Nizar Jetha

The Welfare Cost of Taxation
Its Meaning and Measurement

World Bank Reprints

No. 217  Carl J. Dahlman and Larry E. Westphal, "The Meaning of Technological Mastery in Relation to Transfer of Technology," *Annals of the American Academy of Political and Social Science*


No. 225.  George Psacharopoulos, "The Economics of Higher Education in Developing Countries," *Comparative Education Review*

No. 226.  Katrine Anderson Saito and Delano P. Villanueva, "Transaction Costs of Credit to the Small-scale Sector in the Philippines," *Economic Development and Cultural Change*


No. 230.  Abdun Noor, "Managing Adult Literacy Training," *Prospects*


No. 233.  Keith Bradley and Alan Gelb, "Motivation and Control in the Mondragon Experiment," and "The Replication and Sustainability of the Mondragon Experiment," *British Journal of Industrial Relations*

I. INTRODUCTION

Statements such as “the welfare cost of tax X equals α% of its revenue yield” or that “the welfare gain from substituting tax Y for tax Z would equal β% of GNP” are to be found in a growing number of papers on tax reform. They refer to the cost of distortions caused by taxes through their effects on relative prices. All taxes, except the poll tax, distort the allocation of resources in the sense that the allocation of resources would change if the taxes were removed. The concept of the welfare cost of taxation attempts to translate the cost of tax-induced distortions in money terms. The concept is based on the notion of consumer’s surplus.

This paper provides an introduction to the concept of the welfare cost of taxation and its measurement. In addition, it derives the basic conclusion of the theory of optimal indirect taxation by taking advantage of the close relationship between the concept of welfare loss and the theory of optimal taxation.

II. CONSUMER’S SURPLUS

The consumer’s surplus is simply the difference between the amount a consumer is willing to pay for a good and the amount he actually pays for it. Thus, if one is prepared to pay $1 for a can of beer but its price is 40 cents, the consumer’s surplus is 60 cents. Every point on the ordinary demand curve represents the sum a consumer is willing to pay for various quantities. When the price of a good X is P and the quantity X0 is purchased, the amount the consumer is willing to pay for the X0th unit and the price are equal, but the amounts the consumer is willing to pay for the other units purchased exceeds the (fixed) price of the commodity (see Figure 1). At price P0 the total sum a consumer is willing to pay for the quantity X0 is given by OScX0. The area of the triangle P0Sc, therefore, represents the consumer’s surplus.

Now suppose that a tax of Tx per unit of X is imposed. The price rises to P1 and demand falls to X1. The willingness to pay for X1 is readied by OScX1. The consumer surplus is now P1Sc. The reduction in the consumer surplus due to the increase in price is, therefore, P0P1Sc.

The decrease in consumer surplus due to an increase in price gives the total welfare loss. However, when the price increase is caused by a tax, the additional tax revenues should be subtracted from the total welfare loss because the revenues, when spent, would bring benefits. The tax of Tx per unit of X would bring tax revenues of TxX1, denoted by the area P0P1ab. Hence, we are finally left with the triangle abc as the measure of the welfare cost of taxation. The loss of welfare due to taxation is variously referred to as the “excess burden”, “deadweight loss” and “efficiency cost”.

III. CONSUMER’S SURPLUS AND COMPENSATING VARIATION

It has been suggested that a price-induced change in consumer’s surplus is given by the area to the left of the demand curve between the two prices. Strictly, the change in the consumer’s surplus is not given by the area to the left of the ordinary demand curve, but by the area to the left of a compensated demand curve. The present section explains the reason for this as well as the meaning of the compensated demand curve. In the process, further light will be shed on the concept of consumer’s surplus.

