NBER WORKING PAPER SERIES

THE OPTIMAL ELASTICITY OF TAXABLE INCOME

Joel Slemrod
Wojciech Kopczuk

Working Paper 7922
http://www.nber.org/papers/w7922

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2000

We are grateful to the referees, Julie Cullen, Bev Dahlby, Austan Goolsbee, Jim Hines, and to members of the University of Michigan Public Finance bag lunch workshop, the National Bureau of Economic Research Public Economics workshop, the Harvard Public Finance seminar and the Stanford Institute for Theoretical Economics workshop for comments on an earlier draft.The views expressed are those of the authors and not necessarily those of the National Bureau of Economic Research.

The Optimal Elasticity of Taxable Income
Joel Slemrod and Wojciech Kopczuk
NBER Working Paper No. 7922
September 2000
JEL No. H21, H23, H26

ABSTRACT

The strength of the behavioral response to a tax rate change depends on the environment individuals operate in, and may be manipulated by instruments controlled by the government. We first derive a measure of the social benefit to affecting this elasticity. The paper then examines this effect in the solution to the optimal income taxation problem when such an instrument is available, first in a general model and then in an example when the government chooses the income tax base.

Joel Slemrod
Office of Tax Policy Research
University of Michigan Business School
Ann Arbor, MI 48109
and NBER
jslemrod@umich.edu

Wojciech Kopczuk
Department of Economics
University of Michigan
Ann Arbor, MI 48109
wojtek@umich.edu

1 Introduction

Behavioral elasticities are key to understanding the efficiency cost of non-lump sum taxes. Standard treatments of optimal income taxation take as given behavioral elasticities derived from immutable preferences, and characterize the tradeoff between redistribution and the deadweight loss of progressive taxes. Loosely speaking, larger elasticities imply that less progressive tax systems are optimal. Most of these models have focused on the elasticity of labor supply, in the belief that the relative price of leisure and goods is the most important margin affected by taxes. Empirical evidence suggests that, at least for hours worked, the aggregate (compensated) labor supply elasticity is quite small. However, increasing leisure is by no means the only possible margin of response to marginal tax rates. When personal tax rates on ordinary income rise, evasion may increase, businesses may shift to corporate form, there may be a rise the consumption of deductible activities such as charitable giving, and individual may rearrange their portfolios and compensation packages to receive more income as tax-preferred capital gains. All of these responses to higher taxes will show up in declines in taxable income, and there is a growing body of evidence, that, at least for high-income individuals, the elasticity of taxable income is substantial.

As Feldstein (1999) argues, under certain assumptions it is the compensated elasticity of taxable income, regardless of whether its origin is responsiveness of labor supply or some other behavior, that summarizes the efficiency cost of taxation and therefore is the crucial parameter in models of optimal progressivity. The idea is that all tax-induced behavior entails costs that will be incurred until, at the margin, the private cost equals tax saving. Certain qualifications to this assertion are discussed in Slemrod (1998), including the importance of accounting for shifts across tax bases (individual to corporate, for example) and across time (e.g., a shift to deferred compensation or greater use of retirement accounts).

One critical qualification to the role of the elasticity of taxable income is that it is not only a function of preferences, and is therefore not immutable. With regard to labor supply, it is not too egregious to assume, as it became conventional in the empirical literature, that its elasticity is a primitive value based on preferences.[1] However, this assumption is very dubious when considering the other margins of response to taxation. On the contrary, the magnitude of these behavioral responses depends on a host of policy decisions about such things as the tax base and how it is enforced. For example, the responsiveness of evasion depends on the enforcement regime, the ability to shift the form of business organization depends on the rules governing the choice, the capital gains realization elasticity depends on how carry-

[1]It is more heroic, but standard, to assume that the elasticity is constant across individuals and time. In fact, the labor supply elasticity may depend on labor market institutions such as work-week regulations.

over at death is treated, and the size of the tax shelter industry created by high tax rates depends on the passive loss limitation provisions.[2]

Recognizing that the elasticity of taxable income is subject to policy control has intriguing policy implications. For example, Slemrod (1994) demonstrates that the optimal degree of income tax progressivity given a suboptimal setting of tax enforcement can be below the globally optimal progressivity. The paper adopts the metaphor of Okun (1975), in which redistribution is characterized as transferring water from the rich to the poor in a leaky bucket, the degree of leakage representing the efficiency cost of progressive taxes. Using Okun's language, Slemrod (1994) raises the possibility that the leak can be fixed, albeit at some cost. Moreover, there is an optimal rate of leakage that balances the cost of fixing the leak against the improved efficency of carrying water with a less leaky bucket. Put less metaphorically, there are multiple instruments to an optimal tax system problem, and an important attribute of a non-tax-rate, or administrative, instrument is its effect on the elasticity of taxable income with respect to the tax rate. The judicious choice of these instruments can be thought of as choosing the *optimal* elasticity of taxable income. This paper examines what characterizes this optimal elasticity.

The issue of optimal tax administration and enforcement, and their relationship to the optimal rate structure, has been considered in the prior literature. Several papers address this problem in a representative agent setting. Mayshar (1991) provides general criteria for optimal administration and enforcement. Kaplow (1990) and Cremer & Gahvari (1993) characterize the optimal enforcement and optimal tax rate structure of commodity taxes. Also in regard to commodity taxation, Yitzhaki (1979) and Wilson (1989) model the optimal tax base as a tradeoff between the increased administrative costs, but smaller excess burden, of a larger base. We draw heavily on their framework in Section 3 of this paper.

In a series of papers, Cremer & Gahvari (1994, 1996) examine the same issue in the context of setting optimal income tax progressivity in a world of heterogeneous individuals, and characterize optimal enforcement and the optimal progressivity when tax evasion is present. However, the specific framework they use — in which evasion is a choice made by taxpayers under the prospect of an uncertain audit — makes it difficult to generalize their results to instruments other than the intensity of enforcement and the penalties applied, and focuses on the role of taxpayers' risk preferences. None of the articles in either strand of this literature frames the question — as we do in what follows — around the elasticity of taxable income or

[2]It is interesting to note the relation of our point to the influential critique of macroeconomic models in Lucas (1976). As does Lucas, we raise the possibility that changes in policy affect behavioral elasticities, and, as a result, econometric procedures that ignore this possibility may give misleading results. However, our main argument is different. We stress that affecting behavioral elasticities may be the *goal* of the policy and not just its by-product, because it may make standard tax instruments (e.g., tax rates) more (or less) efficient.

characterizes the central problem as one of choosing its optimal value.

We develop a framework that integrates and systematizes the results of these literatures by identifying three types of effects that any policy instrument has: a direct welfare cost, the revenue effect holding the strength of behavioral response constant and the effect on distortions caused by other tax instruments. Each non-tax-rate, or administrative, instrument will feature each of these effects, but we stress that it is the last effect that makes non-tax-rate policy variables fundamentally different from tax rate instruments. In particular, it is this last effect that is responsible for complementarities between tax design and tax administration, and thus its understanding is key in the process of designing the optimal tax system. The optimal setting of an administrative instrument must balance between its cost and the benefit of reducing distortions; it is this trade-off that makes us interpret it as choosing *the optimal elasticity*.

The paper proceeds as follows. First, we characterize the optimal linear income tax solution when the government has an instrument which affects the elasticity of taxable income. Recognizing that, in general, such an instrument has also a redistributive impact as well as an effect on tax revenue, we isolate the value of just affecting the elasticity. In a special case, when the elasticity is not a function of the tax rate, the benefit of reducing it is directly related to the deadweight loss of taxation. Next, we investigate at length a special case in which the policy instrument which determines the taxable income elasticity is the broadness of the tax base. In this context, we demonstrate some propositions about the optimal tax system and the optimal elasticity of taxable income, in particular how they depend on the primitve parameters of the problem such as the social welfare function and the cost of administration.

