The Distance Function in Consumer Behaviour with Applications to Index Numbers and Optimal Taxation

ANGUS DEATON
University of Bristol

The use of duality concepts has now become widespread in production and consumption theory. In both of these areas, one of the central and most useful concepts is the *cost* or *expenditure function* which represents tastes or technology through the minimum cost to consumer or producer of reaching a particular utility or output level at given prices. Under utility maximizing or cost minimizing assumptions, the value of this function is actual outlay, so that the cost function defines the relation between expenditure, prices and utility or production as the case may be. For many problems, prices and outlay are the natural variables with which to work, and it is this that makes the cost function such a convenient representation of preferences. Even so, the mathematical properties of the cost function, particularly its homogeneity and concavity, give it decisive advantages over either direct or indirect utility functions even in situations where quantities are the more natural variables. For this reason, it is useful to consider the dual of the cost function itself, retaining its mathematical properties, but defined on primal, rather than dual variables. This dual is the *distance function*, sometimes also referred to as the transformation function, the gauge function, or the *direct* cost function.

This function has made a number of distinguished but infrequent appearances in the literature. Wold (1943) uses it to relate quantity bundles to a given reference vector and thus to define a utility function. Debreu (1951) defines a " coefficient of resource utilization " through the distance function while Malmquist (1953) develops a systematic theory of quantity indices based upon it. In production theory, the distance function is discussed by Shephard (1953) and more recently is systematically and extensively used in the forthcoming monograph by Fuss and McFadden (1978), especially in the contributions by McFadden and by Hanoch. In the demand context, the function is briefly discussed by Diewert (1974), but the main contributions are in an unpublished paper by Gorman (1970) and, more briefly, in Gorman (1976). The function is used in a number of recent publications, notably by Blackorby and Russell (1975), Hanoch (1975), Blackorby and Donaldson (1976), Blackorby, Lovell and Thursby (1976) and Diewert (1976*a*) and (1976*b*). The aim of the present paper is to present a reasonably systematic, if informal presentation of the distance function in the context of consumer behaviour. Much of what follows is derived from one or more of the contributions listed above. However, none of these provides anything like a complete treatment, and in view of the wide range of potential applications of the analysis, in particular to demand studies, to rationing theory and to welfare economics generally, a synthesis of such useful material is overdue.

Section 1 of the paper defines the distance function and discusses its properties. Particular attention is focused on the duality between the distance and cost functions and on the use of the former in deriving compensated *inverse* demand functions. These latter seem to have been first systematically treated by Hicks (1956, Chapter XVI) who used them to define " q-complements " and " q-substitutes " in contrast to the now standard " p-complements " and " p-substitutes ". The matrix of q-substitution effects is the Antonelli matrix of integrability theory, Samuelson (1950), and we show how this and the Slutsky matrix of p-substitution effects can be regarded as generalized inverses of one another. Section 2 takes up Malmquist's (1953) analysis and discusses the theory of quantity and utility indices based on the distance function and its dual relation to the price and utility indices based on the cost function. Finally, Section 3 gives a brief foretaste of the application of the analysis. The familiar Ramsey rule for optimal taxation in an equity disregarding society is derived in a new and very simple form. Instead of stating the optimal tax rule in terms of its effects upon quantities, the distance function approach allows a direct characterization of the tax rates themselves.

1. THE DISTANCE FUNCTION

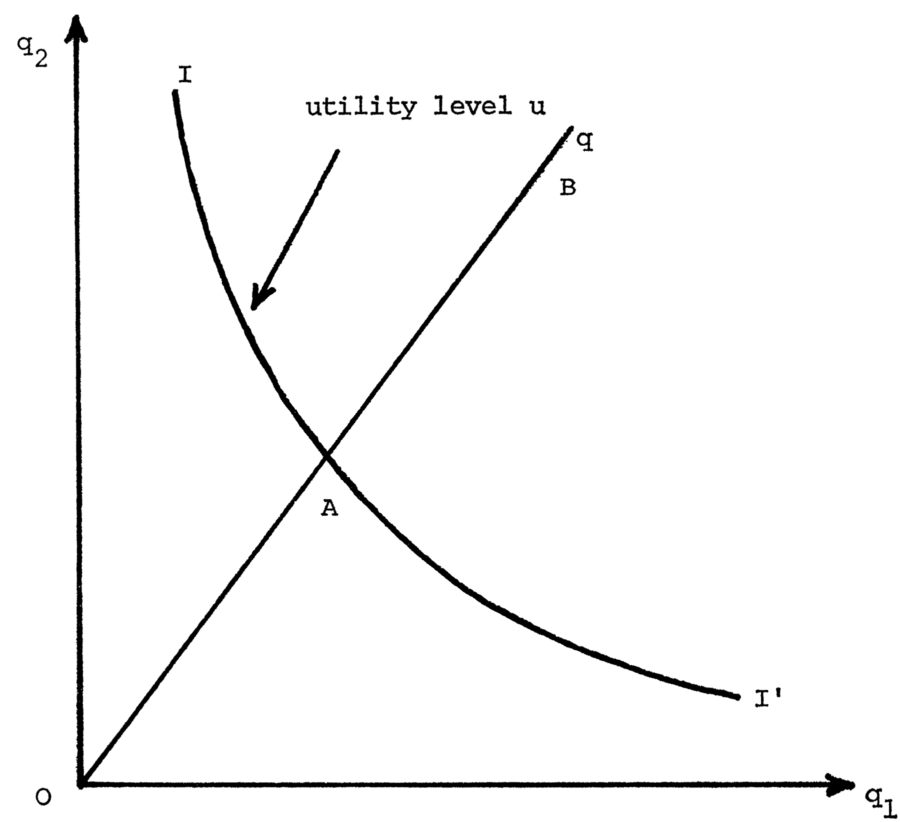

FIGURE 1

Primal space.

Figure 1 illustrates an arbitrary indifference curve II′ in quantity space corresponding to a strictly quasi-concave utility function; according to some normalization of utility, this curve is labelled u. OB is an arbitrary given quantity vector q. The distance function $d(u, q)$, defined on utility u and quantity vector q, gives the amount by which q must be divided in order to bring it on to the indifference curve u. Geometrically, the value of the distance function is the ratio OB/OA. Mathematically, if preferences are represented by the direct utility function $v(\ )$, then the distance function, $d(u, q)$, is defined on arbitrary utility level u and quantity vector q, by the equation

$$v\{q/d(u, q)\} = u. \quad \text{...(1)}$$

Note that in the special case when $u = v(q)$, $d(u, q) = 1$, hence

$$u = v(q) \quad \text{iff } d(u, q) = 1. \quad \text{...(2)}$$

Hence, by appeal to (2) we can always write the direct utility function in the equivalent implicit form $d(u, q) = 1$. Clearly $d(u, q)$ for fixed u is a scalar measure of the magnitude of q and, in this sense, it is simply a quantity index number. Similarly, for fixed q, $d(u, q)$ is an (inverse) measure of utility. Note too that the distance function is entirely *ordinal*; it is defined with reference to an indifference surface and not with respect to any given cardinalization of preferences. Figure 1 should make this clear.

