Demand for “The 1%”: Tax Incidence and Implications for Optimal Income Tax Rates

Richard K. Green & Mark D. Phillips *

Sol Price School of Public Policy, University of Southern California,

Ralph and Goldy Lewis Hall, Los Angeles, CA 90089-0626

Draft: October 31, 2014

Abstract

We develop a model for determining the optimal high income linear tax rate when there exist imperfectly substitutable types of labor. If one type is disproportionately prevalent among higher income taxpayers, then wages adjust in response to more progressive taxation and part of the statutory tax burden is shifted to lower income taxpayers. In principle these wage adjustments could be negated via revenue-neutral differential factor taxation, but we contend that such a policy bears little practical relevance. We estimate the optimal top tax rate under various plausible parameterizations. We reject the notion from the previous literature that wage adjustments are costly enough (from a social welfare perspective) to warrant non-progressive taxation. However, we also estimate that the optimal tax rate may be significantly smaller than when incidence effects are ignored, and may in fact be quite similar to that under current U.S. policy.

*Phillips (corresponding author): `mdphilli@price.usc.edu`; Green: `richarkg@price.usc.edu`.

1 Introduction

Public economics teaches that demand and supply share the burden of any tax. At the same time, the question of socially optimal marginal income tax rates is currently generating substantial interest. We therefore focus in this paper on determining the optimal marginal tax rate for the top 1% of income earners after taking into account incidence. In particular, we present a model in which high tax rates lead to endogenous changes in relative wage rates that disproportionately benefit the 1%. In essence, the incidence of high tax rates on the 1% is partially borne by the remaining 99% of the population. This pecuniary externality, an indirect consequence of taxing high income at a high marginal rate, is socially undesirable (it that it has pernicious effects on both efficiency and equity) and necessarily lowers the optimal top tax rate.

Our optimal tax conditions are derived in a fashion similar to Saez (2001), where the sufficient parameters for optimal tax calculations are readily interpretable. We generalize the Saez (2001) model to allow for imperfect substitution between different types of labor, as well as the possibility that some types of labor are proportionally more prevalent among the 1%. As is the case in Saez (2001), high income taxpayers' labor supply elasticity, their relative social welfare weight, and the shape of the upper tail of the income distribution are critical to the determination of the optimal tax rate. However, our generalization also indicates the relevance of additional parameters, in particular those that impact the extent to which high income taxpayers' wages react to higher tax rates. Wages are more responsive when high income taxpayers are less substitutable, but our framework indicates the relevance of two different concepts of substitutability. The first addresses whether there exist different types of labor that contribute to aggregate production in an imperfectly substitutable manner. The second addresses whether some of these types are disproportionately represented among high income earners.

Solving for the optimal tax rate in terms of readily interpretable parameters enables us to better understand those circumstances in which endogenous wage effects can be ignored.

First, price effects do not occur if all labor is perfectly substitutable, or alternatively, if different types of labor are equally prevalent among high income taxpayers and the population at large. Our review of the available evidence suggests that neither is true. Second, price effects are negated if the government simultaneously institutes the Mirrleesian second-best tax policy along with appropriately set differential factor taxes on different labor inputs. As we discuss later, such counteracting factor taxes do not satisfy the Saez and Diamond (2001) criteria for "policy-relevant" tax analysis. We therefore argue that it is a worthwhile excise to consider the optimal tax rate in a third-best world without government access to these offsetting factor tax instruments.

Another virtue of expressing optimal tax formulas in terms of readily interpretable parameters is that it helps us understand the underlying reasons for previous results from the (small) literature on how incidence considerations affect optimal income tax policy. For instance, Stern (1982) and Jacobs (2012) conclude that high income earners would optimally face a marginal subsidy (i.e. a negative marginal tax rate) in the presence of price effects. However, we show that the necessity of this result relies upon simplifying but restrictive assumptions regarding the distribution of incomes and does not generalize. To our knowledge, ours is the first paper to analyze the implications of general equilibrium price changes in a more realistic environment, with heterogeneous taxpayers that vary with regard to both labor type and skill levels. Our results are substantively different from those that are attained when: (a) skill levels vary but taxpayers supply identical types of labor; or (b) taxpayers supply different types of labor but posses identical (within-type) skill levels.

Our final exercise is to estimate the optimal tax rate under alternative, plausible parameterizations. In our simulations, the incidence of high income taxpayers' marginal tax rates is mostly, but not entirely, borne by high income taxpayers. However, even some tax shifting can significantly lower the optimal tax rate from that which would be obtained in the absence of price effects. On the other hand, we also reject the notion that general equilibrium wage effects are so socially detrimental that high income taxpayers should be subsidized.

Moreover, our numerical estimates uniformly support the notion that marginal tax rates are optimally progressive, though less so than what is suggested by theories that ignore the incidence of income taxation. In fact, our numerical estimates suggest that optimal high income marginal tax rates may be quite close to those under current US policy.

The rest of the paper is organized as follows. In Section 2 we describe the relevant existing literature. In Section 3 we discuss the effects of marginal tax rates on government revenues and labor suppliers' welfare when wages endogenously adjust. In Section 4 we critically analyze and reject the plausibility of those circumstances when incidence considerations are irrelevant to the determination of optimal high income tax rates. In Section 5 we derive the formula for the optimal high income tax rate when accounting for tax incidence. We use the formula to estimate the optimal high income tax rate under various feasible parameterizations in Section 6. We discuss our results and suggest future research in Section 7.

2 Previous Literature

The classical approach to optimal income taxation assumes that individuals differ along a single dimension, their skill level (e.g. Mirrlees (1971), Seade (1977), Atkinson and Stiglitz (1980)). The distribution of skills is considered exogenous and efficiency units of labor supply are perfectly substitutable in the production of aggregate output. The assumption greatly simplifies the analysis since taxes do not affect the price of an efficiency unit of labor supply and the economic incidence of income taxation equals the statutory incidence.

A later literature sought to derive the Mirrleesian optimal tax formulation in terms of more readily observable and interpretable elasticity and income distribution parameters. This contrasts with the classical formulation's reliance on derivatives of the (unobservable) utility function and parameters of the (unobservable) skill distribution (e.g. Dixit and Sandmo (1977), Diamond (1998)). Saez (2001) takes advantage of properties of the labor supplier's indirect utility function, an approach that we also use. Saez does not require that individuals vary along the single skill dimension; however, he does assume that "each taxpayer maximizes a well-behaved individual utility function...which depends positively on consumption...and negatively on earnings. Individual skills or ability are embodied in the individual utility function." (pg. 208) Under this framework, a change in the wage rate would change the individual's utility function itself.

Implicitly assuming that price effects do not occur, Saez (2001) derives both the optimal linear rate on high income earners as well as the more general nonlinear optimal income tax schedule. Focusing on his results regarding the former, he demonstrates the importance of three factors: first, the relative social welfare weight of high income earners; second, the elasticity of high income earners' taxable income with respect to after-tax rates; and third, a parameter that captures the thickness of the upper tail of the income distribution. The first depends on normative judgments, the latter two are measurable. Using this methodology Diamond and Saez (2011) estimate that the optimal top rate in the US is 73%, assuming that high earners' taxable income elasticity is 0.25 and that their social welfare weight is 0. Raising the relative social welfare weight to 0.04 only reduces the optimal tax rate by approximately one percentage point. Their estimates also employ a Pareto parameter (expected earnings of those with income above the relevant income threshold divided by the difference between this conditional expectation and the threshold itself) of 1.5 for the upper tail of the income distribution.[1]

Few papers have considered the optimal design of tax policy in an environment where incidence matters, likely for two reasons. First, the classical framework effectively assumes away price effects by assuming perfect substitution among efficiency units. Second, wage effects do not occur if the government has access to (and fully and properly utilizes) differential factor taxation. As explained in Saez, Slemrod, and Giertz (2012), price effects "are transfers [and] in principle the government can readjust tax rates on each factor to

[1] That the upper tail is characterized by a relatively thick-tailed Pareto distribution is documented in Saez (2001), Atkinson, Piketty, and Saez (2011), and Diamond and Saez (2011).

undo those incidence effects at no fiscal cost." (pg. 27)[2] In Section 4 we explain why we consider it a worthwhile and policy-relevant exercise to consider the optimal tax problem in a "third-best" world where the government does not have access to, or alternatively cannot implement, differential factor taxation.

Feldstein (1973) initiated the small literature on optimal income taxation when price effects occur. He analyzes the optimal linear income tax when there are two types of labor factors (characterized as "skilled" and "not skilled") that contribute to aggregate output via a Cobb-Douglas production function. While the theory generalizes from the standard Mirrleesian treatment by introducing two different factors of production, the article simultaneously assumes that all individuals of a given factor type are identical. That is, each worker of a given type has the same endowment of "skill" as every other worker of his or her type. This assumption limits the extent of heterogeneity in the population, especially with regard to the distribution of incomes. The assumption is maintained in subsequent articles that build off Feldstein's model.

Using numerical methods, Feldstein (1973) concludes that the optimal linear tax rate in the presence of price adjustments is quite similar to that in the case of fixed prices. Allen (1982) shows that Feldstein's result hinges on the assumption of Cobb-Douglas production. Using more general constant elasticity of substitution production functions, Allen (1982) contends that the optimal linear tax rate may be quite different in the case of price adjustments, and even has the potential to be negative (i.e. has the potential to argue for a wage subsidy). Carruth (1982) uses numerical methods and finds that Allen's proposed wage subsidy is actually attained when the elasticity of substitution between factors is sufficiently small, though plausibly so.

Stern (1982) analyzes the optimal nonlinear income tax in the same environment with two factors of production but within-factor homogeneity. The shift from linear to nonlinear taxation allows for individuals of each type to be taxed at a different marginal tax rate. In

[2] Diamond and Mirrlees (1971) formally demonstrates the result.

this case, Stern demonstrates analytically that the factor with higher incomes should necessarily receive a wage subsidy. Using numerical methods, Stern also estimates these optimal subsidies, which tend to be small and vary minimally across different parameterizations. Jacobs (2012) reaches a similar conclusion when considering both optimal tax and education policy; however, the model is similarly restrictive with regards to within-type homogeneity.

Christiansen (1988) considers the optimal linear tax rate in a similar though substantively different environment. Aggregate output is still determined by two imperfectly substitutable labor factors; however, taxpayers can choose which type of labor to supply. On the other hand, taxpayers are necessarily assumed to work - there is no choice being made between how much labor to supply (of either factor type) and how much leisure to consume. Christiansen (1988) shows that under these conditions, "a welfare maximum when tax-induced price changes are neglected at the margin is also a true welfare maximum" (pg. 154), or in other words, that price effects can be ignored in the process of choosing the optimal linear rate.

