APPROXIMATE MEASURES OF SOCIAL WELFARE AND THE SIZE OF TAX REFORM*

Peter J. Hammond, Stanford University

Abstract**: This paper deals with second order approximations to changes of welfare measured by social welfare functions. In the framework of piecemeal policy the impacts of tax reforms to social welfare are considered. Three different kinds of social welfare functions are employed: an arbitrary Bergsonian, a social welfare function based on money metric utility for individuals, and a money metric of social welfare. Furthermore Pareto improving reforms are discussed. If possible, the optimal direction and the optimal size of a tax reform are determined.

1. Introduction

In Hammond (1984), I have noted how welfare economics has recently come very much closer to public finance theory because of its recognition of incentive constraints. In particular, the optimal lump-sum transfers which were often assumed in welfare economics are incentive incompatible. Instead, redistribution of real income to avoid extremes of poverty or to promote equality has to be achieved by distortionary taxes such as commodity or income taxes. Such taxes are not necessarily Pareto inefficient, as is usually alleged, because there may be no incentive compatible procedure for reaching a Pareto superior allocation.

In a "second-best" economy with incentive constraints, it becomes very hard to characterize a welfare optimum, let alone calculate optimal tax rates, etc.. For this reason, theorists as well as practical economists would do well to seek improvements to existing tax structures, rather than the elusive welfare optimum. This calls for a theory of tax reform.

---

* This paper was started during a visit to the Institute of Public Economics at the University of Graz and completed at the Institute for Advanced Studies of the Hebrew University, Jerusalem. Financial support from the Institute of Advanced Studies and the Österreichische Länderbank is gratefully acknowledged, as are the helpful suggestions of conference participants, especially Dieter Bös, and also of Edward Green.

** This abstract was arranged by the editors.

Studies in Contemporary Economics, Vol. 7
Beiträge zur neueren Steuertheorie
Herausgegeben von D. Bös, M. Rose und Ch. Seidl

As I point out in Section 2 below, a basic framework for identifying directions for favourable or welfare improving tax reform is due essentially to Meade (1955). The article which did most to generate a revival of interest in the theory of tax reform is Guesnerie (1977). This work, however, is concerned solely with the direction of tax reform; in this paper, I shall discuss the appropriate size of tax reform in a given favourable direction. This can be done by considering quadratic or second-order approximations to measures of social welfare rather than just the linear or first-order approximations which identify favourable directions of reform. The use of such quadratic approximations underlies numerical algorithms such as Newton's method or conjugate gradient methods, as explained in Hestenes (1980), for example. The quadratic approximations are derived in a manner analogous to that of Hammond (1983). The main difference is that, as Chipman and Moore (1980) point out, measures of equivalent rather than compensated variation are required in order to be able to compare different reforms. Here, of course, I am concerned with reforms of different sizes in the same direction, so I do need to approximate the equivalent rather than the compensated variation. More precisely, I calculate quadratic approximations to a money-metric social welfare function (cf. Samuelson, 1974). I should mention the work of McKenzie (1983), based on McKenzie and Pearce (1976, 1982), which shows how to calculate higher order approximations - notably third order approximations - which could be used to find more accurate measures for the suitable size of a tax reform. Their work, however, rests on knowledge of second order derivatives of demand functions, whereas mine requires knowledge only of the first order effects of the reform on prices and quantities and of consumers' demand responses to income changes. Of course, both they and I also need the interpersonal comparisons which are implicit in a Bergson Social Welfare function in order to construct approximate measures of social welfare.

Section 2 summarizes the usual first-order theory of tax reform. Section 3 introduces second order approximations and shows how they suggest, in some cases, an optimal size for a tax reform in a given direction. Section 4 considers how to calculate the second order approximations used in Section 3, based on money metric utility approximations for individuals. Section 5 derives a money metric approximate measure of social welfare, using an approach similar to that of Hammond (1983). Section 6 considers Pareto improving reforms, and Section 7 discusses the choice of the direction as well as the size of tax reform, based on second order approximations still.

2. First Order Approximations and the Size of Tax Reform

Up to now, the theory of tax reform as in Guesnerie (1977) etc., has concentrated on first order approximations, and on identifying favourable directions of tax reform. For example, there may be a modified Walrasian equilibrium, or "tight semi-market equilibrium" described by a system of equations of the form:

$$\chi(p,t) = 0 \tag{1}$$

where p is the suitably normalized consumer price vector, t is a vector of tax parameters, and $\chi(p,t)$ is the vector excess demand function, expressing the aggregate excess demands for each commodity as a function of consumer prices and taxes. Consumers $i=1,2,\dots,N$ or $i\varepsilon N$ are assumed to face a budget constraint of the form:

$$p.q_i = m_i(p,t) \tag{2}$$

when the price vector is p, where q_i denotes consumer i's net demand vector, and $m_i(p,t)$ is i's unearned income (from dividends and government transfers or tax allowances or tax credits) as a function of the price vector p and the tax system t. In (2), $p.q_i$ denotes the inner product of the price norm p with the quantity column vector q_i.

I shall assume that the economy starts with a tax regime t^o and equilibrium prices p^o, quantities $q_i^o(i\varepsilon N)$ and incomes $m_i^o(i\varepsilon N)$. I also assume that the conditions of the implicit function theorem are met by the function χ in a neighbourhood of (p^o,t^o) so that there exists a unique normalized price vector p(t) for each t in that neighbourhood which satisfies:

$$\chi(p(t),t) = 0 \quad . \tag{3}$$

In addition, p will then be differentiable as a function of t, at t^o, with a matrix P^o of derivatives, so that we have the following first-order or linear approximation:

$$\Delta p \cong P^o \, \Delta t \tag{4}$$

where Δt denotes a <u>tax reform</u> or change in the tax regime t from t^o to t^1, and Δp is the corresponding approximate change in the equilibrium price vector. Assuming that the income functions $m_i(p,t)$ are differentiable at p^o, t^o, and that consumers have demand functions which are differentiable at (p^o, m_i^o) (for each $i \in N$), it follows too that:

