DISCUSSION PAPER SERIES

No. 8675

OPTIMAL TAXATION OF TOP LABOR INCOMES: A TALE OF THREE ELASTICITIES

Thomas Piketty, Emmanuel Saez
and Stefanie Stantcheva

PUBLIC POLICY

Centre for Economic Policy Research

www.cepr.org

OPTIMAL TAXATION OF TOP LABOR INCOMES: A TALE OF THREE ELASTICITIES

Thomas Piketty, Paris School of Economics and CEPR
Emmanuel Saez, University of California, Berkeley, NBER and CEPR
Stefanie Stantcheva, MIT

This Discussion Paper is issued under the auspices of the Centre's research programme in **PUBLIC POLICY**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

CEPR Discussion Paper No. 8675

November 2011

ABSTRACT

Optimal Taxation of Top Labor Incomes:
A Tale of Three Elasticities*

This paper analyzes the problem of optimal taxation of top labour incomes. We develop a model where top incomes respond to marginal tax rates through three channels: (1) the standard supply-side channel through reduced economic activity, (2) the tax avoidance channel, (3) the compensation bargaining channel through efforts in influencing own pay setting. We derive the optimal top tax rate formula as a function of the three elasticities corresponding to those three channels of responses. The first elasticity (supply side) is the sole real factor limiting optimal top tax rates. The optimal tax system should be designed to minimize the second elasticity (avoidance) through tax enforcement and tax neutrality across income forms, in which case the second elasticity becomes irrelevant. The optimal top tax rate increases with the third elasticity (bargaining) as bargaining efforts are zero-sum in aggregate. We then analyze top income and top tax rate data in 18 OECD countries. There is a strong correlation between cuts in top tax rates and increases in top 1% income shares since 1975, implying that the overall elasticity is large. But top income share increases have not translated into higher economic growth, consistent with the zero-sum bargaining model. This suggests that the first elasticity is modest in size and that the overall effect comes mostly from the third elasticity. Consequently, socially optimal top tax rates might possibly be much higher than what is commonly assumed.

Thomas Piketty
Paris School of Economics
48 Boulevard Jourdan
75014 Paris
FRANCE

Emmanuel Saez
University of California, Berkeley
549 Evans Hall, 3880
Berkeley
CA 94720
USA

Email: thomas.piketty@ens.fr

Email: saez@econ.berkeley.edu

Stefanie Stantcheva
Department of Economics
MIT
50 Memorial Drive
Building E52
Cambridge, MA 02142-1347
USA
Email: stefanie@mit.edu

1 Introduction

The share of total pre-tax income accruing to upper income groups has increased sharply in the United States. The top percentile income share has more than doubled from less than 10% in the 1970s to over 20% in recent years (Piketty and Saez, 2003). This trend toward income concentration has taken place in a number of other countries, especially English speaking countries, but is much more modest in continental Europe or Japan (Atkinson, Piketty, Saez, 2011 and Alvaredo et al. 2011). At the same time, top tax rates on upper income earners have declined significantly in many OECD countries, again particularly in English speaking countries. For example, the US top marginal federal individual tax rate stood at an astonishingly high 91% in the 1950s-1960s but is only 35% today (see Figure 1).

While there have been many discussions both in the academic literature and the public debate about the causes of the surge in top incomes, there is not a fully compelling explanation. Most explanations can be classified into market driven changes vs. institution driven changes. The market driven stories posit that technological progress has been skilled-biased and has favored top earners relative to average earners (e.g., Gabaix and Landier (2008) for CEOs as well as Winner-Take-All theories for superstars, following Rosen (1981)). The key problem with those pure market explanations is that they cannot account for the fact that top income shares have only increased modestly in advanced countries such as Japan or Germany or France which are also subject to the same technological forces. The institution driven stories posit that changes in institutions, defined to include labor and financial market regulations, Union policies, tax policy, and also more broadly social norms regarding pay disparity and in particular tolerance for executive pay, have played a key role in the evolution of inequality. Simply put, under that view, the Reagan and Thatcher revolutions ushered new eras in the United States and United Kingdom that favored the rich and significantly increased their bargaining power while other countries were less affected (see Bartels 2008 and Hacker and Pierson 2010 for US analyzes along those lines). The main difficulty is that "institutions" are multi-dimensional and it is difficult to estimate compellingly the contribution of each specific factor.

Related to this issue, there is also a wide empirical literature in public economics analyzing the effects of tax rates on pre-tax incomes (see Saez, Slemrod, and Giertz, 2011 for a recent survey) that reaches two broad conclusions. First, there is compelling evidence that upper incomes respond to tax rates whenever the tax code offers opportunities for tax avoidance.

Such responses can sometime be quite large, especially in the short-run.[1] Second and related, when the tax base is broad and does not offer avoidance opportunities, the estimated elasticities are never large at least in the short or medium-run. In other words, all the compellingly large elasticity estimates obtained in the literature are always due to tax avoidance or short-run re-timing. Hence, no study to date has been able to show convincing evidence in the short or medium-run of large actual economic real economic activity responses of upper earners to tax rates. However, it is difficult to provide compelling estimates of long-run elasticities. As we shall see, international evidence shows a strong correlation between top tax rate cuts and increases in top income shares in OECD countries since 1975. Interestingly, the link between top rate cuts and top income share increases is strong in English speaking countries but much smaller in other countries such as Japan or Sweden which also experienced large top tax rate cuts.

There are three narratives of the link between top tax rates and upper incomes. First, after noting that top US incomes surged following the large top marginal tax rate cuts of the 1980s, Lindsey (1987) and Feldstein (1995) proposed a standard supply-side story whereby lower tax rates stimulates economic activity among top income earners (work, entrepreneurship, savings, etc.). Second, it has been pointed out–originally by Slemrod, 1996–that many of those dramatic responses were actually primarily due to tax avoidance rather than real economic behavior. Although this argument started as a left-wing critique of the supply-side success story, it has more recently been used as a right-wing argument to deny that any real increase in income concentration actually took place (Reynolds, 2007). Under this scenario, the real US top income shares were as high in the 1970s as they are today but a smaller fraction of top incomes was reported on tax returns in the 1970s than today. A third narrative contends that high top tax rates were part of the institutional set-up putting a brake on top compensation through bargaining or rent extraction effects. When top marginal tax rates are very high, the net reward to a highly paid executive for bargaining for more compensation is modest. Under this scenario, some countries such as Japan still have societal or institutional brakes on large top compensation in spite of relatively low top tax rates. In contrast, in countries such as the United

---

[1]For example, in the United States, realized capital gains surged in 1986 in anticipation of the increase in the capital gains tax rate after the Tax Reform Act of 1986 (Auerbach, 1988). Similarly, exercises of stock options surged in 1992 before the 1993 top rate increase took place (Goolsbee, 2000). The Tax Reform Act of 1986 also led to a shift from corporate to individual income as it became more advantageous to be organized as a business taxed solely at the individual level rather than as a corporation taxed first at the corporate level (Slemrod, 1996; Gordon and Slemrod, 2000).

States or the United Kingdom, such (non-tax) brakes are not present–and possibly disappeared at the same time as the high top tax rates during the Reagan or Thatcher revolutions.

The first goal of this paper is to present a very simple model of optimal top labor income taxation that can capture all three avenues of response, the standard supply side response, the tax avoidance response, and the compensation bargaining response to assess how each narrative translates into tax policy implications.[2] We therefore derive the optimal top tax rate formula as a function of the three elasticities corresponding to those three channels of responses. The first elasticity (supply side) is the sole real factor limiting optimal top tax rates. A large tax avoidance elasticity is a symptom of a poorly design tax system. A very high top tax rate within such a system offering many tax avoidance opportunities is counter-productive. Hence, the optimal tax system should be designed to minimize tax avoidance opportunities through a combination of tax enforcement, base broadening, and tax neutrality across income forms. In that case, the second elasticity (avoidance) becomes irrelevant. The optimal top tax rate *increases* with the third elasticity (bargaining) as bargaining efforts are wasteful and zero-sum in aggregate. If a substantial fraction of the behavioral response of top earners comes from bargaining effects and top earners are not paid less than their economic product, then the optimal top tax rate is much higher than the conventional formula and actually goes to 100% if the real supply-side elasticity is very small.[3]

In our view, this is the right model to account for the very high, quasi-confiscatory top marginal rates–80%-90% or more–applied in the United States and the United Kingdom between the 1940s and the 1970s (see Figure 1). That is, policy makers and public opinions at that time probably considered–rightly or wrongly–that at the very top of the income ladder, pay increases reflect mostly greed and socially wasteful activities rather than productive work effort. Whether they were right or wrong is certainly a complicated empirical issue. But in order to address this issue in a meaningful way, we first need a proper conceptual framework within which the various conflicting claims can be rationalized. To our knowledge this is the first paper offering

---

[2]This paper focuses on the optimal taxation of top labor incomes. The optimal taxation of top capital incomes and top wealth holdings (via one-off inheritance taxes or annual property and wealth taxes) raises other issues, which we address in Piketty and Saez (2011). Importantly, if there is a fuzzy frontier between different kinds of flow incomes then it is inefficient to have very different tax rates on labor vs. capital income, a point we will later address in this paper.

[3]The optimal top tax rate is moderate if the supply elasticity is fairly large and top earners are underpaid relative to their product, a situation that is theoretically possible in our model and might exist in countries with very low income concentration.

such a framework.

Needless to say, we would also very much like to be able to provide convincing empirical estimates of the three elasticities e_1, e_2 and e_3. So the second goal of the paper is to use international time series evidence on the evolution of top marginal tax rates, top income shares, and economic growth among about 18 OECD countries since the 1970s to examine broad correlations and determine which scenarios fit best with the data. We deliberately use a macro-approach because the micro-approach (reviewed extensively in Saez, Slemrod, Giertz, 2011) cannot measure compellingly long-term effects that are needed to assess scenarios and draw policy conclusions. We are well aware that the macro-approach requires unduly strong identification assumptions and hence should be seen as an illustrative first step rather than a definitive proof.

We obtain three main results. First, we find a very clear correlation between the drop in top marginal tax rates and the surge in top income shares since 1975. This suggests that the long-run total elasticity of top incomes with respect to the net-of-tax rate is large. That is, $e = e_1 + e_2 + e_3$ appears to be large, around 0.5. Interestingly, there is significant heterogeneity across countries suggesting that the total elasticity varies significantly across countries.

Second, examination of the US case suggests that the tax avoidance response cannot account for a significant fraction of the long-run surge in top incomes because top income shares based on a broad definition of income (that includes realized capital gains and hence a significant part of avoidance channels) has increased virtually as much as top income shares based on a narrower definition of income subject to the progressive tax schedule.[4] That is, the elasticity e_2 appears to be small (say, $e_2 < 0.1$).

Third, we find no evidence of a correlation between growth in real GDP per capita and the drop in the top marginal tax rate in the period 1975 to the present. This evidence, admittedly only suggestive, is consistent with the bargaining model whereby gains at the top come at the expense of lower income earners. This suggests that the first elasticity is modest in size and that the overall effect comes mostly from the third elasticity. Consequently, socially optimal top tax rates might possibly be much higher than what is commonly assumed.

In our preferred estimates, we find an overall elasticity $e = 0.5$, which can be decomposed into $e_1 = 0.2$ (at most), $e_2 = 0$ and $e_3 = 0.3$ (at least). This corresponds to a socially optimal

---

[4]The avoidance scenario cannot explain well either why top income shares have remained relatively low in countries such as Japan where top tax rates have also decreased dramatically and where incentives for tax avoidance are not higher than in the United States.

top tax rate $\tau^* = 83\%$ - as compared to $\tau^* = 57\%$ in the standard supply-side case with $e = e_1 = 0.5$ and $e_2 = e_3 = 0$. This illustrates the critical importance of this decomposition into three elasticities. We hope this will contribute to stimulating further research on obtaining better empirical estimates of e_1, e_2 and e_3 in the future.

The remainder of the paper is organized as follows. Section 2 presents briefly the standard model with real supply-side economic responses. Section 3 introduces tax avoidance and income shifting responses. Section 4 introduces compensation bargaining responses. Section 5 presents an empirical application using international evidence since the 1970s. Section 6 briefly concludes.

2   Standard Model: Real Economic Responses

In the paper, we denote by z taxable earnings and by $T(z)$ the nonlinear tax schedule. We assume a constant marginal tax rate τ in the top bracket above a given income threshold $\bar{z}$.[5] We assume without loss of generality that the number of taxpayers in the top bracket has measure one at the optimum.[6] We refer to this group as top bracket taxpayers. We focus on the determination of the optimal top tax rate τ, taking $\bar{z}$ as given.

We start with the standard Mirrlees (1971) model. We will always assume away income effects for simplicity and tractability, and consider utility functions of the form:

$$u_i(c, z) = c - h_i(z),$$

where z is pre-tax earnings, $c = z - T(z)$ is disposable income, and $h_i(z)$ denotes the labor supply cost of earning z which is increasing and convex in z.[7] Optimal effort choice is given by the first order condition $h_i'(z) = 1 - \tau$ where τ is the marginal tax rate so that individual earnings $z_i(1 - \tau)$ are solely a function of the net-of-tax rate $1 - \tau$. Aggregating over all top bracket taxpayers, we denote by $z(1 - \tau)$ the *average* income reported by top bracket taxpayers, as a function of the net-of-tax rate. The aggregate elasticity of income in the top bracket with

---

[5]For example, in the case of the Federal US individual income tax for year 2011, ordinary taxable income above $\bar{z} = \$379,150$ is taxed at the top marginal tax rate of $\tau = 0.35$. When combining all taxes including state taxes, the total US marginal tax rate is 42.5% in 2011 (see Diamond and Saez, 2011).

[6]Naturally, because of behavioral responses, the number of taxpayers in the top bracket will differ from one when the tax rate changes.

[7]In the standard formulation where individuals differ only through their wage rate w_i and $z = w_i l$ with l labor supply, cost of effort takes the form $h_i(z) = h(z/w_i) = h(l)$. Our formulation nests this standard case and is useful for subsequent extensions.

respect to the net-of-tax rate is therefore defined as

$$e_1 = \frac{1-\tau}{z} \frac{dz}{d(1-\tau)}. \tag{1}$$

This is the standard *first elasticity* that reflects real economic responses to the net-of-tax rate, which can be labeled as labor supply effects, broadly defined (more hours of work, more intense effort per hour worked, occupational choices, etc.)

The government maximizes a standard social welfare function of the form:

$$W = \int G(u_i) d\nu(i),$$

subject to an aggregate budget constraint:

$$\int T(z_i) d\nu(i) \geq T_0$$

where $G(.)$ is an increasing and concave function, and $d\nu(i)$ is the density mass of people of individuals of type i, and $T_0 \geq 0$ is an exogenous tax revenue requirement. Denoting by p the multiplier of the government budget constraint, we define the social marginal welfare weight on individual i as $g_i = G'(u_i)/p$. Because there are no income effects, the average of the social welfare weights, g_i in the population, denoted $E(g)$ is equal to one.[8]

In this paper, we always assume that the average social marginal welfare weight among top bracket income earners is zero. If the social welfare function $G(.)$ has curvature so that $G'(u) \to 0$ when $u \to \infty$, this will be the case when $\bar{z} \to \infty$ and will hence approximately be true for large $\bar{z}$. Considering a zero marginal welfare weight allows us to obtain an upper bound on the optimal top tax rate. We mention briefly how formulas ought to be modified if we instead put a positive social welfare weight g on the marginal consumption of top bracket earners (relative to average).

Diamond (1998) and Saez (2001) show that the optimal top tax rate is given by:

$$\tau^* = \frac{1}{1 + a \cdot e_1}, \tag{2}$$

where $a = z/(z - \bar{z}) = (z/\bar{z})/(z/\bar{z} - 1) > 1$ is the Pareto parameter of the top tail of the

---

[8]This can be seen as follows. If the government increases taxes by \$1 on everybody, the impact on social welfare (measured in units of public funds) is by definition equal to $\int g_i d\nu(i)$. Such a tax change creates no behavioral responses (as we rule out income effects) and hence the net fiscal gain for the government is $\int d\nu(i)$. At an optimal tax schedule, the benefits of such a reform would just equal its cost, so that $E(g) = 1$.

distribution. As the social marginal welfare weight on top bracket income earners is zero, τ^* is also the tax rate that maximizes the tax revenue collected from the top bracket.[9]

Proof: The proof of formula (2) is simple. It is useful to present it as derivations in Sections 3 and 4 build upon the derivation in this simple case. The government chooses τ to maximize tax revenue T from the top bracket (as the government puts no marginal social welfare weight on top bracket earners):

$$T = \tau[z(1 - \tau) - \bar{z}].$$

The first order condition is

$$[z - \bar{z}] - \tau \frac{dz}{d(1 - \tau)} = 0.$$

The first term in square brackets represents the mechanical tax gain of increasing τ while the second term represents the tax revenue loss due to behavioral responses to the tax increase. Introducing e_1 defined in equation (1), we can re-arrange the first order condition as

$$\frac{\tau}{1 - \tau} e_1 = \frac{z - \bar{z}}{z} = \frac{1}{a},$$

which can be rearranged into formula (2). **QED.**

Formula (2) shows that the optimal tax rate can be expressed in terms of two empirically estimable statistics: the Pareto parameter a of the top tail of the income distribution and the elasticity e_1 of income with respect to the net-of-tax rate.

