Commodity Taxation an Social Welfare: The Generalised Ramsey Rule

by

David Coady
International Food Policy Research Unit, Washington, DC

and

Jean Drèze
Delhi School of Economics

The Suntory Centre
Suntory and Toyota International Centres for
Economics and Related Disciplines
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
Tel: (020) 7955 6674

DEDPS 27
September 2000

---

* We are most grateful to Avinash Dixit and Agnar Sandmo for illuminating comments.

Abstract

Commodity taxes have three distinct roles: (1) revenue collection, (2) interpersonal redistribution, and (3) resource allocation. The paper presents an integrated treatment of these three concerns in a second-best general equilibrium framework, which leads to the "generalised Ramsey rule" for optimum taxation. We show how many standard results on optimum taxation and tax reform have a straightforward counterpart in this general framework. Using this framework, we also try to clarify the notion of "deadweight loss" as well as the relation between alternative distributional assumptions and the structure of optimum taxes.

Keywords: Commodity taxes, efficiency, redistribution, shadow prices.

JEL Nos. : H21, H23.

Contact addresses:
Dr David Coady, Food Consumption Nutrition Division, International Food Policy Research Unit, 2033 K St NW, Washington , DC 2006, USA. Email: d.coady@cgiar.org

Dr Jean Drèze, Centre for Development Economics, Delhi School of Economics, Delhi 110007, India. Email: dreze@econdse.org

COMMODITY TAXATION AND SOCIAL WELFARE: THE GENERALISED RAMSEY RULE

1. Introduction

This paper is concerned with the welfare economics of commodity taxation. It focuses both on the "tax reform" problem (i.e. how a given system of commodity taxes can be improved) and on the "optimum taxation" problem (i.e. the choice of optimum taxes). As is already clear from the literature, there is a straightforward connection between the two problems, since an optimum is essentially a point where no welfare-improving reform is possible. In that sense, much of the theory of optimum taxation can be seen as a special application of the theory of tax reform.[1]

The paper has three specific objectives. First, we present a synthesis of the classical Ramsey tax problem (Ramsey, 1927) which attempts to clarify the relation between different approaches to this problem. Second, we show how many of the standard results can be generalised in a second-best general equilibrium framework using the notion of "shadow taxes", defined as the difference between consumer prices and shadow prices. Third, we try to clarify the relation between alternative distributional assumptions and the structure of optimum taxes.

Our analysis makes considerable use of the "generalised Ramsey rule" for optimum taxation, derived in Guesnerie (1979) and Drèze and Stern (1987). This is a general-equilibrium extension of the classical "Ramsey rule", due to Ramsey (1927). The generalised Ramsey rule allows for a wide range of distortions in the economy, which are captured by a vector of "shadow prices". Many of the standard results continue to apply, after replacing actual commodity taxes by "shadow taxes". Conditional on the knowledge of shadow prices, the generalised Ramsey rule is a powerful tool for

---

[1] Most of the theory of tax reform is concerned with <u>marginal</u> reforms. A point where no welfare-improving reform of this type exists is only a <u>local</u> optimum. The optimum taxation literature, however, rarely distinguishes between local and global optima (and makes liberal use of ad hoc assumptions to avoid undesirable local optima).

the analysis of commodity taxation in distorted economies. The issue of how shadow prices are computed is outside the scope of this paper, and belongs to the theory of cost-benefit analysis. The relation between cost-benefit analysis, tax reform and optimum taxation is explored in Drèze and Stern (1987); this paper builds on the general approach presented there.[2]

In general terms, indirect taxes fulfill three distinct roles. First, they generate government revenue. Second, indirect taxes can be used for distributional purposes. Third, they have allocative effects. While the classical Ramsey rule focuses on the revenue-raising objective, the generalised Ramsey rule integrates these three concerns: (1) revenue collection, (2) interpersonal redistribution, and (3) resource allocation. This insight provides an important bridge between the literature generated by Ramsey's original paper, on the one hand, and a somewhat independent literature which focuses on taxes as a means of achieving a better allocation of resources (e.g. the focus on "Pigovian taxes" in environmental economics belongs to the latter tradition).

The outline of the paper is as follows. In the next section, we explore various aspects of the "classical Ramsey tax problem". Section 3 shows how shadow prices can be used to extend the classical results in a general equilibrium framework, based on the "generalised Ramsey rule". Section 4 examines the relation between income distribution and the structure of optimum taxes. The last section concludes.

2. <u>The Classical Ramsey Tax Problem</u>

For expositional clarity, we begin with a brief overview of the classical Ramsey tax problem.

---

[2] On the calculation of shadow prices and their application to tax reform and public policy, see also Ahmad, Coady and Stern (1988); Ahmad and Stern (1991); and Coady (1997).

The results presented in this section are not new, but we hope to clarify the relation between different ways of deriving and interpreting them.[3] We also take our first step towards generalisation, by looking at the problem from the point of view of "tax reform" rather than "optimum taxation".

2.1 The model

The problem initially posed by Ramsey (1927) is as follows: how should commodity taxes be set so as to minimise the loss of utility to the consumer subject to raising a given amount of revenue?[4] Formally, this problem may be written as:

$$\max_{\mathbf{t}} V(\mathbf{p}+\mathbf{t}) \qquad \text{s.t.} \qquad \mathbf{t}.\mathbf{x}(\mathbf{p}+\mathbf{t}) = T \tag{1}$$

where $\mathbf{p} \equiv (p_1,...,p_n)$ is a vector of producer prices, $\mathbf{t} \equiv (t_1,...,t_n)$ a vector of commodity taxes, V(.) the indirect utility function, $\mathbf{x}$(.) the n-dimensional vector of net consumer demands and T the amount of revenue to be raised.[5] The first-order conditions of this optimisation problem lead directly to the "Ramsey rule", on which more below. Two aspects of this formulation of the problem are worth noting: (i) there is a single consumer, and (ii) producer prices are assumed to be fixed. The problem

[3] Important contributions to the analysis of the Ramsey tax problem include Ramsey (1927), Dixit (1970, 1975), Diamond and Mirrlees (1971), Diamond (1975), Atkinson and Stern (1974), Sandmo (1974), Atkinson and Stiglitz (1976), and Deaton and Stern (1986). Useful presentations of the main results can be found in Sandmo (1976), Atkinson and Stiglitz (1980), Auerbach (1985), Dixit (1987), and Myles (1995), among others.

[4] As Ramsey himself put it: “The problem I propose to tackle is this: a given revenue is to be raised by proportionate taxes on some of or all uses of income, the taxes on different uses being possibly at different rates; how should these rates be adjusted in order that the decrement of utility may be a minimum?” (Ramsey, 1927, p.47).

[5] Most of the notation in this paper follows Drèze and Stern (1987), partly to ensure that the relation between the two papers is as transparent as possible. In particular, commodities are indexed by subscripts i=1,...n, and consumers (when there are several) by superscripts h=1,...,H; other subscripts indicate partial derivatives. Also, we use bold type to indicate vectors.

also appears to have a "partial equilibrium" format, but in section 3 below we show that it does have general equilibrium foundations in a specific class of cases.

