New Measures of the Excess Burden of Capital Taxation in Canada

By W. Erwin Diewert[1] (University of British Columbia, Vancouver, Canada)
and Denis A. Lawrence (Tasman Asia Pacific, Canberra, Australia)

Paper presented at the CERF and IRPP Conference held in Ottawa, May 4-6, 2000: *Creating Canada's Advantage in an Information Age*.

1. Introduction

There is a general feeling in some segments of the business press in Canada that taxes are too high in Canada and that this is reducing our competiveness. Our contribution at this conference will be an attempt to determine empirically what the efficiency costs of taxing capital in Canada are.

It is very difficult to explain to the layman what exactly are the efficiency costs of a tax. In section 2 below, we attempt to explain why taxing the return to capital can be expected to reduce the real output of an economy in the context of a very simple production function model. This explanation will still probably not be very convincing for a layman but it should be helpful for the economists present at the conference.

In section 3, we discuss another problem that is not that straightforward: namely, how should we allocate the cost of a durable input across the useful life of the input? This leads us into a discussion of the *user cost of capital*.

Sections 4 to 6 are more technical. These sections gradually relax some of the restrictive assumptions made in section 2. In particular, we need to generalize the model explained in section 2 to cover the case of many (noncapital) inputs and outputs and many capital inputs. We also need to extend the model to an open economy.

Section 7 introduces our preliminary econometric model which is based on relatively recent developments in the theory of flexible functional forms. However, in section 8, we discuss a technical problem with the functional form that is suggested in section 7: namely, it will tend to generate somewhat artificially trending elasticities in many data sets. Given the importance of getting accurate elasticity estimates for the computation of excess burdens, we address this

[1] Our thanks to the Canadian Donner Foundation for partially funding this research over the years.

problem in section 8. We suggest a new functional form that is completely flexible at two data points instead of the usual single data point.

Section 9 presents our empirical estimates for a Canadian model estimated using data for the period 1974-1998. Our model has 8 commodities: a domestic output, an export commodity, an import commodity, an aggregate market sector labour input and three reproducible capital stocks (nonresidential structures, machinery and equipment and inventories). The last input is land.[2]

Section 10 uses the elasticity estimates presented in section 9 and the theory of excess burden measurement developed in section 6 to present empirical estimates of the marginal excess burdens of capital taxation in Canada for the years 1974-1998. However, we regard our estimates as being only very preliminary: there is much more work to be done both on developing better estimates of the allocation of taxes in Canada and in estimating more disaggregated econometric models.

Section 11 concludes.

2. The Excess Burden of Capital Taxation

We illustrate the efficiency costs of taxing capital by considering a very simple model of a closed economy (i.e., we neglect the effects of international trade in goods). We suppose that units of private sector reproducible capital are combined with factors that are held fixed during the short run to produce units of aggregate output that can be used for either consumption C or investment I. Letting L denote the number of units of labour and other factors that are fixed in the short run we have:

$$Y = C + I = f(K, L) \quad (1)$$

where f is the production function, Y denotes output and K denotes the beginning of the period capital stock. Note that we are assuming that units of the investment good I are perfectly substitutable with units of the consumption good C. We also assume that investment goods produced during the current period are added to the reproducible capital stock at the beginning of the following period. Thus, investment goods can be viewed as intertemporal intermediate inputs into the private production sector: I is produced this period so that it can be used as capital input next period and offset this period’s depreciation of the capital stock.

[2] Our data set is basically an update of the data described in Diewert and Lawrence (1999). Of course, we have made heavy use of Statistics Canada data.

We assume that each unit of the capital stock has a physical decline in its efficiency over the period at the rate δ ; ie if K units of the capital stock are in place at the beginning of the period, only $(1-\delta)K$ units are available for further use at the end of the current period.

We consider a steady state capital optimisation problem where investment is set equal to depreciation; ie we replace I in (1) by δK and maximise $C = f(K,L) - I = f(K,L) - \delta K$ with respect to K. Another way of viewing depreciation in this formulation is to regard it as a cost of production; the capital used at the beginning of the period, K, should be assessed a charge equal to the decline in value of the capital stock due to deterioration and a shorter life. Another charge that should be assessed against the starting capital stock is the opportunity cost of capital; ie the interest cost which will be just sufficient to induce owners of the capital stock to hold the capital stock through the period. Thus, if the real interest rate is r^*, then the optimal long run capital stock $K°$ is the solution to the following maximisation problem:

(2) $\max_K \{ f(K,L) - (r^*+\delta)K \}$.

