The Principle of Comprehensive Cost,
Applied to the Relation between Market Structure and Excise Tax Incidence

Perfect competition or Bertrand competition cause price to equal marginal cost, so that excise tax incidence can be illustrated with the most basic supply-demand diagram. Some results are:

- excise taxes are shared by suppliers and demanders according to their relative price elasticities
- holding revenue fixed, it doesn't matter (for supply price, demand price, and quantity traded) whether the excise tax is *ad valorem* or specific[1]
- a tax's incidence is fully summarized its effect on demand price

While price equals marginal cost adequately characterizes the markets for the vast majority of commodities, there are some situations in which price sufficiently exceeds marginal cost that these familiar tax incidence results fail. Furthermore, the latter cases may receive extra weight in policy and academic circles, because they potentially involve firms with market power, which generates lawsuits from competitors, which in turn generate data for empirical academic studies! However, these notes show that a simple and more general principle of price theory which applies to imperfect competition as well, as long as pricing is sufficiently sensitive to demand conditions. It is (what I call) the principle of comprehensive cost:[2] that a commodity's demand price is determined *as if* all taxes, travel costs, and other costs incurred by demanders (as a consequence of purchasing the commodity) were instead incurred directly by suppliers, who treat those costs in their pricing decisions as if they were any other cost of production. Hence, for example, excise tax incidence is different under monopoly than under perfect competition for exactly the same reason these market structures differ in terms of how production costs are passed on to consumers.

This means that measured markup rates are an indicator of how much excise taxes are passed on. The accuracy of this indicator depends on whether marginal and average markups coincide, and whether suppliers markup up their prices over their own costs, over their competitors' costs or some combination.

Throughout this chapter, q denotes the quantity traded per firm, and p the price paid by demanders inclusive of all taxes. v is profits per firm, and $\hat{c}(q)$ is the average cost function of a typical firm. τ and t are the rates of *ad valorem* and specific taxation, respectively, although the definition of the value to which τ applies varies across examples below. In order to understand the various results, the reader is advised to pay careful attention to the distinction between demand price and supply price.

---

[1]*Ad valorem* taxes are a function of the value of the transaction between buyer and seller. Specific taxes are a function of the real quantity exchanged.

[2]It is a simple application of the Coase theorem, which says that the market behaves as if the supplier treats costs as his own, even if the costs fall on third parties.

I. Regardless of Market Structure, Prices are Determined by "Supply" and Demand

In the perfectly competitive case, the distributional effect of an excise tax is fully summarized (at least in the neighborhood of the no-tax equilibrium) by its effect on price. This is not true in general; we have to separately calculate price and distributional effects. When it comes to the price effect, a very general result is available: taxes increase demand price according to the factor $\varepsilon_S/(\varepsilon_S+\varepsilon_D)$, where ε_D is the absolute value of the price elasticity of industry demand. ε_S is the price elasticity of industry "supply," although we have to be careful about the supply concept, because the supply curve need not be the same as the marginal cost schedule.

Let us define industry "supply" as the locus of prices and industry quantities that would be traced out by proportional shifts in demand. This locus is a combination of each firm's marginal cost function, their markup decisions, and entry behavior.[3] In some cases, the locus may slope down.

Since an *ad valorem* tax levied on demanders amounts to a proportional shift in demand from the suppliers' point of view, the equilibrium supply price and demand price under any *ad valorem* tax policy must lie on this locus we call "supply." To calculate $dp/d\tau$, we differentiate the equilibrium condition that supply equals demand, and solve for $dp/d\tau$, as shown below:

$$\bar{D}(p) = \bar{S}\left(\frac{p}{1+\tau}\right)$$

$$\bar{D}'\frac{dp}{d\tau} = \bar{S}'\frac{d}{d\tau}\left(\frac{p}{1+\tau}\right) \tag{1}$$

$$\frac{dp}{d\tau} = \frac{p}{1+\tau}\frac{\varepsilon_S}{\varepsilon_S+\varepsilon_D}$$

where S and D indicate supply and demand, and overbars are used to emphasize that fact that we are looking at industry quantity. Equation (1) applies for all market structures, as long as we use the appropriate concept of supply. It implies that the excise tax is shared as long as supply slopes up and demand is less than perfectly elastic, and that the degree to which taxes show up in equilibrium demand price increases with $\varepsilon_S/\varepsilon_D$.

Another definition of supply is the locus of prices and industry quantities that would be traced out by *parallel* shifts in demand in the price dimension. This locus is also a combination of each firm's marginal cost function, their markup decisions, and entry behavior. This supply curve need not slope up (see below), or be the same as the supply curve defined above, although it is in the competitive case. Since a specific tax amounts to a parallel shift in demand, a derivation like the one shown for equation (1) shows how $dp/dt = \varepsilon_S/(\varepsilon_S+\varepsilon_D)$ where t is the

---

[3]In the competitive case, it is merely the industry marginal cost function.

specific tax measured in dollars per unit quantity sold. Notice the similarity to equation (1),[4] except that ε_S has a different meaning here. Since in either case ε_S may be negative, either tax may be "over-shifted" ($dp/d\tau > p/(1+\tau)$ and/or $dp/dt > 1$).