2 The Optimal Linear Income Tax with a Policy Instrument that Affects the Elasticity of Response

We first consider the general problem of choosing the optimal linear income tax in a multi-person economy. In this problem the government must raise R in revenue.[3] In the standard version of this problem, it can choose two parameters of a linear income tax, the marginal tax rate (t) and a uniform lump-sum tax (T). Raising all of the required revenue using the lump-sum tax would minimize excess burden, but the population is comprised of people with a distribution of abilities to earn income, and the government objective function is such that it would be willing to accept some inefficiency in exchange for less inequality in the distribution of welfare. We

[3]We do not inquire about, nor do we model, how the government disposes of R.

assume that the society consists of a large number of individuals indexed by the parameter y, a measure of ability, with the distribution function $F(y)$ normalized so that $\int F(y)\,dy = 1$.

We modify the standard problem by assuming that the government controls an instrument m, which *inter alia* reduces the elasticity of taxable income with respect to the tax rate, but has a cost of implementation of $A(m)$, $A'(m) > 0$. Given the parameters of the tax system, each individual maximizes the utility function that gives rise to the indirect utility function $v(t, T, m, y)$ and the taxable income function $I(t, T, m, y)$.[4] The total tax paid by the individual is equal to $tI + T$.

The government is assumed to choose its instruments to maximize a generalized utilitarian social welfare function ω, so its problem can be expressed as

$$\max_{t,T,m} \int \omega(v(t, T, m, y))\, dF(y)\,, \tag{1}$$

subject to the budget constraint

$$T + \int tI(t, T, m, y)\, dF(y) = R + A(m)\,. \tag{2}$$

We use the concept of the compensated (Hicksian) taxable income denoted by $I^c(t, T, m, y, u)$ (where u is the utility level), which is the taxable income implied by a solution to the consumer's expenditure minimization problem.[5] Compensated

[4]To keep our framework general, we do not specify the utility function, nor do we define taxable income explicitly. Some examples of problems that fit into this framework are:

- A standard model with consumption (C) - labor (L) choice and an instrument m that affects individual's preferences between these two goods (e.g., a public good or education of taxpayers about services financed through taxes). In this case, the budget constraint is $C = yL - T - tI$. Maximization of the utility function $u(C, L, m)$ gives rise to the indirect utility function and taxable income defined by $I = yL$.
- A model (as in Mayshar (1991) or Slemrod (forthcoming)) with tax sheltering S, that costs $M(S, yL, m)$. The utility function $u(C, L)$ is maximized subject to the budget constraint $C = yL - T - t(yL - S) - M(S, yL, m)$; in this case taxable income is defined as $I = yL - S$.
- Our example of Section 3. An individual consumes a large number (continuum) of commodities, described by a bundle $\{C_i\}_{i\in[0,1]}$ with utility function $u(C)$. Labor supply is present in the model, but it is inelastic. The policy instrument $m \in [0, 1]$ determines the size of the commodity tax base (equivalently, goods $i \in [m, 1]$ are income-tax-deductible), and the budget constraint takes the form of $\int_0^1 C_i\, di = y - T - t(y - \int_m^1 C_i\, di)$. Consequently, $I = y - \int_m^1 C_i\, di$.

Because we allow for primitives other than the utility function (for example, an avoidance technology), it may be impossible to recover taxable income from the indirect utility function, so that they may be "independent."

[5]In the expenditure minimization problem, the consumer minimizes spending, subject to the minimum utility constraint. Consequently, we assume that the dual approach is valid, so that

demand is a well-established concept, and the compensated taxable income is an aggregation of compensated demands for relevant commodities. Analogously, the Slutsky-type equation for taxable income can be derived,

$$\frac{\partial I^c}{\partial(1-t)} = \frac{\partial I}{\partial(1-t)} - I\frac{\partial I}{\partial E}\,, \tag{3}$$

where $\frac{\partial I}{\partial E}$ is the income effect. Denoting the Lagrange multiplier by λ, the three first-order conditions, rearranged using the standard results of consumer optimization (Roy's identity and the Slutsky equation), are as follows

$$T: \qquad 0 = \int \left(-\omega'\mu + \lambda(1 - t\frac{\partial I}{\partial E})\right) dF(y)\,, \tag{4}$$

$$t: \qquad 0 = \int \left(-I\omega'\mu + \lambda I(1 - t\frac{\partial I}{\partial E} - \frac{t}{1-t}\varepsilon)\right) dF(y)\,, \tag{5}$$

$$m: \qquad 0 = \int \left(\omega' v_m + \lambda(t\frac{\partial I}{\partial m} - A'(m))\right) dF(y)\,, \tag{6}$$

where μ denotes the marginal utility of exogenous income (the Lagrange multiplier from the consumer's maximization problem), and $\varepsilon \equiv \frac{1-t}{I^c}\frac{\partial I^c}{\partial(1-t)}$ is the compensated elasticity of taxable income with respect to the net-of-tax rate $(1-t)$. Stating the elasticity with respect to $(1-t)$, rather than with respect to t, is standard in this literature.

For given m, conditions (4), (5) and the budget constraint determine the solution to the standard optimal taxation problem. For example, Atkinson & Stiglitz (1980) combine them to form

$$\frac{t}{1-t} = -\frac{\operatorname{cov}(I, g)}{\int I\varepsilon\, dF(y)}\,, \tag{7}$$

where $g \equiv \frac{\omega'\mu}{\lambda} + t\frac{\partial I}{\partial E}$ is the marginal social valuation of a lump-sum transfer to an individual. This expression implicitly determines the optimal tax rate and suggests that optimal progressivity is inversely related to the compensated elasticity of taxable income which, in the Atkinson-Stiglitz model, is equivalent to the compensated elasticity of labor supply.

In the model presented here, the critical elasticity may depend on the choice of m, so that the optimal progressivity of the tax system cannot be determined independently of m. To pursue this idea further, we can write the first-order condition

the identity $I(t, T, m, y; E(t, T, m, y, u)) = I^c(t, T, m, y; u)$, where $E(\cdot)$ is the expenditure function, holds. Differentiating this identity with respect to the tax rate t, and observing that $\frac{\partial E(\cdot)}{\partial t} = I$, leads to the Slutsky equation in the text. Notation I and I^c distinguishes uncompensated taxable income from compensated taxable income, the difference being that one is a function of exogenous income while the other is a function of the utility level.

for m, expression (6), as

$$\int \left(\frac{v_m}{\mu} g + t \frac{\partial I^c}{\partial m} \right) dF(y) = A'(m)\,, \tag{8}$$

where $\frac{\partial I^c}{\partial m}$ refers to the compensated response of taxable income.[6] It is insightful to further decompose the impact of m on compensated taxable income.[7] To do that, notice that, for any t^*, $t^* I^c(t^*, T, m, y; u) = \int_0^{t^*} \left(t \frac{\partial I^c}{\partial t} + I^c \right) dt = \int_0^{t^*} I^c (1 - \frac{t\varepsilon}{1-t})\, dt$. Increasing the tax rate by dt increases revenue by $I^c dt$, but it also stimulates a substitution response that reduces taxable income by $\frac{dt}{1-t}\varepsilon$ percent for each percent increase in the tax rate. Differentiating this formula with respect to m we can write,

$$t^* \frac{\partial I^c}{\partial m} = \frac{\partial}{\partial m} \left(\int_0^{t^*} I^c (1 - \frac{t\varepsilon}{1-t})\, dt \right) = \int_0^{t^*} \frac{\partial I^c}{\partial m} \left(1 - \frac{t\varepsilon}{1-t} \right) dt - \int_0^{t^*} \frac{t I^c}{1-t} \frac{\partial \varepsilon}{\partial m}\, dt\,. \tag{9}$$

The first term of this expression represents the direct effect on tax revenue of changing m. The second term represents the effect on revenue of decreasing the compensated elasticity of taxable income. The impact of this change is positively related to the tax rate.