The current paper complements the more recent Rothstein (2010) article which compares the social desirability of EITC versus a traditional Negative Income Tax when incidence considerations are taken into account. EITC promotes labor supply among low income individuals which in turn reduces their relative wages; therefore, the positive distributional effects of an EITC program are at least partially offset by equilibrium price changes. Rothstein finds that incidence considerations significantly impact the optimal design of tax policy with respect to low-income individuals. The current analysis mirrors Rothstein's, focusing on tax policy at the upper rather than lower end of the income distribution. Rothstein also acknowledges that he "ignore[s] taxes that would be need to finance the proposed EITC and NIT programs. These would presumably be levied on higher income taxpayers, though their incidence, too, is unclear." (pg. 180) Our analysis addresses precisely this issue.

The evidence on whether income earners bear the full burden of income taxes is limited. Fullerton and Metcalf (2002) note that "for the personal income tax, applied studies have consistently assumed that economic incidence is the same as statutory incidence - on the taxpayer - even though this assumption has never been tested." (pg. 1822) This lack of evidence is attributable, in part at least, to the fact that"such efforts are extremely difficult to convincingly estimate empirically." (Saez, Slemrod, and Giertz (2012), pg. 14). Examples of such efforts include Bingley and Lanot (2002) and Kubik (2004), the former using a structural empirical method on Danish data, the latter a difference-in-difference approach on US data. Both analyses suggest at least partial shifting of the incidence of income taxes. Kubik compares wages before and after the Tax Reform Act of 1986 (the first difference) and across occupations that experienced differential reductions in their median marginal tax rate (the second difference). He concludes that "when assessing the distributional consequences of tax reforms, the wage effects of the tax changes should be considered. The supply shifts in the labor market caused by 1986 tax reform did affect wages." (pg. 1585)[3]

3 Effects of a Marginal Increase in the High Income Tax Rate

We consider an economy in which two different types of labor, denoted H and L, contribute to aggregate production according to a production function with constant returns to scale. Individuals are exogenously assigned to one of these two types. Individuals also vary with regards to their exogenously determined skill level q. Individuals receive disutility from supplying labor l. As is generally the case with optimal tax analyses, l is not merely the hours of labor supply, but also captures things like the individual's effort. An individual's *efficiency units* of labor supply are given by ql. Finally, an individual's pre-tax income is given by $z = qlw_t$, where w_t is the wage rate (per efficiency unit) for labor type t. We maintain the assumption of two types of labor from previous analyses like Feldstein (1973),

[3]Part of the difficulty in estimating equilibrium price effects is that empiricists can observe changes in the hours of work but not changes in more ethereal components of labor supply like a worker's effort level. Therefore observed changes in hourly wages may reflect both changes in the equilibrium wage rate per efficiency unit of labor supply but also changes in the supplied effort per hour. See Blomquist and Selin (2010).

Allen (1982), and Stern (1982); however, we allow for within-type heterogeneity via the introduction of the skill parameter q. Our framework therefore allows for much greater (and realistic) heterogeneity in incomes. Our framework also allows for workers of both types to exist at each pre-tax income amount. Without loss of generality, we discuss these types of labor under the assumption that H-type labor is disproportionately prevalent among higher incomes; however, high income taxpayers need not supply type-H labor exclusively, and low income taxpayers need not supply type-L labor exclusively.

We now wish to derive the optimal linear tax rate on high levels of income above some threshold income $\bar{z}$. We focus our analysis on the case where $\bar{z}$ is the threshold income that defines being in the top 1% of the pre-tax income distribution; we therefore denote the marginal linear tax rate above this amount τ_1. We frame our analysis this way in order for our results to be comparable to those in Saez and Diamond (2011), but note that our methodology generalizes to other values of $\bar{z}$.

Our analytical strategy is similar to that in Saez (2001). We will consider the effects of a marginal increase in the high income tax rate, $d\tau_1$, on government revenues and labor suppliers' welfare. We will then assign social welfare weights to these effects.[4] At the optimal tax rate, the welfare-weighted sum of these effects equals 0.[5]

3.1 Effect on Government Revenues

For simplicity, we assume that income below $\bar{z}$ is taxed at the linear tax rate τ_{99}.[6] For our purposes, we take τ_{99} as given. In so doing, we remain agnostic with respect to whether τ_{99} is optimally determined. It may or may not be.[7] We also make the simplifying assumption,

[4] The assumption of constant returns to scale implies that competitive firms–the employers of both H and L types of labor–do not earn profit.

[5] We thus assume that the first-order condition is sufficient for characterizing the optimal top tax rate; however, we acknowledge that a sizeable literature has established the precariousness of this assumption.

[6] The qualitative results are not significantly affected by a nonlinear tax schedule below z. In that case τ_{99} should be interpreted as an appropriately weighted average of the marginal tax rate for individuals below $\bar{z}$.

[7] Most of the debate in the 2012 presidential election focused on whether and how to change top income tax rates. In the subsequent "fiscal cliff deal" between President Obama and Congress, only high income tax

common in the optimal tax literature, that incremental changes in τ_1 do not induce taxpayers to discretely shift from one part of the marginal tax schedule to the other. Finally, we also make the simplifying assumption that labor supply is not subject to income effects.[8] This implies that the labor supply for an individual with labor type t (where t equals H or L), income category i (where i equals 1 or 99 for taxpayers with income greater or less than $\bar{z}$, respectively), and skill level q only depends on $(1-\tau_i)qw_t$, the individual's after-tax return to labor supply.

A marginal increase in τ_1 will have different effects on the tax revenues collected from different taxpayers. First, it will have a direct effect of increasing the revenue from an $i=1$ taxpayer by $(z-\bar{z})d\tau_1$. The effect does not depend on whether the taxpayer is H or L. Second, it will have an indirect effect of changing each taxpayer's pre-tax income. These indirect effects will depend on both the taxpayer's income and labor type.

For $i=1$ taxpayers, let ϵ_1 be the elasticity of labor supply with respect to the after-tax return to labor supply. For simplicity, we assume that ϵ_1 is common across $i=1$ taxpayers.[9] Each taxpayer's change in the log of pre-tax income is given by

$$d\ln z = \epsilon_1 d\ln(1-\tau_1) + d\ln w_t + \epsilon_1 d\ln w_t. \tag{1}$$

The first addend on the right represents the traditional labor supply disincetive of higher tax rates. The second reflects that the change in τ_1 may change the equilibrium price of an efficiency unit of the taxpayer's labor type – holding constant the taxpayer's labor supply, this will increase (decrease) the taxpayer's pre-tax income if $d\ln w_t > (<)0$. The third addend reflects the wage change's effect on labor supply – pre-tax income increases (decreases) if the price change increases (decreases) labor supply, or alternatively, if $d\ln w_t > (<)0$.

For $i=99$ taxpayers, the change in τ_1 affects pre-tax income indirectly via the induced

rates changed. We interpret this as evidence that it is a policy-relevant exercise to determine the optimal top tax rate taking the rest of the tax schedule as given.

[8]Many other analyses (e.g. Diamond (1998) and Saez (2001)) consider a similar assumption.

[9]If the elasticity is not common across taxpayers, then ϵ_1 should be interpreted as an appropriately weighted average across $i=1$ taxpayers.

change in equilibrium wages. Defining ϵ_{99} to be the elasticity of labor supply among $i = 99$ taxpayers, and again assuming for simplicity that this value is constant across said taxpayers, each taxpayer's change in the log of pre-tax income is given by

$$d \ln z = d \ln w_t + \epsilon_{99} d \ln w_t. \tag{2}$$

Aggregating across all taxpayers, the change in government revenues from a marginal increase in τ_1 is

$$\begin{aligned} dG = {} & (Z_{H1} - \bar{z} N_{H1}) d\tau_1 + (Z_{L1} - \bar{z} N_{L1}) d\tau_1 \\ & + \tau_1 Z_{H1} \left(\epsilon_1 d \ln(1 - \tau_1) + d \ln w_H + \epsilon_1 d \ln w_H \right) \\ & + \tau_1 Z_{L1} \left(\epsilon_1 d \ln(1 - \tau_1) + d \ln w_L + \epsilon_1 d \ln w_L \right) \\ & + \tau_{99} Z_{H99} \left(d \ln w_H + \epsilon_{99} d \ln w_H \right) \\ & + \tau_{99} Z_{L99} \left(d \ln w_L + \epsilon_{99} d \ln w_L \right) \end{aligned} \tag{3}$$

where Z_{ti} is the aggregate income and N_{ti} is the number of taxpayers of labor type t and income category i.

3.2 Effect on Labor Suppliers' Welfare

Each taxpayer chooses his labor supply to maximize utility, where utility depends negatively on labor supply and positively on disposable (i.e. after-tax) income. Employing the envelope theorem to derive well-known properties of the indirect utility function,[10] an $i = 1$ taxpayer's

[10] An $i = 1$ individual with labor type t has the indirect utility function

$$V = \max_{c,l} u(W + qlw_t - \tau_{99} \bar{z} - \tau_1 (qlw_t - \bar{z}), l)$$

where $u(c, l)$ is the utility function, c is after-tax income, l is labor supply, and W is nonlabor income.

equivalent variation from a marginal increase in τ_1 is given by

$$-(z-\bar{z})d\tau_1 + (1-\tau_1)zd\ln w_t. \tag{4}$$

The direct effect of the tax increase is to reduce the taxpayer's welfare in proportion to the amount of income subject to the tax increase, i.e. $(z-\bar{z})$. However, the tax increase can indirectly direct benefit (harm) the taxpayer if he sees his labor type's wage rate increase (decrease). Importantly, this benefit (or harm) is proportional to the entirety of the taxpayer's income z, not just the income in excess of $\bar{z}$.

In contrast, an $i = 99$ taxpayer is only affected by the tax-induced change in wages. Again deriving properties of the indirect utility function,[11] her equivalent variation from a marginal increase in τ_1 is given by

$$(1-\tau_{99})zd\ln w_t. \tag{5}$$

Denoting dV_{ti} to be the aggregate welfare effect for each of the four t and i combinations, we therefore have

$$\begin{aligned}
dV_{H1} &= -(Z_{H1} - \bar{z}N_{H1})d\tau_1 + (1-\tau_1)Z_{H1}d\ln w_H \\
dV_{L1} &= -(Z_{L1} - \bar{z}N_{L1})d\tau_1 + (1-\tau_1)Z_{L1}d\ln w_L \\
dV_{H99} &= (1-\tau_{99})Z_{H99}d\ln w_H \\
dV_{L99} &= (1-\tau_{99})Z_{L99}d\ln w_L.
\end{aligned} \tag{6}$$

[11] An $i = 99$ individual with labor type t has the indirect utility function

$$V = \max_{c,l} u(W + qlw_t - \tau_{99}qlw_t, l).$$

3.3 Comparing the Effects on Tax Revenues and Labor Suppliers

Table 1 disaggregates the effect on government tax revenues into those components that represent transfers and those that represent efficiency effects. The expressions for dG in rows c-f take advantage of the constant returns to scale property that $(Z_{H1}+Z_{H99})d\ln w_H + (Z_{L1}+Z_{L99})d\ln w_L = 0$, i.e. that the weighted change in wage rates must be 0.[12] In the absence of tax-induced wage changes, the tax increase transfers value from $i=1$ taxpayers to the government (rows a and b) and discourages work among these same taxpayers (rows g and h). In that case, the optimal τ_1 is determined by a balancing of these socially beneficial transfers and socially harmful labor supply inefficiencies.

