<u>The Size Distribution of Firms, Cournot and Optimal Taxation</u>

by

Mark Gersovitz[1]
The Johns Hopkins University

[1] I thank James R. Hines Jr. for comments.

Abstract

Tax laws and administrations often treat different sized firms differently. There is, however, little research on the consequences. As modeled here, oligopolists with different efficiencies determine the size distribution of firms. A government that maximizes a weighted sum of consumer surplus, profits and tax receipts can tax firms with different efficiencies differently and provides a reference point for other, more restricted differential tax systems. Taxes include a specific sales tax, an ad valorem tax and a profits tax with imperfect deductibility of capital cost, and a combination of the last two. In general there is a pattern of tax rates by efficiency of firm. It is heavily dependent on the social valuation of tax receipts. For instance, less efficient firms pay higher specific sales taxes if the social value of receipts is below 2 and lower taxes if the social value of receipts is above 2. Other analytic and simulation results are provided. When both ad valorem taxes and the imperfect profits tax are combined, simulations suggest that the former rate is higher and the latter rate is lower when the firm is less efficient.

0. Introduction.

Especially in poor countries, most practitioners would agree that tax laws and administrations often distinguish among firms by their sizes. Gauthier and Gersovitz (1997) provide the only case based on a survey of firms which I know that documents this phenomenon, for the example of Cameroon. Whether tax laws and administrations do or not, it is of interest to ask about the effects of tax treatment that depends on firm size, and in particular when such differentiated treatment may be optimal and whether the correlation between firm size and tax rates should be positive or negative. This topic has not, however, been addressed through formal models with the exception of Keen and Mintz (forthcoming).

The first ingredient for such an analysis is a theory of the size distribution of firms. The literature of industrial organization does not, however, provide strong guidance on what this theory should be. In their analysis of thresholds in taxation and the size distribution of firms, Keen and Mintz (forthcoming) link the size of a firm to the inherent talent of its (single) entrepreneur in co-operating with other factors, an assumption shared with a literature outside public finance (Lucas, 1978 and Kanbur, 1979). These firms behave atomistically, and the interaction between the size distribution of firms and the optimal tax structure arises because sufficiently small firms do not provide enough tax receipts to justify the expenses of gathering it, and are therefore exempt.

The alternative that I explore is that firms are asymmetric Cournot oligopolists with different, constant unit costs of production. The distribution of these costs mediated by the nature of Cournot oligopoly then determines the size distribution of firms.[2] Such situations naturally give rise to pure profits. An important focus of this paper is therefore a profits tax, although one that for reasons of practicality only allows imperfect deductibility of capital costs; I term such a tax a hybrid profits tax because it combines a tax on pure profits with a tax on capital. In contrast to a tax on pure profits, it induces a substitution away from capital and a consequent deadweight loss and is not optimally imposed at one hundred percent. The efficiency of this type of hybrid profits tax does not seem to have been studied even under the assumption of symmetric Cournot oligopoly, but it is ubiquitous and is therefore of interest independently of a concern with the size distribution of firms.[3]

The question of the paper is: What is the relation between the size of firms and the (optimal) taxes that they should pay? The tax authority in the subsequent models can choose a tax that is potentially distinct for each firm to allow for an examination of this question in the least constrained way possible. Of course, the intent is not to propose that actual tax systems subject each firm to a tailor-made tax. Rather the goal is to provide a reference point for tax systems that differentiate tax rates by firm size in one way or another. For instance, when the analysis of this paper shows that smaller (less efficient) firms should be taxed less than larger

---

[2]The literature on industrial organization usually maintains the assumption of symmetric oligopolists. Salant and Shafer (1999) show how asymmetries in costs can arise through a two-stage game in which the first stage determines costs. Lahiri and Ono (1988) consider the benefits of an industrial policy that discriminates against, and even removes, small high-cost firms, and therefore has similarities to the section on specific taxes in this paper although it does not look at the benefits of raising tax receipts. Another theory of the size distribution with implications for taxation is based on credit rationing (for instance, Cabral and Mata, 2003), but it is not the one pursued here.

[3]Desai and Hines (2004) study incidence of a hybrid profits tax.

firms then it lessens concern about the tendency to be lax in going after small firms for reasons of administrative cost. On the other hand, when the model suggests that small firms should be taxed at a high rate then there is some reason to increase efforts to reach small firms despite high administrative costs.

The next section sets out the notation of the paper and some reference cases as well as the structure of the simulations. The three subsequent sections consider in turn the interaction of the size distribution of firms and three taxes, each in isolation: a specific sales tax, an ad valorem sales tax, and a profits tax with incomplete deduction of capital cost. The next section looks at the combination of the two most common of these taxes, the ad valorem sales tax and the tax on profits with imperfect deductibility of capital cost. A final section provides some concluding comments.

1. Asymmetric Oligopolists in an Untaxed Economy.

1.1 The Basic Notation and Structure of the Model:

In this model, consumers purchase the quantity M and pay a price P. The function P(M) is therefore the demand curve with the properties that:

$$(1.1a) \qquad P_M < 0$$

and

$$(1.1b) \qquad P_M + m_j P_{MM} < 0 \qquad \forall\, j,$$

the stability condition in oligopoly models in which m_j is the output of the jth firm. The well-being of consumers is measured by their consumer surplus.

There are n firms each of which has different costs of production. To produce a unit of output at minimum cost, the best practice firm, labeled j=1 with efficiency $\alpha_1 = 1$, minimizes its cost (C_1):

$$(1.3) \qquad C_j = w_j^* l_j + r_j^* k_j \ , \quad j = 1 \ ,$$

subject to the constraint given by its constant-returns-to-scale production function :

$$(1.4) \qquad 1 = F(k_j, l_j)/\alpha_j \ , \qquad j = 1$$

with respect to its factor inputs, capital (k_1) and labor (l_1) that cost r_1* and w_1* per unit inclusive of any tax considerations. Before tax considerations, the market prices of capital and labor are r and w which are exogenous to the industry under consideration, and if there are no tax considerations, $r_1* = r$ and $w_1* = w$. The remaining (n-1) firms are inherently less efficient than the first firm, and solve the same cost minimization problem as the first firm as given by equations (1.3) and (1.4) except that $\alpha_j > 1$ for j>1 and that they may face different tax-inclusive factor costs, r_j* and w_j* . If all firms do face the same tax-inclusive factor prices, however, their costs are $C_j = \alpha_j C_1$.