First-order condition (8) can now be written as,

$$\int \left(\frac{v_m}{\mu} g + \int_0^{t^*} \frac{\partial I^c}{\partial m} \left(1 - \frac{t\varepsilon}{1-t} \right) dt - \int_0^{t^*} \frac{t I^c}{1-t} \frac{\partial \varepsilon}{\partial m}\, dt \right) dF(y) = A'(m)\,. \tag{10}$$

Expression (10) says that, at the optimum, the benefit of increasing m should equal the marginal administrative cost. The benefit is the sum of the direct effect on social welfare (denominated in dollars), the revenue effect holding the elasticity constant, and the benefit of changing the elasticity. Note that all expressions are evaluated at the constant utility level (in other words, integration is along the indifference curve).

The last term of expression (10), $-\int \int_0^{t^*} \frac{t I^c}{1-t} \frac{\partial \varepsilon}{\partial m}\, dt\, dF(y)$, is of particular note. The value to the social planner of a policy instrument is higher (holding its other effects constant), the more an instrument decreases the compensated elasticity of taxable income. Furthermore, the benefit of lowering the elasticity increases with the tax rate and with taxable income. Thus, the more egalitarian is the social objective function (which manifests itself in higher progressivity), the greater is the benefit from using this instrument, because the reduction of the compensated taxable income elasticity is more valuable. This suggests that the elasticity at the optimum, the optimal elasticity if you will, should be lower in more egalitarian societies. We return to this issue in section 3.

[6] In deriving expression (8), we make use of the Slutsky-type equation for the effect of m: $\frac{\partial I}{\partial m} = \frac{\partial I^c}{\partial m} + \frac{\partial I}{\partial E} \frac{v_m}{\mu}$.

[7] Although the decomposition that we consider concerns compensated quantities, a similar approach may be applied to quantify the benefit of a change in the uncompensated elasticity.

An interesting special case occurs when the elasticity does not depend on the marginal tax rate.[8] We can then show that $\int_0^{t^*} \frac{tI^c}{1-t}\,dt = \varepsilon^{-1}\left(\int_0^{t^*} I^c\,dt - t^*I^c\right)$. Interpreting I^c as the demand for taxable commodities and t as their price, the term in brackets can be thought of as the standard (i.e., not including $A(m)$) deadweight loss from taxation: it is (minus) the difference between the amount collected in distortionary taxes and the amount that a "discriminating" (i.e., one that can impose a person-specific, arbitrary non-linear marginal tax schedule) tax authority could collect, with demand evaluated at the constant utility level. Then the benefit of reducing the taxable income elasticity (the third term of the left-hand side of equation (10)) is

$$-\int\int_0^{t^*} \frac{tI^c}{1-t}\frac{\partial\varepsilon}{\partial m}\,dt\,dF(y) = \varepsilon_m \int\left(\int_0^{t^*} I^c\,dt - t^*I^c\right)\,dF(y)\,, \tag{11}$$

where $\varepsilon_m = -\frac{\partial\varepsilon}{\partial m}/\varepsilon$ is the proportionate change in elasticity resulting from an increase in m, defined to be a positive number. Equation (11) says that the value to the social planner of reducing the compensated taxable income elasticity is proportional to the excess burden.

Although this approach provides a clear intuition for the benefit of a change in the elasticity, it may be difficult to calculate in practice, because all terms are evaluated at a constant utility level. In the appendix, an alternative expression is derived with the benefit of a change in elasticity evaluated at a constant income level. The benefit of a change in elasticity can be shown to have a similar form as before:

$$-\int\int_0^{t^*} \frac{tI}{1-t}\frac{\partial\varepsilon}{\partial m}\,dt\,dF(y)\,, \tag{12}$$

but the integration takes place along the budget line instead of an indifference curve.

The model presented in this section isolates in a very general setting the role of "elasticity control" in an optimal tax system. However, the generality of the instrument m makes it difficult to establish insightful propositions about the expanded optimal progressivity problem, although it suggests likely relationships between the optimal taxable income elasticity and the government's goals. Practical examples of m range from the extent of audit coverage, to the definition of a pass-through entity, to broadening the tax base. In addition to its effect on the elasticity of taxable income, any instrument has both efficiency and distributional impacts. To gain further insight, we next examine the optimal elasticity question in a particular setting in which the elasticity is determined by how broadly the tax base is defined.

[8] Even if this is not literally true, it may be a useful approximation if the effect of m on elasticity is much greater than the effect of t. It may also be a useful benchmark for empirical applications, where it is conventional to assume that elasticities do not depend on prices.

3 A Model of Optimal Progressivity and the Optimal Tax Base

The structure of the problem is similar to the general one described before. The government must raise revenue of R, and can choose three instruments: a lump-sum tax T, a marginal tax rate t, and a third instrument that, *inter alia*, affects the elasticity of taxable income. Rather than being very general about this third instrument as we were in the previous section, we are now quite specific: it is the breadth of the tax base. More details follow.

3.1 The Consumer's Problem

Consider an economy in which there is a continuous number of commodities. Commodities are indexed by a number $i \in [0,1]$, and they are denoted by C_i. The bundle of commodities is written as C ($C : [0,1] \to R_+$). In other words, a bundle of consumption goods is a "vector" containing a continuous number of elements.[9]

Each individual is characterized by a wage rate of y, and provides inelastically one unit of labor. She maximizes a Cobb-Douglas utility function given by $u(C) = \exp(\int_0^1 \eta_i \ln C_i \, d\, i)$, where η_i is the weight of good i.[10] Weights are normalized so that $\int_0^1 \eta_i \, di = 1$.

Pre-tax prices of all commodities are set equal to one. The consumer faces a tax system characterized by three parameters: t, T and m, where m, $0 \leq m \leq 1$ denotes the size of the tax base. In a consumption tax framework, this implies that commodity i is taxed as long as $i < m$. Equivalently, in an income tax framework it implies that spending on commodity $i \in [m,1]$ is deductible from taxable income. In a single-period framework like this, these two interpretations are equivalent.

For future reference, following Wilson (1989) and Yitzhaki (1979) we introduce the notation $\gamma(m) \equiv \int_0^m \eta_i \, di$, where $\gamma(m)$ is the expenditure share corresponding to a tax base of m. Note that $\gamma(m)$ is a monotone transformation of m, and as such it

[9] We assume that the consumption space includes only measurable functions, and we employ Lebesgue integrals in what follows.

[10] The setup of our model is closely related to the models of Yitzhaki (1979) and Wilson (1989). Wilson (1989) considered a more general CES function. In an earlier version of this paper we generalize the results that follow to the CES case. Most of the propositions in this paper go through, but some of the proofs are significantly more complicated. The main source of complexity is that, in the CES case, the compensated elasticity of taxable income is a function of the tax rate. One result that is affected is our corollary 1, which must be qualified to state that more egalitarian societies have a lower optimal elasticity of taxable income to the extent that the elasticity of substitution is less than or equal to one. When it is greater than one, it is still true that the tax base increases with egalitarianism (acting to reduce the elasticity), but an adjustment in the tax rate may affect the elasticity in the opposite direction. In the earlier version, we present an example in which the net effect is an increase in the elasticity.

may equivalently be used as a policy instrument. The parameter t is the marginal tax rate imposed on taxable income; we will also use the notation $\tau \equiv 1-t$ to denote the net-of-tax rate. The parameter T is the lump-sum tax. To facilitate the analysis, with no loss of generality, we use the convention that the lump-sum tax is deductible from taxable income, so that taxable income is equal to $I = y - \int_m^1 C_i\,di - T$. Consequently, the budget constraint of an individual is

$$\int_0^1 C_i\,di = y - T - t(y - \int_m^1 C_i\,di - T)\,. \tag{13}$$

The utility function is Cobb-Douglas, so that the solution to the consumer's problem is

$$C_i = \begin{cases} (1-t)\eta_i(y-T) & i < m \\ \eta_i(y-T) & i \geq m\,, \end{cases} \tag{14}$$

and thus taxable income is equal to $I = y - T - \int_m^1 \eta_i(y-T)\,di = \gamma(y-T)$. Substituting the optimal choices of C_i into the utility function we can also derive the indirect utility function: $v(\tau, \gamma, y-T) = \alpha(y-T)\tau^\gamma$, where α is a constant. Henceforth, we normalize α to be one.

