

## RANDOMIZATION OF COMMODITY TAXES

### An Expenditure Minimization Approach

F.R. CHANG and D.E. WILDASIN\*

*Department of Economics, Indiana University, Bloomington, IN 47405, USA.*

Received March 1985, revised version received July 1986

Using an expenditure minimization approach, necessary and sufficient conditions for local random taxation are obtained in terms of the curvature of the compensated demand function, so that intuition from excess burden analysis can be applied. Major findings include: (1) random taxation is locally optimal if the compensated demand function is sufficiently convex; (2) horizontally equitable taxation is locally optimal if the compensated demand function is concave, and (3) local randomization is not optimal if the tax revenue requirement is sufficiently close to zero or to any local maximum. We also derive an inverse elasticity characterization of the optimal random tax structure.

### 1. Introduction

In a recent paper, Stiglitz (1982a) presents an interesting challenge to traditional ideas in public finance by showing that horizontally equitable taxation of identical individuals may conflict with utilitarianism or ex ante utility maximization: The sum or expected value of utilities can sometimes increase by allowing the tax structure to differentiate among identical households. This is an important contribution because it poses a dilemma as to which of these criteria should be regarded as more fundamental when it becomes necessary to choose between them.<sup>1</sup>

Such an intriguing conclusion warrants careful examination. In this paper, our goal is to show the exact conditions under which randomization of the structure is optimal, and to help clarify the intuition underlying this result. More precisely, we present and interpret necessary and sufficient conditions

\*We are grateful to L. Chenault, J. Wilson, and two referees for helpful comments on an earlier version of this paper, but retain responsibility for errors. The first author was partially supported by an Indiana University's Outstanding Young Faculty Award.

<sup>1</sup>See also Weiss (1976), Stiglitz (1982b) and Balcer and Sadka (1982) on random taxation. Those papers focus primarily on randomization of income taxes, whereas we are concerned, as is Stiglitz (1982a), with random commodity taxation. They also deal with informational asymmetries and screening problems which are absent from our analysis. All subsequent citations of Stiglitz's work refer to Stiglitz (1982a).

for a small move toward a random tax structure, i.e. 'local randomization', to be desirable. Our general method is similar to that of Stiglitz, who derives sufficient conditions for local randomization and then argues that these conditions are met if various behavioral parameters take on certain values. We also provide an inverse elasticity characterization of the optimal tax structure, whether it involves 'small' or 'large' departures from uniform taxation.

Unlike Stiglitz, we develop most of our analysis from an expenditure-minimizing (dual) rather than utility-maximizing (primal) perspective. We present this alternative approach for three reasons. First, it is analytically more convenient to deal with than the equivalent social welfare maximization problem. Second, it provides a rigorous and straightforward interpretation of random taxation in terms of consumer's surplus or excess burden. This is most helpful in developing valid intuition for the problem. Third, as explained further below, it allows us to focus on efficiency aspects of the case for randomization independently of ethical judgments about the distribution of welfare, or, equivalently, independently of attitudes toward risk embodied in the cardinal properties of utility functions. In fact, all of our local randomization results are stated, directly or indirectly, in terms of properties of the compensated demand function evaluated at a *given* level of utility.

To summarize the structure and results of the paper, section 2 states the expenditure minimization problem on which the subsequent analysis is based, and presents some preliminary results on the properties of its objective and constraint functions. Section 3 establishes a series of propositions about necessary and sufficient conditions for the optimality of small deviations from a uniform or non-random tax structure. The curvature properties of the compensated revenue function (i.e. the relationship showing the revenue obtained from a household as a function of the tax rate it faces) and the compensated demand curve for the taxed commodity (both evaluated at the level of utility obtained at the uniform tax equilibrium) are of critical importance. If the revenue function is convex, local randomization is desired. Moreover, if the demand curve is sufficiently convex, this condition will be met. As a partial converse, local randomization is not optimal if the compensated demand curve is concave. In the intermediate case, where the revenue function is concave and the demand function is convex, randomization may or may not be desirable. We show that there are two cases, however, where the revenue function is concave and local randomization is definitely not optimal: when the government's revenue requirement is zero or sufficiently close to zero, so that the tax rate under uniform taxation is very small, and when the revenue requirement is sufficiently large, so that uniform taxation would require a tax rate close to one that (globally or locally) maximizes the revenue function.

Since the results of section 3 are derived from an expenditure minimization

problem, they have a direct interpretation in terms of excess burden. In particular, these results rigorously verify an informal diagrammatic analysis of random taxation initially presented by Stiglitz. It is not surprising, of course, that our formal verification of that analysis runs in terms of properties of a compensated curve. Since utility is held constant along such a curve, it is also not surprising that cardinal properties of the utility function play no role in the excess burden argument. Note, however, that this implies in particular that the conditions for randomization derived from such an argument do not, and cannot, depend on the degree of risk aversion that consumers might have.

In section 4, instead of evaluating small or local departures from uniform taxation, we use the expenditure minimization approach to characterize an optimal tax structure – which may involve uniform taxation, or small or large horizontal inequities. Notably, an inverse elasticity formula obtains at an optimum. Also notably, this inverse elasticity formula always holds at a uniform tax structure, even when uniform taxation is not actually optimal. This provides a specific example of a case where necessary conditions for optimal taxation can be satisfied by a non-optimal tax structure, and naturally leads to consideration of second-order conditions for optimal taxation. We show that satisfaction of the second-order conditions depends on curvature properties of the revenue function, and we relate this analysis to that of section 3.