In the event that wages change in response to tax rates, a slew of other transfers and behavioral effects occur. Without loss of generality, suppose that $d\ln w_H > 0$ and $d\ln w_L < 0$. In that case, the w_H increase transfers value from the government back to $t=H$ individuals. Among $i=1$ taxpayers, this means that H taxpayers will regain (row c) at least some of the value that was lost from the tax increase's direct transfer effect. In contrast, L taxpayers with $i=1$ will have their direct losses from the tax increase (row b) further augmented by a tax-induced decrease in w_L. Unlike the case without price effects, $i=99$ taxpayers will also experience transfers due to the wage changes. Those with $t=H$ will benefit from the increase in w_H, with these benefits coming at the expense of the government's tax revenues (row e). The reverse holds for lower income taxpayers with $t=L$ (row f).

The wage changes will also affect the labor supply behavior of the four types of taxpayers. An increase in w_H will encourage labor supply among $t=H$ taxpayers, both those with $i=1$ (row i) and $i=99$ (row k). These promotions of labor supply help offset some of the tax's direct effects on disincentivizing labor supply. On the other hand, a decrease in w_L will disincentive labor supply among $t=L$, both those with $i=1$ (row j) and those with $i=99$ (row l).

[12] The virtue of incorporating this property into the expression for dG is that all transfers can be framed as occuring between groups of taxpayers and the government. Alternatively, the transfers could be framed as occuring between groups of taxpayers and the government, as well as among the taxpayers themselves.

In net, the introduction of equilibrium prices leads to many different effects, some of which appear to be socially beneficial and make a higher τ_1 appealing. In general we expect transfers from higher-than-average-income $i = 1$ taxpayers to the government to be socially desirable and transfers from lower-than-average-income $i = 99$ taxpayers to the government to be socially undesirable. Assuming that tax rates are positive, increases in labor supply are socially beneficial as the difference between the distorted and undistorted labor supply shrinks. Therefore the socially desirable effects arising from wage changes are: transfers from $i = 1$, $t = L$ taxpayers to the government (row d); transfers from the government to $i = 99$, $t = H$ taxpayers (row e); and increased labor supply from $t = H$ taxpayers, both those with $i = 1$ (row i) and $i = 99$ (row k). On the other hand, there are just as many indirect affects that appear to be socially undesirable and make a lower τ_1 appealing. These socially undesirable effects are: transfers from the government to $i = 1$, $t = H$ taxpayers (row c); transfers from $i = 99$, $t = L$ taxpayers to the government; and decreased labor supply from $t = L$ taxpayers, both those with $i = 1$ (row j) and $i = 99$ (row l).

Given these countervailing effects, it is difficult to assess from Table 1 whether or not the net effect of wage changes are socially beneficially. In order to make the net effect more salient it is useful to define the following two wage concepts. First, let $d \ln w_1$ be the income-weighted average change in log wages among those with $i = 1$:

$$d \ln w_1 = \frac{Z_{H1} d \ln w_H + Z_{L1} d \ln w_L}{Z_{H1} + Z_{L1}}. \tag{7}$$

Similarly, let $d \ln w_{99}$ be the income-weighted average change in log wages among those with $i = 99$:

$$d \ln w_{99} = \frac{Z_{H99} d \ln w_H + Z_{L99} d \ln w_L}{Z_{H99} + Z_{L99}}. \tag{8}$$

Using these definitions, the rows from Table 1 can be collapsed to show the total transfers between the government and $i = 1$ taxpayers (both those with $t = H$ and $t = L$), the total transfers between the government and $i = 99$ taxpayers, the total behavioral effects among

$i = 1$ taxpayers, and the total behavioral effects among $i = 99$ taxpayers. Defining Z_i to be the total income and N_i to be the number of taxpayers of income category i, Table 2 shows these partially aggregated effects.

Rows a and d of Table 2 show the effects of $d\tau_1$ in the absence of wage changes. As we will show in the next subsection, an increase in τ_1 leads to $d \ln w_1 \geq 0$ and $d \ln w_{99} \leq 0$. In fact, these inequalities strictly hold (i.e. $d \ln w_1 > 0$ and $d \ln w_{99} < 0$), except under very specific circumstances. Therefore row b shows that in net, the tax increases transfers from the government to $i = 1$ taxpayers, a socially undesirable outcome. Row c shows that the tax increases transfers from $i = 99$ taxpayers to the government, also a socially undesirable outcome. Skipping row e for the moment, row f shows that the wage change disincentives net labor supply among those with $i = 99$, another undesirable outcome assuming that $\tau_{99} > 0$.

The only socially desirable effect of endogenous wage changes is shown in row e. $d \ln w_1 > 0$ promotes the net labor supply among those with $i = 1$, and this promotion of labor increases economic efficiency so long as $\tau_1 > 0$.[13] While this socially desirable effect may appear to make the net social desirability of wage changes ambiguous, this is not the case. This is due to the fact that the desirability of the pro-efficiency effect in row e must be less than the social undesirability of the redistribution in row b. We will demonstrate this fact more formally in Section 5, but explain the intuition here.

The indirect effects of $d \ln w_1 > 0$ operate in an inverse manner to the direct effects of $d\tau_1$. The increase in τ_1 transfers from $i = 1$ taxpayers to the government in a socially beneficial manner, but discourages economic efficiency by disincentivizing their labor supply.[14] Importantly, the size of the socially beneficial transfer from each taxpayer is proportional to the amount of his income that is subject to the higher tax rate, i.e. only his income in excess of $\bar{z}$. In contrast, the increase in w_1 transfers from the government to $i = 1$ taxpayers

[13] In Section 5 we demonstrate that optimality of a negative τ_1 cannot be ruled out by theory alone. In that case, the net supply of labor among $i = 1$ exceeds the efficient amount in the amount of taxation, and the further promotion of labor supply arising from $d \ln w_1 > 0$ leads to an even greater, socially undesirable distortion.

[14] Again assuming $\tau_1 > 0$.

in a socially undesirable manner but encourages economic efficiency by incentivizing their labor supply. In this case, the socially undesirable transfer from the government to each taxpayer is proportional to the taxpayer's entire income. The ratio of indirect transfers to indirect efficiency effects is therefore larger (in magnitude) than the ratio of the tax increase's direct transfers to direct efficiency effects. This implies that in net the indirect effects of the induced increase in w_1 are necessarily undesirable whenever the net direct effects of a tax increase are socially desirable.

3.4 Effects on Wages

Having demonstrated that endogenous wage changes affect the social desirability of increases in the top marginal tax rate, we now wish to examine the determinants of said wage changes. Before proceeding, it will prove useful to define a few useful share concepts. s_1 and s_{99} are the shares of aggregate income going to $i = 1$ and $i = 99$ taxpayers, respectively, with $s_1 + s_{99} = 1$. s_H and s_L are the shares of aggregate income going to $t = H$ and $t = L$ taxpayers, respectively, with $s_H + s_L = 1$. The conditional share $s_{t|i}$ is the share of aggregate category-i income going to taxpayers with labor type t. By construction, $s_{H|1} + s_{L|1} = 1$ and $s_{H|99} + s_{L|99} = 1$. Finally, the conditional share $s_{i|t}$ is the share of aggregate type-t income going to taxpayers in income category i. By construction, $s_{1|H} + s_{99|H} = 1$ and $s_{1|L} + s_{99|L} = 1$. These conditional share concepts are related according to a Bayes Rule analogue – for example $s_{1|H} s_H = s_{H|1} s_1$.[15]

Assuming constant returns to scale, our first labor market equilibrium condition is that firms earn zero profits. Therefore changes in w_H and w_L must offset according to the well-known property

$$s_H d \ln w_H + s_L d \ln w_L = 0. \tag{9}$$

Our second labor market equilibrium condition is that the relative supply of H-to-L

[15] Using these share definitions, the $d \ln w_1$ and $d \ln w_{99}$ concepts in (7) and (8) can be rewritten as $d \ln w_1 = s_{H|1} d \ln w_H + s_{L|1} d \ln w_L$ and $d \ln w_{99} = s_{H|99} d \ln w_H + s_{L|99} d \ln w_L$.

efficiency units must equal the relative demand. We define Q_H^D to be the aggregate efficiency units of H demanded and Q_L^D the aggregate efficiency units of L demanded. Assuming that firms' production technologies exhibit a constant elasticity of substitution σ, then the marginal increase in τ_1 leads to a change in relative demand given by

$$d\ln(Q_H^D/Q_L^D) = -\sigma(d\ln w_H - d\ln w_L). \tag{10}$$

We define Q_H^S to be the aggregate efficiency units of H supplied and Q_L^S the aggregate efficiency units of L supplied. The direct effect of a marginal τ_1 increase is to discourage labor supply among $i = 1$ taxpayers regardless of labor type. However, wage changes will have an indirect effect on both $i = 1$ and $i = 99$ taxpayers depending on the type of the labor they supply. The total effect on type-H supply is given by

$$d\ln Q_H^S = \epsilon_1 s_{1|H}(d\ln(1-\tau_1) + d\ln w_H) + \epsilon_{99} s_{99|H} d\ln w_H \tag{11}$$

and the total effect on type-L supply is given by

$$d\ln Q_L^S = \epsilon_1 s_{1|L}(d\ln(1-\tau_1) + d\ln w_L) + \epsilon_{99} s_{99|L} d\ln w_L. \tag{12}$$

In equilibrium relative labor supply equals relative labor demand. This implies that $d\ln(Q_H^D/Q_L^D) = d\ln Q_H^S - d\ln Q_L^S$. This equilbrium identity, along with the expressions given in (10), (11), and (12) yields the following expressions for the changes in w_H and w_L that arise from changes in τ_1:

$$d\ln w_H = -\left(\frac{\epsilon_1}{\sigma+\tilde{\epsilon}}\right)\left(s_{1|H} - s_1\right) d\ln(1-\tau_1) \qquad \text{and} \tag{13}$$

$$d\ln w_L = -\left(\frac{\epsilon_1}{\sigma+\tilde{\epsilon}}\right)\left(s_{1|L} - s_1\right) d\ln(1-\tau_1) \tag{14}$$

where $\tilde{\epsilon}$ is a weighted average of ϵ_1 and ϵ_{99} defined by

$$\tilde{\epsilon} = \left(s_{1|H}s_L + s_{1|L}s_H\right)\epsilon_1 + \left(s_{99|H}s_L + s_{99|L}s_H\right)\epsilon_{99}.^{16} \tag{15}$$

(13) shows that an increase in τ_1 increases w_H if $s_{1|H} > s_1$, decreases w_H if $s_{1|H} < s_1$, and has no effect on w_H if and only if $s_{1|H} = s_1$. The reasoning is as follows. The direct effect of $d\tau_1$ on aggregate supply of H will be proportionally larger than the effect on aggregate supply of L if H-type income is more prevalent among high income earners. This raises the marginal product of H efficiency units relative to L efficiency units, thus raising the relative wage of H to L. The only exception is if $s_{1|H} = s_1$ and $s_{1|L} = s_1$ and therefore $d\ln w_H = d\ln w_L = 0$. In other words, w_H and w_L are unaffected by tax increases if and only if the types of labor supplied by high income earners are proportionally the same as the types of labor supplied by the population in its entirety.

