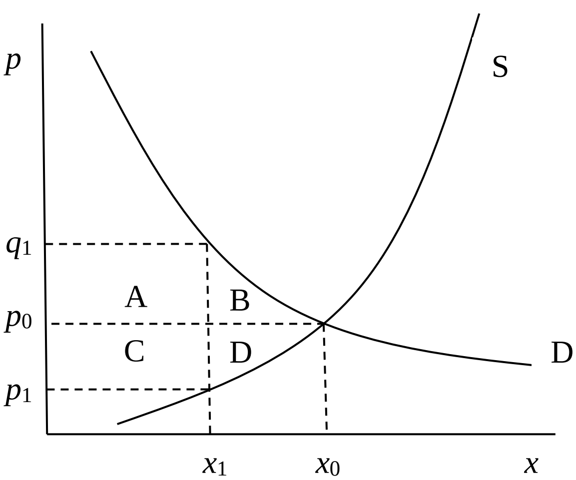
Economics 230a, Fall 2018
Lecture Note 7: Tax Incidence

Tax incidence refers to where the burden of taxation actually falls, as distinguished from who has the legal liability to pay taxes. As with deadweight loss, it is a concept for which the intuition is clear, but for which actual measurement requires the specification of a precise conceptual experiment. It is not enough simply to ask, "What is the incidence of a tax on good x?" We must specify what is done with the revenue, because that will affect incidence, not just through its direct effect on well-being but also through influences on equilibrium product and factor prices.

To illustrate the concept of incidence, consider a small tax introduced in some competitive market, in which the initial price is p_0 and the initial quantity x_0. We introduce a tax, which reduces output, increases the consumer price q, and reduces the producer price p, in the manner shown below. For simplicity, we will assume that the revenue is spent by the government in the same manner that the consumer would spend it. Thus, total demand (by the consumer plus the government) is the same as it would be if the consumer were given the revenue. Starting at an undistorted equilibrium, this is roughly equivalent to compensated demand, since there is no first-order deadweight loss.



The burden of the tax falling on the demand side is the loss of consumer's surplus A+B, while the burden on the producer is the loss of producer's surplus C+D, the sum exceeding revenue (A+C) by the deadweight loss B+D. For a small change starting at a Pareto optimum, the first-order excess burden is small relative to the revenue cost and we can approximate burdens by $x\Delta q$ for the consumer and $-x\Delta p$ for the producer, with the total burdens equal to revenue in this first-order approximation.

The relative burdens on the demand and supply sides will depend on relative elasticities. Defining the term $\hat{z} = d\, log(z)$, and letting the demand and supply elasticities (defined to be non-negative) be η^D and η^S, we know that $\hat{x} = -\eta^D \hat{q} = \eta^S \hat{p}$. Further, if we let $T = (1+\tau)$, where τ is the *ad valorem* tax imposed on the producer price, we have $q = Tp$, so that $\hat{q} = \hat{T} + \hat{p}$. (Also, assuming that we are starting at a value of $\tau = 0$, $\hat{T} = d\tau$.) Thus, setting the two expressions for $\hat{x}$ equal we have $-\eta^D(\hat{T} + \hat{p}) = \eta^S \hat{p} \Rightarrow \hat{p} = \frac{-\eta^D}{\eta^D + \eta^S}\hat{T}$; $\hat{q} = \frac{\eta^S}{\eta^D + \eta^S}$; the ratio of the shares of the burden on consumers and producers is η^S/η^D, i.e., is proportional to the inverse ratio of the respective elasticities – the greater the responsiveness, the lower the burden.

Note: it does not matter whether the tax is imposed on the buyer or the seller, assuming that prices are flexible.

