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: March 9, 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 not only applies to static models, as in previous literature, but also to steady states of dynamic models. One of the three elasticities entering our formula has been estimated using the methods from the elasticity of taxable income literature. We examine the accuracy of these empirical methods and find that they work poorly 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 convincingly 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 widely viewed as relevant to the issue at hand. It seems to be clear that such work, satisfying the trinity of theoretical, empirical and policy relevance, would be celebrated.

The scenario just described is what the sufficient statistic approach hopes to achieve. Chetty (2009) reviews applications 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 agreed that pushing a tax rate beyond the revenue maximizing rate is conterproductive. If so, then an accurate prediction of this rate would usefully narrow 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 already is a widely-used sufficient statistic formula $\tau^* = 1/(1 + a\epsilon)$ that characterizes the revenue maximizing marginal tax rate τ^* that applies beyond an income 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.[1] These formulas are discussed by Saez (2001), Diamond and Saez (2011) and Piketty and Saez (2013) among others. These formulas are incorporated into the Mirrlees Review - an important synthesis of current thinking about what constitutes good tax policy (see Mirrlees et al. (2010)).

A more critical reading of the literature is 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 either the infinitely-lived agent or the overlappping 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 within these models.[2] Also, the widely-used formula does not apply to the class of heterogeneous-agent models that are currently the dominant positive models of the distribution of earnings, income, consumption and wealth.[3]

[1] These two formulas hold within versions of the Mirrlees (1971) model.

[2] Auerbach and Kotlikoff (1987) is an early contribution.

[3] Heathcote, Storesletten and Violante (2009) review this literature. The models that they review feature

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 and to the two workhorse models of modern macroeconomics. It also applies broadly to heterogeneous-agent models. Recently, Badel and Huggett (2014), Guner, Lopez-Daneri and Ventura (2014) and Kindermann and Krueger (2014) analyzed the Laffer curve produced by increasing the marginal tax rate on top earners using specific quantitative models within this class. We show that the formula in Theorem 1 of this paper applies to these types of model economies. Moreover, it accurately predicts the top of the Laffer curve within these models even when the initital steady state tax rate is far from the revenue maximizing tax rate. Thus, the formula builds a bridge between the vast body of quantitative work that analyzes steady states of growth models and the application of a sufficient statistic approach to an important issue in macroeconomics and public economics.

This paper makes two contributions. First, Theorem 1 provides a tax rate formula with wide application. Despite this wide applicability, it is stated in terms of only three elasticities. Thus, the formula in Theorem 1 should replace the widely-used formula in future work. Second, we show that the formula accurately predicts the top of the Laffer curve in a dynamic human capital model. However, when we apply standard methods to estimate a key elasticity, from the large literature on the elasticity of taxable income, to data produced by a tax reform performed within the model, these methods underestimate a key model elasticity. This highlights a key direction for future work.

Our paper is most closely related to three literatures: the optimal tax literature, the elasticity of taxable income literature and the literature on heterogeneous-agent models. These literatures are reviewed by Piketty and Saez (2013), Saez, Slemrod and Giertz (2012) and Heathcote et al. (2009), respectively. Our paper is related to all three literatures because it derives a sufficient statistic formula that applies not only to the static models from the original optimal tax literature but also to the dynamic models from the growth literature and the literature on heterogeneous-agent models. Moreover, we examine the accuracy of methods for estimating a key elasticity drawn from the elasticity of taxable income literature. Our paper is also related to the work of Golosov, Tsyvinski and Werquin (2014). They analyze local perturbations of tax systems of a more general type than the elementary perturbation analyzed here.

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 applies the formula to a quantitative human capital model and examines the accuracy of existing methods for estimating a key elasticity. Section 5 concludes.

2 Tax Rate Formula

Our 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

agents that are ex-ante heterogeneous, face idiosyncratic risk and trade in incomplete financial markets.

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 τ^* is the value of the top tax rate that maximizes revenue.

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

$A2$. *There is $\underline{y} \geq 0$ 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. All of these possibilities are consistent with the additive separability assumption. Assumption $A2'$ states that, for the first component of income y_1, any income above a threshold $\underline{y}$ is taxed at a fixed rate τ. 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 τ^* is the value of the top tax rate that maximizes revenue.

$A1'$. *$(X, \mathcal{X}, \mathrm{P})$ 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 $\underline{y} \geq 0$ 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

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

(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

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

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

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 by assumption $A2'$ it also maximizes $\tau \int_{X_1} (y_1 - \underline{y})dP + \int_{X_1} T_2(y_2,...,y_n)dP + \int_{X_2} T(y_1,...,y_n;\tau)dP$. 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 and, thus, does not require explicit assumptions on functional forms for the model primitives. However, it does require that certain aggregates are differentiable in the tax parameter as stated in assumption $A3$ or $A3'$. The next two sections show how to map equilibria in specific static or dynamic models into the three basic model elements used to state Theorem 1.