Empirically, it is straightforward to estimate a as $z/(z - \bar{z})$ using income distribution data. In a given year and country, a is extremely stable with $\bar{z}$ (see e.g., Diamond and Saez, 2011 for an analysis using recent US data). In recent years, $a \simeq 1.5$ in the United States. Countries with less income concentration than the United States such as continental Europe or Japan have $a \simeq 2$ (see Atkinson, Piketty and Saez 2011).[10]

It is much more difficult to obtain a compelling empirical estimate of the elasticity e_1. Saez, Slemrod, and Giertz (2011) provide a recent survey of the literature to which we will come back later on in detail in Section 5. Formula (2) shows that the supply side elasticity e_1 is the key factor limiting how high the top tax rate can be. For example, with $a = 1.5$, $e_1 = .25$ yields a revenue maximizing tax rate τ^* of 73%, $e_1 = .5$, yields $\tau^* = 57\%$, and $e_1 = 1$ yields $\tau^* = 40\%$.

---

[9]If a positive social weight $g > 0$ is set on top earners marginal consumption, then the optimal top tax rate is $\tau = (1 - g)/(1 - g + ae)$.

[10]Note that a is endogenous to τ only if the elasticity is not constant within the top bracket. But formula (2) continues to apply even when a is endogenous, although it then becomes an implicit formula.

As mentioned above, the top income tax rate on earnings in the United States today is around $\tau = 42.5\%$ when taking into account all taxes (see Diamond and Saez, 2011). As we shall see, countries in continental Europe tend to have significantly higher top marginal tax rates on labor income, sometimes close to or above 60%.

In the remaining of the paper, we want to extend this tax model to account for other behavioral responses, namely tax avoidance (Section 3) and bargaining for pay (Section 4), and analyze how those additional elements alter the basic formula (2).

3   Tax Avoidance Responses

As shown by many empirical studies (see Saez, Slemrod, and Giertz, 2011 for a recent survey), responses to tax rates can also take the form of tax avoidance. We can define tax avoidance as changes in reported income due to changes in the form of compensation but not in the total level of compensation, and while keeping economic output constant. Examples of such avoidance/evasion are (a) reductions in current cash compensation for increased fringe benefits or deferred compensation such as stock-options or future pensions, (b) increased consumption within the firm such as better offices, vacation disguised as business travel, private use of corporate jets, etc. (c) changes in the form of business organization such as shifting profits from the individual income tax base to the corporate tax base, (d) re-characterization of ordinary income into tax favored capital gains, (e) outright tax evasion such as using off-shore accounts. In all those cases, tax avoidance opportunities arise because taxpayers can shift part of their taxable income into another form or another time period that is treated more favorably from a tax perspective.

The key distinction between real and tax avoidance responses is that real responses reflect underlying, deep individual preferences for work and consumption while tax avoidance responses depend critically on the design of the tax system and the avoidance opportunities it offers. While the government cannot change underlying deep individual preferences and hence the size of the real elasticity, it can change the tax system to reduce avoidance opportunities. For example, increased tax enforcement–perhaps through international cooperation–can reduce the use of off-shore accounts for tax evasion. Private consumption within the firm is also conceptually taxable and can be curtailed through tax enforcement efforts. Neutrality in the effective tax rates across organizational forms could also eliminate income shifting. Making fringe benefits

fully taxable instead of tax exempt would eliminate this tax avoidance opportunity as well. Similarly, aligning the tax rates on realized capital gains with those on ordinary income would eliminate the benefits of converting ordinary income into capital gains.

A number of papers have tried and capture avoidance effects for optimal tax design. Saez (2004) and Saez, Giertz, and Slemrod (2011) propose related analyzes but taking avoidance opportunities as given. Slemrod and Kopczuk (2002) endogenize avoidance opportunities in a multi-good model where the government selects the tax base. Finally, a large literature (surveyed in Slemrod and Yitzhaki (2002)) analyzes optimal policy design in the presence of tax evasion. Our model is more basic and tries to capture the key-tradeoffs as simply and transparently as possible.

3.1 Pure Tax Avoidance

We can extend the original model as follows to incorporate tax avoidance. Let us denote by y real income is y and by x sheltered income so that ordinary taxable income is $z = y - x$. The latter is taxed at marginal tax rate τ in the top bracket, while sheltered income x is taxed at a constant and uniform marginal tax rate t lower than τ. For example, in the case of untaxable fringe benefits, $t = 0$. In the case of capital gains conversion, $t > 0$ but is significantly less than τ. The utility function of individual i takes the form:

$$u_i(c, y, x) = c - h_i(y) - d_i(x),$$

where $c = y - \tau z - tx + R = (1 - \tau)y + (\tau - t)x + R$ is disposable after tax income and $R = \tau \bar{z} - T(\bar{z})$ denotes the virtual income coming out of the nonlinear tax schedule. $h_i(y)$ is the utility cost of earning real income y, and $d_i(x)$ is the cost of sheltering an amount of income x. There is a cost to sheltering, since sheltered income such as fringe benefits or deferred earnings are less valuable than cash income. We assume that both $h_i(.)$ and $d_i(.)$ are increasing and convex, and normalized so that $h_i'(0) = d_i'(0) = 0$. This model naturally nests the standard model when the sheltering cost $d_i(x)$ is infinitely large for any $x > 0$.

Individual utility maximization implies that

$$h_i'(y) = 1 - \tau \quad \text{and} \quad d_i'(x) = \tau - t,$$

so that y_i is an increasing function of $1 - \tau$ and x_i is an increasing function of the tax differential $\tau - t$. Aggregating over all top bracket taxpayers, we have $y = y(1 - \tau)$ with real elasticity

$e_1 = [(1 - \tau)/y]dy/d(1 - \tau) > 0$ and $x = x(\tau - t)$ increasing in $\tau - t$. Note that $x(0) = 0$ as there is sheltering only when $\tau > t$.

Hence $z = z(1 - \tau, t) = y(1 - \tau) - x(\tau - t)$ is increasing in $1 - \tau$ and t. We denote by $e = [(1 - \tau)/z]\partial z/\partial(1 - \tau) > 0$ the total elasticity of taxable income z with respect to $1 - \tau$ when keeping t constant. We denote by s the fraction of the behavioral response of z to $d\tau$ due to tax avoidance, and by $e_2 = s \cdot e$ the *tax avoidance elasticity component*:

$$s = \frac{dx/d(\tau - t)}{dy/d(1 - \tau) + dx/d(\tau - t)} = \frac{dx/d(\tau - t)}{\partial z/\partial(1 - \tau)} \quad \text{and} \quad e_2 = s \cdot e = \frac{1 - \tau}{z}\frac{dx}{d(\tau - t)}.$$

By construction, we have $(1 - s)e = (y/z)e_1$, or equivalently $e = (y/z)e_1 + e_2$. If we start from a situation with no tax avoidance ($y = z$), then we simply have $e = e_1 + e_2$, i.e. the total elasticity is the sum of the standard labor supply elasticity and the tax avoidance elasticity component.

Proposition 1 *Partial optimum: For a given t, the optimal top tax rate on taxable income is*

$$\tau^* = \frac{1 + t \cdot a \cdot e_2}{1 + a \cdot e}, \tag{3}$$

where $e = (y/z)e_1 + e_2$. is the total elasticity of taxable income (keeping t constant), $e_1 = [(1 - \tau)/y]dy/d(1 - \tau)$ is the real labor supply elasticity, and $e_2 = [(1 - \tau)/z]dx/d(\tau - t)$ is the tax avoidance elasticity component.

General optimum: If sheltering occurs only within top bracket earners, the optimal global tax policy is to set t and τ equal to

$$t^* = \tau^* = \frac{1}{1 + a \cdot e_1}, \tag{4}$$

Hence, sheltering becomes irrelevant in the full optimum.

Proof: As top bracket earners are of measure 1, the government chooses τ to maximize:

$$T = \tau[z(1 - \tau, t) - \bar{z}] + tx(\tau - t)$$

The first order condition for τ is

$$[z - \bar{z}] - \tau\frac{\partial z}{\partial(1 - \tau)} + t\frac{dx}{d(\tau - t)} = 0,$$

Introducing s, we can rewrite the first order condition as

$$[z - \bar{z}] - \tau\frac{\partial z}{\partial(1 - \tau)} + ts\frac{\partial z}{\partial(1 - \tau)} = 0.$$

10

The first two terms are the same as in the standard model. The third term captures the "fiscal externality" as a fraction s of the behavioral response translates into sheltered income taxed at rate t. Introducing $e = [(1 - \tau)/z]\partial z/\partial(1 - \tau) > 0$, we can rewrite the first order condition as

$$\frac{\tau - ts}{1 - \tau}e = \frac{z - \bar{z}}{z} = \frac{1}{a},$$

which can be rearranged into formula (3) using the fact that $e_2 = se$.

The second part of the proof can be obtained by taking the first order condition with respect to t and recalling that $z(1 - \tau, t) = y(1 - \tau) - x(\tau - t)$,

$$x + [\tau - t]\frac{dx}{d(\tau - t)} = 0.$$

Here we have used the assumption stated in the proposition that sheltering happens only within top bracket taxpayers so that a change in t has no effect on individuals below the top bracket. As $x \geq 0$ and $\tau \geq t$ and $dx/d(\tau - t) \geq 0$, this first order condition can only hold for $t = \tau$ and $x(\tau - t = 0) = 0$. Setting $t = \tau$ in equation (3), and noting that $x = 0$ implies that $z = y$ and hence $(1 - s) \cdot e = e_1$, we immediately obtain (4). Intuitively, as x is completely wasteful, it is optimal to deter x entirely by setting $t = \tau$. **QED.**

Four comments are worth noting about Proposition 1. First, if $t = 0$ then $\tau = 1/(1 + a \cdot e)$ as in the standard model. In the narrow framework where the tax system is taken as given (i.e. there is nothing we can do about tax evasion and income shifting), and where sheltered income is totally untaxed, then whether e is due to real responses vs. avoidance responses is irrelevant, a point made by Feldstein (1999).

Second however, if $t > 0$, then sheltering creates a "fiscal externality," as the shifted income is taxed at rate t and $\tau > 1/(1 + a \cdot e)$. As discussed earlier and as shown in the empirical literature (Saez, Slemrod, Giertz 2011), it is almost always the case that large short-term behavioral responses generated by tax changes are due to some form of income shifting or income re-timing that generates fiscal externalities.

Third and most important, the government can improve efficiency and its ability to tax upper incomes by closing tax avoidance opportunities (setting $t = \tau$ in our model), in which case the real elasticity e_1 is the only factor limiting taxes on upper incomes.[11]

---

[11]Kopczuk (2005) shows that the Tax Reform Act of 1986 in the United States, which broadened the tax base and closed loopholes did reduce the elasticity of reported income with respect to the net-of-tax rate.

Fourth, note that actual tax avoidance opportunities come in two varieties. Some are pure creations of the tax system, such as exemption of fringe benefits or tax exempt local government bonds and hence could be entirely eliminated by reforming the tax system. In that case, t is a free parameter that the government can change at no cost as in our model. Yet other tax avoidance opportunities reflect real enforcement constraints that are costly–sometimes even impossible– for the government to eliminate. For example, it is very difficult for the government to tax profits from informal cash businesses, consumption inside informal businesses, or off-shore tax evasion. Our simple model also ignores that there might be political hurdles to setting $t = \tau$, for example if some types of tax sheltering are fiercely defended by special interests or lobbying groups (Slemrod and Kopczuk 2002 present a model with costs of enforcement). The important policy question is then what fraction of the tax avoidance elasticity can be eliminated by tax redesign and tax enforcement. In a developing country with most economic activity taking place in small informal businesses, the tax avoidance elasticity cannot be reduced to zero. But in a modern economy and with international cooperation, the tax avoidance elasticity could be made minimal especially at the top of the distribution where virtually all economic transactions are recorded and hence verifiable (Kleven, Kreiner, and Saez, 2009).

3.2   Income Shifting

The previous avoidance model assumed that shifting was entirely wasteful so that there was no reason to want to set t lower than τ to start with. In reality, there are sometimes legitimate efficiency or distributional reasons why a government would want to tax different forms of income differently. On efficiency grounds, the classic Ramsey theory of optimal taxation indeed recommends taxing less the most elastic goods or factors.

Let us therefore extend our previous model by considering that there are two sources of income that we will call labor income and capital income for simplicity.[12] Labor income and capital income may respond to taxes differently and individuals can at some cost shift income from one form to the other. For example, in the case of labor vs. capital income in a dual tax system, small business owners can choose whether to get labor income as self-employed laborers or whether to get dividends from the profits of a closely held business.

We assume that labor income z_L is taxed nonlinearly with a top tax rate τ_L above $\bar{z}$, while

[12]Other examples could be individual income vs. corporate income, or realized capital gains vs. ordinary income, or self-employment earnings vs. employee earnings.

capital income z_K is taxed linearly at a constant and uniform tax rate τ_K. We make the simplifying assumption that all capital income is earned by individuals in the top labor income bracket so that changing τ_K has no impact on earners below the top bracket. Hence, all individuals who can potentially shift labor income into capital income have $y_L \geq \bar{z}$. True labor (respectively, capital) income is denoted by y_L, (respectively, y_K) while reported labor (respectively, capital) income is $z_L = y_L - x$ (respectively, $z_K = y_K + x$) where x represents the amount of income shifting between the tax bases. Individual i has utility function:

$$u_i(c, y_L, y_K, x) = c - h_{Li}(y_L) - h_{Ki}(y_K) - d_i(x),$$

with

$$c = R + (1 - \tau_L)z_L + (1 - \tau_K)z_K = R + (1 - \tau_L)y_L + (1 - \tau_K)y_K + (\tau_L - \tau_K)x$$

where $R = \tau_L \bar{z} - T(\bar{z})$ is virtual income created by the nonlinear labor income tax, $h_{Li}(y_L)$ is the cost of producing labor income y_L, $h_{Ki}(y_K)$ is the cost of producing capital income y_K, and $d_i(x)$ is the cost of shifting income from the labor to the capital base. We assume that h_{Li}, h_{Ki}, and d_i are all convex. Note that $d_i(x) \geq 0$ is defined for both positive and negative x. We naturally assume that $d_i(0) = 0$ and $d_i'(0) = 0$ and that $d_i'(x) \gtrless 0$ iff $x \gtrless 0$. This model nests the pure tax avoidance model of Section 3.1 in the case where $y_K \equiv 0$, i.e., there is no intrinsic capital income. Individual utility maximization implies that

$$h_{Li}'(y_L) = 1 - \tau_L, \quad h_{Ki}'(y_K) = 1 - \tau_K, \quad \text{and} \quad d_i'(x) = \tau_L - \tau_K,$$

so that y_{Li} is an increasing function of $1 - \tau_L$, y_{Ki} is an increasing function of $1 - \tau_K$, and x_i is an increasing function of the tax differential $\tau_L - \tau_K$. Aggregating over all top bracket taxpayers, we have $y_L = y_L(1 - \tau_L)$ with real elasticity $e_L > 0$, $y_K = y_K(1 - \tau_K)$ with real elasticity $e_K > 0$, and $x = x(\tau_L - \tau_K)$ increasing in $\Delta \tau = \tau_L - \tau_K$ and $x(0) = 0$.

Whether the elasticity of labor income e_L is larger or smaller than the elasticity of capital income e_K is very much an open issue. Of course a complete analysis of labor vs. capital taxation should also take into account dynamic issues, which we do not consider here (by choice, we focus upon the purely static, income shifting issue).[13]

---

[13]The taxation of capital raises two intrinsically dynamic issues: intergenerational transmission of capital (this tends to push toward higher taxation of capital than labor, assuming meritocratic social preferences) and intertemporal, within-a-lifetime allocation of consumption (this tends to push in the opposite direction). See Piketty and Saez (2011).

Note that $z_L = y_L - x(\Delta\tau)$ is more responsive than y_L to $1 - \tau_L$ when keeping τ_K constant as z_L responds along both the real margin and the avoidance margin. Similarly, z_K is more responsive to $1 - \tau_K$ than y_K. For example, if shifting has low cost, then x is very responsive to $\Delta\tau$. This implies that z_L is very responsive to changes in τ_L and z_K is very responsive to τ_K even if the underlying real corresponding incomes y_L and y_K are fully inelastic.

Finally, we define the Pareto parameter as $a_L = z_L/(z_L - \bar{z})$ for reported labor income and $a = (z_K + z_L)/(z_K + z_L - \bar{z})$ for total income.

Proposition 2 *The optimal tax rates τ_L and τ_K maximizing tax revenue are such that:*

No shifting Elasticity. *If $x \equiv 0$ (no income shifting), then $\tau_L = 1/(1 + a_L \cdot e_L)$ and $\tau_K = 1/(1 + e_K)$. We have: $\tau_L > \tau_K$ iff $a_L \cdot e_L < e_K$.*

Infinite shifting Elasticity. *In the limit where x' is very large and real responses have finite elasticities e_L and e_K, then $\tau_L = \tau_K = 1/(1 + a \cdot \bar{e})$ with $\bar{e} = (y_L e_L + y_K e_K)/(y_L + y_K)$ is the average real elasticity (weighted by income).*

General Case. *In case $a_L \cdot e_L < e_K$, we have: $1/(1 + a_L \cdot e_L) \geq \tau_L > \tau_K \geq 1/(1 + e_K)$. In the opposite case, we have the reverse inequality: $1/(1 + a_L \cdot e_L) \leq \tau_L < \tau_K \leq 1/(1 + e_K)$.*

Proof: See appendix A. **QED**.

