

NBER WORKING PAPER SERIES

WILL THE REAL EXCESS BURDEN
PLEASE STAND UP?
(OR, SEVEN MEASURES IN SEARCH OF A CONCEPT)

Alan J. Auerbach

Harvey S. Rosen

Working Paper No. 495

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge MA 02138

June 1980

The research reported here is part of the NBER's research program in Taxation. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

Will the Real Excess Burden Please Stand Up?
(or, Seven Measures in Search of a Concept)

ABSTRACT

It is well understood that a tax which distorts relative prices generates a welfare cost or "excess burden" in addition to any associated transfer of resources, but there remains considerable controversy and confusion with respect to procedures for measuring this excess burden.

The purpose of this paper is to clarify matters concerning what is one of the most basic concepts in welfare economics. We describe and evaluate a number of alternative conceptual experiments which might lie behind an excess burden calculation, showing how these notions can be represented graphically and algebraically and how they can be approximated numerically.

Harvey S. Rosen
Department of Economics
Princeton University
Princeton, N.J. 08544

National Bureau of Economic Research
1050 Massachusetts Avenue
Cambridge, MA 02138

(609) 452-4022

Alan J. Auerbach
Economics Department
Harvard University
Cambridge, MA 22138

National Bureau of Economic Research
1050 Massachusetts Avenue
Cambridge, MA 02138

(617) 868-3927

I. Introduction

It is well understood that any tax (or subsidy) that distorts relative prices away from a pareto optimum generates a welfare cost or "excess burden" in addition to any associated transfer of resources. With respect to procedures for measuring excess burden, however, there is considerable controversy and confusion. We think this is attributable to two main causes. First of all, various investigators have had different conceptual experiments in mind when computing excess burden. Second, given the choice of conceptual experiment, methods used to represent and numerically approximate excess burden can vary, and few authors have made their choices explicit.

The purpose of this paper is to clarify matters concerning the theory and measurement of what is one of the most basic concepts in welfare economics. In the following section, we describe a number of alternative conceptual experiments which might lie behind an excess burden calculation, how these notions can be represented graphically and algebraically, and appropriate methods for approximating them numerically. In Section III we evaluate the various measures on the basis of several criteria. Section IV contains explicit numerical simulations to illustrate some of the issues raised in Section III. We conclude the paper with a summary and suggestions for future research.

II. Alternative Measures of Excess Burden

Suppose that the government levies a proportional tax on a commodity. The excess burden of the tax equals the loss in welfare due purely to the

tax-induced distortion in relative prices. Though this is the standard definition, it begs two important questions. First, there is no indication of how this loss in welfare ("utils") is to be converted into a dollar measure. Second, though this loss may depend on the disposition of tax receipts, the definition fails to specify how they are used. Although various resolutions of these ambiguities have been suggested, it is not always clear how their differences affect the excess burden measure. This section shows how various concepts of excess burden arise, and how these measures can be represented and approximated.

A. Consumers' Surplus Measures

The simplest representation of any welfare change resulting from a change in price is Marshall's notion of consumers' surplus. Consider an individual with endowment y_0 who consumes some quantity of a commodity x when its price is p_0 . Figure 1 depicts the ordinary demand curve for x , $D(y_0)$, where the presence of y_0 in parentheses indicates that money income is held constant as price varies. Suppose first that, for some reason, the price of x rises to p_τ , with no change in income or other prices. Consumer demand drops to x_τ , and the loss in consumers' surplus is defined by the area $p_\tau ACp_0$. The intuitive idea here is that the height of the demand curve represents a willingness to pay for each successive unit of x , so that the integral over x of the difference between this amount and the actual price represents a rent to the consumer.

Now, suppose further that this price increase from p_0 to p_τ is due entirely to the imposition of a tax; that producer prices are fixed and profits are zero.¹ Under these circumstances, government collects revenue

equal to the area $p_{\tau}ABp_0$ which, by assumption, is disposed of in a way which has no repercussions on the demand for x . Thus, the loss in consumers' surplus exceeds collected revenue by the area ABC , which we may view as the excess burden due to the tax.

As drawn in Figure 1, the demand curve $D(y_0)$ is not linear. Typically, investigators use some second-order Taylor approximation in calculating the area ABC , which has the effect of assuming the demand curve to be linear and therefore makes the measured area a triangle. One such approximation is represented in Figure 1 by the area $A'B'C$, which is obtained by taking the excess burden to be a function of price and expanding around the initial price p_0 . This area equals $-\frac{1}{2} \frac{\partial x(y_0, p)}{p} \Big|_{p=p_0} \cdot (\Delta p)^2$, where $x(\cdot)$ is the ordinary

demand function and $\Delta p = p_{\tau} - p_0$. One could just as easily have expanded around the final price p_{τ} , which would yield the area ABC' , or around one of the quantities x_0 and x_{τ} . This is entirely an arbitrary decision, as is that to use a second-order rather than a higher-order approximation.² In this paper, we will generally follow the common convention of using a second-order expansion around the initial point suggested by the conceptual experiment.

Figure 2 presents the general equilibrium analogue of Figure 1 for the case where there are two goods, consumption x and leisure ℓ , and the consumer has an initial endowment y_0 of labor, which we take as numeraire. The slope of the initial budget line is $-\frac{1}{p_0}$.³ Point c represents the initial consumption point and corresponds to point C in Figure 1. At this initial equilibrium, the consumer attains a utility level U_{00} , as shown by the tangency of the initial budget line and the indifference curve labelled u_{00} . The rise in price to p_{τ} is depicted by a flattening of the budget line, with

utility dropping to $U_{0\tau}$ at point a, corresponding to point A in Figure 1. As the budget line pivots downward, the purchases of x trace out the demand curve $D(y_0)$ between p_0 and p_τ . The tax revenue collected, corresponding to the area $p_\tau ABp_0$ in Figure 1, equals the horizontal distance between the two budget lines starting at point a.

Although the consumer's surplus measure is used fairly often⁴ to represent the extent to which the utility loss between U_{00} and $U_{0\tau}$ exceeds in money terms the revenue collected, it suffers from a number of problems, the most important of which is the way it converts utils to dollars. As is well-known, consumers' surplus is a consistent measure of any price-induced welfare change only when there are no income effects in the demand for the affected good(s). In general, if more than one price changes, the measure of the total welfare effect is not single-valued, but depends on the path of the price changes. For example, if two prices increase, one will find two different measures of the induced welfare loss depending on which price is assumed to increase "first." Related to this problem of path-dependence is the result that if several prices change, leaving utility unaffected, the consumers' surplus change taken by adding over affected markets in any particular order is unlikely to be zero.

In their efforts to "rehabilitate" consumers' surplus, Hicks (1946) and, more recently, Willig (1976) have discussed the potential error involved in using consumers' surplus as a measure of welfare loss. Willig has shown that in many cases, the difference between consumers' surplus and other measures which do not suffer from the above shortcomings is very small relative to the size of the welfare loss itself. However, these results relate to the accuracy of the area $p_\tau ACp_0$ as a measure of total welfare loss, not to the accuracy of area ABC as a measure of the excess burden. Indeed, because

excess burden is a second-order effect, it may well be true that consumers' surplus leads to a very poor measure of excess burden.⁵ We will elaborate upon this point once we have introduced the other methods for measuring excess burden.