2. The widely-used formula $\tau^* = 1/(1 + a\epsilon)$ is effectively a special case of the sufficient statistic formula in Theorem 1. It is useful to understand at an intuitive level what the widely-used formula misses that the formula in Theorem 1 successfully addresses. There are two such problems. The first problem is that when the tax rate τ changes agent types below the threshold, in the set X_2, will in general have their income and expenditures $(y_1, ..., y_n)$ and corresponding tax liabilities change. This can happen, in static or dynamic models, because factor prices change due to the response from agent types above the threshold in the set X_1. In dynamic models this can also happen because agent types in X_2 anticipate being directly

impacted by the tax change later in life with positive probability when they have income above the threshold. Therefore, they adjust their income and expenditure choices before they are directly impacted. These are all circumstances in which 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 problem is that many components of income or expenditure are taxed in practice. If the parameter τ changes then agent types above the threshold, in the set X_1, will adjust other components of income or expenditure $(y_2, ..., y_n)$. The revenue consequences of such adjustments need to be accounted for. The term $a_3\epsilon_3$ will be non-zero when this occurs. 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 allocations in each model into the language of Theorem 1.

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: $(c(x;\tau), l(x;\tau)) \in argmax\ u(c,l)\ s.t.\ c \leq wxl - T(wxl;\tau) + Tr(\tau), l \geq 0$
2. market clearing: $\int_X c(x;\tau)dP = w\int_X xl(x;\tau)dP$
3. government budget: $Tr(\tau) = \int_X T(wxl(x;\tau);\tau)dP$

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

$T(y;\tau) = \tau y$

Equilibrium allocations are straightforward to state:

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

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

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

We now map equilibrium allocations into the elements $((X, \mathcal{X}, P), y(x;\tau), T(y;\tau))$ used to state Theorem 1.

Step 1: Set $(X, \mathcal{X}, P)$ to the probability space used in Example 1.
Step 2: Set $y(x;\tau) = wx[\frac{wx(1-\tau)}{\alpha}]^{\nu}$
Step 3: Set $T(y;\tau) = \tau y$.

We now calculate the terms in the formula. It is clear that $a_1 = 1$ as the threshold is $\underline{y} = 0$. It is also clear that $\int_{X_2} T(y;\tau) dP = 0$ as only agent types with $x = 0$ are in the set X_2. These agent types produce no labor income and generate zero tax revenue for all tax rates τ. Thus, the term $a_2\epsilon_2$ in the formula can be replaced with a zero, consistent with Comment 3 to Theorem 1. Finally, it is clear that $\epsilon_1 = \nu$ by a direct calculation of the elasticity.[4] In summary, the revenue maximizing tax rate depends only on the elasticity ϵ_1. This "high-level" elasticity coincides with the preference parameter ν, given the form of the utility function.

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

3.2 Example 2: Diamond Growth Model

Diamond (1965) analyzes an overlapping generations model with two-period lived agents and a neoclassical production function $F(K, L)$. 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.

(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 $l \in [0,1]$

[4] 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 $\tau \in (0,1)$.

Age 1 agents are heterogeneous in labor productivity $z \in Z \subset R_+$. The distribution of labor productivity is given by $(Z, \mathcal{Z}, \hat{P})$. 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. To apply Theorem 1, an agent type needs to be stated in terms of **exogenous variables.** An agent type does not vary as the tax parameter τ varies.

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

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

$$y_1(x;\tau) = w(\tau)zl(z;\tau) \text{ if } x = (z,1) \text{ and } z \in Z \text{ and } y_1(x;\tau) = 0 \text{ otherwise}$$

$$y_2(x;\tau) = c_1(z;\tau) \text{ if } x = (z,1) \text{ and } z \in Z \text{ and } y_2(x;\tau) = c_2(z;\tau) \text{ otherwise}$$

Step 3: Set $T(y_1, y_2; \tau) = \tau y_1 + \tau_c y_2$. With this choice, 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.

3.3 Example 3: Trabandt and Uhlig

Trabandt and Uhlig (2011) analyze Laffer curves using the neoclassical growth model with 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.

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

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

$l_t \in [0,1]$ 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 $\bar{g} + \bar{b}(\bar{R}^b - 1) + \bar{Tr} = \tau_c \bar{c} + \tau_k \bar{k}\bar{r} + \tau_l \bar{w}\bar{l}$
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 $(\bar{g}, \bar{b})$. Thus, 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: $T(y_1, y_2, y_3; \tau) = \tau y_1 + \tau_c y_2 + \tau_k y_3 - \bar{g} - \bar{b}(\bar{R}^b(\tau) - 1)$

For the tax system in step 3 to satisfy assumption $A2'$, we need (i) separability (i.e. $T = T_1 + T_2$), (ii) T_1 is eventually linear in τ and (iii) T_2 depends on τ only indirectly via $(y_2, ..., y_n)$. Setting $T(y_1, y_2, y_3; \tau) = T_1(y_1; \tau) + T_2(y_2, y_3)$, where $T_1(y_1; \tau) = \tau y_1$ and $T_2(y_2, y_3) = \tau_c y_2 + \tau_k y_3 - \bar{g} - \bar{b}(\bar{R}^b(\tau) - 1)$, is potentially problematic. Without further argument, the tax system does not satisfy property (iii). Specifically, since the government bond return $\bar{R}^b(\tau)$ enters T_2, property

(iii) may be violated. However, $\bar{R}^b(\tau) = 1/\beta, \forall \tau \in [0,1)$ in steady state follows directly from the Euler equation. Thus, the tax function in Step 3 satisfies $A2'$.