Our own starting point is a variant of the problem initially posed by Ramsey. Following the same notation as in (1), consider an economy with a single consumer who has utility V(**p**+**t**,m), where m is a lump-sum transfer from the government.[6] Let T ≡ **t**.**x** - m denote government revenue. Social welfare in this economy may then be written as:

$$V^* \equiv V(\mathbf{p}+\mathbf{t},m) + \ddot{e}(\mathbf{t}.\mathbf{x} - m) \qquad (2)$$

where ë is a suitable (positive) "shadow price" for government revenue.[7] We start from an <u>arbitrary</u> vector of commodity taxes, say $\mathbf{t}^0$, and consider a small change d**t** from $\mathbf{t}^0$ (a "tax reform"). We begin with the simple case where the tax reform consists of changing a single tax, say t_i.[8] Given a small change dt_i in t_i (with $dq_i \equiv dt_i$), the change in social welfare (or the "marginal social value of t_i") is:

$$\frac{\partial V^*}{\partial t_i} = -\hat{a}\; x_i + \ddot{e}\;(x_i + \boldsymbol{t}.\frac{\partial \boldsymbol{x}}{\partial t_i}) \qquad (3)$$

where $\hat{a} \equiv V_m$ is the marginal utility of income. As (3)indicates, a small change dt_i induces: (1) a change in utility equivalent to a lump-sum transfer of $-x_i$ (using Roy's identity), and (2) a change in

[6] Textbook presentations of the Ramsey tax problem often rule out lump-sum transfers (i.e. m≡0); we return to this issue further on. Our formulation of the problem helps to clarify the underlying distributional assumptions.

[7] Note that V^* reflects a particular (arbitrary) cardinalization of V, which also affects ë. Any cardinalization (e.g. money-metric utility) will do.

[8] The generalization to simultaneous marginal changes in all taxes is straightforward and is discussed later on.

revenue of (x_i + **t**.∂**x**/∂q_i). The two components have to be added, <u>with suitable weights</u>, to derive the total change in social welfare. The two components differ by the term **t**.∂**x**/∂q_i, which may be termed the "marginal deadweight loss" (MDL) from the proposed tax change.

2.2 <u>The deadweight loss</u>

Equation (3) may also be written as

$$\frac{\partial V^*}{\partial t_i} \equiv (\text{ë} - \text{â})\; x_i + \text{ë}\; \text{ä}_i \tag{4}$$

where

$$\text{ä}_i \equiv \boldsymbol{t}.\frac{\partial \boldsymbol{x}}{\partial q_i} \tag{5}$$

is the marginal deadweight loss and can be interpreted as the indirect change in revenue arising from the tax change. For convenience we also define the "normalised" MDL as:

$$\text{ó}_i \equiv \frac{\boldsymbol{t}.\dfrac{\partial \boldsymbol{x}}{\partial q_i}}{x_i} \tag{6}$$

which, much like an elasticity, is independent of the choice of units. The first term in (4) is zero <u>if and only if</u> ë and â are the same, i.e. government revenue and private income are deemed to have the same marginal social value. In that case, the MDL is a correct first-order measure of the change in social welfare induced by a small change in t_i. If ë and â are <u>not</u> the same, then MDL can still be thought of as a useful notional quantity which measures some kind of (usually negative) efficiency effect of extra taxes.

The interpretation of $ä_i$ as an efficiency effect can be clarified by relating the marginal deadweight loss to a more general criterion of cost-benefit analysis, the "aggregate benefits criterion" (ABC).[9] The ABC criterion for evaluating a particular policy change (e.g. a public-sector project, a tax reform, or a modification of quantity controls) consists of asking every agent in the economy how much the change is "worth" to him or her in terms of some pre-specified numéraire, and taking the unweighted sum of all these net benefits as a measure of the social worth of the proposed change. By analogy, it is tempting to suggest using the MDL as a measure of the welfare effect of a small tax change. The idea, much as with the ABC criterion, is that if the MDL is positive, then the loser (the consumer) can notionally compensate the gainer (the government) for renouncing the proposed change: for a small change dt_i, x_i is the net loss to the consumer, $(x_i + \mathbf{t}.\partial\mathbf{x}/\partial q_i)$ is the net benefit to the government, and the MDL is the difference. This suggestion, however, capsizes not only on the usual pitfalls of "compensation criteria" but also because the notion that private income and government revenue do not have the same marginal social value (i.e. $â \neq ë$) is a central motivation for commodity taxation in the first place.

Deadweight loss measures, however, do have a role in identifying desirable tax reforms that leave revenue unchanged. To see this, let us consider the trade-off between consumer welfare and government revenue when t_i varies (locally). More precisely, let us ask how much revenue is raised when the tax on commodity i is increased to the extent that the consumer suffers a decline in utility

[9] On this, see Drèze (1998), and the literature cited there.

equivalent to the loss of one unit of income.[10] The latter convention implies $dt_i = 1/x_i$, and the increase in revenue from such a tax change (say Ä_i) is then:

$$\text{Ä}_i \equiv \frac{x_i + \boldsymbol{t}.\frac{\partial \boldsymbol{x}}{\partial q_i}}{x_i} \equiv 1 + \text{ó}_i \qquad (7)$$

Here, the numerator measures the responsiveness of tax revenue (T) to a small change in t_i and the denominator captures the responsiveness of consumer utility. The ratio between the two is the trade-off between revenue and utility. From (7), it immediately follows that, if $\text{ó}_i > \text{ó}_j$, then a (marginal) <u>switch</u> from t_j to t_i is desirable. That is, a small change (dt_i,dt_j) which leaves T unchanged, with $dt_i > 0$, improves consumer utility and therefore (given that T is unchanged) social welfare.[11] Thus, tax reform should aim at switching away from commodities with a high normalized MDL towards those with a low normalized MDL.

In short, the MDL has two possible interpretations. First, it is a measure of change in social welfare under specific (and strong) distributional assumptions. Second, it can be interpreted in terms of a trade-off between revenue and utility, i.e. it is a correct measure of welfare change <u>for tax reforms that leave revenue unchanged</u>.

2.3 <u>An alternative formula</u>

Going back to (3) and (4), the same identities may be written as follows, using the Slutsky

---

[10] Equivalently, one could consider how much consumer welfare declines when the tax on commodity i is raised just enough to raise one unit of extra revenue. In both cases, what matters is the trade-off between utility and revenue.

[11] Alternatively, a small change (dt_i,dt_j) which leaves utility unchanged with $dt_i>0$, increases the tax revenue.

equation:

$$\frac{\partial V^*}{\partial \boldsymbol{t}_{\boldsymbol{i}}} = - b \; x_i + \text{ë} \; \boldsymbol{t}.\frac{\partial \boldsymbol{x}^{\boldsymbol{c}}}{\partial q_i} \qquad (8)$$

where

$$b \equiv \frac{\partial V^*}{\partial m} = \text{â} - \text{ë} + \text{ë} \; \boldsymbol{t}.\frac{\partial \boldsymbol{x}}{\partial m} \qquad (9)$$

and $\mathbf{x}^c(.)$ denotes <u>compensated</u> demand functions. Following Drèze and Stern (1987), we refer to b as the "marginal social value of transfer". Indeed, b measures the change in social welfare induced by a small increase in the lump-sum transfer m. Note that, much as in (3) and (4), the last term in (9) (the "tax propensity" $\mathbf{t}.\partial\mathbf{x}/\partial$m, multiplied by ë) may be thought of as a marginal deadweight loss from a lump-sum transfer. It may seem surprising that lump-sum transfers involve a deadweight loss. However, in a second-best environment, actual lump-sum transfers (as opposed to the "notional" transfers involved in compensation criteria) have non-neutral allocative effects that need to be taken into account. To put it another way, following much the same reasoning as in (7) we may interpret (1+ $\mathbf{t}.\partial\mathbf{x}/\partial$m) as the <u>trade-off</u> between consumer welfare and government revenue when m varies.

