



ELSEVIER

Journal of Public Economics 55 (1994) 41–51

JOURNAL OF
PUBLIC
ECONOMICS

Fixing the leak in Okun's bucket Optimal tax progressivity when avoidance can be controlled

Joel Slemrod*

*School of Business Administration, University of Michigan, Ann Arbor, MI 48109-1234,
USA*

Received September 1992, revised version received June 1993

Abstract

This paper examines optimal income tax progressivity when avoidance responses to taxation are important, and can be controlled at some cost by the government. A simple example shows that ignoring the fact that avoidance can be controlled – that the leak in Okun's bucket can be fixed – can lead to misleading implications about the optimal degree of tax rate progressivity. The paper then generalizes the standard model of the optimal linear income tax to include taxpayer avoidance behavior and the ability of government to control the avoidance, but not the labor supply, response to higher marginal tax rates. The discussion suggests that if the tax burden is to be made more progressive, base broadening and increased enforcement ought to be considered as alternatives, or complements, to increasing tax rates for high-income individuals.

1. Introduction

In *Equality and Efficiency: The Big Tradeoff*, Okun (1975) proposed a hypothetical experiment for testing one's attitudes toward the tradeoff

* Prepared for the Conference on Tax Administration and Tax Policy, co-sponsored by the International Seminar in Public Economics and the Instituto de Estudios Fiscales, held in El Escorial, Spain on 11–12 June 1992. The participants at this conference, and in particular Maurice Marchand and Agnar Sandmo, provided helpful comments on an earlier draft. I am grateful to Jonathan Parker for help with the numerical optimization results.

between equality and efficiency. In the experiment a tax on the most affluent 5 percent of families finances a grant for families in the lowest quintile of the income distribution. Okun notes that the program has 'an unsolved technological problem: the money must be carried from the rich to the poor in a leaky bucket . . . , so the poor will not receive all the money that is taken from the rich' (p. 91). How much leakage – representing the inefficiency due to tax-and-transfer programs – one would tolerate, and still support the program, measures one's (marginal) terms of trading off efficiency and equity. To Okun the inefficiencies of redistribution included primarily the adverse effects on the economic incentives of the rich and the poor, and the administrative costs of tax collection and transfer programs.

The modern theory of optimal taxation has formalized the question of designing a tax system that minimizes the inefficiency cost, and of trading off the inefficiency and the social benefit of a more equal distribution of welfare. Until recently, the optimal taxation literature focused entirely on the incentive effects, to the exclusion of collection costs, although the latter has recently been introduced by Mayshar (1991) and others.

For the most part the incentive cost of taxation has been equated with the behavioral response of individuals to changes in relative prices, in particular the relative price of leisure and goods. It is the substitution away from goods toward leisure caused by taxing labor income that is the source of the inefficiency first studied by Mirrlees (1971), and in most subsequent treatments of optimal income tax progressivity.

The theoretical emphasis on the sensitivity to taxation of 'real' variables such as labor supply is, however, troubling in the light of much recent evidence [Slemrod (1992)] that an equally pervasive kind of response to taxation involves activities which reduce tax liability without requiring a substitution of one good for another, or one input for another. This includes a wide range of activities, such as receiving labor compensation as untaxed fringe benefits or preferentially taxed capital gains via stock options, expending revenues to legally reduce taxable income, or illegally underreporting taxable income. In this paper I will refer to these activities as tax avoidance. The operational distinction between what I call tax avoidance and the standard behavioral response such as reducing labor supply is that the amount of tax avoidance is endogenous to other policy instruments such as the extent of monitoring of taxpayer reports and the extent of 'loopholes' in the definition of taxable income.

How applicable are the standard optimal tax prescriptions to a world where avoidance responses are important? Do the same rules apply with a straightforward reinterpretation, or is a fundamental rethinking of the normative tool kit required? This paper begins a treatment of these issues by reexamining one of the fundamental normative questions of tax policy – optimal income tax progressivity – in a stylized economy characterized by

both kinds of behavioral response to taxation.¹ It begins in section 2 with a simple example of an economy where the optimal income tax schedule must be considered simultaneously with the optimal enforcement regime of that schedule. Section 3 generalizes the example by characterizing the optimal linear income tax cum enforcement system. Section 4 offers some concluding comments.