Section 5 compares and contrasts our analysis with that of Stiglitz. While the difference in approach (primal vs. dual) naturally results in some differences in results, we show that some of the differences are only apparent. Stiglitz argues that randomization of taxes may be optimal when the government's revenue requirement is sufficiently high and the demand curve is sufficiently convex. This claim appears to conflict with our conclusion that randomization is non-optimal near maximum revenue. We show, however, that Stiglitz's requirement that the demand curve be sufficiently convex in the neighborhood of a maximum revenue tax rate is self-contradictory, and that Stiglitz's sufficient conditions for randomization therefore cannot be satisfied at a local (regular) maximum of tax revenue.<sup>2</sup> Indeed, Stiglitz's formal analysis leads to conditions that imply that revenues are at a local (regular) *minimum*. Thus, the revenue function is convex, not concave, just as our excess burden analysis would indicate.

Section 6 concludes with some further comparison of primal vs. dual approaches to the random taxation problem, and some suggestions for further research.

<sup>2</sup>We say that a function  $f(x)$  achieves a *regular* maximum at  $x=x^0$  if  $f'(x^0)=0$  and  $f''(x^0)<0$ . Similarly for a regular minimum.

## 2. The model

For the purposes of our analysis, it is sufficient to examine a very simple two-household economy. Let there be a single consumption good and let labor be the sole factor of production. Taking labor as numeraire, let  $p^i$ ,  $i=A, B$ , be the consumer price faced by household  $i$  for the consumption good, the producer price of which is fixed at unity. Both consumers have identical preferences, as given by the common expenditure function  $e(p, u)$ , where  $u$  is utility. Let  $C(p, u)$  denote the compensated demand function for the consumption good. Define the *revenue function*  $R(p, u) \triangleq (p-1)C(p, u)$ , showing the amount of revenue collected from a household with utility  $u$ , facing the tax rate  $(p-1)$ . The government is assumed to have a fixed revenue requirement of  $R^0$ .

Equal treatment of equals, at least *ex post*, means  $p^A = p^B$  and  $u^A = u^B = \bar{u}$  say. We therefore ask whether  $p^A = p^B$  solves<sup>3</sup>

$$(D) \quad \min_{\langle p^A, p^B \rangle} L(p^A, p^B, \bar{u}) \triangleq e(p^A, \bar{u}) + e(p^B, \bar{u})$$

s.t.

$$R(p^A, \bar{u}) + R(p^B, \bar{u}) \geq R^0. \quad (1)$$

If  $p^A \neq p^B$  at a solution to (D), we say that the tax structure is horizontally inequitable. The basic objective of the analysis in this and the next section is to find conditions under which the tax structure is horizontally inequitable in this sense.

The economic meaning of the problem (D) can be interpreted as follows. Let  $\bar{u}$  be the maximum (common) utility level obtainable under undifferentiated horizontally equitable taxation, subject to the revenue constraint. It follows that a solution  $(\bar{p}^A, \bar{p}^B)$  to (D) has  $\bar{p}^A \neq \bar{p}^B$  if and only if, starting from an initial equal utility, equal price situation ( $p^A = p^B, u^A = u^B = \bar{u}$ ), it is possible to differentiate the tax structure, and to devise appropriate side payments or lump-sum transfers, such that either or both households can be made better off, with neither being made worse off [Harris and Wildasin (1985)]. That is, a solution to (D) with  $\bar{p}^A \neq \bar{p}^B$  implies that there exists a Pareto improvement over an equal utility, equal tax rate situation. This means of course, that horizontally equitable or non-random taxation cannot be utilitarian optimal, since a Pareto improvement must increase the sum of utilities. A much stronger statement can be made, however: if  $\bar{p}^A \neq \bar{p}^B$  at a solution to (D), horizontal equity is undesirable from the viewpoint of any individualistic

<sup>3</sup>Diamond and McFadden (1974) develop the basic expenditure-minimization approach to optimal tax problems. Tresch (1981) provides an extensive treatment from this perspective.

social welfare function, no matter how inequality-averse it may be – including the Rawlsian maximin social welfare function. Furthermore, the tax structure is horizontally inequitable – i.e.  $\bar{p}^A \neq \bar{p}^B$  – if and only if  $p^A = p^B$  does not maximize aggregate consumers' surplus, or minimize excess burden, defined with reference to the (identical) compensated demand curves for each household evaluated at a common level of utility  $u^A = u^B = \bar{u}$  [again, see Harris and Wildasin (1985)]. Thus, analysis of the problem of random taxation using an aggregate expenditure minimization perspective, as in (D), helps to focus attention on the efficiency implications of departures from uniform taxation.<sup>4</sup>

We now develop conditions under which we can say whether or not horizontal equity is optimal from the expenditure-minimizing viewpoint of problem (D). Let  $\bar{L}$  in fig. 1 denote a level contour of  $L(p^A, p^B, \bar{u})$ , while  $R$  is the iso-revenue contour satisfying (1). Note the preference direction in price space indicated by the arrow:  $\bar{L}$  decreases toward the origin. The slope of  $\bar{L}$  is simply

$$\left. \frac{dp^A}{dp^B} \right|_{\bar{L}} = \frac{-C^B(p^B, \bar{u})}{C^A(p^A, \bar{u})}, \tag{2}$$

while the slope of  $\bar{R}$  is

$$\left. \frac{dp^A}{dp^B} \right|_{\bar{R}} = - \frac{C^B(p^B, \bar{u}) + (p^B - 1)C_p^B(p^B, \bar{u})}{C^A(p^A, \bar{u}) + (p^A - 1)C_p^A(p^A, \bar{u})} = \frac{-R_p(p^B, \bar{u})}{R_p(p^A, \bar{u})}, \tag{3}$$

where  $C_p^i = \partial C^i / \partial p^i < 0$  is the own-price derivative of the compensated demand function, and where  $R_p(p, \bar{u}) = \partial R / \partial p$ . At  $p^A = p^B$ , the basic symmetry of our problem implies that both curves have slopes of  $-1$ . Hence, for  $p^A = p^B$  to be globally optimal, it is necessary that  $\bar{R}$  be more convex than  $\bar{L}$  at  $p^A = p^B$ , as shown in fig. 1. This condition is also sufficient for  $p^A = p^B$  to be a local optimum. On the other hand,  $\bar{L}$  more convex than  $\bar{R}$  at  $p^A = p^B$  is sufficient for  $p^A = p^B$  not to be optimal (even locally).