To relate $d(u, q)$ to the cost function, we move to the dual space. Write $\psi(x, p)$ for the indirect utility function giving u as a function of prices and total expenditure x. Then, since $\psi(x, p)$ is homogeneous of degree zero

$$u = \psi(x, p) = \psi\left(1, \frac{p}{x}\right) = \psi^*\left(\frac{p}{x}\right), \text{ say.} \quad \text{...(3)}$$

To define a dual distance function $d^*(u, p)$, say, which corresponds to (1), we write

$$\psi^*\{p/d^*(u, p)\} = u. \quad \text{...(4)}$$

But, by (3),

$$d^*(u, p) = x = c(u, p), \quad \text{...(5)}$$

where $c(u, p)$ is the cost function, the minimum cost of reaching utility u at prices p. Hence $d(u, q)$ and $c(u, p)$ are dual to one another.

Gorman (1976) derives this in another way. Let $d(u, q) = \lambda$, say, and write $q^* = q/\lambda$. Thus

$$d(u, q)c(u, p) = \lambda c(u, p) \leqq \lambda q^*.p = q.p, \quad \text{...(6)}$$

for arbitrary u, p and q. (The inequality in (6) comes from the definition of the cost function, whereby for all q yielding u, $c(u, p) \leqq q.p$.) The inequality can be replaced by an equality for some p and q. Following earlier work by Afriat, Gorman says that price and quantity vectors p and q are *conjugate* at utility u if the cheapest way of reaching u at p is a vector proportional to q. Clearly, in this case, if $\tilde{p}$ and $\tilde{q}$ are conjugates

$$d(u, \tilde{q})c(u, \tilde{p}) = \tilde{q}.\tilde{p}. \quad \text{...(7)}$$

From (6) and (7), it follows immediately that

$$d(u, q) = \min_p \{p'q\colon c(u, p) = 1\}, \quad \text{...(8)}$$

and since the direct utility function can be implicitly written $d(u, q) = 1$, the cost function can be redefined as

$$c(u, p) = \min_q \{q'p\colon d(u, q) = 1\}. \quad \text{...(9)}$$

Equations (8) and (9) make the duality between $d(u, q)$ and $c(u, p)$ absolutely transparent and allow us immediately to prove standard results for the distance function from the well-known corresponding results for the cost function. These are summarized below; proofs are not given since the results are either obvious or follow immediately from the corresponding results for the cost function.

Property 1. *Just as $c(u, p) = x$ defines the indirect utility function $u = \psi(x, p)$, $d(u, q) = 1$ defines the direct utility function $u = v(q)$.*

Property 2. *$d(u, q)$ is increasing in q and* decreasing *in u.*

Property 3. *$d(u, q)$ is homogeneous of degree one in q.* This follows at once from either (1) or (8). In order to avoid problems at the origin, the domain of $v(\ )$ is restricted to the non-negative orthant *excluding* the origin, and we confine attention to $u > v(O)$.

Property 4. *$d(u, q)$ is concave in q.* From (8) we see that the distance function is a minimum value function so that the concavity of $d(u, q)$ follows in the same way as does the concavity of $c(u, p)$ from (9). (Equation (1) also implies the concavity of $d(u, q)$ given quasi-concavity of $v(q)$. We shall discuss the case of non-convex preferences below.)

Property 5. *Whenever they are defined, the partial differentials of $d(u, q)$, which we write $a_i(u, q)$, are the prices normalized with reference to total expenditure x.* Hence, writing r_i for p_i/x,

$$\frac{\partial d(u, q)}{\partial q_i} \equiv a_i(u, q) = r_i \equiv \frac{p_i}{x}. \qquad ...(10)$$

Differentiating the cost function $c(u, p)$ gives the Hicksian compensated demand functions $h_i(u, p)$, say, with quantity demanded as a function of utility and prices. Correspondingly, differentiation of the distance function gives compensated *inverse* demand functions with expenditure normalized price (marginal willingness to pay) as a function of utility and quantity supplied. Just as substitution of $u = \psi(x, p)$ in $h_i(u, p)$ leads from compensated to uncompensated demands, substitution of $u = v(q)$ in $a_i(u, q)$ leads to uncompensated inverse demands. Note also from (10), if p and q are conjugate at u,

$$\frac{\partial \ln d(u, q)}{\partial \ln q_i} = w_i(u, q) = w_i(u, p) = \frac{\partial \ln c(u, p)}{\partial \ln p_i}, \qquad ...(11)$$

where $w_i = p_i q_i/x$ is the value share devoted to good i.

Property 6. *The compensated inverse demand functions $a_i(u, q)$ are homogeneous of degree zero in q.* This follows at once from the linear homogeneity of $d(u, q)$ and from (10). Hence, the functions $a_i(u, q)$ associate with each indifference curve and with each quantity *ray* (only proportions matter) a set of expenditure normalized (shadow prices). Figure 2 illustrates. Note that a_i/a_j is simply the MRS along u at q. In this form, the MRS is a function of both u and q so that we can easily separate changes in the MRS due to changes in welfare from those due to changes in proportions.

Property 7. *Just as the Hessian of the cost function is the Slutsky matrix of compensated derivatives of quantities with respect to price, the Hessian of the distance function is the matrix of compensated derivatives of normalized prices with respect to quantity.* This latter is known in the literature as the *Antonelli* matrix, see e.g. Samuelson (1950). Properties 3–6 imply that, like the Slutsky matrix, the Antonelli matrix is symmetric and negative semi-definite. In observations of markets where quantities, rather than prices, are exogenous, these conditions on the Antonelli matrix would be testable in the same way that Slutsky conditions are often tested in the dual situation.

Formally, write A for the matrix whose i, jth term is $\partial^2 d(u, q)/\partial q_i \partial q_j$. Then

$$\text{(i) } Aq = 0 \quad \text{(ii) } A = A' \quad \text{(iii) } \theta' A \theta \leqq 0 \qquad ...(12)$$

The Antonelli matrix forms the basis for the definition of Hicks' (1956) q-complements and q-substitutes just as the Slutsky matrix defines p-complements and p-substitutes. Thus, goods i and j are q-complements if, and only if, $a_{ij} > 0$ so that, in view of (10), the marginal valuation of i rises with the quantity of j bought *along an indifference curve.* Similarly for substitutes. As Hicks emphasizes, this definition is *not* equivalent to that of p-complements and p-substitutes and is much closer in spirit to the early definitions in terms of marginal utilities by Edgeworth and Pareto.