$$\Delta q_i \cong D_i^o \Delta p + b_i^o \Delta m_i \tag{5}$$

or that

$$\Delta q_i \cong Q_i^o \Delta t \tag{6}$$

where D_i^o is the matrix of responses of consumer i's demand to price changes, and b_i^o is the vector of responses to income changes - i.e., the vector of Engel curve slopes - while Q_i^o is also a matrix. The effect of these changes on a suitably normalized utility function U_i for consumer i is then given approximately by:

$$\Delta U_i \cong p^o.\Delta q_i = \Delta m_i - \Delta p.q_i^1 \tag{7}$$

or by

$$\Delta U_i \cong u_i^o.\Delta t \tag{8}$$

where $u_i^o = p^o Q_i^o$ in this semi-Walrasian framework. Finally, the effect on social welfare is given approximately by:

$$\Delta W \cong \Sigma_i \beta_i^o \Delta U_i = (\Sigma_i \beta_i^o u_i^o).\Delta t \tag{9}$$

where $(\beta_i^o)_{i \in N}$ is a vector of welfare weights representing the "marginal social significance" of each consumer's income. These weights reflect ethical values, of course. Then, substituting for (7) in (9) gives:

$$\Delta W \cong \Sigma_i \beta_i^o \Delta m_i - \Delta p.\Sigma_i \beta_i^o q_i^1 \tag{10}$$

so that $\beta_i^o = \partial W/\partial m_i$, in effect, evaluated at the initial equilibrium.

Such formulae were already available in Meade (1955). Guesnerie's (1977) concern was to find directions of tax reform dt which were Pareto improvements in the sense that (corresponding to (8) above), $dU_i = u_i^o.dt$ is positive for all $i \in N$. He collaborated later in Fogelman, Guesnerie and Quinzii (1978) to develop gradient processes of continuous tax reform which were steadily Pareto improving until some Pareto efficient allocation was reached among those allocations which were achievable given the allowable tax instruments described in the vector t. Diewert (1978) sought directions of small tax perturbations dt which would produce the largest possible welfare improvement $dW = (\Sigma_i \beta_i^o u_i^o).dt$ (corresponding to (9) above) for a given size of the reform dt, and showed that the answer depended upon the units in which the components of the vector t were measured. That work was considerably refined by Tirole and Guesnerie (1981).

These linear or first-order approximations which lead to formulae such as (8) and (9) serve only to identify favourable directions dt of tax reform. No indication is given of how large a reform in a favourable direction dt should be undertaken. If the linear approximations were all exact, an infinite reform in the direction dt would be called for, to produce an infinite improvement. Since infinite improvements are clearly not possible, the linear approximations cannot be exact, and all that this first-order analysis tells us is that small enough reforms in any favourable direction will be favourable. To say more than this higher order approximations must be considered, and I turn next to some relatively simple second order approximations which are quite closely related to those derived in Hammond (1983).

3. Second-Order Approximations and the Size of Tax Reform

Suppose that we extend the first-order or linear approximation (9) to a second-order or quadratic approximation:

$$\Delta W \cong w^o.\Delta t + 1/2 \quad \Delta t.H^o \Delta t \tag{11}$$

where $w^o: = \Sigma_i \beta_i^o u_i^o$ as in (9) and H^o is the Hessian matrix of second-order partial derivatives, evaluated at t^o, of the function F(t) which expresses the dependence of social welfare W on the vector of tax parameters t. H^o also reflects ethical values, just as the vector $(\beta_i^o)_{i \in N}$ does. Now consider a reform:

$$\Delta t = \lambda \ dt \tag{12}$$

of size λ in the direction dt. Suppose that the direction dt is a favourable one, according to first-order analysis, which means that:

$$w^o.dt > 0 \ . \tag{13}$$

Substituting (12) into (11) gives:

$$\Delta W \cong \lambda(w^o.dt) + 1/2\ \lambda^2\ (dt.H^o dt) \quad . \tag{14}$$

Assume that $dt.H^o dt < 0$; conditions to ensure this are discussed later at the end of this section. Then the approximation in (14) reaches a maximum with respect to λ at a step size λ^* given by:

$$\lambda^* = -\ (w^o.dt)\ /\ (dt.H^o dt) \tag{15}$$

which is positive, given our assumptions. This simple calculation not only suggests that λ^* is the optimal step size; it also suggests that any step of size $\lambda > \lambda^*$ is too large even though it may still increase welfare; only steps of size $\lambda \leq \lambda^*$ would appear to merit attention. Remember that smaller steps are more likely to be truly favourable, bearing in mind the errors in the second-order approximation.

To ensure that λ^* is positive, and that the approximation (14) reaches a maximum at λ^*, it is necessary that $dt.H^o dt$ be negative for the direction dt. This would automatically be true if the function F(t) were differentiably strictly concave, of course, for then H^o would be negative definite. But there is no guarantee of this. Indeed, Atkinson and Stiglitz (1980) were able to make a case for random taxation precisely because the function F may well not be even quasi-concave, let alone differentiably strictly concave.

Nevertheless, given the fixed direction dt, and given that F is twice continuously differentiable at t^o, it is possible to apply a suitable sufficiently concave transformation $\bar{W} = \phi(W)$ to the social welfare measure W in order to ensure that $dt.\bar{H}^o dt < 0$ for the new Hessian matrix $\bar{H}^o$ of $\bar{W}$ evaluated at t^o. Indeed, choose a strictly increasing ϕ so that $\phi'(W^o) = 1$ and $\phi''(W^o) = -\ \delta < 0$. Then the gradient vector $\bar{w}^o$ of $\bar{W}$ is the same as w^o, the gradient vector of W, and the Hessian matrix $\bar{H}^o$ is given by:

$$\bar{H}^o = H^o - \delta w^o (w^o)' \tag{16}$$

where $(w^o)'$ is the row vector which is the transposition of the column gradient vector w^o. Given the direction dt, it follows that:

$$dt.\bar{W}^o dt = dt.H^o dt - \delta(w^o.dt)^2 \tag{17}$$

which is certainly negative provided that δ is chosen large enough, because of our assumption that $w^o.dt > 0$. This suggests a step size $\lambda^*(\delta)$ which depends on δ:

$$\lambda^*(\delta) = - (w^o.dt)/[dt.H^o dt-\delta(w^o.dt)^2] \quad . \tag{18}$$

Here $\lambda^*(\delta)$ is positive for all large enough δ and decreases as δ increases, tending to zero as δ tends to infinity. If $dt.H^o dt \geq 0$, which is the case which gives rise to the need for this kind of transformation, then $\lambda^*(\delta)$ can take any value between zero and infinity for suitable values of δ, which is not very helpful.

