Economics 131
Section Notes
GSI: David Albouy

Extensions of Consumer Theory, Taxes in Equilibrium, Welfare, and Optimal Taxation

1 Extensions of Consumer Theory

1.1 Utility Maximization Problem

Recall the problem from the "Micro-Review" notes, where individuals maximize utility subject to budget constraint $px \leq M + wL$ and time constraints $l + L \leq T$. Setting the inequalities to equality, combining the two constraints, and rearranging a bit we get the standard **utility maximization problem** (UMP for short)

$$\max_{l,x} U(l,x) \quad \text{s.t. } px + wl - wT = M \tag{UMP}$$

The solution to this problem is typically found by writing the Lagrangean

$$\mathfrak{L}(l,x,\alpha) = U(l,x) + \alpha(M + wT - px - wl)$$

and taking the first order conditions

$$\frac{\partial \mathfrak{L}}{\partial l} = \frac{\partial U}{\partial l} - \hat{\alpha} w = 0$$
$$\frac{\partial \mathfrak{L}}{\partial x} = \frac{\partial U}{\partial x} - \hat{\alpha} p = 0$$
$$\frac{\partial \mathfrak{L}}{\partial \alpha} = M + wT - px^D - wl^D = 0$$

Solving yields the Lagrange multiplier $\hat{\alpha} = \hat{\alpha}(w,p,M)$ and the demand functions

$$x^D(w,p,M) \quad l^D(w,p,M)$$

To be more general we call these the **uncompensated (or Marshallian or Walrasian) demand functions**. These functions are "uncompensated" since price changes will cause utility changes: a situation that does not occur with compensated demand curves.

Substituting these solutions back into the utility function, the maximand, we get the actual utility achieved as a function of prices and income. This function is known as the **indirect utility function**

$$V(w,p,M) \equiv U\left[x^D(w,p,M), l^D(w,p,M)\right] \tag{Indirect Utility Function}$$

This function says how much utility consumers are getting when they face prices (w,p) and have unearned income M. As we saw in the "Shadow Prices" notes the derivative of the indirect utility function should equal the Lagrange multiplier α on the budget constraint

$$\frac{\partial V}{\partial M} = \hat{\alpha}$$

An interesting fact known as **Roy's Identity** also tells us that the the other two derivatives come close to giving us the uncompensated demand functions

$$\frac{\partial V}{\partial p} = -\hat{\alpha} x^D, \quad \frac{\partial V}{\partial p} = -\hat{\alpha}\left(l^D - T\right) = \hat{\alpha} L^S \tag{Roy's ID}$$

A basic (albeit somewhat flawed) intuition for this identity is straightforward: if p goes up by one dollar then the consumer will lose x^D number of dollars, which each have utility value $\hat{\alpha}$, so that utility drops by

$\hat{\alpha} x^D$. If w goes up by one dollars, the consumer gains L^S dollars, since the consumer is a net supplier of labor to the market, and utility increases by $\hat{\alpha} L^S$. The more correct intuition is that x^D and L^S may actually be changing in response to the price changes, but so long as individuals are maximizing throughout, the utility changes brought about by these changes adds up to zero[1]. Note that with just the indirect utility function we can get $\hat{\alpha}, x^D, L^S$, and $l^D = T - L^S$, just by taking various derivatives of V and combining them appropriately, e.g. $x^D = -(\partial V/\partial p) / (\partial V/\partial M)$.

Example 1 ***Quasilinear utility*** *takes a form where one of the goods consumed enters linearly into the utility function, which in this case we take to be* $U(l,x) = l + v(x)$ *where* $v'(x) < 0$ *and* $v''(x) > 0$. *(The reader may like to try the case where* $T = 1$ *and* $v(x) = \sqrt{x}$*) It is convenient to treat* l *as the "numeraire" setting its price to one,* $w = 1$*, since only relative prices matter this is not an issue. Substituting in the budget constraint* $px = M + L$ *into the time constraint gives* $l = T - L = T + M - px$ *which is then substituted into the utility function to get* $T - px - M + v(x)$. *To find* $\max_x \{T - px - M + v(x)\}$ *we find the FOC* $-p + v'(x^D) = 0$ *which solving gives* $x^D(1,p,M) = (v')^{-1}(p)$ *where* $(v')^{-1}$ *is the inverse function of* v'. *Leisure is given by* $l^D(1,p,M) = T + M - px^D = T + M - p \cdot (v')^{-1}(p)$. *Notice that* x^D *does not depend on income* M *or* T*, changes in these quantities only affect labor supply. Indirect utility is given by* $V(1,p,M) = l^D + v(x^D) = T + M - p \cdot (v')^{-1}(p) + v\left[(v')^{-1}(p)\right]$. *Since we did not use a Lagrangean we can find the shadow price of money in utility terms as* $\hat{\alpha} = \partial V/\partial M = 1$. [2] *(The reader is invited to check Roy's Identity for themselves in the case where* $v(x) = \sqrt{x}$*)*

1.2 Expenditure Minimization Problem

The consumer problem can be approached in a different way which produces some useful tools. Instead of maximizing utility given a certain income, imagine how much income it would take to achieve a certain level of utility. In other words consider the following **expenditure minimization problem** (EMP for short), which as always take prices as given

$$\min_{l,x} px + wl - wT \quad \text{s.t. } U(l,x) = u \tag{EMP}$$

This problem looks very much like the UMP above except that the objective function and constraint have been switched around. We wish to minimize the income $M = px + wl - wT$ needed achieve a fixed level of income u, for given prices (w,p). Our third parameter in parameter in this problem (after w,p) is no longer M, but u. This problem can typically be solved by writing the Lagrangean

$$\mathfrak{L}(l,x,\gamma) = px + wl - wT + \gamma[u - U(l,x)]$$

[1] Here's the proof for the first identity. Differentiating V with respect to p we get

$$\frac{\partial V}{\partial p} = \frac{\partial U}{\partial l}\frac{\partial l^D}{\partial p} + \frac{\partial U}{\partial x}\frac{\partial x^D}{\partial p} = \hat{\alpha} w \frac{\partial l^D}{\partial p} + \hat{\alpha} p \frac{\partial x^D}{\partial p} = \hat{\alpha}\left(w\frac{\partial l^D}{\partial p} + p\frac{\partial x^D}{\partial p}\right)$$

Where the second equality comes from substituting in the FOC. Differentiating the budget constraint with respect to p gets

$$w\frac{\partial l^D}{\partial p} + p\frac{\partial x^D}{\partial p} + x^D = 0 \Rightarrow w\frac{\partial l^D}{\partial p} + p\frac{\partial x^D}{\partial p} = -x^D$$

Substituting in the right hand side of this equation into the parentheses above finishes the proof. The proof using w is quite similar (you can do it yourself).

[2] Roy's identity is more difficult to check

$$\frac{\partial V}{\partial p} = -(v')^{-1}(p) - p\frac{1}{v''(p)} + v'\left[(v')^{-1}(p)\right]\frac{1}{v''(p)} = -(v')^{-1}(p) - p\frac{1}{v''(p)} + p\frac{1}{v''(p)} = -(v')^{-1}(p) = -\alpha x^D$$

and taking the following first order conditions

$$\frac{\partial \mathfrak{L}}{\partial l} = w - \hat{\gamma}\frac{\partial U}{\partial l} = 0$$
$$\frac{\partial \mathfrak{L}}{\partial x} = p - \hat{\gamma}\frac{\partial U}{\partial x} = 0$$
$$\frac{\partial \mathfrak{L}}{\partial \gamma} = u - U\left(l^{CD}, x^{CD}\right) = 0$$

The first two FOC are quite similar to above replacing $\hat{\gamma}$ with $1/\hat{\alpha}$, but the third constraint corresponding to the constraint is very different. Solving these three equations in the three unknowns yields the Lagrange multiplier $\hat{\gamma} = \hat{\gamma}(w, p, u)$, the shadow price in dollars of having to provide an extra unit of utility to this consumer, as well as the **compensated demands**

$$l^{CD}(w, p, u) \quad x^{CD}(w, p, u)$$

which are a function now of the required utility u, not income M. Compensated labor supply can be written as just $L^{CS}(w, p, u) = T - l^{CD}(w, p, u)$. Note here that even though utility stays the same, quantities demanded will change as w and p since the individual is trying to minimize her consumption. Levels of required income M are assumed to automatically adjust to let make sure that individual can still achieve utility u, although not necessarily the bundle of goods previously consumed. The individual is fully compensated for changes in price which could otherwise affect her utility if M were held fixed.