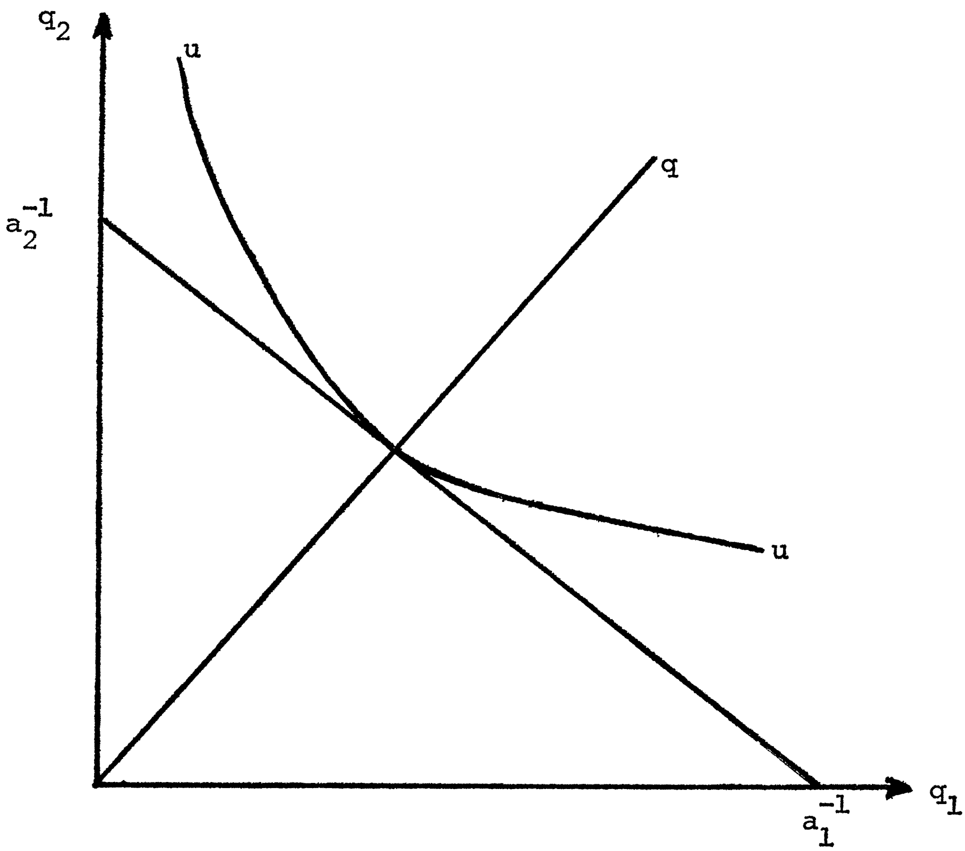

FIGURE 2

Properties 1–7 mean that all the propositions of demand theory can be stated via the distance function just as they can through the cost function. Which is more appropriate depends on the problem at hand.

Property 8. *The Slutsky matrix S and the Antonelli matrix A are generalized inverses of one another* (see also Bronsard *et al.* (1976*b*)). Since this proposition cannot be argued by analogy, we give a brief discussion.

The matrix S is given by $s_{ij} = \partial^2 c/\partial p_i \partial p_j$. Since $c(u, p)$ is linearly homogeneous in p, $c(u, p) = xc(u, r)$. Elementary calculus applied to the derivative property of the cost function gives

$$q_i = h_i(u, r) = \partial c(u, r)/\partial r_i \qquad \text{...(13)}$$

$$s^*_{ij} \equiv xs_{ij} = \partial^2 c(u, r)/\partial r_i \partial r_j. \qquad \text{...(14)}$$

Write $\nabla c(u, r)$ and $\nabla d(u, q)$ for the vectors of first derivatives of $c(u, r)$ and $d(u, q)$, i.e. $h(u, r)$ and $a(u, q)$ respectively. Hence, from (13) and (10) in turn

$$q = h(u, r) = h\{u, \nabla d(u, q)\} = h[u, \nabla d\{u, \nabla c(u, r)\}] \qquad \text{...(15)}$$

$$r = a(u, q) = a\{u, \nabla c(u, r)\} = a[u, \nabla c\{u, \nabla d(u, q)\}] \qquad \text{...(16)}$$

Differentiating (15) with respect to q and (16) with respect to r and repeatedly applying the chain rule

$$S^* = S^*AS^* \qquad \text{...(17)}$$

$$A = AS^*A. \qquad \text{...(18)}$$

These, plus the respective " normalization " rules, $S^*r = Aq = 0$ establish S^* ($\equiv xS$) and A as generalized inverses. In many applications, it is useful to have explicit formulae for AS^* and S^*A, for example when we wish to " invert " S in optimal tax formulae.

From Figure 2,

$$q_i/d(u, q) = h_i\{u, \nabla d(u, q)\} \qquad \text{...(19)}$$

so that multiplying by $d(u, q)$ and differentiating with respect to q_j gives in matrix form

$$S^*A = I - qr'. \qquad \text{...(20)}$$

Similarly, for $p_i/c(u, p) = a_i\{u, \nabla c(u, r)\}$, or by transposition of (20)

$$AS^* = I - rq'. \qquad \text{...(21)}$$

These equations can also be used to check (17) and (18). Note, as always, the close parallels between duality and matrix inversion. Much of the power of duality methods comes from their ability to replace mechanical matrix inversion by elementary algebra with transparent economic interpretations.

We note finally the consequences of *not* assuming quasi-concavity for the direct utility function. The equivalence of the two definitions of $d(u, q)$, i.e. equations (1) and (8), depends upon convexity of preferences. If (1) is used, $d(u, q)$ will be (strictly) concave and continuously differentiable if, and only if, $v(q)$ is (strictly) quasi-concave and continuously differentiable. However, if preferences are non-convex, the cost-function will " bridge " non-convex portions of indifference curves so that if (8) is used to define $d(u, q)$, $c(u, p)$ having been defined as usual, then $d(u, q)$ will be concave, everywhere continuous and first and second differentiable almost everywhere, independently of the quasi-concavity of $v(q)$. The direct utility function which can be " recovered " from the cost function via (8) by setting $d(u, q) = 1$, will not, of course, be the original utility function unless the latter is quasi-concave. However, this synthetic utility function, which has the same indifference curves as the original but with non-convexities bridged, is the economically relevant one. Points where indifference curves are non-convex cannot be supported by shadow prices and the inverse demand functions have no meaning in such situations. Hence, if we confine attention to situations with linear budget sets, the best general procedure would seem to be to use the cost function, defining the distance function by (8) and using $d(u, q) = 1$ rather than the original utility function as the direct representation of preferences.