By analogy with (4), (8) may also be written as:

$$\frac{\partial V^*}{\partial t_i} = - b \; x_i + \text{ë} \; \text{ä}_i^c \qquad (10)$$

where

$$\text{ä}_i^c \equiv \boldsymbol{t}.\frac{\partial \boldsymbol{x^c}}{\partial q_i} \qquad (11)$$

is an alternative measure of the MDL. To distinguish ä^c_i from ä_i, we shall refer to the former as the "marginal excess burden" (MEB), even though the terms "excess burden" and "deadweight loss" tend to be used interchangeably in the literature. Much as before, we define the normalised MEB as $\text{ó}^c_i \equiv \text{ä}^c_i/x_i$. Following the same reasoning as in the preceding sub-section, it is then easy to show from (10) that, if $\text{ó}^c_i > \text{ó}^c_j$, then a switch from tax j to tax i is desirable. This result also follows from noting that ó^c_i and ó_i differ only by a constant. Indeed, using the Slutsky equation again, we have:

$$\text{ó}_i \equiv \frac{1}{x_i}\ \boldsymbol{t}.\frac{\partial \boldsymbol{x}}{\partial q_i} = \frac{1}{x_i}\ \boldsymbol{t}.(\frac{\partial \boldsymbol{x^c}}{\partial q_i} - x_i\ \frac{\partial \boldsymbol{x}}{\partial m}) = \text{ó}_i^c - c \qquad (12)$$

where $c \equiv \mathbf{t}.\partial\mathbf{x}/\partial m$ is the "tax propensity" (more precisely, the marginal propensity to pay taxes out of lump-sum income). Given that ó^c_i and ó_i differ only by a constant, it does not really matter whether the analysis focuses on the marginal deadweight loss (5) or on the marginal excess burden (11).[12] When we extend the analysis to a many-consumer economy, however, this distinction will matter.

For future reference, we note that, using the symmetry of the Slutsky matrix, the normalised marginal excess burden can also be written as:

[12] See Hoff (1994) for a wider discussion of this issue.

$$ó_i^c \;\equiv\; \frac{\boldsymbol{t}.\dfrac{\partial x_i^{\,c}}{\partial \boldsymbol{q}}}{x_i} \qquad (13)$$

In the optimum-taxation literature, the right-hand side is known as the "index of relative discouragement" of commodity i (Mirrlees, 1976). For small taxes, the index of relative discouragement measures the proportionate reduction in the consumption of commodity i induced by the tax system. For non-marginal taxes, an alternative interpretation is also possible: the index of relative discouragement is a first-order approximation to the proportionate change of compensated demand that would be induced by a small “intensification” of the tax system (i.e. a change d**t** such that d**t**=k.**t** for some k>0).

2.4 <u>Vector tax changes</u>

The preceding analysis extends without difficulty to the evaluation of a "tax reform" d**t** which involves simultaneous changes in many taxes. In this case, the change in social welfare is:

$$dV^* \;=\; (ë \;-\; â)\; \boldsymbol{x}.d\boldsymbol{t} \;+\; ë\; \boldsymbol{t}\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{q}}.d\boldsymbol{t} \qquad (14)$$

where the last term continues to be interpreted as the MDL. Similarly, the extension of (10) is simply:

$$dV^* \;=\; -\; b\;\; \boldsymbol{x}.d\boldsymbol{t} \;+\; ë\;\; \boldsymbol{t}.\frac{\partial \boldsymbol{x}^{\boldsymbol{c}}}{\partial \boldsymbol{q}}.d\boldsymbol{t} \qquad (15)$$

where the last term is the MEB. Also as before, a tax reform d**t** which leaves revenue unchanged (i.e. such that $dT/d\mathbf{t} \equiv (\mathbf{x} + \mathbf{t}.\partial\mathbf{x}/\partial\mathbf{q}).d\mathbf{t} = 0$) is socially desirable if and only if it reduces the excess burden.

2.5 Optimum taxes and the Ramsey rule

Using (4), the first-order condition for optimality of t_i, $\partial V^*/\partial t_i = 0$, may be written as:

$$\text{ó}_i \equiv \frac{\boldsymbol{t}.\dfrac{\partial \boldsymbol{x}}{\partial q_i}}{x_i} = \frac{\text{â} - \text{ë}}{\text{ë}} \qquad (16)$$

Let us say that t_i is "unconstrained" if it can be chosen optimally. Recalling (6), (16) states that the normalized MDL <u>should be the same for all unconstrained taxes</u>. This statement can also be seen as a corollary of our earlier result that switches from high-MDL to low-MDL taxes are welfare-improving. Using (10) and (11) and the symmetry of the Slutsky matrix, the first-order condition (16) may also be restated as:

$$\frac{\boldsymbol{t}.\dfrac{\partial x_i^{\,c}}{\partial \boldsymbol{q}}}{x_i} = \frac{b}{\text{ë}} \qquad (17)$$

That is, <u>the index of relative discouragement should be the same for all unconstrained taxes</u>. This is known as the "Ramsey rule". If the amount of revenue to be raised is small, (17) implies that the proportionate reduction in compensated demand induced by the tax system should be the same for all taxable commodities - the formula proposed by Ramsey (1927).[13]

[13] Ramsey's (1927) seminal analysis focuses on the case where $\mathbf{t}^0=\mathbf{0}$ and infinitesimal taxes are introduced. Three points are worth noting about this special case. First, when infinitesimal taxes are introduced at $\mathbf{t}^0=\mathbf{0}$, both MDL and MEB are zero for <u>any</u> $d\mathbf{t}$ (i.e. all tax changes that yield a unit of revenue lead to the same first-order change in

Various interpretations and special cases of the Ramsey rule have been discussed in the literature. For instance, several authors have discussed conditions under which the Ramsey rule boils down to proportional taxes (variously defined). The interested reader is referred to the papers mentioned in footnote 4.

It is instructive to rewrite the right-hand side of (17) as [â+ë**t.**(∂**x**/∂m)-ë]/ë, where the numerator can be interpreted as the difference between the actual marginal social cost of raising revenue through taxes (i.e. ë) and the cost if it were raised via lump-sum taxes.[14] As discussed below, under reasonable assumptions it can be shown that (b/ë) is negative.

2.6 <u>Optimum lump-sum transfer</u>

So far, we have assumed that the consumer's lump-sum income m is exogenous.[15] This is in the spirit of Ramsey's original formulation, which focuses on the use of commodity taxation as a (second-best) <u>substitute</u> for lump-sum taxation. Indeed, if m is a choice variable, then commodity taxes are redundant. To see this, note that the first-order condition for optimum choice of m is simply (from (9)):

---

utility). Second, as Ramsey demonstrated, examining second-order terms leads to his original "rule": "... the taxes should be such as to reduce in the same proportion the production of each taxed commodity" (Ramsey, 1927, p.52). Third, since income effects are second order in this special case, it does not matter whether the proportionate reduction is measured along <u>compensated</u> and <u>uncompensated</u> demands. In general, however, the distinction does matter. While Ramsey stated his original rule in terms of uncompensated demands, the "Ramsey rule" (equation 17) is best expressed in terms of compensated demands. Indeed, outside of this special case, the principle that commodity taxes should lead to an equiproportionate reduction in <u>uncompensated</u> demands makes no intuitive sense.

[14] On this interpretation, see Auerbach (1985, p. 88).

[15] Some readers may find it easier to think in terms of a lump-sum tax L ≡ (-m). The two approaches are, of course, equivalent.

$$b = 0 \qquad (18)$$

With b = 0, $\partial V^*/\partial t_i$ evaluated at $\mathbf{t}_0 = \mathbf{0}$ is zero for each i (see (8)). That is, there is no welfare gain from a small move away from zero commodity taxation. Alternatively, we may note that, if m is a choice variable, then ë = â and $\mathbf{t} = \mathbf{0}$ solve (16) and (18).