Since we are regarding L as fixed, write the production function $f(K,L)$ as $f(K)$. Then the first order necessary condition for $K°$ to solve (2) is:

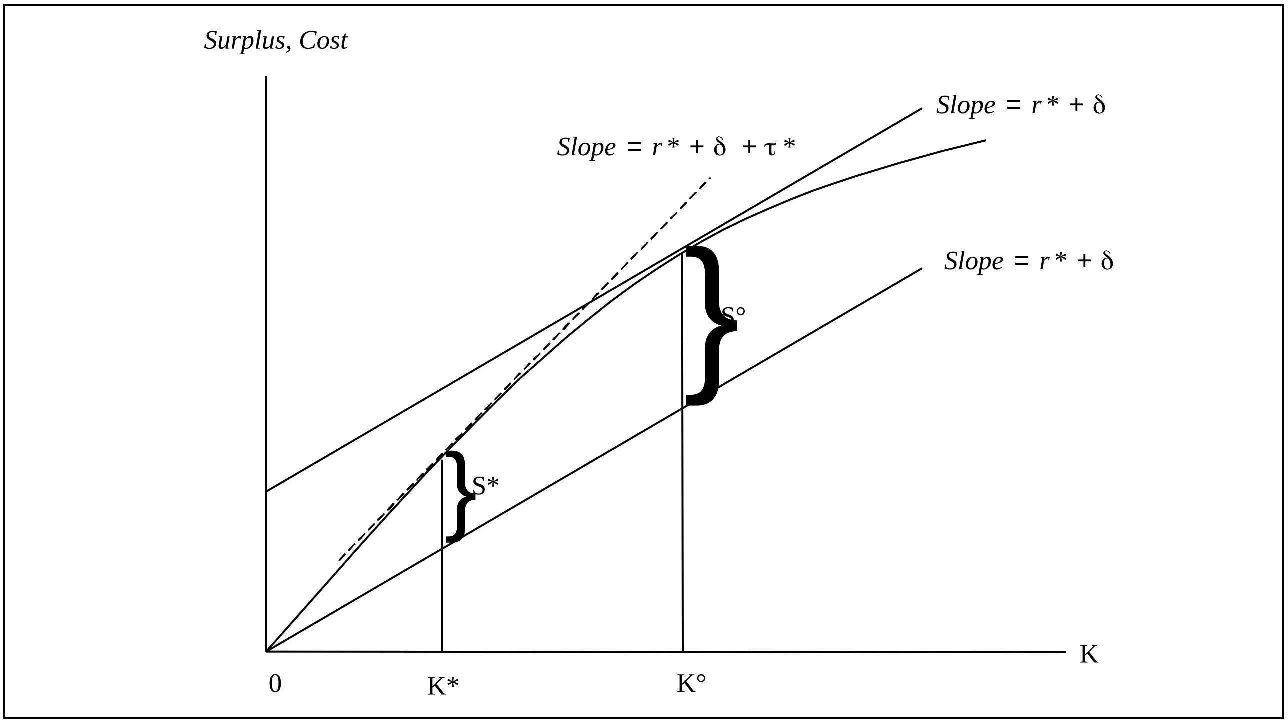
(3) $f'(K^\circ) = r^*+\delta$

where f' denotes the first derivative of f. We assume that the following second order sufficient condition is also satisfied:

(4) $f''(K^\circ) < 0$.

The geometry of the unconstrained maximisation problem (2) is illustrated in Figure 1 below. The curved line through the origin represents the production function constraint, $C+I = f(K)$, while the straight line through the origin represents the depreciation and interest cost of capital. The difference between the two lines represents sustainable consumption (after interest payments) or surplus as a function of the beginning of the period capital stock K. The maximum sustainable surplus $S°$ is achieved at the capital stock $K°$ where the slope of the production function equals the slope of the capital cost function.

