8
The costs of imperfect tax systems

J.A. Kay
St. John's College, Oxford
and London Business School
and
M.J. Keen
University of Essex

Chapter 8. *Mathematical Models in Economics*. MOL Bacharach, MAH Dempster, and JL Enos, eds. University of Oxford (1990).

8

The costs of imperfect tax systems

J.A. Kay
St. John's College, Oxford
and London Business School

and

M.J. Keen
University of Essex

Abstract

The measurement of the inefficiency of tax systems is examined within the general framework proposed by Debreu, which treats inefficiency as the fraction by which inputs could be reduced consistent with the allowd output. This measure is related to more familiar measurements of the deadweight loss from taxation. The analysis can be extended to the measurement of other forms of economic inefficiency, including production inefficiency and inequality.

1. Introduction

The inefficiency of an imperfect tax system may be seen as a measure of the degree to which the results of such a structure fall short of those which could be achieved by a scheme of optimal lump sum taxes (see Auerbach (1985) for a general survey). This inefficiency is of two kinds. There is an excess burden imposed on each individual, who would prefer to pay a lump sum tax of the same amount that he or she is required to pay in distorting taxes. And there is a social loss which occurs if optimal lump sum taxes cannot be employed in pursuit of distributional goals.

Constructions which integrate these two components of inefficiency have been suggested by, for example, Diewert (1985), Jorgenson and Slesnick (1984, 1985, 1986) and King (1983). The purpose of this paper is to demonstrate a simple general framework that facilitates both this integration and its subsequent decomposition. It differs from these earlier approaches by building on the input-based approach to inefficiency proposed by Debreu (1951) which we believe to be clearly more appropriate for the analysis of imperfect tax systems. This approach—which begins from an extremely general concept of inefficiency—illuminates a number of issues in the measurement of both deadweight loss and inequality, and shows the relationship between them. Its properties are explored in more detail in a companion paper (Kay and Keen 1988).

2. The measurement of inefficiency

Consider some process transforming an N-vector z of inputs into an M-vector x of outputs. A feasible pair (x, z) may be described as inefficient if it is possible to produce a larger output x from z or the same output x from a smaller input z. Under weak conditions, these definitions are equivalent, but they point to two alternative natural measures of inefficiency. One is the additional output foregone, the other is the unnecessary input utilized. In the specific context of deadweight loss, this corresponds broadly to the distinction between Hicks–Boiteux and Debreu–Allais measures drawn by Diewert (1984, 1985).

In general, there is no reason to regard either an input or output based measure as superior to the other. If we are measuring the inefficiency of a tax system, however, the output vector x is a measure of social welfare or household utility, the input vector z one of physical commodities. Thus there is a natural metric on z while there is, at best, an ordering on x.

For this reason we prefer an input-based approach. It may be noted that the metric of inputs is not wholly free of ambiguity; in particular, what is the measure of endowment? Is the endowment of leisure twelve, sixteen or twenty-four hours per day?[1] In this paper, we follow the national accounts conventions and define w, the aggregate endowment of the economy, as that part of the endowment which is traded.

The general analysis of input-based measures of inefficiency follows the lines set out by Debreu (1951). Given a description of the technology relating outputs to inputs employed one can associate with any x a set $Z(x)$ consisting of all those input vectors such that (x, z) is feasible;

[1] Somewhat similar difficulties arise if resources are unused—due to unemployment, for instance. Unemployment might itself be treated as an inefficiency, and unemployed resources treated as part of the economy's endowment: alternatively, one might compute inefficiency by reference only to endowments that are actually employed. It is issues of this kind that seem to lie behind the distinction between 'utilized' and 'utilizable' resources made by Debreu (1951, p.285). But while the analysis that follows is perhaps suggestive of possibilities for measuring and interpreting the losses from unemployment, these will not be pursued here: the focus is entirely on inefficiencies in taxation.

the boundary of this input requirements set, denoted $\partial Z(x)$, is analogous to an isoquant or a Scitovsky curve. It is assumed throughout that this technology satisfies the following conditions:[2]

A1	**Convexity**	$Z(x)$ is strictly convex, closed and bounded below, $\forall x$
A2	**Free disposal**	$\{z' > z \in Z(x)\} \Rightarrow \{z' \in Z(x)\}$
A3	**Monotonicity**	$\{z \notin \partial Z(x),\ x' > x\} \Rightarrow \{z \in Z(x')\}$

and further that

A4 $\quad z > 0 \ \ \forall z \in Z(x) \ \ \forall x.$

Conditions A1—A3 are of a familiar kind and require little discussion; their economic substance in any particular application should be transparent. Convexity is essential to the following analysis,[3] though strictness is assumed only for clarity of analysis and statement. Assumption A4 ensures that the measure to be discussed is well defined; though the condition might be weakened, it considerably simplifies the analysis and seems likely to be innocuous in the most obvious applications.

