2\}}$			Yes	Yes		
Instrument 3: $\log(\frac{1-\tau_{t+1}(z_{it})}{1-\tau_t(z_{it})})$					Yes	Yes
Use data for time periods	$t = 2, 3$	$t = 2, 3$	All	All	All	All
Long-run Model Elasticity ϵ_1	0.550	0.550	0.550	0.550	0.550	0.550

Note: The set T_2 is the set of agents in the balanced panel that in model period 2 have earnings above the earnings threshold at which the top tax rate applies. Instrument 1 is the indicator function taking the value 1 when the agent lies in T_2, whereas instrument 2 is the indicator function taking the value 1 when the agent lies in T_2 and the time period is $t = 2$.

Second, the formula from Theorem 1 works well in predicting the top of the Laffer curve in a human capital model. However, state-of-the-art methods for estimating elasticities, drawn from the literature on the elasticity of taxable income, under estimate one of the theoretically-relevant model elasticities when applied to an ideal data set.

These two contributions raise two important issues:

Issue 1: The formula developed in this paper applies to many model frameworks that have been widely used by macroeconomists and public economists. Within a given framework, the formula applies without making parametric assumptions. 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 parameters of the primitives of a specific parametric economic model and then computing policy counterfactuals, for the purpose of predicting the top of the Laffer curve.

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, the results in Table 3 show in a specific context that empirical methods commonly used to estimate a short-term elasticity underestimate the true long-run model elasticity. Third, marginal tax rates applying to US top earners display strong mean reversion. Mertens (2015) uses proxies for exogenous variation in tax rates to argue that shocks to US 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 some of the exogenous variation in top marginal rates highlighted in the data lead to only transitory movements in these tax rates.

References

1. S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, M. Grammie, P. Johnson, G. Myles and J. Poterba (eds.)(2010), Dimensions of Tax Design: The Mirrlees Review, Oxford University Press, Oxford, UK.

2. Altig, D. and C. Carlstrom (1999), Marginal Tax Rates and Income Inequality in a Life-Cycle Model, American Economic Review, 89, 1197-1215.

3. Alvaredo, F., Atkinson, A., Piketty, T. and E. Saez, The World Top Incomes Database, http://topincomes.g-mond.parisschoolofeconomics.eu/ .

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

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

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

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

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

9. 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.

10. Golosov, M., Tsyvinski, A. and N. Werquin (2014), A Variational Approach to the Analysis of Tax Systems, manuscript.

11. Gruber, J. and E. Saez (2002), The Elasticity of Taxable Income: Evidence and Implications, Journal of Public Economics, 84, 1-32.

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

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

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

15. 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.

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

17. 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.

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

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

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

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

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

23. 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.

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

25. 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, compute the relevant moment within each cell and then collapse the dataset so there is a single observation per age-year cell. We put a PSID observation in the (a, y) cell if the interview was conducted during year $y = 1968, 1970, 1971, ..., 1996$ or $y = 1998, 2000, 2002$ with reported head of household's age a in the interval $[a, a + 4]$. 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 Tax Function

The model tax function is described in three steps.

Step 1: Specify the empirical tax function $\hat{T}(x)$:

$$\hat{T}(x) = \begin{cases} R_1[x - q_1] & i(x) = 1 \\ \sum_{n=2}^{i(x)} R_{n-1}[q_n - q_{n-1}] + R_{i(x)}[x - q_{i(x)}] & i(x) > 1 \end{cases}$$

$$i(x) \equiv \max n \; s.t. \; n \in \{1, 2, ..., N\} \; and \; q_n \leq x$$

The values $\{(q_1, R_1), ..., (q_7, R_7)\}$ are set based on the seven tax brackets and tax rates for the 2010 federal income tax schedule for married couples filing jointly. Brackets and rates come from Schedule Y-1 in the IRS Form 1040 Instructions. Adding \$18,700 to each of the taxable income brackets from Schedule Y-1 generates total income cutoffs that produce these taxable income cutoffs in Schedule Y-1 for joint filers without dependents according to the NBER tax program TAXSIM for the 2010 tax year. Total income brackets q_n are stated as multiples of average income in 2010.[12]

Step 2: Fit the 5th order polynomial $P(x; \zeta)$ to $\hat{T}(x)$:

$$\zeta \in \; argmin \sum_{x_i \in X^{grid}} (\hat{T}(x_i) - P(x_i; \zeta))^2 \; subject \; to \; P(0; \zeta) = 0, P'(q_7; \zeta) = 0.35$$

X^{grid} contains 51 points uniformly distributed on the interval $[0, q_7]$.