Substituting in the solutions back into the objective function, the minimand, we get the **expenditure function**

$$e(w, p, u) \equiv px^{CD}(w, p, u) + wl^{CD}(w, p, u) - wT \qquad \text{(Expenditure Function)}$$

which is precisely the amount M needed to maintain utility level u, for given prices w, p. As usual we can differentiate e with respect to u to get $\partial e/\partial u = \hat{\gamma}$. A more interesting result known as **Shepard's Lemma** (the analogue to Roy's Identity) is that

$$\frac{\partial e}{\partial p} = x^{CD} \quad \frac{\partial e}{\partial w} = l^{CD} - T = -L^{CD} \qquad \text{(Shepard's Lemma)}$$

Again the (somewhat misleading) intuition for this is clear. If p changes by a small amount then x^{CD} will not change by very much and so the increased cost of consuming these units is precisely x^{CD}. The better intuition is that there are changes in x^{CD} and L^{CD}, but because of optimizing behavior, the consumer avoids spending any more than x^{CD}, although since she was optimizing before she cannot avoid spending any less.[3]

Example 2 *Continuing with quasilinear utility, use the utility constraint to get $l = u - v(x)$ and substitute in so that we solve $\min_x \{px + u - v(x) - T\}$. The FOC is $p - v'\left(x^{CD}\right) = 0$ which implies $x^{CD}(1, p, u) = (v')^{-1}(p) = x^D(1, p, u)$, and so compensated demand in this special case is the same as uncompensated demand. Compensated leisure demand is not the same as $l^{CD}(1, p, u) = u - v\left(x^{CD}\right) = u - v\left[(v')^{-1}(p)\right]$ which is quite different from l^D. The expenditure function is given by the equation. $e(w, p, u) = px^{CD} + l^{CD} - T = p \cdot (v')^{-1}(p) + u - v\left[(v')^{-1}(p)\right] - T$. A remarkable fact due to the shape of the utility function is that the expenditure function is almost exactly negative of the indirect utility function adding in utility and subtracting income $e(1, p, u) = -V(1, p, M) + M + u$. In this case an extra dollar of income will produce one more unit of utility just as requiring u to rise one unit requires one dollar.*

[3] The proof for this is similar to that for Roy's Identity. Differentiating the expenditure function with respect to p and subsequently substituting in the FOC we get

$$\frac{\partial e}{\partial p} = p\frac{\partial x^{CD}}{\partial p} + w\frac{\partial l^{CD}}{\partial p} + x^{CD} = \hat{\gamma}\frac{\partial U}{\partial x}\frac{\partial x^{CD}}{\partial p} + \hat{\gamma}\frac{\partial U}{\partial l}\frac{\partial l^{CD}}{\partial p} + x^{CD} = \hat{\gamma}\left(\frac{\partial U}{\partial x}\frac{\partial x^{CD}}{\partial p} + \frac{\partial U}{\partial l}\frac{\partial l^{CD}}{\partial p}\right) + x^{CD}$$

Differentiating the utility constraint with respect to p gives

$$\frac{\partial U}{\partial l}\frac{\partial l^{CD}}{\partial w} + \frac{\partial U}{\partial x}\frac{\partial x^{CD}}{\partial w} = 0$$

so the term in parentheses in the first equation is zero, which yields the desired result. The proof is similar with respect to w.

1.3 Slutsky's Equation

A very important relationship between uncompensated demands and uncompensated demands can be derived by first noting that the following identity holds

$$x^{CD}(w,p,u) = x^D(w,p,e(w,p,u))$$

Obviously if a consumer is given income $M = e(w,p,u)$ and solves the UMP, they will at that point get the exact same demand as in the EMP since the prices are exactly the same. Differentiating this identity with respect to p we get

$$\frac{\partial x^{CD}}{\partial p} = \frac{\partial x^D}{\partial p} + \frac{\partial x^D}{\partial M}\frac{\partial e}{\partial p} = \frac{\partial x^D}{\partial p} + \frac{\partial x^D}{\partial M}x^{CD}$$

where the second equation uses Shepard's Lemma. Using the fact that $x^{CD} = x^D$ and rearranging gives us the **Slutsky equation**

$$\frac{\partial x^D}{\partial p} = \frac{\partial x^{CD}}{\partial p} - \frac{\partial x^D}{\partial M}x^D \qquad \text{(Slutsky)}$$

The first term of the right hand side $\frac{\partial x^{CD}}{\partial p}$ is always negative[4] and is commonly known as the **substitution effect**. The second term $-\frac{\partial x^D}{\partial M}x^D$ known as the **income effect** is typically, albeit not always negative, depending on whether $\frac{\partial x^D}{\partial M} > 0$, i.e. x is a normal good.[5]

Example 3 *With quasilinear utility we saw $x^D = x^{CD}$ and this can be attributed partly to the fact that there is no income effect $\partial x^D/\partial M = 0$ and so $\partial x^D/\partial p = \partial x^{CD}/\partial p = 1/v''\left[(v)^{-1}(p)\right] = 1/v''\left(x^D\right)$. For leisure demand $\partial l^D/\partial M = 1$ so all additional income goes to "buying" leisure.*

1.4 Some Related Concepts in the Firm's Problem

IThe firm's **profit maximization problem (PMP)** is given by

$$\max_{x,L} px - wL \quad \text{s.t. } x = f(L)$$

which is solved via the first order conditions given in the "Shadow Prices" notes (2.2.1) to yield the consumption supply function $x^S(w,p)$ and labor demand function $L^D(w,p)$. These can be substituted into the maximand to get the **profit function**

$$\Pi(w,p) = px^S(w,p) - wL^D(w,p) \qquad \text{(Profit Function)}$$

A result known as **Hotelling's Lemma** gives us results similar to Shepard's Lemma that[6]

$$\frac{\partial \Pi}{\partial p} = x^S \quad \frac{\partial \Pi}{\partial w} = -L^D(w,p) \qquad \text{(Hotelling's Lemma)}$$

[4] The fact that $\frac{\partial x^{CD}}{\partial p} < 0$ follows from the fact that $e(w,p,u)$ is a concave function in p and so it's second derivative is negative $\frac{\partial^2 e}{\partial p^2} = \frac{\partial}{\partial p}\frac{\partial e}{\partial p} = \frac{\partial}{\partial p}x^{CD} < 0$. For a proof of why $e(w,p,u)$ is concave please consult a higher level microeconomics textbook, such as Silberberg (1999) or Mas-Colell, Whinston and Green (1995). A relatively intutive demonstration of this fact can be made from a simple graph.