2. THE DISTANCE FUNCTION IN INDEX NUMBER THEORY

The foregoing analysis can be used to bring together a number of apparently unrelated index number concepts. Referring back to Figure 1, Debreu (1951) used the ratio OA/OB, the reciprocal of the distance function, as his " coefficient of resource utilization ". Since Debreu uses maximum potential welfare as reference, the coefficient of resource utilization is a measure of economic efficiency. Similarly, the distance function is a measure of the *inefficiency* of q relative to u. Very similar ideas are involved in Engel's measure of household equivalence scales. If this is interpreted in a utility context, see Muellbauer (1977), family composition effects act to " scale " quantities consumed. Hence if m_h is the number of adult equivalences relative to a reference family, and if the latter has direct utility function $v(q)$, u_h is given by $u_h = v(q_h/m_h)$. Clearly then, m_h, the family equivalence scale, is simply $d(u_h, q_h)$ where $d(\ )$ is the distance function of the reference household. Finally, if $v(\ )$ is interpreted as a Bergson–Samuelson social welfare function and q as a vector of household incomes, $d(u, \iota\bar{q})$, where u is actual social welfare, ι is the unit vector and $\bar{q}$ is average income, is a measure of the extent to which average income could be reduced without loss of social welfare if incomes were to be redistributed perfectly equally. Atkinson's (1970) inequality measure is thus $1 - \{d(u, \iota\bar{q})\}^{-1}$ and can again be interpreted as measure of the inefficiency of the current distribution in achieving social objectives. This relationship between distance functions and inequality indices has recently been explored in detail by Blackorby and Donaldson (1976).

By far the most systematic use of these concepts has been made by Malmquist (1953). Since Konüs' pioneering contributions, (1924) and (1939), it has been realized that for constant utility level, u_R say, $c(u_R, p)$ is an index of prices. Similarly, for u_R fixed, $d(u_R, q)$ is the dual quantity index. If q is chosen at prices p for utility u, i.e. if p and q are conjugate at u_R

$$d(u_R, q) \cdot c(u_R, p) = x, \qquad \ldots(22)$$

or if the indices are Q and P respectively

$$Q \cdot P = x. \qquad \ldots(23)$$

This relation and the derivative property (11) can also be used to define Divisia indices. Differentiating in logs, holding u_R constant

$$d \ln Q + d \ln P = d \ln x, \qquad \ldots(24)$$

or

$$\textstyle\sum_k w_k d \ln q_k + \sum_k w_k d \ln p_k = d \ln x. \qquad \ldots(25)$$

The value shares in this equation are those which occur only when p and q are conjugate at u_R, so that (25) is only exactly valid for infinitesimal changes. Nevertheless, the equation may be a fair approximation for small finite changes and is, for example, the basis for the definition of " real income " in the Rotterdam model, see e.g. Theil (1975).

However, most index numbers are used to compare two situations in time or space; let us index these by the subscripts 0 and 1. The cost function is used to give a Konüs price index by holding utility constant, i.e.

$$P(p^1, p^0; u) = \frac{c(u, p^1)}{c(u, p^0)}, \qquad \ldots(26)$$

while, if prices are held constant, we have a comparison of utility levels,

$$U(u^1, u^0; p) = \frac{c(u^1, p)}{c(u^0, p)}, \qquad \ldots(27)$$

where, in each case, the symbol after the semicolon indicates the information on which the index is based. Equation (27) compares two indifference curves by comparing the expenditure necessary to reach them at constant reference prices; hence Samuelson's (1974) name of " money metric utility ".

The distance function can be used in exactly the same way to give the dual Malmquist quantity indices:

first, a constant utility quantity index

$$Q(q^1, q^0; u) = \frac{d(u, q^1)}{d(u, q^0)}, \qquad \ldots(28)$$

and second, a quantity-reference utility index

$$U(u^1, u^0; q) = \frac{d(u^0, q)}{d(u^1, q)}. \qquad \ldots(29)$$

Thus, while (27) measures indifference curves by the outlays necessary to reach them at reference prices, (29) measures them by the distances from the origin at which they cut the reference quantity vector.

In theory, all these indices are independent of one another but, in practice, reference levels of p, q or u, as appropriate, must be chosen and in most cases, the most natural selection is the relevant variable at either 0 or 1. Thus equations (26) to (29) yield four utility indices and two each of price and quantity indices. Fortunately, these eight indices are not independent of one another.

From (28),

$$Q(q^1, q^0; u^0) = \frac{d(u^0, q^1)}{d(u^0, q^0)} = d(u^0, q^1) = U(u^1, u^0; q^1), \qquad \text{...(30)}$$

and similarly

$$Q(q^1, q^0; u^1) = U(u^1, u^0; q^0), \qquad \text{...(31)}$$

i.e. the base-weight quantity index is the current-weight utility or real-income index, and *vice versa*. The corresponding (well-known) relation between the price and money-metric utility index is

$$P(p^1, p^0; u^0) = \frac{c(u^0, p^1)}{c(u^0, p^0)} = \frac{c(u^0, p^1)}{c(u^1, p^1)} \cdot \frac{x^1}{x^0} = \frac{X}{U(u^1, u^0; p^1)} \qquad \text{...(32)}$$

where X is the expenditure index x^1/x^0. Similarly for $P(p^1, p^0; u^1)$ and $U(u^1, u^0; p^0)$ so that

$$\text{(i) } P(p^1, p^0; u^0)U(u^1, u^0; p^1) = X,$$

and ...(33)

$$\text{(ii) } P(p^1, p^0; u^1)U(u^1, u^0; p^0) = X.$$

There are thus only four distinct indices (provided only reference points from 0 and 1 are allowed): two quantity or real-income indices, and two price indices. The other four indices can be derived from these.

If, and only if preferences are homothetic, utility factors out of both cost and distance functions, i.e. in an appropriate normalization $c(u, p) = a(p).u$ and $d(u, q) = b(q)/u$. In this case it is easy to show that all indices are consistent so that, however defined

$$Q = U = X/P. \qquad \text{...(34)}$$

In the more practical non-homothetic case, there will still be inequality relations between the various indices which are different from, but parallel the inequalities between constant utility and Paasche and Laspeyres price indices.

Note first, from (6)

$$d(u^0, q^1)c(u^0, p^0) \leqq p^0 . q^1 \qquad \text{...(35)}$$

Hence

$$Q(q^1, q^0; u^0) = d(u^0, q^1) \leqq \frac{p^0 \cdot q^1}{c(u^0, p^0)} = \frac{p^0 \cdot q^1}{p^0 \cdot q^0}, \qquad \text{...(36)}$$

i.e. the base-weighted constant-utility quantity index is no greater than the base-weight Laspeyres quantity index. Similarly,

$$Q(q^1, q^0; u^1) = \frac{1}{d(u^1, q^0)} \geqq \frac{c(u^1, p^1)}{q^0 \cdot p^1} = \frac{q^1 \cdot p^1}{q^0 \cdot p^1}, \qquad \text{...(37)}$$

so that the current-weighted constant-utility quantity index is no less than the current-weight Paasche quantity index. These are clearly the exact parallels of the inequalities for the Konüs price index.