Step 3: Set the model tax function(s) $T(e; \tau)$:

$$T(e; \tau) = \begin{cases} \bar{e}P(e/\bar{e}; \zeta) & e \leq q_7\bar{e} \\ \bar{e}P(q_7; \zeta) + \tau[e - q_7\bar{e}] & e > q_7\bar{e} \end{cases}$$

The quantity $\bar{e}$ is average earnings in the model. Note that $P(x; \zeta)$ from Step 2 takes input x stated in multiples of average income and states output in units of average income.

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. The tax system below the threshold is expressed as a function of $\bar{e}$ as described in Appendix A.2. Thus, to compute the Laffer curve we fix the value $\bar{e}$ at its value in the benchmark model.

Algorithm:

1. Guess transfers $Tr(\tau)$ for any top tax rate $\tau \geq 0.35$.

2. Solve Problem DP-1, given $Tr(\tau)$.

[12] The World Top Incomes Database reports that average income per tax unit in 2010 is 53,347 stated in 2014 dollars. Using the CPI, this is equivalent to 49,278 dollars in 2010. The US top bracket starts at a ratio of total income to average income equal to $q_7 = (373,650 + 18,700)/49,278 = 7.96$, where 373,650 is the start of the top taxable income bracket and 18,700 is the standard deduction for joint filers without dependents.

3. Compute the implied values of $Tr(\tau)$ using the optimal decision rules from step 2 and a discretized initial distribution. If the guessed and implied values are within tolerance, then stop. Otherwise, revise the guess and repeat steps 2-3.

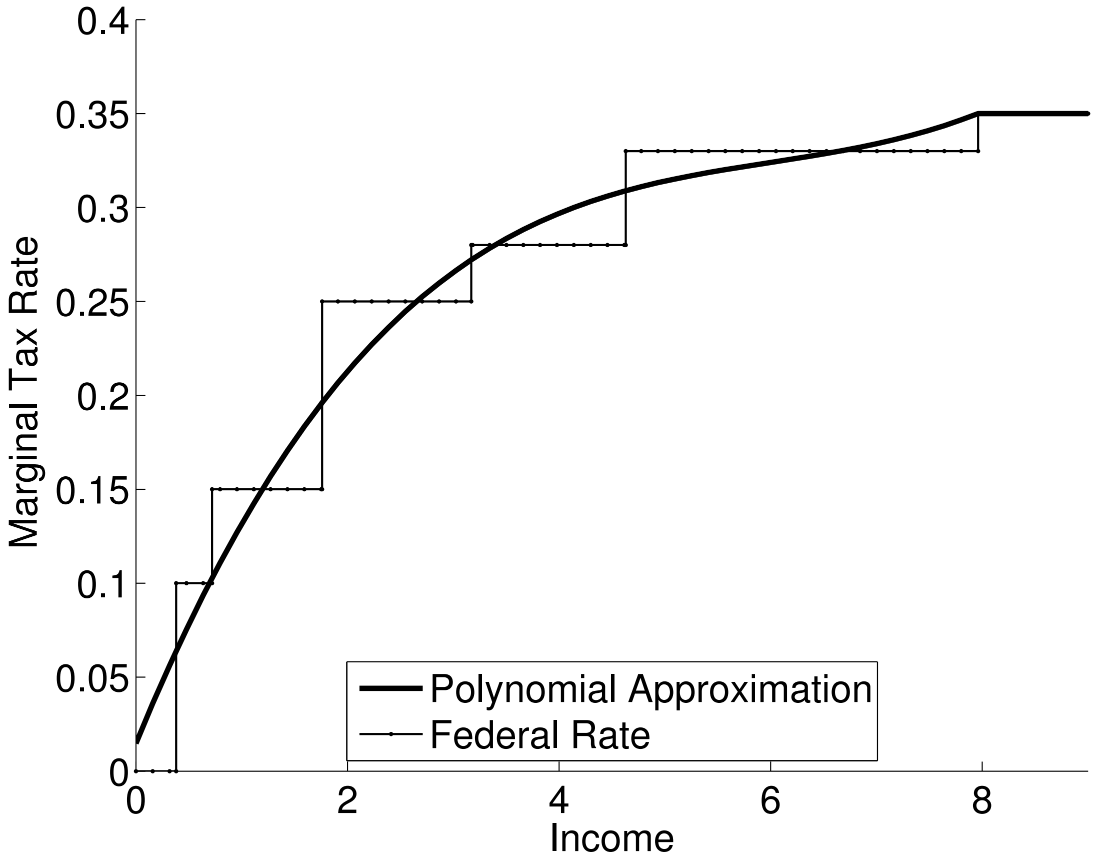
We solve Problem 1 by solving the dynamic programming problem DP-1 below. For each age and state (h, k, a) on a grid, we solve an inner and an outer problem. The inner problem takes choices (k', s) as given and solves for (c, l). The inner problem is easy to solve by bisection. This is true when T is differentiable almost everywhere and T' is increasing. The class of model tax functions satisfies both properties. The outer problem maximizes over (k', s), using solutions to the inner problem. The outer problem is solved by a Nelder-Mead algorithm.