On the supply-side the n firms behave as Cournot oligopolists, taking the output of the other firms as invariant to their own behavior. On the assumption that all firms can make positive profits in equilibrium, and in the absence of government intervention, the jth firm receives total profits (Π_j):

$$(1.5) \qquad \Pi_j \equiv m_j P - m_j C_j \ ,$$

by producing output of m_j and chooses its level of output to satisfy:

$$\text{(1.6)} \qquad \frac{\partial \Pi_j}{\partial m_j} = P + m_j P_M - C_j = 0.$$

Given that all these firms operate in the same market (M is the same) and face the same tax-inclusive prices (r and w because there are no taxes), equation (1.6) implies that there is an inverse relation between m_j and C_j (and therefore α_j) and so less efficient firms produce less. A question addressed in subsequent sections is whether firms with a higher α_j should be taxed less or more, and if less whether so much less as to more than offset their inefficiency.

Unlike perfect competition where there might be an infinite number of firms of each efficiency and only the most efficient firms will produce, the oligopolistic structure of the model provides scope for the inefficient firms to stay in production even when costs are constant for each firm but differ across firms. Although the most efficient firm wants to charge the monopolist's price, the less efficient firms can be profitable below this price and the most efficient firm only finds it profitable to restrict its output to a limited degree. The consequent losses of the most efficient firm relative to its being a monopolist are offset (possibly more than offset) by gains that accrue to consumers and to the less efficient firms, but also are associated with social losses from the higher production costs of the inefficient firms. The existence of the less efficient firms thus has the advantage of mitigating the most efficient firm's market power and the disadvantage of doing so at higher cost.

The social valuation of outcomes (V) in this economy is the sum of total producer surplus (profits, Π) plus consumer surplus (S) plus the social value of total tax receipts (γR in which γ is the social value of a dollar in the government's hands and R is total tax receipts):

$$(1.7) \qquad V = \Pi + S + \gamma R \ .$$

The ideal policy prescription for this economy is clear: The most efficient firm should produce everything, so that all other firms should be banned or otherwise induced to cease production. The now monopolist should be regulated so that it prices at marginal cost plus any sales taxes. If it is desirable to raise tax receipts, the government should tax sales of the commodity using either a specific or ad valorem tax. These outcomes can also be achieved either by a combination of an ad valorem tax on and a specific subsidy to the most efficient firm or by a combination of a profits tax and an ad valorem or specific subsidy (in both cases in conjunction with taxes that induce the other firms to exit).[4]

The whole starting point of this paper is that the Ideal Outcome is not a feasible tax system. Instead, tax systems especially those in poor countries which are relatively simple in scope, basically have two tools to hand: a sales tax and a distortionary profits tax that taxes profits without allowing full deduction of costs, most especially the costs of capital. So the goal is to understand the implications for this type of tax system of a size distribution of firms based in a few firms with asymmetric costs. Some aspects of this problem can only be understood through simulation and before turning to various aspects of the problem, the next sub-section lays out the structure of the simulations.

1.2 Details of the Simulations:

The assumptions used in the numerical analysis are:

[4]These conclusions follow from the observation that either pair of two instruments is enough to ensure that the most efficient firm produces the same quantity as the hypothetical regulated monopolist and earns zero profits. The analysis of Myles (1996) applies to these cases.

- The demand curve is linear: $P = 10 - 0.5\,M$.
- The wage that labor receives is $w = 1$ and the payment to capital is $r = 1$. These payments are the factor prices that firms respond to under the two sales taxes, while under the profits tax the jth firm effectively faces the net of tax factor prices $w^* = (1-\tau_j)\,w$ and $r^* = (1-\delta\tau_j)\,r$ in which τ_j is the rate of tax on profits and δ is the fraction of the cost of capital that is deductible.
- The production function is Cobb Douglas with a capital exponent of $\epsilon = 0.25$:

  $$m_j = (k_j^{\epsilon} l_j^{1-\epsilon})/\alpha_j \ .$$

  A Cobb-Douglas production function with its unitary elasticity of substitution probably overstates the possibilities for substitution between factors and therefore also overstates the scope for distortionary losses from the incomplete deductibility of capital costs. When $\alpha_j = 1$ and the tax system does not affect either the wage or the cost of capital, the cost minimizing combination of factors to produce a unit of output is $l = 1.316$ and $k = 0.439$, so that unit cost is 1.755. When $\alpha_j > 1$ the unit cost is $1.755\,\alpha_j$ because the production function exhibits constant returns to scale.
- If there were many identical firms all with $\alpha_j = 1$ and no taxes, then the competitive price would be $P = 1.755$, the equilibrium quantity would be $M = 16.49$ and social welfare would be $V = S = 67.98$. If there were only one firm with $\alpha_j = 1$ and no taxes, then marginal cost would be 1.755 and the monopoly quantity at which marginal revenue equals this marginal cost would be $M = 8.245$, the corresponding price would be $P = 5.878$ and social welfare would be $V = 50.99$.

- The simulations assume three firms with $\alpha_1 = 1$, $\alpha_2 = 1.20$ and $\alpha_3 = 1.30$. When there are no taxes, these firms altogether produce M=11.928 at a P = 4.036, roughly halfway between a perfectly competitive, fully efficient situation and the monopoly situation. The corresponding value of social welfare is V= 59.58. This welfare outcome is between the outcomes of the polar cases of monopoly and perfect competition, and indicates that in this case the benefits of some competition outweigh the costs of inefficiency, but the reverse could be the case (Lahiri and Ono, 1988). In general, therefore, the oligopolistic situation, although it has only three firms and two of them are distinctly inefficient, is halfway between the case of pure monopoly and the case of perfect competition with price determined by the cost of the most efficient firm.
- The value of γ ranges from 1.15 all the way through 2.5. This range would seem to bracket the range of estimates in the literature.[5]
- Table 1 provides information on the values of γ and the associated specific taxes that maximize social welfare, the level of total output and social welfare for the Ideal Outcome, i.e. when price equals the constant marginal cost of the most efficient firm plus the specific tax. These values provide reference points for the more restricted tax packages of subsequent sections. In particular, the simulations of the different tax policies present the value of an index, Φ, for a given value of γ, of the ratio of the difference between social welfare under the particular policy and social welfare if nothing is done (V=59.58) to the difference between social welfare under the Ideal Outcome (last

[5]See Diewert et al (1998) for a survey of recent estimates of γ and Devarajan and Thierfelder (2001) for some of the few estimates for poor countries. Keen and Mintz (forthcoming) use $\gamma = 1.3$.

column of Table 1) and social welfare if nothing is done. If the Ideal Outcome is thought to be the best plausible policy, then Φ is measuring the proportion of the total improvement achieved by the Ideal Outcome that any more practical policy can realize. By looking at a ratio of differences one gets rid of any artifacts of comparison that might arise from a linear transform of social welfare.