We must therefore determine the curvature of these contours. By (2), we have (suppressing arguments for notational ease)

$$\left. \frac{d^2 p^A}{d(p^B)^2} \right|_{\bar{L}} = \frac{-C_p^B}{C^A} + \frac{C^B}{(C^A)^2} C_p^A \left. \frac{dp^A}{dp^B} \right|_{\bar{L}} = - \frac{C_p^B}{C^A} - \frac{C_p^A}{C^A} \left( \frac{C^B}{C^A} \right)^2.$$

<sup>4</sup>Single-consumer models have often been used in optimal tax analysis to isolate 'efficiency' from 'equity' problems. As a referee observes, one justification for considering such models is to assume that optimal lump-sum interpersonal redistribution is going on behind the scenes. [This is stated explicitly in Diamond and McFadden (1974, n. 6).] Many other studies have explicitly incorporated lump-sum interpersonal transfers in second-best models with heterogeneous consumers. Among them are Marchand (1968), Mohring (1970), Green (1975), Wildasin (1984) and, indeed, Boiteux (1971).

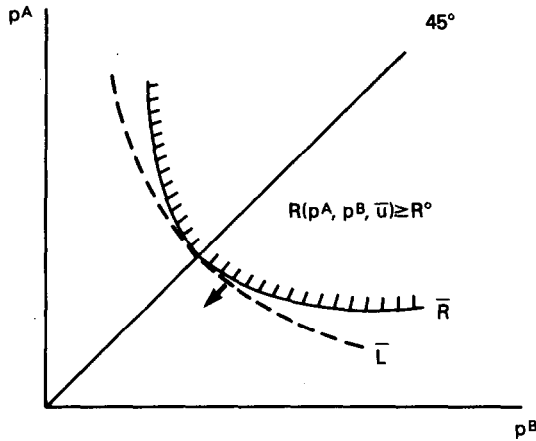


Fig. 1

Thus,

$$\left. \frac{d^2 p^A}{d(p^B)^2} \right|_{\bar{L}} = -2 \frac{C_p}{C} = -\frac{2\varepsilon}{p} > 0, \quad \text{at } p^A = p^B, \tag{4}$$

where  $\varepsilon = (p/C)C_p$  is the compensated price elasticity of demand. Thus,  $\bar{L}$  is necessarily convex around the 45° line.

By (3), letting  $R_{pp} = \partial^2 R / \partial p^2$ ,

$$\left. \frac{d^2 p^A}{d(p^B)^2} \right|_{\bar{R}} = \frac{-R_{pp}(p^B, \bar{u})}{R_p(p^A, \bar{u})} + \frac{R_p(p^B, \bar{u})}{R_p(p^A, \bar{u})} R_{pp}(p^A, \bar{u}) \left. \frac{dp^A}{dp^B} \right|_{\bar{R}}.$$

Thus,

$$\left. \frac{d^2 p^A}{d(p^B)^2} \right|_{\bar{R}} = -2 \frac{R_{pp}(\cdot)}{R_p(\cdot)} = -\frac{2\varepsilon(2 + \tau v)}{p(1 + \tau\varepsilon)}, \quad \text{at } p^A = p^B, \tag{5}$$

where  $\tau = (p-1)/p$  is the proportional tax rate on consumption and where  $v = pC_{pp}/C_p$  is the curvature of the compensated demand curve. Since the sign of  $v$  is unrestricted a priori,  $\bar{R}$  could be either convex or concave around the 45° line. Note for future reference that  $v > 0$  iff  $C_{pp} < 0$  iff the demand curve is concave to the origin. Hence, since  $R_{pp} = C_p(2 + \tau v)$ ,  $R_{pp} > 0$  implies  $v < 0$ , while  $v > 0$  implies  $R_{pp} < 0$ . That is, a convex revenue function implies a convex demand curve, while a concave demand curve implies a concave revenue function. Note particularly that this does not rule out a convex demand curve and a concave revenue function.

### 3. Curvature conditions for random taxation

In this section we characterize the desirability of randomization in terms of the curvature of the compensated demand curve,  $v$ , and relate it to the shape of the revenue function,  $R$ . We restrict attention to 'efficient' levels of taxation, i.e. we assume that  $R_p = C(1 + \tau\varepsilon) > 0$ .

By (4) and (5), a sufficient condition for randomization is

$$-\frac{2\varepsilon}{p} = \frac{d^2p^A}{d(p^B)^2} \Big|_L > \frac{d^2p^A}{d(p^B)^2} \Big|_{\bar{R}} = -\frac{2R_{pp}}{R_p} = -\frac{2\varepsilon(2 + \tau v)}{p(1 + \tau\varepsilon)}, \text{ at } p^A = p^B. \quad (6)$$

Eq. (6) can be simplified to

$$v < \varepsilon - 1/\tau. \quad (6')$$

Except for the borderline case that  $v = \varepsilon - 1/\tau$ , we may take (6') as a *necessary and sufficient* condition for local randomization. Expressed in terms of the compensated demand function, we have

$$\frac{C_{pp}}{C_p} < \frac{C_p}{C} - \frac{1}{p-1}. \quad (7)$$

Since we restrict taxation to 'efficient' levels, i.e.  $R_p = C + (p-1)C_p > 0$ , the above equation implies

$$\frac{C_{pp}}{C_p} < \frac{2C_p}{C}, \text{ or } v < 2\varepsilon. \quad (8)$$

Notice, however, *no* compensated demand with constant elasticity, i.e.  $C = kp^{-\alpha}$ ,  $k, \alpha > 0$ , satisfies (7).