Another approach which is less arbitrary and may be better is to use some "money-metric" measure of welfare (cf. Samuelson, 1974 and McKenzie, 1983) as explained below in section 5. Then W has natural monetary units and should not be subjected to a strictly concave transformation of the form $\bar{W} = \phi(W)$. The step size λ^* is then uniquely determined by (15), provided that $dt.H^o dt < 0$, and will maximize the quadratic approximation to the money-metric measure of welfare for a step in the direction dt. If $dt.H^o dt > 0$, on the other hand, choosing λ^* given by (15) will minimize the quadratic approximation and, since $\lambda^* < 0$, will also produce an unfavourable reform. This suggests making a fairly large step in the direction dt away from the minimum. To say much more about the most appropriate step size would require a third-order analysis along the lines of McKenzie (1983). A cubic approximation may well have a local maximum in the step size λ even when the quadratic approximation (14) does not. The extra work of calculation is likely to be very considerable, however, and to require estimates of second-order derivatives of consumer's demand functions. Moreover, the transformation ϕ considered above can always be chosen so that the third-order term in λ^3 which would be added to the second-order approximation (14) actually vanishes. So a third-order approximation is only helpful when it is applied to a measure of welfare that has some cardinal significance, such as a money-metric measure.

4. Calculating Second-Order Approximations

In order to use the approximation (14), it is necessary to calculate the Hessian matrix H^o of the function $W \equiv F(t)$. While this may seem very complicated at first sight, in fact it is quite straightforward provided one makes just a few simplifying assumptions, as I shall now explain.

Given a general Bergson social welfare function $W \equiv G((U_i)_{i\epsilon N})$ which depends on individual utilities alone, a second order approximation is:

$$\Delta W \cong \Sigma_i \beta_i^o \Delta U_i + 1/2\ \Sigma_i \Sigma_j \Delta U_i . \gamma_{ij}^o \Delta U_j \tag{19}$$

where the weights β_i^o are as in (9), and $\Gamma^o := [\gamma_{ij}^o]$ is the Hessian matrix of the function G evaluated at the initial utility levels $(U_i^o)_{i\epsilon N}$. It is worth repeating that Γ^o and $(\beta_i^o)_{i\epsilon N}$ both depend on the ethical value judgements that determine the Bergson Social Welfare function G.

To convert (19) into the form (14) we used in section 3, it is necessary to express each individual's utility change ΔU_i as an approximate function of the tax reform λ dt. For the second-order approximation we are using, it is clear that it is sufficient to calculate the approximation:

$$\Delta U_i \cong \lambda u_i^o . dt + 1/2\ \lambda^2 dt . H_i^o dt \tag{20}$$

where H_i^o is the Hessian matrix at t^o of the function which expresses consumer i's utility U_i in terms of t, and where u_i^o is as in equation (8) above. In fact substituting (20) into (19) tells us at once that, in the quadratic approximations (11) and (14), H^o is given by:

$$H^o = \Sigma_i \beta_i^o H_i^o + \Sigma_i \Sigma_j\ \gamma_{ij}^o (u_i^o . u_j^o) \quad . \tag{21}$$

It therefore remains only to calculate each individual's Hessian matrix H_i^o, to which the rest of this section is devoted.

Evidently the matrix H_i^o will depend upon the particular utility function which we use to represent i's preferences, as indeed does the magnitude (though not the direction) of the gradient vector $u_i^o = p^o Q_i^o$. An appropriate normalization for this purpose is Allen's (1949) and

Samuelson's (1974) "money-metric" utility function, defined as $E_i(p^o, u_i)$, the expenditure needed at the fixed price vector p^o in order to achieve the utility level u_i. Obviously, in discussing tax reform, the status quo price vector p^o is the most sensible choice of the reference price vector.

For the rest of this section, I shall consider only a typical single individual, and will therefore omit the subscript i throughout. It will be reintroduced later when we consider the set of all individuals once again.

If u^1 denotes the individual's utility with the tax reform, and u^o without, the change in money-metric utility is given by:

$$\Delta U = E(p^o, u^1) - E(p^o, u^o) \tag{22}$$

$$= E(p^o, u^1) - E(p^1, u^1) + m^1 - m^o \tag{23}$$

$$= \int_{p^1}^{p^o} dp.x(p, u^1) + \Delta m \; . \tag{24}$$

(23) follows from (22) because $m^o = E(p^o, u^o)$ and $m^1 = E(p^1, u^1)$. Then (24) follows from the envelope property of the expenditure function E, with $x(p, u^1)$ denoting the vector compensated demand function of the individual, equal to the gradient of E. The integral is a line integral along any path from p^1 to p^o; Slutsky symmetry guarantees that the integral is path indepent. In fact, the integral in (24) is merely minus the equivalent variation of the price change from p^o to p^1 in the sense of Hicks (1942). It is important to use the equivalent variation here rather than the compensating variation I used in Hammond (1983) because, as Chipman and Moore (1980) explain, comparisons of utility with a reform require the equivalent variation to be used.

We need a second-order approximation to (24), for which it suffices to use a first-order approximation to the compensated demand function:

$$x(p, u^1) \cong x(p^1, u^1) + S^1(p - p^1)$$

$$= q^1 + S^1(p - p^1) \tag{25}$$

where q^1 is the vector of quantities with the tax reform, and S^1 is the Slutsky matrix evaluated at (p^1, u^1). The approximation is taken from

p^1 rather than from p^o because $q^1 = x(p^1,u^1)$ depends directly on the eventual outcome of the reform, whereas $x(p^o,u^1)$ does not.