The calculations were done in double-precision FORTRAN and searched a three-dimensional grid of the tax rates with each tax rate ranging from 0.0 to 0.99 in increments of 0.01 for the ad valorem and profits taxes and ranging from 0.0 to 3.5 in increments of 0.03 in the case of the specific tax. I have chosen to bound taxes below by 0.0 to remain consistent with the notion of examining differentiated taxes that might correspond to the way tax codes and administrations often treat different sized firms differently. Actual subsidies, when they exist at all, would seem to be more a feature of specialized programs than general tax laws. (The calculation strategy when there are two taxes is discussed in Section 5.)

2. Asymmetric Oligopolists and the Specific Sales Tax.

The government imposes a specific sales tax, possibly of a different amount, on each of the n firms. After paying the tax, the jth firm makes profits of Π_j:

$$\Pi_j \equiv m_j P - m_j (C_j + t_j) \ , \tag{2.1}$$

in which: C_j is the constant cost to the jth firm of producing a unit of output as determined by cost minimization that is independent of the specific tax and t_j is the specific tax per unit sold that the jth firm pays. In this and the subsequent sections, the government may choose different taxes for different firms. If it does so, these taxes are based on the efficiencies of the firms and not on their sizes or other endogenous attributes that the firms could choose in response to differential taxation, in contrast to Keen and Mintz (forthcoming). Note, however, that because there are discrete differences among a finite number of firms in the economy I consider, if tax rates are not too different relative to efficiencies it should not pay a firm of one efficiency to produce at an otherwise unprofitable size to be eligible for a lower tax rate but I do not provide a formal derivation of this condition.

The jth firm maximizes profits with respect to the amount it produces yielding:

$$(2.2) \qquad \frac{\partial \Pi_j}{\partial m_j} \;=\; P + m_j P_M - (C_j + t_j) \;=\; 0.$$

Differentiation of the jth firm's first-order condition, equation (2.2), with respect to the ith tax rate provides a pair of results useful in the derivations that follow:

$$(2.3a) \qquad i \neq j \;\Rightarrow\; P_M \frac{dm_j}{dt_i} \;+\; m_j P_{MM} \frac{dM}{dt_i} + P_M \frac{dM}{dt_i} \;=\; 0$$

and

$$(2.3b) \qquad i = j \;\Rightarrow\; P_M \frac{dm_i}{dt_i} \;+\; m_i P_{MM} \frac{dM}{dt_i} + P_M \frac{dM}{dt_i} - 1 \;=\; 0\ .$$

For instance, the summation of equations (2.3 a) and (2.3b) over all n firms produces an expression for the change in the total market quantity in response to a change in the ith tax:

$$\text{(2.4)} \qquad \frac{dM}{dt_i} = \frac{1}{(n+1)P_M + MP_{MM}} < 0 \ ,$$

which is independent of i, a result familiar from Bergstrom and Varian (1985).

Differentiation of equation (2.1) with respect to t_i, the use of equations (2.3a) and (2.3b) to substitute for dm_j / dt_i, and summation over the n firms produces:

$$\text{(2.6)} \qquad \frac{d\Pi}{dt_i} = \sum_{j=1}^{n} [2m_j P_M + m_j^2 P_{MM}] \frac{dM}{dt_i} - 2m_i \ .$$

The change in consumer surplus is given by:

$$\text{(2.7)} \qquad \frac{dS}{dt_i} = -MP_M \frac{dM}{dt_i} \ .$$

The total tax receipts raised is the sum of the receipts from each of the n firms:

$$\text{(2.8)} \qquad R = \sum_{j=1}^{n} R_j = \sum_{j=1}^{n} t_j m_j \ .$$

Differentiation of equation (2.8) with respect to t_i, and the use of equations (2.3a) and (2.3b) to substitute for dm_j / dt_i produces:

$$\text{(2.9)} \qquad \frac{dR}{dt_i} = m_i + \frac{t_i}{P_M} - \sum_{j=1}^{n} \left(\frac{m_j P_{MM} + P_M}{P_M} \right) \frac{dM}{dt_i} \ .$$

Equations (2.6), (2.7) and (2.9) provide the components of the change in social welfare from a change in the specific tax on the ith firm. This change set to zero for all i=1... n firms defines the optimal structure of taxation:

$$(2.10) \qquad \frac{dV}{dt_i} = -2m_i + \gamma(m_i + \frac{t_i}{P_M}) + \Omega = 0 \ , \qquad i = 1 \ldots n \ ,$$

in which Ω impounds all the terms from equations (2.6), (2.7) and (2.9) that do not depend on i. The components that depend on i and that will play a crucial role in the subsequent analysis arise from the effects of a change in the ith tax on profits ($2m_i$) and on tax receipts ($m_i + t_i/P_M$) which in turn depend on the effect of a change on the ith firm's first-order condition as given by equations (2.3a and b). Note that the component of the change in social welfare that derives from the change in consumer surplus, equation (2.7), is entirely independent of the firm under consideration via equation (2.4), and is therefore impounded in Ω. This last result is particular to the specific tax as shown in subsequent sections.