[5] For labor supply $L^{CS}(w,p,u) = L^S(w,p,e(w,p,u))$ which differentiated with respect to w gives $\frac{\partial L^{CS}}{\approx w} = \frac{\partial L^S}{\partial w} + \frac{\partial L^S}{\partial M}\frac{\partial e}{\partial w} = \frac{\partial L^S}{\partial w} + \frac{\partial L^S}{\partial M}\left(-L^{CS}\right)$. Rearranging and using $L^{CS} = L^S$ gives the Slutsky equation $\frac{\partial L^S}{\partial w} = \frac{\partial L^{CS}}{\partial w} + \frac{\partial L^S}{\partial M}L^S$ where $\frac{\partial L^{CS}}{\partial w} > 0$ but $\frac{\partial L^S}{\partial M}L^S < 0$ when leisure is a normal good. Thus, a priori it is hard to tell whether or not $\frac{\partial L^S}{\partial w} > 0$, i.e. whether labor supply is upward sloping.

[6] The proof here is quite similar to before using the FOC to find

$$\frac{\partial \Pi}{\partial p} = x^S + p\frac{\partial x^S}{\partial p} - w\frac{\partial L^D}{\partial p} = x^S + \hat{\mu}\frac{\partial x^S}{\partial p} - \hat{\mu}f'\left(L^D\right)\frac{\partial L^D}{\partial p} = x^S + \hat{\mu}\left(\frac{\partial x^S}{\partial p} - f'\left(L^D\right)\frac{\partial L^D}{\partial p}\right)$$

and then differentiating the production constraint with respect to p to get $\frac{\partial x^S}{\partial p} = f'\left(L^D\right)\frac{\partial L^D}{\partial p}$ so that the second term equals zero.

Because there are no income effects with firms, there is no interesting analogue problem to the EMP problem[7] and therefore no distinction between compensated and uncompensated supplies.

1.5 Elasticities of Supply and Demand

An **elasticity** is the ratio between proportional (i.e. percentage) change in one variable to proportional change in another variable. Since each proportion is unit-free so is the elasticity. The **elasticity of supply** of x is the proportional change in x^S to proportional change in price, usually its own price p

$$\eta_x^S \equiv \frac{\partial x^S}{\partial p}\frac{p}{x^S}$$

Heuristically this can be rewritten as $\eta^S \cong \frac{\Delta x^S}{x^S}/\frac{\Delta p}{p}$, the percent change in x^S, to the percent change in p. [8] The elasticity of labor supply η_L^S is similarly defined. The **elasticity of demand** of x is the proportional change in x^D to proportional change in price p, typically (but not always) with a negative sign in front so that the quantity is still positive

$$\eta_x^D = -\frac{\partial x^D}{\partial p}\frac{p}{x^D}$$

The compensated elasticity of demand is the same except with "CD" substituted in for "D". In some cases elasticities with respect to prices for other goods, aka **cross price elasticities**,(see below) are of interest. The **elasticity of demand of x with respect to income** is the proportional change in x^D to proportional change in income M

$$\eta_x^M = \frac{\partial x^D}{\partial M}\frac{M}{x^D}$$

Goods for which $\eta^M > 0$ are known as **normal goods**, with goods $\eta^M > 1$ being singled out as **luxury goods**, while goods for which $\eta^M < 0$ are known as **inferior goods**.

With elasticities the Slutsky equation can be rewritten in terms of elasticities by multiplying both sides by $p/x^D = p/x^{CD}$ and additionally

$$\frac{\partial x^D}{\partial p}\frac{p}{x^D} = \frac{\partial x^{CD}}{\partial p}\frac{p}{x^{CD}} - \frac{\partial x^D}{\partial M}p \Rightarrow \eta_x^D = \eta_x^{CD} + s_x\eta_x^M$$

where $s_x = px/M$ is the share of income spent on x. So we can see that for normal goods $\eta_x^D > \eta_x^{CD}$, i.e. uncompensated demand is *more* elastic than compensated demand. Notice that we can write $\eta_x^{CD} = \eta_x^D - s_x\eta_x^M$ so that the typically unobservable compensated elasticity may be calculated using potentially observable quantities η_x^D, η_x^M and s_x.

Example 4 ***Constant elasticity of demand:*** *ignoring non-labor income and wages, suppose $x^D(p) = p^{-a}$ where $a > 0$ is a constant. Calculating the elasticity $\partial x^D/\partial p = -ap^{-a-1}$ and so $\eta_x^D = -\left(-ap^{-a-1}\right) \cdot p/p^{-a} = a$ which is constant, giving the name to this demand.*

2 Taxes in Equilibrium

2.1 Types of Taxes

There are three types of taxes we consider here

Lump sum tax is a tax T which is simply taken out of income M when levied on persons or taken out of profits Π when levied on firms.

[7] The firm's **cost minimization problem (CMP)** $\min_L wL$ s.t. $x = f(L)$ is similar, albeit not analogous, to the EMP and yields the cost function $C(w,x) = wf^{-1}(x)$ which is similar to the expenditure function. Note that $\frac{\partial C}{\partial w} = f^{-1}(x) = L(x)$ the labor needed to produce x (also known as "conditional labor demand"), a very simple result of the same kind as Shepard's Lemma. However, since output x is taken parametrically and not profits, this problem is not analogous to the EMP.

[8] Note that $\eta_x^S = \partial \log x^S/\partial \log p$

Quantity or "Specific" tax is a tax t which is paid per unit of a good so that consumer pay price $q = p+t$. Who nominally bears this tax is not important so long as equilibrium is determined, i.e. decisions are made, and prices are set after the tax is put in place. The tax can be rewritten as $p = q - t$, so that the "wedge" is the same regardless of whether the tax is paid by producers or consumers.

Ad valorem tax encompasses two main types of tax, a **sales tax**, which is a tax on the sales of a business, and a **value added tax (VAT)** which is a tax which is paid only on the "value added" in a transaction, i.e. a business can deduct from their taxes the VAT taxes that they have already paid for purchasing their inputs. In the cases here there is no distinction. An ad valorem tax τ on buyers of x makes consumers pay a final price of $q = (1+\tau)\,p$, while firms only receive p so that a total of $t = \tau p$ per unit goes to the government. Be aware that an ad valorem tax paid by firms is typically written as $p = \left(1 - \tau^F\right) q$, so that the government claims τ^F percent of the revenue from x. An economically equivalent tax on consumers - i.e. one that introduces the same difference or "wedge" between p and q - can be found by dividing both sides by $\left(1-\tau^F\right)$ to get the final price $q = p/\left(1-\tau^F\right) = (1+\tau)\,p$ implying $(1+\tau) = 1/\left(1-\tau^F\right)$ and therefore $\tau = \tau^F/\left(1-\tau^F\right)$. For example a tax on firms of $\tau^F = 1/3$ implies $\tau = (1/3)\,/\,(2/3) = 1/2$.

2.2 Partial Equilibrium

The first step to analyzing the effect of taxes is to see how it affects supply and demand. We assume that the tax is known *before* purchasing decisions are made, so that equilibrium takes these taxes into account. The simplest analysis, known as **partial equilibrium** analysis, involves looking only at the market for the good in question, x, and treating only the price p as endogenous while leaving wages w and income M fixed, say at values $\bar{w}$ and $\bar{M}$.

2.2.1 Partial Equilibrium Condition

Equilibrium will be achieved at the price where supply equals demand, i.e.

$$x^S(\bar{w}, p) = x^D\left(\bar{w}, q, \bar{M}\right)$$

where q is related to p in one of the ways mentioned above so that there are two equations in two unknowns, q and p. For example with the ad valorem tax we can substitute in q so that $x^S(\bar{w}, p) = x^D\left(\bar{w}, (1+\tau)\,p, \bar{M}\right)$ and we now have a single equation in a single unknown p, with solution p_P where the subscript P stands for "partial equilibrium" (I apologize for the use of two kinds of P's). Tax revenue collected by the government equals $R = (q_P - p_P)\,x_P$, where $x_P = x^S(\bar{w}, p_P)$ is the partial equilibrium quantity with the tax in place.

