Distributional Neutrality and Optimal Commodity Taxation

By David E. Wildasin*

"We are still a long way from having an intuition for resource allocation questions in economies with distorting taxes which parallels the level of intuition in first-best economies."

Peter Diamond [p. 342]

The theory of optimal commodity taxation has been elaborated and refined in recent years by a long list of notable contributors. There are, however, certain areas that need extension and clarification. Much of the discussion of optimal tax rules has focused on the case of an economy with a single consumer or many identical consumers. The restriction to a single consumer is particularly serious, because certain results appearing with great regularity in the literature are not generally valid in the many-consumer case. The more careful studies emphasize this point.[1] Manifestly, consumers are not identical, and the question arises whether we have learned, or ever could learn, anything really interesting by studying optimal taxation under this assumption. It is clearly important to understand why the many-consumer case is different.

In this paper I shall examine in detail two tax rules which have appeared in a number of writings on optimal taxation. One well-known rule states that the ratio of the additional revenue obtained by a unit increase in the tax on a commodity to the quantity of the commodity should be the same for all commodities. Another rule, due to Ramsey, states that taxes should induce (approximately) equal percentage reductions in compensated demand for all commodities. Both of these rules are valid in the identical-consumer case; neither is *generally* valid in the many-consumer case. A major purpose of the discussion below is to study the implications of two restrictions that can be imposed on the social welfare function, each of which might be considered a formalization of the idea of "distributional neutrality." One of these (simple neutrality) states that the marginal social utility of *consumption* of the numeraire good (i.e., the

*Assistant professor of economics, University of Illinois at Chicago Circle. I am grateful to the managing editor and a referee for critical comments on an earlier and much longer version of this paper; neither is responsible for remaining errors.

[1] Among those who have explicitly assumed identical consumers are Frank Ramsey in his pioneering work, Avinash Dixit, Joseph Stiglitz and Partha Dasgupta, Anthony Atkinson and Stiglitz (1972), Agnar Sandmo (1974), and Atkinson and Nicholas Stern. William Baumol and David Bradford are not explicit on the point, but their results in fact hold (generally) only in the case of identical consumers. Dasgupta and Stiglitz, James Mirrlees (1972), and Frank Hahn discuss the case of nonidentical consumers, but only in regard to the problem of profit taxation and production efficiency—matters which I will not discuss here. The many-person commodity tax problem was treated by Marcel Boiteux in his classic 1956 paper. Boiteux, it should be noted, did assume that income can be costlessly redistributed in lump sum fashion. This work was introduced to English-language economists by Jacques Drèze, who explicitly draws attention to the role of redistributive transfers in the Boiteux analysis. This point was not lost on Herbert Mohring, who essentially develops the "Ramsey rule" for many consumers by bringing redistribution directly into the welfare-maximization problem. Martin Feldstein and H.A. John Green also derive the Ramsey rule with many consumers; they impose a distributional assumption and an assumption about Engel curves, discussed below in fn. 2. Diamond and Mirrlees present optimal tax formulae for nonidentical consumers in their joint paper, and both have taken up the many-consumer problem again in separate papers (Mirrlees, 1975, and Diamond). Their approach differs from that taken here in that they do not analyze the reasons why the many-consumer formulae diverge from those encountered in the identical-consumer case. The spirit of the Mirrlees paper is in some respects similar to that of the present discussion, however. (See fn. 7 below.) Finally, a number of very recent articles have expressed concern with the usual restriction to the single-consumer case, and have emphasized that distributional issues must somehow be dealt with in the study of commodity taxation. See Richard Musgrave, Sandmo (1976), and Atkinson and Stiglitz (1976).

derivative of the welfare function with respect to the household's consumption of that good) is the same for all consumers. If one imposes this restriction on the welfare function, the first of the tax rules mentioned above will hold in the many-consumer case (Proposition 1), but the second will not (Proposition 2). The second neutrality assumption (extended neutrality) that might be imposed is that the marginal social utility of *income* (the change in social welfare associated with a one dollar increase in a household's before-tax income) is the same for all consumers. Under the assumption of extended neutrality, the first tax rule mentioned above is not optimal in the many-consumer case (Proposition 3), while the Ramsey rule is (Proposition 4).

Since these neutrality assumptions are assumptions about the specific welfare function being used, one cannot legitimately *assume* that both of these restrictions simultaneously hold. However, if one invokes the simple neutrality assumption, it turns out that a knowledge of the (potentially empirically discoverable) income derivatives of the household demand functions would be sufficient information to compute a lump sum redistribution of pretax income which, if actually performed, would lead to the satisfaction of the extended neutrality assumption. Should this be done, then obviously both tax rules stated above will be optimal (Proposition 5). If this redistribution is not actually carried out, however, both tax rules cannot simultaneously hold. This shows, as already suggested, that the single-consumer case differs fundamentally from the far more relevant many-consumer situation.