Application: Payroll Tax Reform in Greece

The U.S. social security payroll tax is assessed partially on employees and employers, up to an annual earnings ceiling (now \$128,400). A standard prediction is that the payroll tax's incidence should be the same whether assessed on employees or employers. But this has been difficult to test, as the same rules typically apply to all workers at a given time. A natural experiment arose in Greece, which has a similar system and introduced a reform in 1993 establishing two groups of workers based on whether they were employed prior to that date. Those hired thereafter were subject to a system with a much higher earnings ceiling, making those with wages above the old ceiling subject to higher employer and employee tax rates at the margin than those under the old system. Using a regression discontinuity approach, Saez et al. find that posted earnings do not change across the threshold, meaning that employer payroll taxes (added to posted wages) are borne by employers and employee payroll taxes (subtracted from posted wages) are borne by employees, even 15 years after the reform. This suggests an inability of firms to pay different posted wages to otherwise similar workers based on their tax regime.

Application: The Berkeley Soda Tax

"Sin" taxes, as on tobacco and alcohol, may be desirable to deal with externalities or self-control problems. This logic underlay the city of Berkeley's 2014 adoption of a tax on sweetened beverages, the first in the United States. When thinking about this tax, the question arises to what extent a local jurisdiction can have any impact on outcomes. For a very small jurisdiction imposing an excise tax, one might expect that both supply and demand elasticities would be very high, as both consumers and producers can shift to nearby jurisdictions. Thus, one would expect local purchases to fall, but the relative impact on consumer and producer prices is less obvious. Cawley and Frisvold study Berkeley's soda tax, using San Francisco and diet beverages as controls, finding that less than half of the tax was passed on to consumers. But, if the soda tax fell partially on suppliers, this leaves unresolved whether this fell on profits, wages, rents, etc.

Application: VAT Reform in France

Benzarti and Carloni carry out such an analysis, based on a French VAT reform that lowered the tax rate on restaurant meals. Using similar goods as control groups, the estimate the impact of the lower tax rate on workers, firm owners, consumers, and suppliers of intermediate goods, finding that all groups gained, with firm owners capturing roughly half the short-run gains and consumers about one fourth. Surprisingly, the gains to firms persist even after 30 months.

The Harberger Model

To analyze incidence more fully in terms of factor incomes, we introduce a simple, two-sector general equilibrium model that is a standard tool for incidence analysis. Assumptions:

- Two factors of production, K and L, in fixed overall supply, $\overline{K}$ and $\overline{L}$.
- Two competitive sectors of production, X and Y, with CRS production functions
- One representative consumer who spends factor income on the two goods
- Government introduces small taxes and spends revenue just as the household would

The last assumption implies the changes in <u>total</u> (household plus government) demand will lie along the household's initial indifference curve, because there is no first-order deadweight loss.

Basic Equations

By definition,

$$(1) \qquad \hat{X} - \hat{Y} \equiv -\sigma_D(\hat{q}_X - \hat{q}_Y),$$

where σ_D is the demand elasticity of substitution (defined to be non-negative) and q_i is the consumer price of good i. Also, as a consequence of cost minimization by producers, the derivative of the cost function with respect to the price of a factor is the quantity of that factor used in production; competition implies that price equals marginal cost. It follows that for each production sector i, $\hat{p}_i = \theta_{Li}\hat{w} + \theta_{Ki}\hat{r}$, where w and r are the returns to labor and capital and θ_{ji} is the share of payments to factor j in sector i's costs. For example, $\theta_{LX} = wL_X/p_X X$, where L_X is the amount of labor used in sector X. Note that the shares θ in each sector must sum to 1, so that $\hat{p}_i = \theta_{Li}\hat{w} + (1 - \theta_{Li})\hat{r}$ for each sector. If we subtract this expression for sector Y from that for sector X, we get:

$$(2) \qquad \hat{p}_X - \hat{p}_Y = \theta^*(\hat{w} - \hat{r}),$$

where $\theta^* = (\theta_{LX} - \theta_{LY})$ measures the labor intensity of sector X relative to sector Y. If $\theta^* > 0$, the relative price of good X will rise with an increase in the wage relative to the return to capital.

Finally, we can relate factor returns to the production of goods X and Y. Intuitively, we would expect an increase in production of good X to lead to greater demand and a higher relative factor return to whichever factor sector X uses more intensively than sector Y.