In order to say more about the distribution of the tax burden, we first must establish the key industrial organization result which will be the basic ingredient for our calculations –

As long as supply price differs from marginal cost, neither of these supply curves can be used to calculate a producer surplus in the usual fashion, because the producer derives surplus from the gap between supply price and marginal cost. This implies that dp/dt does not fully summarize the distribution of the tax burden.

II. Market Structure I: Profits are Limited by Demand
II.A. Taxes are Like Any Other Cost, Regardless of Who Pays Them

In order to say more about the distribution of the tax burden, we first must establish the key industrial organization result which will be the basic ingredient for our calculations – namely, demand conditions dictate the degree to which demand price is marked up over marginal cost, and the relevant concept of marginal cost does not distinguish between technological costs $\hat{c}(q)$ and tax costs, regardless of who is legally liable for taxes. The result is easy to see from the profit function. First let's consider taxes paid by demanders, and let $\hat{p}$ denote the price received by the supplier, so that $p = (1+\tau)\hat{p} + t$. In other words, when a unit is sold at price $\hat{p}$, the buyer is obligated to pay the treasury a specific tax in the amount t, plus a fraction τ of the value $\hat{p}$ of the transaction.

Let $D(p)$ be the demand function facing the firm, which is shifted by income, prices charged by competitors, number of competitors, etc., although in the partial equilibrium spirit we neglect (and assume that suppliers neglect) the income effects of p.[5] Seller profits are:

$$v(p) \;=\; \left[\hat{p} - \hat{c}(D(p))\right]D(p) \;=\; \left[\frac{p-t}{1+\tau} - \hat{c}(D(p))\right]D(p) \tag{2}$$

where we obtain the second equality from the relation $p = (1+\tau)\hat{p} + t$ between supply and demand price. The second equation (2) can be used to show that, holding revenue constant, the profit function would be the same function of demand price p even if sellers paid all taxes. In this case, $p = \hat{p}$ and revenue per unit would be $\hat{p}(1-\hat{\tau})-t$, which is the same as $(p-t)/(1+\tau)$ as long as $\hat{\tau}$ were set so that the same *ad valorem* revenue is collected regardless of whether buyers or sellers are

---

[4]Remember that τ and t are measured in different units; τ is a fraction of the supply price $p/(1+\tau)$.

[5]Since we view p as a choice variable, our model is literally interpreted as a model like "monopoly" or "monopolistic competition" where each firm has market power (although perhaps other firms will be selling products that are close substitutes), instead of a Cournot competition model in which price depends on the reactions of competitors who sell perfect substitutes. However, we show below that the results hold in the Cournot model too.

liable.[6]

Let us define *comprehensive average cost* $c(q)$ to be the sum of technological costs $\hat{c}(q)$ and tax costs:

$$c(q) \equiv (1 + \tau)\hat{c}(q) + t$$

If price were equal to marginal cost, then this would be combined tax and technological cost per unit sold. Let's also define marginal comprehensive cost $MC(q)$ to be $d[c(q)q]/dq$. With the definition of $c(q)$, the profit function becomes:

$$v(p) = \frac{\left[p - c(D(p))\right]D(p)}{1 + \tau}$$

Hence the profit maximizing demand price depends only on the shape of the market demand and cost functions; taxes affect demand price only according to their contribution to marginal *comprehensive* cost. The first order condition for profit maximizing demand price is:

$$\frac{\mu}{1 + \mu} \equiv \frac{p - MC(D(p))}{p} = \frac{D(p)}{-pD'(p)} \equiv \frac{1}{\eta_D(p)} \tag{3}$$

where η_D is the absolute value of the price elasticity of demand, and μ is the proportional markup of demand price over marginal comprehensive cost. In words, *demand price* is:

(a)   marked up over marginal *comprehensive* cost,
(b)   to a degree which depends only on the elasticity of demand facing the firm.

Hence, taxes contribute to demand price just as any other cost, namely through their contribution to marginal comprehensive cost. Taxes, like any other cost, do not affect the markup rate μ except to the extent that taxes move the market to a different position on the demand curve and the demand elasticity varies.

II.B. Specific Taxes are "More Distortionary"

Even though we assume that the only method of tax avoidance is demand reduction,[7] as opposed to falsifying revenue or real quantity transacted, or firms exiting the industry, specific

---

[6]When demanders pay excise taxes, *ad valorem* revenue per unit sold is $\tau p = \tau(p-t)/(1+\tau)$. When suppliers pay, *ad valorem* revenue is $p\hat{\tau}$. These are equal when $\hat{t} = \frac{\tau}{1+\tau}\left[1 - \frac{t}{p}\right]$.