Figure 1: Stylised Loss from Capital Taxation



When capital is taxed, private producers will face the price $r^* + \delta + \tau^*$ per unit of capital used, where τ^* is the capital tax rate expressed as a fraction of the asset value of capital.[3] Thus, instead of solving (2) in the long run, private producers will be induced to choose the capital stock K^* which solves:

(5) $$\max_K \{ f(K) - (r^* + \delta + \tau^*)K \}.$$

We may regard the K^* which solves (5) as a function of the tax rate τ^*; ie $K^* = K(\tau^*)$. This solution to (5) will satisfy the following first order necessary condition:

(6) $$f'[K(\tau^*)] = r^* + \delta + \tau^* .$$

The fact that producers must pay capital taxes to the government increases the cost of using reproducible capital as an input and the resulting steady state capital stock $K(\tau^*)$ is smaller than the optimal capital stock, $K^o = K(0)$, which solved (2). The tax distorted surplus, $S^* \equiv f(K^*) - (r^* + \delta)K^*$ is smaller than the optimal surplus $S^\circ \quad f(K^\circ) - (r^* + \delta)K^\circ$; (see Figure 1).

[3] We use r* to distinguish the real interest rate from the nominal interest rate r and we use τ* to distinguish the capital asset tax rate from the capital income tax rate τ. In the following section, we will consider more precisely how to define τ*.

Figure 1 illustrates *qualitatively* the effects of taxing reproducible capital — the higher the level of taxation, the lower will be the long run level of capital utilised and the corresponding surplus. In what follows, we indicate how a *quantitative* estimate of the decline in the sustainable surplus can be obtained.

First, differentiate equation (6) with respect to $\tau *$. We obtain the following equation for the change in capital due to a small increase in the tax rate, $K'(\tau *)$:

(7) $K'(\tau *) = 1/\, f''[K(\tau *)]$,

where f'' is the second derivative of the production function and will be negative under the usual assumptions on the production function. Now define producer surplus (or sustainable consumption after interest payments) as a function of the capital tax rate $\tau *$ as follows:

(8) $S(\tau *) \quad f[K(\tau *)] - (r * + \delta)\, K(\tau *)$.

Differentiating (8) with respect to $\tau *$ and using (6) yields the following formula for the rate of change of surplus with respect to the level of capital taxation:

(9) $S'(\tau *) = [\, f'[K(\tau *)] - (r * + \delta)]\, K'(\tau *) = \tau * K'(\tau *)$.

Evaluating (9) at $\tau * = 0$ yields

(10) $S'(\tau *) = 0$.

Differentiating (9) with respect to $\tau *$ and evaluating the resulting derivative at $\tau * = 0$ yields

(11) $S''(0) = K'(0) = 1/\, f''[K(0)] < 0$

where the second equality in (11) follows from (7) evaluated at $\tau * = 0$ and the inequality follows from $K^{\circ} = K(0)$ and (4). We now use (10) and (11) to form the following second order Taylor series approximation to $S(\tau *)$:

$$S(\tau *) \quad S(0) + S'(0)\tau * + (1/2)\, S''(0)\tau *^{2}$$
$$= S(0) + (1/2)\tau *^{2} \;/\; f''[K(0)]. \tag{12}$$

Define $L(\tau *)$ as the loss of sustainable output as a fraction of optimal output $Y(0)$:

(13) $L(\tau *) \quad [S(0) - S(\tau *)] \;/\; Y(0)$.

Using (12), a second order approximation to $L(\tau *)$ is

(14) $A(\tau *) \quad -(1/2)\tau *^{2} \;/\; Y(0)\; f''[K(0)]$.

We need to provide an economic interpretation for the second derivative of the production function, $f''[K(0)]$, evaluated at the optimal capital stock $K(0)$. The first derivative of the production function, $f'[K(0)]$, is the optimal return to one unit of the capital stock or the rental price of capital, $P_K = r^* + \delta$. Thus, the second derivative can be interpreted as the change in the rental price due to a small change in the use of capital, $dp_K(K)/dK$:

(15) $dp_K[K(0)]/dK = f''[K(0)] = f''[K^\circ] < 0$

where the inequality follows from (4). We convert dp_K/dK into a non-negative (inverse) elasticity of demand for capital ε by changing the sign of $f''[K(0)]$ and multiplying by $K(0)/p_K[K(0)] = K(0)/f'[K(0)]$:

(16) $\varepsilon \quad -f''[K(0)]K(0)/f'[K(0)] = -f''[K(0)]K(0)/(r+\delta)$

where the second equality in (16) follows from (3).

It is also useful to define the economy's optimal capital output ratio γ as the ratio of the optimal capital stock *K*(*0*) to the optimal gross output $Y(0) = f[K(0)]$:

(17) $\gamma \quad K(0)/Y(0)$.