The paper is organized as follows. The first section introduces the model and presents a general many-person tax rule. In Section II, I first derive the two tax rules for the single-consumer case. Next the principal results on the extension of the two rules to the many-person economy are presented, along with some comments on the methodology of this and other studies of the optimal taxation problem. Some brief concluding remarks are found in Section III.

I. The Model and the Optimal Taxation Problem

The framework for analysis of optimal tax problems is now quite familiar, so the model can be presented very briefly. In the economy discussed here there are H households ($h = 1, \ldots, H$), a private production sector (distinguished by the superscript f), and a public production sector (superscript g). The government carries out productive activities because it is responsible for public good provision. Like other productive units, it purchases its inputs; this creates the need for taxation. Here we shall neglect the problem of optimal public good provision, and assume simply that the government has exogenously fixed private good inputs, measured negatively in the net *output* vector $y^g \equiv (y_i^g)$, where $-y_i^g$ is the amount used of the ith private good ($i = 0, 1, \ldots, n$).

The private production sector is assumed perfectly competitive and operates subject to the overall technology constraint

$$(1) \qquad \sum_{i=0}^{n} \phi_i^f y_i^f \leq 0$$

where every component of $\phi^f \equiv (\phi_i^f)$ is constant and strictly positive, and $y^f \equiv (y_i^f)$ is the aggregate *net* output vector (with $y_i^f < 0$ for a factor i). Maximized profits will thus be zero in equilibrium. Note for future reference that if $p \equiv (p_i)$ is the price vector facing producers, profit maximization implies that

$$(2) \qquad \frac{p_i}{p_k} = \frac{\phi_i^f}{\phi_k^f} \qquad i,k = 0,1,\ldots,n$$

Note that these relative prices will be invariant due to the constancy of ϕ^f.

Let $x^h \equiv (x_i^h)$ denote the net consumption vector for household h, where $x_i^h \gtrless 0$ as i is a commodity demanded or supplied. Preferences are represented by strictly quasi-concave differentiable utility functions $u^h(x^h)$. Let $q \equiv (q_i)$ denote the vector of consumer prices; we have $q = p + t$ where $t \equiv (t_i)$ is the vector of commodity taxes. Households select consumption vectors which maximize utility subject to budget constraints

$$qx^h = 0 \tag{3}$$

The absence of a term on the right-hand side of (3) reflects the absence of lump sum taxes and subsidies, and of private sector profits. Utility maximization yields the household's demand vector as a function of prices and fixed income I^h (which, as just noted, is zero). Let

$$v^h(q, I^h) \equiv u^h(x^h[q, I^h])$$

be the indirect utility function for h. It satisfies Roy's formula,

$$\frac{v_k^h}{v_I^h} = -x_k^h \tag{4}$$

where $v_k^h \equiv \partial v^h/\partial q_k$, and $v_I^h \equiv \partial v^h/\partial I^h$, the marginal utility of income.

The government attempts to maximize the social welfare function $W(v^1, \ldots, v^H)$ by a proper choice of consumer prices and producer prices. (Optimal taxes are thus determined implicitly.) In doing this it is constrained to obtain the necessary inputs for public good provision y^g, subject to the condition that markets clear. Let $X_i \equiv \sum_h x_i^h$ be the market demand for good i, with $X \equiv (X_i)$. Then market clearing requires that

$$X_i = y_i^f + y_i^g \qquad i = 0, 1, \ldots, n \tag{5}$$

The government's budget must also be balanced. Its revenues are tX, while its expenditures are $-py^g$. Hence budget balance requires that

$$tX + py^g = 0 \tag{6}$$

But if all households satisfy their budget constraints, if maximized profits are zero, and if all markets clear, (6) will automatically be satisfied. Hence (6) is a redundant constraint. Note also that if p and q are vectors of producer and consumer prices that satisfy all of the constraints, so are αp and βq for any positive α and β, *not necessarily equal.* Therefore it is possible to normalize *both* price systems; let us suppose that this is done to achieve $q_0 = p_0 = 1$. This implies, of course, that $t_0 = 0$.

Finally, rather than have the government actually choose producer prices, it is formally convenient to have it select a supply vector for the private sector directly; the producer prices that could be used to support this output vector will then be given by the relations (2). It then becomes necessary, however, to append the private production function (1) as a constraint.