Now consider the case analysed by Ramsey, where m is not a choice variable and all taxes are either unconstrained or zero.[16] Multiplying both sides of (17) by $t_i x_i$ and summing over i, we get:

$$b = (\text{ë}/T)(\mathbf{t}.S.\mathbf{t}) \qquad (19)$$

where $T \equiv \mathbf{t}.\mathbf{x}$ is total tax revenue and S is the Slutsky matrix. Remembering that the Slutsky matrix is negative semi-definite, the second-term on the r.h.s. is negative.[17] With ë > 0 by assumption, (19) implies that b and T have opposite signs. In most sensible applications of this problem, T is positive, implying b < 0. Thus, <u>a switch towards lump-sum taxation is desirable</u>.[18]

If there are constrained non-zero taxes, b need not be negative. For instance, if the consumer is already burdened with large taxes on some commodities, a small increase in lump-sum income (dm>0) may be socially desirable (b>0). Note also that in the event where some taxes are constrained and non-zero, an optimum lump-sum tax (b=0) does not entail equalizing â and ë (the

---

[16] If all taxes are unconstrained, the assumption m=0 is needed to make the problem non-trivial (see next section).

[17] We rule out the trivial case where b=0.

[18] See Atkinson and Stern (1974) for further discussion.

marginal social values of private income and government revenue, respectively).

2.7 Price normalization and the untaxed good

One issue that has caused much confusion in the literature is whether it is permissible to assume that a particular commodity is untaxed (and, if so, whether the choice of untaxed commodity matters). In other words, when does this assumption amount to nothing more than an innocuous price normalization rule?

To begin with, consider the standard textbook case where $m \equiv 0$ and all taxes are "unconstrained". Remember that producer prices are fixed; by choice of units, we can set them equal to one, without loss of generality. Further, this model is homogenous (of degree 0) in consumer prices: if $\mathbf{q}$ and $\mathbf{q'}$ are collinear, nothing changes. It follows that the consumer price of one commodity can be given an arbitrary value, without loss of generality. In particular, we can set $q_i \equiv p_i$ for some pre-specified i, that is, assume that one commodity is untaxed. To put it another way, let $\mathbf{t^*}$ be a solution of the Ramsey tax problem. Then it is easy to verify that $\mathbf{t'} \equiv k.\mathbf{t^*} + (k-1).\mathbf{p}$ is also a solution, for any positive k (indeed, $\mathbf{t^*}$ and $\mathbf{t'}$ correspond to the same consumer prices). For any given commodity i, we can opt for this commodity being untaxed by setting $k \equiv p_i/(p_i+t^*_i)$.

Next, consider the extension where m is not identical to zero. As before, and in keeping with the spirit of Ramsey's model, we assume that m is not a choice variable; instead, it takes an exogenous value.[19] Note that the model is no longer homogeneous in $\mathbf{q}$ (though it is homogeneous in $(m,\mathbf{q})$). Hence, an a priori restriction on $\mathbf{q}$, e.g. $q_1 \equiv 1$ or $q_i \equiv p_i$ for some i, would (in general) be

[19] Alternative assumptions are possible, e.g. $m \equiv M/I(\mathbf{q})$ where M is an exogenous parameter and $I(\mathbf{q})$ is an index of consumer prices. The choice of assumption is not (in general) "neutral", and has to reflect the actual basis on which lump-sum transfers are determined.

a real restriction, not just a normalization rule. However, there is a complication: some restriction is needed to make the problem interesting. Indeed, if all taxes are "unconstrained", the problem becomes trivial. This is because, if m is non-zero, proportional taxes ($\mathbf{t} = k.\mathbf{p}$ for some $k>0$) are a form of implicit lump-sum taxation and present an obvious "solution", though the problem they are meant to solve has effectively vanished.[20] This feature of the model, however, is not a matter of concern: in the real world it is not difficult to identify non-taxable commodities (e.g. home-grown food), and this is enough to make the problem non-trivial.[21] The crucial point is that, in this case, the choice of non-taxable commodity is not just a normalization rule: it is a matter of empirical enquiry and affects the structure of optimum taxes.

2.8 The "inverse elasticity rule"

A popular rule of thumb in the field of commodity taxation is that the tax system should weigh heavily on commodities with relatively inelastic demands. The basic idea is that (1) in Ramsey's model, commodity taxes are a second-best substitute for lump-sum taxation, and (2) a tax on a commodity with inelastic demand works much like a lump-sum tax, and is therefore a natural substitute for it. The intuition can also be developed as follows: the Ramsey rule requires equal "discouragement" of each taxable commodity, and commodities with inelastic demands require higher taxes to achieve the same discouragement. Cross-price effects, of course, potentially invalidate this reasoning.

---

[20] See Dixit (1970), and the rejoinder by Sandmo (1974).

[21] Leisure is often chosen as the untaxed good. In some models this is an innocuous normalization rule, in others it is a substantive restriction.

Formally, suppose that t_i is unconstrained and that $\partial x_j/\partial q_i$ is zero for all j such that $j \neq i$ and $t_j \neq 0$. Equation (16) then implies:

$$(t_i/q_i) = \mu/\varepsilon_i \qquad (20)$$

where ε_i is the own-price elasticity of demand for good i and $\mu \equiv (\beta - \lambda)/\lambda$ is independent of i. Thus, t_i should be inversely proportional to ε_i.[22] An analogous result holds in terms of <u>compensated</u> demand elasticities. This follows immediately from (17), under similar assumptions.

2.9 <u>The many-person Ramsey rule</u>

The results derived so far extend quite easily, but with some important qualifications, to the case where there are many consumers. As far as the structure of the problem is concerned, we simply have to replace the utility function V(.) in (2) with a suitable social welfare function. Assuming, as in much of the literature, that the latter takes the Bergson-Samuelson form, we rewrite (2) as:

$$V^* \equiv W(V^1(\mathbf{p}+\mathbf{t},m^1),\ldots,V^H(\mathbf{p}+\mathbf{t},m^H)) + \lambda(\mathbf{t}.\mathbf{x} - \Sigma_h m^h) \qquad (21)$$

where m^h is a lump-sum transfer to consumer h (h=1...,H). In order to extend the earlier results, we first introduce the following notation:

[22] Note that there is no guarantee (in general) of μ/ε_i being positive. For instance, if the initial position is such that some commodities bear heavy positive taxes that cannot be removed (for political or other reasons), the planner may wish to compensate for this by <u>subsidising</u> the commodities with unconstrained taxes. The inverse elasticity rule then suggests that subsidies should be concentrated on commodities with a low elasticity of demand.

$$\hat{a}^h \equiv \frac{\partial W}{\partial V^h} \frac{\partial V^h}{\partial m^h} \tag{22}$$

$$\hat{a}_i \equiv \sum_h \hat{a}^h \frac{x^h{}_i}{x_i} \tag{23}$$

$$b^h \equiv \hat{a}^h - \ddot{e} + \ddot{e}\, \boldsymbol{t} \frac{\partial \boldsymbol{x^h}}{\partial m^h} \tag{24}$$

$$b_i \equiv \sum_h b^h \frac{x_i^h}{x_i} \tag{25}$$

The term $\hat{a}^h$ can be interpreted as the “social valuation of the marginal utility of income” of individual h; we shall also use the more standard (but less precise) term “social marginal utility of income”.[23] We shall refer to b^h, a straightforward generalization of b (see (9)), as the “marginal social value of a transfer to person h”. The quantities $\hat{a}_i$ and b_i are simply commodity-specific weighted averages of ($\hat{a}^h$) and (b^h), respectively, where the shares of different persons in total consumption of commodity i are used as weights. The term b_i is known as the <u>net distributional</u>

[23] Another standard term is “welfare weight”. On the interpretation and measurement of welfare weights, see Stern (1977).

characteristic of commodity i. With these notational conventions, most of the earlier equations go through after replacing $\hat{a}x_i$ everywhere with $\hat{a}_i x_i$ and bx_i with $b_i x_i$. In particular, the Ramsey rule becomes:

$$\frac{\mathbf{t}.\frac{\partial x_i^{\,c}}{\partial \mathbf{q}}}{x_i} = \frac{b_i}{\text{ë}} \qquad (26)$$

Equation (26) is known as the “many-person Ramsey rule” (Diamond, 1975). Note that the right-hand side is no longer independent of i. Instead of being the same for each commodity, the index of relative discouragement should now be proportional to the net distributional characteristic.