Pursuing an input-based approach, a feasible pair $(x, \overline{z})$ is said to be inefficient iff $\overline{z} \notin \partial Z(x)$, and one is led to measure inefficiency as the distance from $\overline{z}$ to $\partial Z(x)$. There can be no uniquely correct metric for this purpose, but when—as in tax analysis—the inputs are literal commodities a natural procedure is to measure inefficiency as $p^R[\overline{z} - \beta]$ for some choice of reference price vector p^R and for some point β in $\partial Z(x)$ (appropriate transposition in the formation of inner products being taken as read throughout). Debreu's procedure specifies simultaneously both a reference price vector and counterfactual input vector.

Note first that by the support theorem there is associated with each $\beta \in \partial Z(x)$ an N-vector $p(\beta)$ such that

$$p(\beta)[z - \beta] > 0, \quad \forall z \in Z(x) \tag{1}$$

while A1 and A2 further ensure that $p(\beta) \gg 0$. The distance between $\overline{z}$ and β can then be measured in index form as

$$d(\overline{z}, \beta) = p(\beta)[\overline{z} - \beta]/p(\beta)\overline{z} > 0\,. \tag{2}$$

---

[2] The conventions here are: $z \geq 0$ means $z_i \geq 0\ \forall i$: $z > 0$ means $z \geq 0$ and $z \neq 0$: $z \gg 0$ means $z_i > 0\ \forall i$.

[3] Whilst non-convexity could be admitted by defining a measure of inefficiency directly in the manner of equations (5) and (18) below, the interpretation that follow the latter in §4 could not be sustained.

Following the convention of measuring the distance between a point and a set as the least of the distances between the point and elements of the set then leads to the measure of inefficiency

$$D(x, \bar{z}) = \inf_{\beta \in \partial Z(x)} \{d(\bar{z}, \beta)\}. \tag{3}$$

It will be assumed throughout that a minimum is attained in (3), so that

$$D(x, \bar{z}) = \{p(z^*)z^*/p(z^*)\bar{z}\} \tag{4}$$

for some $z^* \in \partial Z(x)$. This is the generic form of the measure developed by Debreu, and will be referred to here as the *Debreu measure of inefficiency.*

The Debreu procedure is thus to measure inefficiency by comparing the actual input with all alternative inputs relative to which the actual output is efficient, calculating in each case the proportionate difference in value between actual and counterfactual input (using the shadow prices associated with the latter) and taking as summary measure the smallest of these fractions. The measure also has a more direct interpretation (Debreu 1951):

$$D(x, \bar{z}) = 1 - \gamma \tag{5}$$

where $\gamma \in (-\infty, 1]$ is such that $z^* = \gamma\bar{z} \in \partial Z(x)$. That is, the Debreu measure is simply the largest fraction by which the actual input vector could be scaled down without rendering the actual output vector infeasible. The scalar γ is Debreu's *coefficient of resource utilization.*

3. Inefficiency in a taxed economy

The Debreu framework can readily be applied to an economy in which commodity taxes are imposed. Assume a competitive market for each of N commodities. In the taxed economy, consumer prices are q and producer prices p, so that the vector of commodity taxes is $q - p$.

There are F firms. A production plan of the f^{th} firm is denoted y_f, its production set Y_f is assumed convex and to contain the origin. Outputs are measured positively; the profit earned by f in the taxed state is thus $\pi_f = py_f$. There are no external economies, so that the aggregate production set is $Y = \Sigma_f Y_f$. Production is efficient, so that all inefficiency is the result of the misallocation of goods to households.

There are H households. The preferences of each household h are assumed strictly quasi-concave, and are equivalently characterized by an expenditure function $e_h(q, u_h)$ defined on utility levels u_h and by a direct utility function $v_h(x_h)$ defined on consumption x_h. Consumption here is defined as the traded endowment, denoted w_h, plus net trades, so avoiding the need to consider any unobserved component of the endowment; labour supply, for instance, simply enters negatively in x_h. Household h receives lump sum income of b_h from the government, and owns a claim to a proportion θ_{hf} of the profits of firm f. Assuming non-satiation, the budget constraint on h implies

$$qx_h(q, u_h) = e_h(q, u_h) = b_h + pw_h + \Sigma_f \theta_{hf} \pi_f, \quad h = 1, \ldots, H. \quad (6)$$