Given this mapping, the coefficients pre-multiplying the elasticities are easy to calculate. The two relevant coefficients are $(a_1, a_3) = (1, \frac{\tau_c \bar{c}(\tau^*) + \tau_k \bar{k}(\tau^*)\bar{r}(\tau^*) - \bar{g} - \bar{b}(\bar{R}^b(\tau^*)-1)}{\bar{w}(\tau^*)\bar{l}(\tau^*)})$. 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.

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

The take-away point is that for each of the 15 quantitative 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.

4 A Human Capital Model

We now apply the sufficient statistic approach within the human capital model developed by Badel and Huggett (2014). This model is calibrated to match a number of features of the US age-earnings distribution. In the model, government lump-sum transfers vary as the tax rate on top earners is varied, producing a Laffer curve.

This section has two goals. First, it examines whether the formula in Theorem 1 accurately predicts the top of the model Laffer curve under almost ideal conditions. Specifically, the three coefficients (a_1, a_2, a_3) and corresponding elasticities $(\epsilon_1, \epsilon_2, \epsilon_3)$ are computed using model data at the original steady state, where the top tax rate does not maximize revenue or transfers. Proponents of the sufficient statistic approach (e.g. Diamond and Saez (2011) and the references cited in the introduction) also apply the widely-used formula $\tau^* = 1/(1 + a\epsilon)$ away from the optimum to predict the top of the Laffer curve. Second, it applies existing methods from the elasticity of taxable income literature to model data in order to gauge the extent to which the estimation methods from this literature recover the theoretically-relevant elasticity ϵ_1.

4.1 Equilibrium

We outline the structure of the Badel-Huggett model. Agents solve decision problem P1. Thus, agents maximize expected utility by choosing consumption c_j, asset holding k_{j+1} and time allocation l_j each period. Period resources come from labor market earnings $e_j = wh_jl_j$, which are the product of a wage rate w, human capital (skill) h_j and work time l_j, and from interest-paying assets $k_j(1+r)$ brought into the period. Next period's human capital h_{j+1} depends on current human capital h_j, learning time s_j, learning ability a and an idiosyncratic shock z_{j+1}. The model tax system $T_j(e_j, rk_j)$ depends on age, earnings and capital income.

Problem P1: $\max E[\sum_{j=1}^{J} \beta^{j-1} u(c_j, l_j + s_j)]$ subject to

$$c_j + k_{j+1} \leq e_j + k_j(1+r) - T_j(e_j, rk_j) \text{ and } k_{j+1} \geq 0, \forall j$$

$$e_j = wh_j l_j \text{ for } j < Retire \text{ and } e_j = 0 \text{ otherwise}$$

$$0 \leq l_j + s_j \leq 1,\ \ h_{j+1} = H(h_j, s_j, z_{j+1}, a) \text{ and } k_1 = 0$$

The consumption decision $c = (c_1, ..., c_J)$ is composed of functions $c_j(h_1, a, z^j)$ that map initial condition (h_1, a), age j and shock history z^j into a period choice. Labor time decisions $l = (l_1, ..., l_J)$, learning decisions $s = (s_1, ..., s_J)$, asset decisions $k = (k_1, ..., k_J)$ and human capital decisions $h = (h_1, ..., h_J)$ are specified similarly. These decisions, population shares μ_j and the distribution ψ of initial conditions define steady-state aggregates (K, L, C, T). The population shares by age satisfy $\mu_{j+1} = \mu_j/(1+n)$, where n is a population growth rate.

$$K = \sum_j \mu_j \int E[k_j|h_1, a]d\psi \textit{ and } L = \sum_j \mu_j \int E[h_j l_j|h_1, a]d\psi$$

$$C = \sum_j \mu_j \int E[c_j|h_1, a]d\psi \textit{ and } T = \sum_j \mu_j \int E[T_j(wh_j l_j, rk_j)|h_1, a]d\psi$$

Definition: A steady-state equilibrium consists of decisions (c, l, s, k, h), factor prices (w, r) and government spending G such that

1. Decisions: (c, l, s, k, h) solve Problem P1, given (w, r).
2. Prices: $w = F_2(K, L)$ and $r = F_1(K, L) - \delta$
3. Government Budget: $G = T$
4. Feasibility: $C + K(n + \delta) + G = F(K, L)$

The tax system in the Badel-Huggett model has three components: a progressive tax imposed on labor income, a social security system and a proportional tax on capital income. Figure 1 presents the progressive component of the income tax system. The progressive tax captures the increasing pattern of marginal tax rates on earnings in the US due to federal and state income taxes, sales taxes and medicare taxes. The top marginal tax rate is 42.5 percent, which equals the top rate calculated by Diamond and Saez (2011). This top rate applies to incomes within the top 1 percent of the income distribution.