B. Income Compensation Measures

As just noted, in the consumers' surplus measure of excess burden, taxes are assumed spent in such a way that there are no repercussions on the demand for the taxed goods. Harberger (1974), reasoning along lines suggested by Hotelling (1938), offers a different conceptual experiment:

When a new tax or set of taxes is imposed, we conceive of it as being counterbalanced by a pattern of lump-sum subsidies (and possibly lump-sum taxes) which keeps the relative distribution of income unchanged. (p. 22)

For our one-consumer, two-good economy, this experiment is illustrated in Figure 3, where the initial equilibrium of point c is again depicted. Here, as the tax is levied, causing the slope of the budget line to decrease, revenues are simultaneously rebated. The consumer takes these rebates to be unrelated to purchases of x. Thus, the equilibrium must be characterized by two conditions:

- (i) the marginal rate of substitution equals the post-tax price ratio p_T ; and
- (ii) the optimizing bundle, shown at point d, lies on the initial budget line through point c.

In effect, then, the equilibrium is found by taking a budget line whose slope is given by post-tax prices, and 'sliding' it down the original budget line

until a tangency with an indifference curve occurs at the intersection of the two lines, as is true at d. Note that at this point, the individual's endowment y_2 is implicitly defined by the expression

$$y_2 = y_0 + \Delta p \cdot x(y_2, p_\tau) . \quad (1)$$

Following our earlier convention, we denote the utility attainable at endowment y_2 and price p_τ as $U_{2\tau}$. Since $U_{00} > U_{2\tau}$, and the government keeps no revenue, there is an excess burden, measured in utility terms by $(U_{00} - U_{2\tau})$. In his writings, Harberger devotes his effort to developing and explaining a second-order Taylor approximation, which is pictured in Figure 4. If we let $D^h(y_0)$ represent the demand curve generated when the government rebates all revenues as p rises, then point D in Figure 4, the point on this demand curve at $p = p_\tau$, corresponds to point d in Figure 3. (Note that, by construction, the regular demand curve $D(y_2)$ would also pass through point D.) The drawing of this income-compensated demand curve as steeper than the ordinary one corresponds to the assumption that x is a normal good.

Harberger measures excess burden by the area DEC, and, following Hotelling, suggests a Taylor approximation around the initial price of $-\frac{1}{2} \frac{\partial x^h(y_0, p)}{\partial p} \Big|_{p_0} \cdot (\Delta p)$ where $x^h(\cdot)$ is the income-compensated demand for x . This approximation yields the familiar "Harberger triangle" D'E'C as a measure of excess burden. However, as we shall see later, the conceptual experiment depicted in Figure 3 lends itself to other interpretations which lead to alternative measures. Unfortunately, Harberger's method suffers from the same difficulties as the simpler consumers' surplus approach. Once again, by taking an area under a demand curve where different points correspond to different levels of utility,

path-dependence and related problems are encountered. These problems are, perhaps, "less bad" here because the utility changes are smaller along D^h than along D , but in general one cannot obtain a quantitative measure of the relative errors.

C. Utility Compensation Measures

Building upon the work of Mohring (1971), Diamond and McFadden (1974) have proposed a definition of excess burden that differs from Harberger's:

...We shall define the deadweight burden, or loss, as the excess of the income we give a consumer to restore him to his pre-tax indifference curve over the tax revenue collected from him... While it is not clear that this is the most intuitive notion, for consistency we measure the tax revenue for this definition as the level collected at the consumer equilibrium after the consumer has been restored to his original indifference curve. (p. 5)

Note that the compensation process is defined in terms of utility, not income as in the Harberger experiment.

Figure 5 presents a general equilibrium representation of the Diamond-McFadden experiment. Again, the initial budget line has slope $-\frac{1}{p_0}$ and passes through c . Its intercept on the horizontal axis is y_0 , the consumer's labor endowment. The new budget line faced by the consumer is one which allows attainment of the initial utility level, at point f , and has slope $-\frac{1}{p_\tau}$. By construction, this new budget line must intersect the horizontal axis at the endowment necessary to attain utility level U_{00} at price p_τ . A convenient algebraic representation of this distance is given by the expenditure function⁶

$$E(U_{00}, p_\tau) = E_{00}^\tau \cdot \quad (\text{We use the convention here that } E_i^j = E(U_i, p_j).)$$

Within the framework, the excess burden is measured by the extent to which tax revenue raised in the new state falls short of the increment in endowment, $(E_{00}^T - y_0)$, needed to attain U_{00} . It is depicted in Figure 5 by the distance between y_0 and the horizontal intercept, z_0 , of the line which has slope $-\frac{1}{p_0}$ and passes through the new equilibrium point f . Since $y_0 = E(U_{00}, p_0)$, we may express this distance as

$$E(U_{00}, p_\tau) - E(U_{00}, p_0) - \Delta p \cdot x^c(U_{00}, p_\tau). \quad (2)$$

where $x^c(\cdot)$ is the utility compensated demand for x .

To interpret this conceptual experiment in terms of the area beneath a demand curve, note that by definition the demand for x along u_{00} as p increases from p_0 to p_τ is $x^c(U_{00}, p)$, represented by the demand curve $D^c(U_{00})$ in Figure 6. Point F corresponds to point f in Figure 5. It is not hard to see that the area $p_\tau F C p_0$ is in fact exactly equal to the increase in expenditure $(E_{00}^T - y_0)$. Since the derivative of the expenditure function $E(U, p)$ with respect to p is $x^c(U, p)$,⁷ it follows that

$$\begin{aligned} \text{area}(p_\tau F C p_0) &= \int_{p_0}^{p_\tau} x^c(U_{00}, p) dp = \int_{p_0}^{p_\tau} \frac{\partial E(U_{00}, p)}{\partial p} dp \\ &= E(U_{00}, p_\tau) - E(U_{00}, p_0) = E_{00}^T - y_0. \end{aligned} \quad (3)$$

Moreover, since the revenue collected is the area $p_\tau F G p_0$, the area of FGC equals the excess burden $(z_0 - y_0)$ shown in Figure 5.⁸ That this measure is not path-dependent is evident from the fact that its expression in (2) depends only on the initial and final equilibrium.⁹

Even though the region FGC is smaller than the "Harberger" area DEC (again, assuming that x is a normal good), its Taylor approximation around p_0 ¹⁰ is the same, since it equals $-\frac{1}{2} \frac{\partial x^c(U_{00}, p)}{\partial p} \Big|_{p_0} (\Delta p)^2$ and the slopes of

income-compensated and utility-compensated demand curves are equal at their point of intersection. Thus, once again, excess burden is approximated by the area of the triangle D'E'C. This result is, in a way, unfortunate, because it obscures the fact that the underlying measures really do differ.