2.2.2 Effect of a Small Tax in Partial Equilibrium

Assume we start with a situation with no taxes, with an initial equilibrium price p_0 and quantity traded x_0. It is useful to analyze the effect of a small quantity tax (t "close to" 0) on the quantities p, x, and the final consumer price q. Using a small tax means we can treat derivatives and elasticities as constants, making the formulas more easily tractable. Treat p as a function of t, $p(t)$ and take the equation $x^S(\bar{w}, p(t)) = x^D\left(\bar{w}, p(t)+t, \bar{M}\right)$ and differentiate this implicitly with respect to t applying the chain rule:

$$\frac{\partial x^S}{\partial p}\frac{dp}{dt} = \frac{\partial x^D}{\partial p}\left[1 + \frac{dp}{dt}\right] \Rightarrow \frac{dp}{dt}\left[\frac{\partial x^S}{\partial p} - \frac{\partial x^D}{\partial p}\right] = -\frac{\partial x^D}{\partial p} \Rightarrow \frac{dp}{dt} = \frac{-\frac{\partial x^D}{\partial p}}{\frac{\partial x^S}{\partial p} - \frac{\partial x^D}{\partial p}} \tag{1}$$

multiplying the denominator and numerator by p_P/x_P and substituting in the formulas for elasticities we get

$$\frac{dp_P}{dt} = -\frac{\eta_x^D}{\eta_x^S + \eta_x^D}$$

and so the change in the producer price is the negative of the elasticity of demand relative to the ratio of the sums of the elasticites. The producer pays more, i.e. bears more of the burden of the tax, the more elastic demand is and the less elastic supply is. The change in the consumer price is easy to find since $dq_P/dt = 1 + dp_P/dt$ and so[9]

$$\frac{dq_P}{dt} = \frac{\eta_x^S}{\eta_x^S + \eta_x^D}$$

The change in quantity sold in the market is given by[10]

$$\frac{dx_P}{dt} = \frac{\partial x^S}{\partial p}\frac{dp_P}{dt} = -\frac{\partial x^S}{\partial p}\frac{\eta_x^D}{\eta_x^S + \eta_x^D} = -\frac{\eta_x^S\eta_x^D}{\eta_x^S + \eta_x^D}\frac{x_P}{p_P}$$

which is the product of the two elasticities divided by their sum times the ratio of x_P^P to p_P^P.

The point of going through all these derivatives is that they give us formulas for changes in prices and quantities. If the terms in the derivatives are "fairly constant," which is typically true for a small tax, then it is okay to approximate changes in prices and quantities from a no tax scenario to a small tax scenario with the formulas

$$\begin{aligned}
\Delta p &= p_P - p_0 \approx \frac{dp_P}{dt}t = -\frac{\eta_x^D}{\eta_x^S + \eta_x^D}t \\
\Delta q &= q_P - p_0 \approx \frac{dq_P}{dt}t = \frac{\eta_x^S}{\eta_x^S + \eta_x^D}t \\
\Delta x &= x_P - x_0 \approx \frac{dx_P}{dt}t = -\frac{\eta_x^S\eta_x^D}{\eta_x^S + \eta_x^D}\frac{x_P}{p_P}t
\end{aligned}$$

Looking at these formulas we can see immediately that generally the producer price p falls, the consumer price q rises, and the quantity x traded falls. Notice that the tax burden or **tax incidence** can be divided into that paid by consumers and that paid by firms as $tx_P = \Delta q \cdot x_P - \Delta p \cdot x_P$. The amount of tax paid by the consumer $\Delta q \cdot x_P$ depends on the size of the elasticity of supply relative to the sum of the two elasticities; the more elastic supply is the more the consumer pays. If for instance supply is perfectly elastic $\eta_x^S = +\infty$ then, the consumer bears the entire tax. Similarly, the amount of tax paid of suppliers $-\Delta p \cdot x_P$ depends on the size of the elasticity of demand relative to the sum of the two.

To see the effect on the total quantity Δx traded, divide both the numerator and denominator by $\eta_x^S\eta_x^D$ and rewrite this as

$$\Delta x = -\frac{1}{\frac{1}{\eta_x^S} + \frac{1}{\eta_x^D}}\frac{x_P}{p_P}t = -\bar{\eta}_x\frac{x_P}{p_P}t \quad \text{where } \bar{\eta}_x = \left(\frac{1}{\eta_x^S} + \frac{1}{\eta_x^D}\right)^{-1}$$

where the $\bar{\eta}_x$ term, two times the harmonic mean of the two elasticities[11], is an overall measure of the responsiveness of the quantity sold x_P to the tax. If either elasticity is zero then the quantity does not change at all. If one the elasticities of supply is infinite, i.e. it is perfectly elastic, then $\bar{\eta}_x = \eta_x^D$ and the change will be determined by the demand schedule alone. In general Δx will be larger the greater are the elasticities η_x^S and η_x^D.

2.3 General Equilibrium

The problem with partial equilibrium analysis is that it ignores the repercussion of a tax on other parts of the economy. Effects we need to consider are (i) how a change in price in one market affects supply and

[9] Note that with an equivalent ad valorem tax so that $t = \tau x$ this implies that $dt = \tau dx + xd\tau \cong xd\tau$ if τ is small. In this case then

$$\frac{dp}{d\tau} = \frac{dp}{dt}\frac{dt}{d\tau} = -\frac{\eta_x^D}{\eta_x^S + \eta_x^D}x$$

which just involves another x term.

[10] For the ad valorem tax we can use the chain rule again to get

$$\frac{dx^P}{d\tau} = \frac{\partial x^S}{\partial p}x^P\frac{dp}{d\tau} = -\frac{\partial x^S}{\partial p}\frac{\eta_x^D}{\eta_x^S + \eta_x^D}\left(x^P\right)^2 = -\frac{\eta_x^S\eta_x^D}{\eta_x^S + \eta_x^D}p^P x^P$$

[11] The harmonic mean $\bar{a}_H$ of two numbers a_1, a_2 is defined by the formula $\frac{1}{\bar{a}_H} = \frac{1}{2}\left(\frac{1}{a_1} + \frac{1}{a_2}\right)$.

demand for other goods in other markets, (ii) changes in non-labor incomes due to taxes through its impact on firm's profits, and also (iii) how taxes can also change demand due to how government chooses to spend its tax revenues. **General equilibrium** analysis requires trying to model all of this.

2.3.1 General Equilibrium Condition

For the simple two good world where each individual owns a firm, there is only one relative price that matters, w/p, and so (i) is not much of a complication as we may just set $w = 1$. (ii) is certainly a complication as now income $M = \Pi(1,p)$ and p will be affected by the tax. We also assume for now that the government does not purchase any of the taxed good with its revenue, but rather purchases the untaxed good so that (iii) remains in the background and does not appear in the supply and demand equation. In this special case we have that supply equals demand

$$x^S(1,p) = x^D(1,q,\Pi(1,p))$$

where again q and p are interrelated by the tax imposed. Thus with the ad valorem tax $x^S(1,p) = x^D(1,(1+\tau)p,\Pi(1,p))$, which has solution p_G where G stands for "general equilibrium". The tax revenue $R = (q_G - p_G)x_G$ is then used to buy labor so in the other market we have $L^D(1,p_G) + R/w = L^S(1,p_G)$ - note that there is no need to solve for this - it is automatic through Walras' Law.