Figure 2 presents the Laffer curve that results from increasing the top tax rate in Figure 1 leaving the tax rate schedule unchanged at lower income levels. The top of the Laffer curve in the human capital model occurs at a 52 percent top tax rate. Figure 2 also plots the Laffer curve in a closely-related model called the exogenous human capital model.[5] The exogenous

[5]Badel and Huggett (2014) describe this model in detail. This model is observationally equivalent to the benchmark human capital model under the benchmark tax system in that earnings, labor hours, consumption, assets and tax payments are indentical. However, when the tax system changes then the two models differ. Human capital decisions change in the human capital model in response to the tax reform via the learning time decision s_j. Human capital decisions in the exogenous human capital model evolve in the same way regardless of the size of the reform. All other decisions (consumption, labor hours, asset choice) are made optimally in the exogenous human capital model.

human capital model has a higher revenue-maximizing top tax rate (66 percent) and produces a substantially larger lump-sum transfer compared to the human capital model. This is because aggregate labor supply is less elastic with respect to a change in the top rate compared to the human capital model. Skills fall in response to an increase in marginal tax rates in the human capital model but not in the exogenous human capital model. The structure of the exogenous human capital model is similar to the structure of the models analyzed by Guner et al. (2014) and Kindermann and Krueger (2014) in that individual labor productivity is invariant with respect to changes in the tax system.

4.2 How to Map Model Allocations into Theorem 1

We map the equilibrium allocation of the Badel-Huggett model 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 a type $x = (h_1, a, j, z^j)$ is determined by the initial conditions (h_1, a), current age j and shock history z^j up to the current age. Under this formulation, an agent type completely determines all decisions.

We now specify the objects in Theorem 1 following the same three steps used in Exampes 1-3.

Step 1: $x = (h_1, a, j, z^j) \in X$
Step 2:

$$y_1(x, \tau) = w(\tau) h_j(h_1, a, z^j; \tau) l_j(h_1, a, z^j; \tau) \text{ for } j < Retire \text{ and } y_1(x, \tau) = 0 \text{ otherwise}$$

$$y_2(x, \tau) = b\bar{e} \text{ for } j \geq Retire \text{ and } y_2(x, \tau) = 0 \text{ otherwise.}$$

$$y_3(x, \tau) = r(\tau) k_j(h_1, a, z^j; \tau)$$

Step 3: $T(y_1, y_2, y_3; \tau) = T_1(y_1; \tau) + T_2(y_2, y_3)$

$$T_1(y_1; \tau) = T^{prog}(y_1; \tau) + \tau_{ss} \min\{y_1, emax\} \; and \; T_2(y_2, y_3) = T^{prog}(y_2) - y_2 + \tau_k y_3$$

In Step 2, (y_1, y_2, y_3) denote labor income, social security income and capital income respectively. The model tax and transfer system depends on these values. Specifically, Step 3 indicates that labor income y_1 is taxed progressively, where T^{prog} was graphed in Figure 1. The top marginal tax rate on earnings is due entirely to the progressive tax function top tax rate τ. This holds because the maximum taxable earnings level in the U.S. social security system, denoted $emax$, is well below the income level of the top tax bracket in Figure 1. Thus, the social security tax rate τ_{ss} does not apply, on the margin, to top earners. Capital income is taxed at a constant tax rate τ_k.

Based on Steps 1-3, we calculate the coefficients (a_1, a_2, a_3) and the elasticities $(\epsilon_1, \epsilon_2, \epsilon_3)$.[6] Table 1 present the results. The predicted top of the Laffer curve $\tau^* = .52$ equals the actual

[6] To calculate elasticity ϵ_1, set $X_1 = \{x \in X : y_1(x; \tau) = w(\tau) h_j(h_1, a, z^j; \tau) l_j(h_1, a, z^j; \tau) > \underline{y}\}$ where the top tax rate is set to the benchmark value $\tau = .425$. Then approximate the elasticity with a difference quotient by calculating steady-state equilibria at a different top tax rate.

top of the Laffer curve up to two-digit accuracy in the endogenous human capital model. In the exogenous human capital model the predicted top of the Laffer curve $\tau^* = .64$ differs from the actual top rate by two percentage points. This holds even though the initial steady state top tax rate, set at $\tau = .425$, is quite far from the revenue maximizing top rate. Based on the results in Table 1, the formula accurately predicts the top of the model Laffer curve.

Table 1: Revenue Maximizing Top Tax Rate Formula

Terms	Endogenous Human Capital Model	Exogenous Human Capital Model
$a_1 \times \epsilon_1$	$1.97 \times .396 = .780$	$1.97 \times .240 = .473$
$a_2 \times \epsilon_2$	$3.06 \times .019 = .059$	$3.06 \times .014 = .044$
$a_3 \times \epsilon_3$	$.043 \times .508 = .022$	$.043 \times .459 = .020$
$\tau^* = \frac{1-a_2\epsilon_2-a_3\epsilon_3}{1+a_1\epsilon_1}$	.52	.64
τ at peak of Laffer curve	.52	.66

Table 1 also provides a quantitative illustration of what the widely-used formula $\tau^* = 1/(1+a\epsilon)$ misses. The two additional terms $(a_2\epsilon_2, a_3\epsilon_3)$ in the numerator in the formula in Theorem 1 are both positive in the economies analyzed in Table 1. Thus, if one used the invalid but widely-used formula to predict the top of the Laffer curve and used the correct values for the remaining inputs, then one would predict that the top of the Laffer curve occurs at a higher top tax rate than the true top of either Laffer curve.