The form of the utility compensation approach just outlined is described by Hicks(1946) as "compensating variation": the additional income that must be given to the consumer to compensate for the introduction of the distortion assuming all tax revenues as rebated. One can also imagine constructing a measure based on Hicks' "equivalent variation": the amount of income the consumer would be willing to relinquish to be rid of the distortion. To derive such a measure,¹¹ we imagine starting out at the post-distortion equilibrium shown by point d in Figure 7 (repeated from Figure 3), where $y = y_2$ and $p = p_\tau$, and ask how much we can lower y if p is lowered to p_0 , given that utility must remain at $U_{2\tau}$. By definition, the income required to attain utility level $U_{2\tau}$ at price p_0 is $E(U_{2\tau}, p_0)$, shown as $E_{2\tau}^0$ in Figure 7. Thus, since tax rebates drop from $(y_2 - y_0)$ to zero as p drops from p_τ to p_0 , and the equilibrium shifts to point h, the consumer is willing to give up an additional sum equal to¹²

$$E(U_{2\tau}, p_\tau) - E(U_{2\tau}, p_0) - \Delta p \cdot x^c(U_{2\tau}, p_\tau) . \quad (4)$$

It is at once evident from comparing (4) to (3) that the two utility-compensation measures of excess burden differ only in the choice of base utility level. In particular, the measure in (4) will equal the area under the compensated demand curve calculated at utility $U_{2\tau}$. This is shown by the area DEH in Figure 8, the left shaded area. Note that the utility-compensated demand curve $D^c(U_{2\tau})$ intersects the income-compensated demand curve $D^h(y_0)$ at

point D, the starting point of the conceptual experiment just described. (The corresponding Taylor approximation is $-\frac{1}{2} \frac{\partial x^c(U_{2\tau}, p)}{\partial p} \Big|_{p_\tau} (\Delta p)^2$). For purposes of comparison, we also shade the area FGC, which we showed represents the compensating variation of the distortion. Note that, as long as x is normal, each of these areas must be smaller than the area DEC, Harberger's measure of excess burden.

Though this equivalent variation measure is also path independent, it is distinct from the compensating variation approach of Diamond and McFadden. This difference may be represented as an index number problem in the choice of which prices to use in measuring a change in real income. Looking first at the compensating variation, we recall that equation (2) defines the distance between y_0 and z_0 in Figure 9. Since the original budget line through point c also passes through point d , we are also measuring the income required to move from point d to point f . These are the points on the indifference curves $U_{2\tau}$ and U_{00} which would be chosen by the consumer when $p = p_\tau$. Now consider our equivalent variation measure, which equals $(y_0 - E_{2\tau}^0)$. This also measures the income required to go from point h to point c , the points on $u_{2\tau}$ and u_{00} which would be chosen at $p = p_0$. Thus, our two measures of excess burden offer two measures of the real income difference between u_{00} and $u_{2\tau}$, one, compensating variation, at post-tax prices and the other, equivalent variation, at pre-tax prices.

We suggested earlier that the problem with Harberger's notion of excess burden lies not in his conceptual experiment, but in its implementation. As we have just demonstrated, both of the measures discussed in this section are nothing more than income measures of the utility loss caused by introducing the price distortion while refunding revenue, which is precisely Harberger's way of

looking at the problem. Unfortunately, his demand curve measure corresponds to neither of the Hicksian variations.

The choice between compensating and equivalent variation is no easier than the choice of a price index. To paraphrase a point first made by Hicks, the compensating variation of introducing a distortion equals the equivalent variation of removing it. We could ask how much the consumer must receive to compensate for a distortion being introduced, or how much he must be given in place of the distortion being removed. Thus, our choice may depend on which we consider to be the logical "initial situation." As pointed out by Hause (1975), if we are starting at some situation α , the equivalent variations of going from α to β and α to γ (where β and γ are two alternative situations) are the same if and only if the consumer's utility is the same at β and γ . The measures are comparable because we use the same set of prices, those effective at α , in the two calculations. Since the compensating variations of these two movements use different price ratios (those prevailing at β and γ , respectively) they are not exactly comparable. Thus, if α is the initial point, we may wish to use equivalent variation as a measure of the welfare change between α and β . However, if β is the initial point, the same argument suggests that we should use the equivalent variation between β and α , which is the compensating variation between α and β .

D. The Disposition of Revenues

In the previous two subsections, we have focused on measures centered around the initial, pre-tax point, labelled C in Figure 8 (and c in Figure 9). We could just as easily have started at point A, assuming the initial condition to be with the tax already imposed. For example, performing Harberger's

experiment, we could have asked how much lower utility is in the post-tax situation at A than it would have been had the same revenue been raised through lump sum taxes. This experiment is depicted in Figure 10. Point a represents the post-tax equilibrium, chosen at income y_0 and price p_τ and yielding utility $U_{0\tau}$ (see Figure 2). Point i is the equilibrium which would have occurred had the government instead collected the same amount of revenue through lump sum taxes. Here, the individual would begin with net endowment y_1 and face price p_0 . Because of the distortion, $U_{0\tau} < U_{10}$ even though government revenue is the same in the two cases. The relationship between points i and a is the same as that between c and d in Figure 3, except that the first two points lie on the post-revenue budget line through y_1 , the latter set on the initial budget line through y_0 .

As before, we can measure the utility loss ($U_{10} - U_{0\tau}$) by compensating variation or equivalent variation,¹³ measuring the distance between the indifference curves at points chosen when $p = p_\tau$ (starting at point a) or at points chosen when $p = p_0$ (starting at point i). In terms of the expenditure function, we express these measures as:

$$\begin{aligned} E(U_{10}, p_\tau) - E(U_{10}, p_0) - \Delta p x^c(U_{10}, p_\tau) & \quad \text{(compensating)} \\ E(U_{0\tau}, p_\tau) - E(U_{0\tau}, p_0) - \Delta p x^c(U_{0\tau}, p_\tau) & \quad \text{(equivalent)}. \end{aligned} \tag{5}$$

We could picture them in Figure 8 by drawing the income-compensated demand curve $D^h(y_1)$ which passes through point A, and then adding the utility-compensated demand curves $D^c(U_{0\tau})$ and $D^c(U_{10})$. The first of these two curves passes through point A, and the second passes through the point where the $D^h(y_1)$ curve intersects the p_0 price line. The corresponding Taylor approximations are

$$-\frac{1}{2} \frac{\partial x^c(U_{10}, p)}{\partial p} \Big|_{p_0} (\Delta p)^2 \text{ for the compensating variation and}$$
$$-\frac{1}{2} \frac{\partial x^c(U_{0\tau}, p)}{\partial p} \Big|_{p_\tau} (\Delta p)^2 \text{ for the equivalent variation.}$$

E. Summary

In Table 1, we collect the various measures of excess burden. (The blanks in the table are due to the fact that, as stressed above, neither simple consumers' surplus nor the area under the income-compensated demand curve lends itself to an interpretation in terms of the underlying preference structure. A natural question at this point is whether the various measures are all that different. We provide below some simulation results which illustrate how the measures and their approximations depend on the size of the distortion and the structure of consumer preferences.

III. Evaluating the Measures

The investigator trying to decide whether or not to use a given measure is likely to ask at least three questions:

- (a) Is the conceptual experiment appropriate?
- (b) Are the informational requirements necessary to implement the measure reasonable?
- (c) How serious are aggregation problems?

We consider each of these questions in turn.

A. The Conceptual Experiment

As Table 1 illustrates, the various measures differ with respect to the "initial" level of income and the set of prices used for comparing alternatives.