By definition of the production elasticities of substitution, σ_X and σ_Y, $\hat{K}_i - \hat{L}_i = \sigma_i(\hat{w} - \hat{r})$ for $i = X, Y$. For convenience, express K and L as ratios of output, e.g., $k_X \equiv K_X/X$. It follows that

$$(3) \qquad \hat{k}_i - \hat{l}_i = \sigma_i(\hat{w} - \hat{r}) \qquad i = X, Y$$

By the envelope theorem, we know that derivatives of the cost function satisfy $d(rk_i + wl_i) = k_i dr + l_i dw$, so $rdk_i + wdl_i = 0$. This implies that

$$(4) \qquad \left(\frac{rk_i}{p_i}\right)\hat{k}_i + \left(\frac{wl_i}{p_i}\right)\hat{l}_i = \theta_{Ki}\hat{k}_i + \theta_{Li}\hat{l}_i = 0, \; i = X, Y.$$

Finally, note that $L_X + L_Y = l_X X + l_Y Y = \bar{L}$; $K_X + K_Y = k_X X + k_Y Y = \bar{K}$; totally differentiating:

$$(5a) \quad (\hat{l}_X + \hat{X})\lambda_{LX} + (\hat{l}_Y + \hat{Y})\lambda_{LY} = 0; \quad \text{also} \quad (5b) \quad (\hat{k}_X + \hat{X})\lambda_{KX} + (\hat{k}_Y + \hat{Y})\lambda_{KY} = 0$$

where $\lambda_{LX} = L_X / \bar{L}$ is the share of the economy's labor that is used in sector X, and the other terms are defined in the same manner.

Now, substitute (4) into (3) for both sectors to get expressions for $\hat{l}_X$ and $\hat{l}_Y$ and (using the fact that the labor and capital cost shares θ add to 1 for each sector, and that $\lambda_{LX} + \lambda_{LY} = 1$) substitute these expressions into (5a) to obtain:

(6a) $\lambda_{LX}\hat{X} + \lambda_{LY}\hat{Y} = (\lambda_{LX}\theta_{KX}\sigma_X + \lambda_{LY}\theta_{KY}\sigma_Y)(\hat{w} - \hat{r})$

Follow the same procedure to get expressions for $\hat{k}_X$ and $\hat{k}_Y$ to substitute into (5b) to obtain:

(6b) $\lambda_{KX}\hat{X} + \lambda_{KY}\hat{Y} = -(\lambda_{KX}\theta_{LX}\sigma_X + \lambda_{KY}\theta_{LY}\sigma_Y)(\hat{w} - \hat{r})$,

and subtract (6b) from (6a) to obtain:

(7) $\lambda^*\left(\hat{X} - \hat{Y}\right) = \left(a_X\sigma_X + a_Y\sigma_Y\right)\left(\hat{w} - \hat{r}\right) = \bar{\sigma}\left(\hat{w} - \hat{r}\right)$

where $a_i (= \lambda_{Li}\theta_{Ki} + \lambda_{Ki}\theta_{Li})$ is a weighted average of sector i's share of production, as measured by its use of labor and capital, λ_{Ki}, and labor, λ_{Li}, and $\lambda^* (= \lambda_{LX} - \lambda_{KX})$ is positive (negative) if sector X is more (less) labor intensive than sector Y. As expected, a shift in production toward X will increase the relative return to the factor that X uses relatively intensively. The effect will be stronger the smaller is the average elasticity of substitution, $\bar{\sigma}$, because it will take larger changes in factor prices to induce the changes in factor intensities needed to clear factor markets.

Note that (2) and (7) combined provide an expression for the production possibilities frontier,

(8) $\hat{p}_X - \hat{p}_Y = \frac{\lambda^*\theta^*}{\bar{\sigma}}\left(\hat{X} - \hat{Y}\right)$. (Note that sgn($\lambda^*$) = sgn($\theta^*$), so the frontier is convex.)