Finally, there are two inequalities which have nothing to do with Paasche or Laspeyres indices and which link the cost function indices with the distance function indices. From (29),

$$U(u^1, u^0; q^0) = \frac{1}{d(u^1, q^0)} \geqq \frac{c(u^1, p^0)}{x^0} = \frac{c(u^1, p^0)}{c(u^0, p^0)} = U(u^1, u^0; p^0) \qquad \text{...(38)}$$

$$U(u^1, u^0; q^1) = d(u^0, q^1) \leqq \frac{x^1}{c(u^0, p^1)} = \frac{c(u^1, p^1)}{c(u^0, p^1)} = U(u^1, u^0; p^1). \qquad \text{...(39)}$$

Thus, combining results:

$$Q(q^1, q^0; u^1) = U(u^1, u^0; q^0) \geqq U(u^1, u^0; p^0) = X/P(p^1, p^0; u^1) \qquad ...(40)$$

$$Q(q^1, q^0; u^0) = U(u^1, u^0; q^1) \leqq U(u^1, u^0; p^1) = X/P(p^1, p^0; u^0). \qquad ...(41)$$

In words, using *current* reference levels, quantity-metric utility (= real income) is *no less than* money-metric utility (= real income); using *base* reference levels, quantity-metric utility is *no greater than* money metric utility. These inequalities simply reflect the fact that money is better than goods. From (41), period 0's quantity vector has to be increased by more than does period 0's income to attain period 1's welfare. Similarly, a proportionate reduction in current consumption levels—with proportions fixed—will decrease welfare more than an identical proportionate reduction in income. If, for example, quantities are constrained by rationing, the difference in the two indices measures the cost of rationing.

Figures 3 and 4 illustrate the two inequalities. In both figures, the same two indifference curves u_0 and u_1 are shown; u_0 is reached in period 0 at E_0 (Figure 3), while u_1 is reached in period 1 at E_1 (Figure 4). Money metric utility, $U(u^1, u^0; p)$ is computed using prices of both periods to draw hypothetical tangents and then taking ratios of the perpendiculars to these tangents. Hence, from Figure 3, $U(u^1, u^0; p^0)$ is OB_0/OA_0, while in Figure 4, $U(u^1, u^0; p^1)$ is OB_1/OA_1. The quantity metric measure proceeds by pro-

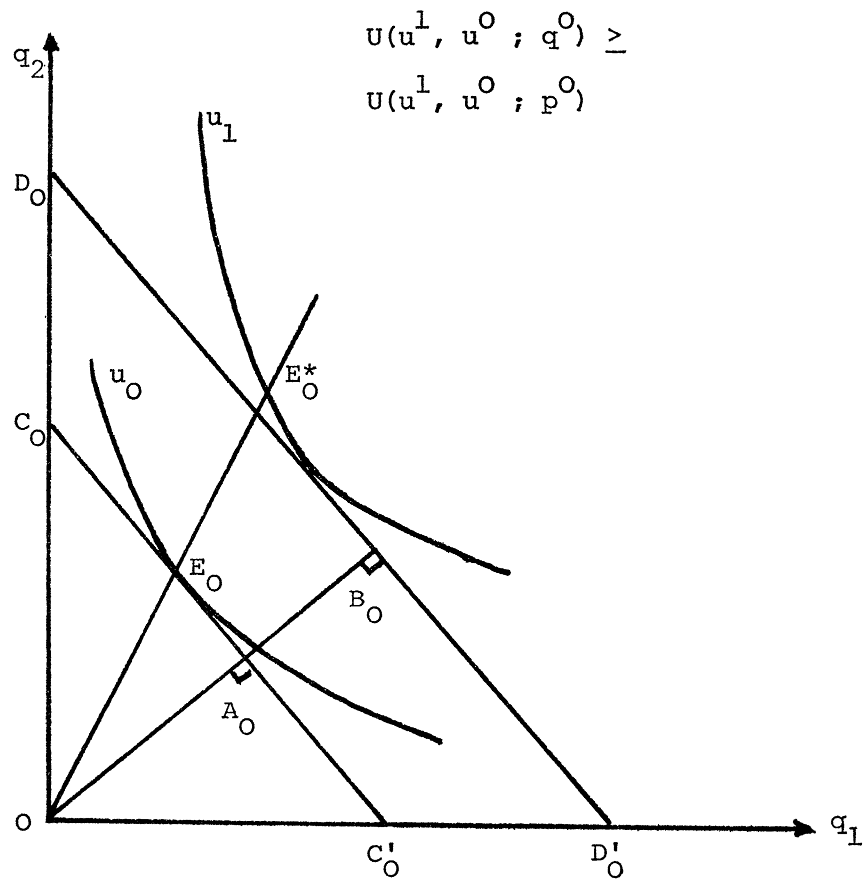

FIGURE 3

Inequality (40).

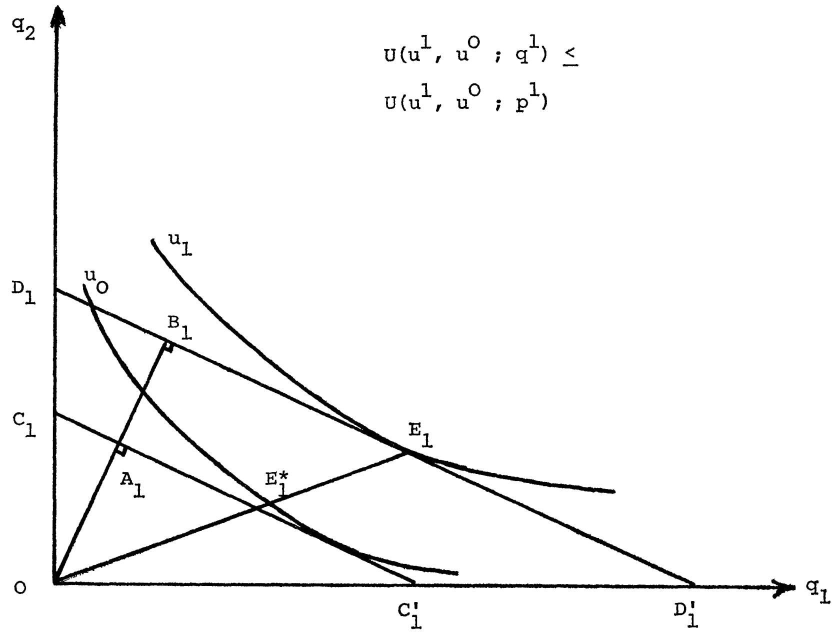

FIGURE 4

Inequality (41).

jecting the base quantity vector till it cuts the new indifference curve, E_0^* in Figure 3 and E_1^* in Figure 4. Thus $U(u^1, u^0; q^0)$ is OE_0^*/OE_0, while $U(u^1, u^0; q^1)$ is OE_1/OE_1^*. It is obvious from the diagrams that it is the convexity of preferences which guarantees the results.

