This article highlights two approaches to tax policy for the top 1 percent of earners. On the one hand are dynamic general equilibrium models requiring complicated calibration and simulation algorithms and strong structural assumptions. On the other hand is the sufficient statistic approach, which attempts to parsimoniously reach the trinity of empirical, theoretical, and policy relevance. The author illustrates ongoing work highlighting explicit connections between these two approaches. (JEL D91, E21, H2, J24)


The fact that the distribution of income exhibits substantial concentration is common to many eras and countries and has been well known for many decades. More recently, the study of the earnings distribution over time and over the life cycle has gained attention in both the economics field and the media.

Regarding the life cycle, Deaton and Paxson (1994) used survey data from the United States, Taiwan, and Great Britain to document the fact that the dispersion of log earnings within a group of people born in the same year increases as they age. Following this tradition, Badel and Huggett (2014; BH hereafter) use data from the Social Security Administration (SSA), tabulated in Guvenen, Ozkan, and Song (2014), to document a similar pattern at the top of the earnings distribution: The ratio of the 99th percentile to the 50th percentile of earnings doubles over the working lifetime, and the mass of resources concentrated in the top 1 percent of the distribution increases substantially. Guvenen et al. (2015) also use SSA data and show that in the United States, individuals with higher lifetime earnings exhibit higher earnings growth over the life cycle. In particular, individuals in the top 1 percent of the lifetime earnings distribution experience an average earnings growth of approximately 1,500 percent over their working lifetime.

The analysis of income inequality over time has also attracted widespread attention. First, a literature based on survey data documents an increase in earnings dispersion in the United States. Heathcote, Perri, and Violante (2010) provide a compendium of survey-based measures...
of economic inequality for the United States showing, among other things, that the variance of log annual male earnings increased steadily, almost doubling in magnitude, between 1967 and 2005. Piketty and Saez (2003) use tax administration data and find a U-shaped pattern of income concentration in the top 1 percent over the twentieth century. The recent resurgence of income concentration within the top 1 percent in the United States documented by Piketty and Saez (2003) has received widespread attention.

Casual observation of the basic fact that income and earnings are substantially concentrated suggests that government intervention might improve social welfare. In particular, taxing some earnings away from (relatively satiated) very high earners and giving it away to (relatively deprived) low earners can improve certain measures of social welfare. The catch is that higher taxes can create a disincentive for earnings generation and ultimately reduce the size of the “income pie.”

As tax rates are increased, the size of the “income pie” shrinks. Thus, the additional government revenue raised by increasing tax rates shrinks as well. In fact, tax rates may reach a point where additional increments actually reduce government revenue. These reductions can lead to a bell-shaped relationship between tax rates and government revenue known as the Laffer curve (for an example, see Figure 1).

The Laffer curve is related to the trade-off between distribution (how the pie is shared) and efficiency (the size of the pie) in two ways: First, if the tax rate to be chosen affects very
high earners (whose well-being is not very sensitive to small changes in after-tax earnings),
then the revenue-maximizing tax rate approximately coincides with the socially optimal tax rate (i.e., in this case more revenue is better, regardless of the tax burden for top earners).

Second, policymakers should avoid setting a tax rate higher than the top of the Laffer curve. Such a tax rate would likely be wasteful, as an equal amount of revenue can be generated at a lower tax rate. Trabandt and Uhlig (2011) consider tax systems consisting of a single flat rate and calculate the top of the Laffer curve within representative-agent growth models calibrated for the United States and several European countries.

Partly because of the connection of the Laffer curve with considerations about distribution versus efficiency, economists have attempted to calculate the top of the Laffer curve. A separate reason for calculating the top of the Laffer curve is the need to gauge the maximum amount of debt sustainable by a particular economy. Whether an economy is fiscally sustainable is a basic practical question faced by treasury departments and investors around the world. Clearly, the maximum amount of debt that can be eventually repaid depends directly on the maximum attainable amount of revenue.

A widely known article by Diamond and Saez (2011; DS hereafter) uses a technique known as the sufficient statistic approach to predict the top of the Laffer curve. The Laffer curve considered is the one that would result from increasing the U.S. top marginal income tax rate—that is, the marginal rate applying to the top 1 percent of earners. DS provide provocative advice stating that the U.S. top marginal income tax rate (including federal, state, local, and other taxes) should be raised from 42.5 percent to 73 percent. A similar approach has been used to predict the top of the Laffer curve for Britain (see Brewer, Saez, and Shephard, 2010, p. 110).