4.3 Comparing Model Elasticities to Estimated Elasticities

Saez, Slemrod and Giertz (SSG) (2012) review the literature that estimates the income elasticity with respect to the net-of-tax rate. Their review is very clear on several points: (1) the elasticities that this literature produces are the key input into the two sufficient statistic formulas (i.e. $\tau^* = 1/(1 + a\epsilon)$ and $\tau^* = (1 - g)/(1 - g + a\epsilon)$) discussed in the introduction, (2) the existing literature has primarily focused on short-run responses, (3) the long-run response to a tax reform is of most interest for policy making and (4) long-run elasticity estimates are plagued by extremely difficult issues so that there are no convincing estimates.[7]

This section applies existing methods from this literature to calculate, using model data, an earnings elasticity for top earners. This exercise is informative about the degree to which those methods approximate the long-run model elasticities that determine the top of the model Laffer curve (see Table 1) and that are of most interest for policy making. To the best of our knowledge, our work is the first in the literature to carry out such an analysis where the relevant long-run elasticity is known in advance.

[7]SSG(2012, p. 13) state *"The long-term response is of most interest for policy making although, as we discuss below, the long-term response is more difficult to identify empirically. The empirical literature has primarily focused on short-term (one year) and medium-term (up to five year) responses ... ."* SSG (2012, p. 43) state *"Estimates of the elasticity of taxable income in the long run (i.e., exceeding a few years) are plagued by extremely difficult issues of identification, so difficult that we believe that there are no convincing estimates of the long-run elasticity..."*.

Much of the literature applies the regression framework developed by Gruber and Saez (2002). They use the regression equation below, where ϵ is the empirical proxy for the elasticity parameter in the formula, z_{it} is income of individual i at time t, $\tau_t(z_{it})$ is the marginal tax rate, $f(z_{it}) = \log z_{it}$ is an income control and α_t are time dummy variables.

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

Indeed, Saez et al. (2012) estimate the elasticity of taxable income using U.S. data from 1991-97 and the regression equation above. Their estimation is based on income variation before and after a tax reform that occurred in 1993. They calculate that this reform increased the average marginal tax rate of the top 1 percent but did not significantly change the average marginal tax rate for the rest of the top 10 percent. Their panel consists of annual income histories for taxpayers in the top 10 percent of the income distribution in 1991. They estimate the elasticity for several specifications of instrumental variables, control variables and sample period.

In particular, Panel B in Table 2 of Saez et al. (2012) estimates the regression equation above using three different choices of instruments for log net-of-tax-rate changes and a two-stage-least-squares estimator. Instrument 1 is the indicator function taking the value 1 if individual i is in the top 1 percent in 1992 (i.e. $1_{\{i \in T_{1992}\}}$). Thus, T_{1992} denotes the set of individuals in the top 1 percent in 1992 which is the pre-reform year. Instrument 2 is $1_{\{i \in T_{1992} \text{ and } t=1992\}}$. Instrument 3 is $\log\left(\frac{1-\tau_{t+1}(z_{it})}{1-\tau_t(z_{it})}\right)$ and represents the log of the ratio of net-of-tax rates across years if income were not to change across years.

We use the same regression equation specification, instrument definitions, sample definitions and estimation techniques (i.e. two-stage-least squares) and apply them to a tax reform in the two models analyzed in Table 1. The benchmark tax system applies in period $t = 1$ and $t = 2$ and agents view this system as being permanent. Agents are surprised to learn that the model tax system is modified permanently at $t = 3$. They learn this at the start of period $t = 3$. Thus, the model periods $t = 1, ..., 7$ loosely correspond to the years $1991 - 1997$ for the US economy. In period 3, the tax system is modified by increasing the top tax rate to $\bar{\tau} = 0.52$, which corresponds to the tax rate at the top of the Laffer curve for the human capital model.

The response of agents and the aggregate transition path in the model is computed in partial equilibrium. This is for two reasons. First, the general equilibrium computation is much more costly in terms of computational resources. Second, the long-run response of factor prices to the reform is very small because aggregate capital and labor inputs decrease in roughly the same proportion.

The results of this exercise are presented in Table 2. Table 2(a) finds that the mean of the estimated elasticity across the six regressions is $\epsilon = 0.257$, whereas the model elasticity that is relevant for determining the top of the model Laffer curve is much larger and is $\epsilon_1 = 0.396$ as reported in Table 1. Thus, in the endogenous human capital model the estimated elasticity averages roughly two-thirds of the true, long-run model elasticity. Table 2(b) finds that the mean of the estimated elasticity across the six regressions is $\epsilon = 0.285$ when the true elasticity is $\epsilon_1 = 0.240$. The Appendix applies methods used in the literature to estimate a "longer-run" elasticity and finds that the results are even less favorable than the results in Table 2. The main conclusion that we draw is that existing methods do not accurately estimate the long-run model elasticity when the data is produced by the endogenous human capital model.