Finally and critically, but demonstrated in the appendix, the compensated elasticity of taxable income with respect to the net-of-tax rate, denoted ε, is simply $1-\gamma$. Thus, with these assumptions, the choice of tax base immediately determines the central elasticity in the model. Intuitively, the larger is the set of untaxed commodities to which the consumer can turn to in the face of taxation, the larger is the elasticity.[11]

3.2 The Government's Problem

The government's problem is to choose t, T, and γ to maximize the social welfare function, subject to its budget constraint. To be precise, its problem is

$$\max \int w(v(t, \gamma, y, T))\,dF(y) \tag{15}$$

subject to

$$T + t\gamma(Y - T) = A(\gamma) + R\,, \tag{16}$$

where $Y = \int y\,dF(y)$. The choice of the tax base determines the administrative cost $A(\gamma)$. Recall that all individuals maximize a homogeneous utility function, and following Wilson (1989, p. 1199) we assume that, for given γ, those commodities which minimize A are included in the tax base, so that $A'(\gamma) > 0$ and $A''(\gamma) > 0$.[12]

[11] This is also true in the CES case, where this elasticity is equal to $\varepsilon = \sigma(1 - \frac{\gamma}{\gamma + (1-\gamma)\tau^{1-\sigma}})$, σ being the elasticity of substitution. This expression is decreasing in γ.

[12] The assumption $A'(\gamma) > 0$ was introduced by Yitzhaki (1979), who provided a simple justification for this shape. We make this assumption for convenience, but it is not required for our

Maximization of (15) subject to (16) leads to the following first-order conditions with respect to T, t and γ, respectively,

$$\int \left[-w'\tau^{\gamma} + \lambda\left(1 - t\gamma\right)\right] dF(y) = 0, \tag{17}$$

$$\int \left[- w'\tau^{\gamma-1}\gamma(y-T) + \lambda\gamma(y-T)\right] dF(y) = 0, \qquad \text{and} \tag{18}$$

$$\int \left[\tau^{\gamma}\ln(\tau)w'(y-T) + \lambda t(y-T)\right] dF(y) = \lambda A'(\gamma). \tag{19}$$

We will denote by $X = t\gamma \int (y-T)\, dF(y) = t\gamma(Y-T)$ the revenue from distortionary taxes, so that the budget constraint becomes $T + X = A(\gamma) + R$. Given the revenue requirement of R, if X and γ are known, the value of T can be easily found. For this reason any point in (γ, X) space represents a unique tax system satisfying the budget constraint. We will concentrate on X as a measure of progressivity, because in this model progressivity depends not only on t, but on γ as well.

The standard, or what we will call the Mirrlees, optimal income taxation problem can be thought of as solving for the optimal t and T , for a given tax base of γ. Its solution is characterized by the first-order conditions (17) and (18), plus the budget constraint. Solving this problem for all values of γ defines a curve in (γ, X) space that we will refer to as a *Mirrlees curve*. It determines the optimal progressivity for a given tax base or, equivalently, for a given compensated elasticity of taxable income.

Contrast this to the problem considered by Yitzhaki (1979) and Wilson (1989), who take the revenue requirement and the lump-sum portion of the tax system (T) as given, and solve for the optimal γ and t. The solution to this problem is characterized by first-order conditions (18) and (19), plus the budget constraint. Solving this problem for all values of T defines a curve in (γ, X) space that we will refer to as a *Yitzhaki-Wilson* (Y-W) *curve*. It determines the optimal tax base for a given degree of progressivity.

Note that using (18) to get an expression for λ, and substituting for λ in (19) yields:

$$A'(\gamma) = \left(t + \tau\ln\tau\right)\left(Y - T\right). \tag{20}$$

This expression is almost identical to Yitzhaki's equation 5, and is a special case of Wilson's equation 18, that characterize the optimal tax base. The only difference is

results. In particular we know that, as long as a global optimum exists, administrative costs must be increasing in its neighborhood — otherwise one could extend the tax base and decrease both distortions and administrative costs. Consequently, without this assumption, our propositions should be read as describing the situation in the neighborhood of the optimum only.

We recognize, but do not address, the possibility that, depending how the tax system is administered, $A(1)$ could be less than $A(\gamma)$, $\gamma < 1$. In other words, a completely broad base may be less costly to administer than one with exclusions.

that the term $Y - T$ replaces y in order to account for multiple individuals and the lump-sum tax. Together with the budget constraint, expression (20) characterizes the optimal choice of γ and X (or t), given T. The right-hand side of this expression is the marginal social benefit of expanding the tax base, which is equal to the marginal reduction in the excess burden from raising a given amount of revenue with a larger tax base. The left-hand side is simply the marginal resource cost of expanding the tax base. In the models of Yitzhaki and Wilson, the amount of revenue to be raised by distortionary taxes is fixed, so that the only problem is whether to use a narrow-base, high-rate system, or a broader-base, lower-rate system. In the problem considered here, though, the government can also reduce distortionary taxes (and, therefore, redistribution) by making greater use of lump-sum taxes. Because with a Cobb-Douglas (or, more generally, a CES) utility function the income elasticity of all goods is exactly one, there are no distributional implications to whether a given amount of distortionary tax is raised via a narrow or broad base; thus, for a given R the value of X is a sufficient statistic for evaluating the tax system's progressivity, as is the value of T.

Equation (20) can be decomposed as in Section 2. Recall that the social benefit of a reduction in elasticity is given by $-\int \int_0^{t^*} \frac{tI}{1-t} \frac{\partial \varepsilon}{\partial \gamma} \, dt \, dF(y)$. Substituting expressions for I and ε, and integrating, this benefit becomes $\gamma(Y-T) \int_0^{t^*} \frac{t}{1-t} \, dt = -\gamma(\ln(\tau) + t)(Y-T)$. It is straightforward to show that this expression is positive and increasing in both γ and t. Consequently, without this effect the benefit of an increase in γ (the right-hand side of expression (20)) would be smaller, and the incentive to extend the tax base to meet the egalitarian goals would be weaker.

How important is the role of a change in the elasticity? To quantify it, we calculate the share of this benefit in the total benefit of a change in γ, the right-hand side of equation (20),

$$\frac{-\gamma(t + \ln(\tau))(Y-T)}{(t + \tau \ln(\tau))(Y-T)} = \gamma B(t) \,, \tag{21}$$

where $B(t) = \left(\frac{t \ln(1-t)}{t + \ln(1-t)} - 1 \right)^{-1}$. One can show that $\partial B / \partial t > 0$, $B(0) = 1$, and $\lim_{t \to 1} B(t) = +\infty$. Consequently, for small values of t, the share of the benefit due to the change in the elasticity is approximately γ. For sufficiently large t (but less than one), we have that $\gamma B(t) > 1$, meaning that the benefit of a change in the elasticity exceeds the total benefit, i.e., that the sum of all other effects is a net loss. Consequently, it is only the change in the elasticity that makes an increase in γ worthwhile (note that the total benefit, the right-hand side of (20), is always positive). This should not be too surprising: if the elasticity does not change, there would be limits to redistribution so that, at some point, increments to progressivity achieved by increasing γ would reduce welfare.