The article by DS follows the key guidelines of the sufficient statistic approach: The quantitative advice therein comes from deriving a simple formula and obtaining the numerical inputs into the formula from existing empirical studies. See Chetty (2009) for a review of the literature on the sufficient statistic approach. (Hereafter the simple formula is referred to as the DS formula.) This formula gives the revenue-maximizing tax rate as a function of only two parameters. One parameter is the Pareto coefficient at the 99th percentile of the income distribution, which captures the magnitude of resources generated by earners in the top 1 percent. The other parameter measures the propensity of top earners to reduce their income in response to higher tax rates. This second parameter is drawn from a large empirical literature on the “elasticity of taxable income” (ETI).²

A direct approach to calculating the top of the Laffer curve with respect to the top marginal tax rate is to use quantitative dynamic models with heterogeneous agents; here we will call this the quantitative macro approach. See Heathcote, Storesletten, and Violante (2009) for a review of this literature. In this tradition, Guner, Lopez-Daneri, and Ventura (2014) consider a reform that increases the marginal tax rate that applies to the top 5 percent of the income distribution, while Kindermann and Krueger (2014) and BH consider reforms that increase the marginal tax rate that applies to the top 1 percent of earners. In the first two articles, the evolution of an agent’s labor productivity over the life cycle is driven by an exogenous probabilistic process. In contrast, BH consider a model where an agent’s labor productivity is determined by the accumulation of human capital, so it can be affected by tax reforms.³
The key features of the quantitative macro approach highlighted here are (i) dynamics through which agents take into account the future consequences of current decisions, (ii) household heterogeneity in initial endowments and shocks, and (iii) a life cycle such that agents live for finite spans of time. In this framework, a combination of incomplete insurance against idiosyncratic shocks and ex ante agent heterogeneity can deliver consumption, hours of work, and earnings distributions by age that resemble several aspects of those in U.S. data. These models are well suited to study issues of inequality and tax progressivity.

How are the sufficient statistic approach and the quantitative macro approach related? Badel and Huggett (2015) provide an answer consisting of two points:

(i) Under fairly general conditions, there exists a revenue maximization formula (analogous to the DS formula) that predicts the top of the Laffer curve in dynamic models as a function of three elasticity parameters and three coefficients. In contrast to the DS formula, the Badel–Huggett formula applies to models with several sources of government revenue besides earnings (e.g., taxes on capital or consumption) and to models featuring certain anticipatory responses (e.g., changes in human capital accumulation or savings that occur in anticipation of higher future tax rates). Badel and Huggett (2015) show how to map decisions in several dynamic and static models into the terms of the formula.

(ii) The popular reduced-form econometric methods used to estimate the key propensity parameter (ETI) that enters the DS formula do not adequately capture the magnitude of the underlying elasticity when applied in some dynamic settings—for example, when human capital takes a long time to readjust in response to a tax reform.

The connection suggested in points (i) and (ii) is that the quantitative macro approach and the sufficient statistic approach are conceptually compatible, but that the existing empirical methods used by the sufficient statistic approach so far are not reliable for capturing the elasticity that holds in dynamic models.

In the remainder of this article, we illustrate points (i) and (ii) in more depth. Readers are referred to Badel and Huggett (2015) for the theorem deriving the Badel–Huggett revenue-maximization formula and several formal examples illustrating the application of the formula.

**AN APPLICATION OF THE SUFFICIENT STATISTIC APPROACH**

We start by posing the formula used to form the DS quantitative advice. Consider a static model in which a fixed marginal rate \( \tau \) applies to earnings beyond a threshold, \( e_0 \). Then the per person tax revenue from earnings above a given earnings level, say \( e_0 \), is given by \( [\bar{e} - e_0] \tau \), where \( \bar{e} = \mathbb{E}[e|e > e_0] \) denotes mean earnings among agents with earnings above the threshold. The revenue-maximization problem is therefore \( \max \{ [\bar{e} - e_0] \tau \} \). The solution to this problem implies that, at a maximum, the top marginal tax rate satisfies the formula below:

\[
\tau^* = \frac{1}{1 + a \varepsilon},
\]
where \( a \equiv \frac{\bar{e}}{(\bar{e} - e_0)} \) is the Pareto coefficient of the earnings distribution at percentile \( e_0 \), an inverse measure of the thickness of the right tail of the earnings distribution above \( e_0 \), and \( \varepsilon \) is the elasticity of \( \bar{e} \) with respect to the net-of-tax rate, defined as \( (1 - \tau) \). That is,

\[
\varepsilon \equiv \frac{d\bar{e}}{d(1-\tau)} \frac{1-\tau}{\bar{e}}.
\]

The DS quantitative advice says that the revenue-maximizing top tax rate for the United States is \( \frac{1}{1+1.5 \times 0.25} = 73 \) percent. This follows from plugging into the formula the value \( a = 1.5 \), which is the Pareto coefficient at the 99th percentile of the income distribution in the United States in 2007, and the value \( \varepsilon = 0.25 \), a midrange value from the literature estimating the ETI with respect to the net-of-tax rate.