2.3.2 Effect of a Small Tax in General Equilibrium

To find out how a small tax ($t = dt \approx 0$) will affect prices and the amount sold depends fundamentally on how the government decides to spend its tax dollars. A common assumption is to assume that the government will spend the dollar in a similar manner as would the consumer. This different treatment of assumption (iii) can be modeled by just giving back the consumer the tax revenue collected so that non labor income $M = \Pi(1,p) + tx^D$. Again we must

$$x^S(1,p(t)) = x^D\left[1,p(t)+t,\Pi(1,p(t))+tx^D\right]$$

Differentiating with respect to t and using the fact that $dq/dt = dp/dt + 1$

$$\frac{\partial x^S}{\partial p}\frac{dp}{dt} = \frac{\partial x^D}{\partial p}\left(1+\frac{dp}{dt}\right) + \frac{\partial x^D}{\partial M}\left(\frac{\partial \Pi}{\partial p}\frac{dp}{dt} + x^D + t\frac{\partial x^D}{\partial p}\left(1+\frac{dp}{dt}\right)\right)$$

since the tax is small we can set set $t = 0$ to give us[12]

$$\begin{aligned}
\frac{\partial x^S}{\partial p}\frac{dp}{dt} &= \frac{\partial x^D}{\partial p}\left(1+\frac{dp}{dt}\right) + \frac{\partial x^D}{\partial M}\left(\frac{\partial \Pi}{\partial p}\frac{dp}{dt} + x^D\right) \\
&= \frac{\partial x^D}{\partial p}\left(1+\frac{dp}{dt}\right) + \frac{\partial x^D}{\partial M}\left(x^S\frac{dp}{dt} + x^D\right) \\
&= \left(\frac{\partial x^D}{\partial p} + \frac{\partial x^D}{\partial M}x^S\right)\left(1+\frac{dp}{dt}\right) \\
&= \frac{\partial x^{CD}}{\partial p}\left(1+\frac{dp}{dt}\right)
\end{aligned}$$

where Hotelling's Lemma is used for the second line and then $x^D = x^S$ and some rearranging produces the third line, and Slutsky's equation is used for the last line. This now looks similar to the partial equilibrium equation 1 with compensated demands in place of uncompensated demands. Carrying through the same analysis we can see that

$$\Delta p_G \approx -\frac{\eta_x^{CD}}{\eta_x^S + \eta_x^{CD}}t,\ \Delta q_G \approx \frac{\eta_x^S}{\eta_x^S + \eta_x^{CD}}t,\ \text{and } \Delta x_G \approx \frac{\eta_x^S\eta_x^{CD}}{\eta_x^S + \eta_x^{CD}}\frac{p_G}{x_G}t \tag{2}$$

[12] Doing so will eliminate any income effect due to deadweight loss, simplifying the analysis a bit, although it will not eliminate the deadweight loss itself.

The one thing that changes fundamentally in this analysis is the use of compensated elasticites for uncompensated ones. This is because income effects wash away as tax dollars are ultimately spent on goods and services, lessening the impact of the tax when goods are normal. The same points should be made as were discussed in the partial equilibrium case, although the general equilibrium is typically more accurate.

3 Welfare and Taxation

3.1 Welfare Changes with the Indirect Utility Function

As we saw above tax policies to affect prices and incomes of individuals an thus the welfare consequences of a tax should in theory be able to be judged through the indirect utility function. For example say prices and incomes begin at (w_0, q_0, M_0) but because of taxes are changed to (w_1, q_1, M_1), (now dropping the subscript G and replacing it with 1 which means "after tax"). The change in utility can be found using the indirect utility function

$$\Delta V = V_1 - V_0 = V(w_1, p_1, M_1) - V(w_0, p_0, M_0)$$

Since M depends on profits we could substitute in $M_0 = \Pi(w_0, p_0)$ and $M_1 = \Pi(w_1, p_1)$ to finish the calculation. However in order to proceed further it is useful to add back in tax revenues to income, as if they were redistributed lump sum. We know already that changes in welfare due to small lump-sum tax R should be just $-\frac{\partial V}{\partial M}R = -\alpha R$ (assuming α is constant, justified by a smal R), but we want to know if there are *any other* changes in welfare due to how taxes change relative prices. Adding back the taxes also simplifies the analysis since we assume that the government will spend its money just as the consumer would had it gotten it back in lump sum form. Therefore we will set $M_1 = \Pi(1, p_1) + tx(1, p_1)$

Therefore in a general equilibrium setting where $w = 1$ we would calculate the loss from imposing a quantity tax on x as

$$\Delta V = V(1, q_1, \Pi(1, p_1) + tx(1, p_1)) - V(1, p_0, \Pi(1, p_0))$$

Differentiating this quantity with respect to t we get

$$\begin{aligned}
\frac{dV}{dt} &= \frac{\partial V}{\partial p}\left(1 + \frac{dp}{dt}\right) + \frac{\partial V}{\partial M}\left[\frac{\partial \Pi}{\partial p}\frac{dp}{dt} + x_1 + t\frac{\partial x^S}{\partial p}\frac{dp}{dt}\right] \\
&= -\alpha x_1 \left(1 + \frac{dp}{dt}\right) + \alpha\left[x_1\frac{dp}{dt} + x_1 + \frac{\partial x^S}{\partial p}\frac{dp}{dt}t\right] \\
&= \alpha x_1 \left(1 + \frac{dp}{dt} - 1 - \frac{dp}{dt}\right) + \alpha\frac{\partial x^S}{\partial p}\frac{dp}{dt}t \\
&= \alpha\frac{\partial x^S}{\partial p}\frac{dp}{dt}t \\
&= -\alpha\frac{\partial x^S}{\partial p}\frac{\eta_x^{CD}}{\eta_x^S + \eta_x^{CD}}t \\
&= -\alpha\frac{\eta_x^S\eta_x^{CD}}{\eta_x^S + \eta_x^{CD}}\frac{x_1}{p_1}t
\end{aligned}$$

making good use of Roy's Identity and Hotelling's Lemma on the second line and substituting in for dp/dt on the fifth line. This expression gives the marginal effect on V of an increase in taxes. Assuming that the terms in front of t are "fairly constant," then we should be able to approximate the full change V due to taxes by integrating both sides with respect to t.[13] Proceeding with the integration gives

$$\Delta V \approx -\alpha\frac{\eta_x^S\eta_x^{CD}}{\eta_x^S + \eta_x^{CD}}\frac{x_1}{p_1}\frac{t^2}{2} \tag{3}$$

so that even when the tax is redistributed through lump sum transfers there is still a welfare loss. This loss is known as the **deadweight loss (or burden) of taxation** as it is a loss in welfare which is not used to

[13] If c is a constant, the integral of $V_1 = \int dV = \int ct \cdot dt = c\int t \cdot dt = c\frac{t^2}{2} + C$ where C is a constant of integration which here we set to V_0 and so $\Delta V = V_1 - V_0 \approx c\frac{t^2}{2}$.

create any kind of transfer at all. Be aware that if taxes are not redistributed lump sum, then the full loss can be approximated by $\Delta V - \alpha \cdot tx_1$.

3.1.1 Deadweight Loss

Losses seen in equation (3) are often seen in the form of a **deadweight loss triangle (aka "Harberger" triangle)** which measures the combined loss in producer and consumer it surplus. This triangle has three sides (1) the supply curve (2) the compensated demand curve, and (3) the line defining the amount x_1 traded after the tax is imposed. The height of the deadweight loss triangle is simply t while the width is Δx_1 from (2) and therefore the area of the deadweight loss triangle

$$DWL = \frac{1}{2} \cdot t \cdot \Delta x_1 = \frac{\eta_x^S \eta_x^{CD}}{\eta_x^S + \eta_x^{CD}} \frac{x_1}{p_1} \frac{t^2}{2} \tag{4}$$

In order to translate the loss dollar loss measured by DWL into the utility loss ΔV , we just multiply by $-\alpha$ (which we hope is relatively constant).

