

# **TAX ANALYSIS**

**14.471 - Fall 2012**

Market initially in equilibrium at price  $p$ . Government introduces a per-unit tax at rate  $\tau$  per unit of the good purchased.

Problem: Describe effects on producers and consumers.

Market Equilibrium:

$$D(p+\tau) = S(p). \quad (1)$$

$p$  = producer price of the taxed commodity

$\tau$  = "specific tax" of amount  $\tau$

$q$  = the consumer price ( $q = p + \tau$ ).

Does it matter if the tax is collected from producers or consumers? NO.  $D(q) = S(q - \tau)$  vs.  $D(p + \tau) = S(p)$ .

Same equilibrium outcome.

A tax at rate  $\tau$  per unit of the taxed good is an example of a specific tax. The tax amount is independent of the producer price of the good. Examples:

- \* federal (18.4 ¢) and state gasoline excise taxes (average 18.2¢ per gallon nationally, 23.5¢ in MA)
- \* federal tax on cigarettes (\$1.01 per pack)
- \* federal tax on distilled spirits (\$13.50 per gallon).

In contrast, ad valorem sales taxes are levied as a percentage of the value of each transaction. When producer prices rise, the amount of the ad valorem tax also rises. Example:

- \* MA sales tax which is levied at a rate of 6.25% on purchases other than food, prescription drugs, fuel, electricity, and clothing costing less than \$175.00.

To analyze a specific tax: find  $\frac{dp}{d\tau}$  from (1):

$$\frac{dp}{d\tau} = \frac{D'}{S' - D'} \quad (2)$$

Rewrite in terms of elasticities: multiply numerator and denominator by  $q/D = (p+\tau)/D$  which yields

$$\frac{dp}{d\tau} = \frac{\eta_D}{\frac{p+\tau}{p} \eta_S - \eta_D} \quad (3)$$

where  $\eta_D = \frac{D'(p+\tau) \cdot (p+\tau)}{D(p+\tau)}$  and  $\eta_S = \frac{S'(p) \cdot p}{S(p)}$ . Evaluated at the point  $\tau = 0$ , i.e. with no initial taxes, we find

$$\frac{dp}{d\tau} = \frac{\eta_D}{\eta_S - \eta_D}. \quad (4)$$

### Special Cases:

i)  $\eta_S = 0$ : Inelastic supply.  $\frac{dp}{d\tau} = -1$ . The consumer price is unaffected by the tax. “Producers bear the tax.”

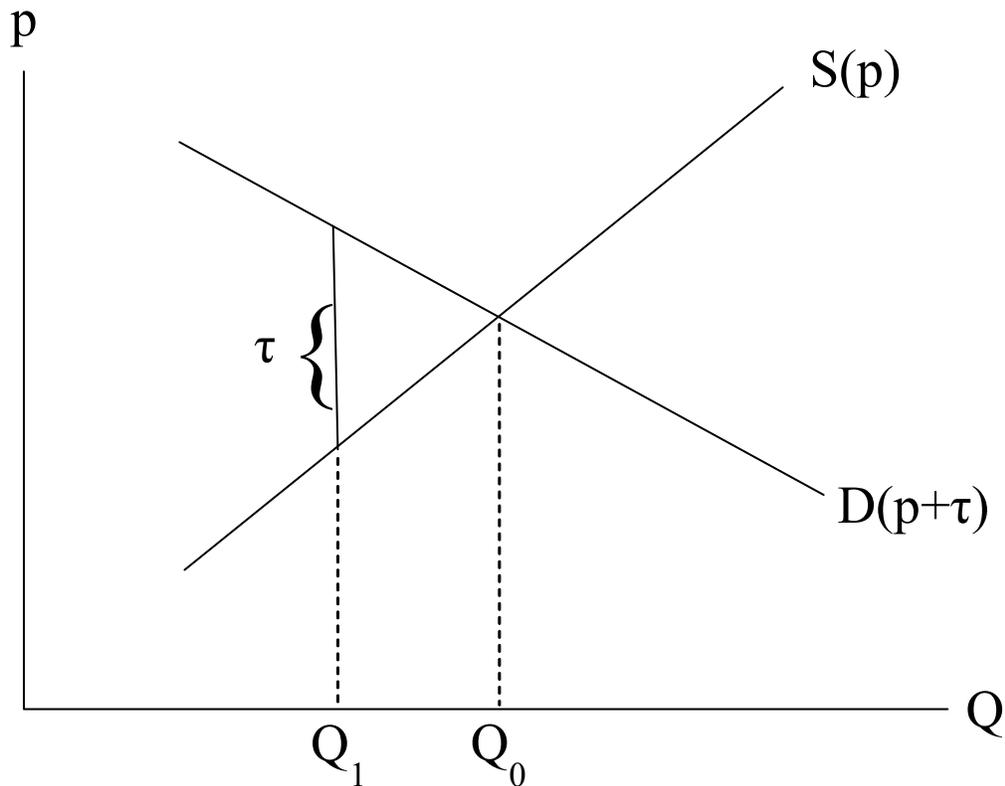
ii)  $\eta_D = 0$ : Inelastic demand.  $\frac{dp}{d\tau} = 0$ . The producer price is fixed, consumer prices rise by the full amount of the tax. “Consumers bear the tax.”