In the many-person case, we can still define the marginal deadweight loss and the marginal excess burden as in (5) and (11), respectively. Note, however, that the difference between (normalised) MDL and MEB is no longer constant across commodities. Further, it is no longer desirable to equate (normalised) MDL or MEB across commodities, unless one is prepared to make the further assumption that the distribution of income is in some sense optimal. More precisely, under the assumption that $\hat{a}^h$ is identical for each h, the normalised MDL should be equalised across commodities; on the other hand, if b^h is identical for each h, the normalised MEB should be equalised across commodities.[24] The distinction between these two distributional assumptions is discussed in section 4.

2.10 Further observations

[24] These results are easily derived from (5), (11) and (26).

We end this section by noting, for future reference, that when lump-sum transfers are feasible it is possible to identify comparatively simple directions of welfare-improving tax reform (Dixit, 1975). To illustrate, let us go back to equation (15), with b=0.[25] Suppose that d**t** = -á.**t** for some small á > 0. Thus, we are considering a small proportionate reduction of all taxes. Substituting into (15), and bearing in mind that the Slutsky matrix is negative semi-definite, we find that the right-hand side is positive: a small proportionate reduction of all taxes is welfare-improving. Note that in the case being considered, optimum commodity taxes are zero. Thus, in this case a small move in the direction of the optimum from an arbitrary initial point $\mathbf{t}_o$ is always welfare-improving.

The above illustration describes what is know as a "radial reform" of commodity taxes. Another example is the "concertina reform", which consists of reducing the highest taxes only. For further discussion of these and other welfare-improving tax reforms under optimum lump-sum transfers, the reader is referred to Dixit (1975).

3. <u>The Generalised Ramsey Rule</u>

3.1 <u>The model</u>

We now consider a generalised version of the problem discussed in the preceding section.[26] To start with, we rewrite the vector of net demands of consumer h as $\mathbf{x}^h$ ($\mathbf{q}$, $\mathbf{x}^h$, m^h), where $\mathbf{x}^h$ is an n-dimensional vector of quantity constraints and $\mathbf{x}^h$ (.) solves:

[25] In the many-person case, the first term on the right-hand side of (15) has to be appropriately re-written, but it is still zero if there are optimum lump-sum transfers (b^h = 0 for each h).

[26] For a more detailed exposition, see Drèze and Stern (1987).

$$\underset{x^h}{\text{Max. }} U^h(\mathbf{x}^h) \quad \text{s.t.} \qquad \mathbf{q}.\mathbf{x}^h = m^h$$

$$x^h_i \qquad x^h_i \ \text{ for each i}$$

Note that, subject to regularity conditions, the constrained demand functions $\mathbf{x}^h(\,.\,)$ have the standard properties with respect to $\mathbf{q}$ and m^h. For instance, Roy's identity and the Slutsky equation continue to apply. Similarly, the Slutsky matrix **S** is symmetric and negative semi-definite and $\mathbf{q}.\mathbf{S} = \mathbf{0}$. The main difference is that the matrix **S** has columns of zero entries for commodities such that the quantity constraint is binding (because a small change in the price of such a commodity works like a change in lump-sum income).

The fixed income m^h of individual h consists of the sum of his or her share in private profits and a lump-sum transfer (r^h) from the government:

$$m^h \equiv r^h + \sum_g \text{è}^{hg} \, . \, \text{Ð}^g \qquad (27)$$

where $\text{Ð}^g \equiv \mathbf{p}.\mathbf{y}^g$ is the profit of firm g (see below) and è^{hg} is the share of individual h in firm g's profits.

Next, we write the net supply vector of firm g (g = 1,...,G) as $\mathbf{y}^g = \mathbf{y}^g(\mathbf{p}, \bar{\mathbf{y}}^g)$, where $\bar{\mathbf{y}}^g$ is a vector of quantity constraints and $\mathbf{y}^g(.)$ solves:

$$\underset{\mathbf{y}^g}{\text{Max. }} \mathbf{p}.\mathbf{y}^g \qquad \text{s.t.} \qquad \mathbf{y}^g \quad Y^g$$

$$y^g_i \quad \bar{y}^g_i \text{ for each i}$$

with Y^g denoting the production set of firm g which is assumed to be convex.[27] Here again, the

[27] Strictly speaking we only need convexity in the space of commodities for which the quantity constraints are not binding.

constrained supply functions $\mathbf{y}^g(.)$ have the standard properties with respect to **p**, the vector of producer prices. Note also that the quantity constraints enable us to treat the net supply vectors of firms operating under constant returns to scale as functions rather than correspondences.

Let $(\mathbf{p}, \mathbf{t}, \{\mathbf{x}^h\}, \{\bar{\mathbf{y}}^g\}, \{r^h\}, \{\grave{e}^{hg}\})$ be the vector of "signals" to which individuals and firms respond (bearing in mind (27) and the identity $\mathbf{q} \equiv \mathbf{p} + \mathbf{t}$). The signals are of two types: exogenous signals or "parameters", and control variables. The idea is that the "planner", who conducts the optimization exercise below, chooses the variables within his or her control, taking other variables as given.[28] Denoting the vectors of parameters and control variables as **ù** and **s** respectively, the planner chooses **s** to solve the following maximization problem:

$$\max_{\mathbf{s}} W(\ldots,V^h(\mathbf{s};\grave{\mathbf{u}}),\ldots) \qquad \text{s.t.} \qquad \sum_h \mathbf{x}^h(\mathbf{s};\grave{\mathbf{u}})-\sum_g \mathbf{y}^g(\mathbf{s};\grave{\mathbf{u}})-\mathbf{z}=0 \qquad \text{(P)}$$

where **z** is the vector of net supplies from the public sector. As before (see (21)), V^h is h's indirect utility function and W is a Bergson-Samuelson social welfare function.

The model is restrictive in some respects (e.g. the dichotomous partitioning of signals between exogenous variables and control variables), but it does encompass a wide range of models in the second-best tradition. The classical Ramsey tax problem, for instance, is a special case of this model. So is Diamond-Mirrlees's classic model of "optimum taxation and public production" (Diamond and Mirrlees, 1971). Similarly, the general equilibrium model underlying the project evaluation manual by Little and Mirrlees (1974) can be expressed in the present format. Note also

---

[28] A more sophisticated model would take into account the possibility of endogenous policy responses by government agencies not under the control of the planner.

that the basic model can be readily extended in various directions, e.g. to take into account Boiteux-type budget constraints in the public sector.