We pointed out above that the second of these differences is an index number problem, with the basis of choice being what the actual initial state of the economy is. The issue of which income level to use is a bit tricky. One might want to argue on a priori grounds that the pre-tax level is appropriate, since in a general equilibrium setting, if the economy as a whole is to stay on its budget constraint, the resources raised through taxation must be returned in some form. From a behavioral point of view, however, a case can be made for the post-tax measures. For many taxes, it is unclear in what form individuals reclaim the receipts, if at all. (See, e.g., Kay 1980, p. 112.) To the extent this is the case, then, measuring excess burden around post-tax income will be more appropriate

A third way in which measures differ is whether they are exact or some Taylor approximation of an exact measure. Since use of the latter must inherently introduce some error (assuming the exact measure is appropriately chosen), there must be some other argument in its favor. Typically this deals with the amount of information necessary to use it.

B. Informational Requirements

An important consideration for the applied welfare economist is how much information is required to implement each measure. In this section we consider this issue in a setting where the investigator has data on the individual's economic behavior (e.g., from a cross-section sample). In section C below we examine the problem when only aggregate demands are known.

The consumers' surplus appears to be the easiest measure to implement. All that is required is an econometric estimate of the ordinary demand curve. With that in hand, it is a simple matter to integrate underneath it, and find the relevant areas.

On the other hand, the "exact" formulations of Table 1 seem to have greater informational requirements. One must "know" the entire utility function into order to implement these measures.¹⁴ But the differences in informational requirements are more apparent than real. In order for the whole welfare economics exercise to make sense, it must be assumed that the ordinary demand curve is generated by some utility function. But if this is the case, then it is in general possible to integrate the demand function to find the utility function that generated it.¹⁵ Alternatively, one can postulate ordinary demand curves based upon a specific utility function from the coefficients of the demand equation.¹⁶

In short, the information requirements for the consumers' surplus measure and the exact measures are virtually identical. Why, then, are the consumers' surplus measures used so much more than expenditure function measures? There are several explanations. In some cases, the functional specification of the ordinary demand curve is sufficiently complicated that integration is infeasible. In other cases, consumers' surplus measures are implemented without an explicit demand function. Rather, 'consensus' values of price elasticities are culled from the literature, and these are used to approximate the excess burden triangle.¹⁷ Thus, there is no function under which to integrate. Finally, it may be that some investigators believe that their econometric estimates are only a rough approximation to the demand curve, and are not willing to view them as having been generated by a utility function.

We turn now to the second-order Taylor approximations of Table 1. All the information required to implement them is also in the ordinary demand curve. The latter allows estimation of income effects and uncompensated price effects, and the Slutsky equation shows how to combine these in order to find

approximations to the compensated price responses needed in Table I. In a sense, then, it is hard to understand why one would ever use an approximation. If the utility function itself can be inferred from the ordinary demand curve, why settle for the approximation? As before, the approximations may be preferred to their exact analogues when either: (a) The investigator has elasticity estimates but no explicit demand curve, or (b) integration of the ordinary demand curve is infeasible.

We conclude that if the ordinary demand function for an individual is available and feasible to integrate, then the expenditure function formulations can and should be implemented. In cases where only price and income effects are available, then the corresponding Taylor approximations should be used.¹⁸ Only when the sole piece of information is the ordinary price elasticity of demand is it necessary to rely upon the consumer's surplus measure.

C. Aggregation Problems

We now turn to the question of the relative merits of the various measures when information regarding behavior comes to the investigator in aggregate form. For example, much of the debate concerning the excess burden of capital income taxation is based upon elasticities from aggregate savings functions.

There are two separate problems which arise when an investigator must deal with aggregate data. First, as is well-known, aggregate demand functions will correspond to some underlying "aggregate" preference ordering or utility function only under very special conditions: in general, the Engel curves of each household must be linear and possess the same slope.¹⁹ If these conditions fail to be met, there exists no "aggregate individual" to whom we can

apply the exact measures derived above. However, even if these conditions do not obtain, it may be feasible to compute an exact utility function by applying Roy's Identity to the aggregate demand equation, either by imposing the assumption that the equation correspond to an underlying utility function, or being fortunate enough for this constraint to be satisfied by unconstrained estimates. Should this "pseudo-utility function" be used to calculate excess burden? The answer is, not necessarily. If preferences differ among households, then by applying one of the exact measures to the aggregate consumer we will misrepresent the actual loss, which should be calculated by summing the losses incurred by individual households. This leaves us in a rather unsatisfactory position. What should we calculate when we lack the disaggregated data necessary for an exact answer? There are two possibilities:

- (1) An exact measure generated by the aggregate individual's expenditure function, if the latter can be computed
- (2) A second-order approximation of (1) (If no utility function can be computed, this is the only option.)

In the next section we evaluate these options using simulations.

IV. Simulation Results

In this section, we seek to answer two questions. First, given that any of the four exact measures of excess burden seem equally preferable a priori, how different can they be? In some sense, we are asking how wide the bounds are on the excess burden being measured. Second, how good are the exact and second-order measures as estimates of the excess burden when applied to aggregate data?

Consider a simple example. Suppose that, as above, there are two commodities, goods and leisure, the former having relative price p . Suppose further that each consumer maximizes a Cobb-Douglas utility function:

$$U^i = (x^i)^{\alpha^i} (\ell^i)^{1-\alpha^i} \quad (6)$$

where x^i and ℓ^i are consumption of goods and leisure by individual i and α^i is a taste parameter. The consumer's budget constraint is:

$$px^i + \ell^i = y^i \quad (7)$$

where y^i is individual i 's endowment or "full income" (including compensating taxes or transfers where appropriate).

To solve for the various measures in Table 1, we must know the compensated demand function for x , the indirect utility function, and the expenditure function which correspond to (6) and (7). These three functions are, respectively (suppressing superscripts)

$$x^c(U, p) = \alpha^{(1-\alpha)} (1-\alpha)^{-(1-\alpha)} p^{-(1-\alpha)} U \quad (8a)$$

$$v(y, p) = \alpha^\alpha (1-\alpha)^{(1-\alpha)} p^{-\alpha} y \quad (8b)$$

$$E(U, p) = \alpha^{-\alpha} (1-\alpha)^{-(1-\alpha)} p^\alpha U \quad (8c)$$

All that remains before applying the formulas in Table 1 is to calculate y_1 and y_2 . The first is less than y_0 by the amount of revenue collected when $y = y_0$ and $p = p_\tau$; the second equals the amount of income which results when the price is raised to p_τ with all revenues rebated. Thus,

$$y_1 = y_0 - \Delta p \cdot x(y_0, p_\tau) = y_0 - \Delta p x^c(U_{0\tau}, p_\tau) \quad 20 \quad (9a)$$

$$y_2 = y_0 + \Delta p \cdot x(y_2, p_\tau) = y_0 + \Delta p x^c(U_{2\tau}, p_\tau) \quad (9b)$$

Substituting (8) into (9a) and (9b), and letting τ equal $\frac{p_\tau - p_0}{p_\tau}$, we obtain:

$$y_1 = y_0(1-\alpha\tau) \quad (10a)$$

$$y_2 = y_0/(1-\alpha\tau) \quad (10b)$$

We can now use (8) and (10) to solve explicitly for the four utility-compensated measures outlined in Table 1. The results are shown in Table 2. The four exact measures can be summarized by the expression

$$[(1-\tau)^{-\alpha}(1-\alpha\tau) - 1] (p_0/p_j)^\alpha y_1 \quad (11)$$

where i and j are the indices that correspond to the subscripts of the utility level, U_{ij} , appropriate for the particular formulation. Because $(1-\tau)^\alpha < (1-\alpha\tau)$ for all α between 0 and 1, the largest estimate of excess burden is compensating variation at pre-tax income (CV_I), the second largest is equivalent variation at pre-tax income (EV_I), with the two post-tax measures following in the same order. From largest to smallest, the measures differ by a factor of $(1-\tau)^\alpha$. For $\alpha = .5$, this means the smallest measure is about 10-1/2 percent lower than the largest if $\tau = .2$, but almost 30 % lower if $\tau = .5$. As one would expect, the bigger the price change, the more severe the index number problem. Similarly, the bigger is α , the share of the taxed good, the bigger the differential. For $\tau = .2$, if α equals .8, the differential is 16%. If α equals .2, it is reduced to 4%. In general, then, there is sensitivity among the measures to tax rates and tastes, but the bounds present for a typical case do not seem unreasonably large. Perhaps what we should gain from this exercise is the realization that even "exact" measures do differ, and substantially so if one is considering substantial tax changes.