It is immediate from (6) that if  $R_{pp} > 0$ , i.e. the revenue function is convex, then randomization is welfare improving. Notice that  $R_{pp} > 0$  if, and only if,  $v < -2/\tau$ , which implies a convex compensated demand function,  $C_{pp} > 0$ . In fact, if the demand curve is kinked and convex, we have  $v = -\infty$  and  $R_{pp} = \infty$ . This special case can easily be illustrated with an excess burden diagram, as in fig. 2(a). Suppose the required revenue can be obtained by taxing both households at a uniform rate  $\hat{\tau}$ , implying a consumer price of  $\hat{p}$  and a total excess burden of  $2 \cdot abc$ . If the demand curve is kinked at  $\hat{p}$ , then the taxes on A and B can be randomized in such a way as to hold revenue constant and reduce total excess burden. For example, in fig. 2(a) we show an increase in  $p^A$ , and a decrease of equal magnitude in  $p^B$ , that reduces aggregate excess burden (because  $b'a'ab < baa''b''$ ) or increases aggregate consumers' surplus (because  $\hat{p}p^A a'a < \hat{p}aa''p^B$ ) and actually increases tax revenues. Note, as illustrated in fig. 2(b), that the kink in the demand curve also implies a convex kink in the revenue function  $R$ , as shown at  $\tau = \hat{\tau}$ . This suggests the desirability of randomization around a tax rate at which revenue is definitely not maximized, a point that will be further elaborated below.

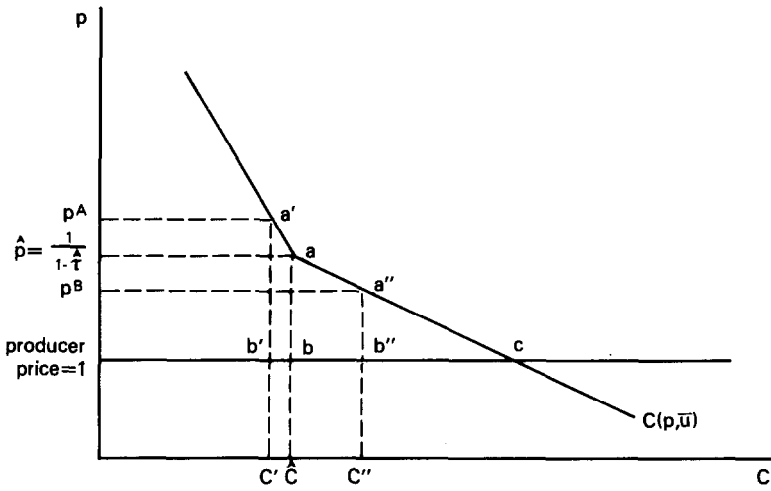


Fig. 2(a)

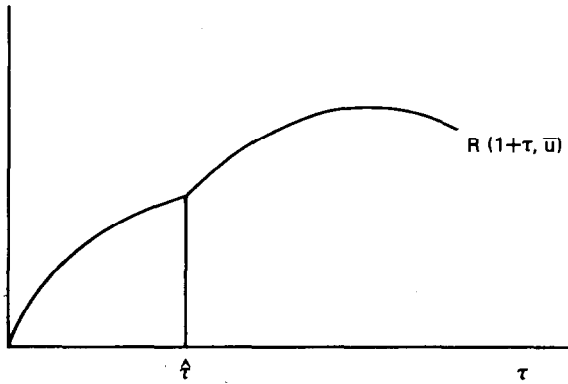


Fig. 2(b)

To summarize,

*Proposition 1.* Suppose that the revenue function  $R$  is convex at all points of efficient uniform taxation, i.e.  $R_{pp} > 0$  and  $R_p > 0$ . Then the optimal tax structure is random, i.e.  $p^A \neq p^B$ .

Next, we observe from (6') that randomization is not optimal, at least locally, if

$$1 + \tau(v - \epsilon) > 0. \tag{9}$$



Clearly, (9) is valid for all  $v \geq 0$ , i.e. it is not desirable to randomize locally for any concave (including linear) compensated demand function ( $C_{pp} \leq 0$ ). Notice  $R_{pp} < 0$  in this case. This result also reaffirms the intuition from the excess burden argument.

*Proposition 2. If the compensated demand function is concave at a point of efficient uniform taxation, i.e.  $C_{pp} \leq 0$  and  $R_p > 0$ , then local randomization of the tax structure is undesirable.<sup>5</sup>*

Propositions 1 and 2 identify two cases where one can make clear-cut statements about the optimality of random taxation, based on convexity of  $R$  in one case and concavity of  $C$  in the other case. This leaves the interesting case in which we have a convex compensated demand function together with a concave revenue function, i.e.

$$-2/\tau \leq v < 0. \tag{10}$$

Combining with (6'), local randomization with a concave revenue function is optimal iff

$$-2/\tau \leq v < \varepsilon - 1/\tau. \tag{11}$$

Hence we have:

*Proposition 3. Even if the revenue function is concave, the compensated demand curve may be sufficiently convex for randomization to be optimal.*

The question of randomization around a tax rate where  $R$  is concave is of particular interest because we can identify two situations where  $R$  definitely will be concave. The first occurs when  $\tau$  is 'sufficiently large', or, more precisely, when  $R$  is at or near a local or global maximum (if one exists). Here,  $R_p \simeq 0$  and  $R_{pp} < 0$ , the right-hand side of (6) is large and positive, and hence local randomization cannot be optimal.<sup>6</sup> The second case occurs when

<sup>5</sup>It is worthwhile noting that (weak) concavity of demand functions is an 'unusual' property in the sense that not all ordinary demand functions can be (simultaneously) concave at any price vector. See Harris (1975) for the proof. The extension to the compensated case is immediate: let  $x^i = x(q^i, \bar{u})$ , for  $i=1,2$ , be the compensated demand vector for consumer prices  $q^i$ ,  $x^1 = x(\lambda q^1 + (1-\lambda)q^2, \bar{u})$ . Concavity of  $x$  (i.e. concavity of each component of  $x$  in  $q$ ) means  $x^1 \geq \lambda x^1 + (1-\lambda)x^2$ . If preferences are strictly quasi-concave,  $u(x^1) = u(x^2) = \bar{u}$  implies  $u(x^1) > \bar{u}$ , a contradiction.