Because deadweight loss is a triangle, it is an area which increases with the *square* of the tax. This means two things: (1) small taxes have small dead-weight losses relative to income and (2) increasing already large taxes typically incurs a larger loss than increasing smaller taxes. Taking the ratio of DWL to tax revenue $R = tx_1$ we get

$$\frac{DWL}{R} = \frac{1}{2} \frac{\eta_x^S \eta_x^{CD}}{\eta_x^S + \eta_x^{CD}} \frac{t}{p_1} = \frac{1}{2} \frac{\eta_x^S \eta_x^{CD}}{\eta_x^S + \eta_x^{CD}} \tau$$

The ratio $t/p_1 = \tau$ is simply the proportion of the tax relative to the price (much like with an ad valorem tax)[14]. So we can see that for a small tax, the proportion of deadweight loss to revenue will be typically small, while for large taxes the opposite is true.

3.2 Welfare Changes with the Expenditure Function

Unfortunately utility functions can never be observed and so it is hard to use these formulae in practice with this term α, the marginal utility of a dollar: it is not known typically not constant for large taxes. The other option is to find welfare losses in terms of monetary equivalents, which is what the expenditure function and the profit function are used for

3.2.1 Compensating and Equivalent Variation

Profit functions pose no problem however expenditure functions can be evaluated at two different utility levels, namely utility before the tax t is imposed $u_0 = V_0$ and utility after the tax is imposed $u_1 = V(w_1, q_1, \Pi(w_1, q_1))$ without adding back in tax revenues to income.

Compensating Variation is the amount of money that needs to be given (i.e. through a lump-sum transfer) to an individual in order to compensate her for the utility loss due to the tax, namely

$$CV = e(w_1, q_1, u_0) - M$$

$e(w_1, q_1, u_0)$ is the total amount of money needed to get u_0, while M is the money the individual already has and so CV makes up for the difference. The after tax prices w_1, q_1 are used since the tax has already been imposed. More money is needed because prices are higher.

Equivalent Variation is the amount of money that needs to be taken away (i.e. in a lump-sum form) from the individual in order to achieve that same utility as would occur with the tax, but *without actually imposing it.* This lump-sum tax is "equivalent" in producing the same utility as the actual tax

$$EV = M - e(w_0, q_0, u_1)$$

[14] For an advalorem tax

$$DWL = \frac{1}{2} \frac{\eta_x^S \eta_x^{CD}}{\eta_x^S + \eta_x^{CD}} p_1 x_1 \tau^2$$

M is original income and $e(w_0, q_0, u_1)$ is the total (lower) amount of money that would lead to utility u_1 and so EV is the difference. Here the pre-tax prices w_0, q_0 are used. Another take on the equivalent variation is that it describes the "willingness to pay" of an individual to avoid getting taxed.

Unfortunately there is no consensus on whether compensating variation or equivalent variation is a better measure of welfare. Notice that income is the actually amount of money spent in the actual scenarios and so we can write $M = e(w_0, q_0, u_0) = e(w_1, q_1, u_1)$ and therefore CV and EV can be rewritten as

$$\begin{aligned} CV &= e(w_1, q_1, u_0) - e(w_0, q_0, u_0) \\ EV &= e(w_1, q_1, u_1) - e(w_0, q_0, u_1) \end{aligned}$$

and so the difference between CV and EV depends on whether we use pre-tax utility or post-tax utility. What the "right" utility level to choose from is hard to say. However if the tax imposed is small then u_0 and u_1 will not be that far apart and CV and EV will be very similar.

Example 5 *In the case of quasilinear utility this problem goes away as utility drops out of the expression*

$$CV = \left\{p_1 \cdot (v')^{-1}(p_1) - v\left[(v')^{-1}(p_1)\right]\right\} - \left\{p_0 \cdot (v')^{-1}(p_0) - v\left[(v')^{-1}(p_0)\right]\right\} = EV$$

no matter how large the tax is there is no worry here. This quantity is also equal to the consumer"s surplus.

3.2.2 Deadweight Loss Revisited

Consider the sum of how much individuals and firms would pay to avoid a distortionary tax on x. This would be the equal to the equivalent variation of consumers EV plus the loss in profits of firms due to the tax $-\Delta\Pi = -[\Pi(w_1, p_1) - \Pi(w_0, p_0)]$. Now compare this total to the revenue of the tax $R = tx^S(w_1, p_1)$. The difference between these numbers is this deadweight loss of taxation:

$$DWL = EV - \Delta\Pi - R = .\ e(w_1, q_1, u_1) - e(w_0, q_0, u_1) + \Pi(w_0, p_0) - \Pi(w_1, p_1) - tx(w_1, p_1)$$

The difference between these two numbers, which is positive, is in theory an "avoidable cost": if the government could get consumers and firms to hand over the taxes in lump sum form, then it could raise more revenue at the same utility cost. This extra amount could be spent on public goods or be given back to individuals. As discussed below lump sum taxes are not practical for a number of reasons.

Consider the effect of a small tax t, so that we make $p(t)$ and set $w_1 = 1$ and write DWL as following

$$DWL(t) = \ e(1, p(t) + t, u_1) - e(1, p_0, u_1) + \Pi(1, p(t)) - \Pi(w, p_0) - tx(1, p_1)$$

Taking the derivative with respect to t and making use of Shepard's Lemma and Hotelling's Lemma, and substituting in dp/dt gives

$$\begin{aligned} \frac{dDWL}{dt} &= \frac{\partial e}{\partial p}\left(1 + \frac{dp}{dt}\right) - \frac{\partial \Pi}{\partial p}\frac{dp}{dt} - x_1 - t\frac{\partial x^S}{\partial p}\frac{dp}{dt} \\ &= x_1\left(1 + \frac{dp}{dt}\right) - x_1\left(1 + \frac{dp}{dt}\right) - \frac{\partial x^S}{\partial p}\frac{dp}{dt}t \\ &= -\frac{\partial x^S}{\partial p}\frac{dp}{dt}t \\ &= \frac{\eta_x^S \eta_x^{CD}}{\eta_x^S + \eta_x^{CD}}\frac{x_1}{p_1}t \end{aligned}$$

Of course this is just the marginal effect of t on DWL. To find the total DWL we integrate with respect to t, hoping that the other terms are constant and get

$$DWL = \frac{\eta_x^S \eta_x^{CD}}{\eta_x^S + \eta_x^{CD}}\frac{x_1}{p_1}\frac{t^2}{2}$$

which is exactly the same expression that we had before in equation (4).

One advantage of working within the framework here is that it avoids use of the utility function and the tricky shadow price of a dollar in utility terms α. One disadvantage is that in more general cases where different people have different α's, like when they have different incomes, we neglect whether the money is being lost by people who have higher α's (i.e. poorer people) or with lower α's (i.e. richer people). In other words this framework potentially neglects distributional concerns the government (or voters) might have.

4 Optimal Taxation

Having discussed the various impacts of taxes on prices, quantities, and welfare, the next step is to consider what the optimal way to tax is. We assume that the government must reach some target revenue R.

4.1 Lump-sum Taxes

As we know already lump-sum taxes, which mean just collecting R from each individual, are the best taxes because they cause no distortion in prices: consumers and producers face the same relative prices and so $MRS_{lx} = MRT_{lx}$, Pareto optimality is achieved and there is zero dead-weight loss of taxation. Lump-sum taxes in the real world are quite problematic however because we do not want to tax everyone identically, unlike in the case here where we might as well since everyone is identical. People vary in their "ability to pay," say through their non-labor income M, their time endowments T, and their market wages w. Therefore, we are forced to tax economic transactions such as through the taxes considered here.