Substituting the approximation (25) into (24) and integrating gives:

$$\Delta U \cong \Delta m - \Delta p.q^1 + 1/2\ \Delta p.S^1\Delta p \quad . \tag{26}$$

But, from the budget equations $m^o = p^o q^o$ and $m^1 = p^1 q^1$, it follows that:

$$\Delta m - \Delta p.q^1 = p^o.\Delta q \tag{27}$$

which leads to:

$$\Delta U \cong p^o.\Delta q + 1/2\ \Delta p.S^1\Delta p \quad . \tag{28}$$

The first term corresponds to equation (7), as is to be expected. The second term depends upon the Slutsky matrix S^1, and this must be calculated next. It is given by:

$$S^1 = D^1 + b^1(q^1)' \tag{29}$$

where D^1 is the matrix of uncompensated demand responses to price changes, and b^1 the vector of responses to income changes, both evaluated at the price-income pair (p^1,m^1) which is reached by the reform. Next, notice that using a first-order approximation to the consumer's uncompensated demand function leads to:

$$\begin{aligned}\Delta q &\cong D^1\Delta p + b^1\Delta m \\ &= S^1\Delta p + b^1(\Delta m-\Delta p.q^1) \\ &= S^1\Delta p + b^1(p^o.\Delta q)\end{aligned} \tag{30}$$

(substituting from (29) and (27) in turn). So, substituting (30) into (28) to eliminate the term in S^1 leads to:

$$\begin{aligned}\Delta U &\cong (1-1/2\ \Delta p.b^1)(p^o.\Delta q) + 1/2\ \Delta p.\Delta q \\ &= (p^o + 1/2\ \Delta p).\Delta q - 1/2(\Delta p.b^1)(p^o.\Delta q) \quad .\end{aligned} \tag{31}$$

All the terms in this approximation, with the exception of b^1, can be calculated from comparative static equations such as (4), (5) and (6). It is true that a genuine second-order approximation requires second-order approximations for Δq instead of the first-order approximations in (5) and (6). Nevertheless, even if such second-order approximations are too complicated to calculate, as is all too likely, we still have a useful if incomplete second-order approximation on the basis of which the appropriate size of a tax reform can be discussed.

The one exception in (31) is the vector b^1 of demand responses to income changes, or of Engel curve slopes. Provided that the consumer's uncompensated demands are twice continuously differentiable in prices and income, we have a linear approximation of the form:

$$b^1 \cong b^o + B_p{}^o \Delta p + b_m{}^o \Delta m \tag{32}$$

for a suitable matrix $[B_p{}^o, b_m{}^o]$ of second-order partial derivatives of the demand function. Then, however, the error in replacing b^1 in (33) by b^o will be of third order in Δt, and so we can use the second-order approximation:

$$\Delta U \cong (p^o + 1/2\ \Delta p).\Delta q - 1/2(\Delta p.b^o)(p^o.\Delta q) \tag{33}$$

which depends on b^o, the vector of demand responses to income changes evaluated at the price income pair (p^o, m^o) without the reform. It is true that b^o cannot be directly calculated from knowledge of the comparative static equations (3) or the consequent linear approximations (4) and (6); nevertheless, there is plenty of empirical evidence on which to base good estimates of b^o, estimates which are likely to be needed anyway to calculate (4) and (6) as suggested by the appearance of b^o in (5).

Substituting from (33) for each separate individual and using (19) gives, to second order:

$$\Delta W \cong (p^o + 1/2\ \Delta p).\Sigma_i \beta_i{}^o \Delta q_i - 1/2\ \Sigma_i \beta_i{}^o (\Delta p.b_i{}^o)(p^o.\Delta q_i) + $$
$$+ 1/2\ \Sigma_i \Sigma_j (p^o.\Delta q_i)\ \gamma_{ij}{}^o\ (p^o.\Delta q_j)\quad . \tag{34}$$

Substituting from (4) and (6) into this then gives the following approximation, which determines the Hessian matrix in (11):

$$\Delta W \cong (p^{o} + 1/2\ P^{o}\Delta t).(\Sigma_i \beta_i{}^{o} Q_i{}^{o})\Delta t -$$

$$- 1/2\ \Sigma_i \beta_i{}^{o}\ [(P^{o}\Delta t.b_i{}^{o})]\ p^{o}.(Q_i{}^{o}\Delta t)\ +$$

$$+ 1/2\ \Sigma_i \Sigma_j [p^{o}.(Q_i{}^{o}\Delta t)]\ \ \gamma_{ij}{}^{o}[p^{o}.(Q_j{}^{o}\Delta t)]\ \ . \qquad (35)$$

For a favourable reform $\Delta t = \lambda$ dt in the given direction dt, an explicit formula for the optimal step size λ^*, as given by (15), can also be calculated.

Define:

$$dp : = p^{o}dt,\ dq_i\ : = Q_i{}^{o}dt\ \text{(all i)} \qquad (36)$$

for the price and quantity changes induced by a reform in the direction dt. Then (35) becomes:

$$\Delta W \cong \lambda\ p^{o}.\Sigma_i \beta_i{}^{o} dq_i + 1/2\ \lambda^2 H \qquad (37)$$

where

$$H: = dp.\Sigma_i \beta_i{}^{o} dq_i - \Sigma_i \beta_i{}^{o}(dp.b_i{}^{o})(p^{o}.dq_i)\ +$$

$$+ \Sigma_i \Sigma_j (p^{o}.dq_i)\ \gamma_{ij}{}^{o}(p^{o}.dq_j)\ \ . \qquad (38)$$

This suggests choosing:

$$\lambda^* = - p^{o}.\Sigma_i \beta_i{}^{o} dq_i / H \qquad (39)$$

provided, of course, that H is negative, which makes λ^* positive because $p^{o}.\Sigma_i \beta_i{}^{o}dq$ is positive for a favourable direction of reform dt. The problems that arise when H is nonnegative were discussed in Section 3.