$$(DP-1) \;\; v_j(h, k, a) = \max \; u(c, 1-s-l) + \beta v_{j+1}(h', k', a) \text{ subject to}$$

$$c + k' \leq whl + k(1+r) - T(whl) \;, c, s, l, k' \geq 0, \; 0 \leq s + l \leq 1 \text{ and } h' = H(h, s, a)$$

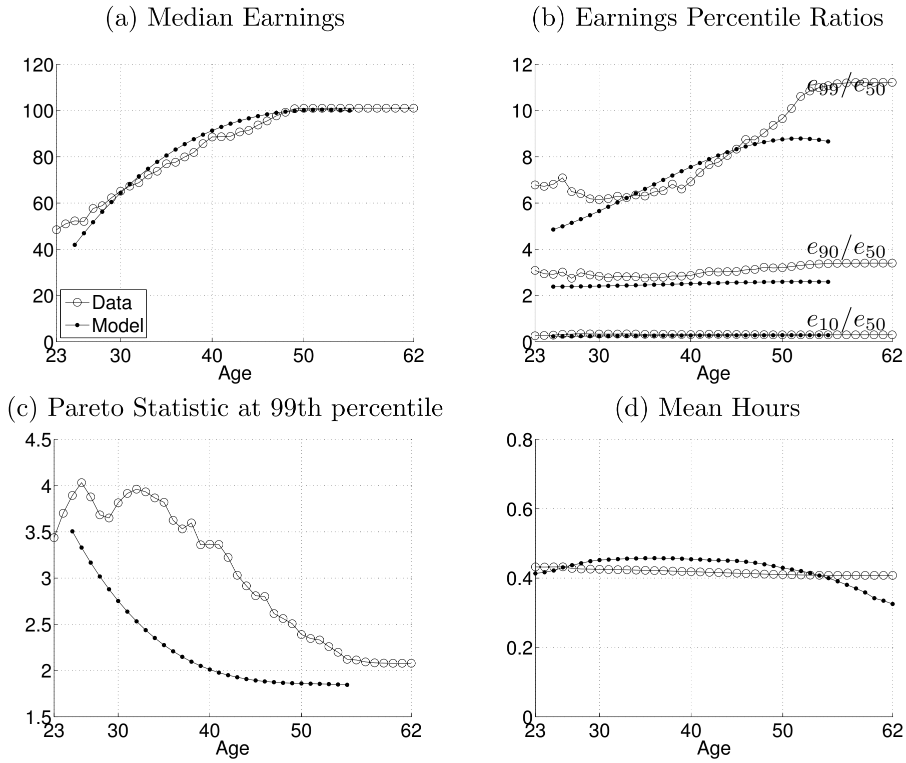
To compute the implied value of $Tr(\tau)$, we use a discrete approximation to the bivariate lognormal $N(\mu, \Sigma)$. We use 16 points for $\log a$ and 61 points for $\log h_1$. The grid points for $\log a$ and $\log h_1$ range from 3.99 standard deviations above and below the mean of the respective marginal distributions. For learning ability, we put 8 points below the 90th percentile and 8 points above the 90th percentile.

Figure 1: **Model Tax System**



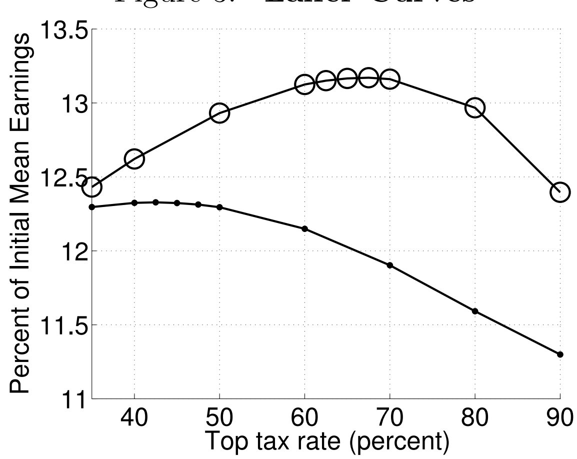
Note: The horizontal axis measures income in multiples of average US income per tax unit. Appendix A.2 discusses data sources for the 2010 US federal tax schedule and approximation methods used to compute the model tax system.

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



Note: Large open circles describe profiles for the model economy. Small solid circles describe profiles for the U.S. economy.

Figure 3: **Laffer Curves**



Note: Small solid circles describe the Laffer curve in the benchmark model with $\nu = 0.25$. Large open circles describe the Laffer curve in the alternative model with $\nu = 0.10$.