We are now ready to show the relationship between the social welfare function

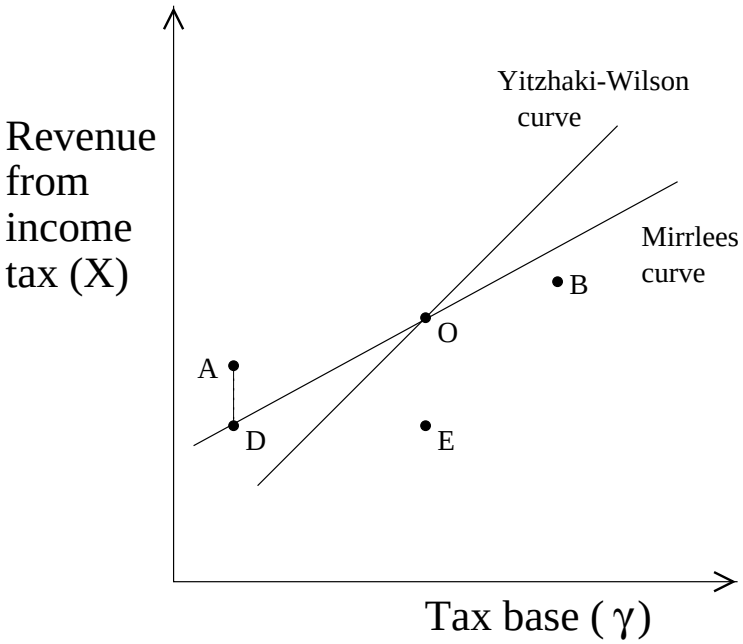

Figure 1: The Mirrlees and Yitzhaki-Wilson curves and the Optimal Tax System

and the optimal tax system.

3.3 Results

In the problem we consider all of the policy variables, t, T, and γ, are endogenous. Thus, the global optimum lies at the intersection of Mirrlees and Yitzhaki-Wilson curves, determining both the progressivity of the tax system and the tax base (which determines the elasticity). This is illustrated in Figure 1.

We seek to understand how the optimal tax system depends on the key aspects of the environment. The first of our propositions depends on the following lemma:[13]

Lemma 1 *The Yitzhaki-Wilson curve is upward sloping.*

We can now state our first proposition as

Proposition 1 *A more egalitarian society will feature a larger tax base (and a more progressive tax system).*

Notice that the Y-W curve does not depend on the social welfare function: it is defined by the budget constraint and equation (20), and both of these are invariant

[13] All proofs are in the appendix.

to changes in the social welfare function.[14] Consequently, an increase in progressivity corresponds to a shift along the Y-W curve. Thus, either both progressivity and the tax base increase or they both decrease.

It is not surprising that more egalitarian societies have a more progressive tax structure. However, in this model, the social planner can increase progressivity by either increasing the marginal tax rate or by increasing the tax base. Both policy changes affect tax revenue in a similar way, but increasing the tax base reduces distortions (and increases administrative costs). Our previous analysis demonstrates that extending the tax base is more than just a complementary way of making taxes more progressive. By reducing the strength of the behavioral response, an increase in the tax base makes a higher marginal tax rate more effective in achieving egalitarian goals. Consequently, without this effect, the benefit of an increase in the tax base (the right-hand side of equation (20)) would be smaller, and it is not clear that the tax base would have to increase at all.

Corollary 1 *More egalitarian societies have a lower compensated elasticity of taxable income.*

Proof: The compensated elasticity of taxable income is equal to $1 - \gamma$, and the previous proposition established that γ is higher in more egalitarian societies. □

Our second proposition depends on the three lemmas which also underlie the pattern of curves in Figure 1. These results assume an isoelastic social welfare function.

Lemma 2 *The Mirrlees curve is upward sloping.*

Lemma 3 *An increase in egalitarianism shifts the Mirrlees curve up.*

Lemma 4 *Near the intersection of the two curves corresponding to the global optimum, the Yitzhaki-Wilson curve has a larger slope.*

Note that these lemmas allow for a diagramatic description of the experiment in Proposition 1. An increase in egalitarianism shifts the Mirrlees curve up and, given the way the curves intersect, it leads to higher progressivity and a broader tax base.

Proposition 2 *An increase in the cost of maintaining a broad tax base (i.e., $A'(\gamma^*)$ increasing while $A(\gamma^*)$ stays constant) will decrease optimal progressivity (as well as the tax base).*

[14] A change in the social welfare function does not affect the tradeoff between t and γ. Because of the homogeneity of the utility function, neither of these instruments has distributional implications.

This is the complement to Proposition 1. In this case, changing the marginal cost of administration does not alter the solution to the Mirrlees problem, so that the optimum (γ^*, X^*) lies on the original Mirrlees curve. The changing administrative cost does, though, shifts the Y-W curve, i.e., it changes the optimal mix of t and g to raise a given amount of distortionary tax. In fact, it must shift upward, so that for a given γ, the optimal X is higher. Thus, the new optimum features a smaller tax base, which is not surprising, as well as a less progressive tax system, because of the higher cost of collecting those revenues.

Corollary 2 *A decrease in the cost of maintaining a broad tax base leads to a lower elasticity of taxable income.*

This corollary is immediate, because the elasticity of taxable income moves inversely with the tax base.

In a sense, corollary 2 is parallel to corollary 1; in the latter the benefit of extending the tax base increases, and in corollary 2 the cost decreases. Both situations naturally lead to a broader tax base being optimal. However, corollary 1 is more fundamental, because when egalitarianism increases a change in the compensated elasticity is a crucial part of the response of the social planner. Consequently, if this effect were not present, the optimal tax base would not have to increase. In the present case, as long as the optimum is interior (as we assume throughout the paper), the instrument that becomes cheaper will be used more, regardless of its effect on the elasticity (although exactly how much more it will be used depends on this effect).

Figure 1 and the model that underlies it may also be used to demonstrate the danger of designing tax policy without regard for whether the existing elasticity of taxable income is appropriate. Consider an initial tax system corresponding to point A. Based on the elasticity of taxable income for that tax base, the government correctly decides to choose the tax system represented by point D, that is, to reduce progressivity. However, the global optimum O corresponds to a *higher* level of progressivity than point A.[15] Even if one recognizes that other instruments are available, but fails to recognize their effect on the taxable income elasticity, the results may still be misleading. The perceived Mirrlees curve would be flatter, but still crossing point D, because the elasticity would be assumed to be unaffected by γ, and thus the perceived cost of progressivity at higher levels of γ would be higher than it is in reality. The perceived Yitzhaki-Wilson curve would be to the left of the actual one, because the benefit of a change in g would be underestimated. The perceived global optimum would correspond to lower progressivity and a lower tax base than the actual one, and the corresponding level of progressivity could be lower than at A.

[15]For completeness, observe that it is also possible that the level of progressivity suggested by the standard analysis may be too high, for example if one starts at a point such as B.

The choice of the income or commodity tax base is an illustration of how policy-makers have some degree of control over the elasticity of taxable income to the tax rate. While the details of other examples may certainly differ from what we have presented here, the underlying intuition will remain.

4 Implications

The optimal progressivity of the tax system depends inversely on the compensated elasticity of taxable income with respect to the net-of-tax rate. That elasticity is not entirely an immutable function of consumer preferences and production technologies, but is subject to manipulation by government policy with respect to tax administration and enforcement as well as the choice of tax base.