5. A Money-Metric Measure of Social Welfare

At the end of Section 3 I argued that a money-metric measure of social welfare would be desirable. The measures derived from (19) in Section 4 do not have this property, except for a very special form of the Hessian matrix Γ^O of G. In this section I shall show how to calculate a money-metric form, arguing along the lines of section 4 of Hammond (1983). Thus, I shall treat the whole society as a single consumer of a profile $\underset{\sim}{x}$ of consumption vectors, one for each individual. Following Sen (1976, 1979), different individuals' consumptions of the same good are effectively treated as different goods, with different prices which vary in proportion to the welfare weights. Thus, society becomes a single fictitious consumer who maximizes the Bergson measure of $W(\underset{\sim}{x})$ subject to a budget constraint of the form $\Sigma_i \beta_i p.x_i \leq M$ where $\beta_i p_g$ is the "virtual" price of individual i's consumption of good g. The allocation $\hat{\underset{\sim}{x}}$ which results from this maximization problem is the same as that in which each individual i maximizes $U_i(x_i)$ subject to $p.x_i \leq m_i$, and the distribution of incomes $\underset{\sim}{m}$ is chosen to maximize the indirect social welfare function $V(p,\underset{\sim}{m})$ subject to $\Sigma_i \beta_i m_i \leq M$. In particular, the welfare weight β_i must be proportional to $\partial V/\partial m_i$, the marginal social welfare of i's income, so we can write:

$$\beta_i = \alpha \; \partial V/\partial m_i \tag{40}$$

for some positive constant α. In order that M really should represent total money income in the economy, α should be chosen so that:

$$M = \Sigma_i m_i = \Sigma_i \beta_i m_i = \alpha \; \Sigma_i m_i (\partial V/\partial m_i) \; . \tag{41}$$

Thus $\alpha = M/\Sigma_i m_i (\partial V/\partial m_i)$ which implies that, for each i:

$$\beta_i = \frac{M(\partial V/\partial m_i)}{\Sigma_j m_j (\partial V/\partial m_j)} \; . \tag{42}$$

Now a money metric measure of social welfare is given by:

$$\begin{aligned} E(\underset{\sim}{\beta}^O, p^O, W) : &= \min_{\underset{\sim}{x}} \{\Sigma_i \beta_i^O p^O . x_i \mid W(\underset{\sim}{x}) \geq W\} \\ &= \min_{\underset{\sim}{m}} \{\Sigma_i \beta_i^O m_i \mid V(p^O, \underset{\sim}{m}) \geq W\} \end{aligned} \tag{43}$$

based on the reference vector of prices $(\beta_i^O p_g^O)$. Arguing as in Section 4, it follows that the following second-order approximation to (43) can be derived analogously to (33):

$$\Delta W \cong \Sigma_i \; [\beta_i{}^o p^o + 1/2(\beta_i{}^1 p^1 - \beta_i{}^o p^o)] . \Delta q_i \; -$$

$$- \; 1/2[\Sigma_i (\beta_i{}^1 p^1 - \beta_i{}^o p^o) . b_i{}^o \mu_i{}^o][\Sigma_k \beta_k{}^o (p^o . \Delta q_k)] \; . \tag{44}$$

Here the vector $\underset{\sim}{\mu}^o$ is given as in Hammond (1983), eq. 4.13) by:

$$\underset{\sim}{\mu}^o = \; (\Gamma^o)^{-1} \underset{\sim}{\beta}^o \; / \; \underset{\sim}{\beta}^o . (\Gamma^o)^{-1} \underset{\sim}{\beta}^o \tag{45}$$

where Γ^o is the Hessian matrix of second order partial derivatives $\partial^2 V / \partial m_i \partial m_j$ of $V(p, \underset{\sim}{m})$ evaluated at $(p^o, \underset{\sim}{m}^o)$. To calculate the other terms in (44), notice first that, for each i:

$$\beta_i{}^1 p^1 - \beta_i{}^o p^o = \Delta\beta_i p^1 + \beta_i{}^o \Delta p \tag{46}$$

so that (44) becomes:

$$\Delta W \cong \Sigma_i \beta_i{}^o \; [(p^o + 1/2\Delta p) . \Delta q_i] + 1/2 \; \Sigma_i \Delta\beta_i (p^1 . \Delta q_i) \; -$$

$$- \; 1/2[\Sigma_i (\beta_i{}^o \Delta p + \Delta\beta_i p^1) . b_i{}^o \mu_i{}^o][\Sigma_k \beta_k{}^o (p^o . \Delta q_k)]$$

$$\cong \Sigma_i \beta_i{}^o \; [(p^o + 1/2\Delta p) . \Delta q_i] + 1/2 \; \Sigma_i \Delta\beta_i (p^o . \Delta q_i) \; -$$

$$- \; 1/2\{\Sigma_i [\beta_i{}^o (\Delta p . b_i{}^o) + \Delta\beta_i] \mu_i{}^o\}[\Sigma_k \beta_k{}^o (p^o . \Delta q_k)] \tag{47}$$

to second-order, using the fact that p^1 can be replaced by p^o in this approximation, and that $p^o . b_i{}^o = 1$ for all i.

To complete the calculation, it is necessary to derive an expression for $\Delta\beta_i$, the change in individual i's welfare weight, for each individual i. For a second-order approximation to ΔW, a first-order approximation to $\Delta\beta_i$ suffices, and this can be calculated by total differentiation of (42). In the appendix it is shown that, to first order:

$$\Delta\beta_i \cong \Sigma_k [\gamma_{ik}{}^o - (\beta_i{}^o / M^o)(\Sigma_j m_j{}^o \gamma_{jk}{}^o)] \; (p^o . \Delta q_k) + \beta_i{}^o \; [(\Delta M / M^o) - (\Delta p . b_i{}^o)] \; -$$

$$- \; (\beta_i{}^o / M^o) \Sigma_k \beta_k{}^o [\Delta m_k - m_k{}^o (\Delta p . b_k{}^o)] \tag{48}$$

$$\text{where } \gamma_{ik}{}^o := \partial^2 V / \partial m_i \partial m_k \; \text{ (all i,k)} \tag{49}$$