Substitution of (16) and (17) into (14) yields the following formula for the approximate loss of producer surplus as a fraction of optimal output:

(18) $A(\tau^*) = (1/2)\tau^{*2}\gamma / \varepsilon(r^*+\delta) > 0$.

For advanced industrial economies, a typical range for the capital tax rate τ^* is 0.01 to 0.03 (ie one percent to three percent of the asset value of capital), for the capital output ratio γ is two to four, for the inverse elasticity of demand for capital ε is 0.5 to 1.0, for the real after tax rate of return *r** is 0.01 to 0.05 and for the depreciation rate δ is 0.04 to 0.08. If we substitute the midrange values for these parameters into the right hand side of (18), we find that the approximate output loss due to capital taxation at the rate $\tau^* = 0.02$ is *A*(0.02) = 0.0089 or 0.89 percent of gross domestic product. Table 1 below indicates how the approximate loss of output $A(\tau^*)$ varies as each parameter varies between the low and high values of its assumed range while letting the remaining parameters equal their midrange values.

Table 1: Percentage Loss of Output

Parameter	*Midrange*	τ^*		r^*		γ		δ		ε	
Value	*Values*	0.01	0.03	0.01	0.05	2.00	4.00	0.04	0.08	0.50	1.00
% loss	0.89	0.22	2.00	1.14	0.73	0.59	1.19	1.14	0.73	1.33	0.67

Note that the approximate loss of output increases as the square of the capital tax rate $\tau *$ so if $\tau *$ increases from 0.02 to 0.03 and the other parameters remain at their midpoint values, the loss of output due to capital taxation increases from 0.89 percent of GDP to two percent of GDP. These output losses persist year after year so that the present value of these annual output losses is substantial.

It should be emphasised that the above efficiency losses induced by the taxation of capital are entirely avoidable: equivalent amounts of revenue could be raised by taxing the final outputs of the private production sector or by taxing primary inputs. We note that the latter two forms of taxation do not involve a loss of productive efficiency for the economy whereas taxing an intermediate input like capital invariably involves a loss of productive efficiency.[4]

The efficiency losses listed above in Table 1 are likely to underestimate substantially the actual losses that are induced by capital taxation in an industrialised economy. The above model assumes only one capital stock with an average tax rate of $\tau *$ which is applied to the asset value of reproducible capital. In actual economies, the system of business income taxation invariably taxes lightly some components of the capital stock and taxes other components very heavily. The efficiency losses associated with the differential taxation of each type of capital will grow approximately as the square of the tax rate. Thus, the large losses associated with the heavily taxed components will not be balanced by the small losses associated with the lightly taxed components and the total loss will be much larger than the loss obtained by applying an average tax rate to the total reproducible capital stock.[5]

Another diagram may be helpful in illustrating the efficiency costs of capital taxation. Note that $K(\tau *)$, the capital stock solution to equation (6), can be regarded as the long run demand for reproducible capital as a function of the tax rate $\tau *$. Now rewrite equation (6) as follows:

(19) $f'[K(r*+\delta+\tau*)] = r*+\delta+\tau*$

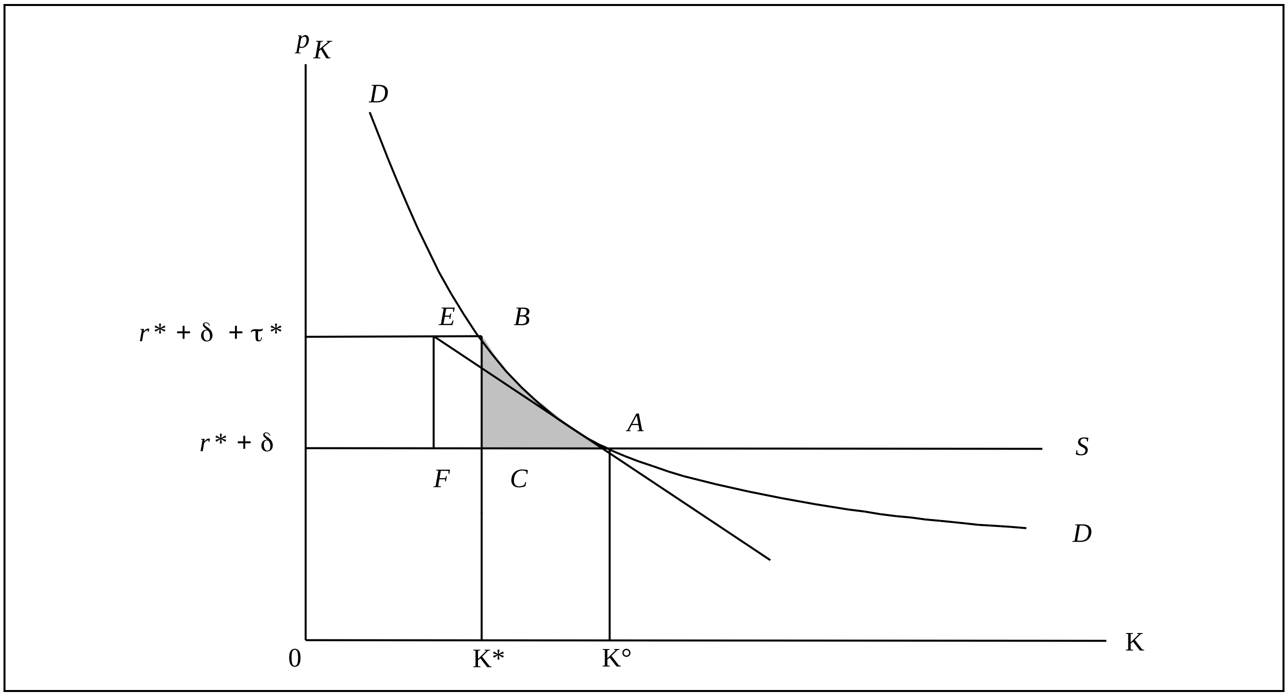
ie the demand for capital $K(r*+\delta+\tau*)$ which solves (19) can be written as a function of the tax distorted rental price of capital, $r*+\delta+\tau*$. In Figure 2 below, the inverse of this demand for capital function is graphed as the curve *DD*. If there were no capital taxes, capital would be supplied to the private production sector at the rental price $r*+\delta$ which would just cover the real interest and depreciation costs of using a unit of capital for the period. This horizontal

[4] See Diewert (1983a) (1983b) for the productive efficiency approach to tax policy. Although capital taxation cannot be justified on productivity or efficiency grounds, it still can be justified on equity grounds.

[5] Reproducible capital stocks are stocks produced by the production sector. There are no efficiency losses associated with taxing capital stock components that are fixed (such as land).

supply of capital curve intersects the demand curve at the point *A*. The imposition of the capital tax τ^* shifts the supply of capital curve up and this tax distorted supply curve intersects the demand curve at *B*. Note that the equilibrium level of capital used has decreased from $K°$ to K^*.

Figure 2: Alternative Representation of Capital Tax Loss



Equation (19) can be integrated to obtain an expression for the gross change in output; ie the change in gross output produced due to capital taxation before deducting depreciation and interest costs; ie we have

(20) $f(K^\circ) - f(K^*) = \int_{K^*}^{K^\circ} f'(K)dK = Area\ BAK^\circ K^*.$

The efficiency cost of capital taxation is defined to be the net change in output after deducting depreciation and interest payments:

$$(21)\quad S(0) - S(\tau^*) \quad [f(K^\circ) - (r^* + \delta)K^\circ] - [f(K^*) - (r^* + \delta)K^*]$$
$$= f(K^\circ) - f(K^*) - (r^* + \delta)(K^\circ - K^*)$$
$$= Area\ BAK^\circ K^* - Area\ CAK^\circ K^*$$
$$= Area\ ABC.$$

Thus, the area of the shaded triangular region under the demand curve is a measure of the efficiency costs of capital taxation. This is a producer surplus measure of deadweight loss.[6]

We now linearise the demand curve around the undistorted equilibrium point *A* and use the triangle *AEF* as an approximation to the exact deadweight loss *ABC*. From (15) and (16), it can be verified that the absolute value of the slope of the linear approximation to *DD* at *A* is $\varepsilon(r^*+\delta)/K^\circ$. The vertical distance *EF* in Figure 2 is equal to τ^* so the horizontal distance of the triangle *AEF*, *AF*, will equal the vertical distance τ^* divided by the slope $\varepsilon(r^*+\delta)/K^\circ$. Thus,

$$(22) \qquad Area\ AEF = (1/2)[\tau^* / \{\varepsilon(r^*+\delta)/K^\circ\}]\tau^*$$
$$= (1/2)\tau^{*2} K^\circ / \varepsilon(r^*+\delta)$$
$$= Y(0)\,A(\tau^*)$$

where we have used (18) to derive the last equality in (22). Thus, the area *AEF* in Figure 2 is equal to the approximate loss $A(\tau^*)$, defined earlier by (18), times optimal GDP, $Y(0)$. We note that for small tax rates τ^*, the approximate loss measure *AEF* should be quite close to the exact loss measure *ABC*.