We use these expressions to characterize how the average wage rate changes among $i = 1$ and $i = 99$ taxpayers. Plugging (13) and (14) into the $d\ln w_1$ and $d\ln w_{99}$ definitions in (7) and (8)

$$d\ln w_1 = -\left(\frac{\epsilon_1}{\sigma + \tilde{\epsilon}}\right)\rho^2 s_{99} d\ln(1-\tau_1) \qquad \text{and} \tag{16}$$

$$d\ln w_{99} = +\left(\frac{\epsilon_1}{\sigma + \tilde{\epsilon}}\right)\rho^2 s_1 d\ln(1-\tau_1) \tag{17}$$

where

$$\rho^2 = \frac{\left(s_{H|1}s_1 - s_H s_1\right)^2}{s_H s_L s_1 s_{99}}. \tag{18}$$

ρ^2 is readily interpretable as the coefficient of determination between whether a dollar of income belongs to a taxpayer in income category $i = 1$ or $i = 99$ and whether that same dollar of income arises from supply of $t = H$ or $t = L$ type labor. To clarify, consider the OLS regression where the explanatory variable is an indicator for whether a given dollar of

[16] $\tilde{\epsilon}$ is interpretable as an average due to the fact that $s_{1|H}s_L + s_{1|L}s_H + s_{99|H}s_L + s_{99|L}s_H = 1$. If ϵ_1 and ϵ_{99} are equal to a common ϵ, then $\tilde{\epsilon} = \epsilon$.

income is earned by a taxpayer in the $i = 1$ income category and the dependent variable is an indicator for whether that dollar arises from supply of $t = H$ type labor.[17] The R^2 from the regression would equal the coefficient of determination above. This makes clear that the relevant issue for determining ρ^2 is not whether $i = 1$ taxpayers tend to be specifically H or specifically L. The relevant issue is instead how well $i = 1$ explains the type of labor supply.[18]

ρ^2 is bound between 0 and 1. ρ^2 equals 0 if and only if $s_{H|1} = s_H$ and accordingly $s_{L|1} = s_L$. In other words, ρ^2 equals 0 if and only if the types of labor that generate income for high income taxpayers are proportionally the same as the types of labor that generate income for the population as a whole. ρ^2 equals 1 if and only if all taxpayers with $i = 1$ supply one type of labor and all taxpayers with $i = 99$ supply the other type of labor.

4 When Can Price Effects Be Ignored?

Saez and Diamond (2011) propose three criteria for assessing when theoretical optimal tax results are relevant to the practical design of tax policy. These conditions provide a constructive framework for assessing whether price effects can and should be ignored. The criteria are:

> "First, the result should be based on an economic mechanism that is empirically relevant and first order to the problem at hand. Second, the result should be reasonably robust to changes in the modeling assumptions... Therefore, we should view with suspicion results that depend critically on very strong homogeneity or rationality assumptions... Third, the tax policy prescription needs to be implementable – that is, the tax policy needs to be socially acceptable and not too

[17] The regression's units of observation are the dollars in the economy, not taxpayers. This illustrates that our analysis allows for the possibility that a single taxpayer may generally provide some combination of H and L efficiency units.

[18] Running OLS on, for example, an indicator for $t = L$ versus an indicator for $i = 1$ would yield a different regression coefficient and intercept, but the same R^2.

complex relative to the modeling of tax administration and individual responses to tax law." (pg. 166)

With these criteria in mind, we now state those conditions when it is appropriate to ignore equilibrium price effects in the determination of optimal income tax rates.

(Proposition 1). *Wage effects can be ignored if:*

a. $\sigma = \infty$*; or,*

b. $\rho^2 = 0$*; or,*

c. the government has access to and properly utilizes offsetting, revenue-neutral, differential factor taxation on the H *and* L *types of labor.*

Any one of these conditions is sufficient for the irrelevance of price effects. We discuss each of the three conditions in sequence, and find that none of them obtain.

The first condition for ignoring price effects, i.e. $\sigma = \infty$, fails the first and second policy-relevance criteria. $\sigma = \infty$ implies that H and L are perfectly substitutable in production. It is the implicit assumption in the Mirrleesian framework in which taxpayers differ only in regards to their skill level. One person may be particularly skilled, but his labor can be perfectly replaced by ten of his fellow workers who each have one-tenth the skill level. This assumes a great deal of homogeneity in terms of the talents and productive contributions of different individuals, and we indeed view this homogeneity with suspicion.

We have intentionally remained agnostic with regards to the precise definition of the H and L labor types. Feldstein (1973) characterizes his two labor types as "skilled" and "unskilled." There are many other ways that one could feasibly think of H and L, for example college-educated and high-school-educated, white collar and blue collar, or entrepreneurs and non-entrepreneurs. However one conceptualizes H and L, the empirical evidence in support of imperfect substitution among labor suppliers is robust. For instance, Katz and Murphy (1992), Card and Lemieux (2001), and Goldin and Katz (2007) are but a few of the many studies that find imperfect substitution between college- and high-school-educated workers.

We do not claim to have the "right" answer as to the value of σ, because it will necessarily depend on how H and L are defined.[19] Even limiting the economy to two types of labor supply seems overly simplistic, as expounded upon in Haskel et al (2012). With more than two types of labor there would be multiple relevant substitution parameters. Nonetheless, we feel confident that an assumption of $\sigma = \infty$ is implausible and empirically refutable.

The second condition for ignoring price effects, i.e. $\rho^2 = 0$, is also an assumption about the degree of heterogeneity (or lack thereof) in the economy. While $\sigma = \infty$ is implicitly a statement about the homogeneity of laborers' productive contributions, ρ^2 is instead a statement about the homogeneity of individuals across the income distribution. Recall that $\rho^2 = 0$ if and only if $s_{H|1} = s_H$ and $s_{L|1} = s_L$, or in other words, if the types of labor supplied by higher income earners are proportionally equivalent to the types of labor supplied by the population as whole. We also view this homogeneity with suspicion.

Using data from federal tax returns, Bakija, Cole, and Heim (2012) (hereafter "BCH") report on the occupations of the top 1% of US income earners. In order to assess whether the top 1% earn income in a proportionally equivalent manner to the rest of the population, we report some of their findings for TY2005 (the last year of their data) in Table 3. Table 3 lists the top 10 occupations from BCH's Table 6 (pg. 41). Columns (1) and (2) in our Table are inferred from the BCH data. Column (1) shows the share of the top 1%'s total income received by each occupation.[20] Column (2) provides the cumulative sum of these shares. Columns (3) and (4) in our Table provide the same occupation-specific information, but for the entire U.S. population. These numbers are calculated using the May 2005 Occupational Employment Statistics from the Bureau of Labor Statistics, along with BCH's occupation

[19] We do not believe that defining H to be equivalent to having a college degree is entirely appropriate for the current application. This would end up classifying such uniquely skilled (and definitively top 1%) earners as Jennifer Aniston, Michael Dell, Robert Downey, Jr., Larry Ellsion, Bill Gates, LeBron James, Ralph Lauren, Steve Jobs, Madison Bumgarner, J.K Rowling, Mark Zuckerberg as L.

[20] BCH reports the share of total income (not just the income of the top 1%) received by top 1% earners in each occupation. Column (1) in our Table is inferred by dividing this number by BCH's estimate of the 1%'s total share in 2005, 16.97%.

definitions (see BCH Appendix A.1, pgs. 49-50).[21]

Occupations are not the ideal way to define an individual's labor type; however, we believe that occupations are sufficiently good proxies for labor type for the purpose of assessing (and rejecting) the null hypothesis that high income taxpayers supply different types of labor in the same proportions as the entire work force. Executives, financial professionals, lawyers, doctors, and real estate professionals account for 37.4, 16.3, 7.2, and 10.9, and 3.4 percent, respectively, of the top 1%'s income. Cumulatively, these occupations account for over three-quarters of the top 1% income. The analagous shares for national income in its entirety are substantially different. The five occupations account for 15.5, 3.9, 1.2, 4.8, and 0.4 percent of national income, respectively, and only about one-quarter of national income cumulatively. The income share of each of these occupations is at least twice at large among the top 1% than among the entire population. If $\rho^2 = 0$ were true, we would expect these shares to be approximately equal.

The final condition for ignoring price effects does not depend on technical, and presumably estimable, parameters regarding the economy's production function, or the distribution of labor types. The third condition instead reflects an assumption on the types of policy instruments the government can effectively access and implement. As explained in Saez, Slemrod, and Giertz (2012), price/incidence effects "are transfers [and] in principle the government can readjust tax rates on each factor to undo those incidence effects at no fiscal cost." (pg. 27) We do not disagree with the truth of this statement, but emphasize that this is possible only in principle. With regards its policy relevance, we are more suspicious and contend that offsetting differential factor taxes fails the third of the Diamond and Saez (2011) criteria for policy-relevance. More specifically, we believe this policy would be neither "implementable" nor "socially acceptable," and furthermore would be "too complex relative to the modeling of tax administration and individual responses to tax law."

[21] We omit one of the BCH occupation categories, "entrepreneur not elsewhere classified," because we cannot obtain an estimate for this "occupation" from publicly available BLS data.

Our reasons for thinking this are numerous. First, it is helpful to describe how such offsetting taxes would work. In the absence of price effects, let's suppose that the optimal τ_1 equals 73%. In the presence of socially undesirable wage changes however, the optimal rate drops to, let's say, 53%. The mechanism underlying the change in optimal tax rates is as follows. The 73% tax rate would reduce the relative supply of H-to-L efficiency units, assuming $\rho^2 > 0$. Assuming that σ is finite the labor market would find itself out of equilibrium, with relative demand for H exceeding relative supply. In response to this relative shortage, the relative price (w_H/w_L) would have to rise. These price changes would in turn raise the equilibrium ratio of H-to-L labor supply above that which was optimally desired when setting $\tau_1 = 73\%$. In order to return the equilibrium ratio of H-to-L efficiency units to its desired level in the absence of price effects, the government would need to discourage firms from hiring H and encourage firms to hire L. In principle, this could be accomplished by taxing firms for employing H-type labor and subsidizing firms for employing L-type labor.