The Lagrangian can now be formulated:

$$L = W(\{v^h\}) + \sum_{i=0}^{n} \rho_i(y_i^f + y_i^g - X_i) - \lambda^f \phi^f y^f$$

The first-order conditions are that

$$\frac{\partial L}{\partial q_k} = \sum_h W_h v_k^h - \sum_{i=0}^{n} \rho_i \frac{\partial X_i}{\partial q_k} = 0 \tag{7a}$$

$$\frac{\partial L}{\partial y_k^f} = \rho_k - \lambda^f \phi_k^f = 0 \tag{7b}$$

$$k = 0, 1, \ldots, n$$

where $W_h = \partial W/\partial v^h$, of course.

From (7b) and (2), $\rho_i/\rho_0 = p_i$; substitution from (4) into (7a) thus yields

$$-\sum_h \frac{W_h v_I^h}{\rho_0} x_k^h = p \frac{\partial X}{\partial q_k}$$

Differentiating the sum of all the budget constraints (3) with respect to q_k and substituting, we have

$$\sum_h \frac{W_h v_I^h}{\rho_0} x_k^h = t \frac{\partial X}{\partial q_k} + X_k \tag{8}$$

With producer prices constant (from (2)), $dt_k = dq_k$, so that (8) becomes

$$\sum_h \frac{W_h v_I^h}{\rho_0} x_k^h = t \frac{\partial X}{\partial t_k} + X_k = \frac{\partial (tX)}{\partial t_k} \tag{9}$$

Equation (9) will serve as a starting point for further derivations.

II. Sufficient Conditions for Two Optimal Tax Rules

While (9) is a general many-person tax rule, it is not expressible as one of the frequently encountered formulae of the single-consumer model, as it stands. Let me now state formally the two common tax rules alluded to earlier:

Optimal Tax Rule 1 (Marginal revenue proportionality): For every commodity, the ratio of the marginal tax revenues associated with a marginal increase in the tax rate on the commodity to the quantity of the commodity purchased is the same; that is,

$$\frac{\partial(tX)/\partial t_i}{X_i} = \frac{\partial(tX)/\partial t_k}{X_k} \qquad i,k = 1,\ldots,n$$

Optimal Tax Rule 2 (Ramsey rule): For all commodities, the percentage reduction in demand (along the compensated demand curve) due to taxation is (approximately) the same.

A. *Results for a Single Consumer*

Before going on to discuss these rules in a many-person context, it may be helpful to present a brief derivation for the single-consumer case. If there is only one consumer, (9) is simplified by eliminating the social welfare function and by identifying individual and market quantities:

$$\text{(10)} \qquad \frac{v_I}{\rho_0} X_k = \frac{\partial(tX)}{\partial t_k}$$

Clearly, division by X_k yields the marginal revenue proportionality rule.

Next, using the Slutsky relation for the derivative of a demand function with respect to a tax rate (= derivative with respect to a consumer price, producer price constant), namely,

$$\text{(11)} \qquad \frac{\partial X_i}{\partial t_k} = \left.\frac{\partial X_i}{\partial t_k}\right|_{\bar{u}} - X_k \frac{\partial X_i}{\partial I}$$

equation (10) yields

$$\frac{v_I}{\rho_0} X_k = X_k + \sum_{i=0}^{n} t_i \left.\frac{\partial X_i}{\partial t_k}\right|_{\bar{u}} - X_k t \frac{\partial X}{\partial I}$$

Using the symmetry of the substitution terms (i.e., $(\partial X_i/\partial t_k)_{\bar{u}} = (\partial X_k/\partial t_i)_{\bar{u}}$) and rearranging, we have

$$\frac{v_I}{\rho_0} - 1 + t\frac{\partial X}{\partial I} = \sum_{i=0}^{n} \frac{t_i}{X_k} \left.\frac{\partial X_k}{\partial t_i}\right|_{\bar{u}}$$

The right-hand side of this expression is the (approximate) percentage reduction in (compensated) demand for commodity k due to taxation. Since the left-hand side is independent of k, we see that this percentage reduction is the same for all commodities: the Ramsey rule. Thus both tax rules are valid and equivalent characterizations of the optimal tax structure for a single-consumer economy.

B. *The Many-Person Case: Simple Neutrality*

From one perspective, the rest of this paper is devoted to the task of finding conditions under which the above tax rules obtain in a many-consumer economy. These conditions will take the form of restrictions on the social welfare function W, which (roughly speaking) state that the distribution of welfare is in some sense equitable. As in so many other cases, there is a tremendous gain in the sharpness of analytical results when one assumes away the distributional problem and focuses (insofar as possible, and not necessarily without ambiguity) on the efficiency issues in isolation. I shall discuss the legitimacy of this approach later.