<sup>6</sup>As discussed in detail in Harris and Wildasin (1985), solutions to (D) correspond to solutions of social welfare maximization problems under certain conditions. However, this correspondence can break down in some cases. In particular, a maximum of the compensated revenue function need not, and generally will not, occur at the same tax rate that maximizes an uncompensated revenue function. If the income elasticity of  $C$ ,  $\eta$ , is positive, then  $R_p > 0$  at a tax rate where the

$\tau$  is 'sufficiently small', i.e. zero or close to zero. If  $\tau=0$ ,  $R_{pp}=(C/p)\varepsilon(2+\tau v)=(C/p)2\varepsilon<0$ , so  $R$  is definitely concave. Moreover, (6) requires  $-2\varepsilon> -4\varepsilon$  when  $\tau=0$ , which is obviously impossible. Hence, randomization is undesirable in this case as well.

To see these results in a different way, note that since  $R$  is concave near  $\tau=0$  and at maximum revenue tax rates, we can apply (11) to check the optimality of randomization. When the tax is sufficiently high in the sense that it is near the maximal feasible tax revenue, i.e.  $\tau=-1/\varepsilon$  approximately, the right-hand side of (11) becomes  $\varepsilon-1/\tau=-2/\tau$ . Hence, the set of values for  $v$  allowing randomization is empty. Moreover, rewriting (11) as

$$0 < 1/(\varepsilon - v) < \tau \leq 2/v, \quad (11')$$

we see that the range of tax rates at which randomization is desirable is bounded away from zero, which implies that one should not randomize near a zero revenue point. Thus, we have shown

*Proposition 4.* *The revenue function is concave at  $\tau=0$  and around (global or local) maximum revenue tax rates. In neither case is local randomization of the tax structure desirable.*

Proposition 4 suggests that randomization is 'most likely' to be optimal for 'intermediate' values of the revenue requirement,  $R^0$ . To avoid misinterpretation, however, note that for any  $R^0 > 0$ , no matter how small, there exists a demand curve that is sufficiently convex that randomization would be optimal. To see this, simply suppose that the tax rate  $\tau$  pictured in fig. 2(a) is 'small'. No matter how small  $\tau$  is, it would be possible for a demand curve to have a convex kink, or for  $v$  to be very large (negative), so that (6') would be satisfied. In other words, although  $R$  is concave at  $R^0=0$ , this is strictly a local property, and cannot be guaranteed over any finite interval. Similarly, although  $R$  is concave around any revenue maximum, this too is a strictly local property. It would be possible for randomization to be optimal near a revenue maximum, so long as  $R^0$  is not 'too' near such a maximum. Thus, although randomization can be shown to be non-optimal for sufficiently small and sufficiently large revenue requirements, these are only local

---

uncompensated revenue function is maximized. (The converse holds when  $C$  is inferior.) Thus, a maximum of  $R$  might correspond to an inefficient uniform tax from the primal perspective, and the duality of the two approaches would break down. Readers may therefore prefer to interpret the results here on maximum tax revenue as pertaining to the special case where the income elasticity of  $C$  is zero: in this case, compensated and uncompensated functions coincide, and none of the above difficulties arises. As we show in section 5, the basic conclusion that randomization is undesirable near maximum feasible revenue can be directly verified in the primal case without imposing the zero income elasticity restriction. Thus, for the purposes at hand, qualifications about income effects are not really critical.

statements, and little more can be established, with generality, about revenue requirements and randomization.

In summary, our analysis shows that convexity of the revenue function,  $R$ , which implies convexity of the demand curve,  $C$ , is sufficient but not necessary for randomization to be optimal. Concavity of  $C$ , which implies concavity of  $R$ , is sufficient but not necessary to rule out the desirability of local randomization. If  $R$  is concave, but  $C$  is convex, local randomization could be desirable. Under no circumstances, however, could local randomization be desirable when the tax rate is near zero or when the tax revenue requirement is sufficiently close to a local maximum.

**4. An inverse elasticity formula and second-order conditions for optimal tax structure<sup>7</sup>**

So far we have described conditions under which horizontally equitable taxation is locally non-optimal. What can we say about the tax structure that is optimal?<sup>8</sup>

From Harris and Wildasin (1985), we know that if  $(\bar{p}^A, \bar{p}^B)$  describes a tax policy that maximizes some social welfare function  $W(u^A, u^B)$ , given optimal lump-sum interpersonal redistribution, and if  $(\bar{u}^A, \bar{u}^B)$  are utility levels obtained by each household at this optimum, then  $(\bar{p}^A, \bar{p}^B)$  solve

$$\begin{aligned} & \min_{\langle p^A, p^B \rangle} e(p^A, \bar{u}^A) + e(p^B, \bar{u}^B) \\ & \text{s.t.} \\ & R(p^A, \bar{u}^A) + R(p^B, \bar{u}^B) \geq R^0. \end{aligned} \tag{12}$$

<sup>7</sup>This section was prompted by the insightful remarks of a referee, to whom we are greatly indebted.