The approximate total efficiency loss or excess burden of capital taxation defined by (22) is not the most interesting number from the viewpoint of economic policy. A more interesting concept, initiated by Browning (1976) (1987), is the *marginal excess burden of capital taxation*. We explain this concept in the context of our highly simplified model in the following section.

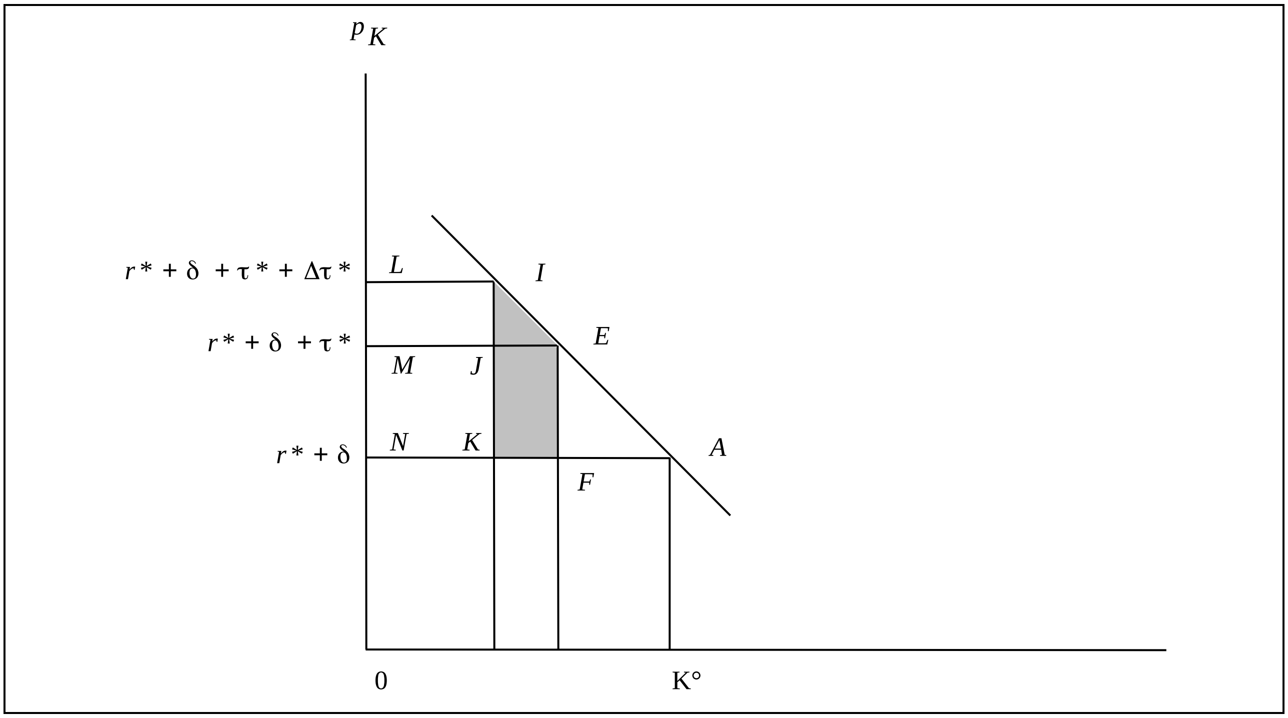
The approximate total efficiency loss or excess burden of capital taxation defined by (22) is not the most interesting number from the viewpoint of economic policy. A more interesting concept, initiated by Browning (1976) (1987), is the marginal excess burden of capital taxation. This concept compares the increase in efficiency loss due to a small increase in the level of capital taxation to the increase in tax revenue that can be attributed to the tax increase. Another diagram may be helpful in explaining the concept.