Given an optimal tax structure, the question is whether low- or high- cost firms are taxed more in the optimal tax equilibrium. This question is answered by treating as constant all terms that do not depend on the firm under consideration, for instance the terms Ω and M, and then comparing the optimal tax rate that is applied to a higher rather than a lower cost firm. Note that this question is different from asking what would happen to the whole equilibrium if one firm's unit cost increased; in this case, such terms as Ω and M would be endogenous and could not be treated as constant. Equation (2.10) and the first-order condition equation (2.2) provide two equations in two unknowns, t_i and m_i:

$$(2.11a) \qquad t_i = \frac{P_M(2-\gamma)}{\gamma} m_i - \frac{P_M \Omega}{\gamma}$$

and

$$(2.11b) \qquad t_i = P_M m_i + (P - C_i) \ .$$

There are two cases, illustrated in Figures 1 and 2.

CASE 1: $(2 - \gamma) > 0$ and $\gamma > 1$. In this case, both equations (2.11a) and (2.11b) have negative slopes, and the slope of the former is less in absolute value than that of the latter. Note that C_i appears only in equation (2.11b). An increase in C_i therefore shifts only the curve for equation (2.11b), and shifts it down. Consequently, firms with a higher C_i face a higher t_i, and for both reasons have a lower m_i (Figure 1).

CASE 2: $(2 - \gamma) < 0$. In this case, equation (2.11a) has a positive slope and (2.11b) has a negative slope. As in the preceding case, an increase in C_i shifts only the curve for equation (2.11b), and shifts it down. Consequently, firms with a higher C_i face a lower t_i, but not so much lower that they do not have a lower m_i; the lower t_i less than offsets the higher C_i (Figure 2).

Table 2 presents the results from simulating the government's problem of maximizing social welfare, V. At relatively low levels of γ, including1.3, some (or all) of the taxes are zero. Recall from section 1.2 that the grid constrains the tax rates to be non-negative, and the binding nature of these constraints in these cases reflects the fact that with imperfect competition and a low valuation of tax receipts, there is the traditional motive to subsidize the output of the relatively efficient firms by analogy to the classic result of the subsidy to a monopolist. Not surprisingly, therefore, a tax constrained to be non-negative does particularly badly (low value of Φ) relative to the Ideal Outcome that allows for such a subsidy when it would be necessary (low values of γ). Even without the possibility of a subsidy, the output of the most efficient firm and its profits rise as γ rises from 1.15 to 1.30 before falling as tax rates rise. The last column of Table 2 provides the value of social welfare when all the specific taxes are constrained to be the same. For the given parameters, social welfare is not very different when this constraint is imposed.

3. Asymmetric Oligopolists and the Ad Valorem Sales Tax.

The government imposes an ad valorem sales tax, possibly at a different rate, on each of the n firms. After paying the tax, the jth firm makes profits of Π_j:

$$(3.1) \qquad \Pi_j \equiv (1-t_j)m_jP - m_jC_j \ ,$$

in which t_j is the ad valorem tax on the value of sales that the jth firm pays. The jth firm maximizes profits with respect to the amount it produces yielding:

$$(3.2) \qquad \frac{\partial \Pi_j}{\partial m_j} = (1-t_j)(P+m_jP_M) - C_j = 0.$$

Differentiation of the jth firm's first-order condition, equation (3.2), with respect to the ith tax rate and substitution of the first-order condition provide a pair of results useful in the derivations that follow:

$$(3.3a) \qquad i \neq j \Rightarrow P_M\frac{dm_j}{dt_i} + m_jP_{MM}\frac{dM}{dt_i} + P_M\frac{dM}{dt_i} = 0$$

and

$$(3.3b) \qquad i{=}j \;\Rightarrow\; P_M\frac{dm_i}{dt_i} \;+\; m_iP_{MM}\frac{dM}{dt_i}+P_M\frac{dM}{dt_i}-\frac{C_i}{(1-t_j)^2} \;=\; 0 \;.$$

Summation of equations (3.3 a) and (3.3b) over all n firms produces an expression for the change in the total market quantity in response to a change in the ith tax:

$$(3.4) \qquad \frac{dM}{dt_i} \;=\; \left[\frac{1}{(n+1)P_M+MP_{MM}}\right]\;\left[\frac{C_i}{(1-t_i)^2}\right] < \; 0 \;.$$

In contrast to the result for a specific tax, this expression is not independent of i but varies directly with C_i (and therefore with α_i) and t_i. Consequently, the expression for the optimal tax structure is inherently much more complicated for the ad valorem than for the specific tax.

As in section 2, the goal in designing taxes is to maximize total social welfare as defined by equation (1.7). Differentiation of equation (3.1) with respect to t_i, the use of equations (3.3a) and (3.3b) to substitute for dm_j / dt_i, and summation over the n firms produces:

$$(3.5) \qquad \frac{d\Pi}{dt_i} \;=\; \sum_{j=1}^{n}[(1-t_j)(2m_jP_M+m_j^2P_{MM})]\frac{dM}{dt_i} \;-m_i\left[\frac{(1-t_i)P+C_i}{(1-t_i)}\right] \;.$$

The expression for the change in consumer’s surplus is given by equation (2.8) which in contrast to the case of the specific tax now depends on the firm under consideration because the expression for dM/dt_i does depend on i via equation (3.4). Total tax receipts are the sum of the receipts from each of the n firms:

$$(3.6) \qquad R \;=\; \sum_{j=1}^{n}R_j \;=\; \sum_{j=1}^{n}t_jPm_j \;.$$

Differentiation of equation (3.6) with respect to t_i, and the use of equations (3.3a) and (3.3b) to substitute for dm_j / dt_i produces:

$$(3.7) \qquad \frac{dR}{dt_i} = m_i P + \frac{t_i P C_i}{(1-t_i)^2 P_M} + \sum_{j=1}^{n} [\ t_j m_j P_M\ ]\ \frac{dM}{dt_i} - \sum_{j=1}^{n} \left[\ t_j P \frac{(P_M + m_j P_{MM})}{P_M}\ \right] \frac{dM}{dt_i}\ .$$

Equations (3.5), (2.8) and (3.7) provide the components to calculate the change in social welfare from a change in the ad valorem tax on the ith firm. This change set to zero for all i=1... n firms defines the optimal structure of taxation:

$$(3.8) \qquad \frac{dV}{dt_i} = (\gamma-1)Pm_i + \left[\frac{C_i}{(1-t_i)}\right] \left[\frac{\gamma t_i P - (1-t_i)P_M m_i}{(1-t_i)P_M}\right] + \Omega \left[\frac{C_i}{(1-t_i)^2}\right] = 0\ , \qquad i = 1 ... n\ ,$$

in which Ω impounds terms from equations (3.5), (2.8) and (3.7) that do not depend on i.