Some of the index numbers discussed above are illustrated for post-war British data in Table I. These estimates are based on predictions of the linear expenditure system fitted to an eight-commodity disaggregation of total non-durable consumers expenditure over the years 1954 to 1974. Columns 1–3 show the Divisia indices for total non-durable expenditure, prices and quantities according to equation (25). These are simply the annual changes in logarithms weighted by last period's value shares; no attempt has been made to improve the approximations. Column 4 shows the percentage increase in total expenditure in each of the years. Other indices are expressed in comparable form so that, for example, the column labelled Q_1 is given by

$$Q_1 = (Q(q^t, q^{t-1}; u^t)-1)\times 100, \qquad \ldots(42)$$

i.e. the base is constantly updated. Thus columns 5 and 6 show that in 1954, substitution in response to price changes meant that a 3·9 per cent increase in quantities would have been required to match a 3·6 per cent increase in total expenditure, and both would have resulted in 1955's welfare level.

Columns 7 and 8 give the same figures using base level welfare levels; 1954's welfare level would have resulted in 1955 either if quantities had been reduced by 3·2 per cent or total expenditure by 3·6 per cent. As one might expect, the four quantity or real-expenditure indices are not very different from one another but this is a consequence of working with time-series rather than with, say, cross-country comparisons. Note however that the two price-based quantity indices, (6) and (8), are much closer together than are the two quantity-based quantity indices, (5) and (7). This appears to be due to the inelasticity of demand inherent in the calculations. Changes in prices cause less than proportionate changes in quantities so that variations in weights cause greater alterations in quantity indices, which essentially use prices as weights. Although one might reasonably expect such inelasticity to characterize a relatively aggregated breakdown of consumers expenditure such as that used in the present study, inelasticity is also an inherent property of the linear expenditure system. Note finally that the whole exercise could be repeated with price indices rather than quantity indices, comparing P_1 with X/Q_1 and P_0 with X/Q_0. All these numbers can easily be constructed from the information in the table.

TABLE I

per cent

	Divisia indices			Expenditure index	Quantity indices			
	$d\log x \simeq d\log P + d\log Q$			X	$Q_1 = U_{q0} \geqq U_{p0} = X/P_1$		$Q_0 = U_{q1} \leqq X/P_0 = U_{p1}$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
55–4	7·07	3·47	3·53	7·32	3·87	3·62	3·20	3·61
56–5	5·69	3·96	1·71	5·86	1·79	1·74	1·67	1·74
57–6	4·55	3·48	1·05	4·65	1·07	1·05	1·04	1·06
58–7	4·10	3·03	1·04	4·18	1·05	1·03	1·03	1·05
59–8	3·81	1·21	2·58	3·89	2·71	2·62	2·50	2·63
60–59	4·50	1·23	3·22	4·60	3·42	3·28	3·14	3·31
61–0	5·10	3·28	1·80	5·23	1·86	1·82	1·79	1·83
62–1	5·09	4·23	0·85	5·22	0·86	0·85	0·85	0·86
63–2	5·13	2·54	2·57	5·27	2·67	2·61	2·55	2·62
64–3	5·56	3·48	2·07	5·72	2·13	2·09	2·05	2·10
65–4	5·98	5·10	0·87	6·16	0·87	0·87	0·86	0·87
66–5	5·94	4·28	1·65	6·12	1·68	1·66	1·65	1·67
67–6	4·15	2·99	1·15	4·24	1·17	1·16	1·15	1·16
68–7	6·65	4·87	1·76	6·87	1·79	1·77	1·76	1·78
69–8	6·42	5·64	0·77	6·63	0·78	0·77	0·77	0·77
70–69	7·49	5·59	1·88	7·77	1·92	1·90	1·89	1·91
71–0	9·33	8·02	1·29	9·78	1·31	1·30	1·30	1·30
72–1	11·07	6·84	4·20	11·71	4·38	4·30	4·20	4·30
73–2	13·07	8·95	4·10	13·97	4·25	4·18	4·08	4·17
74–3	14·93	14·22	0·70	16·10	0·70	0·69	0·69	0·70

3. THE DISTANCE FUNCTION APPLIED TO OPTIMAL TAX THEORY

One of the most celebrated results in optimal tax theory is the formula called the Ramsey rule. This applies to the situation where a government, in an economy where there is effectively only one consumer, is unable to levy lump sum taxes or a non-linear income tax, and wishes to raise a predetermined revenue by ad-valorem taxes while creating as little distortion as possible. We shall treat leisure as good zero (q_0) with price p_0 (= wage rate) and, since demand functions are homogeneous so that one tax can always be zero, we assume leisure is untaxed. We also assume that all income is earned (or that unearned income is taxed at 100 per cent without meeting the government's revenue requirement) so that the consumer's budget constraint is

$$\textstyle\sum_0^n p_k q_k = p_0 T, \qquad \ldots(43)$$

where T is the consumer's time endowment. The relevant cost function is now the " full " cost function $c(u, p_0, p)$ which takes the value $p_0 T$. If we let the government's revenue requirement be $R = \rho p_0 T$, say, then following Mirrlees (1976), the government's problem is

$$\max u \text{ subject to } c(u, p_0, p) = p_0 T \text{ and } \textstyle\sum_1^n t_k q_k = \rho p_0 T. \qquad \ldots(44)$$

The Lagrangean is

$$\phi = u + \lambda\{p_0 T_0 - c(u, p_0, p)\} + \xi\{\rho p_0 T - \textstyle\sum t_k q_k\}, \qquad \ldots(45)$$

which is to be maximized with respect to u and t under the assumption that producer prices, $p-t$, are constant.

The first-order condition with respect to p_i gives immediately the Ramsey rule

$$\frac{\sum_{k=1}^{n} s_{ik}t_k}{q_i} = \text{constant}\left(= -\frac{\lambda+\xi}{\xi}\right) \qquad \ldots(46)$$

i.e. if the taxes are small, the distortion introduced by imposing them should be equiproportionate, or, more exactly, that at the optimum, ignoring income effects, a small intensification of the tax system should lead to equiproportionate reduction in quantities consumed.