In Figure 3 below, we have reproduced the approximate deadweight loss triangle *AEF* as in Figure 2 and this triangle corresponds to the efficiency loss when the capital tax rate is τ^*. We now increase the capital taxation rate by a small amount $\Delta\tau^*$ and we note that the increase in efficiency loss is equal to the triangle *EIJ* plus the rectangle *EFKJ*. The initial tax revenue is equal to the area of the rectangle *EFNM* and the new tax revenue is equal to the area of the

[6] Note that our loss measure does not contain a consumer surplus term. For analogous consumer surplus measures of deadweight loss, see Browning (1976) (1987) and Findlay and Jones (1982).

rectangle *IKNL*. Thus, the change in tax revenue is equal to the area of *IJML* minus the area of *EFKJ*. This change in tax revenue (the incremental benefits of the tax increase) can be compared to the increased efficiency loss, *EFKI*, (the incremental costs of the tax increase). Note that if the initial level of taxation τ^* is very high, then the incremental tax revenue can be negative; ie the induced reduction in the use of capital can outweigh the increased tax revenue per unit of capital used by private producers.

Figure 3: Marginal Excess Burden of Capital Taxation



We now provide an analytic formulation that corresponds to the marginal excess burden measure described by Figure 3. Equation (18) describes the approximate efficiency loss $A(\tau^*)$ as a fraction of optimal output $Y(0)$. Differentiating this function with respect to τ^* gives us the following formula for the marginal efficiency loss (as a fraction of $Y(0)$):

(23) $A'(\tau^*) = \tau^*\gamma \ / \ \varepsilon(r^*+\delta) > 0$.

Define tax revenue as a function of the capital tax rate τ^*, $T(\tau^*)$ as follows:

(24) $T(\tau^*) \quad \tau^* K(\tau^*)$.

Note that $T(0) = 0$ and the first and second order derivatives of $T(\tau^*)$ evaluated at the no distortion point $\tau^* = 0$ are:

(25) $T'(0) = K(0)$

(26) $T''(0) = 2K'(0)$.

Thus a second order Taylor series approximation to $T(\tau^*)$ is

(27) $T(\tau^*) \quad K(0)\tau^* + K'(0)\tau^{*2}$

Define the approximate benefit function $B(\tau^*)$ as the right hand side of (27), divided by the optimal output $Y(0)$:

(28)

$$\begin{aligned} B(\tau^*) &\quad [K(0)\tau^* + K'(0)\tau^{*2}] / Y(0) \\ &= \gamma\tau^* + K'(0)\tau^{*2} / Y(0) && \text{using (17)} \\ &= \gamma\tau^* - \gamma\tau^{*2} / \varepsilon(r^*+\delta) && \text{using (7) and (16).} \end{aligned}$$

Now differentiate (28) with respect to τ^* which gives us a formula for the marginal benefit of increasing capital taxes $B'(\tau^*)$ (as a fraction of optimal output $Y(0)$):

(29) $B'(\tau^*) = \gamma[1 - 2\tau^* / \varepsilon(r^*+\delta)]$.

Finally, define the (approximate) marginal excess burden of capital taxation $MEB(\tau^*)$ as the ratio of the marginal efficiency cost $A'(\tau^*)$ defined by (23) to the marginal benefit $B'(\tau^*)$ defined by (29):

(30) $MEB(\tau^*) = A'(\tau^*) / B'(\tau^*) = \tau^* / [\varepsilon(r^*+\delta) - 2\tau^*]$.

Note that $MEB(\tau^*)$ depends not only on the rate of capital taxation τ^* but it also depends on the inverse elasticity of demand for capital ε, the real interest rate r^* and the depreciation rate δ. However, $MEB(\tau^*)$ does not depend on the capital output ratio γ, in contrast to our earlier formula for the approximate total efficiency loss (as a fraction of optimal output), $A(\tau^*)$, defined by (18).

In Table 2 below, we evaluate $MEB(\tau^*)$ defined by (30) at our midrange estimates for the capital tax rate $(\tau^* = 0.02)$, the real interest rate $(r^* = 0.04)$, the depreciation rate $(\delta = 0.06)$ and the (inverse) elasticity of demand for capital $(\varepsilon = 0.75)$. We also table $MEB(\tau^*)$ as each parameter varies between the low and high values of its assumed range, while letting the other parameter values equal their assumed midrange values.