<sup>8</sup>Our *local* analysis in section 3 can be extended to a large economy with an arbitrary number of individuals,  $n$ , divided into two groups of arbitrary size, say group A with  $(n - m)$  individuals and group B with  $m$  individuals. In that case, the loss function is  $(n - m)e^A + me^B$ , while the revenue constraint is  $(n - m)R^A + mR^B \geq R^0$ . The only effect of this on our formal analysis is to change the curvature of both the iso-expenditure and iso-revenue curves by the same proportion, as is easily verified. This of course leaves the local analysis unaffected. However, a referee has constructed an argument that suggests that the analysis of *globally* optimal taxation can depend in an important way on the size of the economy. In particular, with a large number of households, it may be possible to devise optimal policies which involve very heavy taxation of a very small number (e.g. one) of consumers and minimal or no taxation of all of the others. This policy could be accompanied by very large lump sum transfers from the low-tax group to the high tax individual which could maintain (or actually increase) the welfare level of that individual. (In fact, this individual is able indirectly to impose the tax burden on other households and actually ends up better off than everyone else. One might think of this person as a tax farmer.) As noted above, this argument in no way affects our local analysis. Moreover, the two-person global analysis in this section applies directly to a large economy if we impose the restriction that the population is divided into two equal-size groups. The referee's analysis indicates, however, that new results can be obtained in the large economy case with unequal group sizes, a possibility that deserves attention in future research.

This problem is identical to (D) except that we have substituted the optimal and possibly unequal values of  $u^A$  and  $u^B$  for the common utility level  $\bar{u}$ .

Forming a Lagrangian with multiplier  $\lambda > 0$  for (12), the following necessary conditions characterizing  $(\bar{p}^A, \bar{p}^B)$  are easily derived:

$$\bar{\tau}^i = \frac{1-\lambda}{\lambda} \frac{1}{\bar{\epsilon}^i}, \quad i = A, B, \quad (13)$$

where

$$\bar{\epsilon}^i \triangleq \frac{\partial \log C^i(\bar{p}^i, \bar{u}^i)}{\partial \log p_i}$$

$$\bar{\tau}^i \triangleq \frac{\bar{p}^i - 1}{\bar{p}^i}.$$

Hence,

$$\frac{\bar{\tau}^A}{\bar{\tau}^B} = \frac{\bar{\epsilon}^B}{\bar{\epsilon}^A}. \quad (14)$$

In other words,

*Proposition 5.* *At an optimal tax structure, the tax rate faced by each household  $i$  is inversely proportional to the elasticity of the household's compensated demand curve for the taxed commodity, evaluated at the optimal consumer price and utility level  $(\bar{p}^i, \bar{u}^i)$ .*

As is well known, first-order necessary conditions do not necessarily uniquely determine an optimal tax structure.<sup>9</sup> For example, the usual Ramsey formula for optimal taxation may also be satisfied by other, non-optimal tax structures. While this insufficiency of first-order conditions is widely noted in the literature, there is a dearth of examples of non-optimal tax structures satisfying the first-order conditions.

Such an example is readily available in the present context, however. In particular, it is obvious (from the symmetry of our two consumers) that a horizontally equitable tax policy, with  $\bar{p}^A = \bar{p}^B$  and  $\bar{u}^A = \bar{u}^B$ , must satisfy the inverse elasticity formula (14), whether or not horizontally equitable taxation is actually optimal. Put another way, the necessary inverse elasticity condition for optimal taxation is satisfied by the uniform tax structure, even if randomization of the tax structure is the optimal policy.

<sup>9</sup>For discussion of sufficient conditions for optimal taxation and other programming problems, see, for example, Harris (1975), Green (1975), Atkinson and Stiglitz (1980, pp. 374–375), and Chenault (1985).

This leads one to study the second-order conditions for an expenditure-minimizing tax policy. These conditions require that

$$|D| = \begin{vmatrix} 0 & -\bar{R}_p^A & -\bar{R}_p^B \\ -\bar{R}_p^A & \bar{e}_{pp}^A - \lambda \bar{R}_{pp}^A & 0 \\ -\bar{R}_p^B & 0 & \bar{e}_{pp}^B - \lambda \bar{R}_{pp}^B \end{vmatrix} < 0, \tag{15}$$

where superscripts identify the functions  $R(\bar{p}^i, \bar{u}^i)$ ,  $e(\bar{p}^i, \bar{u}^i)$ ,  $i = A, B$ , and where overbars denote evaluation at  $(\bar{p}^A, \bar{p}^B, \bar{u}^A, \bar{u}^B)$ .<sup>10</sup> Of course,  $\bar{e}_{pp}^i \triangleq \partial^2 e(\bar{p}^i, \bar{u}^i) / \partial (p^i)^2$ ,  $i = A, B$ . Computing the determinant, we have

$$|D| = -(\bar{R}_p^A)^2 (\bar{C}_p^B - \lambda \bar{R}_{pp}^B) - (\bar{R}_p^B)^2 (\bar{C}_p^A - \lambda \bar{R}_{pp}^A). \tag{16}$$

At a horizontally equitable tax structure, i.e. with  $\bar{p}^A = \bar{p}^B$  and  $\bar{u}^A = \bar{u}^B$ , (16) becomes (dropping superscripts)

$$|D| = -2\bar{R}_p^2 (\bar{C}_p - \lambda \bar{R}_{pp}). \tag{17}$$

Note, however, that if  $\bar{R}_{pp} > 0$ , the expression in parentheses is negative, and hence  $|D| > 0$ . From Proposition 1, of course, we know that  $\bar{R}_{pp} > 0$  implies that local randomization is desirable, so the failure of the second-order condition in this case is not surprising. We can now say more, however: since the first-order conditions for an extremum are met under horizontally equitable taxation, convexity of  $R$  is actually a sufficient condition for a uniform tax structure to be locally expenditure *maximizing*.<sup>11</sup>

### 5. Comparison with the Stiglitz analysis

As noted earlier, our expenditure minimization analysis of random taxation focuses on ‘efficiency’ aspects of the problem. By contrast, an analysis such as that of Stiglitz, which maximizes a social welfare function (in his case, a utilitarian one) and which disallows lump-sum transfers, in some sense mixes efficiency and equity considerations. It does so by judging a randomized tax structure to be preferred, if the gain in one household’s

<sup>10</sup>See, for example, Henderson and Quandt (1980, p. 383).