Now envision the legislator who proposes this policy. "I would like to raise marginal tax rates on high income individuals; however, I am concerned about the unintended consequences of this policy. Therefore, I propose that we raise taxes on these individuals and simultaneously implement a tax on those who choose to employ these individuals." Given the recent political discord surrounding President's Obama proposed return to Clinton-era marginal tax rates, the further suggestion that employers should also be taxed for employing high income individuals sounds like political suicide. After all, the voting public has not all taken a public economics course and may not find salient the idea that statutory and economic tax incidences are different.

While we consider the previous discussion a compelling reason to omit differential factor taxation from the government's toolkit, it is not what Diamond and Saez (2011) have in mind for their third policy-relevance criterion. "We do not mean to limit the choice to currently politically plausible policy options. Rather, we mean there should be very widely held normative views that make such policies seem implausible and inappropriate at pretty much all times." (pg. 166) There are several reasons to think that differential factor taxation would in fact "seem implausible and inappropriate at pretty much all times."

In order to administer factor taxes, the government would have to tax employers based on observable characteristics. Examples of such observable characteristics may include an individual's education level, the industry of his employer, or perhaps his occupation; however, neither education nor industry nor occupation would perfectly capture the nature of an employee's productive contributions. In the simple case of two labor types, some workers who are truly type H would have their labor subsidized as if they're L, while others who are truly type L may have their labor taxed. This would benefit the former's and hinder the latter's job market prospects, an outcome that would seem to violate notions of horizontal equity. The problem of errant labor-type classification would grow exponentially with the (realistic) introduction of more than two labor types. It would also grow upon recognizing that individual workers may supply some mix of labor types.

Even if legislators and tax administrators could properly define labor types and pick the appropriate taxes and subsidies on employers, the policy would have to rely on employers accurately reporting their factor tax liabilities. However, employees and employers would have a strong incentive to collude and report that the employee is the type that warrants a subsidy as opposed to a tax. It is unlikely that the IRS would have the necessary resources to deter such noncompliance. After all, the IRS could not feasibly place a revenue officer in each and every cubicle, classroom, stock room, and factory across the country. Furthermore, the perceived unfairness of the tax and subsidy program would likely lead to a low level of what the tax evasion literature has termed "tax morale," and therefore provide little non-pecuniary incentive for compliance.

Finally, it is easy to imagine that the policy would lead to a sizeable amount of lobbying and rent-seeking as employers tried to convince the government that their work force composition warranted subsidization. Such rent-seeking would involve the exhaustion of real economic resources, and further cast doubt on whether the factor taxes and subsidies had actually been set in an appropriate manner to undo the tax-induced wage changes.

5 Characterizing the Optimal High Income Tax Rate

We hope to have convinced our readers of the practical relevance of accounting for tax-induced wage changes. We now move on to characterizing the optimal high income tax rate in an environment of tax incidence considerations.

5.1 Effect of Marginal Tax Increases on Social Welfare

We define λ_1 to be the average marginal social utility of income among the top 1% of earners relative to that of the government. λ_{99} is the analagous relative social welfare weight for the bottom 99%. They are related according to the identity, $0.01\lambda_1 + 0.99\lambda_{99} = 1$ and we generally assume that $\lambda_1 < 1$ and $\lambda_{99} > 1$. Applying these social welfare weights to the dG, dV_1, and dV_{99} effects listed in Table 2, a marginal increase in τ_1 has the following effect on social welfare:

$$\begin{aligned}\frac{dSW}{d\tau_1} = Z_1 \Bigg[& (1-\lambda_1)\left(\frac{Z_1 - \bar{z}N_1}{Z_1}\right) \\ & + \tau_1\epsilon_1 \cdot \frac{d\ln(1-\tau_1)}{d\tau_1} \\ & - (1-\lambda_1)(1-\tau_1) \cdot \frac{d\ln w_1}{d\ln(1-\tau_1)} \cdot \frac{d\ln(1-\tau_1)}{d\tau_1} \\ & + \tau_1\epsilon_1 \cdot \frac{d\ln w_1}{d\ln(1-\tau_1)} \cdot \frac{d\ln(1-\tau_1)}{d\tau_1} \\ & - (1-\lambda_{99})(1-\tau_{99})\left(\frac{Z_{99}}{Z_1} \cdot \frac{d\ln w_{99}}{d\ln(1-\tau_1)}\right)\frac{d\ln(1-\tau_1)}{d\tau_1} \\ & + \tau_{99}\epsilon_{99}\left(\frac{Z_{99}}{Z_1} \cdot \frac{d\ln w_{99}}{d\ln(1-\tau_1)}\right)\frac{d\ln(1-\tau_1)}{d\tau_1}\Bigg]. \end{aligned} \quad (19)$$

At the socially optimal tax rate τ_1^*, $dSW/d\tau_1 = 0$. Taking advantage of the CRS property $Z_1 \frac{d\ln w_1}{d\ln(1-\tau_1)} + Z_{99}\frac{d\ln w_{99}}{d\ln(1-\tau_1)} = 0$ and defining $a > 1$ to be the Pareto parameter of the upper tail of the income distribution (i.e. $a = \mathbb{E}[z|z > \bar{z}]/(\mathbb{E}[z|z > \bar{z}] - \bar{z})$), τ_1^* is implicitly defined by

$$\left[(1-\lambda_1)\cdot\frac{1}{a} - \epsilon_1\frac{\tau_1^*}{1-\tau_1^*}\right]$$
$$= \left[(1-\lambda_1) - \epsilon_1\frac{\tau_1^*}{1-\tau_1^*} + (\lambda_{99}-1)\left(\frac{1-\tau_{99}}{1-\tau_1^*}\right) + \epsilon_{99}\frac{\tau_{99}}{1-\tau_1^*}\right]\left|\frac{d\ln w_1}{d\ln(1-\tau_1)}\right|. \qquad (20)$$

The term on the left is the net social welfare gain in the absence of endogenous wage effects. The optimal tax rate in that case, which we denote τ_1^{**}, is defined by

$$\frac{\tau_1^{**}}{1-\tau_1^{**}} = \frac{1-\lambda_1}{a\epsilon_1}. \qquad (21)$$

This is the optimal tax solution given in Saez (2001) and Diamond and Saez (2011). In the limiting case with $\lambda_1 = 0$, τ_1^{**} is simply the Laffer rate. The term on the right side of (20) is the net social welfare loss associated with endogenous wage changes.

(Proposition 2)**.** $\tau_1^* = \tau_1^{**}$ *if* $\frac{d\ln w_1}{\delta\ln(1-\tau_1)} = 0$. *Otherwise,* $\tau_1^* < \tau_1^{**}$.

If $(d\ln w_1)/(d\ln(1-\tau_1)) = 0$, then the right-hand side of (20) equals 0 and $\tau_1^* = \tau_1^{**}$. However, suppose $|(d\ln w_1)/(d\ln(1-\tau_1))| > 0$. If τ_1^* were set to τ_1^{**} the left-hand side of (20) would equal 0; however, the right-hand side would necessarily be positive and the marginal social costs of a higher τ_1 would exceed the marginal social benefits. To see why, consider the the bracketed term on the right-hand side of (20). It's additive components are all positive with the exception of $-\epsilon_1\cdot(\tau_1^*/(1-\tau_1^*))$.[22] This term reflects the lone positive effect of endogenous wage changes, namely that higher wages encourage labor supply among high income earners and promote economic efficiency. However, the social benefit of this pro-efficiency effect must be smaller in magnitude than the social cost of the higher wage's transfer from the government to high income earners. Formally, this follows from the fact that $((1-\lambda_1)/a) - \epsilon_1(\tau_1^*/(1-\tau_1^*)) = 0$ when $\tau_1^* = \tau_1^{**}$. Therefore $(1-\lambda_1) - \epsilon_1\tau_1^*/(1-\tau_1^*) = ((1-\lambda_1)(a-1))/a > 0$ because $a > 1$, or in other words, $(1-\lambda_1) > \epsilon_1\tau_1^*/(1-\tau_1^*)$.

[22] Recall that $\lambda_1 < 1$ and $\lambda_{99} > 1$. We are also assuming that $\tau_{99} > 0$.

The intution for this result was given in Section 3.3, but we reiterate it here. The direct effect of an increase in τ_1 is to transfer resources from high income earners to the government in a socially desirable manner. Importantly, the magnitude of the transfer is proportional only to the portion of a high earner's income that exceeds the income threshold $\bar{z}$. The drawback is that it also discourages labor supply among those earners. At τ_1^{**}, these direct benefits and costs are equalized. The indirect effects of the tax-induced increase in w_1 work in an inverse manner. The wage increase results in an undesirable tranfer back to high income earners, but also a desirable increase in their labor supply. In this case though, the undesirable transfer is proportional to the entirety of a taxpayer's income, not just the portion that exceeds $\bar{z}$. Therefore, the indirect social welfare effect of endogenous wage changes must be negative whenever the direct social welfare effect of the tax increase is positive, or even when it equals 0.

(Proposition 3). $\tau_1^* < 0$ *if* $\left|\frac{d\ln w_1}{d\ln(1-\tau_1)}\right| > \frac{(1-\lambda_1)\frac{1}{a}}{(1-\lambda_1)+(\lambda_{99}-1)(1-\tau_{99})+\epsilon_{99}\tau_{99}}$.

Proposition 3 arises from evaluation of the social marginal welfare effect at $\tau_1 = 0$. If $dSW/d\tau_1 < 0$ at $\tau_1 = 0$, then the optimal tax policy is to subsidize high income taxpayers' labor supply. While our analysis allows for the theoretical possibility of an optimal subsidy, Stern (1982) and Jacobs (2012) find that a subsidy is necessarily optimal. Proposition 3 illustrates the reason for their result. As in our model, those papers assume the existence of two different types of labor; however, their models assume that all individuals of a given labor type are identical. This implies that all of the H laborers earn the same income, which in turn implies a bounded income distribution. This implies that the a parameter that captures the thickness of the upper tail of the income distribution in infinite, which in turn implies that the inequality in Proposition 3 is necessarily satisfied. If a is finite, and in reality it is, then a subsidy may or may not be theoretically optimal.[23]

(Proposition 4). *All else equal,* τ_1^* *decreases when:*

a. λ_1 *increases and* λ_{99} *decreases;*

[23] In our numerical simulations a positive τ_1 is always optimal.

b. a increases;

c. ϵ_1 increases;

d. $|(d\ln w_1)/(d\ln(1-\tau_1))|$ increases, which in turn occurs when σ decreases or ρ^2 increases;

e. ϵ_{99} decreases, assuming τ_{99} is sufficiently close to 0; and

f. τ_{99} increases, assuming λ_{99} is sufficiently close to 1.