**The Sufficient Statistic Approach within a Dynamic Model**

The revenue-maximization formula used by DS is not generally valid in dynamic models. Intuitively, there are two reasons behind this statement: (i) The DS formula does not account for the reaction of agents below the top 1 percent of the income distribution to the top tax rate reform. Agents who anticipate entering the top 1 percent in the future may take anticipatory behavioral responses. (ii) The DS formula does not account for the potential response of revenue from non-income sources, such as capital income, to the top tax rate reform. For example, agents may hold less financial wealth in reaction to a higher top marginal tax rate for earnings. This would decrease their capital income and, therefore, their capital income tax contribution. Trabandt and Uhlig (2011) emphasize, for example, the degree to which labor tax cuts can be self-financing through higher revenues from other sources. Finally, note that either reason (i) or (ii) can matter either directly or through general equilibrium effects such as factor price adjustments.

Badel and Huggett (2015) provide a revenue-maximization formula that applies to steady-state equilibria of dynamic models. They consider economies in which individual decisions specify \( n \) types of flows or stocks that are subject to taxation, denoted by \( y_1(x, \tau), \ldots, y_n(x, \tau) \). These decisions depend on an agent’s vector of characteristics, \( x \), and the top marginal tax rate \( \tau \). The net tax contribution of an individual with type \( x \) is \( T(y_1(x, \tau), \ldots, y_n(x, \tau); \tau) \). The distribution of characteristics \( x \) across the population is exogenous.

The formula in Badel and Huggett (2015) relies on differentiability assumptions together with the following assumption on the structure of the tax system (Assumption A2’ from Badel and Huggett, 2015): \( T(.; \tau) \) is separable in that \( T(y_1, \ldots, y_n; \tau) = T_1(y_1; \tau) + T_2(y_1, \ldots, y_n) \) and moreover, there is \( y \geq 0 \) such that

(i) \( T_1(y_1; \tau) - T_1(y; \tau) = \tau [y_1 - y] \)

(ii) \( T_1(y_1; \tau) = T_1(y_1; \tau') \) for all \( y_1 < y \) and \( \tau' \in (0,1) \).

Assumption A2’ in Badel and Huggett (2015) begins by stating that the tax system is separable in the component that specifies the value of taxes from the first component of the vector of
decisions, \( y_1 \). Part (i) then states that beyond a cutoff level \( y \) of \( y_1 \), the tax system is characterized by a flat marginal rate \( \tau \) on \( y_1 \). Finally, part (ii) states that an agent’s taxes from source \( y_1 \) are invariant to the top tax rate \( \tau \) if \( y_1 \) is below the cutoff \( y \).

Theorem 1 in Badel and Huggett (2015) states that the revenue-maximizing top tax rate formula is

\[
\tau = \frac{1 - a_x e_2 - a_x e_3}{1 + a_x e_3}.
\]

Detailed definitions for each of the statistics in the formula are provided in Badel and Huggett (2015). This formula is valid in both dynamic and static models. Compared with the DS formula, it includes two additional terms in the numerator. These terms take the value zero when the formula is applied within a version of the Mirrlees model (see Badel and Huggett, 2015, Example 1). Therefore, equivalence with the DS formula holds in that case. The numerator terms capture two types of responses to changes in the top tax rate \( \tau \). The term \( a_x e_2 \) captures the \( y_1 \) response of agents with \( y_1 < y \). The term \( a_x e_3 \) captures the tax implications of the \( y_1 \) decisions (i.e., all decisions except the \( y_1 \) decision) of agents inside and outside of the top 1 percent.

**DYNAMIC MODELS WITH HETEROGENEOUS AGENTS**

Badel and Huggett (2014) consider the top tax reform recommended by DS within a human capital model. In the human capital model, agents differ in their initial human capital and learning ability. Agents start their working life at age 23 and retire at age 62. During their working lifetime, agents choose their savings and how to allocate their time into three activities: learning, working, and leisure. Learning ability, learning time, existing human capital, and idiosyncratic shocks determine an agent’s accumulation of human capital from one year to the next. During their working lifetime, agents pay individual and payroll taxes on their earnings and also pay taxes on the financial returns from saving. During retirement, agents consume out of their savings and their social security income. A steady-state equilibrium of the model features a constant interest rate determined at each point in time by the total savings and labor supply of all the generations alive.