(There is no loss of generality here in thinking of endowments as being traded at producer prices; if they are traded at (say) consumer prices, then b in (6) is to be interpreted as including an offsetting lump sum tax). Households are assumed to be the only holders of endowments and shares, so that

$$\Sigma_h w_h = w \quad (7a)$$

$$\Sigma_h \theta_{hf} = 1, \qquad f = 1, \ldots, F. \quad (7b)$$

It will also prove convenient to define an aggregate expenditure function

$$E(q, u) = \Sigma_h e_h(q, u_h) \quad (8)$$

where u denotes the utility allocation (u_h); $E(q, u)$ gives the minimum total expenditure needed to attain the utility allocation u at consumer prices q. Note too that

$$E(q, u) = qX(q, u) \quad (9)$$

where $X(q, u)$ is an aggregate of compensated demand functions, defined in the manner of (8).

By definition, $x_h - w_h$ is the N-vector of net trades of household h. In the aggregate, households' net trades are matched by those of the production sector, so that

$$X(q, u) - w = y + g \quad (10)$$

where $y = \Sigma_f y_f$ denotes total private production in the taxed state and g denotes public production. For simplicity, the analysis below will not deal with inefficiences relating to public production (and public goods): in the counterfactual exercise of evaluating inefficiency it is to be understood that the taxed state is compared only with allocations involving the same vector of public production. Given this, no generality is lost in assuming $g = 0$. Then (10) becomes

$$w - X(q,u) = y\,. \tag{11}$$

By Walras' Law, it follows in the usual way from (6), (7) and (11) that

$$\Sigma_h b_h = (q-p)X(q,u) \tag{12}$$

so the government's budget constraint is satisfied. For the sake of brevity the framework just described has been kept as simple as possible: taxes are assumed to be the only source of distortion; prices do not vary across households or across firms and there are no quantity constraints in the taxed state; there is no foreign trade. It is not difficult to relax these restrictions.

It is also convenient to note here the following result, which will be useful in interpreting the measure of inefficiency derived below.

Lemma. *Let $\{A_1, A_2, \ldots, A_m\}$ be a finite class of sets and suppose that p supports their sum $A = \Sigma_i A_i$ at a^*, so that*

$$p[a - a^*] \geq 0 \qquad \forall a \in A \tag{13}$$

and that $a^* = \Sigma a_i^*$ with $a_i^* \in A_i$, $\forall i$. Then p also supports each of the component sets at a_i^*.

Proof. If p does not support the j^{th} component set at a_j^* then $\exists\, \beta_j \in A_j$ such that $p[\beta_j - a_j^*] < 0$. Define $\beta = \sum_{i \neq j} a_i^* + \beta_j$. Then $\beta \in A$ *but*

$$p[\beta - a^*] = p[\beta_j - a_j^*] < 0 \tag{14}$$

which contradicts (13).

Loosely speaking, the substance of this is that a hyperplane which supports a sum of sets also supports each of the component sets. The lemma itself is essentially a variant of Proposition 7 in Malinvaud (1985, pp.106–7).

4. The welfare loss from taxation

Welfare losses result from taxation if the same level of social welfare could be attained by using less of the economy's endowment. It therefore reflects the costs of misallocation of resources in two senses: an inappropriate (inegalitarian, if the social welfare function has the relevant convexity properties) distribution of endowments and a further maldistribution from trading at distorted prices.

Define a set $\Omega(u;\Gamma)$ consisting of those aggregate consumption bundles that can be allocated across households to yield a level of social welfare at least equal to that of the taxed state: $\Omega(u;\Gamma) = \{x | \exists\{x_h\}$ such that (i) $\Sigma_h x_h \leq x$ and (ii) $\Gamma(v(x)) \geq \Gamma(u)\}$. It is assumed that Γ is quasi-concave (corresponding to a mild egalitarianism),[4] continuous and strictly increasing in at least one argument, and that every v_h is strictly concave (implying a decreasing marginal utility of income).[5] This ensures that

$$W(u;\Gamma) = \Omega(u;\Gamma) - Y \tag{15}$$

satisfies the analogues of A1–A3,[6] and welfare loss may then be measured as

$$D(u,w;\Gamma) = \min_{\beta \in \partial W(u;\Gamma)} \{p(\beta)[w-\beta]/p(\beta)w\} \tag{16}$$

$$= p^*[w - w^*] \tag{17}$$

where $p^* = p(w^*)/p(w^*)w$. Then from (5) we have

$$D(u,w;\Gamma) = 1 - \rho_\Gamma \tag{18}$$

where $w^* = \rho_\Gamma w$ and $\rho_\Gamma \in (-\infty, 1]$; the scalar ρ_Γ is to be thought of as a generalized coefficient of resource utilization.