One interesting assumption that will be used below states that the marginal social utility of consumption of the numeraire good is the same for every household. I shall call this the "simple assumption of distributional neutrality."

ASSUMPTION 1 (simple neutrality): *$W_h u_0^h$ is the same for all h; that is,*

$$W_h u_0^h \equiv \mu \qquad \textit{for } h = 1,\ldots,H$$

Note that

$$v_I^h = \sum_{i=0}^{n} u_i^h \frac{\partial x_i^h}{\partial I^h} = u_0^h \sum_{i=0}^{n} q_i \frac{\partial x_i^h}{\partial I^h} = u_0^h$$

It will be more convenient to use Assumption 1 in the marginal utility of income form

$$\text{(12)} \qquad W_h v_I^h \equiv \mu$$

This is one possible characterization of neutrality, but it is not the only one, as we shall see below. For the moment, however, let us restrict our attention to Assumption 1, using it to demonstrate two propositions: first, that the simple neutrality assumption is sufficient to establish the optimality of the marginal revenue proportionality rule; second, that simple neutrality is *not* sufficient to establish the Ramsey rule. The fact that this second rule does not generally hold under simple neutrality motivates the search for another definition of neutrality.

Beginning with the general optimal tax rule (9), it is a short step to the marginal revenue proportionality rule under Assumption 1, for substituting (12) into (9), minor rearrangement yields

$$\frac{\partial(tX)/\partial t_k}{X_k} = \frac{\mu}{\rho_0} \qquad k = 1, \ldots, n$$

But this is precisely Optimal Tax Rule 1.

PROPOSITION 1: *Under the simple assumption of neutrality, the marginal revenue proportionality rule is optimal in the many-consumer case.*

On the other hand, using the Slutsky equation (11) and the neutrality condition (12) in (9) produces

$$(13) \quad \frac{\mu}{\rho_0} X_k = \sum_{i=0}^{n} t_i \left.\frac{\partial X_i}{\partial t_k}\right|_{\bar{u}} - \sum_{i=0}^{n} t_i \left(\sum_h x_k^h \frac{\partial x_i^h}{\partial I^h}\right) + X_k$$

where $\partial X_i/\partial t_k \mid_{\bar{u}}$ is the sum of the price derivatives of the compensated demand functions.

Rearranging (13), using the symmetry of the substitution terms, one has

$$(14) \quad \sum_{i=0}^{n} \frac{t_i}{X_k} \left.\frac{\partial X_k}{\partial t_i}\right|_{\bar{u}} = \frac{\mu}{\rho_0} - 1 + \sum_h \frac{x_k^h}{X_k} \frac{\partial tx^h}{\partial I^h}$$

The left-hand side of (14) is the percentage reduction in demand along the compensated demand for good k. Observe that the right-hand side of (14) will vary from commodity to commodity (i.e., will vary with k) because the share of commodity k consumed by h, x_k^h/X_k, will vary with k; hence, the weights attached to the terms $\partial tx^h/\partial I^h$ will vary. Moreover, if income derivatives of demand vary from consumer to consumer, as one would expect, these latter terms will all be different as well. Hence, only in the special case where each household consumes the same fraction of the aggregate amount of each commodity and/or has identical income derivatives of demand, as would be the case with a single consumer or many identical consumers, will the Ramsey rule be valid.[2]

PROPOSITION 2: *Under Assumption 1, the Ramsey rule is not generally valid in the many-consumer case.*

C. *An Alternative Neutrality Concept and the Methodology of Optimal Tax Analysis*

As mentioned above, Assumption 1 does not imply that the welfare maximizer would not, if possible, carry out lump sum redistribution among households. To see this, suppose that although all *net* government expenditures (i.e., expenditures for public good provision) must be commodity tax financed, so that the balanced budget constraint (6) continues to hold, the government is nevertheless permitted to distribute

[2]Feldstein and Green arrive at the Ramsey rule by in effect imposing Assumption 1 (since the Feldstein-Green "distributional characteristic" is the same for every commodity in this case), and by further imposing the Gorman-Nataf condition that the Engel curves for all individuals are parallel straight lines. To see this result, note that the latter condition implies that for each i, $q_i(\partial x_i^h/\partial I^h)$ is the same for all h. But then $(q_i - p_i)(\partial x_i^h/\partial I^h) = t_i(\partial x_i^h/\partial I^h) = \sigma_i$ is the same for all h. Substituting in $\Sigma_i \sigma_i$ for the last term on the right-hand side of (14) produces an expression independent of k, the Ramsey rule. I am concerned here with the *general* many-person problem, however, and do not wish to impose such restrictions on consumer behavior.

to each household a lump sum amount of the numeraire commodity, s_0^h; this will be negative for some households (corresponding to a lump sum tax). It is required that the *sum* of all these be zero, that is,

$$(15) \qquad \sum_h s_0^h = 0$$

Thus one can imagine that the government is allowed to determine the initial distribution of income via costless lump sum transfers; once this is done, it is restricted to use commodity taxation to achieve a net transfer of resources from the private to the public sector. In this situation, the household's budget constraint would become

$$(16) \qquad qx^h = s_0^h$$

Note that $\partial v^h/\partial s_0^h \equiv v_I^h$, $\partial x_i^h/\partial s_0^h \equiv \partial x_i^h/\partial I^h$, etc.