Three comments on Proposition 2 are worth making. First, absent any shifting elasticity, there is no cross elasticity and we obtain the standard Ramsey inverse elasticity rule for each income factor.[14]

Second, a shifting elasticity brings the optimal tax rates τ_L and τ_K closer together (relative to the inverse elasticity rule). When the shifting elasticity is large, optimal tax rates τ_L and τ_K should be close–even if the real elasticities e_L and e_K are quite different. Importantly, the presence of shifting does not necessarily reduce the ability of the government to tax but only alters the relative mix of tax rates. For example, in the case with infinite shifting, the optimum tax rates on labor and capital are equal and should be based on the average of the real elasticities.

Third, in this simple model, deciding whether labor or capital income should be taxed more requires comparing the intrinsic elasticities e_L and e_K. Empirically, this would require increasing simultaneously both τ_L and τ_K to determine which factor responds most keeping the level of income shifting $x(\Delta\tau)$ constant.

---

[14]As we have no income effects, the elasticities are also compensated elasticities.

4 Compensation Bargaining Responses

Pay may not be equal marginal economic product for top income earners. In particular, executives can be overpaid if they are entrenched and can use their power to influence compensation committees. Indeed, a large literature in corporate finance has made those points (see for instance Bebchuk and Fried (2004) for an overview).

In principle, executives could also be underpaid relative to their marginal product if there is social outrage about high levels of compensation. In that case, a company might find it more profitable to under-pay its executives than face the wrath of its other employees, customers, or the public in general.[15] To the extent that top income earners generally have more opportunities to set their own pay than low and middle income earners, the first case seems more likely. But from a theoretical perspective both cases are interesting.

More generally, pay can differ from marginal product in any model in which compensation is decided by on-the-job bargaining between an employer and an employee, as in the classic search model of Diamond-Mortensen-Pissarides (see e.g. Pissarides (2000) for an exposition). In that framework, there is a rent to be shared on the job because of frictions in the matching process and inability to commit to a wage before the match has occurred. Indeed, in such models, the wage rate is not pinned down and can actually be anywhere in a band bounded by the outside options of the employer and the employee (Hall, 2005). Typically, the wage is then determined by the relative bargaining powers of the employer and employee, as is the case with Nash bargaining with exogenous weights. In general, the wage rate is not efficient, unless the so-called Hosios condition is met.[16] Given the substantial costs involved in replacing quits in most modern work environments, especially at management levels where specific human capital is important, it seems reasonable to think that there would be a band of possible compensation levels. In such a context, bargaining efforts on the job can conceivably play a significant role in determining pay. Marginal tax rates affect the rewards to bargaining effort and hence can possibly affect the level of such bargaining efforts.

Let us take an example which will be familiar to academic economists. In many University departments, pay is determined by outside options taking the form of competitive offers from

---

[15]Recent examples of such outrage have arisen in the case of the 2008 and 2009 bailouts of financial firms in the United States–although the effects on executives compensation has remained unclear.

[16]Those standard search models stand in contrast to newer "directed search" models where the wage is negotiated ex-ante in which case efficiency is restored (see e.g., Moen, 1997).

similarly or higher-ranked departments. Because moving costs are difficult to observe by the upper administration of one's home University, a formal competitive offer letter is often sufficient to trigger a pay increase in one's current job. Obtaining an outside offer for the sole purposes of getting a pay raise is costly and time consuming (both for the academic and to potential recruiters). If the pay raise in the home institution does not translate into higher productivity, then this is a pure compensation bargaining elasticity. Obviously, lower tax rates make the pay raise more valuable.[17]

There is relatively little work in optimal taxation that uses models where pay differs from marginal product. A few studies have analyzed optimal taxation in models with labor market imperfections such as search models (see e.g., Boone and Bovenberg, 2002), union models (see e.g., Fuest and Huber, 1997 and Aronsson and Sjogren, 2004), efficiency wages models (see e.g., Koskela and Schöb (2007)). Sorensen (1999) provides a simple overview of those models. The main focus of those papers has been on efficiency issues rather than redistribution issues, with most of the focus on the employment vs. unemployment margin. Therefore, most of those models do not incorporate heterogeneity among workers and hence cannot capture the issue of redistribution between workers as we do here. Fewer papers have addressed redistributive optimal tax policy in models with imperfect labor markets. Hungerbuhler et al. (2006) analyze a search model with heterogeneous productivity, and Stantcheva (2011) considers contracting models where firms cannot observe perfectly the productivity of their employees.

Most closely related to our paper, Rothschild and Scheuer (2011) consider a model with rent-seeking and earnings heterogeneity. The key difference between their core model and ours is that they consider a two-sector model where crowdable rent-seeking activities are limited to a single sector with no direct externalities to the other sector. As a result, in their core model, it is optimal to limit entry in the crowdable rent-seeking activity. High top tax rates discourage labor supply conditional on entry in the rent-seeking sector but may encourage entry which is inefficient. While this is a good model for a clearly segregated sector with a production limit–such as a natural resource extraction sector, it does not capture the notion that the pay of top earners can come at the expense of lower paid workers economy wide as in our model. Importantly and consistent with our analysis, they also obtain higher optimal tax rates when they allow externalities across sectors in an extension of their model. More broadly, their

---

[17]If the productivity of a given academic economist varies with department location and taxes reduce mobility, then the location response to taxes is partly a supply side e_1 elasticity.

approach is theoretical and uses a complex multi-dimensional screening approach. Hence, they are more interested in properties of the optimum, rather than developing simple tractable top rate formulas expressed in terms of estimable elasticities as we do here. Thus, we view our two contributions as complementary.

In this section, we consider the simplest model that can capture such bargaining compensation effects. Let us assume that individual i receives a fraction η of his/her actual product y. Individual i can put productive effort into increasing y or bargaining effort into increasing η. Both types of effort are costly to the individual. Hence, individual i utility is given by

$$u^i(c, \eta, y) = c - h_i(y) - k_i(\eta),$$

where c is disposable after-tax income, $h_i(y)$ is the cost of producing output y as in the standard model, and $k_i(\eta)$ is the cost of bargaining to get a share η of the product. Both h_i and k_i are increasing and convex.[18] We again rule out income effects as this simplifies substantially the derivations. Note that this model nests the standard model in the case where there is no cost to have $\eta = 1$ and infinite cost to the individual to pushing η above 1.

Let $b = (\eta - 1)y$ be bargained earnings defined as the gap between received earnings ηy and actual product y. Note that the model allows both overpay (when $\eta > 1$ and hence $b > 0$) and underpay (when $\eta < 1$ and hence $b < 0$). Let us denote by $E(b)$ the average bargained earnings in the economy. In the aggregate, it must be the case that aggregate product must be equal to aggregate compensation. Hence, if $E(b) > 0$, average overpay $E(b)$ must come at the expense of somebody. Symmetrically, if $E(b) < 0$, average underpay $-E(b)$ must benefit somebody. For simplicity, we assume that any gain made through bargaining comes at the expense of everybody else in the economy uniformly. Hence, individual incomes are all reduced by a uniform amount $E(b)$ (or increased by a uniform amount $-E(b)$ if $E(b) < 0$). A simple but admittedly unrealistic scenario in which our uniformity assumption holds is a situation where firms are owned equally in the population and bargaining pay comes at the expense of profits. We describe such a simple model fully in Appendix B. In reality, bargaining pay likely comes at the expense of other employees or shareholders in the same company. Some of the bargaining overpay might also be partly passed on to prices of the goods produced.[19]

---

[18]We could consider a general non separable cost of effort function $h_i(y, \eta)$ to allow for example for substitution between productive vs. bargaining effort. The optimal tax formula would be identical but the comparative statics would be less transparent and would require additional assumptions.

[19]We discuss below how relaxing our simplifying uniformity assumption would affect our results.

Because the government uses a nonlinear income tax schedule, it can adjust the demogrant intercept $-T(0)$ to fully offset $E(b)$. Effectively, the government can always tax (or subsidize) $E(b)$ at 100% before applying its nonlinear income tax. Hence, we can assume without loss of generality that the government absorbs one-for-one any change in $E(b)$. Therefore, we can simply define earnings as $z = \eta y = y + b$ and assume that those earnings are taxed nonlinearly. This simplification is possible because of our key assumption that $E(b)$ affects all individuals uniformly.

Individual i chooses y and η to maximize:

$$u^i(c, \eta, y) = \eta \cdot y - T(\eta \cdot y) - h_i(y) - k_i(\eta),$$

which leads to the first order conditions

$$(1 - \tau)\eta = h_i'(y) \quad \text{and} \quad (1 - \tau)y = k_i'(\eta),$$

where $\tau = T'$ is the marginal tax rate. This naturally defines y_i, η_i as *increasing* functions of the net-of-tax rate $1 - \tau$. Hence $z_i = \eta_i \cdot y_i$ and $b_i = (1 - \eta_i) \cdot y_i$ are also functions of $1 - \tau$.

Let us consider as in the previous section the optimal top tax rate τ above income level $\bar{z}$. We assume again that there is a mass of measure one of top bracket taxpayers. Let us denote by $z(1 - \tau)$, $y(1 - \tau)$, $b(1 - \tau)$ average reported income, productive earnings, and bargained earnings across all taxpayers in the top bracket. We can then define, as above, the real labor supply elasticity e_1 and the total compensation elasticity e as:

$$e_1 = \frac{1 - \tau}{y} \frac{dy}{d(1 - \tau)} \geq 0 \quad \text{and} \quad e = \frac{1 - \tau}{z} \frac{dz}{d(1 - \tau)}$$

We define s, the fraction of the marginal behavioral response due to bargaining and by $e_3 = s \cdot e$ the *bargaining elasticity component*:

$$s = \frac{db/d(1 - \tau)}{dz/d(1 - \tau)} = \frac{db/d(1 - \tau)}{db/d(1 - \tau) + dy/d(1 - \tau)} \quad \text{and} \quad e_3 = s \cdot e = \frac{1 - \tau}{z} \frac{db}{d(1 - \tau)}.$$

This definition immediately implies that $(y/z)e_1 = (1 - s) \cdot e$. By construction, $e = (y/z)e_1 + e_3$. If we start from a situation where top taxpayers are paid their marginal product ($y = z$), then we simply have $e = e_1 + e_3$ (in the same way as with the tax avoidance elasticity component e_2). Importantly, s (and hence e_3) can be either positive or negative but it is always positive if individuals are overpaid (i.e., if $\eta > 1$). If individuals are underpaid (i.e., $\eta < 1$) then s (and hence e_3) can be negative, as long as η satisfies a condition described in the following Lemma, the proof of which is straightforward.

Lemma 1 *We have:*

$$s = 1 - \frac{e_1}{\eta(e_\eta + e_1)} = 1 - \frac{y \cdot e_1}{z \cdot e} \le 1 \quad \text{with} \quad e_\eta = \frac{1-\tau}{\eta} \frac{d\eta}{d(1-\tau)} = e - e_1 \ge 0.$$

$$s \le 0 \quad \Leftrightarrow \quad \eta \le \frac{e_1}{e_1 + e_\eta}.$$
$$\text{If} \quad \eta > 1 \quad \text{then} \quad s > 0.$$

We can now state our main proposition.

Proposition 3 *The optimal top tax rate is*

$$\tau^* = \frac{1 + a \cdot e_3}{1 + a \cdot e} = 1 - \frac{a(y/z)e_1}{1 + a \cdot e}, \tag{5}$$

where $e = (y/z)e_1 + e_3$ is the total elasticity of taxable income, $e_1 = [(1-\tau)/y]dy/d(1-\tau) = (z/y)(1-s)e$ is the real labor supply elasticity, and $e_3 = s \cdot e = [(1-\tau)/z]db/d(1-\tau)$ is the compensation bargaining elasticity component.

τ^* decreases with e (keeping e_3 constant) and increases with e_3 (keeping e constant). τ^* decreases with the real elasticity e_1 (keeping e and y/z constant) and increases with the level of overpayment $\eta = z/y$ (keeping e_1 and e constant)

If $e_1 = 0$ then $\tau^* = 1$.

If $z \ge y$ (top earners are overpaid) then $\tau^* > 1/(1 + a \cdot e_1)$

Proof: The government aims at maximizing taxes collected from the top bracket. Taxes collected from the top bracket are $\tau[z - \bar{z}]$ but the top bracket tax rate τ also impacts $E(b)$ and hence the government budget (as the government absorbs one-to-one any change in $E(b)$). Since the total size of the population is N (recall top earners are of measure one), the government chooses τ to maximize:

$$T = \tau[y(1-\tau) + b(1-\tau) - \bar{z}] - N \cdot E(b),$$

Importantly, if $d\tau$ triggers a change in b in the top bracket, that change is then reflected one-to-one in $NE(b)$. Hence we have $db/d(1-\tau) = NdE(b)/d(1-\tau)$ and the first order condition for τ is:

$$[y + b - \bar{z}] - \tau\frac{dy}{d(1-\tau)} - \tau\frac{db}{d(1-\tau)} + \frac{db}{d(1-\tau)} = 0,$$

which can be rewritten as

$$\tau \left(\frac{db}{d(1-\tau)} + \frac{dy}{d(1-\tau)} \right) - \frac{db}{d(1-\tau)} = z - \bar{z}$$

$$[\tau - s]\frac{dz}{d(1-\tau)} = z - \bar{z}$$

$$\frac{\tau - s}{1 - \tau} \cdot e = \frac{z - \bar{z}}{z} = \frac{1}{a}$$

which can be rearranged into (5) using the fact that $e_3 = se$. The rest of the proposition is straightforward. **QED.**

Proposition 3 shows that it is possible to obtain a simple optimal tax formula that nests the standard model in the case $e_3 = 0$ (no bargaining elasticity). Implementing the formula requires knowing the total elasticity e and the bargaining elasticity component e_3 defined as the fraction of the behavioral response (at the margin) due to bargaining earnings changes. e_3 can also be indirectly be obtained by substraction from e using the real labor supply elasticity e_1 and the ratio of product to pay y/z. Hence, implementing the formula requires to be able to know not only how compensation responds to tax changes but also how real economic product responds to tax changes, which is considerably more difficult than estimating the standard taxable income elasticity e.

Trickle-Up. In the case where top earners are overpaid relative to their productivity ($z > y$), then $s > 0$ and hence $e_3 > 0$ and the optimal top tax rate is higher than in the standard model (i.e., $\tau^* > 1/(1 + a \cdot e)$). This corresponds to a "trickle-up" situation where a tax cut on upper incomes shifts economic resources away from the bottom and toward the top. Those effects can be quantitatively large. For example, if we assume that most of the surge in top income shares in the United States has been driven by top rate tax cuts then $e = 1$ as we shall see in Section 5.1. With $a = 1.5$ for the Unites States in recent years (Piketty and Saez, 2003), the conventional model obtains an optimal top tax rate of $\tau = 1/(1 + 1.5 \cdot 1) = 40\%$, which is about equal to the current top tax rate equal to 42.5% when including all taxes (Diamond and Saez, 2011). However, if you assume that the real labor supply elasticity of top earners is only 0.5 (an upper bound based on standard labor supply estimates focusing on hours of work), and that top earners in the United States are paid at their marginal product today ($y = z$), then $\tau = 1 - 1.5 \cdot 0.5/(1 + 1.5) = 70\%$ considerably larger than the conventional rate of 40%. As is well known from the executive compensation literature (see e.g., Murphy 1999), US executives

20

are paid significantly more than executives in Japan or Continental Europe even controlling for characteristics of the company. Hence, if we are further willing to assume that US top earners are overpaid, for example paid twice as much as their actual product, then $y/z = 0.5$ and $\tau = 1 - 1.5 \cdot 0.5 \cdot 0.5/(1 + 1.5) = 85\%$. This discussion shows that bargaining effects, which are plausible but not easy to measure, can substantially alter optimal tax policy recommendations. We return to this issue when we present our empirical estimates of e_1, e_2 and e_3 in section 5 below.

Trickle-Down. In the case where top earners are underpaid relative to their productivity ($z < y$) then it is possible that $s < 0$ and hence $e_3 < 0$ (see the Lemma), in which case the optimal top tax rate is lower than in the standard model (i.e., $\tau < 1/(1 + a \cdot e)$). This corresponds to a "trickle-down" situation where a tax cut on upper incomes also shifts economic resources toward the bottom, as upper incomes work in part for the benefit of lower incomes. Coming back to the discussion above, another interpretation of the differential in executive pay between the United States vs. Japan and Continental Europe is that executives in Japan and Continental Europe are underpaid relative to their product. Consider for example the case of Japan where top income have not responded much to the substantial cuts in top marginal tax rates (Moriguchi and Saez, 2008) so that $e = .25$ is a reasonable elasticity and $a = 2$. The conventional optimal top tax rate would be $\tau = 1/(1 + 2 \cdot .25) = 67\%$. Suppose further that $e_1 = .25$ and that η is rigid, with $e_\eta = 0$ as there might be little scope for bargaining for top executives in Japan, perhaps because of rigid executive pay structures. Assume that $y/z = 2$, so that top earners in Japan are paid only 50% of their product. Then, the optimal top tax rate would be only $\tau = 1 - 2 \cdot 2 \cdot 0.25/(1 + 2 \cdot .25) = 33\%$. In effect, the optimal top rate is so low because 50% of upper incomes are transferred to the rest of society. In that context, it is particularly inefficient to discourage their labor supply. This again shows that it is critical to know the extent to which upper earners are overpaid vs. underpaid in order to determine the optimal top tax rate.