The formal expressions for these comparative statics are included in the appendix. We discuss only the intuition for each result here.

The optimal tax rate in the absence of endogenous wage changes, τ_1^{**}, depends only upon λ_1, a, and ϵ_1. A larger value of any of these three decreases τ_1^{**}, and these negative impacts remain in the more general treatment with wage adjustments. A larger value of λ_1 decreases the attractiveness of transfers from the 1% to the government; therefore, τ_1^* decreases. A larger value of a implies a thinner upper tail of the income distribution; therefore, the socially attractive transfers from high income taxpayers to the government are smaller and τ_1^* decreases. In the case of fixed wages, a higher ϵ_1 implies a larger labor market inefficiency associated with high tax rates. This effect remains in the case with wage adjustments; however, a larger ϵ_1 also implies that $|(d\ln w_1)/(d\ln(1-\tau_1))|$ is larger, or in other words, that high income taxpayers bear a smaller portion of the incidence of their marginal tax rate. This aligns with the general phenomenon that tax incidence falls more heavily upon the more inelastic side of the market. This increased tax-shifting reinforces the social undesirability of high τ_1.

A lower σ implies H and L are less substitutable in production. This implies that, all else equal, the relative change in wages must be larger in order to eliminate the relative shortage of H-to-L labor supply induced by high top tax rates. The relative wage change is socially undesirable and therefore τ_1^* decreases. Similarly, a high ρ^2 implies that the types of labor supplied by high income taxpayers have fewer substitutes among lower income taxpayers. This also leads to a larger increase in relative wages, and therefore τ_1^* decreases.

ϵ_{99} and τ_{99} are irrelevant to the determination of τ_1^* when wages are fixed. This is not true when wages adjust. A smaller ϵ_{99} implies that a larger portion of τ_1 is borne by lower income taxpayers via endogenous wage changes. Once again, the more inelastic side of a market bears a larger portion of the tax incidence. On the other hand, a smaller ϵ_{99} corresponds to a smaller inefficiency associated with the decrease in lower income taxpayers' labor supply that arises from their decreased wages. This inefficiency effect is proportional to the size of the distortionary wedge in the lower income taxpayers' labor supply. If τ_{99} is sufficiently small then, the former effect dominates and τ_1^* decreases.

A higher τ_{99} also has competing effects on the desirability of higher τ_1^*. On the one hand, a higher τ_{99} increases the magnitude of the inefficiency associated with distorted labor supply among low income taxpayers. On the other hand, the higher τ_{99} also shrinks the magnitude of the socially undesirable transfer from low income taxpayers to the government that is generated by endogenous wage changes. However, the social value of these transfers are proportional to $(\lambda_{99} - 1)$. Therefore, for λ_{99} close to 1, the former effect dominsates and τ_1^* decreases.

The fact that τ_1^{**} does not depend on τ_{99} implies that the optimal design of high income tax rates does not depend on the design of the rest of the tax schedule when wages are fixed. This is not true in the more general case when wages adjust. We have taken τ_{99} as given, but the dependence of τ_1^* on the value of τ_{99} emphasizes the value of an endeavor to determine the entire optimal tax schedule when wages adjust. This is beyond the scope of our paper; however, we discuss briefly how we expect such an analysis to pan out. On the one hand, price effects lead to a τ_1^* that is smaller than τ_1^{**}. Choosing a lower τ_{99} would help bring τ_1^* closer in line with τ_1^{**}. On the other hand, a lower τ_{99} would imply that fewer revenues are collected from even high income taxpyayers. After all, high income taxpayers pay a marginal rate of τ_{99} on their first $\bar{z}$ of income. Furthermore, a lower τ_{99} would itself tend to decrease low income taxpayers' wages and increase high income taxpayers' wages. In general we expect that optimal tax rates will be lower (relative to the optimal rates with fixed wages) for those incomes at which marginal tax increases lead to socially undesirable wage adjustments and higher for those incomes at which marginal tax increases lead to socially desirable wage adjustments.

6 Estimating the Optimal High Income Tax Rate

The previous analysis shows that tax incidence considerations necessarily lower the optimal tax rate on high income taxpayers. However, it remains to be seen whether these impacts are large or small. To address this issue, we estimate the optimal τ_1^* under alternative assumptions. We first fix a few parameters to reasonable values. The Pareto parameter a is set to 1.5, per Diamond and Saez (2011). We pin down the income shares of H workers and high income earners using data from Bakija, Cole, and Heim (2012), supplemented with our own analysis of the May 2005 Occupational Employment Statistics (OES) from the Bureau of Labor Statistics. BCH estimates that the top one percent earned 16.97% of national income in 2005.[24] We define H-type labor to be labor in one of the first five occupations in BCH Table 6; therefore, our estimate of the share of H within the top 1% (i.e. $s_{H|1}$) is 79.8%.[25] Based on our analysis of the OES data, the share of total income for these five occupations is 25.9%. Together, these estimates imply that the coefficient of determination is $\rho^2 = 0.310$. We believe that this is a conservative estimate of the true coefficient of determination, since occupations are likely to be a poor proxy for the differences in the type of labor supplied by high income taxpayers. We also assume that the tax rates under the status quo are $\tau_{99} = 0.21$ and $\tau_1 = 0.425$. The former is the CBO's estimate of the average federal tax rate in 2005.[26] The latter is Diamond and Saez's estimate of the effective marginal tax rate for

[24] This estimate is the share of national income excluding capital gains.

[25] The first first occupations are "executives, managers, and supervisors (non-finance)," "financial professionals, including management," "lawyers," "medical," and "real estate." We add to this value the income for those taxpayers in BCH's Table 6 who are "not working or deceased" or "unknown" in order to make our data on the top 1% more comparable to the population-wide OES data.

[26] See www.taxpolicyccenter.org/taxfacts/displayafact.cfm?Docid=456.

the top 1% of earners.[27]

Table 4 presents our results for various values of ϵ_1, ϵ_{99}, λ_1, and σ. Panel A gives τ_1^* when $\epsilon_1 = \epsilon_{99} = 0.25$. We start with this case to enable comparison to Diamond and Saez (2011). Based on our reading of the literature, we agree with their assessment that 0.25 is a reasonable (though by no means definitive) midrange estimate for the top 1%'s labor supply elasticity with respect to the after-tax rate. The elasticity of the bottom 99% is irrelevant to their analysis but not ours. Early empirical analyses suggested that high income earners were much more elastic than lower income earners, but more recent evidence suggests that a great deal of high earners' responsiveness reflects things like the timing of compensation. Therefore, it seems reasonable to consider a case where high and low earners are equally responsive. Panel B instead considers the case where $\epsilon_1 = 0.4$ and $\epsilon_{99} = 0.1$. Both values are plausible according to McClelland and Mok (2012), and their rank ordering is supported by Gruber and Saez (2002).

Within each panel are two rows. The first estimates τ_1^* assuming $\lambda_1 = 0$, the second asssuming $\lambda_1 = 0.04$. These values are again chosen in order to make our results more directly comparable to Diamond and Saez (2011), who estimate τ_1^{**} under both scenarios.[28] Finally, the three columns of the Table estimate τ_1^* for three values of σ: $\sigma = \infty$, $\sigma = 1$, and $\sigma = 0$. The first value implies that all labor is perfectly substitutable. The second assumes that the two labor types contribute to aggregate production via a Cobb-Douglas production function, as assumed in Feldstein (1973). The third assumes that the two labor types are perfect complements. The assumption is extreme and should be interpreted as a boundary; however, we do not consider it any more extreme than an assumption of perfect substitutability.

Comparing τ_1^* across rows, it appears to be minimally affected by switching from $\lambda_1 = 0$

[27]The estimate of τ_{99} is necessary because its value directly affects the optimal value τ_1^*. An estimate of the status quo τ_1 is necessary due to the endogeneity of optimal τ_1^* determinants like the share of the top 1%. That is, we have to know the τ_1 that applied when our income shares were observed in order to estimate how those shares would change when the tax changes to τ_1^*.

[28]In order to satisfy the identity that the population-weighted average of λ_1 and λ_{99} equal 1, $\lambda_{99} = 1/0.99$ when $\lambda_1 = 0$ and $\lambda_{99} = (1 - (0.04)(0.01))/0.99$ when $\lambda_1 = 0.04$.

to $\lambda_1 = 0.04$. In Diamond and Saez (2001), the switch only reduces τ_1^{**} by approximately one percentage point. Table 2 shows that this one percentage point difference does not appear to depend on the values of ϵ_1, ϵ_{99}, or σ.

Comparing within-row outcomes, the value of τ_1^* only drops by approximately one percentage point when we shift from $\sigma = \infty$ to $\sigma = 1$. This result supports the Feldstein (1973) conclusion that endogenous wage effects have little impact on the design of optimal tax policy. On the other hand, the value of τ_1^* drops considerably more when we shift from $\sigma = 1$ to $\sigma = 0$. This supports the Allen (1982) critique of Feldstein (1973), namely that the latter's result depended critically upon the assumption of Cobb-Douglas technology. In the case of $\epsilon_1 = \epsilon_{99} = 0.25$ and $\lambda_1 = 0$ (row A), τ_1^* drops approximately six percentage points, from 72.7 to 66.8, as we move from column 1 to column 3. Framed alternatively, τ_1^* decreases by nearly 10 percent and the after-tax rate $(1 - \tau_1^*)$ increases by nearly 25 percent when we shift from an assumption of perfect substitutes to perfect complements. The impacts of changes in σ are even larger in the case of $\epsilon_1 = 0.4$, $\epsilon_{99} = 0.1$, and $\lambda_1 = 0.04$ (row D). In that case, τ_1^* drops approximately 18 percentage points, from 61.5 to 43.7, as we move from column 1 to column 3. Framed alternatively, τ_1^* decreases by nearly 30 percent and the after-tax rate increases by nearly 50 percent when we shift from an assumption of perfect substitutes to perfect complements.

In net, our results suggest that optimal tax rates may be signifiantly lower when wages endogenously adjust compared to when wages are fixed. On the other hand, the impact of wage adjustments does not appear to be nearly as large as suggested in the previous literature. We do not find support for the Stern (1982) and Jacobs (2012) conclusion that high income earners should be subsidized. We could not reject the possibility based on theory alone; however, our various parameterizations suggest that the theoretical possibility has little practical relevance. Furthermore, our parameterizations also support the notion that progressive income taxation is desirable even when undesirable wage adjustments occur. Our lowest estimated τ_1^* of 43.7 is still double our estimate of τ_{99}.