[7]In our notation, "demand reduction" is consistent with substitution away from legal markets toward illegal markets, as long as we view the goods sold in the illegal market as imperfect substitutes.

taxes are more distortionary than *ad valorem* taxes. To see this, consider a substitution between τ and t that holds marginal comprehensive cost constant, and thereby has no effect on demand price or quantity transacted.

$$0 = \hat{MC}\,d\tau + dt \tag{4}$$

where the hat (^) indicates technological rather than marginal comprehensive cost. In words, quantity and demand price are unaffected when the specific tax and *ad valorem* tax are substituted in proportion to the technological marginal cost.

Total excise tax revenue per unit transacted is $t + \tau(p\text{-}t)/(1+\tau) = t + \tau\hat{p}$. If t and τ are substituted in proportion to technological marginal cost, revenue will not be constant unless $\hat{p}$ were the same as technological marginal cost. $\hat{p}$ exceeds technological marginal cost (at least if both profits and tax revenue were are nonnegative so that marginal comprehensive cost is at least as big as technological marginal cost – see equation (3)), so revenue can be added without affecting the market by substituting *ad valorem* taxation for specific taxation. Algebraically, follow the variation specified in equation (4), an calculate the change in revenue per unit:

$$d(T/q) = \frac{\hat{p} - \hat{MC}}{1 + \tau}\,d\tau \tag{5}$$

The reason for the dominance of the *ad valorem* tax is that it is a combination of a specific excise tax and a tax on pure profits. This is also the reason why specific and ad valorem taxes are equivalent under perfect competition and other cases when $\hat{p}$ equals technological marginal cost – there are no pure profits to tax.

One way to see how ad valorem combines specific excise and pure profit taxes is to follow the variation specified in equation (4) – remember that it holds quantity and demand price constant – and calculate the change in profits per unit sold:

$$d(v/q) = \left. -\frac{dt + \hat{p}\,d\tau}{1 + \tau}\right|_{\hat{MC}\,d\tau + dt = 0} = -\frac{\hat{p} - \hat{MC}}{1 + \tau}\,d\tau = -d(T/q)$$

In other words, substituting toward the *ad valorem* tax adds exactly as much to tax revenue as it takes from profits.

Equation (5) applies even when $t = 0$, which means that revenue can be enhanced by a specific *subsidy* financed with additions to the *ad valorem* tax. Indeed, with a large enough specific subsidy, and a large enough *ad valorem* tax rate, the government can capture any fraction of pure profits in the market without distorting quantity transacted any more than the firm would in the absence of taxes. Figure 1 illustrates with solid black and dashed blue curves showing the tastes and technology, respectively. The solid blue curve is the technological marginal cost function marked up according to demand conditions, in particular at the quantity

q it is marked up over marginal cost in the proportion $1/(\eta\text{-}1)$, so the equilibrium demand price is p as indicated in the Figure.[8] Since there are no taxes for this initial equilibrium, $p = \hat{p}$.

Now consider introducing an *ad valorem* tax at rate $\tau' > 0$. This would shift the demand curve down *proportionally* (by factor $1/(1+\tau')$) to the dashed red curve. This by itself would decrease the quantity traded (not shown in the Figure), but consider adding a specific subsidy in the amount τ' times the initial marginal cost. The specific subsidy shifts the demand curve back up, this time in a *parallel* fashion, by the amount $-\tau'/(1+\tau')$.[9] If q were the equilibrium quantity, then supply price would be marked up over technological marginal cost by $\mu/(1+\tau')$ rather than μ.

In fact, it is profit maximizing for suppliers to mark up $\mu/(1+\tau')$ rather than μ because specific taxes (subsidies) decrease (increase) the absolute value of the elasticity of demand with respect to $\hat{p}$. Algebraically,

$$\hat{\eta}_D \;=\; \eta_D \, \frac{d\ln p}{d\ln\hat{p}} \;=\; \eta_D \, \frac{p - t'}{p} \tag{6}$$

where η_D and $\hat{\eta}_D$ denote the absolute value of the elasticity of demand with respect to p and $\hat{p}$, respectively, evaluated at price p. This shifts "supply" to the dash-dot line, so that q is still an equilibrium. In other words, taxes (and other cost incurred by demanders) change demand conditions from the firm's point of view (see equation (6)), so that the firm changes its markup over marginal technological cost in such a way that the markup of demand price over market comprehensive cost is the same. The "invisible hand" of the marketplace is not fooled by conventions as to who nominally pays each cost: it is the combination of all costs that matters, and how that comprehensive total various with market behaviors (e.g., which costs are variable and which are fixed).

The differential incidence of specific and *ad valorem* taxes raises some interesting questions. To what extent has the shift over time to value-added taxes substituted for specific excise taxation and thereby taxed pure profits? Why do specific taxes persist in a number of

---

[8]The solid blue curve can be interpreted as a supply curve, in the sense that it determines supply price as a function of q. This supply curve can be traced out by proportionally shifting the demand curve, and for each demand curve position calculating equilibrium quantity and supply price.

However, unlike the competitive case, every point on the supply curve requires some knowledge of the demand curve for its calculation (namely, the demand curve determines the optimal markup over marginal cost). For the same reason, nonproportional demand shifts will shift the supply curve.