Equations (1), (2), and (7) are a system in four unknowns, $(\hat{w} - \hat{r})$, $(\hat{p}_X - \hat{p}_Y)$, $\left(\hat{X} - Y\right)$ and $(\hat{q}_X - \hat{q}_Y)$. We add a fourth equation by introducing a tax. We begin with a tax on good X, setting $q_X = T_X p_X$, so that:

(9) $\hat{q}_X - \hat{q}_Y = \hat{p}_X + \hat{T}_x - \hat{p}_Y$

Solving this system of equations, we obtain:

(10) $\hat{p}_X - \hat{p}_Y = -\frac{\sigma_D}{\frac{\bar{\sigma}}{\lambda^*\theta^*} + \sigma_D}\hat{T}_x$; and (11) $\hat{q}_X - \hat{p}_Y = \frac{\frac{\bar{\sigma}}{\lambda^*\theta^*}}{\frac{\bar{\sigma}}{\lambda^*\theta^*} + \sigma_D}\hat{T}_x$

Expressions (10) and (11) say that, if we take good Y as the numeraire (i.e., $\hat{p}_Y = 0$), the burden of the tax is borne on the demand and supply sides of X according to the values of terms that relate to demand and supply. As will now be demonstrated, these expressions are basically equivalent to those derived in the simple partial equilibrium example based on demand and supply elasticities.

Note that the term $\frac{\bar{\sigma}}{\lambda^*\theta^*}$ comes from the expression for the production possibilities frontier, (8). Under profit maximization, $p_X dX + p_Y dY = 0 \Rightarrow \hat{Y} = -\frac{p_X X}{p_Y Y}\hat{X}$, so (8) implies:

(8′) $\hat{X}\left(1 + \frac{p_X X}{p_Y Y}\right) = \frac{\bar{\sigma}}{\lambda^*\theta^*}(\hat{p}_X - \hat{p}_Y)$

With good Y as numeraire, $\hat{p}_Y = 0$ and (8′) may be rewritten:

$$(12) \qquad \frac{\bar{\sigma}}{\lambda^*\theta^*} = \frac{\hat{X}}{\hat{p}_X}\left(1 + \frac{p_X X}{p_Y Y}\right) = \eta_X^S\left(1 + \frac{p_X X}{p_Y Y}\right),$$

where η_X^S is the elasticity of supply of good X with respect to its producer price. Now, consider consumer demand, which is determined by the elasticity of substitution, σ_D, according to (1). Under utility maximization, $dU = q_X dX + q_Y dY = 0 \Rightarrow \hat{Y} = -\frac{q_X X}{q_Y Y}\hat{X}$, so (1) implies:

$$(1') \qquad \hat{X}\left(1 + \frac{q_X X}{p_Y Y}\right) = -\sigma_D(\hat{q}_X - \hat{p}_Y)$$

Again using the fact that good Y is numeraire, (1′) may be rewritten:

$$(13) \qquad \sigma_D = -\frac{\hat{X}}{\hat{q}_X}\left(1 + \frac{q_X X}{p_Y Y}\right) = \eta_X^D\left(1 + \frac{q_X X}{p_Y Y}\right)$$

where η_X^D is the elasticity of demand of good X with respect to its consumer price. Substituting (12) and (13) into the incidence expression (11), and noting that $q_X = p_X$ in the initial equilibrium, we have:

$$(14) \qquad \hat{q}_X - \hat{p}_Y = \frac{\eta_X^S}{\eta_X^S + \eta_X^D}\hat{T}_x,$$

which is precisely the partial-equilibrium expression for the impact on the taxed good's consumer price.

Returning to the general incidence solution, we combine (10) and (2) to obtain:

$$(15) \qquad (\hat{w} - \hat{r}) = -\frac{1}{\theta^*}\frac{\sigma_D}{\frac{\bar{\sigma}}{\lambda^*\theta^*} + \sigma_D}\hat{T}_x.$$

This expression says that the tax on good X, which lowers the producer price of good X, will also lower the ratio w/r if sector X is labor intensive – a tax on the labor-intensive good is relatively bad for labor. How would we measure the share of the burden borne by labor? Intuitively, if w/r is fixed, i.e., $\hat{w} - \hat{r} = 0$, then the tax is borne in proportion to each factor's share of income – since relative rates of return don't change, and factor supplies are fixed, an increase in the consumer price of good X will lower real factor incomes of labor and capital by the same proportion. More generally, we can ask what fraction, ψ, of the tax revenue we would have to give back to labor in order to keep labor's share of *gross* income (including the tax), $\frac{wL + \psi(T_X - 1)p_X X}{wL + rK + (T_X - 1)p_X X}$, constant. Clearly, if w/r doesn't change as the tax is imposed, $\psi = \frac{wL}{wL + rK}$. If $\hat{w} - \hat{r} < (>) 0$, ψ is larger (smaller).