The notion that the elasticity of taxable income can be chosen via the setting of certain policy instruments has noteworthy implications both for policy and the empirical analysis of taxation. Although Section 3 focused on cross-jurisdictional implications, it is of interest to speculate about its implications for a single country's tax policy over time. For example, consider its implications for the last major reform in the United States, the Tax Reform Act of 1986 (TRA). TRA lowered statutory tax rates and at the same time broadened the tax base and restricted the use of tax shelters, both of which arguably reduced the elasticity of taxable income. This is consistent with the fact that empirical estimates of taxable income elasticities based on the behavioral response to TRA (Feldstein 1995, Moffitt & Wilhelm 1998, Auten & Carroll 1999) are larger than those based on the behavioral response to subsequent (1990 and 1993) tax rate changes (Carroll 1998). Although differences in methodology and offsetting biases due to unobserved trends may be partly responsible for the decrease, this paper raises the possibility that the base broadening provisions of TRA reduced the *true* taxable income elasticity between the two periods. More troublingly, this paper raises doubt about longitudinal analyses of tax reforms which feature changes in both tax rates and (in the case of TRA, as an objective) changes in the elasticity of taxable income. Comparing pre- and post-reform behavior to identify *the* elasticity is problematic, to say the least, because the elasticity after the reform is different from the elasticity before the reform.

If one takes seriously the idea that actual tax policy decisions reflect a solution to a maximization problem like that posed here, many intriguing predictions arise. For example, one could argue that, once the Tax Reform Act of 1986 was passed and policy makers became aware that the taxable income elasticity (and the marginal social cost of increasing progressivity) had declined, they appropriately increased the progressivity of the tax system by raising the top income tax rates in 1990 and 1993.[16] We are aware that many factors are at work in producing tax legislation,

[16]In terms of Figure 1, TRA moved the income tax system from D to E, and the top rate

and we are certainly not suggesting that this factor dominates the others, but we are raising the possibility that a policy maker who controls the elasticity can affect the resolution of the progressivity issue. The principal message to policy makers is to be wary of decisions about progressivity based on the existing elasticity of taxable income: this elasticity is itself subject to control, and may be suboptimal. Progressivity, and all of the other aspects of a tax system, must be jointly optimized.

Another natural research agenda is to test the cross-jurisdictional implications of this model. For example, exogenous measures of egalitarian social preferences should be positively correlated not only with effective tax progressivity, but also the breadth of the tax base, and be negatively correlated with the observed elasticity of taxable income. The simultaneous presence of multiple taxes — such as income tax and value added tax — means that the underlying normative model will have to be extended in straightforward ways.

increases of 1990 and 1993 moved it from E to O.

A Appendix

Alternative expression for the benefit of a change in the elasticity Write the first-order condition (18) as $\int \left[\frac{\omega' v_m}{\lambda} + t\frac{\partial I}{\partial m}\right] dF(y) = A'(m)$. Similarly, as in the decomposition in the text, $t^*I = \int_0^{t^*} \left(t\frac{\partial I}{\partial t} + I\right) dt$, but because it is uncompensated taxable income that we deal with here, the integration takes place at a constant income level. Using the Slutsky equation and rearranging leads to $t^*I = \int_0^{t^*} I\left(1 - t\frac{\partial I}{\partial E} - \frac{t}{1-t}\varepsilon\right) dt$. Differentiating this with respect to m and substituting into the first-order condition yields,

$$\int \left\{ \frac{\omega' v_m}{\lambda} - \frac{\partial}{\partial m}\int_0^{t^*} tI\frac{\partial I}{\partial E}\, dt \;+\; \int_0^{t^*} \frac{\partial I}{\partial m}\left(1 - \frac{t}{1-t}\varepsilon\right) dt \right. \\ \left. - \int_0^{t^*} \frac{tI}{1-t}\frac{\partial \varepsilon}{\partial m}\, dt \right\} dF(y) = A'(m)\,. \qquad (1)$$

The last term on the left-hand side is the social benefit of a change in the elasticity. The third term is the benefit, holding elasticity constant. The second term accounts for a change in income, while the first is the direct effect on welfare. When the elasticity is constant, the benefit of a change in elasticity may be expressed as

$$-\int_0^{t^*} \frac{tI}{1-t}\frac{\partial \varepsilon}{\partial m}\, dt = \varepsilon_m \left(\int_0^{t^*} I\, dt - t^*I - \int_0^{t^*} tI\frac{\partial I}{\partial E}\, dt \right). \qquad (2)$$

The term in the brackets on the right-hand side is the deadweight loss of taxation, corrected by potential revenue effects.

Derivation of the compensated elasticity of taxable income. Standard Slutsky equations for individual commodities are

$$\left.\frac{\partial C_i}{\partial \tau}\right|_u = \frac{\partial C_i}{\partial \tau} - \frac{\partial C_i}{\partial E} I\,, \qquad (3)$$

where $\frac{\partial C_i}{\partial \tau}|_u$ is the compensated response of C_i and E denotes exogenous (non-taxable) income. Since, by assumption, the lump-sum tax T affects taxable income, its effect on C_i is not the same as the effect of E, but they are linked through the relation $\frac{\partial C_i}{\partial T} = -\tau\frac{\partial C_i}{\partial E}$. Using these relations, $\frac{\partial I}{\partial \tau}|_u = -\int_m^1 \frac{\partial C_i}{\partial \tau}|_u\, di = I(1-\gamma)\tau^{-1}$, and the compensated elasticity of taxable income is $\epsilon = \frac{\tau}{I}\frac{\partial I}{\partial \tau}|_u = 1 - \gamma$.

The Yitzhaki-Wilson (Y-W) curve is upward sloping with $\frac{d\gamma}{dT} > 0$ and $\frac{dX}{dT} > 0$. The Y-W curve is defined by the budget constraint and equation (20). The budget line may be used to define $t(\gamma, T) = \frac{A(\gamma)+R-T}{\gamma(Y-T)}$. Differentiating this expression and then simplifying yields $\frac{\partial t}{\partial T} = -\frac{1-t\gamma}{\gamma(Y-T)} < 0$ and $\frac{\partial t}{\partial \gamma} = \frac{A'(\gamma)-t(Y-T)}{\gamma(Y-T)} < 0$. (To see that the numerator is negative, observe that it is equal to $\tau \ln(\tau)(Y-T)$ from equation (20)). Total differentiation of equation (20) (with $t = t(\gamma, T)$) yields $\frac{d\gamma}{dT} =$

$-\frac{t+\tau\ln(\tau)+(Y-T)\ln(\tau)\frac{\partial t}{\partial T}}{A''(\gamma)+(Y-T)\ln(\tau)\frac{\partial t}{\partial \gamma}}$ (it is negative because both the numerator and denominator are positive). Finally, from the budget constraint, $X = R + A(\gamma) - T$, so that $\frac{dX}{dT} = A'(\gamma)\frac{d\gamma}{dT} - 1 < 0$.

Proof of Proposition 1: An increase in egalitarianism increases X and γ. Consider increasing the concavity of the welfare function, i.e., replacing w by $f(w)$, where $f' > 0$, $f'' < 0$. The global optimum moves along a constant Y-W curve (because neither the budget constraint nor equation (20) are affected) and, given that it is upward sloping, there are two possible cases: both X and γ increase, while T falls, or both X and γ decrease, while T increases.

To prove the proposition, we need to show that the second case is not possible. Suppose, on the contrary, that it was the case. Let $v_1(y)$ and $v_2(y)$ represent the level of utility of person y with the optimal tax system under the initial and modified social welfare function respectively. There must be some individuals who are better off under one system, and some who are better off under the alternative, or else one of the these tax systems could not be optimal for any social welfare function.