iii)  $\eta_S = \infty$ : Infinitely elastic supply.  $\frac{dp}{d\tau} = 0$ . The consumer price adjusts by the full amount of the tax and again “consumers bear the tax.”

### Efficiency Analysis

The classic diagram depicting the deadweight loss (DWL) of a tax helps motivate the corresponding algebra. The area of the DWL triangle is  $(1/2) * \tau * (Q_0 - Q_1)$ . Note

that this is a positive value; by convention below we set  $DWL < 0$ .



We can write the (negative of) the area of the triangle as:

$$DWL = \frac{1}{2} \tau dQ = \frac{1}{2} \frac{dQ}{d\tau} \cdot \tau^2.$$

This measure is sometimes called the “Harberger Triangle” after Arnold Harberger of Chicago (now UCLA). The term  $dQ$  denotes the change in the quantity of the good traded as a result of the specific tax  $\tau$ . For a

small tax, we can evaluate various expressions at  $\tau = 0$  and set  $d\tau = \tau$  in equation (4), and find

$$dQ = S'(p) \cdot dp = \frac{Q}{p} \cdot \tau \cdot \frac{\eta_S \eta_D}{\eta_S - \eta_D}. \quad (5)$$

DWL is therefore

$$DWL = \frac{\eta_S \eta_D}{\eta_S - \eta_D} \cdot \frac{Q}{p} \cdot \frac{\tau^2}{2}. \quad (6)$$

A more revealing expression normalizes the deadweight burden by the revenue raised,  $R = \tau Q$ :

$$\frac{DWL}{R} = \frac{1}{2} \cdot \frac{\eta_S \eta_D}{\eta_S - \eta_D} \cdot \left( \frac{\tau}{p} \right). \quad (7)$$

We can also find the marginal DWL per unit of revenue raised,  $\partial DWL / \partial R = (\partial DWL / \partial \tau) / (\partial R / \partial \tau)$ , by differentiating (6) and  $R = \tau \cdot Q$ . At  $\tau = 0$ , this yields

$$\frac{\partial DWL}{\partial R} = \frac{\eta_S \eta_D}{\eta_S - \eta_D} \cdot \left( \frac{\tau}{p} \right) = 2 * (DWL / R).$$

## Points to Note About DWL:

- (i) DWL is increasing in both  $\eta_s$  and  $\eta_d$ . If either elasticity is zero, there is no deadweight burden. This is the basis for claims that an efficient tax system places high tax rates on inelasticity supplied and demanded goods.
- (ii) DWL as a share of revenue raised is increasing in the ad valorem tax rate  $\tau / p$ . This is the basis for claims that roughly equal tax rates across goods are desirable and that high tax rates lead to high deadweight losses relative to revenue yield.
- (iii) If  $\eta_s = \infty$  for each of  $N$  goods, and there are no cross-price demand effects for these goods, then if the tax rate is equal for all taxed goods, the sum of the deadweight losses in all markets will be a sales-weighted arithmetic mean of the demand elasticities for the taxed goods. (Verify this!) Raising the tax rate on goods with lower demand elasticities, and lowering the tax rate on goods with higher elasticities, while keeping total revenue constant, will reduce the sum of deadweight burdens. This is a simple form of an "optimal tax" problem.

## Tax Incidence

Incidence analysis describes how the burden of a tax is divided between different market participants. In the case of an excise tax in a single market, this corresponds to the division between producers and consumers. In the absence of pre-existing taxes, we can approximate the producers' burden as the number of units sold times the change in the producer price ( $Qdp$ ), and the consumers' burden as the number of units purchased times the change in the consumer price ( $Qdq$ ). (By convention to make both values negative, since  $dp < 0$  and  $dq > 0$ , we set consumers burden equal to  $-Qdq$ .)