Table 2 - Elasticity Estimates

(a) Endogenous Human Capital Model

	(1)	(2)	(3)	(4)	(5)	(6)
Mean Elasticity ϵ	0.1536	0.2527	0.2688	0.2888	0.2806	0.2982
S.D.	(0.0254)	(0.0243)	(0.0251)	(0.0245)	(0.0228)	(0.0221)
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: $1_{\{i \in T_2 \text{ and } t=2\}}$			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	0.396	0.396	0.396	0.396	0.396	0.396

(b) Exogenous Human Capital Model

	(1)	(2)	(3)	(4)	(5)	(6)
Mean Elasticity ϵ	0.2128	0.2961	0.2925	0.3131	0.2892	0.3075
S.D.	(0.0238)	(0.0254)	(0.0239)	(0.0241)	(0.0233)	(0.0234)
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: $1_{\{i \in T_2 \text{ and } t=2\}}$			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	0.240	0.240	0.240	0.240	0.240	0.240

Note: (1) We draw 100 balanced panel data sets of 30,000 agents. Each data set mimics the structure of the data set used by Saez et al. (2012, Table 2). The agents in each balanced panel have labor income above the 90th percentile of earnings at $t = 1$. We follow these agents from periods $t = 1$ to $t = 7$. The tax reform occurs at $t = 3$. (2) We report means and standard deviations of the point estimates of ϵ across 100 randomly drawn balanced panels.

4.4 Why Do Existing Methods Work Poorly?

Why do state-of-the-art methods underestimate the theoretically-relevant elasticity in the endogenous human capital model but overestimate the elasticity in the exogenous human capital model? The proximate answer to this question comes from examining the transition path for aggregate capital K and labor L after the tax reform.

Figure 3 shows that aggregate labor input drops sharply in the human capital model in the first period of the reform and then gradually declines to the new steady state value after roughly 40 model periods - the length of the working lifetime in the model. Less than half of the overall change in labor input occurs in the initial reform period. Thus, regression methods that compare an earnings or income measure before and a year or so after the reform will at best pick up how hours movements impact earnings but not the impact of skill change on earnings. In the human capital model more than half of the change in the aggregate labor input is due to skill change across the two steady states.

Figure 3 shows that there is a steep fall in labor input in the initial period of the reform in the exogenous human capital model. Most of the change in the steady state aggregate labor input occurs in this model period. In fact, aggregate labor input falls by more in the first period of the reform than the total eventual decline in the steady-state aggregate labor input. Thus, it is not then surprising that the estimates in Table 2 for the exogenous human capital model often overestimate the true long-run model elasticity.

The movement in aggregate labor input in both models is largely determined by changes in behavior of agents with very high learning ability who have the greatest chance of becoming a top earner later in life. In the human capital model, agents with high learning ability decrease their time investments in human capital production over the lifetime in response to the reform. The fall in the labor input which is due to these lower investments takes a long time to be fully realized as the impact on skill accumulation is greatest in percentage terms at the end of an agent's working lifetime. This accounts for the substantially different dynamics across the two models.

5 Discussion

This paper makes two main points. First, the formula in Theorem 1 applies broadly to static models and to steady states of dynamic models and yet is stated in terms of only three elasticities. Thus, this sufficient-statistic formula should replace the widely-used formula in future work. Second, the formula from Theorem 1 works well in predicting the top of the Laffer curve in two quantitative models. Accurate prediction is a key roll for sufficient statistic formulas in practice. However, state-of-the-art methods for estimating elasticities, drawn from the literature on the elasticity of taxable income, do not accurately estimate one of the theoreticallly-relevant model elasticities even on an ideal data set drawn from a tax experiment performed within the model.

We think that these two points warrant two views:

First, economists should be skeptical of existing attempts in the literature that apply the sufficient statistic approach to predict the top of the Laffer curve. Taking the widely-used formula $\tau^* = 1/(1 + a\epsilon)$ at face value and then inputting values for (a, ϵ) into this formula, based on state-of-the-art techniques and an ideal data set, leads to inaccurate predictions. Using $a = 1.97$ from Table 1 and a range of values for ϵ estimated in Table 2, the predicted top of the model Laffer curve ranges from a low of $\tau^* = 0.63$ to a high of $\tau^* = 0.77$ when the actual top of the Laffer curve is 52 percent.

Second, this skepticism should not extend to the sufficient statistic approach in principle. The sufficient statistic formula in Theorem 1 applies to a wide range of relevant static and dynamic models. It also accurately predicts the top of the Laffer curve in the two dynamic models that we examine. The remaining problem is one of methods. New methods for estimating the high-level elasticities in the formula need to be developed with the structure of dynamic models in mind and applied convincingly to data. After this occurs celebrations can begin.

References

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

2. Auten, G, Carroll, R. and G. Gee (2008), The 2001 and 2003 Tax Rate Reductions: An Overview and Estimate of the Taxable Income Response, National Tax Journal, 61, 345-64.