Regulation vs. Taxation. We have taken as given the bargaining opportunities in the economy. Conceivably, the government can affect bargaining opportunities through regulations. A large literature in corporate finance analyzes whether regulations can impact executive compensation (see e.g., Frydman and Jenter 2010 and Murphy 2011 for recent discussions). In a reduced form way, regulations would impact the cost of bargaining $k_i(\eta)$ but our analysis of the optimal tax would remain valid taking regulations are given.

Ideally, as bargaining is a wasteful effort that shifts resources without any real productive effect, the government would want to completely discourage it, so that pay would always be equal to real economic product. In that case, bargaining effects disappear and we naturally revert to the standard model. However, as long as some bargaining effects subside, our analysis is relevant. We leave an optimal regulation analysis taking into account both benefits and costs of regulation versus taxation for future work.

Differentiated Taxation. Conceivably, some economic sectors or industries might be more prone to bargaining effects than others. For example, less competitive industries have higher rents and hence more scope for bargaining effects. In that case, differentiated tax rates across industries could conceivably be desirable. The same argument calls for differentiated tax rates in the standard model if some sectors have a higher labor supply elasticity. In practice, there are two important arguments against differentiated taxation. First, it would be difficult to measure bargaining effects for each sector. This uncertainty might allow the better paid lobbyists to argue in favor of preferential tax rates for their industry. Second, differentiated tax rates create additional distortions if there are opportunities to shift income from one sector to another.

Non uniform external effects. We have made the key assumption that aggregate external effects $E(b)$ are spread in a uniform and lumpsum fashion among all individuals. That simplifies the formula as the social value of the external effect created by bargaining effects by the tax change $d\tau$ is one-to-one. This is the case because the government can exactly undo the external effect by simply shifting the schedule and adjusting the demogrant. In general, the external effects will not be uniformly distributed. In that case, if a negative external effect is tilted toward lower income earners, it has more social cost and the optimal tax rate would be higher than formula (5). Conversely, if the external effect is tilted toward upper income earners, it has less social cost and the optimal tax rate would be lower than formula (5).[20]

Charitable giving. Charitable giving is conceptually related to the bargaining model. Charitable giving is equivalent to giving away a fraction of one's income to society. That fraction responds to the tax rate as higher tax rates encourage charitable giving when charitable contributions can be deducted from taxable income as in the United States (see Andreoni 2006 for a survey of the empirical literature). In other words, if compensation bargaining reflects

---

[20]This is somewhat complicated by the fact that the external income is taxable (in the non-uniform case, the government can no longer undo the external effect).

greed, charitable giving reflects altruism.[21] Charitable giving also has an external effect on society. If we make the assumption that the external effect of charitable giving is lumpsum and uniformly distributed in the population, then formula (5) applies where e_3/e is the fraction of the taxable income elasticity due to charitable giving responses.

However, an important difference between charitable giving and bargaining effects is that charitable giving is observable and hence it is possible to subsidize charitable giving independently of the income tax rate. As discussed above this is not possible for bargaining as the real product is not directly observable by the government.

Putting the three elasticities together. We can also put the three elasticities together in a single formula. If we have at the same time tax avoidance effects and compensation bargaining effects, then we can write the total elasticity of taxable income e as the sum of three terms: $e = (y/z)e_1 + e_2 + e_3$. In case we start from a situation where there is no tax avoidance activity and incomes are equal to marginal products, then $y = z$ and we simply have: $e = e_1 + e_2 + e_3$. For a given tax rate t on sheltered income, we have:

$$\tau^* = \frac{1 + t \cdot a \cdot e_2 + a \cdot e_3}{1 + a \cdot e}$$

If the tax administration can choose t optimally to fully eliminate tax avoidance, we have:

$$\tau^* = t^* = \frac{1 + a \cdot e_3}{1 + a \cdot (e_1 + e_3)}$$

If government puts a social welfare weight $0 \leq g < 1$ on marginal consumption of top earners (relative to the average), then the optimal top rate formula becomes

$$\tau^* = \frac{1 - g + t \cdot a \cdot e_2 + a \cdot e_3}{1 - g + a \cdot e}$$

5   Empirical Exploration and Policy Implications

In this section, we want to use our model to account for the evolution of top tax rates and top incomes in OECD countries. We first lay out in Section 5.1 the key empirical facts using US evidence and then international evidence on top income shares gathered in the *World Top Incomes Database* and top income tax rates. Next in Section 5.2, we lay out the various scenarios that have been proposed to explain the key facts and their tax policy consequences.

---

[21]Ironically, both can happen in the same person such as the "Robber Barons" of the Gilded Age developing monopoly power to extract rents and then bequeathing their fortunes to charitable causes.

5.1 Key Facts

US Evidence

US evidence is depicted in graphical form in Figure 2 and key estimates are presented in Table 1. Panel A of Figure 2, taken from Piketty and Saez (2003) depicts the top 1% income shares including realized capital gains (pictured with full diamonds) and excluding realized capital gains (the empty diamonds).[22] Both top income shares, whether including or excluding realized capital gains, display an overall U-shape over the century. Panel A also displays (on the right y-axis) the top marginal tax rate for the Federal individual income tax for ordinary income (dashed line) and for long-term realized capital gains (dotted line). Two important lessons emerge from this panel.

Considering first the top income share excluding realized capital gains which corresponds roughly to income taxed according to the regular progressive schedule, there is a clear negative overall correlation between the top 1% income share and the top marginal tax rate: (a) the top 1% income share was high before the Great Depression when top tax rates were low (except for a short period from 1917 to 1922), (b) the top 1% income share was consistently low between 1932 to 1980 when the top tax rate was uniformly high, (c) the top 1% income share has increased significantly since 1980 after the top tax rate has been greatly lowered. This clear visual correlation suggests that the overall elasticity of reported incomes is high. For the recent period that is of most interest for current policy debates, the top 1% income share more than doubled from around 8% in the late 1970s to around 18% in last five years, while the net-of-tax (retention) rate increased from 30% (when the top marginal tax rate was 70%) to 65% (when the top tax rate is 35%). If we attribute the entire surge in the top income share to the decline in the top tax rate, this translates into an elasticity of top incomes with respect to the net-of-tax rate around one, as shown in column (1), Panel A of Table 1. Column (1) of Panel B in Table 1 also shows a strong correlation between the net-of-tax rate and the top income share with a basic time series regression of the form

$$\log(\text{Top 1\% Income Share}) = a + e \cdot \log(1 - \text{Top MTR}) + \varepsilon$$

---

[22]Those series are based on the family unit (and not the individual adult). Income includes cash market income before individual taxes and credits, and excludes government transfers (such as Social Security benefits, unemployment insurance benefits, or means-tested transfers) as well as non-cash benefits (such as employer or government provided health insurance).

This link remains the same when including a linear time trend in the regression.[23] The implied elasticity is around 0.25-0.30 and very significant. Importantly, as the average marginal tax rate faced by the top 1% was smaller than the statutory top rate before the 1970s, our elasticity estimate is likely to be downward biased.[24]

Second, the correlation between the top shares and the top tax rate also holds for the series including capital gains. Realized capital gains have been traditionally tax favored (as illustrated by the gap between the top tax rate and the tax rate on realized capital gains in the figure) and have constituted the main channel for tax avoidance of upper incomes.[25] Under the tax avoidance scenario, taxable income subject to the progressive tax schedule should be much more elastic than a broader income definition that also includes forms of income that are tax favored. Indeed, in the pure tax avoidance scenario, total real income should be completely inelastic. However, both the graphical analysis of Panel A and the estimates presented in Table 1, column (2) show that the link between the top tax rate is as strong for income including realized capital gains as it is for income excluding capital gains. The implied elasticity for the late 1970s vs. today is slightly in excess of one for income including realized capital gains. The time series regressions also generate virtually identical estimates as the series excluding capital gains. This suggests that income shifting responses do not account for much of the evolution in top income shares documented in Figure 2.[26] In the short-run, to be sure, there is strong evidence on Panel A of large tax avoidance responses in various tax reform episodes with clear differential responses for top incomes including vs. excluding realized capital gains.[27] But in the long run

[23]Naturally, the correlation disappears when additional polynomials in time are added as identification is based solely on time variations.

[24]The solution would be to instead use the actual average marginal tax rate faced by the top 1% instrumented with the top marginal tax rate (as in Saez, 2004). Unfortunately, actual top 1% marginal tax rate series are not available before 1960 and would be very time consuming to construct.

[25]When individual top tax rates are high (relative to corporate and realized capital gains tax rates), it becomes more advantageous for upper incomes to organize their business activity using the corporate form and retain profits in the corporation. Profits only show up on individual returns as realized capital gains when the corporate stock is eventually sold. See e.g., Vickrey (1947) for an early analysis and Gordon and Slemrod (2000) for a more modern analysis, yet strikingly similar in its conclusions.

[26]In future work, it would be useful to sharpen this test by (a) subtracting deductions–such as charitable giving or interest paid on debt–from the narrow income definition to come closer to taxable income, (b) adding forms of income that are non-taxable–such as tax exempt interest, capital gains unrealized till death, or fringe benefits to further broaden the broader income definition. There is no easy route to do this as most of those items are not reported consistently and continuously in income tax statistics (Fack and Landais (2008) have constructed homogeneous series for charitable giving).

[27]For example, in 1986, realized capital gains surged in anticipation of the increase in the capital gains tax rate from 20 to 28% (Auerbach (1988)), a clear spike in the series including capital gains in the Figure. From 1986 to 1988, income excluding realized capital gains surged as closely held businesses shifted from the corporate form to the individual form, and as many business owners paid themselves accumulated profits as wages and salaries

the income shifting elasticity e_2 (as estimated long the ordinary income vs capital gains margin) appears to be small (say, $e_2 < 0.1$).

Clearly, capital gains are not the only channel through which income avoidance can occur. Off-shore accounts and perquisites also come to mind. However, if anything, it seems that those have increased at the same time as top rates have declined. For the former channel, Zucman (2011) for example shows that a growing fraction of Swiss fiduciary deposits are recorded as belonging to tax havens since the 1970s. For the latter, it is notoriously hard to find historical data, as disclosure rules for perquisites have only recently been imposed[28] but perquisites would have had to be huge pre-1970 to generate a high elasticity of avoidance through that channel. Indeed, there is both anecdotal[29] and more rigorous evidence on the high level of perks today (Yermack (2006), Grinstein, Weinbaum and Yehuda (2008)), which are several times the $0.74 million median total pay of top executives pre-1970s (Frydman and Saks (2010)). Despite the lack of data, it is hard to believe that perks could possibly have been so high in the past.

The most difficult question to resolve is whether this large responsiveness of top incomes to tax rates is due to supply side effects generating more economic activity as in the standard model of Section 2 or whether it is due to a zero-sum game transfer from the bottom 99% to the top 1% as in the bargaining model of Section 4. This is the critical question to decompose the total elasticity e into e_1 and e_3 effects. Panel B of Figure 2 tries to cast preliminary light on this issue by plotting the evolution of top 1% incomes and bottom 99% incomes adjusting for price inflation.[30] The graph shows clearly that income growth for the bottom 99% was highest

---

(Slemrod (1996), Saez (2004)). Such shifting increased reported ordinary income at the expense of realized capital gains, explaining why there is a big discontinuity in income excluding realized capital gains but not in income including realized capital gains. Finally, there is a clear surge in incomes in 1992 in anticipation of the increase in the top tax rate on ordinary income in 1993 due to re-timing in the exercise of stock-options for executives (Goolsbee (2000)). See Saez, Slemrod, and Giertz (2011) for a much more detailed discussion.

[28]Regulation introduced in December, 1978 required firms to disclose only the total amount of remuneration distributed in the form of securities or property, insurance benefits or reimbursement, and personal benefits. Only in 1993 were perquisites and other personal benefits (above a minimum threshold) separately reported. Even then, the data poses problems in terms of transparency and accuracy.

[29]For example, research firm Morningstar reports that Diller, CEO of Expedia spent a combined 1.3 million on personal flights in 2010, almost twice the pre-1970 total CEO pay. D.Blankenship, former CEO of Massey Energy received as post-employment perks a secretary for 5 years, two year health care benefits, a two-year consulting contract, a house and free company land. The New York Times ('From Coffee to Jets, Perks for Executives Come Out in Court', Feb 22 2004) reports that two of Tyco International's executives got paid their children's tuition fees, a $1 million birthday party in Sardinia, as well as the now infamous $6,000 shower curtain and $15,000 umbrella stand.

[30]To control for changes in the number of adults per family, we plot income per adult (aged 20 and over) assuming that the top 1% income share at the individual adult level is the same as at the family level. This assumption holds true in countries such as Canada where top income shares can be constructed both at the individual and family levels (Saez and Veall, 2005).

in the 1933 to 1973 period when top income tax rates were high and the growth of top 1% was modest. Conversely, the growth of bottom 99% incomes has slowed down since the 1970s when top tax rates came down and top 1% incomes grew very fast. Those graphical impressions can be captured by a basic regression analysis of the form:

$$\log(\text{Real Income}_{gt}) = a + b \cdot \log(1 - \text{Top MTR}_t) + c \cdot t + \varepsilon$$

where g indexes either the Bottom 99% or the Top 1% or the overall average income and t denotes the year. We naturally control for time to capture overall exogenous growth independent of tax policy. The estimate b, reported in Table 2 Panel C, is positive and highly significant for the top 1% incomes, with a magnitude around .25 very similar to the time series elasticity estimation of Panel B. In contrast, the estimate b is negative (and just significant at the 5% level with a t-stat around 2) for the bottom 99%, and close to zero and insignificant for the overall average income. Again, the estimates are very similar for income excluding capital gains in column (1) and for income including capital gains in column (2).

This evidence is consistent with the bargaining model where gains at the top have come at the expense of the bottom. In principle, the estimate b obtained for the overall average income can be used to compute e_1. I.e. if the model is well identified we have: $b = \pi \cdot e_1$, where π is the initial income share of top marginal tax rate taxpayers. That is, if we take $\pi = 10\%$,[31] then a doubling of the net-of-top-marginal-tax-rate should lead to a $b = 10\%$ rise in the average real income of the economy if the real supply side elasticity e_1 was 1. But since we find that b is close to zero and insignificant for the overall average income, it is tempting to conclude that e_1 is also small and insignificant, and that the overall elasticity e comes mostly from bargaining effects through e_3.

Note also that this evidence can also be used to rule out the possibility of significant unrecorded tax avoidance effects. That is, assume that in the 1950s-1970s top income earners were escaping high top rates via consumption within the firm (counted as intermediate consumption within corporations -fancy restaurants, corporate jets, etc. - and therefore unrecorded in GDP estimates) or tax havens (again unrecorded in GDP estimates, at least partly).[32] If such tax

---

[31]The exact fraction of taxpayers falling in the top marginal rate bracket varies over time. In recent decades, it is generally larger than the top 1% (it is often closer to the top 2%-3%), so $\pi = 10\%$ should be viewed as a lower bound (implying that the estimates for e_1 should be viewed as upper bounds).

[32]Other forms of tax avoidance such as deferred compensation or legal income shifting would be recorded in GDP.

avoidance had declined significantly in the recent period, then this should show up as extra economic growth. I.e. in presence of such unrecorded tax avoidance activities, then the estimate b obtained for the overall average income should actually be equal to: $b = \pi \cdot (e_1 + e_2)$. In any case, this suggests that the overall elasticity e comes mostly from e_3 effects.

However, this evidence relies on the strong OLS assumption that any deviation of growth from trend (captured by the error term ε) is uncorrelated with the top marginal tax rate. It is conceivable that economic growth could have slowed down in the 1970s for reasons unrelated to the top tax rate decreases. This could have driven down the bottom 99% income growth as well. In that case, the cut in top tax rates could have increased top incomes growth as in the supply side scenario without negatively impacting bottom 99% incomes. Indeed, growth slowed down in many OECD countries after the oil shocks of the 1970s. Therefore, this evidence based on a single country is at best suggestive. Hence, to make further progress, we now turn to international evidence.

International Evidence

To analyze international evidence, we use data on the income shares of the top 1% from 18 OECD countries, gathered in the *World Top Incomes Database* combined with top income tax rate data since 1975. We focus on the period since 1975 because this allows us to include more countries (a number of countries in the top income database have data only for recent decades) and to gather top tax rate data relatively easily. In addition, focus on the recent period is interesting because of the very divergent trends across countries in both top income shares and top tax rates.[33] Our top income tax rates series include both the central and local government top tax rates on ordinary income. We do not include payroll taxes as those taxes apply only to wage earnings which constitute only a fraction of top 1% incomes and are often capped. We do not include consumption taxes either.[34] We provide complete details on the construction of top tax rates and on the variables' construction in Appendix C. Top incomes are defined as cash market income excluding capital gains and subject to the regular income tax (see Atkinson and Piketty 2007, 2010 volumes for complete country level details).