When $\sigma = 1$ and $\epsilon_1 = \epsilon_{99} = 0.25$, we estimate that $|(d \ln w_1)/(d \ln(1-\tau))|$ is approximately 0.05. When ϵ_1 goes up to 0.4 and ϵ_{99} goes down to 0.1, $|(d \ln w_1)/(d \ln(1-\tau))|$ increases to approximately 0.08. Since the magnitude of the tax-shifting is not very large in these cases, it is not surprising that the τ_1^* differs minimally from the case where $\sigma = \infty$. In constrast, when $\sigma = 0$ and $\epsilon_1 = \epsilon_{99} = 0.25$, we estimate that $|(d \ln w_1)/(d \ln(1-\tau))|$ is approximately 0.24. When ϵ_1 goes up to 0.4 and ϵ_{99} goes down to 0.1, $|(d \ln w_1)/(d \ln(1-\tau))|$ increases to approximately 0.47. Tax-shifting is much more substantial, and τ_1^* is accordingly very different from the case where $\sigma = \infty$.

As mentioned in Section 2, there is little empirical evidence on the incidence effects of income taxation. The few studies that exist are subject (through no fault of the authors) to considerable skepticism regarding their internal validity. Even if internal validity is attained, the external validity of the results is questionable. For instance Kubik (2004) concludes with the caveat that his incidence results "should not be interpreted as predictions of how the labor market will respond to other potential tax changes...Other tax reforms that affect different groups of workers might have somewhat different effects." (pg. 1586) Nonetheless we are comforted that our methodology's implied tax-shifting rates under the $\sigma = 0$ scenario are comparable to his estimates.[29]

7 Discussion and Conclusion

We consider it noteworthy that plausible parameterizations lead to τ_1^* values in the low-to-mid 40% range.[30] These values are remarkably close to the current effective marginal tax

[29]Kubik (2004) estimates that a one percentage point increase in the median marginal tax rate for an occuapation leads to an increase of 0.2514 in the log median wage of the occupation. Evaluated at the mean of the median marginal tax rate across occupations (20.92), this implies that the elasticity of wages with respect to the after-tax rate in his study equals $(0.2514)(1 - 0.2092) \approx 0.2$. If Kubik's observed changes in workers' hourly wages compound both changes in workers' wage *per efficiency unit* as well as behavioral changes in the quantity of efficiency units supplied per hour, then the elasticity of wage *per efficiency unit* with respect to the after-tax rate exceeds 0.2, perhaps by a sizeable amount.

[30]Strictly postive values of σ could also lead to similar estimates of τ_1^* if, for instance, we have underestimated the coefficient of determination ρ^2.

rate in the United States.[31] That being said, we also heed the advice of Professors Atkinson and Stiglitz:

> "One must bear firmly in mind the purpose of this kind of literature. The aim is not to provide a definite numerical answer to the question, 'how progressive should the income tax be?'... The results are therefore qualitative, rather than quantitative. At the same time they serve to identify the key factors, and to provide counter-examples to certain popular views." (Atkinson and Stiglitz (1980), pgs. 422-423)

As such, we do not wish to overstate the importance of the specific values in Table 4. Instead, we have demonstrated that accounting for tax incidence satisfies the first Diamond and Saez (2011) policy-relevance criterion, namely that "the result should be based on an economic mechanism that is empirically relevant and first order to the problem hand." (pg. 166) Table 4 shows that endogenous wage changes can have first order effects on the problem at hand. The empirical relevance of the issue is more debateable, but only because the topic of income tax incidence is understudied, not because of any overwhelming evidence that the economic incidence of income taxation equals the statutory incidence. We hope that our analysis motivates future empirical endeavors that examine the wage effects of income tax reforms and how the labor of high income earners differs (or does not) from that of low income earners.

We also acknowledge several simplifications in our model that we hope we hope will be addressed in future research. First, it is worth expanding the model to more than two types of labor. Second, we have omitted non-labor inputs from our model. We hypothesize that the introduction of a third input, capital, may have significant impacts on the results. If capital is complementary to the labor supplied by high income earners, then some of the burden of high marginal income tax rates may be shifted to capital income earners. Since

[31] Diamond and Saez (2011) estimated the rate, inclusive of average state income and federal non-income taxes, at 42.5% prior to the tax increases under President Obama's fiscal cliff deal.

high income taxpayers' disproportionately receive capital income, then the equilibrium price effects of high marginal tax rates may be less undesirable. Third, we have assumed that wages are determined in a competitive market setting. It would be worthwhile to assess how noncompetitive pricing affects the results, especially given evidence on the market for superstars (Rosen (1981), Kaplan and Rauh (2013)), the market for executives (Gabaix and Landier (2008)), and rent-seeking (Piketty, Saez, and Stantcheva (2014)). That being said, we caution against reading too much into results that rely on another homogeneity assumption, namely that all high income earners' wages are determined by the same pricing mechanism, competitive or otherwise. Accounting for noncompetitive pricing may be particularly relevant once the model is expanded to allow for more than two types of labor. We are huge fans of Harry Potter, and find it hard to imagine that any number of individuals (or even monkeys typing) could perfectly substitute for J.K. Rowling; we also find it hard to imagine that Ms. Rowling's pay is set under perfectly competitive conditions.

Our research demonstrates that tax incidence considerations can substantially alter the optimal design of tax policy at the top of the income distribution. The Rothstein (2010) analysis of the EITC finds that the same can be said of tax policy at the bottom of the income distribution. We believe that the insights of both papers can be useful towards deriving the complete nonlinear optimal income tax schedule in an environment of general equilibrium wage adjustments.

References

Allen, Franklin, 1982. "Optimal linear income taxation with general equilibrium effects on wages." Journal of Public Economics 17, 135-143.

Atkinson, Anthony B., and Joseph E. Stiglitz, 1980. "Lectures on Public Economics." McGraw-Hill Book Company, New York.

Atkinson, Anthony B., Thomas Piketty, and Emmanuel Saez, 2011. "Top incomes in the long run of history." Journal of Economic Literature 49, 3-71.

Bakija, Jon, Adam Cole, and Bradley T. Heim. "Jobs and income growth of top earners and the causes of changing income inequality: evidence from U.S. tax return data." Working paper.

Bingley, Paul, and Gauthier Lanot, 2002. "The incidence of income tax on wages and labour supply." Journal of Public Economics 83, 173-194.

Blomquist, Soren, and Hakan Selin, 2010. "Hourly wage rate and taxable labor income responsiveness to changes in marginal tax rates." Journal of Public Economics 94, 878-889.

Card, David, and Thomas Lemieux, 2001. "Can falling supply explain the rising return to college for younger men? A cohort-based analysis." Quarterly Journal of Economics 116, 705-746.

Carruth, Alan A., 1982. "On the role of the production and consumption assumptions for optimum taxation." Journal of Public Economics 17, 145-155.

Christiansen, Vidar, 1988. "Choice of occupation, tax incidence, and piecemeal tax revision." The Scandinavian Journal of Economics 90, 141-159.

Diamond, Peter, 1988. "Optimal income taxation: an example with a U-shaped pattern of optimal marginal tax rates." American Economic Review 88, 83-95.

Diamond, Peter A., and James A. Mirrlees, 1971. "Optimal taxation and public production I: production efficiency." American Economic Review 61, 8-27.

Diamond, Peter, and Emmanuel Saez, 2011. "The case for a progressive tax: from basic research to policy recommendations." Journal of Economic Perspectives 25, 165-190.

Dixit, Avinash, and Agnar Sandmo, 1977. "Some simplified formulae for optimal income taxation." The Scandanavian Journal of Economics 4, 417-423.

Feldstein, Martin, 1973. "On the optimal progressivity of the income tax." Journal of Public Economics 2, 357-376.

Fullerton, Don, and Gilbert E. Metcalf, 2002. "Tax Incidence," in: Auerbach, Alan J., and Martin Feldstein (eds.), Handbook of Public Economics. Elsevier, Amersterdam, 1787-1872.

Gabaix, Xavier, and Augustin Landier, 2008. "Why has CEO pay increased so much?" Quaterly Journal of Economics 123, 49-100.

Goldin, Claudia, and Lawrence F. Katz, 2007. "Long-run changes in the wage structure: narrowing, widening, polarizing." Brookings Papers on Economic Activity 2, 135-165.

Gruber, Jon, and Emmanuel Saez, 2002. "The elasticity of taxable income: evidence and implications." Journal of Public Economics 84, 1-32.

Jacobs, Bas, 2012. "Optimal redistributive tax and education policies in general equilibrium." International Tax and Public Finance 20, 312-337.

Katz, Lawrence F., and Kevin M. Murphy, 1992. “Changes in relative wages, 1963-1987: supply and demand factors.” Quarterly Journal of Economics 107, 35-78.

Kaplan, Steven N., and Joshua Rauh, 2013. “It’s the market: the broad-based rise in the return to top talent.” Journal of Economic Perspectives 27, 35-56.

Kubik, Jeffrey D., 2004. “The incidence of personal income taxation: evidence from the tax reform act of 1986.” Journal of Public Economics 88, 1567-1588.

Mirrlees, James A., 1971. “An exploration in the theory of optimal income taxation.” Review of Economic Studies 38, 175-208.

McClelland, Robert, and Shannon Mok. “A review of recent research on labor supply elasticities.” Congressional Budget Office working paper.

Piketty, Thomas, Emmanuel Saez, and Stefanie Stantcheva, 2014. “Optimal taxation of top labor incomes: a tale of three elasticities.” American Economic Journal: Economic Policy 6, 230-271.

Rosen, Sherwin, 1981. “The economics of superstars.” American Economic Review 71, 845-58.

Rothstein, Jesse, 2010. “Is the EITC as good as an NIT? Conditional cash transfer and tax incidence.” American Economic Journal: Economic Policy 2, 177-208.

Saez, Emmanuel, 2001. “Using elasticities to derive optimal income tax rates.” Review of Economic Studies 68, 205-229.

Saez, Emmanuel, Joel Slemrod, and Seth H. Giertz, 2012. “The elasticity of tax income with respect to marginal tax rates: a critical review.” Journal of Economic Literature 50, 3-50.

Seade, J.K., 1977. “On the shape of optimal tax schedules.” Journal of Public Economics 7, 203-235.

Stern, Nicholas, 1982. “Optimum taxation with errors in administration.” Journal of Public Economics 1982, 181-211.

Appendix: Proof of Proposition 4

Setting $dSW/d\tau_1$ (given in (19)) equal to 0 yields the first-order condition for the optimal τ_1. Defining $D = d^2SW/d\tau_1^2$, the second-order condition is $D < 0$. Implicit differentiation of the first-order condition yields the comparative statics in Proposition 4.