[9]Remember that we have assumed that the specific tax revenue is excluded from the *ad valorem* tax base, so a specific tax at rate t' only costs $t'/(1+\tau')$.

Figure 1 ignores demand shifts resulting from changes in the demand prices charged by competitors because, by construction, all demand prices remain constant.

markets, like tobacco, alcohol, and gasoline?[10] We can speculate that specific taxes are better when (a) the market is not fully competitive (so that the form of the tax actually matters) and (b) trade in the market generates negative externalities, and the amount of the externality depends on the real quantity consumed and not the value consumed, as it would, for example, if vehicles consumed gallons of gasoline and damaged the highways in fixed proportions.[11] Another possibility is that supplier profits play a useful allocative role, as in the Kay and Keen (1983) model of entry into a market subject to excise taxation. Yet another possibility is that specific taxes are easier to collect from suppliers than are *ad valorem* taxes, because the final value of the transaction is not known until the retail sale. *Ad valorem* taxes seem easy to collect at the retail level (most states have an *ad valorem* retail sales tax), although many states may not have the infrastructure to collect them at rates that vary by commodity.

II.C. Tax Incidence I: Demand Price

We already have the formula (1) for the effect of excise taxes on demand price. Our task now is to interpret the industry supply and demand curves featured in (1) in terms of the demand and technological costs functions D and $\hat{c}$ facing the typical firm. The demand curve facing a typical firm is $D(p; \tilde{p})$, where $\tilde{p}$ is the average price of competing products. The industry price elasticity of demand is the elasticity of $D(p; \tilde{p})$, assuming that $d\tilde{p}=dp$. $\eta_D > \varepsilon_D$, except for a monopoly in which case they are equal.[12] For simplicity, we assume that $\tilde{p}$ does not affect the elasticity with respect to p.[13]

Regarding supply, consider the locus of points in the $[q,\hat{p}]$ plane traced out by proportional shifts of D for the firm and its competitors. Consider the locus as the "supply curve" for a typical firm, and multiply by the (fixed) number of firms to obtain an industry supply curve. In order to calculate the elasticity of the industry supply curve, which we denote ε_S, we differentiate the optimal markup in the absence of taxes, using the fact that proportional shifts do not affect the demand elasticity, and thereby only effect the optimal markup via the changes in quantity traded associated with movements along the demand curve

---

[10]If we allow the potential for falsifying the real quantity traded (which would affect specific tax revenue without affecting *ad valorem* tax revenue), eg., by changing the number of cigarettes in a pack, or changing the length of a cigarette, that seems to further weaken the efficiency case for specific taxes. See Barzel (1976) for a theoretical analysis of this kind of substution, and an application to cigarettes, and Ljunge (2003) for an application of the effect of per bottle taxes on wine on the "quality" of wine consumed.

[11] Of the three examples, tobacco seems the least competitive, although maybe the case for large per cigarette externalities is also pretty weak.

[12]The firm's demand elasticity η_D can be interpreted as a market share weighted average of the industry elasticity ε_D and the elasticity of substitution in utility between the firm's product and the products of competitors.

[13]If it did, we would have to consider a slightly different concept of supply in order to derive the formula (1), because proportional demand shifts for the industry would not be the same as proportional demand shifts for a typical firm.

$$\ln p = \ln \hat{MC}(q) + \ln[1 + \mu(q)]$$

$$\varepsilon_S^{-1} = \frac{\hat{MC}'}{\hat{MC}}q + \frac{\mu' q / \eta_D}{\mu} \qquad (7)$$

In words, the elasticity of inverse supply is the elasticity of marginal cost adjusted for changes in markup due to movements along the demand curve. In regions of the demand curve where the price elasticity is constant, this adjustment is zero. When movement down the demand curve raises the price elasticity, this "supply curve" can slope down.[14]

We can rederive the formula (1) from the optimal markup we've calculated for the case in which profits are limited by demand. In this case, demand price is marked up over marginal comprehensive cost:

$$p = [1 + \mu(D(p;\tilde{p}))] MC(D(p;\tilde{p})) = [1 + \mu(D(p;\tilde{p}))][(1 + \tau)\hat{MC}(D(p;\tilde{p})) + t]$$

Recall from equation (3) that the shape of the optimal markup μ depends only on the elasticity of demand facing the firm, so that the schedule $\mu(q)$ does not shift with taxes. Totally differentiating this relation between demand price and comprehensive cost in the neighborhood of $t = 0$, assuming that $d\tilde{p} = dp$:

$$dp = \frac{\varepsilon_S}{\varepsilon_S + \varepsilon_D}[\hat{p}d\tau + (1 + \mu)dt] = \frac{\varepsilon_S}{\varepsilon_S + \varepsilon_D}\hat{p}d\tau + \frac{dt}{1 - \frac{\varepsilon_S - \varepsilon_D(\eta_D - 1)}{\varepsilon_S \eta_D}} \qquad (8)$$

where the first equality follows from the definition of the supply elasticity ε_S, and the final equality from the relation between markup and demand elasticity. ε_D is the absolute value of the price elasticity of *industry* demand (namely, when $d\tilde{p} = dp$). With $dt = 0$, this is exactly the formula from the competitive case; the shifting of the *ad valorem* excise tax depends on supply and demand elasticities exactly as it does in the competitive case.