---

[33]Roine, Vlachos, and Waldenstrom (2009) have used the database to explain the long-run determinants of inequality over the full century, including the top tax rate as an explanatory variable among many others. They present overall regressions without focusing specifically on the recent decades as we do here.

[34]As a robustness check, we have constructed and done the analysis including all uncapped payroll and consumption taxes as well. The results were very similar and available from the authors upon request.

We start in Figure 3 by showing the link between the top tax rate (on the x-axis) and the top 1% income share (on the y-axis) for 1975-9 (Panel A) and 2004-8 (Panel B) across countries. If the country does not have data for those years, we select the first available five years after 1975 and the most recent 5 years (also described in Appendix C). Panel A shows that there was a very wide dispersion in top tax rates across OECD countries in the late 1970s with rates as low as 40% for Spain and Switzerland and above 75% for Sweden, the United Kingdom, or Japan. The United States had fairly high top tax rates, around 70%, significantly above those of France and Germany. Top 1% income shares were uniformly fairly modest with a maximum just above 10% for Germany and a minimum slightly below 5% for Sweden and Denmark. The graph shows that there was a weak negative correlation between top shares and top tax rates. Panel A1 in Table 2 confirms that the implied elasticity obtained from regressing the log of the top income share on the log of the net-of-tax rate across those 18 countries is around .33 and only marginally significant.

Panel B of Figure 3 shows a dramatic shift by 2004-8. Top tax rates are much lower than they were in the late 1970s with no country above 60% and a number of countries clustering around 40% including the United States, or the United Kingdom. There is also much more heterogeneity in top income shares which vary from a low of 4% for Denmark to a high of almost 18% for the United States. Importantly, there is also a much stronger negative correlation between top tax rates and top income shares in 2004-8 than in 1975-9. As reported in Table 2, Panel A1, the implied elasticity for 2004-8 is extremely large around 1.40 and highly significant.

In order to extend the 1970s vs. today comparison we did for the United States to our 18 OECD countries, Panel A of Figure 4 plots the change in top income shares from 1975-9 to 2004-8 (on the y-axis) against the change in the top marginal tax rate (on the x-axis) for all the countries. The figure shows a very clear and strong correlation between the cut in top tax rates and the increase in the top 1% income share with interesting heterogeneity. Countries such as France, Germany, Spain, Denmark or Switzerland which did not experience any significant top rate tax cut did not experience changes in top income shares. Among the countries which experienced significant top rate cuts, some experience a large increase in top income shares (all 5 English speaking countries but also Norway and Finland) while others experience only modest increases in top income shares (Japan, Italy, Sweden, Portugal, or the Netherlands). Interestingly, no country experiences a significant increase in top income shares

without implementing significant top rate tax cuts.

This graph provides two lessons. First, it shows a very strong link between top tax rates and top income shares. The implied elasticity reported in Table 2, Panel A2 using a regression of the log-change in top income shares on the log-change in the net-of-tax rate over those 18 countries generates a fairly large and highly significant elasticity around .5. Additional regressions in Panel A3 using the complete time series of the form

$$\log(\text{Top 1\% Income} : \text{Share}_{it}) = a + e \cdot \log(1 - \text{Top MTR}_{it}) + \varepsilon_{it}$$

also generate estimates around .5, and are robust to the introduction of an overall time trend or country fixed effects.[35] Second, the implied elasticity varies significantly across countries with strong effects in English speaking countries, and particularly the United States where the elasticity is around one,[36] and much more modest effects in other countries such as Japan, Sweden, or Italy, where the elasticity is only around .2. This suggests that the elasticity likely depends on the institutional set-up of each country.

To tell apart the supply side vs. the bargaining scenario, Panel B of Figure 4 plots the annual real GDP per capita growth from 1975-9 to 2004-8 (on the y-axis) against the change in the top marginal tax rate (on the x-axis) for all the countries. Under the supply-side scenario, a cut in top rates translates into additional economic activity among upper incomes, hence higher top income shares but also higher economic growth. In contrast, under the bargaining scenario, a cut in top tax rates generates a "trickle-up" transfer from lower to upper incomes with an increase in top income shares but no additional economic activity.

The graph displays no visible correlation between the change in top tax rates and growth rates. The countries experiencing the largest increases in top income shares (United States, United Kingdom, or Canada) have growth rates that are comparable to those of France, Germany, or Denmark who did not experience any significant top rate cuts and top income share increases. Panel B of Table 2 provides additional regression evidence using the complete time

---

[35]Estimates using both country and time fixed effects generate smaller elasticities as they rely on year-to-year variation for identification. Our analysis focuses instead on long-run effects of top tax rates.

[36]Canada has an even larger implied elasticity but as argued by Saez and Veall (2005), part of the surge in Canadian top income shares might have been driven by brain drain threats following the surge in US top compensation rather than internal Canadian tax policy.

series and specifications of the form:

$$\log(\text{Real GDP per Capita}_{\text{it}}) \ = \ a + b \cdot \log(1 - \text{Top MTR}_{\text{it}}) + c \cdot t + \varepsilon_{it}$$

$$\text{with:} \qquad b = \pi \cdot e_1$$

Both regressions excluding or including country fixed effects do not reveal any significant effect of the top tax rate on real GDP per capita.[37] The magnitude and precision of the estimates also rule out large effects. For example, doubling the retention rate (as done in the US for top incomes since 1980) would only increase GDP per capita by .1 percentage point in the case of the estimate with fixed effects. If we take $\pi = 10\%$, as in the previous sub-section, and an absolute upper bound $b = 0.2$ (accounting for the standard errors), then assuming that these regressions are properly identified we can exclude the possibility of a real supply side elasticity e_1 larger than 0.2. Given that the overall elasticity e is about 0.5, this would suggest that the compensation bargaining elasticity e_3 is at least 0.3.[38] Importantly and as mentioned above, as the top statutory rates in 1970s sometimes applied to less than the top 1%, the average marginal tax rate effectively faced by the top 1% was likely smaller than the statutory top rate. Hence our elasticity estimate e and the growth effect b are likely to be biased downward, but by the same factor so that the ratio of the two estimates should be unbiased. All this tends to imply that e_1 is at most 40% of the total elasticity.

Naturally, those regressions are only suggestive as they rely on a very strong identifying assumption: any deviation of GDP growth from its trend (not due to top tax rate factors) is uncorrelated with the evolution of top tax rates.[39] The goal of this analysis is not to provide full-fledged evidence of the bargaining scenario but rather to show that, a first pass macro-level

_____

[37]We have also run regressions using GDP per adult, GDP per working age adult, or GDP per worked hour. In all cases, we find small and insignificant effects.

[38]As was noted above, in the presence of unrecorded tax avoidance effects e_2 we would have $b = \pi(e_1 + e_2)$, in which case $e_1 + e_2 < 0.2$. In any case, this corresponds to an elasticity e_3 of at least 0.3.

[39]Many potential factors could invalidate this assumption. For example, if countries cut top tax rates when their growth is expected to slow down (for example if Anglo-Saxon countries during the 1970s feared to be overtaken by Continental Europe and Japan, and opted for the Thatcher-Reagan revolution as a way to fight relative decline), that would generate a spurious negative correlation between growth and the net-of-tax top rate - thereby implying that then elasticity e_1 is under-estimated (one would however have to explain the reason for this structural decline). Factors going in the other direction include the voluntary reduction in working hours that took place in a number of Continental European countries since the 1970s-1980s - which in principle should have reduced their growth performance relatively to Anglo-Saxon countries with long working hours. It could also be that there exists a spurious positive correlation between top income growth and net-of-tax top rates (both being caused by changing social norms and beliefs about top end inequality), thereby implying that we are overestimating the total elasticity e.

analysis does not reject the non-conventional bargaining model in favor of the standard model used in tax analysis.

At the same time, there is ample direct micro evidence in favor of a bargaining view of top earners', and especially CEOs', compensation (see Frydman and Jenter (2010), Section 4, for a good summary of both sides of the debate). First, parts of compensation packages are hidden from shareholders which should not be the case if pay was set competitively (Bebchuk and Fried (2004), Kuhnen and Zwiebel (2009)). Consequently, new disclosure rules almost certainly lead to an increase in reported earnings. Secondly, CEOs frequently receive rewards for good outcomes that are not the result of their own effort (but instead for example, of a booming economy) and are not symmetrically punished for unlucky events (Bertrand and Mullainathan (2001), Garvey and Milbourn (2006)). CEOs do better when there is a lack of competition, since their pay increases following reductions in a takeover threat (Bertrand and Mullainathan (1998)). On the other hand, their compensation decreases after regulatory changes aimed at improving board control (Chhaochharia and Grinstein (2009)). Furthermore, there is widespread malpractice in compensation setting which seems to indicate rent-extraction. For example, 30% of firms from 1996 to 2005 seem to have used 'options backdating' (which consists in choosing the "grant dates" ex-post to allow for the minimal strike price of at-the-money options) (Lie (2005), Heron and Lie (2007, 2009), Narayanan and Seyhun (2008)). Bebchuk et. al (2010) further show that this practice occurred more in firms with weak boards. Another practice suggestive of rent-extraction is spring loading (Yermack (1997)) which happens when CEOs are awarded options right before the release of positive news. Finally, it is also possible that the increased use of stock options, has helped managers to disguise their increased rent extraction, since few people might have thoroughly understood option valuation (Hall and Murphy (2003), Bebchuk and Fried (2004)).

5.2   Scenarios

We can now bring together our theoretical and empirical analysis to evaluate the plausibility and policy consequences of each of the key scenarios we outlined in introduction that have been put forward to explain the surge in top incomes in recent decades. Those scenarios and tax implications are summarized in Table 3.

(0) **Skill-Biased Technological Change** This scenario posits that technological progress

has been skill-biased and has favored top earners relative to average earners. The labor economics literature has debated the merits of the skill-biased hypothesis (see Katz and Autor, 1999 for a survey). In the case of top earners, this hypothesis takes the form of "Winner-Takes-All" theories whereby highly talented individuals can deploy their skills on a broader and worldwide market, hence increasing the marginal product of any given unit of talent. Gabaix and Landier (2008) propose a model of the market for CEOs along those lines. The theory of skill-biased technological progress is largely independent of behavioral responses to taxation.

As we have seen, this scenario cannot explain why only some OECD countries have experienced a surge in top income shares and why that surge has been highly correlated with the drop in top marginal tax rates. Indeed, all OECD countries have been subject to similar technological and globalization forces and hence should have experienced the same change in income concentration under the basic skill-biased technological change scenario.

(1) Supply Side Tax Effects (e_1). This scenario posits that the drop in top tax rates has led to an increase in top income shares through a standard supply side effect whereby highly skilled individuals work and earn more. In this case, the standard model is valid and there is no avoidance nor bargaining elasticity. If this scenario is correct, then we can interpret the overall cross-country elasticity $e = 0.5$ as deriving from standard supply side effects: $e = e_1 = 0.5$ and $e_2 = e_3 = 0$. With $a = 1.5$ (the Pareto coefficient currently prevailing in the U.S.), the top tax rate maximizing tax revenue would then be $\tau^* = 1/(1+a \cdot e) = 57\%$ (see Table 3). With $a = 2.0$ (prevailing in many European countries), the top tax rate maximizing tax revenue would be only $\tau^* = 50\%$. This is less than the current top tax rate (combining all taxes) currently applied in a number of European countries. Hence, decreasing top tax rates would be a desirable policy both from the point of view of top earners but also from the point of view of the bottom 99% as taxes collected from upper incomes would increase. This would also imply that the high top US tax rates of the 1970s were set well above the revenue maximizing rate.

However, this scenario creates three major difficulties. First, it somewhat strains credibility to believe that top 1% earners in the U.S. had enough leeway to be able to drastically increase their work effort. Any objective measure of labor supply such as hours of work or those based on retirement behavior does not show any such large increase.[40]

---

[40]For example, Moffitt and Wilelhm (2000) show that, while top incomes surged after the Tax Reform Act of 1986, hours of work of top earners measured in the Survey of Consumer Finances did not change.

Second, the link between the surge in top income shares and top rate cuts is not perfect. Some countries, such as Japan, have cut their top tax rates about as much as the United States and yet have experienced no surge in top shares. This suggests that the behavioral response to taxes might depend on the tax system and institutions rather than on some universal preferences on work and leisure.

Thirdly, and most importantly, the supply side scenario implies that the surge in top incomes is due to additional economic activity and does not come at the expense of lower incomes. Therefore, countries who cut their top tax rates should have experienced more economic growth than other countries a prediction that is not borne out by a simple cross-country comparison. The critical problem with the supply-side scenario is simply that the United States or the United Kingdom did not experience significantly higher economic growth than Germany or France or Japan since the 1970s. On the basis of these cross-country comparisons, a large e_1 does not seem plausible and it is likely that $e_1 \leq 0.2$.

(2) Tax Avoidance Effects (e_2) This scenario posits that the link between top income shares and top tax rates is due to a large avoidance elasticity. When tax rates are high, top income earners find ways to exploit loopholes and report less of their taxable income. As shown in the survey by Saez, Slemrod, and Giertz (2011), there is indeed evidence that income shifting responses–where shifting can occur either over time or across tax bases–can be important in the short-run whenever such tax avoidance opportunities are created by tax changes. This tax avoidance channel was first presented as a left-wing critique of the supply-side scenario, since part of the extra individual income tax came at the expense of tax revenues from other tax bases or other time periods. More recently, this tax avoidance argument has been used to deny that any real increase in US income inequality has taken place (Reynolds (2007)).

Under this scenario, the avoidance elasticity is large while the standard supply side elasticity is modest. Under the current US tax regime with its existing loopholes, the optimal tax rate should be $\tau = (1 + t \cdot a \cdot e_2)/(1 + a \cdot e)$. It is difficult to estimate t accurately but tax avoidance exploits primarily deferral and the favorable treatment of capital gains, so that a marginal tax rate t of 20% is perhaps reasonable. If we assume $e = 0.5$, $e_1 = 0.2$, $e_2 = 0.3$, $e_3 = 0$, $a = 1.5$ and $t = 20\%$, then we get $\tau^* = 62\%$ which is somewhat larger than the 57% optimal tax arising from the pure supply side scenario due to the "fiscal externality" (see Table 3). More importantly however, the deeper policy implication is that one needs to first close tax avoidance

opportunities, in order to reduce the shifting elasticity and only then increase the top tax rate.[41] As shown in Table 3, if the government can broaden the base and reduce the avoidance elasticity from 0.3 to 0.1, then the optimal top tax rate increases to 71%.

The weakness of the tax avoidance scenario is that taxable income subject to the progressive tax schedule should be much more elastic than a broader income definition that also includes forms of income that are tax favored. Our US long-run analysis rejects this scenario as we found that income including realized capital gains is exactly as elastic as income excluding them, using a full century long data series. Our US and cross-country analysis also seem to reject the possibility of large unrecorded tax avoidance effects.

(3) Compensation Bargaining Effects (e_3) Under this scenario, the high top tax rates of the 1970s were part of the institutional setup putting a brake on top compensation through bargaining or rent extraction effects. Lower top tax rates induces top earners to bargain more aggressively for higher pay.

Under this scenario the bargaining elasticity is large while the supply side (and avoidance) elasticities are modest. The optimal tax rate is sensitive to (a) how large the supply side elasticity is, (b) whether top earners are overpaid vs. underpaid. If the supply side elasticity is very small, then the optimal top tax rate will be large (100% in the limit where there is no supply side elasticity). If the supply side elasticity is not large and top earners are not underpaid, then the optimal top tax rate is also high. The optimal top tax rate will not be high only if the supply elasticity is large *and* top earners are underpaid (since then it would be desirable to induce top earners to work harder for the benefit of others).

The main difficulty with this scenario is that it is difficult to obtain compelling direct evidence that the surge in top incomes did come at the expense of lower earners. The US evidence over a century is consistent with this scenario. International evidence since 1975 is also consistent with this scenario. From this evidence, the most reasonable estimates would be $e = 0.5$, $e_1 = 0.2$ (at most), $e_2 = 0.0$, $e_3 = 0.3$ (at least), which together with $a = 1.5$ would imply $\tau^* = 83\%$ (see Table 3). Of course these estimates are far from being well identified. But they illustrate the critical importance of the decomposition of the overall elasticity into three elasticities.

As mentioned above, the fact that the overall elasticity seems to vary across countries,

---

[41]The study of Kopczuk (2005) convincingly shows that the Tax Reform Act of 1986 in the United States was indeed successful in reducing the elasticity of taxable income with base broadening and loopholes closing.

being high in English speaking countries and small in Continental Europe or Japan, could also mean that the bargaining elasticity depends on the country's institutional setup. Presumably, in Europe or Japan, it might still be difficult for top executives to bargain for higher pay even though top tax rates are no longer very high. In contrast, it is conceivable that in the United States and the United Kingdom, the Reagan and Thatcher conservative revolutions also weakened other institutional barriers to top pay such as Unions or social intolerance for pay inequality.