4.2 Taxing All Goods

A typical assumption in this type of analysis is that there is one good which *cannot* be taxed. If every good could be taxed then a sort of lump-sum tax (proportional to income) would be available to us. Say we could tax both consumption x and leisure l - a tax on labor is different from a tax on leisure) on individuals with ad valorem taxes τ_x and τ_l. Then individuals face the budget constraint budget constraint we get $p(1+\tau_x)x + w(1+\tau_l)l = wT + M$ and will set $MRS_{lx} = \frac{w(1+\tau_l)}{p(1+\tau_x)}$ and firms which pay no taxes will set $MRT_{lx} = \frac{w}{p}$. In order to achieve Pareto optimality, set $\tau_x = \tau_l = \tau$, in which case $MRS_{lx} = \frac{w}{p} = MRT_{lx}$ and so no distortion is caused. The budget constraint can be rearranged to produce $px + wl = (wT + M)/(1+\tau)$, which simply reduces total income (which includes the entire value of the time endowment) proportionally, similar to a lump sum tax.[15]. However, leisure taxes are hard to collect since there is no obvious market transaction involved in consuming leisure. A way around this would be to impose a tax on the market value of people's time endowment equal to τwT and then *subsidize* labor at rate τ, but this is likely to be as problematic as imposing a lump sum taxes to begin with. Given these difficulties, we assume that we cannot tax leisure (or labor), albeit mainly for simplicity.

4.3 The Laffer Rate

With a two good world where only one good cannot be taxed, then there is not much choice over taxes. t must be set so that $tx_1^S = R$, which will impose a burden on society similar to the one discussed above. However we have to be sure that we can even raise R with this single tax. In fact the maximum amount that can be raised on a tax can be found by taking the first order condition of R with respect to t.

$$\frac{dR}{dt} = x_1 + t\frac{\partial x^S}{\partial p}\left(\frac{dp}{dt}\right) = 0$$

[15] The equivalent lump sum tax R is found as

$$wT + M - R = \frac{wT+M}{1+\tau} \Rightarrow R = \frac{\tau}{1+\tau}(wT+M)$$

Substituting in for dp/dt and rearranging we get the solution for the **Laffer Rate**

$$x_1 = t\frac{\partial x^S}{\partial p}\frac{\eta_x^{CD}}{\eta_x^S + \eta_x^{CD}} \Rightarrow x_1 = t\frac{\eta_x^S \eta_x^{CD}}{\eta_x^S + \eta_x^{CD}}\frac{x_1}{p_1} \Rightarrow \tau^* = \frac{t^*}{p_1} = \frac{\eta_x^S + \eta_x^{CD}}{\eta_x^S \eta_x^{CD}} = \frac{1}{\eta_x^S} + \frac{1}{\eta_x^{CD}} = \frac{1}{\bar{\eta}_x}$$

This implies that the lower the elasticities of supply and demand the larger the maximum revenue that can be raised.

4.4 The Simplified Two Consumption Goods Problem

To make the problem more interesting let us introduce a second consumption good y with price p_y and quantity tax t_y and calling now the price of x, p_x giving it a quantity tax t_x. To make things easier we can normalize the producer prices of x and y to one, $p_x = p_y = 1$, and by assuming that taxes all terms other than t_x and t_y are constant (a big assumption). Consider the problem of minimizing the deadweight loss DWL_x and DWL_y of x and y which we assume to be independent (another big assumption), subject to the constraint of raising revenue $t_x x + t_y y \geq R$. In other words

$$\min_{t_x t_y} \bar{\eta}_x x \frac{t_x^2}{2} + \bar{\eta}_y y \frac{t_y^2}{2} \qquad \text{s.t.} \quad t_x x + t_y y \geq R$$

Now form the Lagrangian with multiplier λ

$$\mathfrak{L}(t_x, t_y, \lambda) = \bar{\eta}_x x \frac{t_x^2}{2} + \bar{\eta}_y y \frac{t_y^2}{2} + \lambda \left[R - t_x x + t_y y\right]$$

Then the FOC are given by

$$\bar{\eta}_x t_x^* x - \lambda x = 0$$
$$\bar{\eta}_y t_y^* y - \lambda y = 0$$

Divide by x and y in each equation gets $\bar{\eta}_x t_x^* = \bar{\eta}_y t_y^* = \lambda$ which then means

$$\frac{t_x^*}{t_y^*} = \frac{\bar{\eta}_y}{\bar{\eta}_x}$$

i.e. the ratio of the tax on x to the tax on y should be equal the ratio of the inverse elasticities. Only if the two elasticities are the same $\bar{\eta}_x = \bar{\eta}_y$ should the taxes on these different goods be equal. So for instance a uniform sales tax may not be optimal for a large variety of goods. This result also tells us that it rarely makes sense to impose a tax on one good, but not the other other: $t_x^* = 0$ implies $\bar{\eta}_y = 0$, or $\bar{\eta}_x = +\infty$. If $\bar{\eta}_y = 0$ we should tax only the perfectly inelastically supplied good y (with no deadweight loss) or if $\bar{\eta}_x = +\infty$ we cannot raise any revenue from a tax on x, only dead weight loss, and hence it should not be taxed. In general this rule tells us to spread out taxes but only in proportion to the inverse elasticities of supply and demand.

4.5 The Two-Good Ramsey Problem

The derivation of the optimal tax rule above, was a bit too fast and loose to really trust. A more careful treatment is warranted, and therefore we turn to Ramsey's (1927) formulation of the optimal tax problem in a two taxable good setting (although there are three goods, total as leisure is not taxable by assumption).

4.5.1 The Set-Up and Assumptions

For the simple version of the Ramsey problem assume that producer prices are fixed at $p_x = p_y = 1$. Second we assume that there is no non-labor income $M = 0$. This assumption can be justified if firms have constant returns to scale and perfect competition holds, in which case they make zero profits. The indirect utility function will then take the form $V(1, 1 + t_x, 1 + t_y, 0)$. Recall that there is no tax on labor. The government seeks to maximize utility through the indirect utility function subject to its revenue constraint.

$$\max_{t_x, t_x} V(1, 1+t_x, 1+t_y, 0) \qquad \text{s.t. } t_x x + t_y y \geq R$$

The Lagrangean is then given by

$$\mathfrak{L}(t_x, t_y, \lambda) = V(1, 1+t_x, 1+t_y, 0) + \lambda \left[t_x x^D + t_y y^D - R \right]$$

The first order conditions can be given as, remembering that x^D and y^D depend on the tax through prices

$$\begin{aligned}
\frac{\partial \mathfrak{L}}{\partial t_x} &= \frac{\partial V}{\partial p_x} + \lambda \left[x^D + t_x^* \frac{\partial x^D}{\partial p_x} + t_y^* \frac{\partial y^D}{\partial p_x} \right] = 0 \\
\frac{\partial \mathfrak{L}}{\partial t_y} &= \frac{\partial V}{\partial p_y} + \lambda \left[y^D + t_x^* \frac{\partial x^D}{\partial p_y} + t_y^* \frac{\partial y^D}{\partial p_y} \right] = 0
\end{aligned}$$

Using Roy's identity and rearranging

$$\begin{aligned}
\lambda \left[x^D + t_x^* \frac{\partial x^D}{\partial p_x} + t_y^* \frac{\partial y^D}{\partial p_x} \right] &= \alpha x^D \\
\lambda \left[y^D + t_x^* \frac{\partial x^D}{\partial p_y} + t_y^* \frac{\partial y^D}{\partial p_y} \right] &= \alpha y^D
\end{aligned}$$

Now solving for $1/\lambda$ in each equation and setting the two equal

$$\frac{1}{\lambda} = \left[x^D + t_x^* \frac{\partial x^D}{\partial p_x} + t_y^* \frac{\partial y^D}{\partial p_x} \right] / \alpha x^D = \left[y + t_x^* \frac{\partial x^D}{\partial p_y} + t_y^* \frac{\partial y^D}{\partial p_y} \right] / \alpha y^D$$

Cancelling out the α's and simplifying we get

$$t_x^* \frac{\partial x^D}{\partial p_x} \frac{1}{x^D} + t_y^* \frac{\partial y^D}{\partial p_x} \frac{1}{x^D} = t_x^* \frac{\partial x^D}{\partial p_y} \frac{1}{y^D} + t_y^* \frac{\partial y^D}{\partial p_y} \frac{1}{y^D} \tag{5}$$

We pause here to consider how to deal with the cross derivative terms $\frac{\partial y^D}{\partial p_x}$ and $\frac{\partial x^D}{\partial p_y}$.