Recall that the indirect utility function is $v_1(y) = (y-T)\tau^\gamma$. The case that we are considering is an increase in T, accompanied by a fall in γ and an adjustment in t so that X falls. Observe that τ^γ needs to go up, or else everybody would be worse off as T increases. It follows that the difference between the utility of an individual before and after this change is an increasing function of y, i.e., $v_2(y) - v_1(y)$ is increasing in y. In particular, it means that individuals at the the bottom of the distribution are worse off, while individuals at the top are better off. As a result there must exist y^* such that $v_2(y^*) = v_1(y^*)$, while for all $y < y^*$ we have $v_2(y) < v_1(y)$ and for $y > y^*$ it is true that $v_2(y) > v_1(y)$. The same inequalities hold when we compare $w(v_1(y))$ and $w(v_2(y))$. The concavity of f implies that

$$\begin{aligned} f(w(v_2(y))) &< f(w(v_1(y))) + f'(w(v_1(y)))\,[w(v_2(y)) - w(v_1(y))] \\ &= f(w(v_1(y))) + f'(w(v_1(y^*)))\,[w(v_2(y)) - w(v_1(y))] \\ &+ [f'(w(v_1(y))) - f'(w(v_1(y^*)))]\,[w(v_2(y)) - w(v_1(y))]\,. \end{aligned} \tag{4}$$

Integrating this expression yields

$$\begin{aligned} &\int f(w(v_2(y)))\,dF(y) \\ &\quad< \int f(w(v_1(y)))\,dF(y) + f'(w(v_1(y^*)))\int \Big(w(v_2(y)) - w(v_1(y))\Big)\,dF(y) \\ &\quad+ \int \Big[f'(w(v_1(y))) - f'(w(v_1(y^*))\Big]\Big[w(v_2(y)) - w(v_1(y))\Big]\,dF(y)\,. \end{aligned} \tag{5}$$

The second term is negative because $\int w(v_1(y))\,dF(y)$ is the maximized welfare under the initial welfare function. Notice that for $y < y^*$, we have $v_2(y) < v_1(y)$ and

$f'(w(v_1(y))) > f'(w(v_1(y^*)))$,[17] and both inequalities change their signs for $y > y^*$; thus, the third term in negative.

Thus, $\int f(w(v_2(y)))\,dF(y) < \int f(w(v_1(y)))\,dF(y)$. This is a contradiction, however, because the left-hand side is the maximized welfare under the modified social welfare function. Thus, the second case is not possible.

The Mirrlees curve is upward sloping. Based on the first-order conditions (17) and (18) we define two terms: $MCT^+(T, X(t,T,\gamma);\gamma) \equiv (1-t\gamma)^{-1}\int w'\tau^\gamma\,dF(y)$, and $MCX(T, X(t,T,\gamma);\gamma) \equiv (Y-T)^{-1}\int w'\tau^{\gamma-1}(y-T)\,dF(y)$. MCX is the social cost, in utils, per dollar of revenue raised by increasing the marginal tax rate, while holding constant the revenue from the lump-sum tax. MCT^+ is the marginal social cost, in utils, per dollar of tax revenue raised by changing the lump-sum tax. Note that, because of the income effect on revenue, increasing revenue by one dollar requires raising T by $\frac{1}{1-t\gamma}$ dollars, and causes a reduction in the revenue from distortionary taxes of $\frac{t\gamma}{1-t\gamma}$. We can combine these two conditions to get the marginal cost of increasing the revenue from lump-sum taxes by a dollar, while holding the revenue from distortionary taxes constant. We denote it as $MCT(T, X(t,T,\gamma);\gamma)$, which equals $MCT^+(1-t\gamma) + MCX\,t\gamma$.

At the optimum MCT is equal to both MCX and MCT^+, because the marginal social cost of collecting an additional dollar by any method must be the same. This is just a restatement of the first-order conditions (17) and (18).

Now consider a perturbation in the given value of γ. Totally differentiating the first-order condition $MCT = MCX$ and the budget constraint, we can find expressions for $\frac{\partial X}{\partial\gamma}$ and $\frac{\partial T}{\partial\gamma}$, as follows

$$\frac{\partial X}{\partial\gamma} = \frac{MCT_\gamma - MCX_\gamma}{SOC} + A'(\gamma)\frac{MCT_T - MCX_T}{SOC} \tag{6}$$

$$\frac{\partial T}{\partial\gamma} = \frac{MCX_\gamma - MCT_\gamma}{SOC} + A'(\gamma)\frac{MCX_X - MCT_X}{SOC}\,, \tag{7}$$

where SOC (which, by the second-order condition, must be positive) equals $MCX_X - MCX_T - MCT_X + MCT_T > 0$. These formulae decompose the total effect of a change in γ on X and T into substitution and revenue effects. The former refers to the fact that a change in γ alters the relative attractiveness of T and X as ways to raise revenue; the latter refers to the fact that increasing γ has a social cost of A', which has to be financed by T and X. Intuitively, an increase in γ reduces the elasticity of taxable income and lowers the relative cost of distortionary taxation. As a result, one expects that it induces substitution from the lump-sum tax to the distortionary taxes. In order to formally sign these effects, it will prove useful to consider the ratio $\frac{MCT}{MCX}$, which equals $\Psi\tau + t\gamma$, where $\Psi \equiv \frac{\int w'\,dF(y)\int(y-T)\,dF(y)}{\int(y-T)w'\,dF(y)}$ is a

[17] This is implied by the concavity of f, and v_1 increasing in y.

measure of the "distributional characteristic" of the tax system.[18] It is higher the more concave is the social welfare function, and the more unequal is the distribution of abilities. We will make use of two properties of Ψ:

Property 1 *A monotone, concave transformation of the social welfare function increases* Ψ.

Proof: Let the original welfare function be w and the transformed one be $g(w)$ where $g' > 0$, $g'' < 0$. To evaluate how Ψ changes, it is enough to compare $\frac{\int w'(y-T)\,dF(y)}{\int w'\,dF(y)}$ and $\frac{\int g'w'(y-T)\,dF(y)}{\int g'w'\,dF(y)}$. Defining $W(x) = \frac{\int_0^x w'\,dF(y)}{\int w'\,dF(y)}$ and $G(x) = \frac{\int_0^x g'w'\,dF(y)}{\int g'w'\,dF(y)}$, these expressions can be written as $\int (y-T)\,dW(y)$ and $\int (y-T)\,dG(y)$. We also obviously have $\lim_{y\to\infty} W(y) = 1$ and $\lim_{y\to\infty} G(y) = 1$. Concavity and monotonicity of g implies that $g'(y_1) > g'(y_2)$ if and only if $y_1 < y_2$. It follows that $\int_0^x g'w'\,dF(y) > g'(x)\int_0^x w'\,dF(y)$ and $\int_x^\infty g'w'\,dF(y) < g'(x)\int_x^\infty w'\,dF(y)$, so for any x:

$$\begin{aligned} G(x) &= \frac{\int_0^x g'w'\,dF(y)}{\int_0^x g'w'\,dF(y) + \int_x^\infty g'w'\,dF(y)} = \left(1 + \frac{\int_x^\infty g'w'\,dF(y)}{\int_0^x g'w'\,dF(y)}\right)^{-1} \\ &> \left(1 + \frac{g'(x)\int_x^\infty w'\,dF(y)}{g'(x)\int_0^x w'\,dF(y)}\right)^{-1} = \frac{\int_0^x w'\,dF(y)}{\int_0^x w'\,dF(y) + \int_x^\infty w'\,dF(y)} = W(x)\ . \end{aligned} \quad (8)$$

This condition says that $W(x)$ first-order stochastically dominates $G(x)$ (although the context is different). Thus $\int (y-T)\,dW(y) > \int (y-T)\,dG(y)$, which directly implies the desired result. □

To demonstrate the next property, we will use the following notation: $r \equiv y - T$, $p \equiv \tau^\gamma$, so that, the indirect utility of individual is $v(t, T, \gamma; y) = rp = p(y - T)$. Notice that p and T are the same for all individuals. Using this notation, $\Psi(p, T) = \frac{\int r \int w'(pr)\,dF(y)}{\int w'(pr)r\,dF(y)}$. Define also $\xi(x) \equiv \frac{w''(x)}{w'(x)}x$.