<sup>11</sup>Because of the symmetry of our problem, it is obvious that optimal random tax structures come in pairs: in one case, A faces a high tax and B a low tax, and in the other case the two are interchanged. Of course, as noted, a uniform tax structure also provides an extremum for our problem. This illustrates a general result, obtained by Chenault (1985): the number of critical points for a well-behaved non-linear programming problem must be odd. Incidentally, uniform taxation can be expenditure *minimizing* only if the revenue constraint (12) is written in equality form. Obviously the expenditure maximization problem is not well set if we impose a *minimum* revenue constraint.

utility sufficiently exceeds the loss in utility to the other household. Other social welfare functions, with different degrees of inequality aversion, i.e. with different equity implications, would evaluate randomized tax structures differently. In the limit, for example, Rawlsian maximin would prohibit any departure from uniform taxation.

Both approaches to the random taxation problem – an expenditure-minimizing approach exploiting the feasibility of lump-sum transfers, and a utilitarian or other social function-maximizing approach, in which such transfers are ruled out – would appear to be of some value. Moreover, one would expect the results in each case to bear a certain resemblance to each other, if not to be identical in detail. Indeed, our characterization of necessary and sufficient conditions for local randomization is certainly consistent with Stiglitz's informal excess burden presentation of the problem. Our results differ from those in Stiglitz's formal analysis, however.

Some of these differences are quite minor. Stiglitz's results, for example, are stated in terms of ordinary demand elasticities rather than compensated elasticities, as ours are. Given the different (expenditure-minimization) approach that we take, however, this difference is not surprising. A second difference is that while risk aversion appears in the Stiglitz formulae, no cardinal properties of the utility function are required for our analysis. But since we hold utility fixed throughout our discussion, whereas cardinality intrinsically involves comparison of different utility levels, this feature of our analysis is also understandable.

What is more striking is the fact that we find that *convexity* of the (compensated) revenue function  $R(p, \bar{u})$ , i.e.  $R_{pp} > 0$ , is sufficient for randomization. Moreover, while randomization can sometimes be optimal when  $R$  is concave, it is definitely not optimal near a (local) maximum of  $R$ . By contrast, Stiglitz derives a sufficient condition for randomization in the neighborhood of a tax rate at which, it is claimed, tax revenue is maximized. This would imply that the uncompensated revenue function<sup>12</sup>

$$R^*(p) = (p-1)C^*(p, 0), \quad (18)$$

where  $C^*(p, 0) = C^*(p, I)_{I=0}$  is the ordinary demand function, is *concave*, and in fact would be strictly concave ( $R_{pp}^* < 0$ ) at a *regular* local maximum. This conclusion is all the more puzzling given that Stiglitz's sufficient condition [his (14)] requires that  $v^*$ , the curvature of the ordinary demand function, must be sufficiently negative, i.e. the ordinary demand curve must be sufficiently convex. At least for the compensated case, we saw above that  $v$  sufficiently negative implies that  $R_{pp} > 0$ .

<sup>12</sup>We will henceforth use an asterisk to identify parameters characterizing ordinary (uncompensated) demand or revenue functions.

A simple calculation, however, shows that Stiglitz's (14), i.e.  $2 + \tau^* \nu^* < 0$ , actually implies  $R_{pp}^* > 0$ , as our analysis of the compensated case would lead one to expect. This 'sufficient convexity' of the demand curve thus contradicts the assumption that revenue is at a maximum ( $R_{pp}^* < 0$ ). Hence, Stiglitz's formal analysis does not verify the claimed optimality of randomization around a point of maximum tax revenue, such as the tax rate  $\tau^*$  in fig. 3, and thus does not contradict our conclusions.<sup>13</sup> Indeed, strictly speaking, Stiglitz's analysis only applies to situations where tax revenue is at a local *minimum*, as for example at tax rate  $\tau'$  in fig. 3. We note, however, that a uniform tax structure yielding a local minimum of tax revenue cannot be welfare-maximizing. Thus, under a correct interpretation, Stiglitz's argument, strictly speaking, only shows that randomization is desirable when the uniform tax is not a welfare-maximizing uniform tax.

The conclusion that local randomization is undesirable around a maximum revenue point is thus valid not only for expenditure-minimization problems such as we considered in section 3, but also for a utilitarian social welfare maximization problem with no lump sum transfers. Indeed, this finding is not unique to utilitarian social welfare functions: it holds for all social welfare functions, even ones that are less inequality-averse than the utilitarian (i.e. even for inequality-preferring social welfare functions). Because of its similarity to our analysis in section 2, we only sketch the proof of this assertion. First, note that the necessary and sufficient condition for local randomization requires that the social indifference contour in price space be less concave (more convex) than the iso-revenue contour around  $p^A = p^B$ . The curvature of the latter, however, is  $-2R_{pp}^*/R_p^*$ . [This is equivalent to Stiglitz's (10), and directly parallels (6) above.] Near the maximum tax revenue,  $R_p^* \approx 0$ ,  $R_{pp}^* < 0$ , and hence this iso-revenue contour becomes L-shaped. No indirect social indifference contour can be more convex than this.

<sup>13</sup>Stiglitz (pp. 9–10) presents an example in which it appears that it is optimal to randomize taxes when the tax revenue requirement is sufficiently high, but not when it is low. For interested readers, we have two comments on this example.

