Economics 230a, Fall 2018
Lecture Note 5: Excess Burden and Basic Optimal Taxation

Deadweight loss measures the economic cost of market distortions; when one is referring to the distortions caused by taxation, the deadweight loss is referred to as the excess burden of taxation, because it is the economic cost to taxpayers over and above the tax revenue collected.

Although deadweight loss is an intuitive concept, it may be defined in more than one way, depending on the conceptual experiment one has in mind. Consider the case of a representative consumer. Let y be the consumer's initial income endowment and $\boldsymbol{p_0}$ be the initial price vector the consumer faces. Assume that this price vector represents the undistorted prices charged by producers, and that these prices are fixed, for example set by world markets. Now, suppose government imposes a tax vector $\boldsymbol{t}$ to raise revenue, with the resulting new price vector $\boldsymbol{p_1} = \boldsymbol{p_0} + \boldsymbol{t}$. One definition of deadweight loss is the amount that one would have to give the consumer to compensate for the taxes, net of the revenue the government collects. Assuming the individual's indirect utility function is $V(\boldsymbol{p}, y)$, this leads to the following expression for deadweight loss:

(1) $\quad \text{DWL}_1 = E(\boldsymbol{p_1}, V(\boldsymbol{p_0}, y)) - y - \boldsymbol{t}'\boldsymbol{x}^c(\boldsymbol{p_1}, V(\boldsymbol{p_0}, y))$

where $E(\cdot)$ is the household's expenditure function and $\boldsymbol{x}^c(\cdot)$ is the vector of the household's compensated demands. The first two terms indicate how much the individual must be compensated to remain at the initial level of utility (the Hicksian measure known as compensating variation), and the third term indicates how much tax revenue is available for compensation. Note that, for consistency, we calculate this revenue assuming that the compensation is occurring, i.e., that the consumer remains on the original indifference curve.

Or, we could define deadweight loss as the amount we could take from the consumer, beyond the revenue we give up; to offset gains from removing the tax (the Hicksian equivalent variation):

(2) $\quad \text{DWL}_2 = y - E(\boldsymbol{p_0}, V(\boldsymbol{p_1}, y)) - \boldsymbol{t}'\boldsymbol{x}(\boldsymbol{p_1}, y)$

How do these measures relate? Using the identities $y = E(\boldsymbol{p_0}, V(\boldsymbol{p_0}, y)) = E(\boldsymbol{p_1}, V(\boldsymbol{p_1}, y))$, $\boldsymbol{x}(\boldsymbol{p_1}, y) = \boldsymbol{x}^c(\boldsymbol{p_1}, V(\boldsymbol{p_1}, y))$, and $\boldsymbol{t} = \boldsymbol{p_1} - \boldsymbol{p_0}$, we may rewrite these two expressions as:

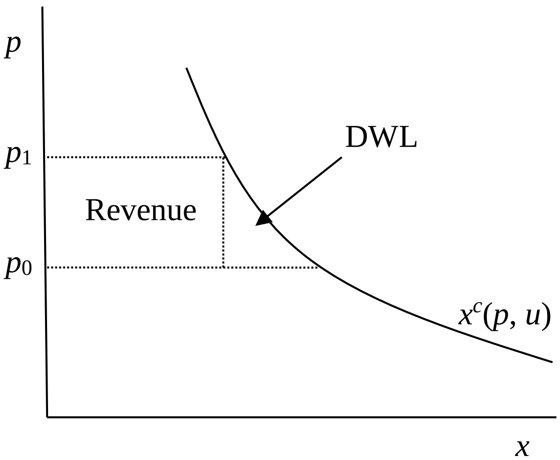
(1′) $\quad \text{DWL}_1 = E(\boldsymbol{p_1}, V(\boldsymbol{p_0}, y)) - E(\boldsymbol{p_0}, V(\boldsymbol{p_0}, y)) - (\boldsymbol{p_1} - \boldsymbol{p_0})'\boldsymbol{x}^c(\boldsymbol{p_1}, V(\boldsymbol{p_0}, y))$

(2′) $\quad \text{DWL}_2 = E(\boldsymbol{p_1}, V(\boldsymbol{p_1}, y)) - E(\boldsymbol{p_0}, V(\boldsymbol{p_1}, y)) - (\boldsymbol{p_1} - \boldsymbol{p_0})'\boldsymbol{x}^c(\boldsymbol{p_1}, V(\boldsymbol{p_1}, y))$