---

[14]As an example, consider a firm with constant marginal cost and demand curve

$$q = p^{-\eta_o} + p^{-\eta_\infty} \quad , \quad \eta_o > \eta_\infty > 1$$

As we move down this demand curve, the price elasticity rises from η_∞ to η_o, so the optimal markup falls.

Although the specific tax is more likely to be over-shifted, ε_S can be negative, so either tax can be over-shifted. In other words, if we observed a market with over-shifting, in which the only shocks were taxes and other proportional demand shifts, we would observe a negative correlation between q and $\hat{p}$, and yet might interpret that the observations as tracing out a single supply curve.

Equation (8)'s first equality also shows how there are two kinds of over-shifting. One is associated with $\varepsilon_S < 0$ – price elasticity of demand increases (in magnitude) as we move down the demand curve – and is common to the specific and *ad valorem* taxes. The second type of over-shifting applies only to the specific tax, and occurs merely because there *is* a markup, regardless of whether that markup rises or falls.

Equation (8) shows that a specific tax increase bringing in the same revenue as an *ad valorem* tax increase will have a bigger impact on demand price as long as the markup is positive. This is just another symptom of the fact that the specific tax is more distortionary, and means that the specific tax may be "over-shifted" ($dp/dt > 1$) even when the *ad valorem* tax is not – namely when $\varepsilon_S > \varepsilon_D(\eta_D\text{-}1)$ and $\varepsilon_S > 0$. In this case, proportional demand shifts trace out an upward sloping supply curve, while parallel shifts trace out an downward sloping curve.[15]

II.D. Taxes are Markup up Like other Costs, but whose Markup and whose Cost?

The various examples above illustrate that excise taxes are marked up like other costs. But whose markup and whose costs? It is possible that industry prices are driven by the costs of a small group of fringe competitors, so that the ratio of industry average price to industry average cost does not accurately represent how taxes and other costs are passed on to consumers. To see this, consider a competitive fringe model with a specific tax. A small share of the industry consists (or, if pricing conditions were right, would consist) of a firm or group of firms whose price $\tilde{p}$ equals their constant technological marginal cost $\tilde{c}$ plus the specific tax rate t. The market share of the competitive fringe is $\max\{0, b\ln[p/(z\tilde{p})]\}$, with constants $b > 0$ and $z > 0$. In words, consumers in the market shift their demand toward the fringe as $p/\tilde{p}$ rises.

In order to starkly illustrate the point, we suppose that industry demand is inelastic, so that the dominant firm's profit function is:

$$v(p,\tilde{p}) = \begin{cases} p - c - b(p - c)\ln\frac{p}{z\tilde{p}} & \text{if } p \geq z\tilde{p} \\ p - c & \text{if } p < z\tilde{p} \end{cases}$$

If $b > 1$, the dominant firm may desire to set $p = z\tilde{p}$ in order to keep the entire market to itself, but it would never set $p < z\tilde{p}$. With $p = z\tilde{p} = z(\tilde{c}+t)$, taxes are passed through according to the demand parameter z:

$$\frac{dp}{dt} = z = (1 + \mu)\frac{c + t}{\tilde{c} + t}$$

---

[15]Remember from Section I that $dp/dt = \varepsilon_S/(\varepsilon_S+\varepsilon_D)$ as long as we adjust our concept of supply to be the locus of points in the $[q,\hat{p}]$ plane traced out by *parallel* demand shifts for the firm and its competitors.

where μ denotes the rate at which the dominant firm marks up its price over comprehensive marginal cost.[16] Since the dominant firm has 100% market share, μ is also the industry markup rate. However, μ is not the rate at which tax costs are markup up, because the dominant firm's price is marked up over *the fringe's cost* at rate $z = (1+\mu)(c+t)/(\tilde{c}+t)$. Indeed, a marginal change in the dominant firm's cost, holding the fringe's cost constant, has no effect on p. The reason why taxes are shifted is because they would be paid by the fringe competitors had they entered the market.

Note that the measured industry markup rate μ might over- or understate the amount of over-shifting. With $z < 1$, there is no over-shifting at all, even though the industry markup rate μ might be strictly positive. The confounding factor in this model is that the industry markup rate μ is measured relative to the wrong costs.

If b were small enough ($b < 1$ is sufficient), then the dominant firm does not price as low as $p = z\tilde{p}$. The dominant firm has less than 100% market share, and both dominant firm and fringe firm costs determine the dominant firm's price. I leave it to the reader to show that p is independent of z, and that specific excise taxes are shifted according to

$$\frac{dp}{dt} \in \left((1 + \mu)\frac{c + t}{\tilde{c} + t}, 1 + \mu \right) \quad \text{if } c \leq \tilde{c}$$

$$\frac{dp}{dt} \in \left(1 + \mu, (1 + \mu)\frac{c + t}{\tilde{c} + t} \right) \quad \text{if } c \leq \tilde{c}$$

where, as before, μ denotes the rate at which the dominant firm marks up its price over comprehensive marginal cost.