In setting up the new welfare-maximization problem, it is necessary to add (15) as a constraint. The Lagrangian is therefore

$$L = W(\{v^h\}) + \sum_{i=0}^{n} p_i(y_i^f + y_i^g - X_i) - \lambda^f \phi^f y^f + \theta \sum_h s_0^h$$

The first-order conditions are as before (equations (7)) with the addition of

$$(17) \quad W_h v_I^h - \sum_{i=0}^{n} p_i \frac{\partial x_i^h}{\partial I^h} + \theta = 0 \qquad h = 1, \ldots, H$$

It would seem reasonable to interpret (17) as an alternative formulation of the idea of "equal social marginal utilities of income"; or perhaps the phrase "equal social marginal net benefits of income" would be better. Equation (17) says that one equates the social marginal utility or benefit of income for all households, where this benefit is the marginal welfare of the marginal utility of income (or consumption) to the household *less* the marginal welfare cost of the additional consumption that the household would carry out as the result of an addition to income. Alternatively, differentiating (16) with respect to s_0^h, one can rewrite (17) as

$$\frac{W_h v_I^h}{p_0} + \sum_{i=0}^{n} (q_i - p_i) \frac{\partial x_i^h}{\partial I^h} + \theta - 1 = 0$$

or

$$(18) \qquad \frac{W_h v_I^h}{p_0} + \frac{\partial(tx^h)}{\partial I^h} = 1 - \theta$$

Equation (18) suggests the interpretation that if possible the government would (*ceteris paribus*) redistribute income in favor of those whose additional spending from another dollar of income would generate the most additional taxes.

Obviously equation (18) will typically *not* be satisfied if the government is not permitted to carry out this ideal redistributive scheme. Therefore, to assume that (18) is satisfied in the absence of such a scheme is an assumption about the welfare function W. This assumption has a claim to be considered as a formalization of the notion of distributional neutrality: if met, the welfare maximizer would *choose* not to carry out any lump sum redistribution of income, even if empowered to do so. On the other hand, we note that equation (18) is not equivalent to Assumption 1, simple neutrality. Clearly, $W_h v_I^h = \mu$ for all h neither implies, nor is implied by, equation (18). Thus, even if one assumes simple neutrality, the government would still wish to engage in lump sum redistribution, if it could. This suggests that (12) may not be the best (and certainly is not the only) definition of distributional neutrality. In attempting to isolate and suppress the purely distributional aspects of the problem of optimal commodity taxation, then, one might wish to invoke the following assumption:

ASSUMPTION 2 (extended neutrality): *Even if empowered to do so, the government will not carry out lump sum redistribution; that is, the welfare function W is such as to satisfy equation* (18).

Before going on to indicate the application of Assumption 2, I would like to discuss its relationship to Assumption 1, and to raise some troubling methodological points.

First, having motivated Assumption 2 by noting that it would be satisfied as a consequence of lump sum redistributional measures, it is useful to reconsider Assumption 1 from the same perspective. Suppose that households possess given consumption bundles, and that the government is able to redistribute the numeraire good in lump sum fashion subject to the condition that no household is permitted to trade away from its posttransfer consumption bundle. Then clearly Assumption 1 will characterize the optimal set of transfers. Thus, one can think of Assumption 1 as a neutrality assumption corresponding to the situation after all trades have taken place and markets are closed, whereas Assumption 2 characterizes distributional neutrality in an *ex ante* sense, allowing for the effects of redistribution on the purchases of households.[3]

Note that in the absence of commodity taxes, all t_i's are zero, (18) reduces to (12), and the distinction between simple and extended neutrality disappears. The possibility of two different neutrality concepts depends essentially on the existence of commodity taxes.