(1′) and (2′) differ only in the level of utility (pre-tax or post-tax) at which the calculation is made; the two measures will generally differ. Figure 3 in Auerbach-Hines provides a graphical illustration of DWL_2, and a similar measure can be drawn for DWL_1. It is also customary to illustrate these measures graphically in price-quantity diagrams. Assume that there are two goods, a numeraire commodity that is untaxed and has a price of 1, and a taxed commodity with price p and quantity x. Then measure DWL_1 (similarly for DWL_2) reduces to an expression that replaces the price and quantity vectors with scalars relating to the taxed good:

(3) $\quad \mathrm{DWL}_1 = E(p_1, V(p_0, y)) - E(p_0, V(p_0, y)) - (p_1 - p_0)'x^c(p_1, V(p_0, y))$

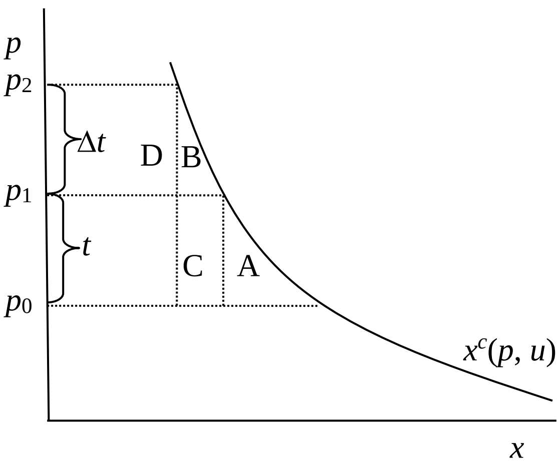
Using the fact that $E(p_1, u) - E(p_0, u) = \int_{p_0}^{p_1} x^c(p,u)dp$, we graph expression (3) as:



It is important to note that deadweight loss relates to the compensated demand curve, because the distortion is to relative prices; responses to lump-sum taxes also have income effects, but there is no associated distortion. The deadweight loss area, approximately triangular in shape, is known as a Harberger triangle; its size is approximately $-\frac{1}{2}t\Delta x$. One may also derive the vector version of this approximation for many taxes imposed simultaneously from expression (1′) or (2′) as a second-order Taylor approximation, equal to $-\frac{1}{2}\boldsymbol{t}'\Delta\boldsymbol{x}$, around the undistorted equilibrium.

Some Observations about Deadweight Loss

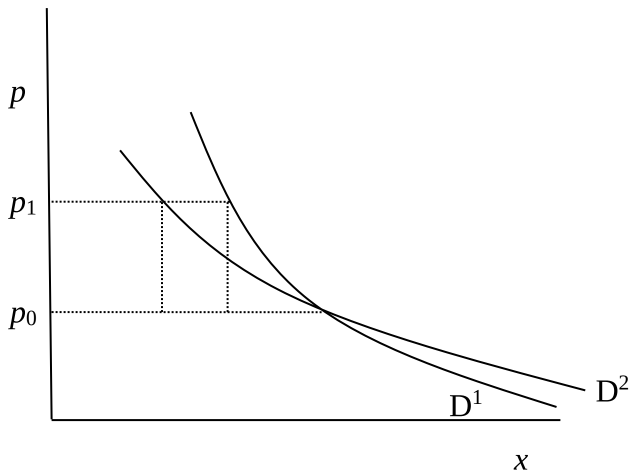
1. Deadweight loss rises roughly with the square of the tax. From the approximation we can see this, since when the tax rate t is twice as large, so, roughly, is the reduction in x, Δx (this is exact only if the demand curve is linear). For intuition, consider doubling the tax in the above graph:



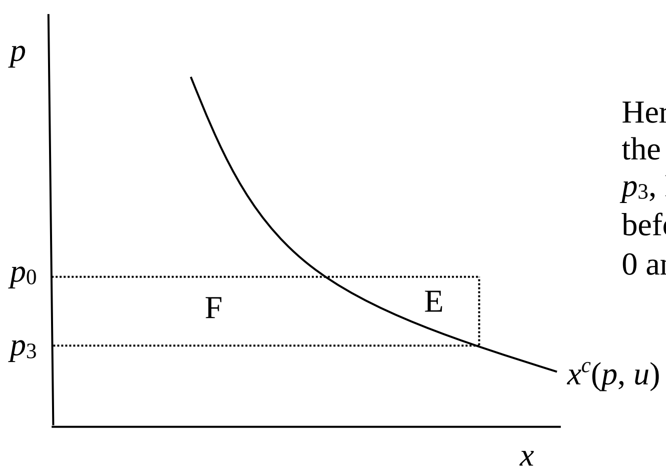
Increasing the tax so that the price rises from p_1 to p_2 incurs not only another Harberger triangle, B, but also a rectangle, C; whereas A and B are second-order terms, C is a first-order term – it does not vanish as the additional tax becomes small. This is because the additional tax is imposed starting at a distorted point. In other words, it is more costly to exacerbate an existing distortion. For the additional tax, say Δt, the second-order DWL approximation is $-(t\Delta x + \frac{1}{2}\Delta t\Delta x)$.

Another way to view the extra distortion, C, is that it is the revenue loss due to a decline in taxed-goods consumption. That is, additional revenue from the tax increase is D – C, not D. It is even possible for area C to exceed area D in size, in which case revenue would decline with the increase in the tax, indicating that the tax rate is higher than the revenue-maximizing rate, i.e., to the right of the peak of revenue as a function of the tax rate (also known as a Laffer curve).

2. Deadweight loss rises roughly in proportion to the elasticity of demand. Since Δx is approximately $t\cdot \partial x^c/\partial p = (t/p)\cdot p\cdot \partial x^c/\partial p = -(t/p)\cdot x\cdot \varepsilon$, where ε is the own compensated demand elasticity (defined so that it is nonnegative), we can rewrite the second-order approximation for DWL as $(t/p)^2\cdot px\cdot\varepsilon$. Graphically, we can see that a larger elasticity of demand increases deadweight loss and reduces revenue for a given tax rate.



3. Excess burden applies for subsidies as well as taxes. With a subsidy, an individual will be better off, but by less than if the government had transferred the revenue directly:



Here, area E represents DWL, equal to the revenue cost of reducing the price to p_3, E+F, net of the individual gain, F. As before, DWL $\approx -\frac{1}{2}\Delta t \Delta x$, where now $\Delta t < 0$ and $\Delta x > 0$.

4. Excess burden applies to supply distortions as well. Consider, for example, the labor supply decision, which we can represent using a horizontal labor demand curve and an upward sloping compensated labor supply curve. We can also measure deadweight loss in markets where both demand and supply are not infinitely elastic. (See Auerbach and Hines, sec. 2.2.) The basic formula still applies; in particular the distortion is proportional to the change in quantity, Δx.

Can't we avoid excess burden by imposing taxes at the same rate on all commodities? If the consumer's budget constraint is $\boldsymbol{p}'\boldsymbol{x} = y$, why not impose taxes at a constant rate, θ, so that the household's budget constraint becomes $(1+\theta)\boldsymbol{p}'\boldsymbol{x} = y \Rightarrow \boldsymbol{p}'\boldsymbol{x} = y/(1+\theta)$ – effectively, a lump-sum income tax? The problem is that most of what we call "income" results from individual choices, for example how much labor to supply. That choice would be distorted by the proposed scheme. Suppose that income equals $w(\bar{L} - l)$, where the term in parentheses is labor endowment less leisure; then we can rewrite the budget constraint as $\boldsymbol{p}'\boldsymbol{x} + wl = w\bar{L}$. Taxing everything on the left-hand side at a uniform rate would give us a nondistortionary tax, but it would require that we be able to tax leisure, l, separately from the labor endowment, $\bar{L}$. If the government taxes leisure *net of* labor endowment at a constant rate, this feasible tax, applied to $\boldsymbol{p}'\boldsymbol{x} + w(l - \bar{L})$, will raise no revenue. Given that a realistic tax system will involve distortions, how should we choose taxes to minimize deadweight loss? The discussion above suggests that we should avoid high rates of tax on any one commodity, and be especially concerned about taxing commodities with high response elasticities. But this intuition is based on analysis of a tax on a single margin.