2. A simple numerical example

Consider an economy consisting of two persons whose abilities (assumed equal to wage rates) differ. The government is presumed to maximize a concave social welfare function, and can levy a linear income tax on the high-income individual (denoted H) and transfer the revenue collected from H, net of administrative costs, in a lump-sum fashion to the low-income individual (denoted L). There is no labor–leisure choice, so that before-tax incomes are fixed. The high-income individual can, at some cost, reduce her tax liability by an amount A , representing tax avoidance. The cost of tax avoidance depends on government expenditure on enforcement, denoted E .

Let the cost of avoidance be

$$C = \frac{1}{2} \left(\frac{EA^2}{a} \right), \tag{1}$$

so that the marginal cost of avoidance is increasing and is proportional to E . In this example the high-income individual faces the following problem:

$$\max_A Y_H = w_H - t(w_H - A) - \frac{1}{2} \left(\frac{EA^2}{a} \right). \tag{2}$$

The optimal level of A is then equal to at/E . Let $Y_H^*(E, t)$ be net income, given that an optimal amount of A is chosen. Then

$$Y_H^*(E, t) = (1 - t)w_H + \left(\frac{a}{2E} \right) t^2. \tag{3}$$

Revenue, net of enforcement costs, is equal to

$$R = t(w_H - A) - E = t \left(w_H - \frac{at}{E} \right) - E. \tag{4}$$

This is the amount that can be transferred to the low-income individual, whose income, Y_L , is simply $w_L + R$.

The government’s problem can now be stated explicitly as

¹ An early treatment of this issue is Sandmo (1981), who examines optimal progressivity and enforcement in an economy with tax evasion.

$$\max_{t,E} W = \frac{1}{\alpha} \left\{ \left[(1-t)w_H + \left(\frac{a}{2E} \right) t^2 \right]^\alpha + \left[w_L + t \left(w_H - \frac{at}{E} \right) - E \right]^\alpha \right\}, \quad (5)$$

where α is an (inverse) measure of the taste for egalitarianism embedded in the social welfare function.

The nature of the optimal progressivity problem of expression (5) is similar to the standard one where there is no avoidance behavior, except that there is no labor-leisure choice. Increasing t transfers welfare from H to L, up to a point,² but it also causes a deadweight loss to the economy. In this model the deadweight loss is $C + E$, or $(at^2/2E) + E$. If α were equal to one, so there is no preference for equality, social welfare would be maximized at $t = 0$, $E = 0$, and W would equal $w_H + w_L$.

Table 1 presents the solutions to this problem for the case where $w_H = 3$ and $w_L = 1$. I investigate values of α equal to -1 , -2 , or -3 , and values of a equal to 0.5 , 1.0 , or 1.5 . Note that the parameter a is a proportionality constant in the cost of avoidance function—the higher is a , the lower the cost of avoidance; T refers to the tax paid by individual H.

In this example, for a constant value of a , increasing egalitarianism (lower α) calls for both a higher tax rate and higher expenditure on tax

Table 1

Solutions to a numerical example of the optimal linear tax problem when avoidance can be controlled

Parameters		Solutions		Characteristics of solution					
α	a	t	E	A	C	Y_H	Y_L	R	T
-1	0.5	0.374	0.221	0.85	0.16	2.04	1.59	0.59	0.81
-1	1.0	0.358	0.319	1.12	0.20	2.13	1.35	0.35	0.67
-1	1.5	0.266	0.306	1.30	0.17	2.38	1.15	0.15	0.46
-2	0.5	0.411	0.243	0.85	0.17	1.94	1.64	0.64	0.88
-2	1.0	0.437	0.390	1.12	0.24	1.93	1.43	0.43	0.82
-2	1.5	0.424	0.489	1.30	0.28	2.00	1.23	0.23	0.72
-3	0.5	0.429	0.253	0.80	0.18	1.89	1.67	0.67	0.92
-3	1.0	0.475	0.424	1.12	0.27	1.84	1.33	0.47	0.89
-3	1.5	0.498	0.573	1.30	0.32	1.83	1.13	0.27	0.84
-2	1.1	0.438	0.415	1.16	0.25	1.94	1.39	0.39	0.82
-2	1.1	0.425	0.390 ^a	1.20	0.25	1.98	1.33	0.38	0.77
-3	1.5	0.247	0.253 ^a	1.47	0.18	2.44	1.13	0.13	0.38