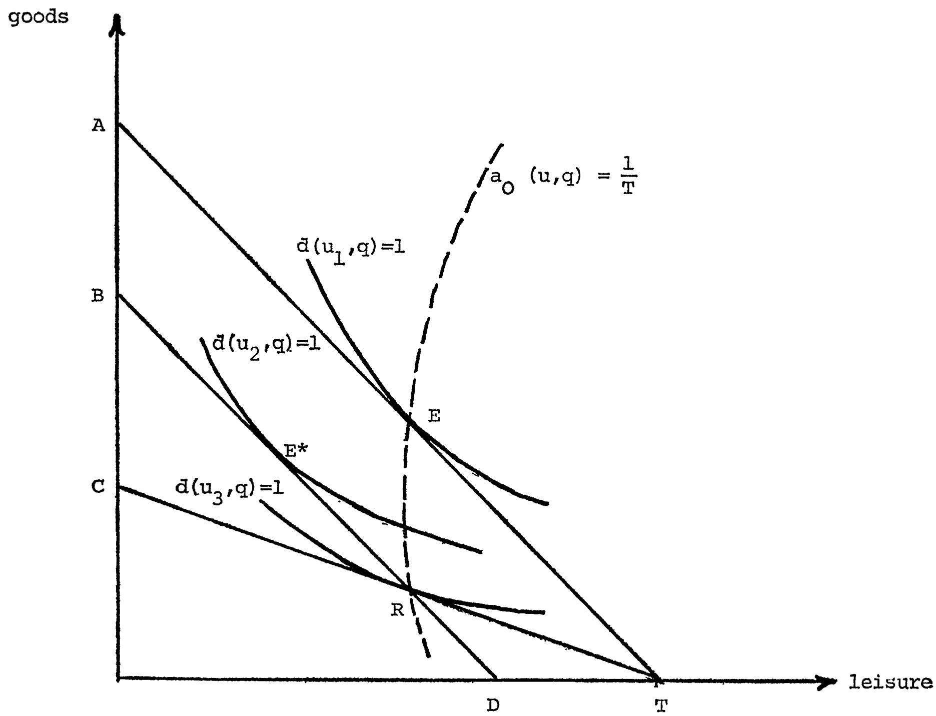

FIGURE 5

The use of the cost function to formulate the problem as (44) produces a very rapid and simple way of deriving the Ramsey rule (46). However, this simplicity is really due to the mathematical properties of the cost function rather than to the fact that either (44) or (46) are the natural ways to formulate and solve the problem. In particular, (46) characterizes the tax indirectly, in terms of its consequences for quantities, rather than giving an explicit, direct result. This is because the original problem is formulated in price space leading to a solution in quantity space. If we turn to the distance function we can reformulate the original problem in quantity space and, with equal simplicity, derive a solution in price space.

Figure 5 illustrates the tax problem in quantity space. AT is the budget line with no tax revenue; equilibrium at E on u_1 is reached. With a revenue requirement, the feasible set is BD with first best optimum E* on u_2. Without the possibility of lump sum transfers the government is restricted to the offer curve, here characterized by

$$a_0(u, q) \equiv \partial d/\partial q_0 = 1/T,$$

i.e. the price of leisure, p_0, divided by full income, p_0T, must be on the compensated inverse demand curve. Second-best equilibrium is thus at R on u_3.

Formally, the government must maximize u by choosing quantities of goods and leisure subject to the offer curve and to its own revenue requirements:

Maximize u, subject to ...(47)

(i) $d(u, q) = 1$

(ii) $\partial d(u, q)/\partial q_0 = T^{-1}$

(iii) $1 - q_0 \partial d/\partial q_0 - \sum_1^n z_k q_k = \rho$,

where z_k is the producer price of good k as a proportion of p_0T. According to 47(iii), the government chooses the q's so that the consequent prices are such as to yield the necessary revenue. It proves convenient to take u and $q_1, \ldots, q_n$ as instruments, allowing (47)(ii) to determine q_0 as a function of these, i.e. $q_0 = \zeta(u, q)$. Differentiating (47)(ii) and using the implicit function theorem gives

$$\frac{\partial \zeta}{\partial q_i} = -\frac{a_{0i}}{a_{00}} \qquad \ldots(48)$$

where, as before, a_{ij} is the i, jth element of the Antonelli matrix. The Lagrangean is

$$L = u + \lambda\{1 - d(u, \zeta, q)\} + \xi\{\rho - 1 + q_0 a_0(u, \zeta, q) + z.q\} \qquad \ldots(49)$$

Differentiating with respect to q_i holding u constant, using (10) and rearranging, gives

$$\frac{t_i}{p_i} = \left(\frac{\xi - \lambda}{\xi}\right)\left(1 - \frac{a_{0i}a_0}{a_{00}a_i}\right) \qquad \ldots(50)$$

where t_i is the tax on good i so that the LHS of (50) is the tax rate. Applying the revenue constraint gives at once

$$\frac{t_i}{p_i} = \rho\left(1 + \frac{\eta_{i0}}{\eta_{00}}\right) \qquad \ldots(51)$$

where $\eta_{i0} \equiv \partial \log r_i/\partial \log q_0$ and $\eta_{00} = -\partial \log r_0/\partial \log q_0$ are *compensated* inverse price elasticities, or more appropriately, compensated flexibilities.

This extremely elegant form is, as we shall check in the Appendix, precisely equivalent to the Ramsey rule (46). For example, when labour supply is inelastic, η_{00} is infinite so that all taxes are ρ. Similarly if good i is in inelastic demand so that η_{i0} is infinite, then that good attracts all the tax. In those two cases, there is no distortion. More generally, of course, distortion is inevitable, and since, on average, η_{i0} is positive, while η_{00} is always positive, the average tax rate must be above ρ. Equation (51) shows very clearly through the two flexibilities how it is that substitution gives rise to this distortion. We can also see that individual tax rates will deviate from ρ according to the substitutability or complementarity with leisure of the good concerned. The distortion which the revenue requirement induces takes the form of a fall in labour supply so that the tax system should attempt to remedy this by taxing relatively heavily those goods which are ($q-$) complementary with leisure. This distinction is even clearer if we write (50) in the form

$$\frac{t_i}{p_i} - \frac{t_j}{p_j} \propto \frac{\partial \log (a_i/a_j)}{\partial q_0} \qquad \ldots(52)$$

The RHS of (52) is the effect of a change in leisure on the MRS between goods i and j *along an indifference surface* and is a measure of the relative complementarity between i and j and leisure. (This formula is very close to but different from the optimal tax formula derived by Atkinson and Stiglitz (1976) in the context of non-linear taxation in a many person economy.)

Equation (52) gives us very simply the conditions for uniform commodity taxes in those cases where the Ramsey rule holds. The RHS of (52) is zero if, and only if, the distance function takes the form

$$d(u, q_0, q) = D(u, q_0, d_1(u, q)) \qquad \ldots(53)$$

which implies an identical structure for the cost-function, see Gorman (1976),

$$c(u, p_0, p) = C(u, p_0, c_1(u, p)). \qquad \ldots(54)$$

This functional structure is known as *implicit* or *quasi*-separability between leisure and goods, see Gorman (1970), Blackorby and Russell (1976). This condition for uniform taxation has been derived in the form of equation (54) by Simmons (1974). Leisure is singled out because the offer curve in Figure 5 is the main determinant of the possibilities open to the government. Neither (53) nor (54) are particularly restrictive and although obviously precluding close substitutability or complementarity between individual goods and leisure, they do not impose any clearly objectionable empirical requirements. This is not the case if they are combined with *weak* separability between goods and leisure. Weak and implicit separability are only compatible in the separable homothetic case when total expenditure elasticities are unity. However, such stringent conditions are not required for uniform taxes under the Ramsey rule.