Turning to the Taylor approximations, the four measures are summarized by the equation

$$\left(\frac{\tau}{2}\right)\alpha(1-\alpha)(p_{\tau}/p_j)^2 y_i \quad (12)$$

where i and j are, as before, the indices of the utility level appearing in the relevant expression. Here, the differences can be quite substantial, even for reasonable parameters. For example, if $\tau = .2$, the CV_I measure is more than 50% larger than the EV_{II} measure. Clearly, the linearization of demand curves makes the choice of initial point important.

We now consider the aggregation question. First note that the i^{th} individual's demand for x is

$$x^i(y^i, p) = \frac{\alpha^i y^i}{p} \quad (13)$$

However, the investigator using aggregate data estimates

$$x(\bar{y}, p) = \frac{\bar{\alpha}\bar{y}}{p} \quad (14)$$

where \bar{y} is aggregate income and $\bar{\alpha}$ is the average value of α , weighted by income. Clearly, it is easy to apply Roy's Identity to (14) and find the associated utility function. But a glance at equation (13) indicates that the income slopes of our two individuals differ. Therefore, as noted above, the use of equation (14) to compute excess burden for the aggregate individual will generate an erroneous result.

In light of that fact, an investigator faces several questions: (a) If he uses an expenditure function based on (14), how far will his exact measure of excess burden depart from the "true" one?²¹ (b) If all he has are price and income responses from the aggregate demand function, (and cannot work back

to the "aggregate individual" expenditure function), how far will measures based upon the second order approximation depart from the true one? (c) When there is a choice between (a) and (b), which should be used? Question (c) is non-trivial: Given the fact that approximation error is involved in either case, the additional error caused by the second-order approximation might be unimportant or, conceivably salutary.

In Table 3 we present simulation results for the first measure of compensating variation for the case where $\alpha^1 = .3$ and $\alpha^2 = .7$, so that $\bar{\alpha} = .5$. Reading from left to right, the columns present the tax rate τ , the correct measure of excess burden (expressed as a fraction of measured income), the measure calculated from the aggregate demand curve,²² and their respective Taylor approximations. For small tax rates, the aggregation error is small, as is the approximation error. For example, for $\tau = .05$, the aggregation error is only 5.6 %, while the approximation error is -5.00 %. Making both errors actually helps; the approximation of the aggregate measure is only .4 % above the correct value.

When taxes increase, the aggregation error remains roughly constant at 5.6 % of the actual excess burden. However, though the second-order approximations differ little from each other, they both misrepresent substantially the actual excess burden.²³ Other simulations not presented corroborate this qualitative result. While this is of course a rather specific example, it does suggest that even though the aggregate individual expenditure function does not "really" exist, it may nevertheless be a very useful tool for welfare cost measurement. When available, it should be given as serious consideration as second order approximations.

V. Concluding Remarks

Our intent has been to provide for the practitioner a guide to the theory and estimation of the excess burden of taxation. The first step was to explain precisely the differences between the several measures that have appeared in the literature. We then indicated how the choice of a measure depended upon the conceptual experiment relevant to the investigator's concern.

A major practical problem we discussed was how much information is required to implement each measure. This question is particularly important because for a number of issues, the relevant data come to us only in aggregate form (say, in a time series), and we therefore lack information on the underlying structure of preferences in the population. Some simulation results suggest that treating an aggregate demand function "as if" it were generated by a single utility function yields quite reasonable estimates of excess burden.