^a For these cases the value of E was fixed at the value given and the objective function was maximized with respect to t only.

² The transfer-maximizing value of t is $(Ew_L/2a)$.

enforcement.³ In fact, for this example the ratio of t to E is invariant to α , so that t and E increase proportionally as α declines. The deadweight loss also increases with egalitarianism.

A change in the cost of avoidance, for a given value of α , does not have a consistent effect on optimal policy. For all values of α , facilitated avoidance (higher a) reduces both the optimal amount of tax and the optimal transfer, and therefore makes the low-income individual worse off. However, with relatively low egalitarianism ($\alpha = -1$), the optimal reaction to a higher value of a is to give in by reducing tax rates (although enforcement is increased), and allowing the high-income individual to reap some of the benefits of increased avoidance expenditure. With much greater egalitarianism ($\alpha = -3$), the optimal reaction to facilitated avoidance is to sharply increase enforcement and also increase tax rates, somewhat offsetting the otherwise sharp decline in welfare of the low-income individual, but at the cost of reducing the welfare of the high-income individual.

The bottom three rows of Table 1 offer further insight into this example. First focus on the third-to-last row, where $\alpha = -2$ and $a = 1.1$. Compared with the case of $\alpha = -2$, $a = 1$, this represents an exogenous decrease in the cost of avoidance. If both t and E are re-optimized, they both rise. But, as the second-to-last row shows, if E is held constant at 0.390, the optimal response is to reduce, not raise, the tax rate. The last row offers a similar exercise. Holding α at -3 when a goes from 0.5 to 1.5 the optimal response is to increase both t and E . However, if E is held constant at 0.253, the optimal response of t is a sharp decline.

These examples illustrate the critical difference between an optimal progressivity model with avoidance and one with only a labor–leisure tradeoff – that the former kind of leakage can be controlled by government policy. In the standard model an exogenous increase in the elasticity of substitution inevitably leads to a decline in optimal progressivity, as the social cost per unit of redistribution becomes higher. In these examples, an exogenous increase in leakiness (higher a) decreases t only if E is held constant. Note, though, that even though t increases in the new optimum, the transfer to the low-income individual is less than when a is lower, i.e. under a less leaky tax system.

3. The optimal linear income tax and optimal tax base enforcement

In this section I generalize the now-standard model of the optimal linear income tax, with notation adapted from Atkinson and Stiglitz (1980), to

³ It is also true, though not shown in Table 1, that the optimal value of E for given t is increasing in t ; also, the optimal value of t for given E is increasing in E .

include taxpayer avoidance behavior. In this expanded model the individual maximizes utility $U(Y, L)$ subject to

$$Y = wL - t(wL - A) + G - C(A, E, w), \quad (6)$$

where Y is after-tax income, w is the wage rate, L is labor supplied, t is the constant marginal tax rate, and G is the demogrant. Individuals can reduce their taxable income by an amount A , at a cost to them of C , which depends not only on the amount of avoidance achieved but also on government expenditures on enforcement, denoted E ,⁴ so that $C_A > 0$ and $C_E > 0$.

The first-order conditions for L and A , assuming an interior solution, are

$$L: (1 - t)wU_Y + U_L = 0 \quad (7)$$

and

$$A: t = C_A. \quad (8)$$

The government's revenue constraint, normalized by setting $\int_{w_L}^{w_H} dF = 1$, is

$$E + G + R = t \int_{w_L}^{w_H} (wL - A) dF, \quad (9)$$

where w_L and w_H are, respectively, the lower and upper bound on wage rates and R is the exogenous per capita revenue requirement.