3.2 <u>Policy reform</u>

Let V^*(**ù**,**z**) denote the maximum value function of (P). From the envelope theorem, we know that the gradient of V* is the same as the gradient of the Lagrangean. The latter may be written as:

$$= W(\ldots,V^h(\mathbf{s};\mathbf{\grave{u}}),\ldots) - \mathbf{\acute{\imath}}\ [\mathbf{x}(\mathbf{s};\mathbf{\grave{u}}) - \mathbf{y}(\mathbf{s};\mathbf{\grave{u}}) - \mathbf{z}]$$

where $\mathbf{x} \equiv \text{Ó}_h \mathbf{x}^h$ and $\mathbf{y} \equiv \text{Ó}_g \mathbf{y}^g$ denote aggregate (net) consumer demands and aggregate (net) producer supplies, respectively, and **í** is a vector of Lagrangean multipliers or <u>shadow prices</u>. If $\grave{u}_k$ is a particular component of the vector **ù** of parameters (e.g. a tax, a lump-sum transfer, or a quantity constraint), the marginal social value of $\grave{u}_k$, say MSV_k, is:

$$MSV_k \equiv \frac{\partial V^*}{\partial \grave{u}_k} \equiv \frac{\partial}{\partial \grave{u}_k} = \sum\nolimits_h \frac{\partial W}{\partial V^h}\frac{\partial V^h}{\partial \grave{u}_k} - \mathbf{\acute{\imath}}.\frac{\partial(\boldsymbol{x}-\boldsymbol{y})}{\partial \grave{u}_k} \tag{28}$$

The first term on the r.h.s. of (28) is the direct effect on social welfare, and the second term is the social value of the additional excess demands generated by the proposed reform. Applying the same reasoning to a small change d**z** in the public production plan, we find that the “project” d**z** is welfare improving if **í**.d**z**>0, i.e. if it makes a profit at shadow prices. Of course, the shadow prices depend on the specification of "choice variables".[29]

[29] For an early discussion of this point, see Sen (1972).

As shown in Drèze and Stern (1987), using Walras' law, the Lagrangean may be rewritten as:

$$(\boldsymbol{s};\textbf{ù}) \equiv W(\ldots,V^{h}(\boldsymbol{s};\textbf{ù}),\ldots) + \text{ë}\ R \qquad (29)$$

where R is the "shadow revenue" of the government. The latter is defined as:

$$R \equiv \textbf{í}^{*}.\boldsymbol{z} + \hat{\textbf{o}}.\boldsymbol{x} + \hat{\textbf{o}}^{\boldsymbol{p}}.\boldsymbol{y} + \textstyle\sum_{g} \text{è}^{0g}.\text{Đ}^{g} - \textstyle\sum_{h} r^{h} \qquad (30)$$

where $\textbf{í}^{*}\equiv\textbf{í}/\text{ë}$ is a vector of normalized shadow prices, $\hat{\textbf{o}}\equiv(\textbf{q}-\textbf{í}^{*})$ is a vector of <u>shadow consumer taxes</u> (or "shadow taxes" for short), $\hat{\textbf{o}}^{p}\equiv(\textbf{í}^{*}-\textbf{p})$ is a vector of <u>shadow producer taxes</u>, and $\text{è}^{0g}\equiv(1-\sum_{h} \text{è}^{gh})$ is the government's share in the profits of firm g. Note that ë is basically a normalization parameter: a different cardinalization of the social welfare function W(.) leads to a different ë, but leaves $\textbf{í}^{*}\equiv\textbf{í}/\text{ë}$ unchanged. Equation (29) is extremely useful in so far as it converts this complex general-equilibrium model into the standard format of a trade-off between consumer welfare and (shadow) government revenue, as in section 2.

To illustrate, let us consider a specific "parameter", say r^h (the lump-sum transfer going to individual h). From (28), we may write the "marginal social value" of r^h (say b^h) as:

$$b^{h} \equiv \frac{\partial V^{*}}{\partial r^{h}} = \text{â}^{h} - \textbf{í}.(\frac{\partial \boldsymbol{x}^{\boldsymbol{h}}}{\partial m^{h}}) \qquad (31)$$

where, as before, â^h denotes the social marginal utility of h's income (see section 2.9). Thus, the marginal social value of a lump-sum transfer to consumer h is simply the difference between his or her social marginal utility and the shadow value of the additional consumer demands generated by a transfer. Equivalently, we can derive b^h from (29) and (30) as:

$$b^h \equiv \frac{\partial V^*}{\partial r^h} = \hat{a}^h - \ddot{e}\ (1 - \mathbf{\hat{o}}.\frac{\partial \boldsymbol{x}^{\boldsymbol{h}}}{\partial m^h}) \qquad (32)$$

The latter expression also follows from (31), the definition of **í** and the adding-up property. Note that (32) is exactly the same as (24), with shadow taxes replacing actual taxes.

3.3 Commodity taxation

Turning to the specific issues of tax reform and optimum taxation in this general model, the key point to note is that the gradient of (29) with respect to **t** is much the same as in the classical Ramsey tax problem (see (3)). Indeed, substituting **q**≡**p+t** and **ô**≡**(q-í*)** in (29) and using the standard properties of $\mathbf{x}^h(.)$ and $V^h(.)$, it is easy to derive:

$$\frac{\partial}{\partial t_i} \equiv -\hat{a}_i\ x_i + \ddot{e}\ (x_i + \mathbf{\hat{o}}.\frac{\partial \boldsymbol{x}}{\partial t_i}) \qquad (33)$$

Comparing (33) with (3), there are two differences. First, the term $\hat{a}x_i$ is replaced with $\hat{a}_i x_i$, exactly as in the extension from the one-person to the many-person Ramsey rule (see section 2.9) - nothing new here. Second, in the last term on the right-hand side, the vector **ô** of shadow taxes replaces the vector **t** of actual taxes. Hence, the earlier results are readily extended, by simply replacing actual taxes with shadow taxes where appropriate. For instance, setting the right-hand-side of (33) equal to zero and following the usual steps (as with the derivation of (17) and (26)), the first-order conditions for optimum taxation lead to the generalised Ramsey rule:

$$\frac{\mathbf{ô}.\dfrac{\partial x_i^{\,c}}{\partial \boldsymbol{q}}}{x_i} = \frac{b_i}{ë} \qquad (34)$$

where, as before, $b_i \equiv \sum_h (x^h_i/x_i)b^h$ and b^h is the marginal social value of a transfer to individual h (discussed in the previous section).

By analogy with (13) in section 2, we can refer to the l.h.s. of (34) as the "generalised index of relative discouragement". The generalised index of discouragement is a useful tool to restate the general principles of optimum taxation, bearing in mind the three core objectives stated in the introduction: (1) resource allocation, (2) revenue collection, and (3) interpersonal distribution. The implications of the first objective can be seen by assuming optimum lump-sum transfers (i.e. b^h = 0 for each h). In that case, the generalised Ramsey rule states that the generalised index of discouragement should be zero for each commodity. If all taxes are unconstrained, this is achieved by bringing consumer prices in line with shadow prices (**ô**=á.**q** for some á≠0, so that the left-hand side of (34) is zero for each i).[30] To illustrate the implications, in a labour-surplus economy this would typically require subsidising labour-intensive commodities, since the latter would tend to have a low shadow price (Drèze and Stern, 1987). Similarly, in a foreign-exchange-scarce economy, it may require taxing imported goods.

Note that this derivation is based on differentiating the Lagrangean with respect to **t** as if producer prices (**p**) were constant. This, however, does not require the assumption that prices are

[30] The situation where consumer prices and shadow prices are collinear is known as "C-C efficiency"; for further discussion, see Guesnerie (1979).

actually constant. If producer prices are among the "control variables", the derivation remains valid, by the envelope theorem. Further, it must be remembered that the "control variables" formally include the market-clearing variables, i.e. the variables that implicitly adjust to clear the scarcity constraints.[31] Seen in that light, the device of holding producer prices constant is much more general than it appears at first sight. More precisely, the derivation is valid if any of the following holds: (1) producer prices are fixed (as in Diamond and Mirrlees, 1975); (2) producer prices adjust endogenously to clear the scarcity constraints; (3) producer prices are directly controlled by the planner.[32] If none of these assumptions apply, the mechanism through which producer prices are determined has to be specified in the form of additional constraints in the optimisation problem (e.g. the requirement that marginal revenue equals marginal cost, in the case of a commodity supplied by a monopoly). If the vector **t** enters in these additional constraints, the first-order conditions for optimum taxation have to be correspondingly modified.