Figure 1
Consumers' Surplus

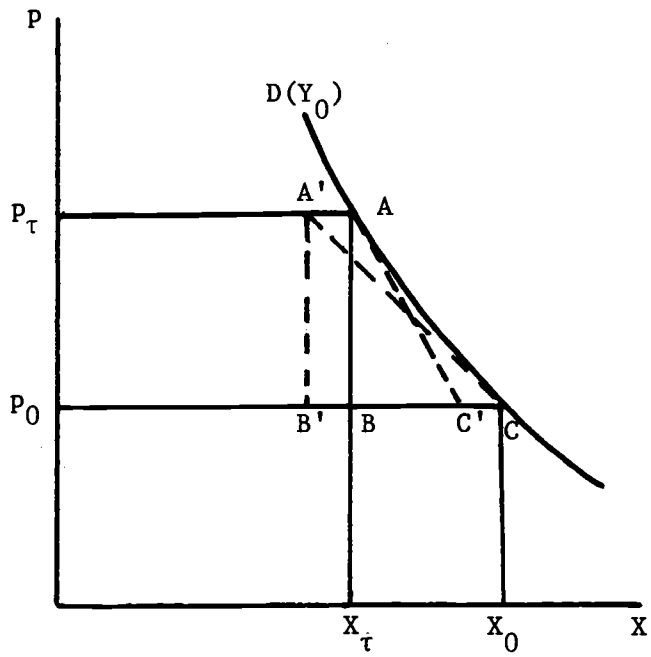


Figure 3
Income Compensation Illustrated

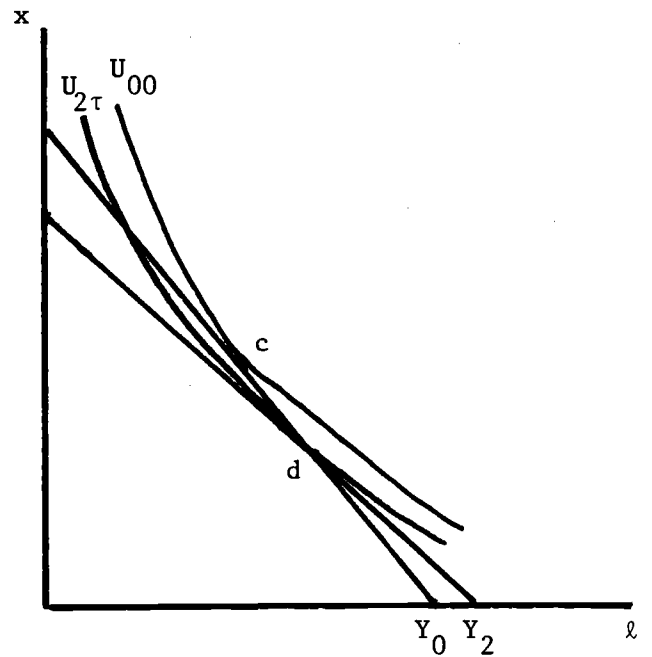


Figure 2
A General Equilibrium Analogue

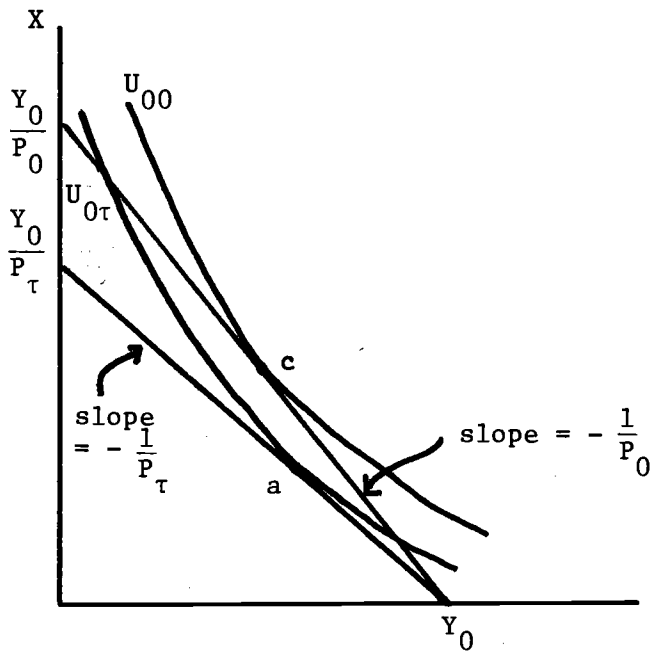


Figure 4
The "Harberger Triangle"