Property 2 *For any isoelastic welfare function,* $\frac{\partial \Psi}{\partial T} > 0$.

Proof: Consider a change in T of dT and notice that $w(p(z - dT)) = g(w(pz))$ where $g(x) = w(h(x) - p\,dT)$ and $h(x) = w^{-1}(x)$. Using the properties of the inverse function that $w'(h(x))h'(x) = 1$ and $w''h'h' + w'h'' = 0$, we can show that $g'(x) = w'h' > 0$ and $g''(x) = w''(h(x) - p\,dT)h'(x)^2 + w'(h(x) - p\,dT)h''(x)$. Substituting for h'', $g'' = w''(h - p\,dT)h'^2 - w'(h - p\,dT)w''(h)h'^2w'(x)^{-1} = \frac{h'^2w'(h-p\,dT)}{(h-p\,dT)h}\left(\xi(h - p\,dT)h - \xi(h)(h - p\,dT)\right)$. For an isoelastic welfare function, the term in brackets reduces to $\xi p\,dT$, so that

[18] The inverse of Ψ is the (normalized) Feldstein's (1972) measure of the distributional characteristic of a tax.

$g'' < 0$ (because $\xi < 0$ when $w'' < 0$), and an increase of T is a concave transformation of the social welfare function. By property 1, Ψ therefore must increase, so $\frac{\partial \Psi}{\partial T} > 0$. □

Now, consider the term $\frac{MCT}{MCX}$, and recall that its arguments are X, T and γ. In order to calculate the partial derivative with respect to γ, we need to keep $X = t\gamma(Y - T)$ constant, so that t must vary inversely with γ; to be precise, $\frac{dt}{d\gamma} = -t/\gamma < 0$. Because the social welfare function is isoelastic, one can show that Ψ is not affected by changes in t or γ, and thus $\partial \frac{MCT}{MCX}/\partial \gamma = \Psi t/\gamma > 0$: a larger tax base reduces the relative social marginal cost of raising revenue via the marginal tax rate, as opposed to the lump-sum tax. This implies that $MCT_\gamma MCX - MCX_\gamma MCT > 0$. Given that, at the optimum, $MCX = MCT > 0$, we must have $MCX_\gamma < MCT_\gamma$. Now recall equations (6) and (7), in which the first term in both formulae is a substitution effect and the second one is a revenue effect. We have just signed the numerators of the substitution terms so, given that $SOC > 0$, the substitution response is to increase X and decrease T. From property 2, it immediately follows that $\partial \frac{MCT}{MCX}/\partial T > 0$: the relative marginal social cost of raising revenue via the lump-sum tax increases the more it is used. This implies, analogously as above, that $MCT_T - MCX_T > 0$. Thus, the revenue effect on X is also positive, concluding the proof: X increases when γ increases.

An increase in egalitarianism shifts the Mirrlees curve up. Following the line of the previous proof, consider $\frac{MCT}{MCX}$. For given X and T, the transformation of the social welfare function affects this ratio only through its effect on Ψ. Property 1 guarantees that Ψ increases when the social welfare function becomes more egalitarian, thus increasing the relative marginal cost of T versus X; as before, this implies a substitution response from T to X. It is the only effect, because the transformation of the social welfare function does not have a revenue aspect concluding the proof.

Near the intersection of the two curves corresponding to the global optimum, the Yitzhaki-Wilson curve has a larger slope. Recall that an increase in egalitarianism does not affect the Y-W curve, shifts the upward sloping Mirlees curve up and increases both X and γ. This is only possible if the Mirrlees curve is flatter than the Y-W curve.

Proof of Proposition 2. The previous optimum (γ^*, X^*) still lies on the orignal Mirrlees curve. As a result, there are only second-order changes in the solutions of the Mirrlees problem in a neighborhood of the optimum. The changing administrative cost does, though, affect the Y-W curve; it changes the optimal mix of t and γ to raise a given amount of distortionary tax. In fact, it must shift upward, so that for a given γ, the optimal X is higher. To see this, consider if it were otherwise, so that the optimal X declines which implies, via the budget constraint, that the corresponding value of T increases. Recall that $X = t\gamma(Y - T)$, so that, even though an increase in T affects X, t must fall to keep gross revenue $(T + X)$ constant. But the right-hand side of expression (20), which has to be satisfied on the Y-W curve, is

increasing in t, and so an increase in T and a decline in t implies that the right-hand side of this expression needs to fall, as well. Yet, by assumption, the left-hand side of (20) increases, leading to a contradiction. Intuitively, an increase in the marginal cost of expanding the tax base must, in equilibrium, be accompanied by an increase in its marginal social benefit, which occurs in a more progressive tax system. Thus, γ^* must now correspond to a lower T and a higher X; in other words, the Y-W curve shifts upwards. Because the Mirrlees curve is flatter than the Y-W curve, the global optimum shifts to the left along the Mirrlees curve, corresponding to lower progressivity and a smaller tax base.

References

Atkinson, A. B. & Stiglitz, J. E. (1980), *Lectures on Public Economics*, McGraw-Hill Book Co., London; New York.

Auten, G. & Carroll, R. (1999), 'The effect of income taxes on household behavior', *Review of Economics and Statistics* **81**(4), 681–693.

Carroll, R. (1998), Do taxpayers really respond to changes in tax rates? Evidence from the 1993 Act, Office of Tax Analysis Working Paper 78, U.S. Department of Treasury.

Cremer, H. & Gahvari, F. (1993), 'Tax evasion and optimal commodity taxation', *Journal of Public Economics* **50**(2), 261–275.

Cremer, H. & Gahvari, F. (1994), 'Tax evasion, concealment and the optimal linear income tax', *Scandinavian Journal of Economics* **96**(2), 219–39.

Cremer, H. & Gahvari, F. (1996), 'Tax evasion and the optimum general income tax', *Journal of Public Economics* **60**(2), 235–249.

Feldstein, M. S. (1972), 'Distributional equity and the optimal structure of public prices', *American Economic Review* **62**(1/2), 32–36.

Feldstein, M. S. (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, M. S. (1999), 'Tax avoidance and the deadweight loss of the income tax', *Review of Economics and Statistics*.

Kaplow, L. (1990), 'Optimal taxation with costly enforcement and evasion', *Journal of Public Economics* **43**(2), 221–236.

Lucas, R. E. (1976), 'Econometric policy evaluation: A critique', *Journal of Monetary Economics* **1**(2), 19–46.

Mayshar, J. (1991), 'Taxation with costly administration', *Scandinavian Journal of Economics* **93**(1), 75–88.

Moffitt, R. & Wilhelm, M. (1998), Taxation and the labor supply decisions of the affluent, Office of Tax Policy Research Working Paper 98-8, University of Michigan.

Okun, A. M. (1975), *Equality and Efficiency: The Big Tradeoff*, The Brookings Institution, Washington, D.C.

Slemrod, J. (1994), 'Fixing the leak in Okun's bucket. Optimal progressivity when avoidance can be controlled', *Journal of Public Economics* **55**(1), 41–51.

Slemrod, J. (1998), 'Methodological issues in measuring and interpreting taxable income elasticities', *National Tax Journal*.

Slemrod, J. (forthcoming), 'A general model of the behavioral response to taxation', *International Tax and Public Finance*.

Wilson, J. D. (1989), 'On the optimal tax base for commodity taxation', *American Economic Review* **79**(5), 1196–1206.

Yitzhaki, S. (1979), 'A note on optimal taxation and administrative costs', *American Economic Review* **69**(3), 475–480.