The government is assumed to maximize the social welfare function

$$\int_{w_L}^{w_H} \psi(U) dF. \quad (10)$$

Forming the Lagrangian,

$$L = \int_{w_L}^{w_H} \{ \psi + \lambda [t(wL - A) - E - G - R] \} dF, \quad (11)$$

we can derive the first-order conditions as follows:

$$G: \int_{w_L}^{w_H} \left\{ \psi' \frac{\partial U}{\partial G} + \lambda \left[t \left(w \frac{\partial L}{\partial G} - \frac{\partial A}{\partial G} \right) - 1 \right] \right\} dF = 0, \quad (12)$$

⁴ Although I refer to E as enforcement, it can also be thought of as including such things as a costly broadening of the tax base.

$$t: \int_{w_L}^{w_H} \left\{ \psi' \frac{\partial U}{\partial t} + \lambda \left[(wL - A) + t \left(w \frac{\partial L}{\partial t} - \frac{\partial A}{\partial t} \right) \right] \right\} dF = 0, \quad (13)$$

$$E: \int_{w_L}^{w_H} \left\{ \psi' \frac{\partial U}{\partial E} + \lambda \left[t \left(w \frac{\partial L}{\partial E} - \frac{\partial A}{\partial E} \right) - 1 \right] \right\} dF = 0. \quad (14)$$

Using the avoidance-modified Slutsky relationships,

$$\frac{\partial L}{\partial t} = -wS_{LL} - (wL - A) \left(\frac{\partial L}{\partial M} \right), \quad (15a)$$

$$\frac{\partial A}{\partial t} = S_{At} - (wL - A) \left(\frac{\partial A}{\partial M} \right), \quad (15b)$$

where S_{LL} and S_{At} are the substitution terms, and the facts that $\partial U/\partial t = -\alpha(wL - A)$, $\partial U/\partial G = \alpha$ and $\partial U/\partial E = -\alpha C_E$, where α is the private marginal utility of income, allows us to rearrange (12), (13) and (14) as follows:

$$G: \int_{w_L}^{w_H} \left[\psi' \frac{\alpha}{\lambda} + t \left(w \frac{\partial L}{\partial M} - \frac{\partial A}{\partial M} \right) - 1 \right] dF = 0, \quad (16)$$

$$t: \int_{w_L}^{w_H} \left[\frac{\psi' \alpha}{\lambda} (wL - A) + wL \left(-1 + \frac{twS_{LL}}{L} \right) + tw(wL - A) \frac{\partial L}{\partial M} + A + tS_{At} - t(wL - A) \frac{\partial A}{\partial M} \right] dF = 0, \quad (17)$$

$$E: \int_{w_L}^{w_H} \left[\frac{-\psi' \alpha C_E}{\lambda} + t \left(w \frac{\partial L}{\partial E} - \frac{\partial A}{\partial E} \right) - 1 \right] dF = 0. \quad (18)$$

The first-order condition for E says that, at the margin, the avoidance cost saving from one dollar of increased enforcement, weighted by the marginal social valuation of the dollars saved, should equal to minus any extra revenue collected because the increased enforcement increases the tax base by causing either additional labor supply or reduced avoidance).

Following Atkinson and Stiglitz, we now define b to be the net social marginal valuation of income (modified to include the income effect of the revenue loss from avoidance):

$$b \equiv \frac{\psi' \alpha}{\lambda} + t \left(w \frac{\partial L}{\partial M} - \frac{\partial A}{\partial M} \right). \quad (19)$$

Using this definition, the conditions characterizing optimal tax policy become

$$G: \bar{b} = 1, \quad (20)$$

$$t: \frac{t}{1-t} = \frac{-\text{cov}[b, wL - A]}{\int wL \varepsilon_{LL} dF + \int A \varepsilon_{At} dF}, \quad (21)$$

where \bar{b} is the mean value of b , ε_{LL} is the compensated wage elasticity of labor ($=w(1-t)S_{LL}/L$), and ε_{At} is the compensated tax elasticity of avoidance⁵ expressed in wage-equivalent units ($=(1-t)S_{At}/A$).