The requirements of revenue collection can be seen by continuing to assume that the interpersonal distribution of income is optimal ($b^h = b$ for each h, for some b), and also assuming that a switch from indirect taxation to lump-sum taxation would be welfare-improving ($b > 0$). In that case, the generalised Ramsey rule states that consumer prices should move away from shadow prices in such a way that the generalised index of relative discouragement remains the same for each commodity. In the special case where one commodity can be assumed to have zero shadow tax (through suitable normalization) and compensated cross-price demand elasticities among other commodities are zero, we obtain the generalised inverse elasticity rule:

---

[31] For further discussion on this point and related issues, see Drèze and Stern (1987, 1990).

[32] Combinations of these different assumptions for different commodities are also possible (e.g. fixed prices for traded commodities and market-clearing prices for domestic commodities).

$$\hat{o}_i = \frac{b}{e_i}$$

where $e_i \equiv (\partial x^c_i/\partial q_i)(q_i/x_i)$ is the compensated price elasticity of demand for good i.

Finally, to see the implications of concern for interpersonal equity, we relax the assumption that b^h is the same for each h. In that case, there has to be a further adjustment of consumer prices to bring the vector of generalised indices of discouragement in line with the vector of net distributional characteristics. Typically, this would mean shifting taxes away from commodities heavily consumed by transfer-deserving individuals.

All these principles make good intuitive sense. The results illustrate the power of shadow prices in acting as "sufficient statistics" for a general equilibrium model: different applications of the general model outlined in section 3.1 lead to different shadow prices but, subject to the calculation of shadow prices, the basic rules of tax reform and optimum taxation remain the same.[33] Shadow prices also provide a useful bridge between the theory of taxation, the theory of cost-benefit analysis, and the theory of the second-best.

In the special case where shadow prices are proportional (without loss of generality, equal) to producer prices, shadow taxes are the same as actual taxes and the results derived in section 2 apply as stated there. Thus, the classical Ramsey tax problem, which appears to have a partial-equilibrium format, does have a general-equilibrium foundation, as long as shadow prices and

[33] Shadow prices, of course, are not (generally) independent of the tax system. Strictly speaking, all the relevant control variables and shadow prices are simultaneously determined. Nevertheless, the shadow-price approach has much to commend in terms of conceptual clarity. Further, in the context of "tax reform", shadow prices can be taken as given even though they may depend on the <u>initial</u> vector of taxes.

producer prices coincide.[34]

The preceding discussion focuses on optimum taxation, but similar extensions of the standard results apply in the context of tax reform. For example, the conditions for radial and concertina reforms to be welfare improving still go through but with tax reductions suitably defined in the space of shadow taxes. The details are left to the reader.

4. Optimum Distribution: Interpretation and implications

Policy reforms are often evaluated from the point of view of "efficiency" by looking at their implications as if distribution did not matter.[35] But what is meant by the latter assumption? At least two interpretations are possible:

(A1) Equal marginal social utilities: $\hat{a}^h = \hat{a}$ for all h (for some $\hat{a}$).

(A2) Equal marginal social values of transfer: $b^h = b$ for all h (for some b).

Assumption (A1) states that the social valuation of the marginal utility of income is the same for all individuals. (A2) states that a small income transfer between any two individuals would leave social welfare unchanged. The two assumptions have often been used interchangeably in the literature. However, they are not identical, bearing in mind that actual transfers have allocative implications

[34] The conditions under which shadow prices are proportional to producer prices are discussed in Drèze and Stern (1987, 1990). In the opening paragraph of his paper, Ramsey (1927) shows some awareness of this feature of his model: "...I shall suppose that, in Professor Pigou's terminology, private and social net products are always equal or have been made so by State interference not included in the taxation we are considering" (p.47).

[35] For one example (among many), see Dixit (1975).

(captured in b^h, but not in $\hat{a}^h$ - see sections 2.9 and 3.2). To illustrate, it is sometimes argued that otherwise welfare-improving transfers from the rich to the poor may not be desirable in a savings-constrained economy, if the rich have a higher marginal propensity to save than the poor (see Drèze and Stern, 1987, p.966). In other words, equalizing marginal social values of transfer need not entail equalizing marginal social utilities (and vice-versa).

It is also useful to introduce a stronger version of (A2):

(A2') <u>Optimum transfers</u>: $b^h = 0$ for all h.

Note that (A2') involves the further assumption of optimum distribution between consumers and the government: a small transfer from the government to any individual leaves social welfare unchanged.

When are (A1) and (A2) equivalent? Recalling the definition of b^h (see (31) and(32)), we have:

$$\hat{a}^h - b^h \equiv \boldsymbol{\acute{\imath}} \frac{\partial \boldsymbol{x^h}}{\partial m^h} \equiv \ddot{e}\left(1 - \boldsymbol{\hat{o}} \frac{\partial \boldsymbol{x^h}}{\partial m^h}\right) \qquad (35)$$

Thus, $\hat{a}^h$ and b^h essentially differ by a constant minus the "shadow tax-propensity" (the last term in (35)), so that (A1) and (A2) are equivalent if and only if the latter is the same for all individuals.[36] It is easy to see that a sufficient condition for this to hold is the following:

[36] The shadow tax-propensity is essentially a weighted average of shadow tax rates, with marginal propensities to consume as weights. With suitable normalization, the shadow tax-propensity may also be interpreted as the covariance between the vector of shadow tax rates and the vector of marginal propensities to consume.

(B1) <u>Proportional shadow taxes</u>: **í** = ã.**q** for some ã > 0 (i.e. $ô_i/q_i$ is the same for all i).

Condition (B1) is also <u>necessary</u> for (A1) and (A2) to be equivalent <u>if</u> the (nxH) matrix **R** with elements $R_{ih} \equiv (\partial x^h_i / \partial m^h)$ has rank n. To see this, first write (35) in vector form as follows:

$$\mathbf{â} - \boldsymbol{b} \equiv \mathbf{í}\ \boldsymbol{R} \qquad (36)$$

where **â** ≡ {$â^h$} and **b** ≡ {b^h}. Note also that **q**.**R** = **1**, where **1** indicates a vector of ones. If both (A1) and (A2) hold, then there is a constant c such that **b** = **â** - c.**1**. Thus

$$\mathbf{í}.\mathbf{R} \equiv \mathbf{â} - \mathbf{b} = c.\mathbf{1} = c.\mathbf{q}.\mathbf{R} \qquad (37)$$

If **R** has rank n, (38) implies that **í** and **q** are collinear, i.e. shadow taxes are proportional.

For the rank condition to be satisfied, there have to be at least as many consumers as there are commodities, and the H vectors of marginal propensities to consume should span a space of dimension no less than n. In other words, there has to be adequate diversity of (marginal) spending patterns. When the rank condition holds, any divergence between **í** and **q** makes it impossible for <u>both</u> (A1) and (A2) to be satisfied.

The following is a corollary of the reasoning so far:

<u>Proposition 1</u>: If **R** has rank n, then (B1) is necessary and sufficient for (A1) and (A2) to be equivalent; further, any two of (A1), (A2) and (B1) imply the third.

When the rank condition is not satisfied, assumptions other than (B1) can be made to ensure that shadow tax-propensities are identical (so that (A1) and (A2) are equivalent). One such assumption is that everyone has the same marginal propensity to consume, for each commodity (the extreme opposite of the assumption of "adequate diversity"). Another one is that, for each individual, marginal propensities to consume are uncorrelated with shadow taxes. Obviously, all these assumptions are restrictive.