One immediate objection to either of these assumptions is that it is strange to suppose that the government can redistribute income in lump sum fashion but cannot raise *net* revenues in a nondistortionary way. It would seem that if the government could devise optimal lump sum redistributions of income, it should be able to avoid distortionary taxation altogether. On the other hand, one might turn this argument around and observe that if the government has the ability to compute optimal commodity taxes—if it can impose a tax system satisfying condition (9), for example—then it should be able to compute optimal lump sum taxes. For to raise a given amount of revenue in lump sum fashion, the government needs "only" to know the marginal social utility of income or consumption (the two are the same with lump sum taxes) for every consumer, which is clearly less information than is needed for optimal commodity taxes (in the general case). Once given a completely specified social welfare function which can be used to resolve all issues of ethical deservingness, it *is* possible to devise ideal taxes: if it is desired that Paul should be made better off at the expense of Peter, then tax Peter relatively heavily because he is Peter and tax Paul relatively lightly because he is Paul. Such taxes are unrelated to the economic behavior of the individuals and hence lump sum.

I would conjecture that the discussion of optimal commodity taxation has been motivated by the observation that most widely practiced forms of taxation (a) are based on the economic behavior of the individual (for example, income taxes) and (b) seem to be designed, at least in part, to achieve distributional objectives (for example, many income tax systems are progressive; income supplements of various kinds are usually directed toward the poor, etc.). The latter point suggests that somewhere the ethical problems have all been worked out, so that a social welfare function exists; lump sum taxes would certainly be used to maximize this function if feasible; (a) suggests that they are not used; hence, lump sum taxes are infeasible.

But why should lump sum taxes be infeasible? Certainly the physical act of taxation on a lump sum basis is feasible; on the administrative level, lump sum taxes would seem to involve smaller collection costs than a complex system of commodity taxation.[4] And as just noted, the informational requirements for optimal lump sum taxation are less severe than those for optimal commodity taxes. I think it should be clear that there is nothing infeasible about optimal lump sum taxes except the construction of the social welfare function needed to

[3]I should note here that Diamond also distinguishes between the marginal social utilities of income and consumption, which underlies the distinction between simple and extended neutrality; he does not, however, introduce any neutrality assumption.

[4]Walter Perrin Heller and Karl Shell have discussed the costs of commodity taxation and their implications for production efficiency.

compute them; but then optimal *commodity* taxes may also be infeasible.

This discussion obviously raises serious questions about the whole approach to the study of optimal commodity taxation as an exercise in social welfare maximization.[5] On the other hand, setting up a welfare-maximization problem does permit one to characterize the set of Pareto (quasi-) optimal tax systems,[6] and it may be possible to find particular systems which can be implemented without specific reference to an underlying social welfare function. This offers hope that the welfare-maximization technique can be used to obtain relatively practicable tax rules which at least permit the attainment of Pareto optimality.[7]

I would hasten to point out however that tax rules which do not explicitly involve the social welfare function have at best tenuous claims to special attention as ethically neutral "efficient" tax rules. Each of the two neutrality assumptions introduced above leads to a different tax formula, as reference to Propositions 1 and 4 (below) will verify; which formula, if either, describes the efficient tax system? Each rule leads to a particular point on the Pareto frontier, and there is no efficiency basis for choosing between them. There is simply no unambiguous way to segregate the efficiency and distributional aspects of the optimal tax problem; however much one is tempted to apply the intuitive insights derived from the study of the single-person economy, they are of limited value in the more interesting many-person context.

The foregoing discussion has raised some perplexing questions about the proper approach to the study of optimal taxation problems and about the interpretation of certain results. Unfortunately I shall not be able to satisfactorily resolve these questions here, and in the remainder of this paper, I shall proceed uncritically within the standard framework. Assumptions 1 and 2 have a number of implications which shed some light on the critical differences between single-consumer and many-consumer economies; these should be of interest in themselves, however one feels about their interpretation in view of the above remarks.

D. *Extended Neutrality and Optimal Taxation*

Some of the implications of the simple neutrality assumption have already been examined and summarized in Propositions 1 and 2. Let us now exploit the extended neutrality assumption in a similar fashion.

Substituting from equation (18) into (9), and dividing through by X_k yields

$$(19) \quad (1 - \theta) - \sum_h \frac{x_k^h}{X_k} \frac{\partial(tx^h)}{\partial I^h} = \frac{\partial(tX)/\partial t_k}{X_k}$$

which does not reduce to the marginal revenue proportionality rule because the weight x_k^h/X_k that is attached to each term in the sum on the left-hand side will depend on k. Only if every household purchases the same fraction of each commodity or has the same income derivatives of demand for all commodities, as would be true with one consumer or many identical consumers, will the extended neutrality assumption be sufficient to establish Optimal Tax Rule 1.[8] This demonstrates

PROPOSITION 3: *Under Assumption 2, the*

[5]Hahn's comments, pp. 105–06, are particularly to the point here.