Compare expression (16) for the optimal marginal tax rate with the one found in Atkinson and Stiglitz (pp. 405–408), for the case without avoidance:

$$\frac{t}{1-t} = \frac{-\text{cov}(b, wL)}{\sum wL \varepsilon_{LL} dF}. \quad (22)$$

Clearly expression (22) is a special case of expression (21) in which the cost of avoidance is infinite, at a zero or infinitesimal value of E , so that A and ε_{At} are both zero. In the numerator of expression (21) $-\text{cov}[b, wL - A]$ replaces $-\text{cov}[b, wL]$, because $wL - A$ and not wL is pre-tax income. This implies, for example, that if A is more concentrated among high-income taxpayers than wL , then $|\text{cov}[b, wL - A]| < |\text{cov}[b, wL]|$, so that the optimal tax rate is lower. This reflects the fact that t redistributes income through its effect on the tax base, and to the extent that the skewness of A erodes the effectiveness of the tax base as a measure of well-being, t is a less effective instrument for redistributing income.

The denominator is also altered. In the Atkinson–Stiglitz version, it is the compensated labor supply elasticity, weighted by labor income. In the expanded version, it is that plus the compensated avoidance elasticity, weighted by the amount of avoidance. Both kinds of leaks increase the costliness of redistributing income. Ceteris paribus, a higher avoidance elasticity lowers the optimal marginal tax rate because it increases the marginal resource cost per dollar of revenue raised for an increase in t .

Although ε_{LL} depends on the shape of individuals' preferences, the value of $A\varepsilon_{At}$ in general depends on government policy, represented by the parameter E . In the numerical example of section 2, it is inversely proportional to E , so that when E increases, $A\varepsilon_{At}$ declines, increasing the optimal value of t , ceteris paribus. This illustrates that the problem of setting

⁵ The simple form of the maximization problem of (1) dictates that $S_{At} = (1/C_{AA})$.

the optimal progressivity of the income tax cannot be separated from the problem of optimally enforcing whatever tax system is in place.

In the problems considered so far there has been no direct interaction between the terms of the labor–leisure tradeoff and the terms of the avoidance decision. Such a direct interaction could be introduced in a variety of ways. For example, in a graduated income tax system increased avoidance could reduce the marginal tax rate and thereby increase the relative price of leisure. Alternatively, private expenditure on avoidance could be modelled as reducing the *fraction* of income that is subject to tax, rather than as reducing the amount of taxable income. In this case the marginal tax rate on an incremental dollar of labor income is $t(1 - A')$, where A' is the fraction of taxable income that is avoided. In this sense avoidance directly affects the terms of the labor–leisure tradeoff, and could affect the deadweight loss due to an inefficiently low labor supply.⁶ The implications of this alternative formulation are worth pursuing in future research.

4. The appropriate policy reaction to evidence of increased behavioral responsiveness: A reinterpretation of the 1980s

One of the most stunning tax developments of the 1980s in the United States was the steep decline in the statutory progressivity of the income tax schedule; the U.S. income tax rate at the highest incomes fell from 70 percent in 1980 to 28 percent by 1988. During the decade many other countries also flattened their income tax rate structure, although most did so in a less extreme way. Although some of the decline in the graduation of rates was offset by base-broadening measures, most observers agree that during this period in the United States the burden of taxation shifted away from the highest income groups.

Two, not mutually exclusive, explanations have been offered for this phenomenon. One is that political power shifted toward the parties that represented the interests of the wealthy, and once in power these parties effected a transfer of resources (compared with earlier tax regimes) toward their constituency. An alternative explanation is that by 1980 economists had convinced policy-makers that the incentive costs of progressive income taxation were high enough that, at the margin, they outweighed whatever benefits of redistribution they provided – ‘supply-side economics’. In the extreme version of this view, associated with Arthur Laffer, the incentive effects at the margin were so large that decreases in tax rates would actually

⁶ Alternatively, the avoidance cost function could be written as $C(A/wL, E)$

bring in greater revenues, and would therefore potentially benefit all taxpayers.