The upper part of Figure 2 shows income on the vertical axis and purchases of a good X on the horizontal axis. U1 and U2 are indifference curves, and YZ is the initial budget line corresponding to price P1 for X. At the price P1, the consumer spends ¥X of his income on X, purchasing the quantity X1. This gives us one point on the ordinary and compensated demand curves, denoted by a in the lower part of the diagram. When the price rises to P2, the budget line changes to YW, giving a new equilibrium on the lower indifference curve U2. The purchases of X fall to X2 and a second point on the ordinary demand curve (b) is obtained. To obtain the corresponding point on the compensated demand curve, we need to ask what quantity of X the consumer would purchase at the higher price P2 if his income were increased so that he could remain on the original indifference curve U1. The answer can be found by “moving” the budget line Y W to the right, keeping its slope constant, until it touches U1. As the diagram shows, the answer is that the quantity X3 would be purchased, and this gives us the second point on the compensated demand curve (c). Other points can be derived by considering different prices. It should be clear that the compensated demand curve is a relationship between prices and quantities, when a consumer remains on a given indifference curve (that is, when his utility is held constant).


1. A simple mathematical derivation of a compensated demand curve may help. Let the utility function be of the form U = Q1Q2, where Q1 and Q2 are the quantities of goods 1 and 2 consumed, respectively. Let P1 and P2 be the prices of goods 1 and 2, respectively. Minimizing the expenditure on the two goods (P1Q1 + P2Q2) subject to the initial level of utility (say U) being maintained will give the compensated demand curves. Minimizing the Lagrangean:

\[ L = P1Q1 + P2Q2 + \varepsilon (U - Q1Q2) \]

we get

\[ \delta L = P1 - \varepsilon Q2 = 0 \]

and

\[ \delta L = P2 - \varepsilon Q1 = 0 \]

\[ \delta \varepsilon = U - \varepsilon Q1Q2 = 0 \]

which give

\[ Q1 = \left( \frac{P1}{\varepsilon} \right) ^{1} \]

and

\[ Q2 = \left( \frac{P2}{\varepsilon} \right) ^{1} \]

The amount of compensation that must be paid to the consumer in order to enable him to remain on \( U_t \) when the price of \( X \) rises from \( P_1 \) to \( P_t \) is given by \( YY^* \). In general, the amount of compensation paid or received that will leave the consumer in his initial welfare position following a change in price is known as "compensating variation". To put it differently, compensating variation measures the change in utility due to a price change in money terms. But this is precisely what the change in consumer's surplus is supposed to represent. It can be shown, but will not be shown here, that the compensating variation equals the area to the left of the compensated demand curve between the two prices. In terms of the diagram, \( YY^* \) corresponds to \( P_t P_i X \). This provides the rationale for the use of compensated demand curves.

In light of the above discussion, we can define excess burden as compensating variation minus tax receipts. In practice, the excess burden is often calculated directly by estimating the area of the relevant triangle under the compensated demand curve (see in Figure 2).

### IV. MEASUREMENT OF WELFARE COST OF TAXATION

Suppose that a tax of \( T_x \) per unit of good \( X \) is levied. Referring to Figure 1 but treating the demand curve as the compensated demand curve, the excess burden of the tax \( dP/ X \) may be expressed algebraically as:

\[
W = -\frac{1}{2} \left( \frac{dX}{dP} \right) \left( \frac{dP}{P} \right)
\]

(1) and (2) are two equivalent formulae for computing the excess burden of a single tax. These partial equilibrium formulations ignore the existence of other taxes. Moreover, frequently we wish to know the welfare gain or loss of increasing or decreasing one tax and changing another tax to keep the tax revenue constant. We, therefore, need a general expression for the excess burden of a given set of taxes. The general expression for the excess burden of \( N \) taxes is:

\[
W = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} S_{ij} T_i T_j
\]

(3)

where \( S_{ij} = \left( \frac{\partial X_i}{\partial P_j} \right) \) is the compensated derivative of the demand for the \( i \)th commodity with respect to the price of the \( j \)th commodity, and \( T_i \) is the amount of tax per unit of good \( i \).