[6]Since $W_h > 0$ for all h, any tax system satisfying the general condition (9) must be Pareto optimal. (These are second best optima, of course.) The selection of a particular W function singles out one of the points along the Pareto frontier as best; as we vary the form of the welfare function, we trace out what Paul Samuelson would call the utility-feasibility frontier.

[7]As Hahn has said, "Optimum tax formulas are either guides to action or nothing at all" (p. 106). Will tax rules that depend on the partial derivatives of a social welfare function ever be suitable guides to action? In this regard compare the remarks of Mirrlees in discussing his two-class rule: "The particular appeal of the result is that it does not refer to the relative social marginal utilities of the two classes" (1975, p. 30). Diamond's rule—"the change in aggregate compensated quantity demanded is proportional to the covariance between individual quantities demanded and social marginal utilities of income" (p. 335)—though perhaps simpler in form than other statements of the general many-person rule, does not appear to lend itself to ready application.

[8]See (14) above. Note that the Gorman-Nataf condition of fn. 2 will again do the trick; making the appropriate substitution in the left-hand side of (19) gives an expression independent of k.

marginal revenue proportionality rule is not valid in the general many-person case.

Next, using the Slutsky equation in the right-hand side of (19), we have

$$1 - \theta - \sum_h \frac{x_k^h}{X_k} \frac{\partial (tx^h)}{\partial I^h} = 1 + \sum_{i=0}^{n} \frac{t_i}{X_k} \frac{\partial X_i}{\partial t_k}\bigg|_{\bar{u}} - \sum_h \frac{x_k^h}{X_k} \frac{\partial (tx^h)}{\partial I^h}$$

By virtue of the symmetry of the substitution terms,

$$\sum_{i=0}^{n} \frac{t_i}{X_k} \frac{\partial X_k}{\partial t_i}\bigg|_{\bar{u}} = -\theta \tag{20}$$

Thus we have demonstrated

PROPOSITION 4: *Under Assumption 2, the Ramsey rule is optimal in the many-person case.*

To digress momentarily, it is interesting to note from (20) that the percentage reduction in demand along the compensated demand curve should equal the negative of the Lagrange multiplier for the constraint (15). In general, a Lagrange multiplier shows the marginal effect on the objective function of a slight easing of the associated constraint;[9] in the present case, this means that θ is the marginal welfare associated with a one dollar increase (from zero) of the amount of *net* revenues raised via lump sum taxes. It is the "distortion" of commodity taxes that causes the divergence from the first best situation, and so it is not surprising to find a relationship between changes in demand and the welfare cost of commodity taxes; but, intuitively, it is not clear that one would expect to find this precise relationship. In any case, the result would seem to be sufficiently striking to merit special mention.[10]

PROPOSITION 4′: *Under Assumption 2, the percentage reduction in demand along the compensated demand curves is equal to minus the welfare gain associated with a marginal increase in the amount of net revenues that may be raised via lump sum taxes.*

Finally, suppose that one invokes both Assumptions 1 and 2. Of course, both of these cannot hold as *assumptions* about the welfare function, since each would specify a *different* function. Rather, one can think of assuming simple neutrality and then (in line with our earlier remarks) actually carrying out a pretax redistribution of income so as to achieve the satisfaction of (18). In this situation, the following result can easily be established:

PROPOSITION 5: *When both Assumptions 1 and 2 hold, both the marginal revenue proportionality and Ramsey rules are optimal (and equivalent) in the general many-consumer case.*

III. Conclusion

The major objective of this paper has been to analyze the problem of optimal taxation in a many-consumer economy, to see whether and under what circumstances certain familiar results from the single-consumer case apply in this more interesting context. One might expect, perhaps, that the set of such circumstances is empty due to the inherent complexities introduced by nonhomogeneous tastes. Alternatively, one might think that the restriction to a single consumer has been imposed in an attempt to analyze the efficiency aspects of the optimal tax problem, leaving aside the equity or distributional issues. As we have seen, the latter view is more nearly the correct one: under certain purely distributional assumptions, the standard one-consumer results do apply in the many-person case. To be more precise, both the marginal revenue proportionality and Ramsey rules will result in Pareto (quasi-) optimal allocations, given that commodity taxes must be used for raising revenues. However, there is no efficiency criterion for choosing between these two rules, nor is there any reason to believe that either will result in a particularly desirable

[9]See, for example, Michael Intriligator, pp. 36–38, 60–62.

[10]Atkinson and Stern recognize that θ is the marginal welfare gain associated with an increase in net lump sum revenues (see their Appendix, pp. 126–27). However, they do not appear to notice the connection with the Ramsey rule.

allocation (defining "desirable" in terms of some ethical criterion).