II.E. Tax Incidence II: Deadweight Cost Rectangles

The demand price effect is not sufficient to describe excise tax incidence, even as a first order approximation. Figure 2 illustrates the incidence of the *ad valorem* tax in the case that prices of competing products do not effect the *elasticity* of demand facing the firm. In this case, neither τ nor competitor prices affect the optimal markup and thereby the supply curve.[17] The no tax equilibrium has quantity q, price p, and deadweight cost equal to the area of the triangle E. Demand, marginal cost, and supply are displayed as solid and dashed curves as in Figure 1, except now we interpret the demand curve as a demand for the *industry*, which is steeper than the demand curve facing a particular firm. The marginal cost and supply curves are those for a typical firm in the industry, and the markup over marginal cost reflects the demand curve

---

[16]The $\varepsilon_S/(\varepsilon_S+\varepsilon_D)$ term is one in this example, because proportional shifts in demand have no effect on price.

[17]Figure 2 illustrates any case in which markup rates are constant. For example, markup rates might be set by "rules of thumb," by regulation, etc.

facing a typical firm (not shown in the Figure). Profits for a typical firm are equal to the area of the trapezoid vertically bounded by the marginal cost curve and $\hat{p}$, and horizontally bounded by 0 and q.

When an *ad valorem* tax is levied on demanders at rate $\tau' > 0$, the demand curve shifts down proportionally. The new quantity is $q' < q$, with corresponding prices p' and $\hat{p}'$ for demanders and suppliers, respectively. $p' = (1+\tau')\hat{p}'$, so tax revenue is equal to the area of the rectangle vertically bounded by p' and $\hat{p}'$, and horizontally bounded by 0 and q'. The firm loses profits, in an amount equal to the six-side polygon bounded above by p, and below by $\hat{p}'$ (to the left of q') and the marginal cost curve (to the right of q').

Notice that, even if marginal cost were constant (ie, drawn as a perfectly horizontal line in Figure 2), the firm would be losing the pure profits accruing to the units demanded above q'. Hence, even though the relative elasticity formula $\varepsilon_S/(\varepsilon_S+\varepsilon_D)$ describes the quantitative effect of taxes on demand price as it does under competition, it does not describe the share of the marginal tax burden borne by demanders (namely, in the form of foregone consumer surplus). This share can be calculated algebraically, and in the neighborhood of $\tau = 0$ is just a function of the three elasticities:

$$\text{demanders}'\text{ share of AV tax burden} = \frac{\varepsilon_S\,\eta_D}{\varepsilon_S\,\eta_D + \varepsilon_D\,(\varepsilon_S + \eta_D)} < \frac{\varepsilon_S}{\varepsilon_S + \varepsilon_D} = \frac{1}{\hat{p}}\frac{dp}{d\tau}$$

In other words, knowing the industry supply and demand elasticities, or the effect of taxes on demand prices, only puts an upper found on demanders' share of the tax burden; a complete calculation requires knowledge of the degree of competition η_D. This derives from the fact that the *ad valorem* tax is in part a tax on pure profits.

The demanders' share of the specific tax burden can be larger than $\varepsilon_S/(\varepsilon_S+\varepsilon_D)$, but is still less than dp/dt:

$$\text{demanders}'\text{ share of specific tax burden} =$$

$$\frac{\varepsilon_S\,\eta_D}{\varepsilon_S\,\eta_D + \varepsilon_D\,(\varepsilon_S + \eta_D) - (\varepsilon_S + \varepsilon_D)} < \frac{\varepsilon_S}{\varepsilon_S + \varepsilon_D}\frac{\eta_D}{\varepsilon_S + \eta_D} < \frac{dp}{dt}$$

The reason for the inequality is the same for both kinds of excise taxes: the shifting of the tax is offset by the loss of profit on the units above q'.

These formulas also show that the elastic side of the market tends to pay less, as in the competitive case. For example, the demanders' share is monotonic increasing in ε_S and monotonic decreasing in ε_D. However, suppliers still bear some of the burden even when supply is perfect elastic.