6   Conclusion

Our paper has analyzed the problem of optimal taxation of top labor incomes by considering three channels of behavioral responses to taxation: (1) the standard supply-side channel through reduced work effort, (2) the tax avoidance channel of income shifting for tax minimization purposes, (3) the compensation bargaining channel through efforts in influencing own pay setting. We have derived simple optimal top tax rate formulas as a function of the three elasticities corresponding to those three channels of responses. We have shown that the models generate very different predictions for the optimal top marginal tax rate. Outside of the standard supply-side mode, the elasticity of taxable income with respect to the net-of-tax rate is no longer a sufficient statistics. Hence, it is critical to understand the channel of the behavioral responses. Our first pass empirical analysis looking at US and international top income shares and top tax rates shows that there is a strong link between top tax rates and top income shares implying that the overall elasticity is large. Even though the data are consistent with the bargaining effect scenario, it is difficult to compellingly distinguish empirically the standard supply-side channel from the compensation bargaining channel, let alone evaluate to what extent top earners are overpaid vs. underpaid relative to their product. We hope that future research using a more micro-approach and more credible identification sources will make progress on this key issue.

Our paper has focused primarily on the case where the government sets a zero social marginal welfare weight on top earners. This is useful to determine the upper bound revenue maximizing tax rate on upper incomes. In reality, the welfare weight put on top earners by society is likely to depend on perceptions of whether top pay is fair or not. In the supply side scenario, pay is fair by definition and hence a zero weight can only be justified by strong redistributive motives–which might hold in some but not all OECD countries. In the tax avoidance scenario, the public might

36

perceive upper incomes as taking unfair advantage of the tax system and hence this might lower the marginal welfare weight that society puts on top earners. Finally, in the bargaining model, if top earners are overpaid at the expense of lower paid workers, then top pay would naturally be considered as unfair and that could translate into a very low social welfare weight on top earners as the rich would be perceived as undeserving.[42] Indeed, one of the central argument in the recent "Occupy Wall Street" protests in the United States is that pay going to the top 1% is unfair because it has come at the expense of the remaining 99% percent. Historically, the perception of unfair pay at the top has certainly played a key role in the development of very progressive taxation in the first part of the 20th century in most advanced countries that we documented in Figure 1.[43] Therefore, social perceptions are likely to further widen the differences in the socially desired level of top earners taxation across the different models relative to our analysis with zero welfare weights. It is also possible that higher income shares raise the ability of top income groups to influence social perceptions (e.g. by funding think tanks or medias that are more pro-rich), thereby creating some reverse causality between income inequality, perceptions and policies. Economists can play a key role in enlightening those perceptions by evaluating empirically which economic model of top incomes determination accounts best for the facts.

---

[42]In contrast, if top earners are perceived as being underpaid relative to their contributions, they would naturally be seen as a deserving group working in part for the benefits of others, and hence could have a high social welfare weight.

[43]See Brownlee (2004) for a US political history of Federal taxation and Piketty (2001) for a historical analysis in the case of France.

A   Proof of Proposition 2

Let us now consider again the tax rates τ_L, τ_K that maximize tax revenue raised from the top bracket earners (recall that top bracket earners are of measure 1).

$$T = \tau_L[y_L(1 - \tau_L) - x(\tau_L - \tau_K) - \bar{z}] + \tau_K[y_K(1 - \tau_K) + x(\tau_L - \tau_K)]$$

The first order conditions with respect to τ_L and τ_K are:

$$(z_L - \bar{z}) - \tau_L y'_L - \tau_L x' + \tau_K x' = 0$$

$$z_K - \tau_K y'_K + \tau_L x' - \tau_K x' = 0,$$

which, using $\Delta\tau = \tau_L - \tau_K$, can be rewritten as:

$$\frac{z_L - \bar{z}}{y'_L} - \Delta\tau \cdot \frac{x'}{y'_L} = \tau_L$$

$$\frac{z_K}{y'_K} + \Delta\tau \cdot \frac{x'}{y'_K} = \tau_K.$$

Taking the difference of those two equations, we can express $\Delta\tau$ as

$$\Delta\tau = \tau_L - \tau_K = \frac{(z_L - \bar{z})/y'_L - z_K/y'_K}{1 + x' \cdot (1/y'_L + 1/y'_K)},$$

Hence, as the denominator is always positive,

$$\Delta\tau > 0 \Leftrightarrow \frac{z_L - \bar{z}}{y'_L} > \frac{z_K}{y'_K} \Leftrightarrow \frac{y'_K}{z_K} > \frac{y'_L}{z_L} \frac{z_L}{z_L - \bar{z}}.$$

Introducing $e_K = (1 - \tau_K)y'_K/y_K$, $e_L = (1 - \tau_L)y'_L/y_L$, and $a_L = z_L/(z_L - \bar{z})$, we can rewrite this condition as:

$$\Delta\tau > 0 \Leftrightarrow \frac{e_K}{1 - \tau_K} \frac{y_K}{z_K} > \frac{a_L e_L}{1 - \tau_L} \frac{y_L}{z_L}.$$

Recall that $\Delta\tau = \tau_L - \tau_K$, $z_K = y_K + x(\Delta\tau)$, $z_L = y_L - x(\Delta\tau)$, hence we have

$$\Delta\tau > 0 \Leftrightarrow e_K > a_L e_L \cdot \frac{1 + x(\Delta\tau)/y_K}{1 - x(\Delta\tau)/y_L} \cdot \frac{1 - \tau_K}{1 - \tau_K - \Delta\tau}.$$

As $x(\Delta\tau) > 0$ iff $\Delta\tau > 0$, we have $\Delta\tau > 0 \Rightarrow e_K > a_L e_L$. Conversely, $\Delta\tau < 0$ implies $x(\Delta\tau) < 0$ and hence $e_K < a_L e_L \frac{1 + x(\Delta\tau)/y_K}{1 - x(\Delta\tau)/y_L} \frac{1 - \tau_K}{1 - \tau_K - \Delta\tau} < a_L e_L$. Finally, $\Delta\tau = 0 \Rightarrow e_K = a_L e_L$. Hence, we have:

$$\Delta\tau \lesseqgtr 0 \Leftrightarrow e_K \lesseqgtr a_L e_L.$$

We can rewrite the first order conditions as

$$\frac{\tau_L}{1 - \tau_L} y_L e_L = z_L - \bar{z} - \Delta\tau \cdot x'$$

$$\frac{\tau_K}{1-\tau_K} y_K e_K = z_K + \Delta\tau \cdot x'$$

Hence if $a_L e_L < e_K$ then $\Delta\tau > 0$ and hence $e_L \tau_L/(1-\tau_L) \leq (z_L - \bar{z})/y_L = [(z_L - \bar{z})/z_L] \cdot z_L/y_L = [y_L - x(\Delta\tau)]/(y_L a_L) \leq 1/a_L$ which implies $\tau_L \leq 1/(1 + a_L e_L)$. Similarly, $e_K \tau_K/(1 - \tau_K) \geq z_K/y_K = (y_K + x(\Delta\tau))/y_K \geq 1$ which implies $\tau_K \geq 1/(1 + e_K)$. This completes the proof of the general case.

If $x \equiv x' \equiv 0$ then $y_K = z_K$ and $y_L = z_L$ and the first order conditions above immediately imply that $\tau_L = 1/(1 + a_L \cdot e_L)$ and $\tau_K = 1/(1 + e_K)$

If x' is very large relative to y'_L and y'_K, then the expression above for $\Delta\tau$ implies that $\Delta\tau \simeq 0$ and hence $\tau_L \simeq \tau_K$. In that case, summing the two first order conditions above leads to

$$\frac{\tau_L}{1-\tau_L} y_L e_L + \frac{\tau_K}{1-\tau_K} y_K e_K = z_L + z_K - \bar{z} = y_L + y_K - \bar{z},$$

Using the fact that $\tau_L \simeq \tau_K$, we have:

$$\frac{\tau_L}{1-\tau_L} \simeq \frac{\tau_K}{1-\tau_K} = \frac{y_L + y_K - \bar{z}}{y_L e_L + y_K e_K} = \frac{1}{a \cdot \bar{e}},$$

as $a = (z_K + z_L)/(z_K + z_L - \bar{z}) = (y_K + y_L)/(y_K + y_L - \bar{z})$ which completes the proof of the third part of the proposition. **QED**

B A Simple Bargaining Model

In this appendix, we present a very simple micro-founded model that is consistent with our bargaining effect specification in Section 4, i.e., with a payoff of the type:

$$u^i(c, \eta, y) = \eta \cdot y - T(\eta \cdot y) - h_i(y) - k_i(\eta) - E(b) \tag{6}$$

where, as in the main text, y is productive output, η is the factor by which income is scaled up or down through bargaining, $T()$ is the nonlinear tax schedule with a lump-sum demogrant $T(0)$, $h_i(y)$ the cost of producing output, and $k_i(\eta)$ the cost of bargaining. All agents were assumed to bear the cost of the average bargaining in the economy, $-E(b)$ uniformly. The goal is simply to show that it is possible to propose a micro-foundation for the formulation in (6). For simplicity, and to be able to focus exclusively on the bargaining channel, we abstract from the intensive labor supply margin.

Consider an economy made out of firms and N workers. Workers are engaged in a search for jobs with matching frictions, as in Pissarides (2000) which lasts for only one period. At the beginning of the period, all workers start in the unemployed state and firms open vacancies at a cost c. The number of vacancies is determined in equilibrium by the free-entry condition. Given the N unemployed workers and V vacancies, there are $M = m(N, V)$ matches formed, where $m()$ is a standard constant returns to scale matching function. Denote by $\mu = \frac{V}{N}$ the market

tightness variable and let $q(\mu) = \frac{M}{V} = m\left(\frac{1}{\mu}, 1\right)$ be the probability that a vacancy is filled (note also that $\mu q(\mu)$ is the probability that a worker finds a job). We further assume that each of the N workers in the economy owns a fraction $\frac{1}{N}$ of every firm.

Once a firm and a worker match, they determine the worker's pay, w, by engaging in Nash bargaining over the surplus of the match. Workers are indexed by i and have heterogeneous productivities θ_i so that output is simply $y_i = \theta_i$. Let J be the expected value of a filled vacancy for a firm which is given by:

$$J = E(\theta - w) \tag{7}$$

The free entry condition then implies that:

$$c = q(\mu) J \tag{8}$$

Let W_i denote the value of employment for a worker of type θ_i, which in our static case is simply equal to the wage:

$$W_i = w_i \tag{9}$$

If bargaining breaks down, the worker receives an exogenous outside option w_{0i}, and the firm gets $-c$. Denoting by β_i and $(1 - \beta_i)$ the (potentially heterogeneous) bargaining weights of worker i and the firm which is matched to him respectively, the wage rate w_i is set to maximize:

$$\max_{w_i} (W_i - w_{0i})^{\beta_i} (J_i)^{1-\beta_i}$$

subject to (7) and (9).

From the first-order conditions we have that:

$$W_i - w_{0i} = \beta_i (W_i + J_i - w_{0i}) = \beta_i (\theta_i - w_{0i})$$

So that the wage equation is:

$$w_i = (1 - \beta_i) w_{0i} + \beta_i \theta_i$$

Furthermore, we assume that the outside option is proportional to the worker's productivity (for example because more productive people can better start their own business, immigrate abroad for a higher wage, etc..), so that $w_{i0} = \gamma_i \theta_i$. Thus we can rewrite the wage equation as:

$$w_i = \eta_i \theta_i$$

where $\eta_i = (1 - \beta_i) \gamma_i + \beta_i$. Hence, we find that $z_i = \eta y_i$ as in the main text. Note that if the worker has sufficiently high bargaining power (i.e., β_i is high enough) and the outside option is very attractive, then it is possible to have $\eta_i \geq 1$ even though both β_i and γ_i are less than 1, so that the worker can earn more than his marginal product θ_i.[44] Suppose that in order to

---

[44]However, on average, the expected profit of the firm conditional on having found a match (J in equation (7)) is positive, so as to compensate the firm for opening the costly vacancy.

affect η_i, the worker needs to incur a certain cost $k_i(\eta_i)$. Also note that there are potentially two channels through which η_i can be changed, namely the pure bargaining weight β_i and the outside option γ_i.[45]

In addition, the number of vacancies is implicitly determined in equilibrium by the equation:

$$\frac{c}{q(\mu)} = E\left((1 - \eta_i)\,\theta_i\right)$$

where the expectation is over all workers. The firm employing worker i receives a profit, denoted π_i:

$$\pi_i = (1 - \beta_i)(\theta_i - w_{0i}) = (1 - \eta_i)\,\theta_i$$

Denote total profits in the economy by Π. They are equal to:

$$\Pi = \sum_{i=1}^{N} I_i \pi_i - Vc = \sum_{i=1}^{N} I_i (1 - \eta_i)\,\theta_i - Vc$$

where I_i is an indicator variable taking the value 1 if worker i is has found a match and is employed. Each worker receives $\frac{1}{N}\Pi$ from his firm portfolio. Let $b_i = (\eta_i - 1)\,\theta_i$ denote the bargaining part of the wage (noting that $w_i = y_i + b_i$). For unemployed workers, $b_i = 0$. Hence we can rewrite:

$$\frac{1}{N}\Pi = -\frac{1}{N}\sum_{i=1}^{N} b_i - \frac{V}{N}c$$

Thus, a worker receives from his firm ownership: $\frac{\Pi}{N} = -E(b) - \frac{V}{N}c$ where $E(b)$ denotes the average bargaining in the economy, where the use of the expectation operator is a slight abuse of notation.[46] We can ignore the last constant $\frac{V}{N}c$ which can be fully absorbed by the demogrant transfers in our model. Hence, we obtain equation (6) that we used in the main text.

C   Data Sources

C.1   Tax Data

Income tax data (at the national and local level), as well as payroll taxes on employer and employee, come mainly from the OECD annual "Taxing Wages" publication which covers the

---

[45]We could alternatively have taken $\gamma_i = 0$ and let all the bargaining effect go through the Nash bargaining weight β.

[46]Note that by the free-entry condition, in *expectation* $J = E(\theta - w) = -E(b) = c\frac{V}{M}$ so that the expected $\frac{\Pi}{N}$ is ex ante: $cV\left(\frac{1}{M} - \frac{1}{N}\right) \geq 0$ which means that on average, workers are underpaid conditional on matching (to allow the firm to recoup the costs of their vacancies in expectation). However, for any realization of the matching process, and depending on which workers get matched, $\frac{\Pi}{N}$ could be either negative or positive. The government demogrant will capture it in any case. In addition, it leaves open the possibility that all top earners are overpaid (have a positive b) and that all others are underpaid (have a negative b) and that on average profits remain positive conditional on matching.

period from the early 1980s to the present. For the period 1975-1983 taxes are summarized in the publication "Personal income tax systems for the period 1975-1983." Consumption tax data comes from the OECD publications "Consumption Tax Trends" (for 2010, 2008, 2006, 2004, 2001, 1999) which summarize consumption taxes since 1975. For specific countries, additional sources were used.

United States: Federal tax parameters from the Tax Policy Center. The average state tax rate is estimated using actual top statutory state income tax rates weighted by the fraction of high income tax returns in each state (as of 2007).

Japan: Local taxes were taken from the National Tax Administration data.[47]

Canada: The tax rates series were taken from and described in Saez and Veall (2007).

All our results were performed using two measures of the effective marginal top tax rate: the first measure, reported in the text uses only the national and local income taxes. The second measure, included in addition payroll taxes on both the employee and employer and consumption taxes (VAT or sales taxes). Our results were broadly similar using both measures (and are available on demand), but the higher confidence we have in the quality of the income tax data led us to report results using the first measure in the text.

The available time spans vary for each country and we use the largest possible time period. They are as follows: Australia (1975-2007), Canada (1975 - 2000), Denmark (1980 - 2005), Finland (1975 - 2004), France (1975 - 2006), Germany (1975 - 1998), Ireland (1975 - 2000), Italy (1975 - 2004), Japan (1975 - 2004), The Netherlands (1975 - 1999), New Zealand (1975 - 2005), Norway (1975 - 2008), Portugal (1976 - 2005), Spain (1981- 2008), Sweden (1975 - 2009), Switzerland (1975 - 1995), United Kingdom (1975 - 2005), United States (1975 - 2008).

C.2   GDP, Population and Top Incomes Share Data

GDP series in constant US dollars are taken from the OECD statistics database (available at stats.oecd.org). Our results were also cross-checked with income data from tax returns, provided in the World Top Incomes Database.

Population Series for various age groups, used to construct the number of adults were also taken from the OECD Statistics.

Total employment and average hours of work were taken from the OECD Statistics, except for countries for which too many missing entries were found. In particular, for Ireland, the Netherlands, Switzerland, Denmark and Germany data from the International Labor Office on employees and hours was used instead.

Data on the Top 1% income shares comes from the World Top Incomes Database (Alvaredo et al. 2011).

---

[47]We thank Yusuke Narita for kind help with the translation of the Japanese files.

References

Alvaredo, Facundo, Anthony Atkinson, Thomas Piketty, and Emmanuel Saez. 2011. *The World Top Incomes Database*, online at http://g-mond.parisschoolofeconomics.eu/topincomes/

Andreoni, James. 2006. "Philanthropy." in *Handbook on the Economics of Giving, Reciprocity and Altruism* vol. 1. Elsevier. Chapter 18, 1201–1269.

Aronsson, Thomas, and Tomas Sjogren. 2004. "Is the Optimal Labor Income Tax Progressive in a Unionized Economy?", *Scandinavian Journal of Economics*, 106(4), 661-675.