The model is calibrated to match several features of the U.S. earnings distribution by age. The calibrated model can replicate the increase in earnings dispersion over the life cycle observed within the top 1 percent of earners and that observed between the 50th and the 99th percentile of the U.S. earnings distribution.\(^4\) Also, the model is calibrated so that an instrumental variables (IV) regression applied to data from the model matches the labor supply regression coefficient calculated by MaCurdy (1981).\(^5\) This coefficient is conventionally viewed as an important measure of the sensitivity of the labor supply to changes in wages. The model tax system captures key features of the U.S. tax system such as the progressive nature of the individual income tax schedule and the flat rate in the top bracket.

The Laffer curve that holds in the model is calculated. Each point of the Laffer curve corresponds to a steady-state equilibrium of the model with each equilibrium featuring a different
The additional government revenue obtained under top tax rates higher than the baseline (42.5 percent, as calculated by DS) is rebated in equal lump-sum transfers. The Laffer curve plots the magnitude of these transfers as a share of baseline gross domestic product. Figure 1 displays the Laffer curve from the baseline tax reform and calibration in BH.

Two things stand out from the baseline Laffer curve in BH: (i) The revenue-maximizing tax rate is only 52.2 percent, substantially below the DS quantitative guidance of 73 percent, and (ii) the maximum revenue increase from the reform is 0.05 percent of initial gross domestic product.

To illustrate the tax reform, Figure 2A plots the marginal tax schedule. The note under Figure 2A explains how the model tax schedule is based on the U.S. tax code in 2010. To trace the Laffer curve in Figure 1, the top marginal tax rate is increased from 42.5 to 61.9 percent, leaving the tax rate unchanged below the top bracket.

Figure 2B shows how increasing the top marginal tax rate affects the average life cycle profiles of earnings for agents with high learning ability. Good learners have a strong incentive to invest time learning early in the life cycle and reap the benefits by working more later on. Growing work time and growing human capital due to learning imply that earnings grow strongly over the life cycle for high-ability agents. Such strong growth implies that many high-ability agents start their working life below the top 1 percent but reach the top 1 percent later.
in life. For these agents, a policy that increases the top tax rate lowers the return to human
capital later in life but does not affect the opportunity cost of learning early in life. A fall in the
return to learning time without a fall in its cost leads to less learning time and the clockwise
flattening of earnings profiles shown in Figure 2B. An important consequence of the human
capital mechanism is that the full effect of the top tax reform is complete only after several
years because it depends on human capital accumulation choices made by agents over their
working lifetime.

**Applying the Sufficient Statistic Approach within a Dynamic Model**

The sufficient statistic approach to taxing top earners consists of two parts. The first is a
revenue-maximizing tax rate formula. The second is a set of estimated values for the parameters or statistics in the formula.

A key parameter in the DS formula is $e$, the elasticity of top earnings with respect to the
net-of-tax rate. DS consider a range of estimates and a preferred value of this parameter from
the extensive literature that estimates an elasticity of taxable earnings with respect to the net-
of-tax rate. This literature is reviewed by Saez, Slemrod, and Giertz (2012).

In this section, we summarize two different ways to apply the sufficient statistic approach
within the human capital model. The experiments are as follows:

(i) Use the Badel and Huggett (2015) formula to predict the top of the model Laffer curve
under almost ideal conditions. The three coefficients $(a_1, a_2, a_3)$ and the three elastic-

---

**Figure 2B**

*Average Earnings by Age Among Agents with High Learning Ability*

![Graph showing average earnings by age among agents with high learning ability.](image)
(i) Use both parts of the DS approach to predict the top of the Laffer curve using data
(i) from the baseline steady state of the human capital model (to calculate the Pareto
coefficient $a$) and (ii) from a simulated tax reform conducted within the human capital
model (to calculate an estimate of the elasticity $e_i$). We use artificial panel data from a
simulated tax reform to estimate the elasticity of top earnings with respect to the net-
of-tax rate using reduced-form econometric methods from Saez, Slemrod, and Giertz
(2012). The model-generated dataset emulates the key features of the empirical dataset
used by Saez, Slemrod, and Giertz (2012) to provide estimates of $\epsilon_i$. Their dataset
exploits the top tax rate increase embedded in the 1993 Omnibus Budget Reconciliation
Act. The instrumental variables methods from Saez, Slemrod, and Giertz (2012)
are followed closely. The results and methods are summarized in Table 5 in BH. The
key result is that for the variety of empirical specifications considered by Saez, Slemrod,
and Giertz, the regression coefficient obtained underestimates the true $\epsilon_i$, which is $\epsilon_i = 0.396$. Quantitatively, this implies that the range of top tax rate recommendations
produced by the DS approach within the human capital model is above the rate that
actually maximizes revenue in the model.