The argument presented in this paper suggests a competing explanation for the 1980s, one that is in the same economics tradition as the supply-side story, but which has radically different policy implications. To see the change in perspective, consider the following thought experiment. Suppose that definitive econometric evidence arrives to show that the behavioral response to taxation is greater than had been previously thought. What is the appropriate policy reaction, assuming that before the arrival of this information the income tax was optimally progressive? The answer in the Atkinson–Stiglitz framework is straightforward. According to expression (17), if ε_{LL} increases, then the optimal value of t falls, necessitating a fall in G to recover budget balance.⁷ The tax system becomes less progressive.

But what is the appropriate policy response to new evidence about an increased *avoidance* response? At first glance at expression (16), a higher value of avoidance response reduces optimal t in the same way that an increased value of ε_{LL} would. It increases the marginal excess-burden-per-dollar-raised, and through this channel reduces the optimum value of t .

The message of this paper is that this first-glance answer may be incorrect. First of all, an increased avoidance response will be accompanied by higher values of A , for any given tax system. Because the change in A will likely have a different distribution than $wL - A$, this will affect the value of $\text{cov}[b, wL - A]$, as will the induced changes in b . Second, a change in avoidance responsiveness in general will change the optimal value of E , which will affect the avoidance responsiveness. Thus, the apparent decline in the optimal value of t when avoidance response increases, *holding E constant*, may be a misleading indicator of how optimal t changes when all the other instruments at the government's disposal can change as well. How important a factor this is depends on the tax avoidance technology, i.e. the $C(A, E, w)$ function. There is no doubt, though, that increasing the avoidance responsiveness increases the social cost of redistribution, and therefore reduces the optimal amount of redistribution, for any given social preferences.

5. Conclusion: Income tax progressivity in the 1990s

Did the flattening of the income tax rate structure in the 1980s go too far? This is likely to be one of the major fiscal policy questions of the 1990s, and debate on it has already begun. In the United States, the Clinton Adminis-

⁷ A changed labor supply elasticity will also change the pattern of labor supply, potentially changing $\text{cov}[b, wL]$.

tration tax proposals feature an increase in the top statutory income tax rate from 31 to 39.6 percent. In the United Kingdom the Labour Party platform featured an even larger increase in the top rate.

Traditionally, the principal contribution that economists could offer to this debate is measurement of how leaky is the bucket that would carry water (income) from high-income taxpayers to everyone else, or to the public sector. To be sure, there is compelling, though controversial, evidence that high-income people in the United States responded to the tax cuts in 1981 [Lindsey (1991)]. But an unresolved question is how much of the response can be categorized as avoidance and, to continue the analogy, at what cost the leak can be fixed. This discussion suggests that if the tax burden is to be made more progressive, base broadening and increased enforcement ought to be considered as alternatives, or at least complements, to increasing tax rates for high-income individuals.

References

- Atkinson, A. and J. Stiglitz, 1980, *Lectures on public economics* (McGraw-Hill, New York and Maidenhead).
- Lindsey, L., 1991, *The growth experiment* (Basic Books, New York).
- Mayshar, J., 1991, Optimal taxation with costly administration, *Scandinavian Journal of Economics* 73, 75–88.
- Mirrlees, J., 1971, An exploration in the theory of optimum income taxation, *Review of Economic Studies* 38, 175–208.
- Okun, A.M., 1975, *Equality and efficiency: The big tradeoff* (The Brookings Institution, Washington, DC).
- Sandmo, A., 1981, Income tax evasion, labor supply and the equity–efficiency trade-off, *Journal of Public Economics* 16, 265–288.
- Slemrod, J., 1992, Do taxes matter? Lessons from the 1980's, *American Economic Review* 82, 250–256.