To see why the divergence of shadow tax propensities is an issue, consider any feasible "reform" such that the gainers gain more than the losers lose (where any person's net gain or loss is measured in terms of income equivalent).[37] If (A1) holds, then the reform is welfare-improving. Sometimes it is argued that even if (A1) does not hold, the reform can be justified on efficiency grounds, in the sense that "the gainers could compensate the losers". However, unless shadow tax-propensities are identical, transfers from the gainers to the losers have efficiency implications that need further examination, so that the status of the compensation argument is unclear. At best, it is a "notional compensation" argument of dubious ethical relevance.

To see the problem from another angle, we have noted that under assumption (A2'), simple directions of welfare-improving tax reform can be identified. Further results along those lines (with straightforward generalisations in the space of shadow taxes) are presented in Dixit (1975). As the author puts it, the results hold "assuming that there are no distributional problems". However, in general (A2') conflicts with (A1), in which case there is "a distributional problem" even if (A2') holds.

[37] There is no need to distinguish here between "compensating" and "equivalent" variations, since we are looking at marginal changes.

We now turn to the respective implications of (A1), (A2) and (A2') for the structure of optimum taxes.[38] Under (A2), the one-person Ramsey rule applies: for optimum taxes, the generalised index of relative discouragement has to be the same for all commodities. This follows from (34), with $b_i = b$ for all i. A corollary is that, under the stronger assumption of optimum transfers (A2'), optimum taxes are such that shadow taxes are proportional (i.e. C-C efficiency); this has already been discussed in the preceding section. Finally, returning to (A1), the correct optimum tax rule in this case is that the normalised MDL (evaluated using shadow taxes) should be the same for all commodities, i.e.:

$$\frac{\hat{\mathbf{o}}.\dfrac{\partial \boldsymbol{x}}{\partial q_i}}{x_i} = c$$

for some non-zero constant c (this follows from setting the right-hand side of (33) equal to zero).

In short, in a second-best world there is no simple way of avoiding distributional issues, even by assumption (i.e. there is no unique way of stating the assumption that "distributional issues have been resolved"). Distributional assumptions have to be made explicit, bearing in mind that they lead to different tax rules.

5. Concluding Remarks

In a wide class of models, the problem of optimum taxation is usefully split into two parts: the calculation of "shadow prices" and the application of the generalised Ramsey rule. The relevant

[38] The argument focuses on first-order conditions; as noted earlier, second-order conditions have rarely been examined in the literature.

shadow prices depend on the structure of the economy, the range of choice variables and the social welfare function, but the generalised Ramsey rule itself is quite robust. This way of approaching the problem is more transparent and productive than the compilation of a long catalogue of tax rules for different types of distorted economies, as has been done in much of the literature on the second-best. In models where the generalised Ramsey rule applies, the main issue is to identify the relevant shadow prices. In other models, the basic task is to identify the reasons why the rule does not apply (e.g. interdependence between consumer demands and producer prices, independently of consumer prices), and to derive suitable extensions of the general rule.

Similar principles apply in the field of tax reform. Here again, the problem is usefully split between calculation of shadow prices and application of general principles for the identification of welfare-improving reforms in the space of "shadow taxes". One of these principles is that, under specific distributional assumptions, "radial" reforms in the space of shadow taxes (i.e. small tax changes that bring consumer prices closer to shadow prices) are welfare-improving. Care must be taken, however, in identifying and assessing the relevant distributional assumptions. In particular, we have noted a possible tension between the assumptions of "equal marginal social utility" and "equal marginal social values of transfer", which are often used interchangeably in the literature.

References

Ahmad, E., Coady, D., Stern, N., 1988. A Complete Set of Shadow Prices for Pakistan, Pakistan Development Review, 27, 7-43.

Ahmad, E., Stern, N., 1991. The Theory and Practice of Tax Reform in Developing Countries, Cambridge University Press, Cambridge.

Atkinson, A., Stern, N., 1974. Pigou, Taxation and Public Goods, Review of Economic Studies, 41, 119-28.

Atkinson, A., Stiglitz, J., 1976. The Design of Tax Structure: Direct Versus Indirect Taxation, Journal of Public Economics, 6, 55-75.

Atkinson, A., Stiglitz, J., 1980. Lectures on Public Economics, McGraw-Hill: New York.

Auerbach, A., 1985. The Theory of Excess Burden and Optimal Taxation, in Auerbach, A., Feldstein, M. (Eds), Handbook of Public Economics, vol. 1, North-Holland:Amsterdam.

Coady, D., 1997. Agricultural Pricing Policies in Developing Countries: An Application to Pakistan, International Tax and Public Finance, 4, 39-57.

Deaton, A., Stern, N., 1986. Optimally Uniform Commodity Taxes, Taste Differences, and Lump-sum Grants, Economics Letters, 20, 263-266.

Diamond, P., 1975. A Many-Person Ramsey Rule, Journal of Public Economics, 4, 335-342.

Diamond, P., Mirrlees, J., 1971. Optimal Taxation and Public Production I: Production Efficiency, and II: Tax Rules, American Economic Review, 61, 8-27 and 261-278.

Diamond, P., Mirrlees, J., 1975. Private Constant Returns and Public Shadow Prices, Review of Economic Studies, 43, 41-47.

Dixit, A., 1970. On the Optimum Structure of Commodity Taxes, American Economic Review, 60(2), 295-301.

Dixit, A., 1975. Welfare Effects of Tax and Price Changes, Journal of Public Economics, 4, 103-23.

Dixit, A., 1987. Tax Policy in Open Economies. In Auerbach, A., Feldstein, M. (Eds), Handbook of Public Economics, vol. 2, North-Holland: Amsterdam.

Drèze, J., 1998. Distribution Matters in Cost-Benefit Analysis, Journal of Public Economics, 70(3), 485-88.

Drèze, J., Stern, N., 1987. The Theory of Cost-Benefit Analysis. In Auerbach, A., Feldstein, M. (Eds), Handbook of Public Economics, vol. 2, North-Holland: Amsterdam.

Drèze, J., Stern, N., 1990. Policy Reform, Shadow Prices, and Market Prices, Journal of Public Economics, 42, 1-45.

Guesnerie, R., 1979. General Statements of Second-Best Pareto Optimality, Journal of Mathematical Economics, 6, 169-94.

Hoff, K., 1994. The Second Theorem of the Second Best, Journal of Public Economics, 54(2), 223-42.

Kay, J., 1980. The Deadweight Loss From a Tax System, Journal of Public Economics, 13, 111-20.

Little, I., Mirrlees, J., 1974. Project Appraisal and Planning in Developing Countries, Heinemann: London.

Mirrlees, J., 1976. Optimal Tax Theory: A Synthesis, Journal of Public Economics, 6, 327-58.

Myles, G., 1995. Public Economics, Cambridge University Press: Cambridge.

Ramsey, F., 1927. A Contribution to the Theory of Taxation, Economic Journal, 47-61.

Sandmo, A., 1974. A Note on the Structure of Optimal Taxation, American Economic Review, 62(2), 701-6.

Sandmo, A., 1976. Optimal Taxation: An Introduction to the Literature, Journal of Public Economics, 6, 37-54.

Sen, A., 1972. Control Areas and Accounting Prices: An Approach to Economic Evaluation, Economic Journal, 82, 486-501.

Stern, N., 1977. Welfare Weights and the Elasticity of the Marginal Valuation of Income. In Artis, M., Nobay, R., (Eds), Modern Economic Analysis, Blackwell: Oxford.