II.F. *Review of the Main Results*

When firms set prices in order to maximize profits, taking as given the prices of competing products, the same principle of comprehensive cost applies as in the competitive

situation. Demand price is determined *as if* all taxes incurred by demanders (as a consequence of purchasing the commodity) were instead incurred directly by suppliers, who treat those costs in their pricing decisions as if they were any other cost of production. What differs from the competitive case is the way in which costs are passed on to consumers, and the fact that there are deadweight losses even in the absence of taxes. Accounting for these differences when applying the principle of comprehensive cost delivers eight results:

- holding revenue fixed, it doesn't matter whether suppliers or demanders are legally liable for the tax payment – as in the competitive case
- *ad valorem* excise taxes are shifted to consumers in the form of higher demand prices according to the relative industry supply and demand elasticities – as in the competitive case
- holding revenue constant, specific excise taxes have a larger effect on demand price
- excise taxes can be over-shifted in the sense that they increase supply prices
- variable markup rates are one potential source of over-shifting, common to specific and *ad valorem* taxes
- positive markup rates are another potential source of over-shifting, but only for specific taxes
- excise taxes increase demand price, but the amount of the price increase overstates the fraction of the excise tax burden borne by demanders, because the demand reduction erodes supplier profits
- more elastic industry demand means less of the burden for demanders – as in the competitive case
- more elastic supply means less of the burden for suppliers, but suppliers still bear some burden even in the limit of perfectly elastic supply
- even small excise taxes have first order deadweight costs

III. Market Structure II: Profits are Limited by Entry
III.A. *General Model of Entry*

The demand facing a typical firm is the industry demand divided by the number of firms in the industry N. N was fixed above, so a proportional shift in industry demand resulted in the same proportional shift in the demand facing a typical firm. Here we endogenize the number of firms, and show how entry implies a larger industry supply elasticity, and hence the excise tax has a greater impact on demand price. Since entry is limiting profits, there is less scope for excise taxes to reduce profits, so that the distribution of the excise tax burden is in between the competitive model and the model in which profits are limited only by demand.

Regardless of whether there is the possibility of entry, assume that firms choose their prices taking as given the number of firms, and (as before) the prices of competing products. As we shift industry demand, this shifts demand facing the typical firm (although probably by a lesser amount, because some of the additional demand is accommodated by entry), and we can trace out a firm-level "supply curve" in the $[q,\hat{p}]$ plane as the locus of points consistent with profit maximization, and each point corresponding to a different level of industry demand and the same marginal cost function. As before, we cannot measure the producer surplus as the area above the supply curve. We also must remember that the shape of the supply curve depends on the nature of the demand shifts. As before, equation (7) describes the firm-level supply curve when the demand shifts are proportional, and we denote that supply function as $S(\hat{p})$.

The industry supply function is the product of $S(\hat{p})$ and the number of firms $N(\hat{p})$, which probably increases with the price suppliers can obtain for their product. Hence industry supply is more elastic than before (add the elasticity of N to the previous formula for the supply elasticity), but the formula (1) still describes the effect of *ad valorem* taxation on demand price as long as we continue to interpret ε_S as the elasticity of industry supply, which is a combination of each firm's marginal cost function, their markup decisions, and entry behavior.

III.B. *The Monopolistic Competition Model of Entry*

We can obtain more information about the entry function $N(\hat{p})$ by assuming "free entry" in the sense the equilibrium profits for the marginal entrant are zero. Free entry implies perfect competition if average costs are not falling at the firm level, and imperfect competition (a.k.a., "monopolistic competition") if average costs are falling over some range. We consider the latter case, in particular where there is a fixed startup or entry cost, and thereafter a nondecreasing marginal cost schedule (which is the same for all firms). In this case, profits are limited by the startup cost of the marginal entrant.

Let $f(N)$ denote the startup cost of the marginal entrant, as a function of the number of firms already in the industry. Obviously, low startup cost firms are more willing to enter, so $f' \geq 0$. The supply of entrants $N(\hat{p})$ is implicitly defined by the zero profit condition for the marginal entrant:

$$f(N(\hat{p})) = [\hat{p} - \hat{c}(S(\hat{p}))]S(\hat{p})$$

Hence the elasticity of $N(\hat{p})$ is the ratio of the nonnegative elasticity of profits per firm[18] to the nonnegative elasticity of f.

An interesting special case has a single startup cost for all firms, so that $f' = 0$. The industry supply is perfect elastic so that excise taxes are fully shifted to demanders, but not over-shifted. Furthermore, the area above this supply curve (namely, zero) measures the producer surplus as in the competitive case. Hence, producers do not bear any excise tax burden, at least to the extent that their startup cost f was paid in anticipation of the right amount of excise taxation. Compare zero burden with the uncontested monopoly case (see the formula in section II.D, with $\eta_D = \varepsilon_D$ and $\varepsilon_S = 0$), in which the burden on suppliers is as much as the burden on demanders. It matters whether profits are limited by entry, or just by industry demand.

IV. References

Barzel, Yoram. "An Alternative Approach to the Analysis of Taxation." *Journal of Political Economy*. 84(6), Dec. 1976: 1177-97.

Kay, J. A. and M. J. Keen. "How should commodities be taxed?: Market structure, product heterogeneity and the optimal structure of commodity taxes." *European Economic Review*.

---

[18]Obviously, firms benefit from favorable demand shifts. The elasticity of profits with respect to supply price is $\dfrac{S(\hat{p})}{f(N(\hat{p}))} \dfrac{\varepsilon_S - \varepsilon_N + \eta_D}{\eta_D}$, where ε_N is the elasticity of $N(\hat{p})$.

23(3), September 1983: 339-358.

Ljunge, Martin. "Do Taxes Produce Better Wine?" manuscript, University of Chicago, December 2003.