The general expression can be reduced to (2) in the case of a tax on a single good, as follows:

\[
W = -\frac{1}{2} \left( \frac{dX_i}{dP_i} \right) \left( \frac{dP_i}{P_i} \right)^2
\]

(4)


3. Note that there is no unique measure of utility change in money terms. Equivalent variation could be used instead of compensating variation. See Henderson and Quandt, pp. 49-50.

4. This is not difficult to show using indirect utility functions and expenditure functions, but a discussion of these concepts will take us too far afield. See Henderson and Quandt, pp. 41-45, 49-52.

5. A simple numerical illustration of a compensated demand curve, compensating variation and excess burden will be found in P. B. Dixon, S. Bowles and D. Kendrick, Notes and Problems in Microeconomic Theory (Amsterdam: North Holland, 1980), pp. 144-146.

6. The expression is derived in a number of different ways in Arnold C. Harberger, "Taxation, Resource Allocation and Welfare", reprinted in Taxation and Welfare (Boston: Little, Brown and Company, 1974), pp. 25-62. It should be noted that none of the assumptions underlying the derivations is that outputs are produced at constant costs.

7. It will be recalled that the relationship between the derivatives of an ordinary (or uncompensated) demand curve and a compensated demand curve is given by the Slutsky equation.

\[
\frac{\partial X_i}{\partial P_j} = \frac{\partial X_i}{\partial P_j} - \frac{\partial X_i}{\partial x} \cdot \frac{\partial x_i}{\partial y}
\]

8. The expression on the left is the derivative of the ordinary demand curve and the first term on the right is the derivative of the compensated demand curve. We denote the compensated derivative

\[
\frac{\partial X_i}{\partial P_j} \text{ utility constant}
\]

by \( S_{ij} \). It should be noted for later reference that \( S_{ij} = S_{ji} \). For a derivation of the Slutsky equation and its interpretation, see Henderson and Quandt, pp. 25-32.
V. AN APPLICATION

Laidler has estimated the welfare cost of income tax incentives for owner-occupied housing in a simple partial equilibrium framework. The main features of his approach are illustrated in Figure 3. D represents the demand for housing services. $S_2$ is the supply curve and denotes the marginal private cost of providing housing services. The subsidy given through the tax system lowers the private marginal cost of housing to the levels represented by $S_1$. As a result of the subsidy, the consumption of housing services expands by $Q_2Q_1$.

The net loss of welfare is denoted by the shaded triangle $cdg$. This can be seen as follows. The expanded consumption of housing services increases the consumer surplus by $cBg$. But the cost of the subsidy ($bdc$) must be set against it to arrive at the net change in welfare. The net loss of welfare, therefore, given by $cdg$. Since $cdg$ and $fcg$ are similar triangles, the area of the triangle $fcg$ also gives the welfare loss. We obviously need to know the size of the subsidy and the increase it induces in the quantity of housing in order to estimate the size of welfare loss.

If a homeowner were taxed in the same manner as other investors, his taxable income would be given by:

$$\text{Taxable income} = (GR - PT - D - M - M_I)V$$

where $GR$ = gross rent, $PT$ = property taxes, $D$ = depreciation, $M$ = maintenance, and $M_I$ = mortgage interest, all expressed as a percentage of the value of the house ($V$). Taking the numerical values used by Laidler:

$$\text{Taxable income} = (0.11 - 0.015 - 0.0225 - 0.0125 - 0.06)V$$

This states that if the homeowner were to rent his house, he would have to charge a gross rent of 11% of the value of the house if he is to meet all expenses. However, tax laws in the United States and many other countries permit the homeowner to exclude gross rent from his taxable income and to deduct property taxes and mortgage interest from his other income. Taxable income from owner-occupied housing is, therefore, given by:

$$\text{Taxable income} = (-PT - M_I)V$$

Subtracting (5) from (4) would give the reduction in taxable income due to the special tax treatment of owner-occupied housing. Multiplying this reduction in taxable income by the taxpayer’s marginal tax rate then gives the tax saving or subsidy associated with the tax incentive for owner-occupied housing. Thus,

$$\text{Subsidy} = (GR - PT - D - M - M_I)V$$

where the marginal rate of personal income tax (t) has been taken to be 26%. Hence the amount of the subsidy is $0.0195$ per $S$ of the cost of the house, and the rate of subsidy is 17.7% ($0.0195V/0.11V$).