Now, consider a partial factor tax on capital used in sector X, which is how Harberger conceived of the corporate income tax – as an additional tax on capital used in the corporate sector. (Note that a general tax on capital income in this model is simply borne by capital, as capital is in fixed overall supply, so the only interesting factor-tax incidence question involves the differential tax in one sector.) Intuitively, we should expect this tax to have two effects. The first will be to raise the cost of good *X*, just like the excise tax. (The fact that the tax is levied on the production side, rather than on the transaction with the consumer, is irrelevant.) The second will be to discourage the use of capital in production, which should shift the incidence further onto capital. These are sometimes referred to as the excise tax effect and the factor substitution effect of the partial factor tax.

To solve for the effects of this tax, we replace r with rT_{KX} in any equations involving the return to capital in sector *X*. Thus, we get $\hat{p}_X = \theta_{LX}\widehat{w} + \theta_{KX}(\hat{r} + \widehat{T}_{KX})$, which implies:

(2′) $$\hat{p}_X - \hat{p}_Y = \theta^*(\widehat{w} - \hat{r}) + \theta_{KX}\widehat{T}_{KX}$$

This expression picks up the excise tax effect. Also, equation (7) is modified as follows:

(7′) $$\lambda^*\left(\hat{X} - \hat{Y}\right) = a_X\sigma_X\left(\hat{w} - \hat{r} - \hat{T}_{KX}\right) + a_Y\sigma_Y\left(\hat{w} - \hat{r}\right) = \overline{\sigma}\left(\hat{w} - \hat{r}\right) - a_X\sigma_X\hat{T}_{KX},$$

which picks up the factor substitution effect, showing, for example, that even if *X/Y* doesn't change, $\widehat{w} - \hat{r} > 0$.

Solving (1), (2′), and (7′) (and using the fact that consumer prices q and producer prices p are equal – the tax is imposed on producers and hence already included in p), we get the analogue for (15) above:

(15′) $$(\widehat{w} - \hat{r}) = \frac{-\frac{1}{\theta^*}\sigma_D\theta_{KX} + \frac{a_X\sigma_X}{\lambda^*\theta^*}}{\frac{\overline{\sigma}}{\lambda^*\theta^*} + \sigma_D}\widehat{T}_{KX},$$

in which the two terms in the numerator of the right-hand side account for the excise tax effect (which can be positive or negative) and the factor substitution effect (which is non-negative). Harberger showed that under a variety of reasonable assumptions (such as all three elasticities being equal), capital bears exactly 100 percent of the tax. Note that this is the burden on *all* capital – as capital flees the corporate sector, this movement depresses capital returns in the noncorporate sector as well.

Both the realism of the Harberger model for studying corporate tax incidence and the characterization of the corporate income tax as an extra tax on corporate capital are subject to question, as will be discussed in further lectures.

Application: The Incidence of U.S. State Corporate Income Taxes

Although Harberger's analysis applies to national corporate taxes, most U.S. states impose corporate taxes as well. One might expect the incidence of these taxes to fall largely on fixed local factors, such as land and labor, given the high mobility of capital across states. But this fails to account for the fact that firms may have location-specific advantages from locating in a particular state, and also that workers may be mobile. Taking these factors into account, Suárez Serrato and Zidar estimate a structural spatial equilibrium model using annual county-level data, finding that firm owners (as opposed to suppliers of capital) bear roughly 40 percent of local corporate taxes, with the remainder roughly equally divided between land-owners and workers.