Figure 1: *Ad Valorem* Taxes Combine Specific Excise and Pure Profit Taxes

T' = government revenue
v' = profits w/ taxes
E = dwc of monopoly

firm's supply curve

demand price
$= p = \hat{p} =$

$(1+\mu)\hat{MC}(q)$

T'

firm's marginal cost curve

$\hat{p}'$

E

$v' = \dfrac{v}{1+\tau'}$

firm's demand

p

q

q

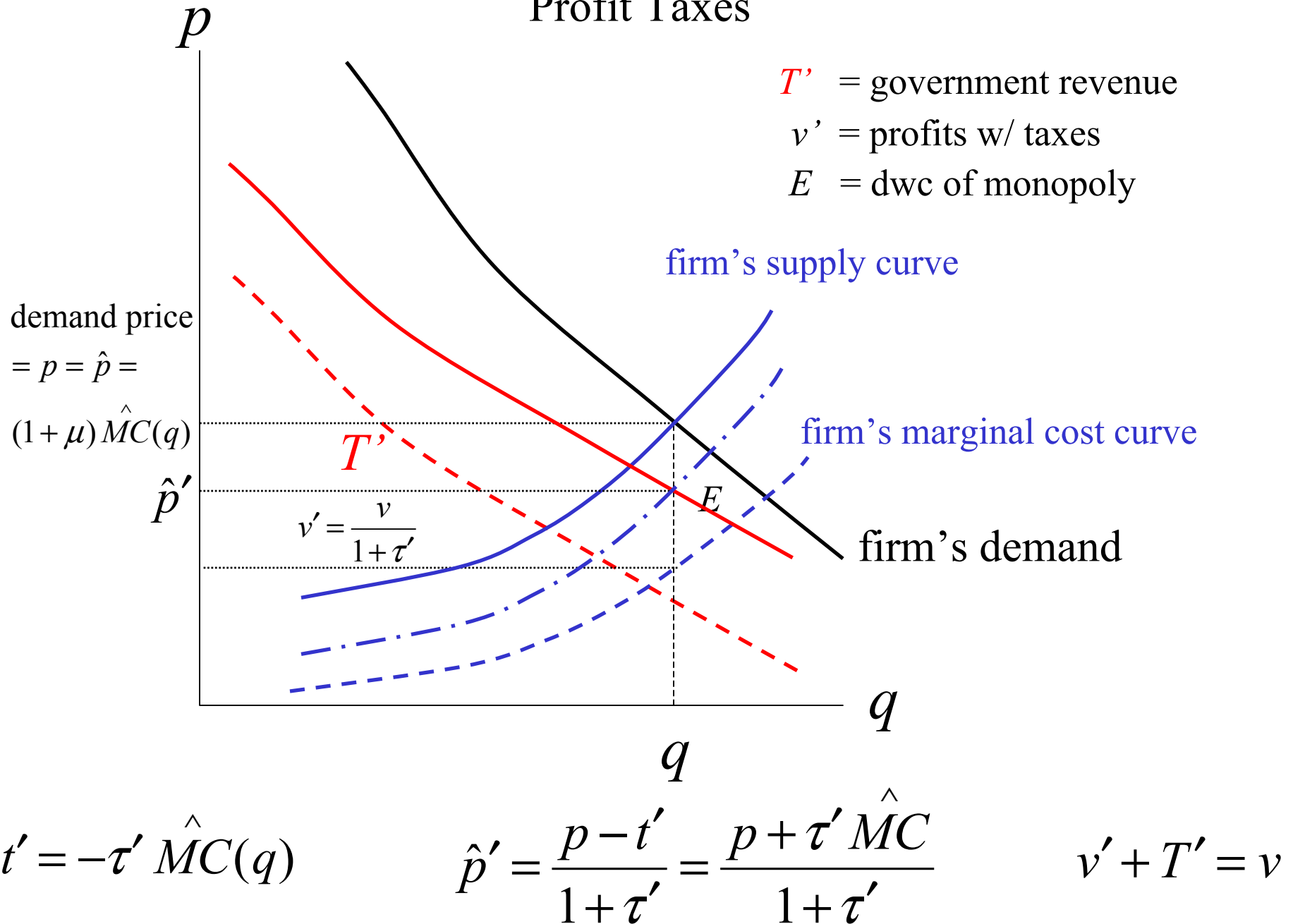$$t' = -\tau'\,\hat{MC}(q) \qquad \hat{p}' = \frac{p-t'}{1+\tau'} = \frac{p+\tau'\,\hat{MC}}{1+\tau'} \qquad v' + T' = v$$

Figure 2: Tax on Demanders

p

after-tax price
$= (1+\tau)(1+\mu)c$

no-tax price
$= (1+\mu)c$

c

T

D

Δv

E

$T+D$ = borne by demanders
Δv = lost profits
T = government revenue
$D+\Delta v$ = dwc of taxes
E = dwc of monopoly

industry supply curve

industry cost curve

industry demand

q' q

q