Producers' burden is

$$B^P = S(p) \frac{dp}{d\tau} \tau = \left( \frac{\eta_D}{\eta_S - \eta_D} \right) \cdot Q\tau . \quad (8)$$

Note when  $\eta_D = 0$  the producer bears none of the tax. In the same notation, the consumers' burden is

$$B^C = -D(q) \frac{dq}{d\tau} \tau = -D(q) \left[ \frac{dp}{d\tau} + 1 \right] \cdot \tau = - \left( \frac{\eta_S}{\eta_S - \eta_D} \right) Q\tau \quad (9)$$

When  $\eta_S = 0$  the consumer bears no burden.

## Illustrative Application: Cigarette Taxation

Partial equilibrium tax analysis can be used to compute the efficiency costs and the distributional effects of the cigarette excise tax. In 2012, average retail price of a pack of cigarettes was about \$6.00. U.S. consumers purchased approximately 15 billion packs of cigarettes. In 2000, annual consumption was 24 billion packs; rising tax burdens on tobacco are one of the factors that are widely cited as a factor in this decline.

Federal excise tax: \$1.01/pack

Average state cigarette excise tax: \$1.49

Average state sales tax: \$.20

Average producer price of cigarettes: \$3.30 (includes about \$0.60 paid to distributors/retailers and \$2.70 paid to manufacturer; treat \$2.70 as producer price)

Total excise tax on cigarettes: \$2.70/pack.

In some states the taxes are much higher; in Massachusetts, for example, the total tax is \$3.52. In New York City, the total state and local tax burden, including a \$1.50/pack city tax, is \$5.85/pack. (compare with \$1.60 in Pennsylvania)

National average:  $\tau/p \approx 0.82$  (2.70/3.30).

Tobacco Tax Revenue ( $\tau Q$ ) about \$40 billion/year.

### Elasticity Assumptions:

Demand: Very horizon-sensitive. At "intermediate" (five year?) horizon, central tendency of estimates suggests  $\eta_D \approx -0.50$ . Short-run elasticity of consumption is lower, but of purchasing may be higher.

Example of empirical work: Hu, Ren, Keeler, and Bartlett (1995, Health Economics) estimate an overall elasticity of -0.46, which is the combined effect of an elasticity of -0.22 for consumption by those who smoke and -0.33 for the decision of whether or not to smoke at all.

Supply: If the U.S. were a small open economy, one might assume a supply elasticity of infinity, since the country might be a price-taker in the world market. In practice the supply curve of cigarettes to the U.S. economy is likely to be upward sloping. Limited empirical evidence on supply elasticity, so evaluate DWL with different values of  $\eta_S$ .

$$\underline{\eta_D = -.50, \eta_S = \infty} \quad \underline{\eta_D = -.50, \eta_S = 1}$$

DWL/Revenue	0.210	0.137
Marginal DWL ( $\partial \text{DWL} / \partial R$ )	0.420	0.274
Producers' Burden	0	\$13.3B
Consumers Burden	\$40B	\$26.7B.

## Empirical Issues in the Analysis of Cigarette Taxation:

### 1. Horizon Issues in Price Elasticity of Demand

- Need to distinguish short run vs. long run (effect of changes in quantity for current smokers vs. changes in number of smokers)
- Example: Chris Carpenter & Phil Cook, NBER WP 2007, study National Youth Risk Behavior Survey (YRBS) 1991-2005 using state tax rate changes. 100k high school kids. Price elasticity of smoking participation is -0.56 for high school kids.
- Difference across groups: Price elasticity appears to be largest at lower income levels, so burden of higher taxes may fall more on higher income smokers: Income Quartile: 1 (-1.09 elasticity), 2 (-.70), 3 (-.53), 4(-.39).
- VERY short run: “hoarding” effects (California 50 cent per pack increase example)
- Unintended consequences - increase in residential fires

2. Addiction and Smoking: How do we model demand for a potentially addictive good? Key issue is whether addiction is "rational" (Becker-Murphy JPE 1988) or a manifestation of irrationality ("behavioral economics?")