Given this optimal tax structure, the question is again whether low- or high-cost firms are taxed more. As in section 2, this question is answered by treating as constant all terms that do not depend on the firm under consideration, for instance the terms Ω and M. Equation (3.8) and the first-order condition provide two equations in two unknowns, t_i and m_i. Differentiation establishes that the slope of the locus between t_i and m_i given by equation (3.8) is:

$$(3.9) \qquad \frac{dm_i}{dt_i} = -\left[\frac{(2s_i\eta + s_i^2 + \gamma\eta^2)}{[(\gamma-2)\eta^2 - s_i^2 - 2s_i\eta]}\right] \left[\frac{(P_M m_i + P)}{P_M(1-t_i)}\right]$$

in which $s_i \equiv m_i/M$ and $\eta \equiv P / (P_M M) < 0$ is the elasticity of demand.

The expression (3.9) may be either positive or negative but when $\gamma = 1$, it is negative and the part of the right -hand side that is in the left brackets equals -1 regardless of the value of i. Furthermore, the relation between t_i and m_i implicit in equation (3.2) always has a negative slope, and is equal to the part of the right-hand side of equation (3.9) that is in the right brackets. Therefore, at $\gamma = 1$ the slopes of the two loci between t_i and m_i are equal. Now the expression in the part of the right -hand side of equation (3.9) that is in the left brackets increases with an increase in γ and is strictly positive at $\gamma = 2$, again regardless of the value of i. Thus for values of γ near one, the configuration of these two loci looks like Figure 1, but near $\gamma = 2$ the configuration looks like Figure 2. As in section 2, increases in C_i shift only the locus given by equation (3.2), and shift it down. Thus near $\gamma = 1$, less efficient firms face a relatively high tax rate while near $\gamma = 2$ the reverse occurs but not to such an extent that the output of a less efficient firm exceeds that of a more efficient one. For intermediate values of γ it seems that the tax rate may not even be monotonic in the efficiency of firms, because the signs of the numerator and denominator in the left-hand pair of brackets in equation (3.8) depend on the value of i. These findings seem to be the limit of properties on the tax structure that can be derived analytically, and further progress depends on simulation.

Table 3 presents the simulation results. All tax rates are positive, even for the lowest values of γ and the output and profits of all firms fall as γ rises, in contrast to the results for the specific tax for the lowest values of γ. As in Table 2 for the case of the specific tax, these ad valorem tax rates are initially higher the higher is the firm's cost, but <u>only</u> for the lowest social valuation of tax receipts, $\gamma = 1.15$. For the next two levels of γ the tax rates are approximately the same across firms. With even higher levels of γ the pattern of taxes reverses; the tax rate is lower the higher is the firm's costs. This reversal occurs as predicted theoretically at a value of γ

< 2 and, in fact, at much lower levels than 2, the value of γ at which the reversal occurs for the specific tax. Values of social welfare are uniformly higher under the ad valorem tax than under the specific tax, consistent with the results on taxation of symmetric oligopolists (Delipalla and Keen, 1992). In fact, in comparison to the specific tax, the ad valorem tax realizes a much higher proportion of the gain in social welfare relative to the Ideal Outcome (compare the values of Φ in Tables 2 and 3). The last column of Table 3 provides the value of social welfare when all the specific taxes are constrained to be the same. For the given parameters, social welfare is not very different when this constraint is imposed.

4. Asymmetric Oligopolists and the Hybrid Profits Tax.

When an industry has a few disproportionately large firms, it is natural to expect pure profits. It is then very tempting to try to tax these pure profits because a tax that only affected pure profits would not cause any distortions and because tax receipts are valued more than funds in private hands. In practice, however, there is no tax on pure profits and therefore no way that is free of distortions to get at this tax base. The most usual obstacle to a pure profits tax is the infeasibility of designing a profits tax that has full deductibility of the cost of capital. Thus the tax examined in this section is a hybrid tax that falls both on pure profits and on the use of capital as an input.

The two sales taxes in the preceding sections did not affect the incentives to use different factors of production. There was, therefore, nothing analytical to gain from endogenizing unit factor cost. In the case of a tax that affects the cost of capital, however, it is necessary to model input choices explicitly. Unit cost does depend on the tax rate because the tax rate affects the net of tax factor prices that the firm faces.

As in the preceding sections, there are n firms. The jth firm makes after-tax profits of Π_j:

$$(4.1) \qquad \Pi_j \equiv (1-\tau_j)m_jP - m_jC_j = (1-\tau_j)m_jP - m_j\alpha_j[(1-\tau_j)wl_j + (1-\tau_j\delta)rk_j],$$

in which C_j is now the minimum cost of producing a unit of output at after-tax factor prices, $\alpha_j\, l_j$ and $\alpha_j\, k_j$ are the amounts of labor and capital that the jth firm uses to produce a unit of output at minimum cost, α_j is the firm-specific efficiency factor, w and r are the wage rate and rental rate for capital before any tax considerations, and τ_j is the tax on the value of profits as defined by the tax code. Wage costs are fully deductible but only a fraction of capital costs, $\delta < 1$, is deductible. The jth firm maximizes profits with respect to the amount it produces yielding:

$$(4.2) \qquad \frac{\partial \Pi_j}{\partial m_j} = (1-\tau_j)(P + m_jP_M) - C_j = 0.$$

Differentiation of the jth firm's first-order condition, equation (4.2), with respect to the ith tax rate and substitution of the first-order condition yields:

$$(4.3a) \qquad i \neq j \;\Rightarrow\; P_M\frac{dm_j}{d\tau_i} + m_jP_{MM}\frac{dM}{d\tau_i} + P_M\frac{dM}{d\tau_i} = 0$$

and

$$(4.3b) \qquad i = j \;\Rightarrow\; P_M\frac{dm_i}{d\tau_i} + m_iP_{MM}\frac{dM}{d\tau_i} + P_M\frac{dM}{d\tau_i} - \frac{(1-\delta)}{(1-\tau_i)^2}\alpha_ik_ir = 0 \; .$$

Summation of equations (4.3 a) and (4.3b) over all n firms produces an expression for the change in the total market quantity in response to a change in the ith tax:

$$(4.4) \qquad \frac{dM}{d\tau_i} = \left[\frac{(1-\delta)r}{(n+1)P_M + MP_{MM}}\right] \left[\frac{\alpha_i k_i}{(1-\tau_i)^2}\right] < 0 \ .$$

In contrast to the result for a specific tax but similar to the case of the ad valorem tax, this expression is not independent of i, but depends positively on α_i, τ_i, and k_i which in turn depends negatively on τ_i.