To estimate the change in the quantity of housing, we need to know the price elasticity of demand for housing and the stock of owner-occupied housing. Laidler used an elasticity of $-1.5$ for his computations. The value of owner-occupied housing owned by persons in the income group $10,000 - 15,000$ was $61,956 million. Multiplying the rate of subsidy (17.7%) by elasticity of demand (1.5) will give the percent over-invested in housing of 26.6%. This translates into the amount over-invested in housing of $13,011 million [(26.6/(100 + 26.6)) x 61,956].

The welfare loss may now be computed as follows: W = 0.5 $\Delta P \Delta Q$

$$= 0.5 \text{(subsidy)} \text{(amount over-invested)}$$

$$= 0.5(0.0195)(13,011) = $137 million$$

Making similar calculations for other income classes, Laidler arrived at the total welfare cost of the tax concessions for owner-occupiers of $300 million per annum.

VI. EXCESS BURDEN AND OPTIMAL INDIRECT TAXATION

The concept of excess burden leads naturally to the theory of optimal taxation, and perhaps provides the simplest way of deriving a basic proposition in optimal taxation. Optimal tax rates are defined as the rates that will raise the required revenue by minimizing the excess burden of taxation. We wish to investigate the main characteristic of the structure of optimal indirect taxation. The traditional theory of taxation concluded that an ad valorem tax at a uniform rate on all commodities would minimize the excess burden of taxation. Recent research has shown that this conclusion would only be valid if all final commodities (including leisure) could be taxed. The modern approach is based on the assumption that leisure cannot be taxed. The question that the theory of optimal indirect taxation then poses is: which tax rates will minimize the excess burden of indirect taxation given that a specific amount of revenue must be raised and that at least one commodity cannot be taxed?

As will be shown, the answer is that the tax rates should be so chosen that the demand for all taxed commodities is reduced by the same proportion. A single rate for all goods will, therefore, not be optimal.


11. The minus sign vanishes because a subsidy is a negative tax.

To find optimal rates of indirect taxes, we minimize (with respect to tax rates):

$$ W = - \frac{1}{2} \sum_{i=1}^{K} \sum_{j=1}^{K} S_{ij} T_i T_j \quad K < N $$

subject to

$$ \sum_{j=1}^{K} T_j X_j = R $$

where $S_{ij}$ and $T$ have already been defined, $X_j$ is the quantity of the $j$th good purchased, and $R$ is tax revenue. $K < N$ denotes that at least one good cannot be taxed. The constraint (the second expression) simply states that the chosen tax rates must bring revenue equal to $R$.

We can simplify this problem without loss of generality. First, define $T_i = t_iC_i$, where $t_i$ is percentage rate of tax on the $i$th good and $C_i$ is the producer's price (cost) for the $i$th good. The solution will then contain ad valorem rates, which are of greater interest than specific rates ($T_i$). Second, let producers' prices for all commodities equal 1 (that is $C_i = 1$ for all $i$). This will simplify the mathematics by reducing the number of symbols. Note that the relationship between the consumers' and producers' prices will be given by $P_i = (1 + t_i)C_i = (1 + t_i)$, where $P_i$ is the consumer's price for the $i$th good. Third, assume that there are only three commodities, two goods ($X_1$ and $X_2$), which can be taxed, and leisure ($X_3$), which cannot be taxed. This will make it possible to write out all expressions in full, thereby making the derivation clearer.