Addiction models :

$$U_t = u(c_t, S_t, x_t) \quad x = \text{all other good}, U_c > 0, U_s < 0$$

$c_t$  = consumption of addictive good

$$S_t = (1-d) * S_{t-1} + c_t$$

Consumers are forward looking, recognize future costs but still decide to smoke; key prediction is that anticipated future increase in cigarette prices should reduce smoking today

Models that allow for irrationality: Gruber-Koszegi (QJE 2001, JPubE 2004)

Time-inconsistent consumers do not recognize future health costs. Example of model could be Laibson-style hyperbolic discounting

$$V_t = U_t + b * \sum_k d^k * U_{t+k}$$

Tax policy may help potential smokers commit not to smoke and therefore may have very different welfare effects.

### 3. Imperfect Competition in Cigarette Market

Barnett, Keeler, and Hu (JPubE 1995) show that retail price reaction to a federal tax is larger than reaction to a state or local tax, suggesting price coordination. This means changes in tax rates could have effects on degree of market power or facilitate collusion - with additional welfare effects.

Geography may matter: Harding, Leibtag, Lovenheim 2010 paper: pass-through of beer and cigarette taxes to consumer prices depends on proximity of retail establishment to state border. Nearby no-tax state reduces passthrough (<50% near borders).

### 4. Smuggling.

If higher tax rates increase smuggling then effect on revenue will be estimated incorrectly (this matters for optimal tax calculation). Also we'll mis-estimate the demand effects and the welfare effects. (Gruber, Sen, Stabile, JHE, 2003): huge interstate differences in tax rates, estimate of 6% of total U.S. cigarette consumption not paying taxes (smuggled), costs states about \$1 billion/year in foregone revenue (role of Indian reservations); note in Canada when taxes rose sharply, "legal exports" rose from 1.5% of sales to 50% of sales.

(Cigarettes were being shipped to US and illegally re-imported).

How elastic is smuggling, what are revenue effects, how does smuggling affect estimates of price elasticity?

## 5. Distribution of Burden from Tobacco Taxes

Consumption of cigarettes accounts for a larger budget share at low income levels. Poterba AER 1989:

Quantile	Income	Consumption
Lowest	4.6%	2.2%
Second	2.0	2.0
Third	1.3	1.7
Fourth	0.9	1.3
Fifth (top)	0.5	0.7

## 6. Externalities Associated with Smoking

How large are the social costs associated with cigarette smoking? What are they? U.S. Center for Disease Control estimates that smoking in the US costs \$96B/year in health outlays, and \$97B/year in lost productivity (> \$12/pack). This is average. What is the marginal per pack cost? How much is external?

- Costs in group insurance and social insurance programs: \$0.33 per pack 1995 CRS estimate
- Secondhand smoke external effects: very controversial, upper estimate \$0.70/pack.
- Low-birthweight babies: \$0.40-\$0.70 per pack (huge costs for small set of individuals)
- Reductions in workplace productivity of non-smokers, cleaning charges, fires...
- Does taxing cigarettes improve matters? Substitution of Higher Tar and Nicotine Cigarettes (Evans and Farelly 1998 RAND – level of inhaled tar / nic changes very little when cigarette tax increases)

### Reminder About Pre-Existing Distortions:

Our analysis of deadweight burdens has assumed that markets are "neoclassical" except for the tax we are analyzing. But pre-existing market distortions such as nominal price rigidities, unionized workers who raise production costs and create factor market distortions, uninformed consumers, imperfect competition can compound (or attenuate) the efficiency effects of taxes. In some cases small taxes that would have only second-order effects may have first-order effects with such distortions.

## Multi-market Analysis of Excess Burden:

Goulder and Williams (2003 JPE) point out that a change in an excise tax in one market can matter through its effect in other markets (i.e. by changing demands).

Original Harberger analysis:

$$DWL = .5 * t_k^2 * \frac{dX_k}{dt_k} + \sum_{j \neq k} t_j t_k \frac{dX_j}{dt_k}$$

Most analyses ignore effects in other markets, usually because we lack information on  $(dX_j/dt_k)$ . BUT these effects can be important: example of effect of an excise tax on the real product wage. Consumer price index  $q^*$  is a budget-share ( $s_j$ ) weighted average of consumer prices:

$$q^* = \sum_j s_j * p_j * (1 + t_j)$$

Real product wage is  $w(1-t_y)/q^*$ . Raising an excise tax on a good with a non-trivial budget share can distort the labor market by reducing the real product wage.

Is this effect important? Goulder and Williams suggest that the under-estimate for range of elasticities we have used could be on the order of 50%.