is the typical element of Γ^o. Substituting (48) into (47) then gives:

$$\Delta W \cong \Sigma_i \beta_i{}^o[(p^o+1/2\Delta p).\Delta q_i] - 1/2\ \Sigma_i \beta_i{}^o(\Delta p.b_i{}^o)(p^o.\Delta q_i) +$$

$$+ 1/2\ \Sigma_i[(p^o.\Delta q_i) - \mu_i{}^o\Sigma_k\beta_k{}^o(p^o.\Delta q_k)][\Sigma_j\gamma_{ij}{}^o(p^o.\Delta q_j) + \beta_i{}^o D] \quad (50)$$

$$\text{where } D := \{\Delta M - \Sigma_j\Sigma_k m_j{}^o\gamma_{jk}{}^o(p^o.\Delta q_k) - \Sigma_k\beta^o[\Delta m_k - m_k{}^o(\Delta p.b_k{}^o)]\}/M^o\ . \quad (51)$$

$$\text{Thus } \Delta W \cong \Sigma_i \beta_i{}^o[(p^o+1/2\Delta p).\Delta q_i - 1/2(\Delta p.b_i{}^o)(p^o.\Delta q_i)] +$$

$$+ 1/2\ \Sigma_i\Sigma_j(p^o.\Delta q_i)\gamma_{ij}{}^o(p^o.\Delta q_j) +$$

$$+ 1/2\ \Sigma_k\beta_k{}^o(p^o.\Delta q_k)\{D(1-\Sigma_i\beta_i{}^o\mu_i{}^o) - \Sigma_i\mu_i{}^o\Sigma_j\gamma_{ij}{}^o(p^o.\Delta q_j)\}\ . \quad (52)$$

Notice that the first two lines of the right hand side of (52) correspond to the earlier formula (34). For the reform $\Delta t = \lambda\ dt$ in a given direction dt, (52) can be written as:

$$\Delta W \cong \lambda p^o.\Sigma_i\beta_i{}^o dq_i + 1/2\ \lambda^2\hat{H} \quad (53)$$

where now, in contrast to (38), $\hat{H}$ is given by:

$$\hat{H} := dp.\Sigma_i\beta_i{}^o dq_i - \Sigma_i\beta_i{}^o(dp.b_i{}^o)(p^o.dq_i) + \Sigma_i\Sigma_j(p^o.dq_i)\gamma_{ij}{}^o(p^o.dq_j) +$$

$$+ \Sigma_k\beta_k{}^o(p^o.dq_k)\{D^*(1-\Sigma_i\beta_i{}^o\mu_i{}^o) - \Sigma_i\mu_i{}^o\Sigma_j\gamma_{ij}{}^o(p^o.dq_j)\} \quad (54)$$

with D^* as the differential form of D:

$$D^* := \{dM - \Sigma_j\Sigma_k m_j{}^o\gamma_{jk}{}^o(p^o.dq_k) - \Sigma_k\beta_k{}^o[dm_k - m_k{}^o(dp.b_k{}^o)]\}/M^o\ . \quad (55)$$

The first line of the right hand side of (54) corresponds to H in (38), of course. (53) suggests choosing a tax reform of size:

$$\lambda^* = -\ p^o.\Sigma_i\beta_i{}^o dq_i/\hat{H} \quad (56)$$

provided that dt is a favourable direction (which implies that $p^o.\Sigma_i\beta_i{}^o dq_i > 0$) and provided that $\hat{H}$ is negative. Again, the problem of what to do if $\hat{H} \geq 0$ was discussed in Section 3.

6. Pareto Improving Reforms

A direction of Pareto improvement dt is one that satisfies:

$$u_i^o.dt > 0 \quad \text{(all i)} \tag{57}$$

where $u_i^o = p^o Q_i^o$ as in Section 2. For a reform $\Delta t = \lambda$ dt in this direction, the money metric measure of individual i's gain is given by (33):

$$\begin{aligned} \Delta U_i &\cong (p^o + 1/2\ \Delta p).\Delta q_i - 1/2\ (\Delta p.b_i^o)(p^o.\Delta q_i) \\ &= \lambda p^o.dq_i + 1/2\ \lambda^2\ [dp.dq_i-(dp.b_i^o)(p^o.dq_i)] \qquad (58) \\ &= \lambda u_i^o.dt + 1/2\ \lambda^2 dt.H_i^o dt \qquad (59) \end{aligned}$$

for the Hessian matrix:

$$H_i^o: = (P^o)'Q_i^o - (P^o)'b_i^o(p^o)'Q_i^o \ . \tag{60}$$

This suggests choosing:

$$\begin{aligned} \lambda_i^*: &= -\ (u_i^o.dt)/(dt.H_i^o dt) \\ &= -\ (p^o.dq_i)/[dp.dq_i-(dp.b_i^o)(p^o.dq_i)] \qquad (61) \end{aligned}$$

for each individual i for whom $dt.H_i^o dt < 0$; let $\lambda_i^*: = +\infty$ for all other individuals. Then, using our quadratic approximations, $\Delta U_i \geq 0$ whenever the step size λ satisfies $0 \leq \lambda \leq 2\lambda_i^*$, and ΔU_i is still increasing as a function of λ whenever λ satisfies $0 \leq \lambda < \lambda_i^*$. Let $\underline{\lambda}^*$ denote the smallest of the step sizes λ_i^*; and let $\bar{\lambda}^*$ denote the largest. Then, if $\lambda < \underline{\lambda}^*$, all individuals benefit from an increase in the step size for the direction dt. If $\lambda > \bar{\lambda}^*$, all individuals benefit from a decrease in the step size. And if $\lambda > 2\underline{\lambda}^*$, at least one individual experiences a utility decrease, $\Delta U_i < 0$, so that the reform $\Delta t = \lambda dt$ is no longer a Pareto improvement. It follows that the step sizes λ which produce Pareto undominated Pareto reforms in the direction dt are those which satisfy $\underline{\lambda}^* \leq \lambda \leq \bar{\lambda}^*$ and $\lambda < 2\underline{\lambda}^*$. These inequalities determine a (possibly trivial) connected interval of step sizes that is non-empty provided only that $dt.H_i^o dt < 0$ for at least one individual. The interval is then trivial only if λ_i^* is the same for all individuals i.