Optimal Taxation

The basic optimal tax problem seeks to maximize a representative agent's utility given that the government must raise a certain amount of revenue, R, using proportional commodity taxes, which may include taxes on supplies of factors, such as labor. (As Auerbach and Hines discuss, this is equivalent to minimizing one of the definitions of deadweight loss derived above.) We also assume, initially, that producer prices, $\boldsymbol{q}$, are fixed, and that the household has no truly exogenous income y. This means that the initial budget constraint is $\boldsymbol{q}'\boldsymbol{x} = 0$, and hence that proportional taxes on all commodities raise no revenue, as just discussed. We can therefore arbitrarily set one tax rate equal to zero, say for good 0. (The same analysis would apply for $y > 0$ due to the existence of pure profits, if we assumed that such pure profits could be taxed away, except that the government would then face the task of raising the remaining R-y rather than R.) We also choose this good as numeraire: $p_0 = q_0 = 1$, where p is the price the consumer faces.

The Lagrangian for the problem is $L = V(\boldsymbol{p},0) - \mu[R - (\boldsymbol{p} - \boldsymbol{q})'\boldsymbol{x}]$. We can maximize with respect to $\boldsymbol{p}$ directly; since $dt_i/dp_i = 1$, choosing $\boldsymbol{t}$ is the same as choosing $\boldsymbol{p}$.) First-order conditions are:

$$-\lambda x_i + \mu\left[x_i + \sum_j t_j \frac{dx_j}{dp_i}\right] = 0 \quad \forall i$$

where λ is the marginal utility of income. Using the Slutsky equation, $\frac{dx_j}{dp_i} = s_{ji} - x_i \frac{dx_j}{dy}$, and grouping terms in x_i, we get:

$$(4) \qquad \left[\mu - \left(\lambda + \mu \textstyle\sum_j t_j \frac{dx_j}{dy}\right)\right] x_i + \mu\left(\textstyle\sum_j t_j s_{ji}\right) = 0 \;\Rightarrow\; -\textstyle\sum_j t_j s_{ji} = \frac{\mu-\alpha}{\mu} x_i \qquad \forall i$$

where α can be thought of as the "social" marginal utility of income. As Auerbach and Hines discuss, $\mu \geq \alpha$, i.e., the marginal shadow cost of revenue must be at least as high as the marginal social utility of income – increasing revenue entails additional deadweight loss. To interpret expression (4), which is referred to as the <u>Ramsey rule</u>, note that the term $-\sum_j t_j s_{ji}$ is the excess burden introduced by the additional tax on good i (the $-t\Delta x$ terms). Also, revenue collected from an additional tax on good i, holding utility fixed, is $x_i + \sum_j t_j s_{ji}$. Thus, (4) says that, at an optimum, where small feasible variations in tax instruments have no first-order effects on utility,

$$dDWL/dt_i = (\mu - \alpha)(dR/dt_i + dDWL/dt_i)/\mu \quad \Rightarrow \quad dDWL/dt_i = (\mu - \alpha)(dR/dt_i)/\alpha$$

That is, we should choose taxes so that the marginal deadweight loss associated with each tax is the same proportion of marginal revenue, $(\mu - \alpha)/\alpha$; put another way, the <u>marginal cost of public funds</u> per dollar of revenue, which taking account of the excess burden of taxation, should be equal for all taxes when we are trying to meet a revenue target with minimum deadweight loss. What do optimal taxes look like? Consider a three-commodity model, with two consumption goods and labor as the untaxed numeraire. Stacking the two first order conditions,

$$\begin{pmatrix} s_{11} & s_{21} \\ s_{12} & s_{22} \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} = -\frac{\mu-\alpha}{\mu} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

we invert the matrix in Slutsky terms to obtain the following expression for the ratio of tax rates, in terms of compensated cross-elasticities of demand, ε_{ij},

(5) $$\frac{t_1/p_1}{t_2/p_2} = \frac{\varepsilon_{22}-\varepsilon_{12}}{\varepsilon_{11}-\varepsilon_{21}} = \frac{\varepsilon_{20}+\varepsilon_{12}+\varepsilon_{21}}{\varepsilon_{10}+\varepsilon_{12}+\varepsilon_{21}}$$

where the second version of the expression follows from the condition (implied by the envelope theorem) that $\sum_i p_i s_{ij} = 0$. (To see this, note that $dU/dp_j|_u = \sum_i U_i s_{ij} = \lambda \sum_i p_i s_{ij} = 0$.)