Finkelstein (2009 QJE):

- compares rate of change of bridge & tunnel tolls before and after adoption of "EZ Pass"
- faster increases afterward, consistent with less consumer or voter visibility once charged automatically

Implication: Marshallian demand may not be  $x(p + t, y)$  but rather is  $x(p, t, y)$ . One example:  $x(p + \theta t, y)$ . In this case

$$\frac{dp}{d\tau} = \theta \frac{\eta_D}{\eta_S - \eta_D}.$$

where  $\eta_D$  denotes  $x_p^*(p+t)/x$ , the "usual" demand elasticity, in this case defined as the elasticity with respect to the producer price.  $\theta \approx 0.35$  in posted prices example, near zero for beer excise tax.

Key implications:

- \* Incidence: If a tax has low salience ( $\theta \approx 0$ ) then consumers bear most of the burden (low salience is like low elasticity)
- \* Legal Incidence May Matter - which side of the market

\* Efficiency: Budget constraint still must be satisfied. If the taxed good doesn't respond to the tax, what does? There must be adjustments elsewhere - THOSE are the source of the distortions. Empirical analysis of efficiency requires evidence of what DOES respond to the tax.

### Other Behavioral Insights for Tax Analysis:

Taxpayers may not understand marginal tax rates at particular points on tax schedule: confusion of average tax rate and marginal (what implications for behavior)?

Contrast (if possible) behavioral reasons for failure to adjust to taxes with rational "adjustment costs" or fixed costs of search - may affect long-run steady state effects, but look similar in short run.

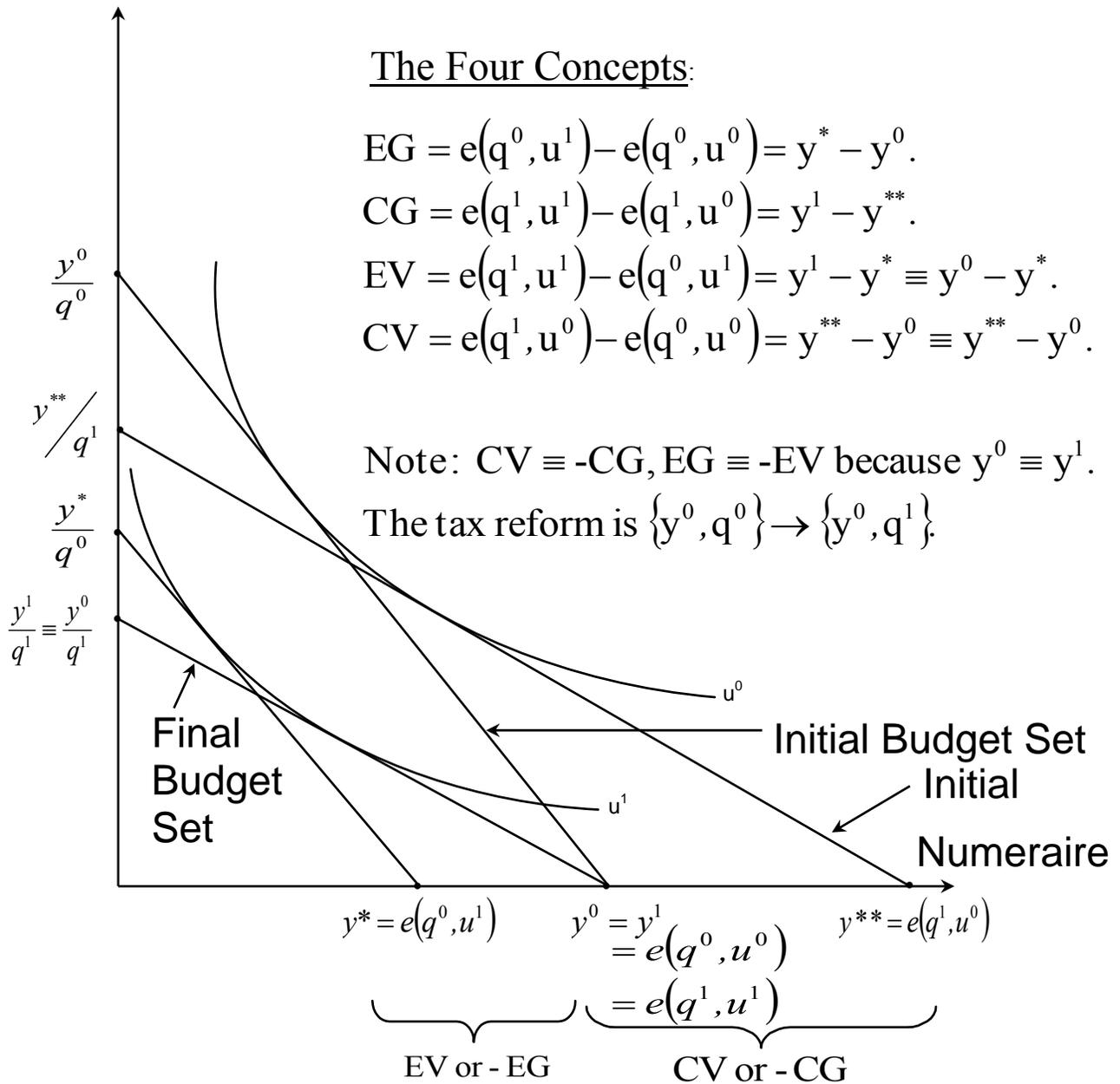
## Beyond Harberger Triangles: Exact Welfare Analysis