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

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

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

6. 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.

7. Giertz, S. (2010), The Elasticity of Taxable Income During the 1990's: New Estimates and Sensitivity Analyses, Southern Economic Journal, 77, 406-33.

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

9. Goolsbee, A. (2000), What Happens When You Tax the Rich? Evidence from Executive Compensation, Journal of Political Economy, 108, 352- 78.

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

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

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

13. Kindermann, F. and D. Krueger (2014), The Redistributive Benefits of Progressive Labor and Capital Income Taxation, manuscript.

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

15. Mirrlees, J. and S. Adams, 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.

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

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

18. 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.

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

A Appendix

A.1 Longer-Run Elasticity

The elasticity of taxable income literature acknowledges that comparing measured income a year before and a year after a tax reform may be misleading because the relevant response for policy is the long-run response to a tax change. One approach to estimate a longer-run elasticity is to use the Gruber and Saez (2002) regression equation from section 4 of this paper but to measure income and net-of-tax rate changes over a longer horizon. For example, Auten, Carroll and Gee (2008) consider three-year and five-year differences, while Giertz (2010) considers one, three and six-year differences. We follow this approach below.

Another approach, used by Goolsbee (2000), argues that some of the measured response may be due to income shifting across years when there is advanced information of a pending reform and provides evidence for such income shifting. As advanced information is not a problem in the ideal data sets constructed in section 4, we will not pursue the modifications of the regression equation from section 4 suggested by the work of Goolsbee (2000) and others.

Table A1 - Longer-Run Elasticity Estimates

(a) Endogenous Human Capital Model

	(1)	(2)	(3)	(4)	(5)	(6)
Mean Elasticity ϵ	0.1998	0.1114	-0.0250	0.2405	0.1647	-0.1750
S.D.	(0.0223)	(0.0392)	(0.0606)	(0.0290)	(0.0492)	(0.0781)
Difference order (k)	1	3	6	1	3	6
Age Polynomial	no	no	no	yes	yes	yes
z_{it} Weights	no	no	no	yes	yes	yes
Long-run Model Elasticity ϵ_1	0.396	0.396	0.396	0.396	0.396	0.396

(b) Exogenous Human Capital Model

	(1)	(2)	(3)	(4)	(5)	(6)
Mean Elasticity ϵ	0.2053	0.1199	-0.0237	0.2279	0.1571	-0.1706
S.D.	(0.0228)	(0.0320)	(0.0603)	(0.0269)	(0.0430)	(0.0770)
Difference order (k)	1	3	6	1	3	6
Age Polynomial	no	no	no	yes	yes	yes
z_{it} Weights	no	no	no	yes	yes	yes
Long-run Model Elasticity ϵ_1	0.240	0.240	0.240	0.240	0.240	0.240

Notes : (1) All regressions include time effects and income control $f(z_{it}) = ln(z_{it})$. (2) We draw 100 balanced panel data sets of 30,000 agents. Each data set mimics the structure of the data set used by Saez et al. (2012, Table 2). The agents in each balanced panel have labor income above the 90th percentile of earnings at $t = 1$. We follow these agents from periods $t = 1$ to $t = 7$. The tax reform occurs at $t = 3$. (3) We report means and standard deviations of the point estimates of ϵ across 100 randomly drawn balanced panels.

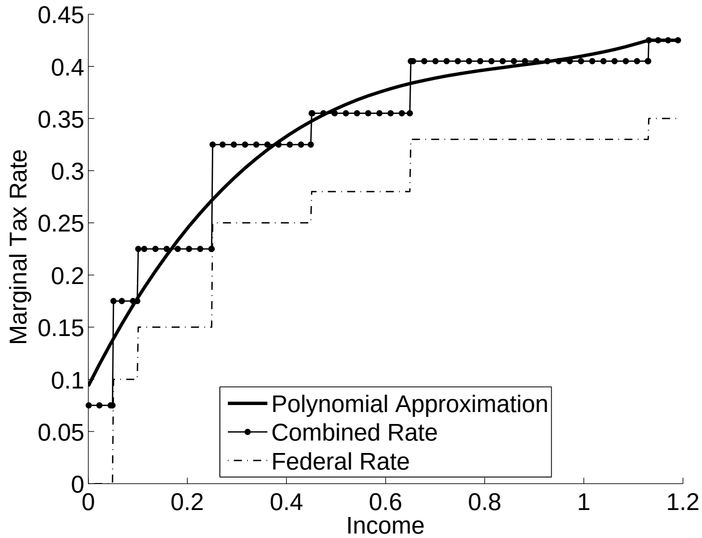
Table A1 estimates the regression equation below using $k = 1, 3$ and 6 year horizons. The regression and instrument specification follows those in Giertz (2010 Table 2, row A).

$$\log\left(\frac{z_{it+k}}{z_{it}}\right) = \epsilon \log\left(\frac{1-\tau_{t+k}(z_{it+k})}{1-\tau_t(z_{it})}\right) + \beta f(z_{it}) + \alpha_t + X_i'\gamma + \nu_{it+k} \quad (1)$$

The regression equation is a straightforward extension of the regression equation from section 4. The instrument specification is analogous to that used in column (6) of our Table 2. It consists of the counterfactual growth $\log\left(\frac{1-\tau_{t+k}(z_{it+k}^P)}{1-\tau_t(z_{it})}\right)$ of the marginal net of tax rate between t and $t+k$. Here z_{it+k}^P equals z_{it} times the growth factor of average earnings in the sample between t and $t+k$ so that earnings are assumed to be constant relative to trend. Columns (4)-(6) of Table A1 report estimations that include a fourth-order polynomial in age in the term X_i and use a weighted estimator where weights are equal to current earnings z_{it}. These practices are advocated by Giertz (2010).