First, and most importantly for our purposes, the example uses the properties of the utilitarian social welfare function in a specific way. In particular, the degree of concavity of the utility function plays a fundamental role in the example, which is based on Stiglitz's (13). By contrast, our results obtain for any social welfare function, no matter how inequality averse, and hinge on the properties of the revenue function.

Second, and more specifically, Stiglitz assumes an indirect utility function that generates a demand function  $C = kp - \beta L^\gamma$ , where  $k > 0$ ,  $\beta$  and  $\gamma$  are parameters. This is derived from an underlying utility maximization problem with  $pC = \bar{L} - L$  as a budget constraint, where  $L \geq 0$  is leisure. Thus,  $pC$  cannot exceed  $\bar{L}$ . For randomization to be optimal in this example, Stiglitz requires (among other restrictions)  $\beta < 1$  and  $p$  'sufficiently large'. It is clear, however, that  $\beta < 1$  and  $pC \leq \bar{L}$  require that  $p \leq \bar{p}$ , where  $\bar{p}^{1-\beta} = k^{-1} \bar{L}^{1-\gamma}$ . That is, these particular indirect utility and demand functions are really only meaningfully defined for values of  $p$  'sufficiently small'. Perhaps further restrictions on parameter values are required for this example to work.

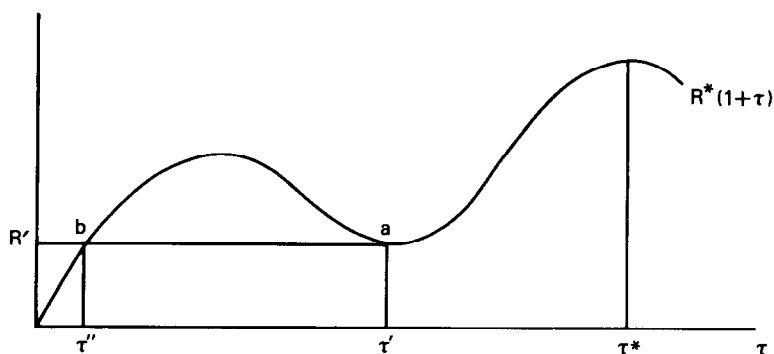


Fig. 3

## 6. Conclusion

In conclusion, we return to a comparison of our expenditure-minimization analysis with a social welfare maximization approach. As noted in section 2, we have found necessary and sufficient conditions for randomization of the tax structure to be Pareto improving, if accompanied by lump-sum interpersonal transfers. Even the most inequality-averse social welfare functions, such as Rawlsian maximin, would approve a move away from an equal tax rate, equal utility situation under the conditions we have derived. (Equivalently, randomization would raise ex ante utility for consumers under these conditions, no matter how risk averse they might be.) Obviously, allowing for a willingness to trade off one household's utility against another's through the use of less inequality-averse social welfare functions (such as the utilitarian one) would expand the range of parameter values under which randomization of the tax structure (accompanied by appropriate lump-sum transfers) would be desirable. In this sense, we have described a very stringent set of conditions for randomization – ones which do not involve any interpersonal tradeoffs, and which therefore can be stated entirely without reference to any cardinal welfare magnitudes, such as marginal social utilities of income. It would be interesting to explore further the possibilities for randomization under weaker conditions.

The possibility of lump-sum transfers plays a critical role in our analysis. For example, Pareto-improving differentiation of the tax structure generally requires such transfers since they permit compensation from low- to high-tax households who would otherwise necessarily be made worse off (ex post). In this respect, we depart significantly from the Stiglitz analysis, which disallows interpersonal transfers. The similarity in the nature of the results that we obtain given this structural difference in the two approaches is rather surprising. This, too, deserves further exploration.



## References

- Atkinson, A.B. and J.E. Stiglitz, 1980, *Lectures on public economics* (McGraw-Hill, New York).
- Balcer, Y. and E. Sadka, 1982, Horizontal equity, income taxation and self-selection with an application to income tax credits, *Journal of Public Economics* 19, no. 3, 291–310.
- Boiteux, M., 1971, On the management of public monopolies subject to budgetary constraints, *Journal of Economic Theory* 3, no. 3, 219–240.
- Chenault, L.A., 1985, Uniqueness in nonlinear programming, unpublished.
- Diamond, P. and D. McFadden, 1974, Some uses of the expenditure function in public finance, *Journal of Public Economics* 3, 3–21.
- Green, H.A.J., 1975, Two models of optimal pricing and taxation, *Oxford Economic Papers* 27, no. 3, 352–382.
- Harris, R., 1975, A note on convex-concave demand systems with an application to the theory of optimal taxation, Queen's University D.P. no. 197.
- Harris, R. and D. Wildasin, 1985, An alternative approach to aggregate surplus analysis, *Journal of Public Economics* 26, no. 3, 289–302.
- Henderson, J.M. and R.E. Quandt, 1980, *Microeconomic theory* (McGraw-Hill, New York).
- Marchand, M., 1968, A note on optimal tolls in an imperfect environment, *Econometrica* 36, nos. 3–4, 575–581.
- Mohring, H., 1970, The peak load problem with increasing returns and pricing constraints, *American Economic Review* 60, no. 4, 693–705.
- Stiglitz, J.E., 1982a, Utilitarianism and horizontal equity: The case for random taxation, *Journal of Public Economics* 18, no. 1, 1–34.
- Stiglitz, J.E., 1982b, Self-selection and Pareto-efficient taxation, *Journal of Public Economics* 17, no. 2, 213–240.
- Tresch, R., 1981, *Public finance: A normative approach* (Business Publications Inc., Plano, Texas).
- Weiss, L., 1976, The desirability of cheating incentives and randomness in the optimal income tax, *Journal of Political Economy* 84, no. 6, 1343–1352.
- Wildasin, D.E., 1984, On public good provision with distortionary taxation, *Economic Inquiry* 22, no. 2, 227–243.