7. The Direction of Tax Reform

So far I have discussed what is a suitable size for a tax reform $\Delta t = \lambda\ dt$ in a given favourable direction dt. The quadratic approximation (11), however, suggests suitable directions for tax reform as well, beyond the first order requirement that $w^o.dt > 0$. For if:

$$\Delta W \gtrsim w^o.\Delta t + 1/2\ \Delta t.H^o\Delta t \tag{62}$$

for all small tax reforms Δt, as in (11), and if the Hessian matrix H^o happens to be negative definite, then the quadratic approximation to ΔW reaches a global maximum for a reform Δt^* given by:

$$\Delta t^* = -\ (H^o)^{-1}w^o\ . \tag{63}$$

If the quadratic approximation were exact, the reform Δt^* would take the economy directly to the welfare optimum; as it is, undertaking the reform Δt^* is like using Newton's method for computing the maximum of a differentiably strictly concave function. If social welfare were a differentiably strictly concave function $W = F(t)$ of the tax parameters t, a sequence of reforms satisfying (63) at each step would converge quite rapidly to the welfare optimum.

As I have previously pointed out, however, there is no reason to believe that H^o will be negative definite. Indeed, the end of Section 3 discussed the problems that arise when $dt.H^o dt$ is not negative for a specific direction of reform dt. If H^o is not negative definite, then Δt^* in (63), even if it is well defined because H^o has an inverse, will not be a maximum at all of the quadratic approximation (62) to ΔW. If the direction as well as the size of the tax reform can be chosen, some other procedure should be followed.

As a symmetric matrix, H^o has all real eigenvalues, and can be diagonalized by applying a rotation matrix T. That is, there exists a matrix and a diagonal matrix D such that:

$$T^{-1} \quad = \quad T' \tag{64}$$

and

$$T'H^oT \quad = \quad D\ . \tag{65}$$

The diagonal elements of D are the eigenvalues of H^o. If H^o is not negative definite, some of these eigenvalues are non-negative.

Write $d_1, d_2, \ldots, d_r$ for the diagonal elements of D. Given any reform Δt, write:

$$\Delta\tilde{t}: \ = T' \ \Delta t \tag{66}$$

and let $\tilde{w}^o: \ = w^o.T$. (67)

Then (62) can be written as:

$$\begin{aligned}\Delta W &\cong w^o.\Delta t + 1/2 \ \Delta t.H^o\Delta t \\ &= \tilde{w}^o.\Delta\tilde{t} + 1/2 \ \Delta\tilde{t}.D \ \Delta\tilde{t} \\ &= \Sigma_{k=1}^{r}[\tilde{w}_k{}^o\Delta\tilde{t}_k + 1/2 \ d_k \ (\Delta\tilde{t}_k)^2] \ .\end{aligned} \tag{68}$$

Write $K:=\{k|d_k \neq 0\}$. Then:

$$\Delta W = 1/2 \ \Sigma_{k\varepsilon K} \ \{d_k[\Delta\tilde{t}_k + \tilde{w}_k^o \ / \ d_k]^2 - (\tilde{w}_k^o)^2/d_k\} + \\ + \Sigma_{k\notin K} \ \tilde{w}_k^o \ \Delta t_k \ . \tag{69}$$

If all the diagonal elements d_k (k=1 to r) are negative, then H^o is negative definite and formula (63) should be applied. But if some of the diagonal elements d_k are non-negative, then the quadratic approximation (62) or (69) has no maximum unless $\tilde{w}_r^o = 0$ for all k and K. A reform which increases the approximate value of ΔW significantly can be found by setting:

$$\Delta\tilde{t}_k = \begin{cases} - \ \tilde{w}_k{}^o/d_k & (\text{if } d_k < 0) \\ 0 & (\text{if } d_k = 0) \\ \lambda \ \tilde{w}_k{}^o/d_k & (\text{if } d_k > 0) \end{cases} \tag{70}$$

where λ is a large positive number and $\tilde{w}^o$ is given by (67). Then take:

$$\Delta t = T\Delta\tilde{t} \ . \tag{71}$$

Notice that in each case:

$$w^o.\Delta t = \tilde{w}^o.\Delta\tilde{t} > 0 \tag{72}$$

so that the reform is in a favourable direction, as well as leading to an increase in the quadratic approximation to ΔW.

APPENDIX

To calculate $\Delta\beta_i$ as in (48) it is first necessary to derive from (42):

$$\beta_i = \frac{M(\partial V/\partial m_i)}{\Sigma_j m_j(\partial V/\partial m_j)} \tag{A.1}$$

the following partial derivatives (evaluated at (p^o,m^o)):

$$\frac{\partial\beta_i}{\partial m_k} = \frac{\frac{\partial V}{\partial m_i} + M\frac{\partial^2 V}{\partial m_i \partial m_k}}{\Sigma_j m_j(\partial V/\partial m_j)} - \frac{M\frac{\partial V}{\partial m_i}\left(\frac{\partial V}{\partial m_k} + \Sigma_j m_j \frac{\partial^2 V}{\partial m_j \partial m_k}\right)}{[\Sigma_j m_j(\partial V/\partial m_j)]^2}$$

$$= \gamma_{ik}{}^o + (\beta_i{}^o/M^o) - (\beta_i{}^o/M^o)(\beta_k{}^o + \Sigma_j m_j{}^o\gamma_{jk}{}^o) \tag{A.2}$$

$$\frac{\partial\beta_i}{\partial p_g} = \frac{M\frac{\partial^2 V}{\partial m_i \partial p_g}}{\Sigma_j m_j(\partial V/\partial m_j)} - \frac{M\frac{\partial V}{\partial m_i}\left(\Sigma_j m_j \frac{\partial^2 V}{\partial m_j \partial p_g}\right)}{[\Sigma_j m_j(\partial V/\partial m_j)]^2}$$

$$= \frac{\partial^2 V}{\partial m_i \partial p_g} - \frac{\beta_i{}^o}{M^o}\Sigma_j m_j{}^o \frac{\partial^2 V}{\partial m_j \partial p_g} \quad . \tag{A.3}$$