As in sections 2 and 3, the goal in designing taxes is to maximize total social welfare as defined by equation (1.7). The total tax receipts are the sum of the receipts raised from each of the n firms:

$$(4.5) \qquad R = \sum_{j=1}^{n} R_j = \sum_{j=1}^{n} [\tau_j m_j \ (P - \alpha_j w l_j - \alpha_j \delta r k_j)] \ .$$

In contrast to the cases of the two sales taxes, a calculation of the change in social welfare with respect to each tax does not seem to lead to any tractable algebraic results on whether more or less efficient firms should face a higher tax rate. To investigate this question for the profits tax, I therefore present some numerical results for the case of three firms.

Table 4 provides the three optimal tax rates imposed on three firms for different values of γ ranging from 1.15 to 2.5. The fraction of the capital cost that is deductible is δ = 0.5. Other details of the calculations are given in section 1.

Table 4 suggests several conclusions about the hybrid profits tax rates: First, despite the cost of the distortion caused by the taxation of capital, the profits tax rate is much higher than conventional tax rates on profits, in practice often set at perhaps 30 percent. Second, the profits tax rate is significantly lower than the one hundred percent rate that would be applied to pure profits if they could be isolated from the cost of capital ($\delta = 1$). By way of comparison, if there were perfect deductibility of the cost of capital, $\delta = 1$, and $\gamma = 1.3$, then V = 66.71 and $\Phi = 0.583$ versus V = 62.91 and $\Phi = 0.272$ recorded in Table 4 for the case of $\delta = 0.5$ and $\gamma = 1.3$ so the imperfect deductibility of capital does make a big difference even for moderate values of the social valuation of tax receipts. Third, the tax rate is approximately constant for a particular value of γ, regardless of the size of the firm except when the tax rate has become so high that the least efficient firm has exited the industry (cases $\gamma \geq 1.9$). Conformably, social welfare and the implied value of Φ are not very different under the constraint that the tax rates be constant across firms (last column of Table 4), except for values of γ implying that the least efficient firm exits the market under an optimally differentiated tax structure . The approximate constancy of optimally differentiated rates contrasts with the results for the two sales taxes. It seems to be a fairly general result for the case of linear demand and Cobb-Douglas production. The last two rows in Table 4 keep all parameters the same as in the base case except for ϵ and shows that for the extreme values of $\epsilon = 0.05$ and $\epsilon = 0.5$ the optimal tax rate tends to be roughly constant as well.

Social welfare is always higher with the profits tax than with either of the sales taxes. This result is quite striking and contrasts with the usual prejudice against a tax on business income and in favor of a tax on sales. It is especially striking because the simulated economy is halfway between a monopolistic and a competitive one (section 1). Only for the highest values of γ do the values of Φ for the ad valorem tax and profits tax lie near each other.

Finally, the outputs of all firms fall as γ rises until the least efficient firm exits the industry. At this point the outputs of the other firms jump up. Nonetheless, the profits (net of taxes) of all the firms fall as γ rises.

5. Asymmetric Oligopolists and the Hybrid Profits and Ad Valorem Taxes

The standard tax package in most countries is an ad valorem sales tax combined with a hybrid profits tax. How does this package compare in terms of social welfare either to each of the three taxes alone discussed in sections 2-4 or to the Ideal Outcomes in which the single most efficient firm is regulated so that it prices at constant marginal cost plus a specific tax (Table 1)? On the one hand, this two-tax package must dominate any of the three taxes alone. After all, the specific tax is in any case the worst and the package contains the best and second-best performing taxes, but how much better is the package? On the other hand, the imperfect deductibility of the cost of capital under the hybrid tax is potentially a material restriction relative to the Ideal Outcome. How serious is this restriction for the actual parameterization of the simulations? Finally, how does the pattern of taxes vary with the efficiency of a firm?

The formulae for total after-tax profits and total tax receipts used in the simulations are:

$$\Pi \equiv \sum_{j=1}^{n} \Pi_j = \sum_{j=1}^{n} m_j \{(1-\tau_j)(1-t_j)P - \alpha_j[(1-\tau_j)wl_j + (1-\tau_j\delta)rk_j]\} \tag{5.1}$$

and

$$R = \sum_{j=1}^{n} R_j = \sum_{j=1}^{n} \{\tau_j m_j \; [(1-t_j)P - \alpha_j wl_j - \alpha_j \delta rk_j)] \; + t_j m_j P\} \; . \tag{5.2}$$

in which t_i is the ad valorem tax rate and τ_i is the profits tax rate and all other variables are as previously defined. The simulation was done over the six taxes (of two kinds for three firms) first using a grid of taxes running from 0 to 0.99 in increments of 0.11 and then searching in the parts of the grid adjacent to and within the grid unit that produced the highest results in increments of 0.01. Table 5 reports the values of the variables corresponding to the highest value of social welfare found.

Table 5 gives the results of the simulations. First, the tax package does better than the best single tax, the hybrid profits tax of Table 4, as it must but not by that much, increasing Φ by a little over 10% at the highest values of γ. Conformably, the tax package still falls materially short of the Ideal Outcomes reported in Table 1, very much so for low values of γ but even by almost 30% for the highest value of γ . The profits tax rates fall as firm efficiency falls, but the ad valorem tax rates rise as firm efficiency falls and uniformly so regardless of γ in contrast to theoretical and simulation results for the ad valorem tax alone (section 3). The sum of the profits and ad valorem taxes paid (T_i, i = 1,2,3) per unit output falls as firm efficiency falls. Finally, the outputs of the two less efficient firms fall as γ rises but the output of the most efficient firm rises up to $\gamma = 1.6$ the point at which the ad valorem tax that it pays becomes positive for the first time; after-tax profits of all firms fall as γ rises.