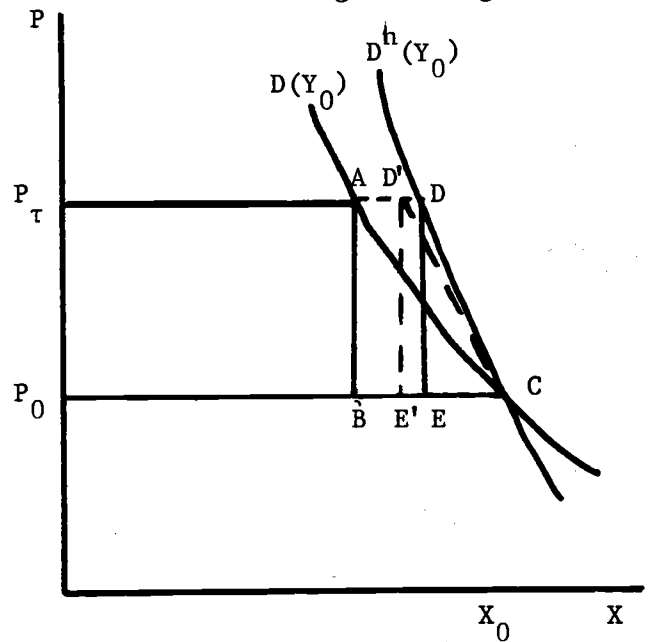


Figure 5
Utility Compensation Illustrated

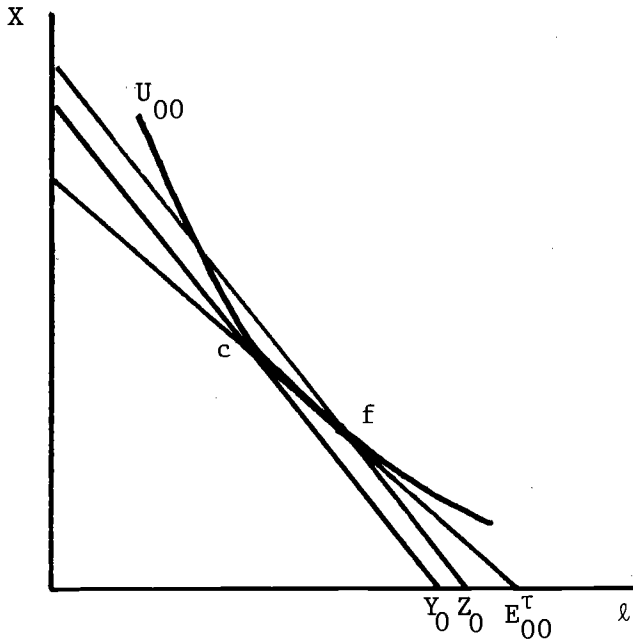


Figure 7
An Equivalent Variation Measure

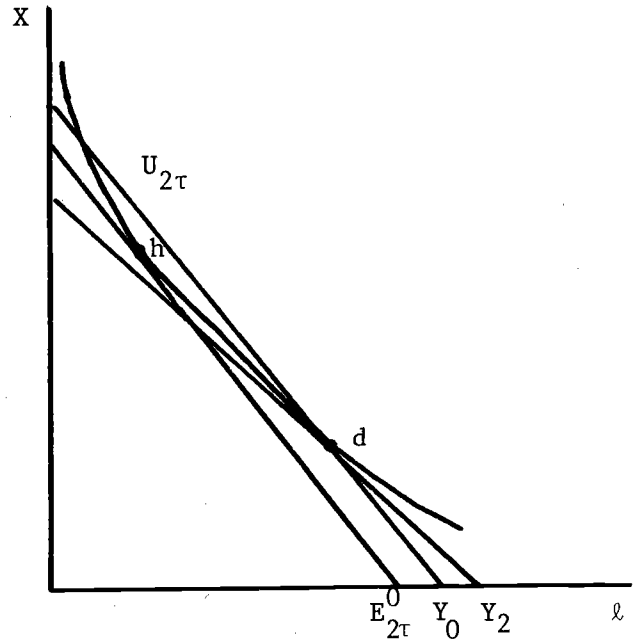


Figure 6
An Exact "Partial Equilibrium" Measure

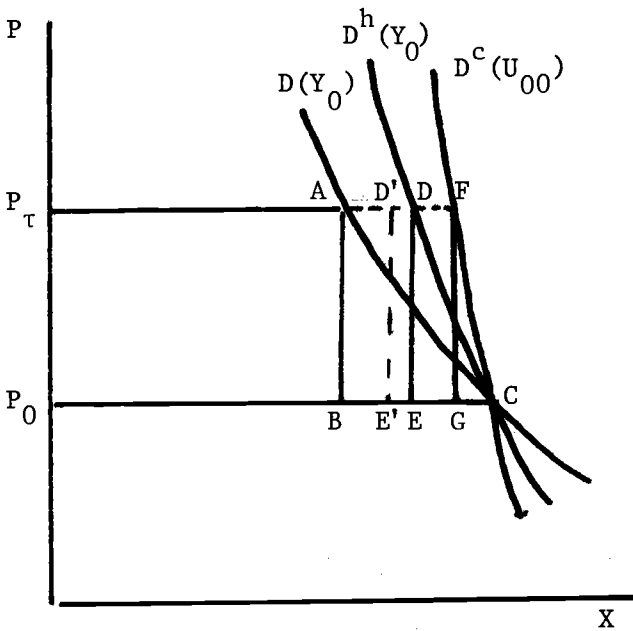


Figure 8
A Comparison of Approaches

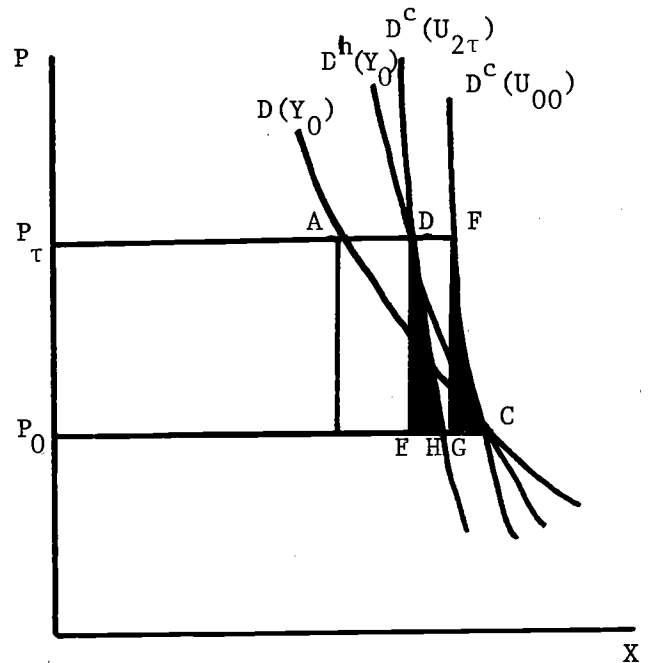


Figure 9
An Index Number Problem

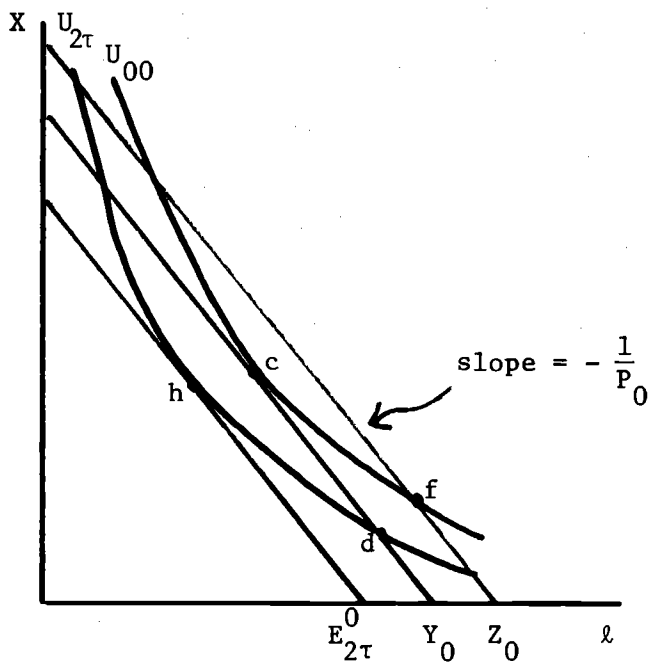


Figure 10
Post-Tax Measures

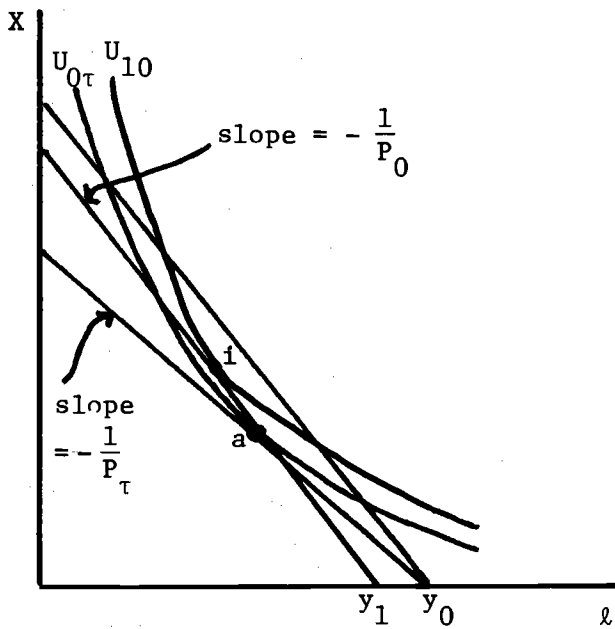


Table 1

Concept	Excess Burden Measures		
	Initial Income	Initial Prices	Exact Measure
Consumers' Surplus (Laidler 1969)	-	-	-
Consumers' Surplus w/income compensation (Harberger 1964)	pre-tax	-	-
Consumers' Surplus w/ income compensation	post-tax	-	-
Compensating Variation (Diamond and McFadden 1974)	pre-tax	post-tax	$E(U_{00}, p_{\tau}) - E(U_{00}, p_0)$ $-\Delta px^C(U_{00}, p_{\tau})$
Compensating Variation	post-tax	post-tax	$E(U_{00}, p_{\tau}) - E(U_{00}, p_0)$ $-\Delta px^C(U_{10}, p_{\tau})$
Equivalent Variation	pre-tax	pre-tax	$E(U_{2\tau}, p_{\tau}) - E(U_{2\tau}, p_0)$ $-\Delta px^C(U_{2\tau}, p_{\tau})$
Equivalent Variation (Kay 1980)	post-tax	pre-tax	$E(U_{0\tau}, p_{\tau}) - E(U_{0\tau}, p_0)$ $-\Delta px^C(U_{0\tau}, p_{\tau})$

	Taylor Approximation
$-\frac{1}{2} \frac{\partial x(y_0, p)}{\partial p} \Big _{p_0} (\Delta p)^2$	$-\frac{1}{2} \frac{\partial x(y_0, p)}{\partial p} \Big _{p_0} (\Delta p)^2$
$-\frac{1}{2} \frac{\partial x^h(y_0, p)}{\partial p} \Big _{p_0} (\Delta p)^2$	$-\frac{1}{2} \frac{\partial x^h(y_0, p)}{\partial p} \Big _{p_0} (\Delta p)^2$
$-\frac{1}{2} \frac{\partial x^h(y_1, p)}{\partial p} \Big _{p_0} (\Delta p)^2$	$-\frac{1}{2} \frac{\partial x^h(y_1, p)}{\partial p} \Big _{p_0} (\Delta p)^2$
$-\frac{1}{2} \frac{\partial x^C(U_{00}, p)}{\partial p} \Big _{p_0} (\Delta p)^2$	$-\frac{1}{2} \frac{\partial x^C(U_{00}, p)}{\partial p} \Big _{p_0} (\Delta p)^2$
$-\frac{1}{2} \frac{\partial x^C(U_{10}, p)}{\partial p} \Big _{p_0} (\Delta p)^2$	$-\frac{1}{2} \frac{\partial x^C(U_{10}, p)}{\partial p} \Big _{p_0} (\Delta p)^2$
$-\frac{1}{2} \frac{\partial x^C(U_{2\tau}, p)}{\partial p} \Big _{p_{\tau}} (\Delta p)^2$	$-\frac{1}{2} \frac{\partial x^C(U_{2\tau}, p)}{\partial p} \Big _{p_{\tau}} (\Delta p)^2$
$-\frac{1}{2} \frac{\partial x^C(U_{0\tau}, p)}{\partial p} \Big _{p_{\tau}} (\Delta p)^2$	$-\frac{1}{2} \frac{\partial x^C(U_{0\tau}, p)}{\partial p} \Big _{p_{\tau}} (\Delta p)^2$