- *With respect to* λ_1

Implicit differentiation of the first-order condition gives

$$\frac{d\tau_1^*}{\lambda_1} = \frac{Z_1}{(-D)} \cdot \left(-\frac{1}{a} + \left(1 - \frac{d\lambda_{99}}{d\lambda_1} \cdot \frac{1-\tau_{99}}{1-\tau_1} \right) \left| \frac{d\ln w_1}{d\ln(1-\tau_1)} \right| \right). \tag{22}$$

λ_1 and λ_{99} are related according to $0.01\lambda_1 + 0.99\lambda_{99} = 1$; therefore, any increase in λ_1 correspondingly implies that λ_{99} decreases according to

$$\frac{d\lambda_{99}}{d\lambda_1} = -\frac{\lambda_{99}-1}{1-\lambda_1} \quad . \tag{23}$$

Plugging this into (22) and using the first-order condition, (22) can be rewritten as

$$\frac{d\tau_1^*}{d\lambda_1} = \frac{-Z_1}{(-D)(1-\lambda_1)} \cdot \left(\epsilon_1 \cdot \left(\frac{\tau_1^*}{1-\tau_1^*} \right) \cdot \left(1 - \left| \frac{d\ln w_1}{d\ln(1-\tau_1)} \right| \right) + \epsilon_{99} \cdot \left(\frac{\tau_{99}}{1-\tau_1^*} \right) \left| \frac{d\ln w_1}{d\ln(1-\tau_1)} \right| \right). \tag{24}$$

Therefore $d\tau_1^*/d\lambda_1 < 0$ stems from the fact that $|(d\ln w_1)/(d\ln(1-\tau_1))| \in (0,1)$ and the assumptions that τ_1^* and τ_{99} are both positive.

- *With respect to* a

Implicit differentiation of the first-order condition gives

$$\frac{d\tau_1^*}{\lambda_1} = \frac{-Z_1(1-\lambda_1)}{(-D)a^2}. \tag{25}$$

Therefore $d\tau_1^*/da < 0$.

- *With respect to ϵ_1*

Implicit differentiation of the first-order condition gives

$$\frac{d\tau_1^*}{\epsilon_1} = \frac{-Z_1}{(-D)} \cdot \left(\frac{\tau_1^*}{1-\tau_1^*} \cdot \left(1 - \left| \frac{d \ln w_1}{d \ln(1-\tau_1)} \right| \right) + \left((1-\lambda_1) \frac{1}{a} - \epsilon_1 \frac{\tau_1^*}{1-\tau_1^*} \right) \cdot \frac{d \ln \left| \frac{d \ln w_1}{d \ln(1-\tau_1)} \right|}{d\epsilon_1} \right), \tag{26}$$

where

$$\frac{d \ln \left| \frac{d \ln w_1}{d \ln(1-\tau_1)} \right|}{d\epsilon_1} = \frac{1}{\epsilon_1} \cdot \frac{\sigma + \left(s_{99|H} s_L + s_{99|L} s_H \right) \epsilon_{99}}{\sigma + \tilde{\epsilon}} > 0. \tag{27}$$

Therefore $d\tau_1^*/d\epsilon_1 < 0$ stems from the fact that $|(d \ln w_1)/(d \ln(1-\tau_1))| \in (0,1)$, the fact that $(1-\lambda_1)/a > \epsilon_1 \tau_1^*/(1-\tau_1^*)$, and the assumption that τ_1^* is positive.

- *With respect to $|(d \ln w_1)/(d \ln(1-\tau_1))|$, σ, and ρ^2*

Defining $\gamma \equiv |(d \ln w_1)/(d \ln(1-\tau_1))|$, implicit differentiation of the first-order condition gives

$$\frac{d\tau_1^*}{d\gamma} = \frac{-Z_1 \left((1-\lambda_1) \frac{1}{a} - \epsilon_1 \frac{\tau_1^*}{1-\tau_1^*} \right)}{(-D)\gamma}. \tag{28}$$

Therefore $d\tau_1^*/d\gamma < 0$ stems from the fact that $(1-\lambda_1)/a > \epsilon_1 \tau_1^*/(1-\tau_1^*)$. Furthermore, differentiation of γ yields

$$\frac{d\gamma}{d\sigma} = \frac{-\gamma}{\sigma + \tilde{\epsilon}} < 0 \tag{29}$$

and

$$\frac{d\gamma}{d\rho^2} = \frac{\gamma}{\rho^2}. \tag{30}$$

- *With respect to ϵ_{99}*

Continuing to utilize the prior definition of γ, implicit differentiation of the first-order condition gives

$$\frac{d\tau_1^*}{d\epsilon_1} = \frac{-Z_1}{(-D)} \cdot \left(\frac{\tau_{99}}{1-\tau_1} \gamma + \left((1-\lambda_1) - \epsilon_1 \frac{\tau_1^*}{1-\tau_1^*} \right) \frac{d \ln \gamma}{d\epsilon_{99}} \right) \tag{31}$$

where

$$\frac{d\ln\gamma}{d\epsilon_{99}} = \frac{-\left(s_{99|H}s_L + s_{99|L}s_H\right)}{\sigma+\tilde{\epsilon}} < 0. \tag{32}$$

In general, the sign of $d\tau_1^*/d\epsilon_{99}$ is theoretically ambiguous. However, if τ_{99} is sufficiently close to 0, then $d\tau_1^*/d\epsilon_{99} > 0$.

- *With respect to* τ_{99}

Continuing to utilize the prior definition of γ, implicit differentiation of the first-order condition gives

$$\frac{d\tau_1^*}{d\tau_{99}} = \frac{Z_1\gamma\left((\lambda_{99}-1)-\epsilon_{99}\right)}{(-D)(1-\tau_1^*)}. \tag{33}$$

In general, the sign of $d\tau_1^*/d\tau_{99}$ is theoretically ambiguous. However, if λ_{99} is sufficiently close to 1, then $d\tau_1^*/d\tau_{99} < 0$ assuming τ_1^* is positive.

Affected Parties	dG	dV_{H1}	dV_{L1}	dV_{H99}	dV_{L99}
I. Direct transfers from change in τ_1					
a. $G \leftrightarrow H1$	$(Z_{H1} - \bar{z}N_{H1})d\tau_1$	$-(Z_{H1} - \bar{z}N_{H1})d\tau_1$			
b. $G \leftrightarrow L1$	$(Z_{L1} - \bar{z}N_{L1})d\tau_1$		$-(Z_{L1} - \bar{z}N_{L1})d\tau_1$		
II. Indirect transfers from tax-induced changes in w_H and w_L					
c. $G \leftrightarrow H1$	$-(1-\tau_1)Z_{H1}d\ln w_H$	$(1-\tau_1)Z_{H1}d\ln w_H$			
d. $G \leftrightarrow L1$	$-(1-\tau_1)Z_{L1}d\ln w_L$		$(1-\tau_1)Z_{L1}d\ln w_L$		
e. $G \leftrightarrow H99$	$-(1-\tau_{99})Z_{H99}d\ln w_H$			$(1-\tau_{99})Z_{H99}d\ln w_H$	
f. $G \leftrightarrow L99$	$-(1-\tau_{99})Z_{L99}d\ln w_L$				$(1-\tau_{99})Z_{L99}d\ln w_L$
III. Direct behavioral effects from change in τ_1					
g. $H1$	$\tau_1 Z_{H1}\epsilon_1 d\ln(1-\tau_1)$				
h. $L1$	$\tau_1 Z_{L1}\epsilon_1 d\ln(1-\tau_1)$				
IV. Indirect behavioral effects from tax-induced changes in w_H and w_L					
i. $H1$	$\tau_1 Z_{H1}\epsilon_1 d\ln w_H$				
j. $L1$	$\tau_1 Z_{L1}\epsilon_1 d\ln w_L$				
k. $H99$	$\tau_{99} Z_{H99}\epsilon_{99} d\ln w_H$				
l. $L99$	$\tau_{99} Z_{L99}\epsilon_{99} d\ln w_L$				

Table 1: Disaggregated Effects of $d\tau_1$

Affected Parties	dG	dV_1	dV_{99}
I. Direct transfer from change in τ_1			
a. $G \leftrightarrow 1$	$(Z_1 - \bar{z}N_1)d\tau_1$	$-(Z_1 - \bar{z}N_1)d\tau_1$	
II. Indirect transfers from tax-induced changes in w_1 and w_{99}			
b. $G \leftrightarrow 1$	$-(1-\tau_1)Z_1 d\ln w_1$	$(1-\tau_1)Z_1 d\ln w_1$	
c. $G \leftrightarrow 99$	$-(1-\tau_{99})Z_{99} d\ln w_{99}$		$(1-\tau_{99})Z_{99} d\ln w_{99}$
III. Direct behavioral effect from change in τ_1			
d. 1	$\tau_1 Z_1 \epsilon_1 d\ln(1-\tau_1)$		
IV. Indirect behavioral effects from tax-induced changes in w_1 and w_{99}			
e. 1	$\tau_1 Z_1 \epsilon_1 d\ln w_1$		
f. 99	$\tau_{99} Z_{99} \epsilon_{99} d\ln w_{99}$		

Table 2: Effects of $d\tau_1$ Aggregated by Income

	Share of Top 1% Income		Share of All Income	
	Share (%)	Cumulative Share (%)	Share (%)	Cumulative Share (%)
	(1)	(2) [=sum(1)]	(3)	(4) [=sum(3)]
Executives, managers, and supervisors (non-finance)	37.4	37.4	15.5	15.5
Financial professions, including management	16.3	53.7	3.9	19.4
Lawyers	7.2	60.9	1.2	20.6
Medical	10.9	71.8	4.8	25.4
Real Estate	3.4	75.2	0.4	25.9
Skilled sales (except finance or real estate)	3.1	78.3	3.6	29.5
Arts, media, sports	2.5	80.8	1.2	30.7
Computer, math, engineering, technical (nonfinance)	3.5	84.3	7.3	38.0
Business operations (nonfinance)	2.8	87.2	3.0	41.0
Professors and scientists	1.4	88.5	2.3	43.2

Table 3: Share of Top 1% and Total Income by Population (2005)

Labor Supply Elasticities	**Elasticity of Substitution** $\sigma = \infty$	$\sigma = 1$	$\sigma = 0$
I. $\epsilon_1 = \epsilon_{99} = 0.25$			
A. $\lambda_1 = 0$	72.7	71.9	66.8
B. $\lambda_1 = 0.04$	71.9	71.0	65.9
II. $\epsilon_1 = 0.4$**,** $\epsilon_{99} = 0.1$			
C. $\lambda_1 = 0$	62.5	61.2	44.6
D. $\lambda_1 = 0.04$	61.5	60.2	43.7

Table 4: Estimates of τ_1^*