---

[4] See for instance Sen (1973).

[5] Differentiating the familiar necessary conditions for the consumer's optimization problem with respect to income m gives (assuming an interior solution)

$$v_{xx}(x)x_m(q,m) = \lambda_m(q,m)q \tag{a1}$$

where v_{xx} is the Hessian matrix of the direct utility function, $x_m(q,m)$ is the vector of income effects and $\lambda_m(q,m)$ is the derivative of the marginal utility of income with respect to income. Premultiplying (a1) by $x_m(q,m)$ and using the adding up condition $x_m(q,m)q = 1$, it follows that $\lambda_m(q,m) < 0$ if v_{xx} is negative definite.

[6] Somewhat more precisely, the compact set

$$\widehat{W}(u) = \{\widehat{w} \in W(u) \text{ and } \widehat{w} \leq w\}$$

satisfies appropriately redefined analogues of A1–A3 (and the condition of the lemma); see for instance corollary 4.3 of Arrow and Hahn (1971).

This measure can be interpreted in terms of more familiar concepts. By the lemma of §3, p^* supports Y at some y^*; it is also straightforward to show that there exists some welfare-equivalent allocation u^* such that $\Gamma(u^*) = \Gamma(u)$ and p^* supports $\Omega_h(u_h^*)$ for all h, where $\Omega_h(u_h) = \{x_h | v_h(x_h) \geq u_h\}$. Thus

$$w^* = X(p^*, u^*) - y^* \tag{19}$$

and so, recalling (11) and (17),

$$D(u, w; \Gamma) = p^* X(q, u) - E(p^*, u^*) + p^*[y^* - y]. \tag{20}$$

This may be rewritten as

$$\begin{aligned} &D(u,w;\Gamma) \\ &\quad = \{E(q,u) - E(p^*,u)\} - (q-p)X(q,u) \\ &\qquad + \{\pi^* + p^*w - [\pi + pw]\} + \{E(p^*,u)\} - E(p^*,u^*)\} \end{aligned} \tag{21}$$

where $\pi^* = p^* y^*$.

Consider each component of this total measure D in turn. The first is the amount which consumers would be willing to pay to have distorting taxes abolished and be permitted to trade at the producer prices p^* which would prevail under a lump sum tax system. The second term is the revenue which is in practice raised by the imperfect taxes.[7] Thus the difference between these two terms is a generalization of the deadweight loss measure proposed by Kay (1980) for the case of the single consumer economy with constant producer prices.

The third term—which is ambiguous in sign—is the difference between the level of aggregate profits π and the value of the endowment pw evaluated using producer prices of the taxed and untaxed states. An inegalitarian distribution of income might increase the demand for caviar, for example, which might be welfare reducing if there were severe diseconomies of scale in caviar production and/or the economy's endowment of caviar were limited: a high tax on caviar might be welfare enhancing for the same reasons. If producer prices are constant, this third term disappears.

[7] This is not quite accurate if tax is raised on net purchases $x_h - w_h$ rather than on consumption x_h. But the discussion is easily amended by offsetting adjustments to both the first and second terms in (21).

The final term in (21) is the sum of households' equivalent gains (evaluated at reference prices p^*) in moving from the utility allocation of the taxed state to that of the welfare-equivalent allocation. Since

$$E(p^*, u) - E(p^*, u^*) = p^*\{X(p^*, u) - X(p^*, u^*)\} \qquad (22)$$

and p^* supports $\Omega(u, \Gamma)$ at $X(p^*, u^*)$ while $X(p^*, u) \in \Omega(u; \Gamma)$, this term is nonnegative. Intuitively, it is a measure of the waste implied by an excessively costly allocation of utilities in the taxed state. For by its construction $E(p^*, u^*)$ is the minimum total expenditure needed to attain social welfare of $\Gamma(u)$,[8] and in this sense the additional expenditure (at constant prices) associated with the actual utility allocation is socially unproductive.