### Table 1
Predicting the Top of the Laffer Curve

<table>
<thead>
<tr>
<th>Parameter values</th>
<th>Implied revenue-maximizing top tax rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1 \times \epsilon_1$</td>
<td>$1.97 \times 0.396 = 0.78$</td>
</tr>
<tr>
<td>$a_2 \times \epsilon_2$</td>
<td>$3.07 \times 0.019 = 0.059$</td>
</tr>
<tr>
<td>$a_3 \times \epsilon_3$</td>
<td>$0.043 \times 0.508 = 0.22$</td>
</tr>
<tr>
<td>$\tau^* = \frac{1 - a_2 \epsilon_2 - a_3 \epsilon_3}{1 + a_1 \epsilon_1}$</td>
<td>0.52</td>
</tr>
<tr>
<td>$\tau$ at peak of Laffer curve</td>
<td>0.52</td>
</tr>
</tbody>
</table>

**SOURCE:** Badel and Huggett (2015).

CONCLUSION

In this article, we review ongoing work connecting the application of the sufficient statistic
approach and the quantitative macro approach to finding the revenue-maximizing marginal
tax rate for top earners. Two key points are highlighted: First, Badel and Huggett (2015) derive
a revenue-maximizing top tax rate formula that is valid within steady states of dynamic models.
Moreover, in a human capital model the Badel-Huggett formula predicts the top of the Laffer curve well. Second, when the inputs of the formula are obtained using existing reduced-form empirical methods for the ETI applied to model-generated data, the top of the Laffer curve can be overestimated. A likely reason for this result is that the ETI methods do not capture the long-run effects of human capital accumulation.

NOTES

1 To illustrate this, consider the narration in Mandelbrot and Hudson (2004, pp. 153-4): Italian economist Vilfredo Pareto “gathered reams of data on wealth and income through different centuries, through different countries: the tax records of Basel, Switzerland, from 1454 and from Augsburg, Germany, in 1471, 1498, and 1512; contemporaneous rental income from Paris; personal income from Britain, Prussia, Saxony, Ireland, Italy, Peru. What he found—or thought he found—was striking. When he plotted the data on graph paper, with income on one axis and number of people with that income on the other, he saw the same picture nearly everywhere in every era. Society was not a ‘social pyramid’ with the proportion of rich to poor sloping gently from one class to the next. Instead, it was more of a ‘social arrow’—very fat on the bottom where the mass of men live, and very thin at the top where sit the wealthy elite.”

2 The parameter is known in the literature as the ETI with respect to the net-of-tax rate.

3 A look at these three articles reveals widely different conclusions about the shape of the Laffer curve. The reasons for these differences can be clearly traced back to model structure, calibration strategy, and details of the specific reforms considered in each article. This is a well-known advantage of the quantitative macro approach. However, we leave a comparative analysis of these studies for future work.

4 One of the novel features in BH is the introduction of a bivariate distribution based on the Pareto lognormal for the joint distribution of initial human capital and learning ability. The thicker tails of this distribution, compared with those of the bivariate lognormal used in Huggett, Ventura, and Yaron (2010), enable the model to match some aspects of the top tail of the U.S. earnings distribution by age.

5 MacCurdy (1981) used Panel Study of Income Dynamics data for males 25 to 55 years of age. The regression equation in MacCurdy (1981) is \( \Delta \log(hours_i) = \alpha_i + \alpha_r \Delta \log(wages_i) + \nu_i \), where \( i \) denotes an individual and \( \nu \) is a random disturbance.

6 The conditions are “almost” ideal because, even though the inputs of the formula are measured directly within the model, they are measured at the benchmark equilibrium (away from the optimum), while the formula is derived using a local argument at the optimum.

REFERENCES


Guvenen, Fatih; Karahan, Fatih; Ozkan, Serdar and Song, Jae. “What Do Data on Millions of U.S. Workers Reveal about Life-Cycle Earnings Risk?” Federal Reserve Bank of New York Staff Report No. 710, February 2015; 
http://www.newyorkfed.org/research/staff_reports/sr710.pdf.


http://pubs.aeaweb.org/doi/pdfplus/10.1257/jel.50.1.3.