4.5.2 Cross-terms are zero

First, consider the case where both these cross-terms are zero, $\frac{\partial y^D}{\partial p_x} = \frac{\partial x^D}{\partial p_y} = 0$. This which implies that equation (5) becomes

$$t_x^* \frac{\partial x^D}{\partial p_x} \frac{1}{x^D} = t_y^* \frac{\partial y^D}{\partial p_y} \frac{1}{y^D}$$

Recalling that $p_x = p_y = 1$ we have two uncompensated elasticities and so

$$\frac{t_x^*}{1+t_x^*} \eta_x^D = \frac{t_y^*}{1+t_y^*} \eta_y^D \Rightarrow \frac{\frac{t_x^*}{1+t_x^*}}{\frac{t_y^*}{1+t_y^*}} = \frac{\eta_y^D}{\eta_x^D}$$

This condition along with the revenue constraint (which holds with equality) $t_x^* x^D + t_y^* y^D = R$ are the two equations which define the two optimal tax rates $\left(t_x^*, t_y^*\right)$ although do not ignore that x^D and y^D themselves depend on $\left(t_x^*, t_y^*\right)$.

Notice now that if we solve for $1/\lambda = \left(1 - t_x^* \eta_x^D\right) / \alpha \Rightarrow$

$$\lambda = \frac{\alpha}{1 - t_x^* \eta_x^D} > \alpha$$

which means that the shadow price of a dollar of revenue λ is greater than the shadow price of a dollar of income α, which is the also the shadow price of a dollar of lump-sum tax. This means that even optimal commodity taxation does worse than lump-sum taxation in efficiency terms.

4.5.3 Cross-terms are not zero (Optional)

In the more complicated case we need to introduce the concept of a **cross-price elasticity**, i.e. the percent change in demand of x for a percent change in the price of y.

$$\eta_{xy}^{D} = -\frac{\partial x^{D}}{\partial p_y}\frac{p_y}{x^{D}}$$

The definition for η_{yx}^{D} is symmetric, and we now take care to write **own-price elasticities** with a double xx, i.e. $\eta_{xx}^{D} = \eta_{x}^{D}$. Compensated demand elasticities are defined in the obvious way. The next result we need is to state the Slutsky equation for cross terms: taking $y^{CD}(p_x, p_y, u) = y^{D}(p_x, p_y, e(p_x, p_y, u))$ and differentiating with respect to p_x gives

$$\frac{\partial y^{CD}}{\partial p_x} = \frac{\partial y^{D}}{\partial p_x} + \frac{\partial y^{D}}{\partial M}\frac{\partial e}{\partial p_x} = \frac{\partial y^{D}}{\partial p_x} + \frac{\partial y^{D}}{\partial M}x^{D}$$

which rearranging gives

$$\frac{\partial y^{D}}{\partial p_x} = \frac{\partial y^{CD}}{\partial p_x} - \frac{\partial y^{D}}{\partial M}x^{D}$$

The last result needed is **Slutsky symmetry** which is that the cross-derivative of compensated demand of y with respect to price p_x is equal to the cross-derivative of the compensated demand of x, with respect to y,

$$\frac{\partial y^{CD}}{\partial p_x} = \frac{\partial x^{CD}}{\partial p_y}$$

This surprising result follows from a fact from calculus (known as "Young's Theorem") that cross-partial derivatives of a function are equal. Therefore, with the expenditure function $\frac{\partial^2 e}{\partial p_x p_y} = \frac{\partial^2 e}{\partial p_y p_x}$ or $\frac{\partial}{\partial p_x}\left(\frac{\partial e}{\partial p_y}\right) = \frac{\partial}{\partial p_y}\left(\frac{\partial e}{\partial p_x}\right)$. Inserting Shepard's Lemma gives the result $\frac{\partial}{\partial p_x}y^{CD} = \frac{\partial}{\partial p_y}x^{CD}$.

Now we can simplify the expressions in equation we need into an intelligible form. First

$$\frac{\partial x^{D}}{\partial p_x}\frac{1}{x} = \frac{\partial x^{CD}}{\partial p_x}\frac{1}{x} - \frac{\partial x^{D}}{\partial M} = -\eta_{xx}^{CD} - \frac{\partial x^{D}}{\partial M}$$

For the second term we have to use the Slutsky equation for a cross elasticity and use the fact that which implies ,

$$\frac{\partial y^{D}}{\partial p_x}\frac{1}{x} = \frac{\partial y^{CD}}{\partial p_x}\frac{1}{x} - \frac{\partial y^{D}}{\partial M} = \frac{\partial x^{CD}}{\partial p_y}\frac{1}{x} - \frac{\partial y^{D}}{\partial M} = -\eta_{xy}^{CD} - \frac{\partial y^{D}}{\partial M}$$

Plugging this into equation (5) gives and cancellling out all of the income terms gives

$$\frac{t_x^*}{1+t_x^*}\eta_{xx}^{CD} + \frac{t_y^*}{1+t_y^*}\eta_{xy}^{CD} = \frac{t_x^*}{1+t_x^*}\eta_{yx}^{CD} + \frac{t_y^*}{1+t_y^*}\eta_{yy}^{CD}$$

Simplifying further

$$\frac{\frac{t_x^*}{1+t_x^*}}{\frac{t_y^*}{1+t_y^*}} = \frac{\eta_{yy}^{CD} - \eta_{xy}^{CD}}{\eta_{xx}^{CD} - \eta_{yx}^{CD}}$$

Using the fact that $\eta_{xl} + \eta_{xx} + \eta_{xy} = 0$ and $\eta_{yl} + \eta_{yx} + \eta_{yy} = 0$ we can substitute in and rearrange to get[16]

[16] This fact comes from the fact that if all prices rise by the same amount, then compensated demand does not change (it is homogenous of degree zero), i.e. for $\lambda > 0$ $x^{CD}(\lambda w, \lambda p_x, \lambda p_y, u) = x^{CD}(w, p_x, p_y, u)$. Differentiating this equation with respect to λ and setting $\lambda = 1$ we get

$$\frac{\partial x^{CD}}{\partial w}w + \frac{\partial x^{CD}}{\partial p_x}p_x + \frac{\partial x^{CD}}{\partial p_y}p_y = 0$$

Dividing by x, these turn into the elasticites.

$$\frac{\frac{t_x^*}{1+t_x^*}}{\frac{t_y^*}{1+t_y^*}} = \frac{\eta_{xl}^{CD} + \left(\eta_{xx}^{CD} + \eta_{yy}^{CD}\right)}{\eta_{yl}^{CD} + \left(\eta_{xx}^{CD} + \eta_{yy}^{CD}\right)}$$

Since the terms in parentheses $\left(\eta_{xx}^{CD} + \eta_{yy}^{CD}\right)$ are identical for each commodity, and therefore if $\eta_{xl}^{CD} > \eta_{yl}^{CD}$ then $t_x^* > t_y^*$. Here a low or negative cross elasticity implies a high complementarity with leisure, the untaxed good. We can indirectly shift some of the burden onto the untaxed leisure market by taxing its complement, e.g. say y were holiday vacations, just as we might tax cigarettes by taxing lighters.