6. Conclusions.

Tax laws and administrations treat different sized firms differently but there is next to no analysis of the consequences of these differences. This paper has used one assumption about the size distribution of firms, asymmetric efficiencies among oligopolists, to investigate the pattern of taxation by firm size. Under the assumptions of the model, the government can impose a different tax on each firm, but such flexibility is only meant to establish a reference point for much more restrictive schemes that are actually put into practice.

There is no one lesson from the theory and simulations of this paper. Even for one tax imposed all by itself, such as the specific sales tax in section 2, the pattern of taxes may rise or fall with firm size depending on the social valuation of tax receipts. When the two most prevalent taxes are combined, an ad valorem sales tax and a profits tax with imperfect deductibility of capital cost, the former tax falls with the size of firm while the latter tax rises regardless of the social valuation of tax receipts for the particular parameter values in the simulations.

There is really neither theoretical nor simulation evidence of a hump-shaped pattern to the optimal taxes such as that observed for Cameroon by Gersovitz and Gauthier (1997) where intermediate-sized firms paid the most taxes. But this paper is exploring only one determinant of the size distribution of firms and the implications for optimal taxation. At the least, a practical proposal for tax policy would have to be based on a combination of these considerations with the ones raised by Keen and Mintz (forthcoming) who stress that the costs of collecting taxes from small firms may justify their exemption. Presumably, the Keen-Mintz considerations combined with any of the cases in this paper exhibiting taxes that decline with firm size would result in a hump-shaped burden of taxes.

References

Bergstrom, T. C. and H. R. Varian, "Two Remarks on Cournot Equilibria," Economic Letters, vol. 19 (1985), pp. 5-8.

Cabral, L. M. B. and J. Mata, "On the Evolution of the Firm Size Distribution: Facts and Theory," American Economic Review vol. 92 (2003), pp. 1075-1090.

Delipalla, S. and M. Keen, "The Comparison Between Ad Valorem and Specific Taxation Under Imperfect Competition," Journal of Public Economics, vol. 49 (1992), pp. 351-367.

Desai, M. A. and J. R. Hines, "Market Reactions to Export Subsidies," NBER Working Paper no. 10233, January, 2004.

Devarajan, S. and K. E. Thierfelder, "The Marginal Cost of Public Funds in Developing Countries," in A. Fossati and W. Wiegard (eds.), Policy Evaluations with Computable General Equilibrium Models (New York: Routledge, 2001).

Diewert, W. E., D. A. Lawrence, and F. Thompson, "The Marginal Costs of Taxation and Regulation," in F. Thompson and M. T. Green (eds.), Handbook of Public Finance (New York: Marcel Dekker, 1998), pp. 135-171.

Gauthier, B. and M. Gersovitz, "Revenue Erosion through Exemption and Evasion in Cameroon, 1993," Journal of Public Economics vol. 64, no. 3 (1997), pp. 407-424.

Kanbur, S. M., "Of Risk Taking and the Personal Distribution of Income," Journal of Political Economy, vol. 87, no. 4 (1979), pp. 769-797.

Keen, M. and J. Mintz, "The Optimal Threshold for a Value-Added Tax," Journal of Public Economics, (forthcoming).

Lahiri, S. and Y. Ono, "Helping Minor Firms Reduces Welfare," Economic Journal, vol. 98 (1988), pp. 1199-1202.

Lucas, R. E., "On the Size Distribution of Business Firms," Bell Journal of Economics, vol. 9 (1978), pp. 508-523.

Myles, G. D., "Imperfect Competition and the Optimal Combination of *ad valorem* and Specific Taxation," International Tax and Public Finance, vol. 3 (1996), pp. 29-44.

Salant, S. W. and G. Shaffer, "Unequal Treatment of Identical Agents in Cournot Equilibrium," American Economic Review, vol. 89 (1999), pp. 585-604.

Table 1			
Most Efficient Firm Pricing at Marginal Cost Plus a Specific Tax to Yield the Ideal Outcome			
Social Value of Taxes	Specific Tax	Total Output	Social Welfare
γ	t	M	V
1.15	0.96	14.57	69.16
1.30	1.56	13.37	71.81
1.45	1.95	12.59	75.23
1.60	2.25	11.99	79.11
1.75	2.46	11.57	83.28
1.90	2.64	11.21	87.65
2.05	2.79	10.91	92.16
2.20	2.91	10.67	96.78
2.35	3.00	10.49	101.47
2.50	3.09	10.31	106.23

Table 2

Use of a Specific Tax Alone

Social Value of Taxes	Specific Tax by Firm			Quantity Produced by Firm			After- Tax Profits by Firm			Social Welfare	Fraction of Max ΔSW	Memo: V When $t_1 = t_2 = t_3$
γ	t_1	t_2	t_3	m_1	m_2	m_3	Π_1	Π_2	Π_3	V	Φ	
1.15	0.00	0.00	0.00	4.56	3.86	3.51	10.40	7.45	6.16	59.58	0.000	59.58
1.30	0.00	0.36	0.57	5.03	3.60	2.83	12.63	6.50	4.01	59.79	0.017	59.67
1.45	0.78	0.99	1.11	4.44	3.32	2.73	9.86	5.51	3.72	60.78	0.077	60.73
1.60	1.32	1.44	1.50	4.05	3.11	2.64	8.21	4.83	3.48	62.58	0.154	62.56
1.75	1.71	1.77	1.80	3.78	2.96	2.55	7.15	4.38	3.25	64.86	0.223	64.85
1.90	2.01	2.04	2.04	3.59	2.82	2.47	6.43	3.99	3.06	67.44	0.280	67.44
2.05	2.22	2.22	2.22	3.45	2.75	2.40	5.96	3.78	2.88	70.23	0.327	70.23
2.20	2.43	2.40	2.37	3.30	2.66	2.37	5.45	3.54	2.81	73.17	0.365	73.17
2.35	2.58	2.52	2.49	3.20	2.61	2.32	5.11	3.42	2.70	76.22	0.397	76.22
2.50	2.70	2.64	2.61	3.14	2.55	2.26	4.92	3.26	2.56	79.36	0.424	79.35