In (5), were we to ignore the cross-elasticities ε_{12} and ε_{21}, the first version would call for tax rates that are inversely proportional to the own demand elasticities. This <u>inverse elasticity rule</u> is consistent with the intuition developed earlier when looking at a single distortion, but it does not hold when the distortions interact (i.e., when $s_{12} \neq 0$). The second version says that we should tax more heavily the good with the smaller value of ε_{i0} – the good that is more complementary to leisure. The logic is that taxing goods 1 and 2 discourages labor supply (since either tax lowers the real wage – the wage relative to the price of consumer goods), so taxing more heavily the good that is complementary to leisure helps lessen this distortion, but at the cost of a new distortion, between goods 1 and 2. This illustrates the general <u>principle of "second-best"</u> – that once we have one distortion, in this case the labor-leisure distortion, we may improve welfare by introducing another distortion, in this case to the margin of choice between goods 1 and 2.

Application: The Taxable Income Elasticity

Sometimes, we observe a taxpayer response that reflects several decisions. For example, when we see taxable income respond to the income tax rate, this reflects not only the decision of how much to work, but also the mix of compensation between taxed and untaxed forms (e.g., fringe benefits like health insurance) and the mix of expenditures between tax-deductible forms (e.g., mortgage interest, charitable contributions, etc.) and non-deductible ones. In such a case, the relevant elasticity will, under certain assumptions, be the overall response. Following Feldstein (1999), consider a household with a utility function, $U(C, l, E, D)$, facing the budget constraint:

$$C + (1\text{-}t)D = (1\text{-}t)[\mathrm{w}(\bar{L}\text{-}l) - E]$$

where E is the portion of compensation taken in the form of goods that are excluded from taxation and D is tax-deductible household spending. We can rewrite the budget constraint as:

$$C = (1\text{-}t)[w(\bar{L}\text{-}l) - E - D]$$

from which it is clear that the relevant elasticity affecting the deadweight loss of an income tax is the elasticity of demand for non-deductible consumption, C, which equals the <u>taxable income elasticity</u> – the elasticity of taxable income $[\mathrm{w}(\bar{L}\text{-}l) - E - D]$ with respect to the tax rate. As Feldstein emphasized, this elasticity may be substantially larger in magnitude than the labor supply elasticity, suggesting more deadweight loss for a given tax rate.

Note that the taxable income elasticity matters here only because we have assumed that tax rates cannot be set separately on E and D. Indeed, the taxable income elasticity depends on the tax structure (Kopczuk, *J. Pub. Econ.* 2005); for example, if no spending were deductible ($D \equiv 0$)

and all compensation taxable ($E \equiv 0$), then the taxable income elasticity would equal the labor supply elasticity. It might seem obvious that such "base broadening" would reduce deadweight loss, as it would leave only labor-supply responses to taxation. That conclusion would hold if labor-supply responsiveness to a change in the tax rate, t, were invariant to base broadening, but this is generally not the case: the labor-supply response to an increase in t, when there are deductions and exclusions, equals the direct response to a reduction in the real wage, $(1-t)w$, plus the cross-effects from reductions in the prices of E and D, $(1-t)$, which effectively increase the real wage. Intuitively, if an individual expects to pay tax on only a portion of marginal labor income (because of additional deductions and exclusions from such income), the labor supply elasticity with respect to a change in tax rates may be smaller in magnitude. In a model where taxes can be set independently on all goods, having a tax base equal to labor income will be optimal precisely when the optimal taxes on the other three goods, c, E, and D, are equal (and hence can all be set to zero in the presence of a tax on labor income).

Application: Internet Sales

Most individual US states rely heavily on sales taxes as a source of revenue. But, until a 2018 Supreme Court decision overturned existing restrictions, states faced stringent limits on the extent to which they could require out-of-state vendors to collect sales tax on remote (e.g., internet and mail-order) purchases by state residents. This meant that residents faced sales tax on direct purchases from retail stores and remote sellers but effectively not on remote purchases from out-of-state sellers. The paper by Einav et al. finds online purchases to be quite sensitive to state tax rates using data from eBay, where potential buyers find out whether a seller is in the same state (and hence required to collect sales tax) only after expressing interest in an item. Their results also show that when a state sales tax is higher, residents purchase *more* on the internet (relative to in-state retail purchases) but *less* from remote in-state vendors.

From an optimal-tax perspective, one might see this case as one with two very closely-related commodities, direct purchases and remote out-of-state purchases, with a high cross-elasticity of demand, where the state can impose tax only on one good. The high demand elasticity of the taxable commodity is likely to limit the extent to which the state might wish to tax it. Also, there may be large efficiency gains from the recent decision relaxing the restrictions states faced on internet sales.