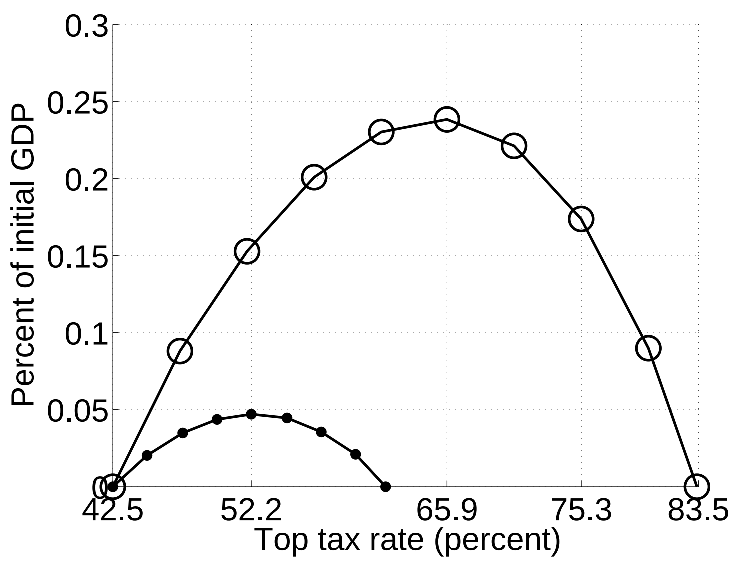
The specification in column (1) of Table A1 resembles the specification used in column (6) of Table 2 in that age controls are not used and earnings weighting is not used. However, the estimates differ slightly because the instrument used in column (1) of Table A1 is based on predicted future earnings z_{it+1}^P while those in column (6) of Table 2 only depend on current earnings z_{it}.

Figure 1. Progressive Component of Tax System, $T^{prog}(.)$



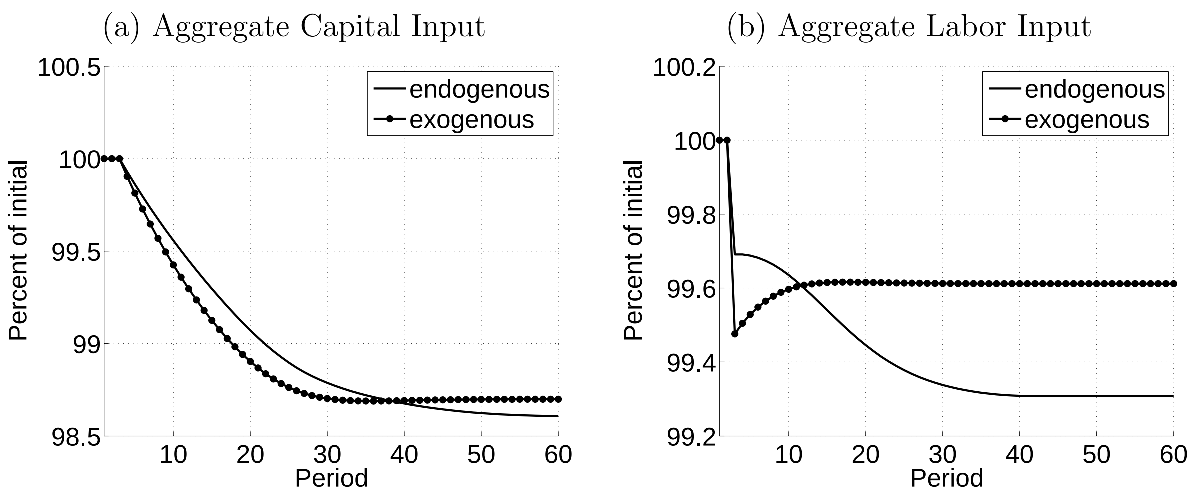
Note: The horizontal axis measures income in multiples of the 99th percentile of income. The income distribution data and the federal tax rate data is for 2010. The curve labeled Polynomial Approximation is the schedule of marginal tax rates in the model. See Badel and Huggett (2014) for full details behind the construction of Figure 1.

Figure 2. Laffer Curves: Endogenous and Exogenous Human Capital



Note: Large circles describe the Laffer curve for the exogenous human capital model. Dots describe the Laffer curve for the endogenous human capital model.

Figure 3. Transition Paths



Note: Aggregate capital and labor input are normalized to equal 100 in period 1. The tax reform occurs in model period 3.