Table 3												
Use of an Ad Valorem Tax Alone												
Social Value of Taxes	Ad Valorem Rate by Firm			Quantity Produced by Firm			After-Tax Profits by Firm			Social Welfare	Fraction of Max ΔSW	Memo: V When $t_1 = t_2 = t_3$
γ	t_1	t_2	t_3	m_1	m_2	m_3	Π_1	Π_2	Π_3	V	Φ	
1.15	0.03	0.08	0.13	4.74	3.78	3.12	10.91	6.58	4.22	59.67	0.009	59.64
1.30	0.28	0.29	0.30	4.46	3.40	2.81	7.15	4.10	2.77	61.15	0.128	61.14
1.45	0.38	0.38	0.38	4.29	3.16	2.59	5.71	3.10	2.09	63.68	0.262	63.68
1.60	0.44	0.43	0.42	4.11	2.99	2.51	4.74	2.55	1.83	66.74	0.367	66.73
1.75	0.48	0.46	0.46	4.00	2.95	2.30	4.16	2.35	1.43	70.11	0.444	70.08
1.90	0.51	0.49	0.48	3.89	2.79	2.28	3.70	1.99	1.35	73.68	0.502	73.64
2.05	0.52	0.50	0.49	3.86	2.75	2.22	3.57	1.89	1.26	77.39	0.547	77.32
2.20	0.54	0.52	0.51	3.80	2.66	2.12	3.32	1.69	1.10	81.20	0.581	81.10
2.35	0.55	0.53	0.52	3.77	2.61	2.06	3.19	1.60	1.02	85.08	0.609	84.96
2.50	0.56	0.53	0.52	3.63	2.65	2.11	2.91	1.65	1.06	89.03	0.631	88.87

Table 4												
Use of a Hybrid Profits Tax Alone												
Social Value of Taxes	Profits Tax Rate by Firm			Quantity Produced by Firm			After-Tax Profits by Firm			Social Welfare	Fraction of Max ΔSW	Memo: V When $\tau_1 = \tau_2 = \tau_3$
γ	τ_1	τ_2	τ_3	m_1	m_2	m_3	Π_1	Π_2	Π_3	V	Φ	
1.15	0.49	0.47	0.48	4.51	3.76	3.36	5.18	3.75	2.93	60.63	0.110	60.63
1.30	0.63	0.61	0.62	4.47	3.70	3.27	3.70	2.67	2.03	62.91	0.272	62.91
1.45	0.69	0.68	0.69	4.46	3.65	3.19	3.09	2.13	1.58	65.58	0.383	65.58
1.60	0.73	0.73	0.74	4.47	3.60	3.12	2.69	1.75	1.27	68.45	0.454	68.45
1.75	0.76	0.75	0.77	4.45	3.60	3.06	2.37	1.62	1.08	71.45	0.501	71.45
1.90	0.83	0.86	0.98	5.47	4.27	0.00	2.55	1.28	0.00	75.09	0.553	74.54
2.05	0.83	0.86	0.98	5.47	4.27	0.00	2.55	1.28	0.00	78.84	0.591	77.68
2.20	0.83	0.86	0.98	5.47	4.27	0.00	2.55	1.28	0.00	82.60	0.619	80.88
2.35	0.83	0.86	0.98	5.47	4.27	0.00	2.55	1.28	0.00	86.35	0.639	84.11
2.50	0.83	0.86	0.98	5.47	4.27	0.00	2.55	1.28	0.00	90.10	0.654	87.37
Memo: When ϵ = 0.05												
1.30	0.90	0.88	0.88	4.65	4.15	3.89	1.08	1.03	0.91	75.44		
Memo: When ϵ = 0.5												
1.30	0.48	0.48	0.51	4.43	3.46	2.85	5.10	3.11	1.99	57.86		

Table 5

Simultaneous Use of a Hybrid Profits Tax and an Ad Valorem Tax

Social Value of Taxes	Profit Tax Rate by Firm			Ad Valorem Tax Rate by Firm			Total Tax Payments per Unit of Output			Quantity Produced by Firm			After-Tax Profits by Firm			Social Welfare	Fraction of Max ΔSW
γ	τ_1	τ_2	τ_3	t_1	t_2	t_3	T_1/m_1	T_2/m_2	T_3/m_3	m_1	m_2	m_3	Π_1	Π_2	Π_3	V	Φ
1.15	0.49	0.47	0.47	0.00	0.00	0.01	1.26	1.06	1.01	4.51	3.77	3.33	5.20	3.76	2.92	60.63	0.110
1.30	0.64	0.57	0.49	0.00	0.10	0.20	1.78	1.61	1.61	4.82	3.62	2.64	4.17	2.53	1.42	63.04	0.283
1.45	0.72	0.59	0.49	0.00	0.20	0.30	2.13	1.98	1.99	5.07	3.36	2.19	3.60	1.85	0.86	66.08	0.415
1.60	0.74	0.60	0.50	0.06	0.27	0.35	2.39	2.26	2.24	5.11	3.13	2.00	3.19	1.43	0.65	69.50	0.508
1.75	0.74	0.61	0.47	0.14	0.31	0.40	2.63	2.48	2.47	5.03	3.07	1.79	2.83	1.27	0.51	73.17	0.573
1.90	0.75	0.61	0.48	0.18	0.35	0.42	2.80	2.66	2.62	5.00	2.89	1.73	2.56	1.06	0.45	77.03	0.622
2.05	0.75	0.62	0.48	0.22	0.37	0.44	2.94	2.79	2.75	4.94	2.83	1.64	2.38	0.96	0.39	81.01	0.658
2.20	0.75	0.60	0.50	0.25	0.40	0.45	3.06	2.90	2.86	4.90	2.71	1.60	2.25	0.88	0.35	85.09	0.686
2.35	0.75	0.61	0.50	0.28	0.41	0.46	3.16	2.99	2.94	4.80	2.68	1.58	2.08	0.83	0.34	89.23	0.708
2.50	0.75	0.62	0.50	0.30	0.42	0.47	3.24	3.07	3.01	4.76	2.64	1.54	1.98	0.77	0.31	93.43	0.726