Atkinson, Anthony, and Thomas Piketty. 2007. *Top Incomes over the Twentieth Century.* Oxford: Oxford University Press: 585 pp.

Atkinson, Anthony, and Thomas Piketty. 2010. *Top Incomes: A Global Perspective.* Oxford: Oxford University Press.

Atkinson, Anthony, Thomas Piketty, and Emmanuel Saez. 2011. "Top Incomes in the Long-Run of History", *Journal of Economic Literature* 3-71.

Auerbach, Alan. 1988. "Capital Gains Taxation in the United States." *Brookings Papers on Economic Activity*, 2: 595-631.

Bartels, Larry M. 2008. *Unequal Democracy: The Political Economy of the New Gilded Age*, Princeton University Press: Princeton.

Bebchuk, Lucian, and Jesse Fried. 2004. *Pay without Performance: The Unfulfilled Promise of Executive Compensation,* Harvard University Press: Cambridge.

Bebchuk, Lucian, Grinstein, Yaniv, and Urs Peyer. 2010. "Lucky CEOs and Lucky Directors," *Journal of Finance,* 65(6), 2363-2401

Bertrand, Marianne, and Sendhil Mullainathan. 2001. "Are CEOs Rewarded for Luck? The Ones without Principals Are," *Quarterly Journal of Economics,* 116(3), 901-932

Bertrand, Marianne, and Sendhil Mullainathan. 1998. "Executive Compensation and Incentives: The Impact of Takeover Legislation," NBER Working Paper No. 6830.

Boone, Jan, and Lans Bovenberg. 2002. "Optimal Labor Taxation and Search," *Journal of Public Economics,* 85, 53–97.

Brownlee, Eliott W. 2004. *Federal Taxation in America: A Short History*, Cambridge University Press: Cambridge, MA.

Chhaochharia, Vidhi, and Yaniv Grinstein. 2009. "CEO Compensation and Board Oversight", *Journal of Finance* 64(1), 231-261.

Diamond, Peter. 1998. "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates", *American Economic Review* 88, 83-95.

Diamond, Peter, and Emmanuel Saez. 2011. "The Case for a Progressive Tax: From Basic Research to Policy Recommendations," CESifo Working Paper No. 3548, August 2011, forthcoming *Journal of Economic Perspectives.*

Fack, Gabrielle, and Camille Landais. 2010. "Charitable Giving and Tax Policy in the Presence of Tax Cheating: Theory and Evidence from the US and France," SIEPR Working

Paper.

Feldstein, Martin. 1995. "The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act." *Journal of Political Economy*, 103(3): 551-572.

Feldstein, Martin. 1999. "Tax Avoidance and the Deadweight Loss of the Income Tax." *Review of Economics and Statistics*, 81(4): 674-680.

Frydman, Carola, and Dirk Jenter. 2010. "CEO Compensation," *Annual Review of Financial Economics* 2, 75-102.

Frydman, Carola, and Raven Saks. 2010. "Executive Compensation: A new View from a Long-Term Perspective, 1936-2005", *Review of Financial Studies* 23, 2099-2138.

Fuest, Clemens, and, Bernd Huber. 1997. "Wage Bargaining, Labor-tax Progression, and Welfare," *Journal of Economics* 66, 127–150.

Gabaix, Xavier, and Augustin Landier. 2008. "Why has CEO Pay Increased So Much?," *Quarterly Journal of Ecoonomics* 123, 49-100.

Garvey, Gerald, and Todd T. Milbourn. 2006. "Asymmetric Benchmarking in Compensation: Executives are Rewarded for Good Luck but not Penalized for Bad," *Journal of Financial Economics* 82(1), 197-225.

Goolsbee, Austan. 2000. "What Happens When You Tax the Rich? Evidence from Executive Compensation." *Journal of Political Economy*, 108(2): 352-378.

Gordon, Roger, and Joel Slemrod. 2000. "Are 'Real' Responses to Taxes Simply Income Shifting Between Corporate and Personal Tax Bases?" In *Does Atlas Shrug? The Economic Consequences of Taxing the Rich*, ed. Joel Slemrod, 240-288. New York: Russell Sage Foundation and Harvard University Press.

Grinstein, Yaniv, David Weinbaum, and Nir Yehuda. 2008. "Perks and Excess: Evidence from the New Executive Compensation Disclosure Rules." Working paper, Cornell University, Johnson Graduate School of Management, Cornell University

Hacker, Jacob, and Paul Pierson. 2010. *Winner-Take-All Politics: How Washington Made the Rich Richer–and Turned Its Back on the Middle Class*, Simon and Schuster.

Hall, Robert E. 2005. "Employment Fluctuations with Equilibrium Wage Stickiness," *American Economic Review*, 95 (1), 50–65.

Hall, Brian, and Kevin J. Murphy. 2003. "The Trouble with Stock Options," *Journal of Economic Perspectives*, 17 (3), 49–70.

Heron Randall A., and Erik Lie. 2007. "Does Backdating Explain the Stock Price Pattern around Executive Stock Option Grants?" *Journal of Financial Economics,* 83(2), 271–295.

Heron Randall A., and Erik Lie. 2009. "What Fraction of the Stock Option Grants to Top Executives Have Been Backdated or Manipulated?" *Management Science,* 55(4), 513–525.

Hungerbuhler, Mathias, Lehmann, Etienne, Parmentier, Alexis, and Van Der Linden, Bruno. 2006. "Optimal Redistributive Taxation in a Search Equilibrium Model," *Review of Economic Studies,* 73, 743–767.

International Labor Office Labor Statistics Database. http://laborsta.ilo.org/

Japan National Tax Administation Data.

http://www.soumu.go.jp/main_sosiki/jichi_zeisei/czaisei/czaisei_seido/ichiran06_h17.html

Katz, Lawrence F., and David H. Autor. 1999. "Changes in the Wage Structure and Earnings Inequality," in O. Ashenfelter and D. Card (eds.) *Handbook of Labor Economics,* volume 3A. Amsterdam: North-Holland.

Kleven, Henrik, Claus Kreiner, and Emmanuel Saez. 2009. "Why Can Modern Governments Tax So Much? An Agency Model of Firms as Fiscal Intermediaries," NBER Working Paper No. 15218.

Kopczuk, Wojciech. 2005. "Tax Bases, Tax Rates and the Elasticity of Reported Income." *Journal of Public Economics*, 89(11-12): 2093-2119.

Kuhnen, Camelia, and Jeffrey Zwiebel. 2008. "Executive Pay, Hidden Compensation and Managerial Entrenchment." Working Paper, Northwestern University

Koskela, Erkki, and Ronnie Schöb. 2007. "Tax Progression Under Collective Wage Bargaining and Individual Effort Determination," CESifo Working Paper No. 2024, forthcoming *International Relations: A Journal of Economy and Society*

Lie, Erik. 2005. "On the Timing of CEO Stock Option Awards." *Management Science*, 51(5): 802-812.

Lindsey, Lawrence. 1987. "Individual Taxpayer Response to Tax Cuts: 1982-1984, with Implications for the Revenue Maximizing Tax Rate." *Journal of Public Economics*, 33(2): 173-206.

Mirrlees, James. 1971. "An Exploration in the Theory of Optimal Income Taxation." *Review of Economic Studies*, 38: 175-208.

Moen, Espen R. 1997. "Competitive Search Equilibrium" *Journal of Political Economy* 105(2), 385-411.

Moffitt, Robert, and Mark Wilhelm. 2000. "Taxation and the Labor Supply Decisions of the Affluent." In *Does Atlas Shrug? The Economic Consequences of Taxing the Rich*, ed. Joel Slemrod, 193-234. New York: Russell Sage Foundation and Harvard University Press.

Moriguchi, Chiaki, and Emmanuel Saez. 2008. "The Evolution of Income Concentration in Japan, 1885-2005: Evidence from Income Tax Statistics," *Review of Economics and Statistics* 90(4), 713-734.

Murphy, Kevin J. 1999. "Executive Compensation," in Orley Ashenfelter and David Card (eds.), *Handbook of Labor Economics,* Vol. 3b, Elsevier Science North Holland, Chapter 38, 2485-2563.

Murphy, Kevin J. 2011. "The Politics of Pay: A Legislative History of Executive Compensation", USC Marshall School of Business Working Paper FBE 01.11.

Narayanan, M.P., and Seyhun H. Nejat. 2008. "The Dating Game: Do Managers Designate Option Grant Dates to Increase their Compensation?," *Review of Financial Studies* 21(5), 1907-1945.

OECD Statistics. http://stats.oecd.org

OECD. (annual). *Taxing wages*, OECD, Paris.

OECD. (annual). *Consumption Tax Trends*, OECD, Paris.

OECD. (1986). *Personal income tax systems*, OECD, Paris.

Piketty, Thomas. 2001. *Les Hauts Revenus en France au 20ème sicle,* Paris: Grasset.

Piketty, Thomas, and Emmanuel Saez. 2003. "Income Inequality in the United States, 1913-1998," *Quarterly Journal of Economics* 118(1), 1-39, series updated to 2008 in July 2010.

Piketty, Thomas, and Emmanuel Saez. 2011. "A Theory of Optimal Capital Taxation," PSE Working Paper.

Pissarides, Christopher A. 2000. *Equilibrium Unemployment Theory,* 2nd ed., Cambridge, MA: MIT Press.

Reynolds, Alan. 2007. "Has U.S. Income Inequality Really Increased?" *Policy Analysis* 586, 1–24.

Roine, Jesper, Jonas Vlachos, Daniel Waldenstrom. 2009. "The Long-Run Determinants of Inequality: What Can We Learn from Top Income Data?" *Journal of Public Economics*, 93(7-8): 974-988.

Rosen, Sherwin. 1981. "The Economics of Superstars," *American Economic Review* 71(5), 845-858.

Rothschild, Casey, and Florian Scheuer. 2011. "Optimal Taxation with Rent-Seeking" NBER working paper No. 17035.

Saez, Emmanuel. 2001. "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies* 68, 205-229.

Saez, Emmanuel. 2004. "Reported Incomes and Marginal Tax Rates, 1960-2000: Evidence and Policy Implications." in James Poterba, ed., *Tax Policy and the Economy*, 18: 117-174.

Saez, Emmanuel, Joel Slemrod, and Seth Giertz. 2011. "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review," forthcoming *Journal of Economic Literature.*

Saez, Emmanuel, and Michael Veall. 2005. "The Evolution of High Incomes in Northern America: Lessons from Canadian Evidence," *American Economic Review,* 95(3), 831-849.

Saez, Emmanuel, and Michael Veall. 2007. "The Evolution of High Incomes in Canada, 1920-2000", Chapter 6 in A. B. Atkinson and T. Piketty, eds. *Top Incomes over the Twentieth Century: A Contrast Between European and English Speaking Countries*, Oxford: Oxford University Press, 226-308

Slemrod, Joel. 1996. "High Income Families and the Tax Changes of the 1980s: The Anatomy of Behavioral Response." In *Empirical Foundations of Household Taxation*, eds. Martin Feldstein and James Poterba, University of Chicago.

Slemrod, Joel, and Wojciech Kopczuk. 2002. "The Optimal Elasticity of Taxable Income." *Journal of Public Economics*, 84(1): 91-112.

Slemrod, Joel, and Shlomo Yitzhaki. 2002. "Tax Avoidance, Evasion and Administration." in Alan Auerbach and Martin Feldstein, ed., *Handbook of Public Economics*, v3,

North-Holland: 1423-1470.

Sorensen, Peter B. 1999. "Optimal Tax Progressivity in Imperfect Labour Markets," *Labour Economics* 6, 435-452.

Stantcheva, Stefanie. 2011. "Optimal Taxation with Adverse Selection in the Labor market," MIT Working Paper.

U.S. Treasury Department, Internal Revenue Service. 2009. "Statistics of Income: Individual Statistical Tables by Tax Rate and Income Percentile," Table 1 available online at http://www.irs.gov/taxstats/indtaxstats/article/0,,id=133521,00.html

Vickrey, William. 1947. *Agenda for Progressive Taxation.* Ronald Press Company: New York.

Yermack, David. 1997. "Good Timing: CEO Stock Option Awards and Company News Announcements,"*Journal of Finance* 52(2), 449-476.

Yermack, David. 2006. "Flights of Fancy: Corporate Jets, CEO Perquisites, and Inferior Shareholder Returns," *Journal of Financial Economics* 80, 211-242.

Zucman, Gabriel. 2011. "The Missing Wealth of Nations: Are Europe and the U.S. net Debtors or net Creditors?", Paris School of Economics Working Paper
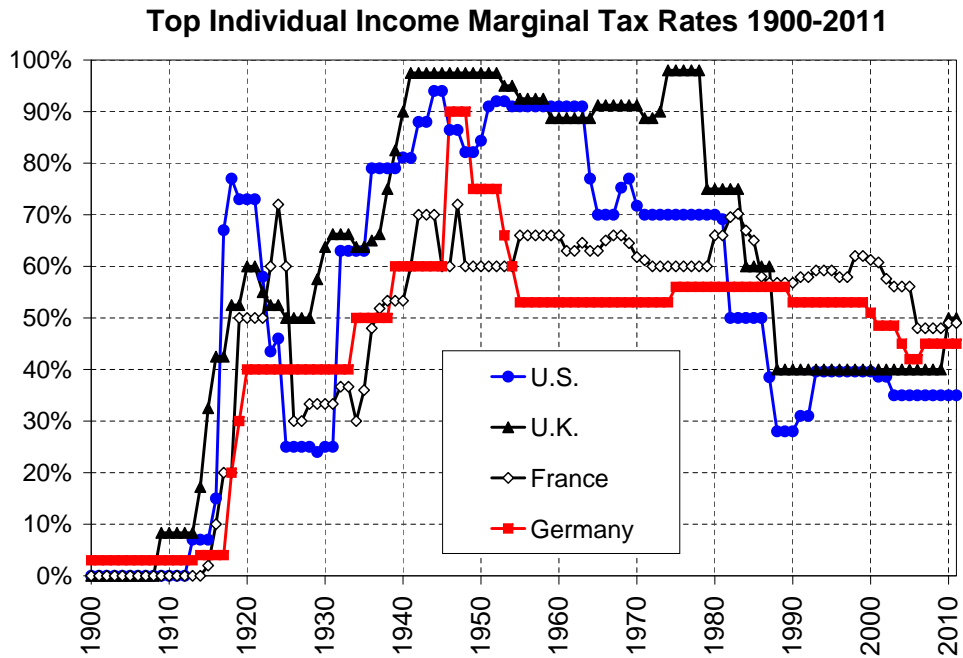
Figure 1: Top Marginal Income Tax Rates in the US, UK, France, Germany

This figure depicts the top marginal individual income tax rate in the US, UK, France, Germany since 1900. The tax rate includes only the top statutory individual income tax rate applying to ordinary income with no tax preference. State income taxes are not included in the case of the United States. For France, we include both the progressive individual income tax and the flat rate tax "Contribution Sociale Généralisée".
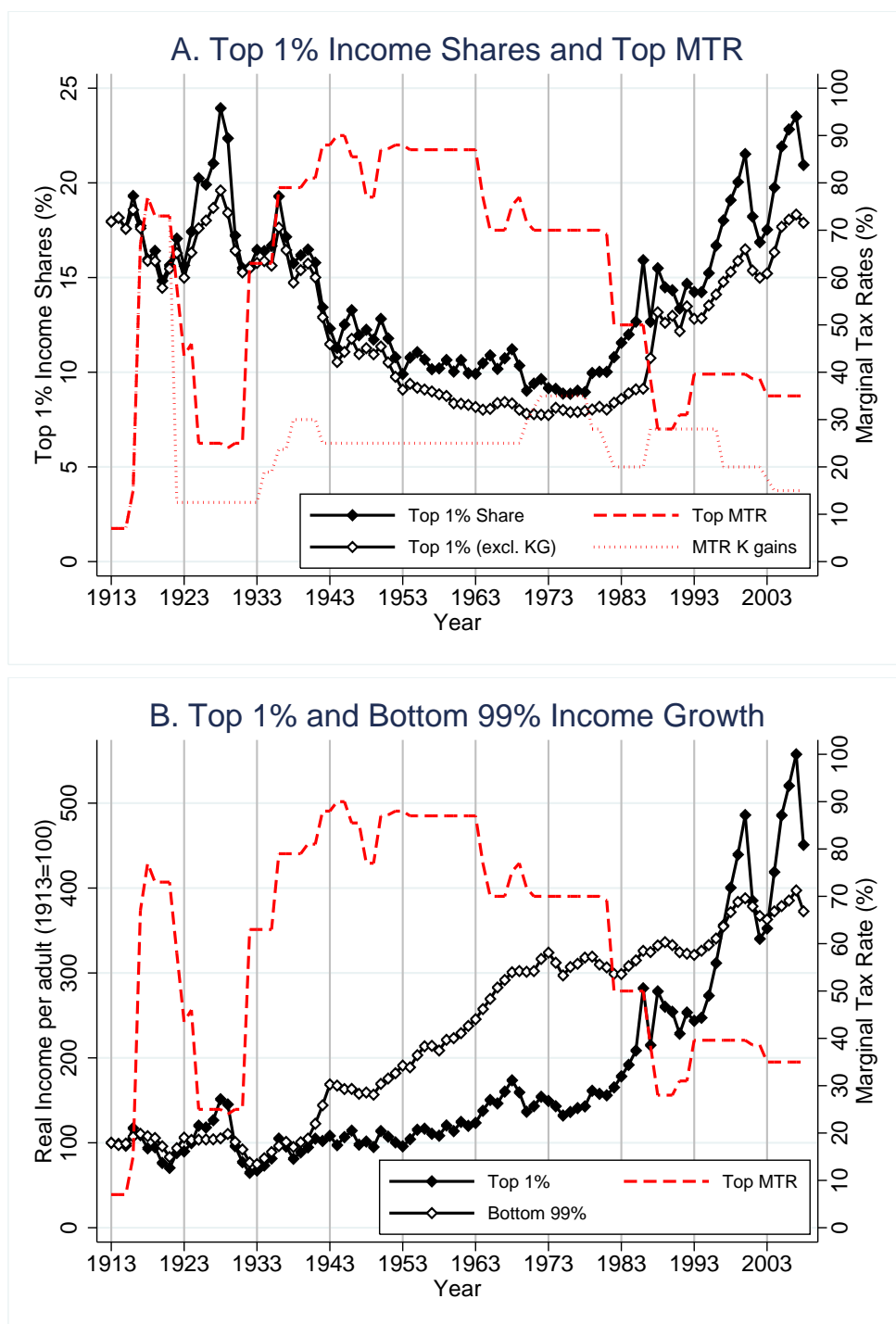
Figure 2: Top Marginal Tax Rates, Top Incomes Shares, and Income Growth: US Evidence 1913-2008