Our problem now is to minimize:

$$ W = - \frac{1}{2} \sum_{i=1}^{3} \sum_{j=1}^{3} S_{ij} t_i t_j $$

subject to

$$ \sum_{j=1}^{3} t_j X_j = R. $$

Recalling that $t_3 = 0$ and noting that $S_{12} = S_{31}$, minimization of $W$ subject to the revenue constraint is equivalent to the minimization of the Lagrangian:

$$ L = - \frac{1}{2} [S_{11} t_1 + 2S_{12} t_2 + 2S_{31} t_3] + g (t_1 X_1 + t_2 X_2 - R) $$

We obtain

$$ \frac{\partial L}{\partial t_1} = - [S_{11} + S_{12} t_2] + g t_1 \frac{\partial X_1}{\partial t_1} + 2 \frac{\partial X_2}{\partial t_1} = 0 \quad (6) $$

$$ \frac{\partial L}{\partial t_2} = - [S_{12} + 2S_{32} t_3] + g t_1 \frac{\partial X_1}{\partial t_2} + 2 \frac{\partial X_2}{\partial t_2} + \frac{\partial X_3}{\partial t_2} = 0 \quad (7) $$

The next step is to transform expressions such as $\frac{\partial X_1}{\partial t_1}$ into $S_{ij}$. This can be done by noting that:

$$ \frac{\partial X_j}{\partial t_j} = \frac{\partial X_j}{\partial P_j} \frac{\partial P_j}{\partial t_j} = \frac{\partial X_j}{\partial P_j} (\text{since } P_j = 1 + t_j) $$

and recalling the Slutsky equation

$$ \frac{\partial X_j}{\partial P_j} = S_{ij} - X_j \frac{\partial X_j}{\partial y} $$

Substituting the right hand side of the Slutsky equation for $\frac{\partial X_j}{\partial P_j}$ into (6)

$$ - [S_{11} t_1 + S_{12} t_2] + g [t_1 (S_{11} - X_3) \frac{\partial X_1}{\partial y} ] $$

$$ + t_2 (S_{12} - X_3) \frac{\partial X_2}{\partial y} + X_3 = 0 $$

Collecting terms in $S_{11}$ and $S_{12}$:

$$ (g-1) [S_{11} t_1 + S_{12} t_2] = g [t_1 \frac{\partial X_1}{\partial y} + t_2 \frac{\partial X_2}{\partial y} - 1] X_1 $$

or $S_{11} t_1 + S_{12} t_2 = C X_1$ \quad (8)

where $C = \frac{g}{g-1} (t_1 \frac{\partial X_1}{\partial y} + t_2 \frac{\partial X_2}{\partial y} - 1)$

Similarly (7) can be reduced to:

$$ S_{12} t_2 + S_{32} t_3 = C X_2 $$

(9)

(8) gives an estimate of the proportion by which the demand for $X_1$ would decline as a result of the taxes on $X_1$ and $X_2$ (9) states that the demand for $X_3$ should decline by the same proportion. It follows that (8) and (9) together imply that tax rates should be so chosen that the demands for both $X_1$ and $X_2$ decline by the same proportion (along compensated demand curves). This proposition, known as the Ramsey rule, is what we had set out to establish.

A special case of the Ramsey rule arises when the demands are independent (that is, when $S_{12} = S_{32} = 0$). The implications of the solution will be more clearly seen if derivatives are replaced by elasticities, and the amounts of tax payable by ad valorem tax rates. To do this, we divide both sides of (8) by $X_1$ and multiply the left hand side by $P_j/P_i$. Similarly, we divide both sides of (9) by $X_2$ and multiply the left hand side by $P_j/P_i$. We get

$$ \left( \begin{array}{c} S_{11} P_i \\ S_{12} P_i \end{array} \right) \left( \begin{array}{c} t_1 P_i \\ t_2 P_i \end{array} \right) = C $$

or

$$ \left( \begin{array}{cc} S_{11} P_i & 1 \\ S_{12} P_i & 1 \end{array} \right) \left( \begin{array}{c} t_1 P_i \\ t_2 P_i \end{array} \right) = C $$