Indeed, distributional questions are really of the essence of the commodity taxation problem. After all, if one were merely concerned to achieve a Pareto optimum, one could dispense with commodity taxes altogether: a uniform head tax ensures a distortionless *first* best optimum. Thus the real relevance of the frequently cited propositions of the efficiency-oriented single-person model remains obscure. Moreover, as my earlier remarks indicate, the incorporation of the distributional aspects of the problem in the familiar social welfare function may also leave no natural motivation for the introduction of commodity taxes. Perhaps the whole approach to the study of optimal commodity taxation as an exercise in social welfare maximization of the usual type is due for reconsideration.

REFERENCES

A. B. Atkinson and N. H. Stern, "Pigou, Taxation, and Public Goods," *Rev. Econ. Stud.*, Jan. 1974, *41*, 119–28.

——— **and J. E. Stiglitz,** "The Structure of Indirect Taxation and Economic Efficiency," *J. Publ. Econ.*, Apr. 1972, *1*, 97–119.

——— **and** ———, "The Design of Tax Structure: Direct Versus Indirect Taxation," *J. Publ. Econ.*, July/Aug. 1976, *6*, 55–75.

W. J. Baumol and D. F. Bradford, "Optimal Departures from Marginal Cost Pricing," *Amer. Econ. Rev.*, June 1970, *60*, 265–83.

M. Boiteux, "Sur la gestion des Monopoles Publics astreints à l'équilibre budgétaire," *Econometrica*, Jan. 1956, *24*, 22–40; translated by W. J. Baumol and D. F. Bradford as "On the Management of Public Monopolies Subject to Budgetary Constraints," *J. Econ. Theory*, Sept. 1971, *3*, 219–40.

P. Dasgupta and J. E. Stiglitz, "On Optimal Taxation and Public Production," *Rev. Econ. Stud.*, Jan. 1972, *39*, 97–103.

P. Diamond, "A Many Person Ramsey Tax Rule," *J. Publ. Econ.*, Nov. 1975, *4*, 335–42.

——— **and J. A. Mirrlees,** "Optimal Taxation and Public Production," *Amer. Econ. Rev.*, Part I, Mar. 1971, *61*, 8–27; Part II, June 1971, *61*, 261–78.

A. Dixit, "On the Optimum Structure of Commodity Taxes," *Amer. Econ. Rev.*, June 1970, *60*, 295–301.

J. Drèze, "Some Postwar Contributions of French Economists to Theory and Public Policy, with Special Emphasis on Problems of Resource Allocation," *Amer. Econ. Rev.*, June 1964, Suppl., *54*, 1–64.

M. S. Feldstein, "Distributional Equity and the Optimal Structure of Public Prices," *Amer. Econ. Rev.*, Mar. 1972, *62*, 32–36; "Errata," Sept. 1972, *62*, 763.

H. A. J. Green, "Two Models of Optimal Pricing and Taxation," *Oxford Econ. Pap.*, Nov. 1975, *27*, 352–82.

F. H. Hahn, "On Optimum Taxation," *J. Econ. Theory*, Feb. 1973, *5*, 96–106.

W. P. Heller and K. Shell, "On Optimal Taxation with Costly Administration," *Amer. Econ. Rev. Proc.*, May 1974, *64*, 338–45.

Michael D. Intriligator, *Mathematical Optimization and Economic Theory*, Englewood Cliffs 1971.

J. A. Mirrlees, "On Producer Taxation," *Rev. Econ. Stud.*, Jan. 1972, *39*, 105–11.

———, "Optimal Taxation in a Two-Class Economy," *J. Publ. Econ.*, Feb. 1975, *4*, 27–33.

H. Mohring, "The Peak Load Problem with Increasing Returns and Pricing Constraints," *Amer. Econ. Rev.*, Sept. 1970, *60*, 693–705.

R. A. Musgrave, "ET, OT and SBT," *J. Publ. Econ.*, July/Aug. 1976, *6*, 3–16.

F. P. Ramsey, "A Contribution to the Theory of Taxation," *Econ. J.*, Mar. 1927, *37*, 47–61.

P. A. Samuelson, "Evaluation of Real National Income," *Oxford Econ. Pap.*, Jan. 1950, *2*, 1–29.

A. Sandmo, "A Note on the Structure of Optimal Taxation," *Amer. Econ. Rev.*, Sept. 1974, *64*, 701–6.

———, "Optimal Taxation—An Introduction to the Literature," *J. Publ. Econ.*, July/Aug. 1976, *6*, 37–54.

J. E. Stiglitz and P. Dasgupta, "Differential Taxation, Public Goods, and Economic Efficiency," *Rev. Econ. Stud.*, Apr. 1971, *38*, 151–74.