Panel A depicts the top 1% income shares including realized capital gains in full diamonds and excluding realized capital gains in empty diamonds. Computations are based on family market cash income. Income excludes government transfers and is before individual taxes (source is Piketty and Saez, 2003, series updated to 2008). Panel A also depicts the top marginal tax rate on ordinary income and on realized long-term capital gains (source is Tax Policy Center). Panel B depicts real cash market income growth per adult of top 1% incomes and bottom 99% incomes (base 100 in 1913), assuming that individual adult top 1% and bottom 99% shares are the same as top 1% and bottom 99% family based shares.
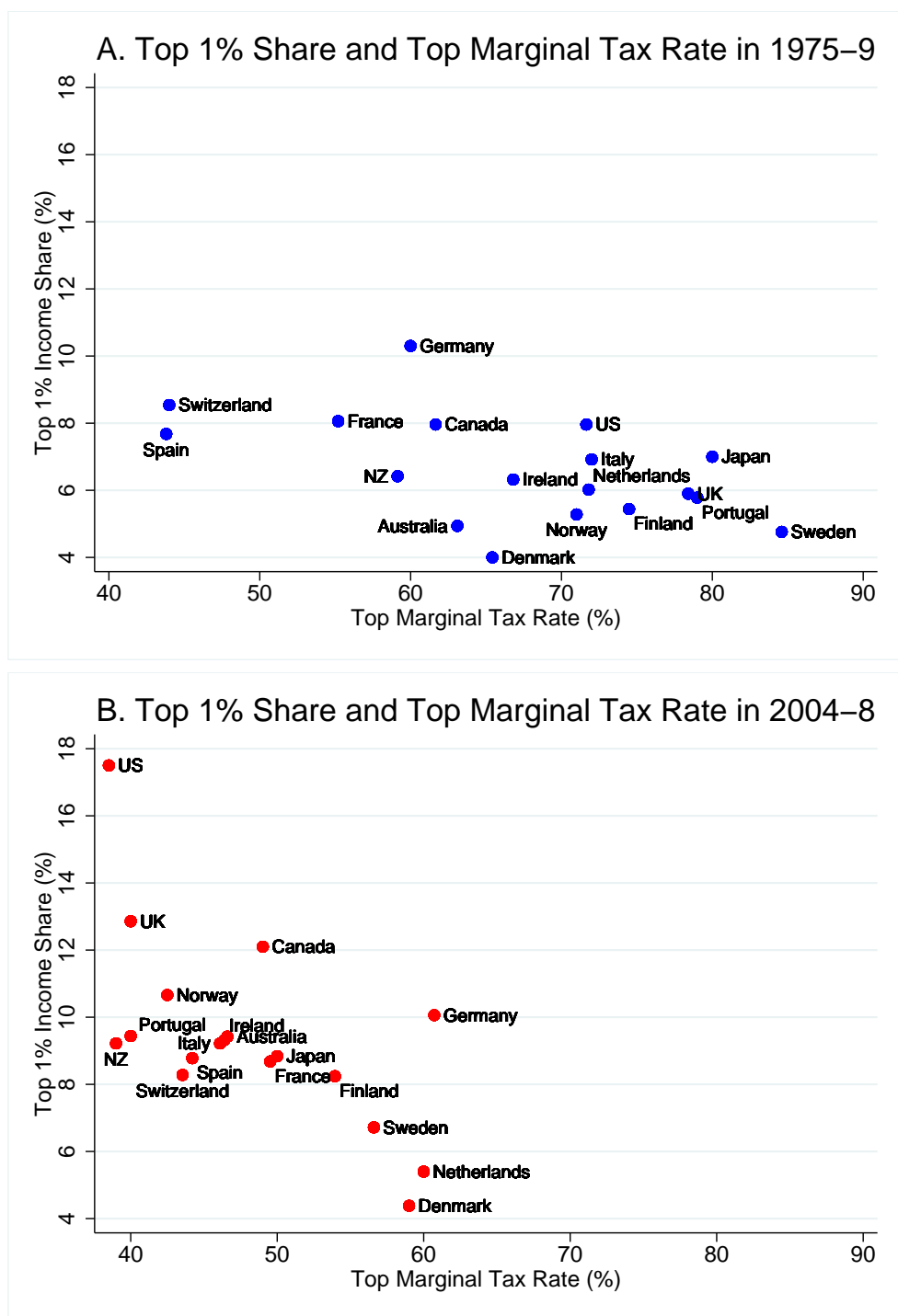
Figure 3: Top Income Shares and Top Marginal Tax Rates: International Evidence 1975-9 and 2004-8

The figure depicts the top 1% income shares and top income tax rates (including both central and local government individual income taxes) across 18 OECD countries in 1975-9 (Panel A) and 2004-8 (Panel B). Source for top income shares is the *World Top Incomes Database*. Source for top income tax rates is OECD. If the country does not have data for those years, we select the first five years after 1975 available and the most recent 5 years (full details in appendix C). The correlation between top tax rates and top income shares is much stronger in 2004-8 than in 1975-9 (see Table 2 for regression estimates).
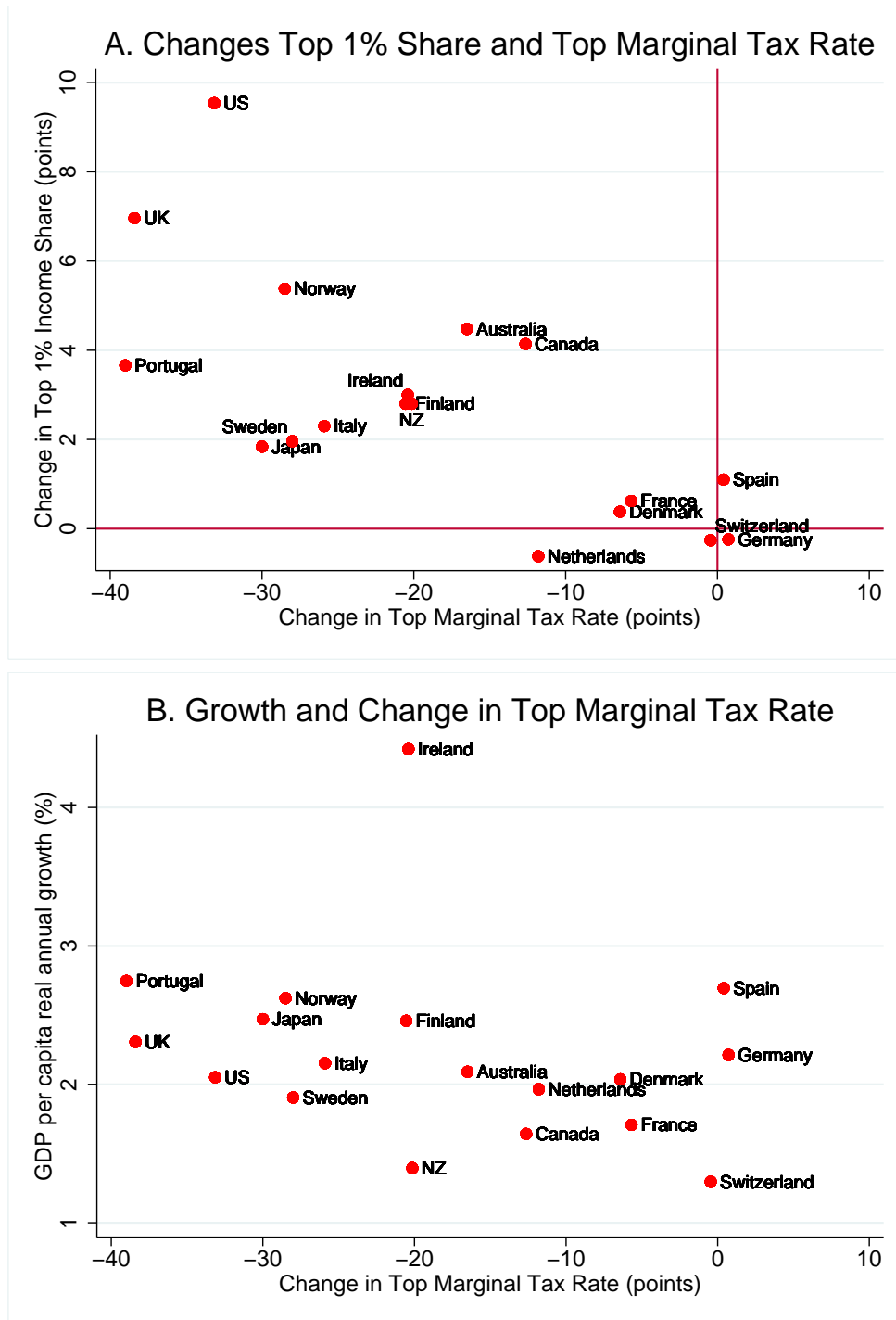
Figure 4: Top Income Shares, Top Marginal Tax Rates, and Growth

Panel A depicts the change in top income shares against the change in top income tax rate from 1975-9 to 2004-8 based on Figure 2 data for 18 OECD countries. The correlation between those changes is very strong (see Table 2 for regression estimates). Panel B depicts the change in real GDP per capita annual growth rate from 1975-9 to 2004-8 against the change in top marginal tax rate. The correlation is virtually zero and insignificant (but imprecise) suggesting that cuts in top tax rates do not lead to higher economic growth. Table 2 reports more precise estimates based on the complete time series.

Table 1: US Evidence on Top Tax Rates, Top Income Shares, and Income Growth

		Income excluding capital gains	Income including capital gains (to control for tax avoidance)
		(1)	(2)
A. 1975-1979 vs. 2004-2008 Comparison			
Top Marginal Tax Rate (MTR)	1975-9	70%	70%
	2004-8	35%	35%
Top 1% Income Share	1975-9	8.0%	9.1%
	2004-8	17.7%	21.8%
Elasticity estimate:			
Δ log (top 1% share) / Δ log (1-Top MTR)		1.03	1.12

B. Elasticity estimation (1913-2008): log(share) = a + e*log(1-Top MTR) + c*time + ε

	Income excluding capital gains	Income including capital gains
No time trend	0.25	0.26
	(0.07)	(0.06)
Linear time trend	0.30	0.29
	(0.06)	(0.05)
Number of observations	96	96

C. Effect of Top MTR on income growth (1913-2008): log(income) = a + b*log(1-Top MTR) + c*time + ε

	Income excluding capital gains	Income including capital gains
Top 1% real income	0.265	0.261
	(0.047)	(0.041)
Bottom 99% real income	-0.080	-0.076
	(0.040)	(0.039)
Average real income	-0.027	-0.027
	(0.018)	(0.034)
Number of observations	96	96

Estimates from Panel A are obtained using series from Figure 2 (source is Piketty and Saez, 2003 for top income shares and Tax Policy Center for top marginal tax rate). If the surge in top income shares is explained solely by the reduction in the top marginal tax rate, then the elasticity is very large around one. The elasticity is the same for income excluding capital gains and income including capital gains. As capital gains are treated more favorably and are the main channel of avoidance for top incomes, this implies that tax avoidance plays no role in the surge of top incomes in the long-run.

Estimates from Panels B and C are obtained by time-series regressions over the period 1913-2008 (96 observations) and using standard errors from Newey-West with 8 lags. Panel B shows significant elasticities of top 1% income shares with respect to the net-of-tax rate (using the top MTR). Elasticities are virtually the same when excluding or including capital gains and are robust to including a linear time trend in the regression. This shows that there is a strong link in the time-series between top income shares and top MTR as evidenced in Figure 2A.

Panel C shows that real income growth of top 1% is strongly related to the net-of-tax rate (using the top MTR), confirming the results of Panel B. Bottom 99% incomes are negatively related to the net-of-tax rate (using the top MTR) suggesting that top 1% income gains came at the expense of bottom 99% earners. Average incomes (including both the top 1% and bottom 99%) are not significantly related to the net-of-tax rate. Those results suggest that most of the elasticity of top incomes is due to bargaining effects and not real supply side effects.

Table 2: International Evidence on Top Tax Rates, Top Income Shares, and Income Growth

A. Effect of the Top Marginal Income Tax Rate on Top 1% Income Share

A1. Cross Country Cross-Sectional Comparisons:
Regression: log(Top 1% share) = a + e*log(1-Top MTR) + ε

Elasticity in 1975-9	0.329
	(0.148)
Elasticity in 2004-8	1.396
	(0.381)
Number of obs.	18

A2. Cross Country Changes from 1975-9 to 2004-8:
Regression: Δ log(Top 1% share) = a + e*Δ log(1-Top MTR) + ε

Elasticity	0.490
	(0.144)
Number of observations	18

A3. Full Time Series analysis (1975-2008):
Regression: log(Top 1% share) = a + e*log(1-Top MTR) + ε

No controls	0.561
	(0.034)
Time trend control	0.512
	(0.039)
Country fixed effects	0.455
	(0.029)
Number of observations	518

B. Effect of the Top Marginal Income Tax Rate on real GDP per capita
Regression: log(real GDP per capita) = a + b*log(1-Top MTR) + c*time + ε

No country fixed effects	0.027
	(0.036)
Country fixed effects	0.012
	(0.013)
Number of observations	518

Panel A presents estimates the elasticity of top 1% incomes with respect to the net-of-tax top marginal tax rate. Panel A1 presents regression estimates using cross-country cross sectional data from Figure 3A (1975-9) and Figure 3B (2004-8). The regressions are based on 18 OECD countries. Panel A2 presents a regression estimate using the change from 1975-9 to 2004-8 for those 18 OECD countries. Panel A3 presents regression estimates using complete panel from 1975 to 2008. Estimates are not

Panel B presents regressions of the log real GDP per capita on the log net-of-tax rate. The regressions include a time trend. The second regression include country fixed effects to improve precision. The effect of the top MTR on GDP per capita growth is small and insignificant. As discussed in the text, the estimate b is equal to $\pi*e_1$ where π is the top 1% income share (around 10%) and e_1 is the real supply side elasticity of top incomes. Hence those estimates imply that e_1 is not very large.

Table 3: Synthesis of Various Scenarios

Total elasticity $e = e_1 + e_2 + e_3 =$	0.5

Scenario 1: Standard supply side tax effects	Scenario 2: Tax avoidance effects		Scenario 3: Compensation bargaining effects
	(a) current narrow tax base	(b) after base broadening	
$e_1 =$ 0.5	$e_1 = 0.2$	$e_1 = 0.2$	$e_1 =$ 0.2
$e_2 =$ 0.0	$e_2 = 0.3$	$e_2 = 0.1$	$e_2 =$ 0.0
$e_3 =$ 0.0	$e_3 = 0.0$	$e_3 = 0.0$	$e_3 =$ 0.3

Optimal top tax rate $\tau^* = (1 + tae_2 + ae_3)/(1 + ae)$

Pareto coefficient a =	1.5
Alternative tax rate t =	20%

Scenario 1	Scenario 2		Scenario 3
	(a) $e_2 = 0.3$	(b) $e_2 = 0.1$	
$\tau^* =$ 57%	$\tau^* = 62\ \%$	$\tau^* = 71\ \%$	$\tau^* =$ 83%

This table presents optimal top tax rates in the case where the overall elasticity of reported taxable income is e=0.5 in three scenarios depending on how this total elasticity breaks down into the standard labor supply elasticity (e_1), the tax avoidance elasticity (e_2), the compensation bargaining elasticity (e_3). In scenario 1, the only elasticity is e_1. In scenario 2, both e_1 and e_2 are present, income shifted away from the regular tax is assumed to be taxed at rate t=20%. Scenario 2(a) considers the case of the current narrow base with avoidance opportunities and Scenario 2(b) considers the case where the base is first broadened so that e_2 falls to 0.1 (end hence the total elasticity e falls to 0.3). In scenario 3, both e_1 and e_3 are present. In all cases, top tax rates are set to maximize tax revenue raised from top bracket earners.