But $\frac{\partial V}{\partial p_g} = -\Sigma_k \frac{\partial V}{\partial m_k} h_{kg}(p,m_k)$, by Roy's identity, so:

$$\frac{\partial^2 V}{\partial m_j \partial p_g} = -\Sigma_k \frac{\partial^2 V}{\partial m_j \partial m_k} q_{kg} - \frac{\partial V}{\partial m_j}\partial h_{jg}/\partial m_j$$

$$= -\Sigma_k\, \gamma_{jk}{}^o\, q_{kg}{}^o - \beta_j{}^o b_{jg}{}^o \tag{A.4}$$

and:

$$\frac{\partial\beta_i}{\partial p_g} = -\Sigma_k\gamma_{ik}{}^o q_{kg}{}^o - \beta_i{}^o b_{ig}{}^o + \frac{\beta_i{}^o}{M^o}\Sigma_j m_j{}^o[\Sigma_k(\gamma_{jk}{}^o q_{kg}{}^o) + \beta_j{}^o b_{jg}{}^o]\ . \tag{A.5}$$

So, to first order, $\Delta\beta_i$ is approximately equal to:

$$\Sigma_k(\partial\beta_i/\partial m_k)\ \Delta m_k + \Sigma_g(\partial\beta_i/\partial p_g)\ \Delta p_g$$

$$= \Sigma_k\ \gamma_{ik}{}^{o}(\Delta m_k - \Delta p.q_k{}^{o}) + \beta_i{}^{o}(\Delta M/M^{o}) - \beta_i{}^{o}(\Delta p.b_i{}^{o}) -$$

$$- (\beta_i{}^{o}/M^{o})\{\Sigma_k(\beta_k{}^{o}+\Sigma_j m_j{}^{o}\gamma_{jk}{}^{o})\Delta m_k - \Sigma_j m_j{}^{o}[\Sigma_k\gamma_{jk}{}^{o}(\Delta p.q_k{}^{o})+\beta_j{}^{o}(\Delta p.b_j{}^{o})]\}$$

$$= \Sigma_k[\gamma_{ik}{}^{o} - (\beta_i{}^{o}/M^{o})\Sigma_j m_j{}^{o}\gamma_{jk}{}^{o}](p^{1}.\Delta q_k) + \beta_i{}^{o}(\Delta M/M^{o}) - \beta_i{}^{o}(\Delta p.b_i{}^{o}) -$$

$$- (\beta_i{}^{o}/M^{o})\Sigma_k\beta_k{}^{o}[\Delta m_k - m_k{}^{o}(\Delta p.b_k{}^{o})] \tag{A.6}$$

$$= \Sigma_k[\gamma_{ik}{}^{o} - (\beta_i{}^{o}/M^{o})(\Sigma_j m_j{}^{o}\gamma_{jk}{}^{o})](p^{o}.\Delta q_k) + \beta_i{}^{o}[(\Delta M/M^{o}) - (\Delta p.b_i{}^{o})] -$$

$$- (\beta_i{}^{o}/M^{o})\Sigma_k\beta_k{}^{o}[\Delta m_k - m_k{}^{o}(\Delta p.b_k{}^{o})] \tag{A.7}$$

because $p^1.\Delta q_k = p^o.\Delta q_k$ to first order.

REFERENCES

R.G.D. Allen (1949), "The Economic Theory of Index Numbers", Economica, 16, 197-203.

A.B. Atkinson and J.E. Stiglitz (1980), Lectures on Public Economics (New York: McGraw-Hill).

J.S. Chipman and J.C. Moore (1980), "Compensating Variation, Consumer's Surplus, and Welfare", American Economic Review, 70, 933-949.

W.E. Diewert (1978), "Optimal Tax Perturbations", Journal of Public Economics, 10, 139-177.

F. Fogelman, R. Guesnerie and M. Quinzii (1978), "Dynamic Processes for Tax Reform Theory", Journal of Economic Theory, 17, 200-226.

R. Guesnerie (1977), "On the Direction of Tax Reform", Journal of Public Economics, 7, 179-202.

P.J. Hammond (1983), "Approximate Measures of the Social Welfare Benefits of Large Projects", Stanford University, Institute of Mathematical Studies in the Social Sciences, Economics Technical Report No. 410, to appear in S.B. Dahiya (ed.) Theoretical Foundations of Development Planning.

P.J. Hammond (1984), "Welfare Economics",Hebrew University of Jerusalem, Institute of Advanced Studies Report No. 2/83, to appear in G. Feiwel (ed.), Issues in Contemporary Microeconomics and Welfare (Macmillan).

M.R. Hestenes (1980), Conjugate Direction Methods in Optimization (New York: Springer-Verlag).

J.R. Hicks (1942), "The Rehabilitation of Consumer's Surplus", Review of Economic Studies, 9, 126-137.

G.W. McKenzie (1983), Measuring Economic Welfare: New Methods (Cambridge: Cambridge University Press).

G.W. McKenzie and I.F. Pearce (1976), "Exact Measures of Welfare and the Cost of Living", Review of Economic Studies, 43, 465-468.

G.W. McKenzie and I.F. Pearce (1982), "Welfare Measurement: A Synthesis", American Economic Review, 72, 669-682.

J.E. Meade (1955), Trade and Welfare (Mathematical Supplement) (Oxford: Oxford University Press).

P.A. Samuelson (1974), "Complementarity - An Essay on the 40th Anniversary of the Hicks-Allen Revolution in Demand Theory", Journal of Economic Literature, 12, 1255-1289.

A.K. Sen (1976), "Real National Income", Review of Economic Studies, 43, 19-39.

A.K. Sen (1979), "The Welfare Basis of Real Income Comparisons: A Survey", Journal of Economic Literature, 17, 1-45.

J. Tirole and R. Guesnerie (1981), "Tax Reform from the Gradient Projection Viewpoint", Journal of Public Economics, 15, 275-293.