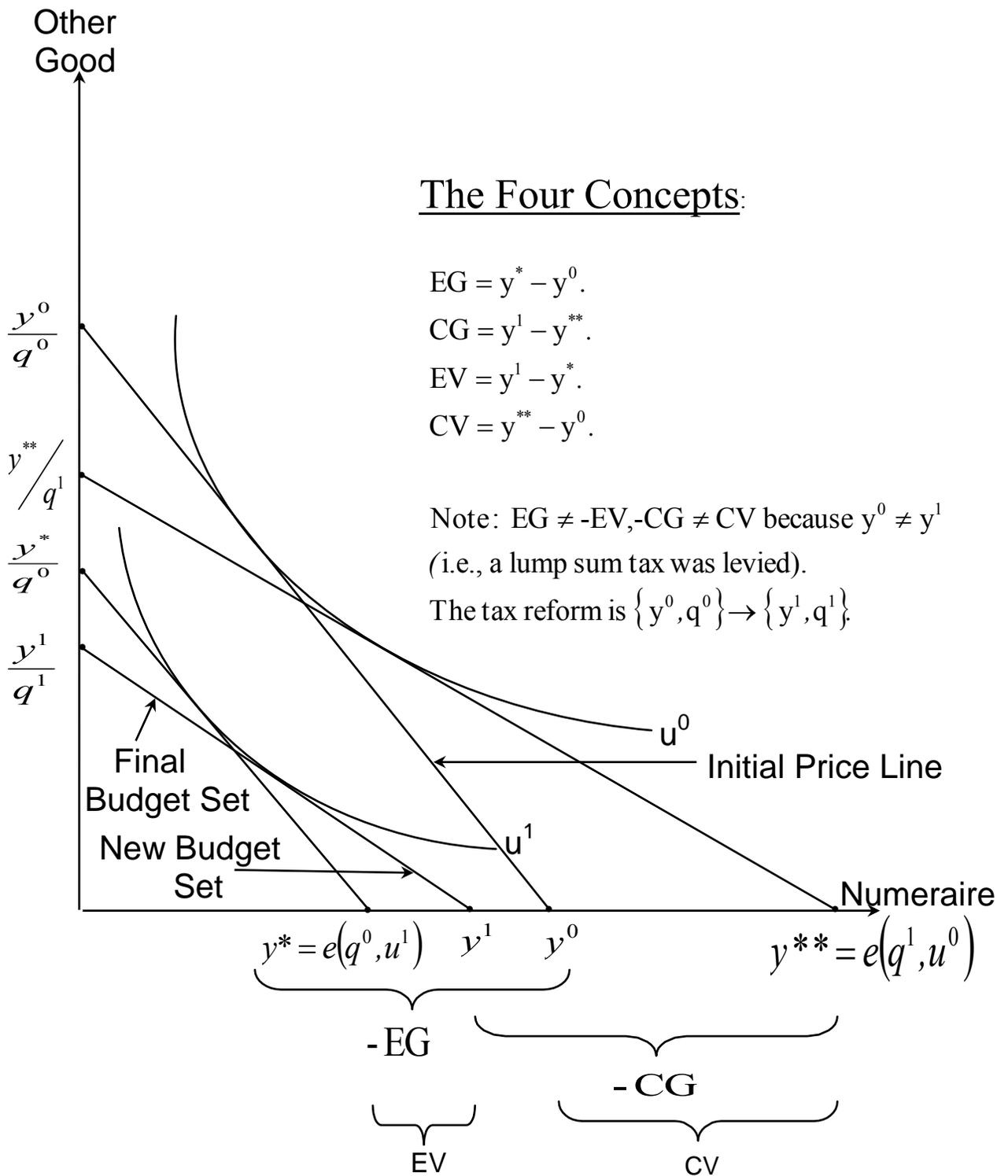
The consumers' burden measures above are an approximation to a utility-based measure of the costs of a tax-induced price change. They suffer from all the limitations of Marshallian consumer surplus measures. A more theoretically grounded measure such as the compensating variation can overcome these shortcomings:

$$CV(q_1, q_0, u_0) = e(q_1, u_0) - e(q_0, u_0). \quad (10)$$

In this expression  $e(q_1, u_0)$  denotes the expenditure function evaluated at utility level  $u_0$  and prices  $q_1$ . The CV measures the cost to a consumer of a shift in prices from  $q_0$  to  $q_1$ . There are several related measures that can be used to evaluate consumer welfare, as this diagram shows:

Other  
Good





Diamond and McFadden (JPubEc 1974) adapt the compensating variation to measure the deadweight burden of a tax-induced price change from  $q_0$  to  $q_1 = q_0 + \tau$ . They define the DWL of a tax vector  $\tau$  as

$$CV(q_1, q_0, u_0) - R(q_1, q_0, u_0) \quad (11)$$

where  $R(q_1, q_0, u_0)$  denotes the compensated revenue associated with a vector of specific taxes  $\tau$ .  $R(q_1, q_0, u_0)$  is defined using Hicksian rather than Marshallian demand functions:

$$R(q_1, q_0, u_0) = \sum_j \tau_j * h_j(q_1, u_0). \quad (12)$$

To implement the Diamond-McFadden approach, one needs estimates of the expenditure function. Most empirical work delivers estimates of Marshallian demand curves but not the underlying utility functions that are needed to construct  $e(q_1, u_0)$  and  $h(q_1, u_0)$ .

Hausman (AER 1981) demonstrates that in many cases it is possible to recover preferences from commonly-used demand specifications, and he then argues that there is no excuse for using “triangle approximations” in place of exact expressions for welfare loss. Consider as an example a linear demand curve for some good:

$$x_1 = \gamma + \alpha q_1 + \delta y. \quad (13)$$

where  $q_1$  is the consumer price and  $y$  is household income. What preferences would generate this demand function? We need to solve the partial differential equation that follows from Roy's Identity

$$-\frac{\frac{\partial v(q, y)}{\partial q_1}}{\frac{\partial v(q, y)}{\partial y}} = x_1(q, y) = \gamma + \alpha q_1 + \delta y \quad (14)$$

to find  $v(q, y)$ , the indirect utility function. In this case,

$$v(q, y) = e^{-\delta q_1} \left[ y + \frac{\alpha}{\delta} q_1 + \frac{\alpha}{\delta^2} + \frac{\gamma}{\delta} \right]. \quad (15)$$

The analogous expenditure function is

$$e(q, u) = e^{\delta q_1} u - \frac{\alpha}{\delta} q_1 - \frac{\alpha}{\delta^2} - \frac{\gamma}{\delta}. \quad (16)$$

It is straightforward to check that (16) implies (13).

Another common demand specification is the double logarithmic model:

$$\log x_1 = \gamma + \alpha \log q_1 + \delta \log y. \quad (17)$$

In this case, the indirect utility function is

$$v(q, y) = -e^\gamma \cdot \frac{q_1^{1+\alpha}}{1+\alpha} + \frac{y^{1-\delta}}{1-\delta}. \quad (18)$$

The corresponding expenditure function is

$$e(q, u) = \left[ (1-\delta) \left( u + e^\gamma \frac{q_1^{1+\alpha}}{1+\alpha} \right) \right]^{\frac{1}{1-\delta}}. \quad (19)$$

From estimates of linear or double-log demand curves, it is straightforward to find the associated preferences and then to compute welfare measures such as CV and the Diamond-McFadden measure of efficiency loss. But what if the demand curves that fit the data best do not take these convenient functional forms?

The foregoing methods rely on parametric assumptions about the structure of demand curves to obtain exact analytical solutions for preferences. These can in turn be used to obtain analytical expressions for measures of

consumer burden, such as the equivalent and the compensating variation.

Two Challenges:

\*Household Heterogeneity. This can be addressed either by modeling preferences as a function of observable household attributes (demographic attributes would be the natural example) or in some cases by allowing for unobserved heterogeneity in the estimation process.