Table 2

A Cobb-Douglas Example

<u>Concept</u>	<u>Initial Income</u>	<u>Initial Prices</u>	<u>Exact Measure</u>	<u>Taylor Approximation</u>
C.V. I	pre-tax	post-tax	$[(1-\tau)^{-\alpha}(1-\alpha\tau)-1]y_0$	$(\frac{\tau}{2})\alpha(1-\alpha)y_0 \cdot \frac{1}{(1-\tau)^2}$
C.V. II	post-tax	post-tax	$[(1-\tau)^{-\alpha}(1-\alpha\tau)-1]y_0 \cdot (1-\alpha\tau)$	$(\frac{\tau}{2})\alpha(1-\alpha)y_0 \cdot \frac{(1-\alpha\tau)}{(1-\tau)^2}$
E.V. I	pre-tax	pre-tax	$[(1-\tau)^{-\alpha}(1-\alpha\tau)-1]y_0 \cdot \frac{(1-\tau)^\alpha}{(1-\alpha\tau)}$	$(\frac{\tau}{2})\alpha(1-\alpha)y_0 \cdot \frac{1}{1-\alpha\tau}$
E.V. II	post-tax	pre-tax	$[(1-\tau)^{-\alpha}(1-\alpha\tau)-1]y_0 \cdot (1-\tau)^\alpha$	$(\frac{\tau}{2})\alpha(1-\alpha)y_0$

Table 3

Aggregation Error - Simulation Results

<u>Tax Rate</u>	<u>Exact Disaggregate</u>	<u>Exact Aggregate</u>	<u>SOA Disaggregate</u>	<u>SOA Aggregate</u>
0.0500000	0.0003114	0.0003289	0.0002958	0.0003125
0.1000000	0.0013139	0.0013679	0.0011833	0.0012500
0.1500000	0.0031275	0.0033034	0.0026625	0.0028125
0.2000000	0.0058992	0.0062306	0.0047333	0.0050000
0.2500000	0.0098129	0.0103630	0.0073958	0.0078125
0.3000000	0.0151001	0.0159443	0.0106500	0.0112500
0.3500000	0.0220577	0.0232866	0.0144958	0.0153125
0.4000000	0.0310723	0.0327956	0.0189333	0.0200000
0.4500000	0.0426579	0.0450098	0.0239625	0.0253125
0.5000000	0.0575131	0.0606602	0.0295833	0.0312500

Footnotes

¹The assumption of fixed prices and zero profits allow us to ignore the possibility of efficiency losses on the production side of the economy.

²See Green and Sheshinski (1979) for further discussion of third-order vs. second-order measures.

³The linear budget line corresponds to the fixed-producer price assumption made above: we continue to focus on this case rather than the more general one where relative prices depend, through a convex production possibilities frontier, on the mix of production. This framework is probably adequate for most problems in applied welfare economics, although it is clearly inapplicable in some special cases (e.g., a Leontief technology).

⁴See, e.g., Laidler (1969).

⁵This point has been made by Rosen and Small (1979) and Hausman (1979).

⁶A concise and elegant treatment of expenditure functions is given in Diamond and McFadden (1974).

⁷See Diamond and McFadden (1974).

⁸It is now easy to demonstrate the point made earlier about the possible inadequacy of consumers' surplus as an excess burden measure even when it is a good approximation to the total welfare loss. While the area $p ACp_0$ may be a reasonable approximation of the area $p FCp_0$, the same cannot be said of ABC as an approximation of FGC.

⁹For the case where there are several goods and price changes, x and p can be interpreted as vectors, with no change in the result.

¹⁰It might seem more sensible to expand around p_τ , since we are compensating in the presence of the distortion. However, we follow common practice here.

¹¹We are unaware of any previous use in the literature on excess burden.

¹²By the definition of y_2 in (1), we could rewrite (4) as $E(U_{00}, p_0) - E(U_{2\tau}, p_0)$; however, the current expression permits easier comparison with the previous measure using compensating variation.

¹³The equivalent variation measure has been used by Rosen (1978) and Kay (1980).

¹⁴More precisely, for each of the measures "all" that is needed is the shape of one indifference curve. (See Haise 1975). In most contexts, however, an investigator infers the shape of any given indifference from the overall utility function.

¹⁵To do so, one need only take advantage of Roy's identity, which relates the ordinary demand curve to partial derivatives of the indirect utility function. See Hausman (1979).

¹⁶See, for example, Wales and Woodland (1976).

¹⁷See, for example, Laidler (1969) for Harberger (1964).

¹⁸Which Taylor approximation is chosen will depend in part on the points at which income and price effects are known,

¹⁹See Gorman (1953).

²⁰Here, we use the fact that $s(y, p) = x^c(v(y, p), p)$.

²¹Of course, "truth" is conditional on choice of conceptual experiment.

²²If anything, we are understating the error normally incurred by using the aggregate demand curve, since we have taken the distribution of income over α , and hence $\bar{\alpha}$, to be constant. Normally, $\bar{\alpha}$ would vary, and the use of a constant aggregate taste parameter when computing the underlying utility function would introduce additional error.

²³The negative bias is due to the positive second derivative of the compensated demand curve which we ignore in using a second-order approximation.

References

- Diamond, P. and D. McFadden, 1974, "Some Uses of the Expenditure Function in Public Finance," Journal of Public Economics, February, pp. 3-21.
- Gorman, W.M., 1953, "Community Preference Fields," Econometrica, January, pp. 63-80.
- Green, J. and E. Sheshinski, 1979, "Approximating the Gains from Welfare Reforms," Journal of Public Economics, April, pp. 179-195.
- Harberger, A.C., 1964, "Taxation, Resource Allocation, and Welfare," in J. Due, ed., The Role of Direct and Indirect Taxes in the Federal Revenue System (Princeton: NBER).
- _____, 1974, Taxation and Welfare. (Boston: Little, Brown).
- Hause, J.C., 1975, "The Theory of Welfare Cost Measurement," Journal of Political Economy, December, pp. 1145-1182.
- Hausman, J., 1979, "Exact Consumers' Surplus," mimeo, MIT.
- Hicks, J.R., 1946, Value and Capital, 2nd ed. (Oxford: Oxford University Press).
- Hotelling, H., 1938, "The General Welfare in Relation to Problems of Taxation and of Railway and Utility Rates," Econometrica, July, pp. 242-269.
- Kay, J.A., 1980, "The Deadweight Loss From a Tax System," Journal of Public Economics, February, pp. 111-120.
- Laidler, D., 1969, "Income Tax Incentives for Owner-Occupied Housing," in A.C. Harberger and M.J. Bailey, eds., The Taxation of Income from Capital (Washington: Brookings).
- Mohring, H., 1971, "Alternative Gain and Loss Measures," Western Economic Journal, December, pp. 349-368.
- Rosen, H.S., 1978, "The Measurement of Excess Burden with Explicit Utility Functions," Journal of Political Economy, April, pp. S121-S135.
- Rosen, H.S., and K. Small, 1979, "Applied Welfare Economics with Discrete Choice Models," NBER Working Paper #319, February.
- Wales, T.J., and A.D. Woodland, 1976, "Estimation of Household Utility Function and Labor Supply Response," International Economic Review, June, pp. 397-410.
- Willig, R., 1976, "Consumers' Surplus Without Apology," American Economic Review, September, pp. 589-597.