(10)

(11)

But $S_{11} P_i = \frac{\partial X_1}{\partial y}$

$$ \left( \begin{array}{c} t_1 P_i \\ t_2 P_i \end{array} \right) \left( \begin{array}{c} P_i \\ X_1 \end{array} \right) $$

is nothing but the compensated elasticity of demand for the $i$th good with respect to its price (to be denoted by $e_i$) and $t_i P_i$ is simply the ad valorem rate of tax (amount of tax as a proportion of the gross-of-tax price) for the $i$th good (to be denoted by $r_i$). Combining (10) and (11)

$$ \begin{array}{r} r_1 = \frac{e_1}{1+e_1} \\ r_2 = \frac{e_2}{1+e_2} \end{array} $$

In words, this states that the tax rates should be inversely related to (compensated) price elasticities. This result is referred to as the inverse elasticity rule.

---

One special case where the conclusion of the traditional theory will hold when leisure cannot be taxed is worth noting. This occurs when the supply of labor (and consequently the consumption of leisure) is fixed. Then, a tax at a uniform rate on all commodities other than leisure will not distort the structure of demand since the relative prices of taxed goods will remain unchanged, and the change in the relative price of leisure and other goods would not matter because the consumption of leisure could not be varied. The assumption that taxes do not affect the work-leisure choice lacks realism, however.

Optimal taxes have been discussed in relation to the allocation of resources. Equity considerations can also be introduced in the framework of optimal taxation.

VII. CONCLUDING REMARKS

The increasing use of welfare cost calculations in applied public finance has stimulated renewed interest in better measures of consumer's surplus. The usefulness of the concept of consumer's surplus for policy purposes has, however, always been surrounded by controversy. The reader alone must, therefore, decide how much importance to attach to excess burden calculations in judging the merits of different tax reform proposals.

14. We have suggested, but not demonstrated, that a uniform rate of tax would be optimal when all final goods (including leisure) can be taxed. An equivalent expression for (5) when \( T_i = \tau C_i \) and \( C_i = 1 \) is:

\[
W = \sum_{i} \sum_{j<i} (1 - \tau) t_i C_{ij}
\]

It follows that the welfare loss will be minimum (in this case, nil) when tax rates on all goods (including leisure) are equal. See Arnold C. Harberger, "Taxation, Resource Allocation and Welfare", Taxation and Welfare, p. 39.


No. 243. Emmanuel Jimenez, "The Value of Squatter Dwellings in Developing Countries," Economic Development and Cultural Change

No. 244. Boris Pleskovic and Marjan Dolenc, "Regional Development in a Socialist, Developing, and Multinational Country: The Case of Yugoslavia," International Regional Science Review


Issues of the World Bank Reprint Series are available free of charge from the address on the bottom of the back cover.
THE WORLD BANK

Headquarters:
1818 H Street, N.W.,
Washington, D.C. 20433, U.S.A.

Telephone: (202) 477-1234
Telex: WUI 64145 WORLDDBANK
RCA 248423 WORLDDBK
Cable address: INTBAFRAD
WASHINGTONDC

European Office:
66, avenue d'Téna
75116 Paris, France

Telephone: (1) 723-54.21
Telex: 842-620628

Tokyo Office:
Kokusai Building
1-1, Marunouchi 3-chome
Chiyoda-ku, Tokyo 100, Japan

Telephone: (03) 214-5001
Telex: 781-26838

The full range of World Bank publications, both free and for sale, is described in the Catalog of Publications; the continuing research program is outlined in Abstracts of Current Studies. The most recent edition of each is available without charge from:

PUBLICATIONS DISTRIBUTION UNIT
THE WORLD BANK
1818 H STREET, N.W.
WASHINGTON, D.C. 20433
U.S.A.

ISSN 0253-2131