\*"Easy" Functional Forms May not Fit Well. Simple linear and log-linear forms may not accurately capture the structure of demand. In that case, it may be attractive to use a nonparametric specification for the demand model. While that does not permit an exact solution to the problem of recovering preferences, it is still possible to develop a numerical estimate of the efficiency cost of taxes. Hausman and Newey (Econometrica 1995) demonstrate how this can be done.

A nonparametric estimate of the demand curve can be used to estimate the local change in demand for a small change in price at each price. This can in turn be integrated over the path of price change to provide an estimate of the aggregate compensating variation for a given tax policy change.

When we have only nonparametric estimates of  $x(q, y)$  we cannot find  $e(q, u)$  by integration but we can still find numerical approximations to the expenditure function evaluated at various price vectors. Recall that  $e(q_0, u_0) = y_0$ . We can approximate  $e(q_1, u_0)$  as:

$$e(q_1, u_0) = e(q_0, u_0) + \frac{\partial e}{\partial q}(q_0, u_0) \cdot (q_1 - q_0). \quad (20)$$

Since

$$\frac{\partial e}{\partial q}(q_0, u_0) = h(q_0, v(q_0, y_0)) = x(q_0, y_0) \quad (21)$$

a one-step approximation to  $e(q_1, u_0)$  is

$$e(q_1, u_0) = e(q_0, u_0) + x(q_0, y_0) \cdot (q_1 - q_0). \quad (22)$$

For large changes in  $q$ , where income effects may become important, we would prefer a closer approximation. Let's divide  $(q_1 - q_0)$  into  $N$  intervals and define

$$\Delta = (q_1 - q_0) / N. \quad (23)$$

We will now evaluate  $e(q_0 + j \cdot \Delta, u_0)$  for  $j=1, \dots, N$  to obtain a better approximation.

Step 1:

$$e(q_0 + \Delta, u_0) = e(q_0, u_0) + x(q_0, y_0) \cdot \Delta = y_0 + x(q_0, y_0) \cdot \Delta$$

The key insight arises in moving from this single step to a multi-step iteration:

Step 2:

$$\begin{aligned} e(q_0 + 2\Delta, u_0) &= e(q_0 + \Delta, u_0) + h(q_0 + \Delta, u_0) \cdot \Delta \\ &= e(q_0, u_0) + x(q_0, y_0) \Delta + h(q_0 + \Delta, u_0) \Delta. \end{aligned}$$

The challenge is to find  $h(q_0 + \Delta, u_0)$ .

We know from step 1 that

$$e(q_0 + \Delta, u_0) = y_0 + x(q_0, y_0) \cdot \Delta. \quad (24)$$

Define the new level of required income as  $\tilde{y}_{(1)}$ :

$$\tilde{y}_{(1)} \equiv y_0 + x(q_0, y_0) \cdot \Delta. \quad (25)$$

Since  $v(q_0 + \Delta, \tilde{y}_{(1)}) = u_0$ , we know that

$$h(q_0 + \Delta, u_0) = x(q_0 + \Delta, \tilde{y}_{(1)}). \quad (26)$$

This provides the key iterative step:

$$e(q_0 + 2\Delta, u_0) \equiv \tilde{y}_{(2)} = \tilde{y}_{(1)} + x(q_0 + \Delta, \tilde{y}_{(1)}) \cdot \Delta$$

At each point in the iteration we evaluate the Marshallian demand curve at a different point based on the previous steps, which generate  $\tilde{y}_{(j)}$ .

The general algorithm for finding  $e(q_0 + N \cdot \Delta, u_0) = \tilde{y}_{(N)}$  can now be written:

$$\tilde{y}_{(0)} = y_0$$

$$\tilde{y}_{(j+1)} = \tilde{y}_{(j)} + x(q_0 + j \cdot \Delta, \tilde{y}_{(j)}) \cdot \Delta \quad j = 1, \dots, N$$

By construction

$$\tilde{y}_{(N)} \equiv e(q_0 + N \cdot \Delta, u_0) = e(q_1, u_0). \quad (27)$$

This numerical procedure can be applied regardless of whether  $x(q,y)$  is estimated parametrically or non-parametrically - all that is needed is an evaluation of  $x(q,y)$  at a sequence of  $N$  points.

MIT OpenCourseWare  
<http://ocw.mit.edu>

14.471 Public Economics I  
Fall 2012

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.