This analysis is merely illustrative of the use of the distance function in optimal tax theory. We have made no attempt to discuss the more interesting cases where there are many consumers and equity issues are taken seriously. These problems are left for another paper.

4. CONCLUSIONS

I believe that the examples given above demonstrate that the distance function has important uses in economics but they do not begin to exhaust its potential. Just as the cost function can be used in dozens of areas of economic analysis, see particularly Gorman (1976), so can its dual. Armed with both concepts we can choose between them, not on the basis of mathematical convenience, but according to which is better suited to the *economics* of the problem at hand.

APPENDIX

DIRECT PROOF OF THE EQUIVALENCE OF THE TWO TAX RULES

Starting from the conventional Ramsey rule (46), since $t_0 = 0$, (46) can be rewritten, $i = 0, \ldots, n$

$$\textstyle\sum_{k=0}^{n} s_{ik} t_k = \alpha q_i + \beta \delta_{i0} \qquad \ldots(\text{A}.1)$$

where α and β are independent of i and δ_{ij} is the Kronecker delta. But, from (21),

$$x \textstyle\sum_i a_{ji} s_{ik} = \delta_{jk} - r_j q_k. \qquad \ldots(\text{A}.2)$$

Hence, since $\sum_i a_{ji} q_i = 0$ by (12)

$$x^{-1} \textstyle\sum_k (\delta_{jk} - r_j q_k) t_k = \beta a_{j0}. \qquad \ldots(\text{A}.3)$$

But $x^{-1} \sum t_k q_k = \rho$ and since $t_0 = 0$, $\beta = -a_0 \rho / a_{00}$, so that, finally

$$\frac{t_j}{p_j} = \rho \left\{ 1 - \frac{a_{j0} a_0}{a_j a_{00}} \right\} \qquad \ldots(\text{A}.4)$$

as before. Similarly, using $xSA = I - qr'$, we can proceed from (A.4) to (A.1).

First version received May 1977; *final version accepted July* 1978 (*Eds.*).

I should like to thank Charles Blackorby, Erwin Diewert, Terence Gorman, Frank Hahn, Michio Morishima, John Muellbauer, Agnar Sandmo and Denis Sargan for helpful comments. The sight of two

unpublished papers by Camille Bronsard and his collaborators (1976*a*) and (1976*b*) was the initial impetus to my working on this topic.

REFERENCES

ATKINSON, A. B. (1970), " On the measurement of inequality ", *Journal of Economic Theory*, **2**, 244-263.

ATKINSON, A. B. and STIGLITZ, J. E. (1976), " The design of tax structure: direct versus indirect taxation ", *Journal of Public Economics*, **6**, 55-75.

BLACKORBY, C. and DONALDSON, D. (1976), " Measures of equality and their meaning in terms of social welfare " (Department of Economics, University of British Columbia, Discussion paper 76-20, mimeo).

BLACKORBY, C., LOVELL, C. A. K. and THURSBY, M. C. (1976), " Extended Hicks neutral technical change ", *Economic Journal*, **86**, 845-852.

BLACKORBY, C. and RUSSELL, R. R. (1975), " Implicit separability, cost-of-living subindices, and flexible functional forms " (Discussion paper no. 75-08, University of British Columbia, Department of Economics, presented at the 3rd World Congress of the Econometric Society, Toronto, August 1975).

BRONSARD, C., BRONSARD, L. S. and DELISLE, D. (1976*a*), " Computing optimal tolls in a money economy " (Presented to IEA conference on Econometric Contributions to Public Policy, Urbino, Italy).

BRONSARD, L. S., LEBLANC, D. and BRONSARD, C. (1976*b*), " Estimating demand functions, the converse approach " (Montreal, mimeo).

DEBREU, G. (1951), " The coefficient of resource utilization ", *Econometrica*, **19**, 273-292.

DIEWERT, W. E. (1974), " Applications of duality theory ", Chapter 3 in Intriligator, M. D. and Kendrick, D. A. (eds.) *Frontiers of Quantitative Economics, Vol. II* (North-Holland/American Elsevier).

DIEWERT, W. E. (1976*a*), " Exact and superlative index numbers ", *Journal of Econometrics*, **4**, 115-145.

DIEWERT, W. E. (1976*b*), " Harberger's welfare indicator and revealed preference theory ", *American Economic Review*, **66**, 143-152.

FUSS, M. and McFADDEN, D. (eds.) (1978) *Production Economics: A Dual Approach to Theory and Applications* (North-Holland).

GORMAN, W. M. (1970), " Quasi-separable preferences, costs, and technologies " (University of North Carolina, Chapel Hill, mimeo).

GORMAN, W. M. (1976), " Tricks with utility functions ", in Artis, M. and Nobay, R. (eds.) *Essays in Economic Analysis* (CUP).

HANOCH, G. (1975), " Production and demand models with direct or indirect implicit additivity ", *Econometrica*, **43**, 395-420.

HICKS, J. R. (1956) *A Revision of Demand Theory* (Oxford University Press).

KONÜS, A. A. (1924), " The problem of the true index of the cost-of-living " (in Russian), *The Economic Bulletin of the Institute of Economic Conjuncture*, Moscow, **9-10**, 64-71 (English translation: *Econometrica* (1939), **7**, 10-29).

KONÜS, A. A. (1939), " On the Theory of Means ", *Acta Universitatis Asiae Mediae* (*Tashkent*), Series Va, *Mathematica*, Fasc. 24, 3-10.

MALMQUIST, S. (1953), " Index numbers and indifference surfaces ", *Trabajos de Estadistica*, **4**, 209-241.

MIRRLEES, J. A. (1976), " Optimal tax theory: a synthesis ", *Journal of Public Economics*, **6**, 327-358.

MUELLBAUER, J. (1977), " Testing the Barten model of household composition effects and the cost of children ", *Economic Journal*, **87**, 460-487.

SAMUELSON, P. A. (1950), " The problem of integrability in utility theory ", *Economica*, **17**, 355-385.

SAMUELSON, P. A. (1974), " Complementarity—an essay on the 40th Anniversary of the Hicks–Allen revolution in demand theory ", *Journal of Economic Literature*, **12**, 1255-1289.

SHEPHARD, R. (1953) *Cost and Production Functions* (Princeton University Press).

SIMMONS, P. (1974), " A note on conditions for the optimality of proportional taxation " (mimeo).

THEIL, H. (1975), (1976) *Theory and Measurement of Consumer Demand, Vol. I, Vol. II* (North-Holland).

WOLD, H. (1943), " A synthesis of pure demand analysis, Parts I and II ", *Skandinavisk Aktuarietidskrift*, **26**, 85-144 and 220-272.