It is tempting to interpret this second component of distributional inefficiency as a descriptive measure of the extent of inequality in the taxed state. The difficulty here is that if preferences differ across households then maximization of the most well-behaved social welfare function need not imply perfect equality (even when, as here, the relevant utility possibility set is convex). Though utilities can always be cardinalized so as to ensure that a welfare maximum is also a point of perfect equality, this would be to reverse the standard procedures of welfare analysis: in an exchange economy endowed only with five star brandy and cheap cigarettes, for instance, one might reasonably expect the non-smoking drinker to fare rather better at a welfare maximum than the teetotal smoker. If however $v_h = v$ for all h and Γ is symmetric, then $u^* = (\overline{u}, \overline{u}, \ldots, \overline{u})$, where $\overline{u}$ is the equally distributed equivalent of utility defined by $\Gamma(u^*) = \Gamma(\overline{u}, \overline{u}, \ldots, \overline{u})$, In this case $E(p^*, u) - E(p^*, u^*)$ is indeed just a money metric measure of inequality (see, for instance, Deaton and Muellbauer 1980). Translated into index form,

$$\{E(p^*, u) - E(p^*, u^*)\}/E(p^*, u) = A(p^*, u; \Gamma) \qquad (23)$$

where

$$A(p^*, u; \Gamma) = 1 - \frac{e(p^*, \overline{u})}{(1/H)\Sigma_h\, e(p^*, u_h)} \qquad (24)$$

is an index of inequality in the manner of Atkinson (1970), though note that unless preferences are homothetic $A(p^*, u; \Gamma)$ will generally differ

[8] Thus $E(p^*, u^*)$ gives the value of the social expenditure function (in the sense of Pollak 1981) evaluated at prices p^* and social welfare $\Gamma(u)$.

from the Atkinson index evaluated at prices of the taxed state (reflecting the possibility of price-dependence discussed in Roberts 1980). The appearance here of an Atkinson index points to an underlying similarity between the methodologies of Atkinson (1970) and Debreu (1951): in each case the strategy is to measure inefficiency in terms of wasted input, and indeed the original Atkinson index can be derived as the Debreu measure of inefficiency for a process transforming a given total of lump sum incomes into a level of social welfare (Kay and Keen 1988).

5. Conclusions

The Debreu framework is a powerful tool for illuminating contested issues in the measurement of welfare loss. The general principles outlined can be readily applied in contexts where output-based measures seem more appropriate although, in the context of imperfect taxation, the input-based approach of Debreu appears to us the more relevant one. Our analysis has assumed production efficiency but it is clear that the general approach can be readily applied to inefficiency in production (and, indeed, this was the procedure adopted by Farrell (1957)) and to the many other sources of inefficiency in real economies: misconceptions by economic agents, missing markets, untradeable commodities, or monopoly power.

References

Arrow, K. J., and Hahn, F. H. (1971). *General competitive analysis.* Holden Day, San Francisco.

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory*, **2**, 244–63.

Auerbach, A. J. (1985). The theory of excess burden and optimal taxation. In *Handbook of public economics*, Vol. I, (ed. A. J. Auerbach and M. S. Feldstein). North Holland, Amsterdam.

Deaton, A. S., and Muellbauer, J. (1980). *Economics and consumer behaviour.* Cambridge University Press.

Debreu, G. (1951). The coefficient of resource utilisations. *Econometrica*, **19**, 273–92.

Diewert, W. E. (1984). The measurement of deadweight loss in an open economy. *Econometrica*, **51**, 23–42.

Diewert, W. E. (1985). The measurement of waste and welfare in applied general equilibrium models. In *New developments in applied general equilibrium analysis*, (ed. J. Piggott and J. Whalley). Cambridge University Press.

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society, Series A*, **120**, 253–78.

Jorgenson, D. W., and Slesnick, D. T. (1984). Aggregate consumer behaviour and the measurement of inequality. *Review of Economic Studies*, **51**, 369–92.

Jorgenson, D. W., and Slesnick, D. T. (1985). General equilibrium analysis of economic policy. In *New developments in applied general equilibrium analysis*, (ed. J. Piggott and J. Whalley). Cambridge University Press.

Jorgenson, D. W., and Slesnick, D. T. (1986). General equilibrium analysis of natural gas price deregulation. Discussion paper 1215. Harvard Institute of Economic Research.

Kay, J. A. (1980). The deadweight loss from a tax system. *Journal of Public Economics*, **13**, 111–20.

Kay, J. A., and Keen, M. J. (1988). Measuring the inefficiences of tax systems. *Journal of Public Economics*, **35**, 265–87.

King, M. A. (1983). Welfare analysis of tax reforms using household data. *Journal of Public Economics*, **21**, 183–214.

Malinvaud, E. (1985). *Lectures on microeconomic theory*, (revd. edn). North Holland, Amsterdam.

Pollak, R. A. (1981). The social cost of living index. *Journal of Public Economics*, **15**, 311–36.

Roberts, K. W. S. (1980). Price-independent welfare prescriptions. *Journal of Public Economics*, **13**, 277–97.

Sen, A. K. (1973). *On economic inequality*. Oxford University Press.