Table 2: Marginal Excess Burdens of Capital Taxation

Parameter	Midrange	τ^*		r^*		δ		ε	
Values	*Values*	0.01	0.03	0.01	0.05	0.04	0.08	0.50	1.00
MEB	0.727	0.211	4.000	1.600	0.471	1.600	0.471	4.000	0.400

From Table 2, the marginal excess burden of capital taxation when $\tau^* = 0.02$, $r^* = 0.03$, $\delta = 0.06$ and $\varepsilon = 0.75$ is 72.7 percent. This means that if the government is contemplating financing a new recurring program expenditure by increasing capital taxation, then for each dollar of tax revenue spent on the program, its benefits should exceed 1.727 dollars; ie the loss of productive efficiency that is induced by a tax increase that yields an extra dollar of revenue is 72.7 cents. In contrast to the rather small numbers in Table 1, the numbers in Table 2 are rather large. For example, if the level of capital taxation increases from two percent to three percent, then the marginal excess burden increases from 72.7 percent to 400 percent; i.e., the marginal benefits that accrue to an incremental program that is financed by the increased level of capital taxation should exceed five dollars for each dollar of revenue raised. Of that five dollars, one dollar of benefits is required to make up for the one dollar of tax revenue that is diverted from private uses and the other four dollars of benefits are required to offset the loss of output that the increased level of capital taxation induces in the private production sector.

Table 2 indicates that the marginal excess burden of capital taxation is very sensitive to the parameter values that were inserted into formula (30). This is unfortunate, because it is difficult to determine τ^*, r^*, δ and ε with great precision in actual economies. Hence relatively small errors in these parameters can translate into relatively large errors in the associated excess burdens. However, our qualitative assessment of the numbers presented in Table 2 is that the marginal excess burdens generated by the taxation of reproducible capital are likely to be considerably larger than the marginal excess burdens generated by taxing consumption or labour.[7] Our reason for this a priori expectation is that even though τ^* is a relatively small fraction, it is a relatively large *proportion* of the undistorted rental price of capital $r^* + \delta$ and hence has a relatively large effect on the allocation of resources.

3. The User Cost of Capital

In the national accounts, interest (or more generally, the return to capital) is treated as a transfer (or as a distribution out of surplus) and not as a cost of production. Thus, the only cost associated with the use of reproducible capital in the national accounts is depreciation. The costs of using nonreproducible capital inputs like land are totally ignored in the national accounts. However, economic theory regards interest as a cost of production – it is the cost of inducing

[7] Diewert and Lawrence (1994) (1996) found that the marginal excess burdens for labour and consumption taxes were in the 10 percent to 20 percent range for the New Zealand economy.

investors to defer consumption for the period under consideration. Thus, an appropriate cost of capital from the viewpoint of production theory and the measurement of deadweight costs is the user cost of capital which includes both interest and depreciation costs. This concept dates back at least a century to the economist Walras (1954; 269) and the industrial engineer Church (1901; 907-908). In more recent times, it was generalized to deal with the complications of the business income tax by Jorgenson (1996). We review this literature below.

We begin by deriving the user cost of capital in a world without the taxation of capital. We suppose that a firm purchases a capital asset (or durable input) such as a machine, a computer, a building, an inventory item or a plot of land at the beginning of an accounting period at the price P. Since a durable asset by definition lasts longer than one period, the firm cannot simply charge the entire cost of the asset to the first accounting period: it must distribute the cost over the useful life of the asset. During the period, the asset declines in value according to the depreciation rate δ, so at the end of the accounting period, if there were no inflation during the period, the asset would be worth $(1-\delta)P$. However, normally there will be some change in the price of the asset over the period. Let the inflation rate for a unit of the asset be denoted by α, so that the end of the period price for the depreciated asset will be $(1-\delta)(1+\alpha)P$. Now we are ready to work out what the net cost of using the asset is for the first accounting period. The beginning of the period user cost of the asset, B, is defined to be the asset's purchase cost P minus the discounted end of period market value of the asset:

(31) $$B = P - (1-\delta)(1+\alpha)P/(1+R)$$

where R is the average cost of capital that the firm faces during the period; i.e., it is an average of the bond interest rate and the equity cost of capital that it faces at the beginning of the accounting period.

The above user cost of capital is the one that economists are most familiar with since they are used to working with discounted values. However, it is also possible to work with the end of the period user cost U, which simply multiplies B by $1+R$:

(32) $$U = P(1+R) - (1-\delta)(1+\alpha)P$$

(33) $$= [R - \alpha + \delta(1+\alpha)]P.$$

Formula (32) for the end of period user cost of an asset should be intuitively appealing to accountants. As part of the cost of using the asset during the period, we need to charge not only the purchase price P of the asset, but also the direct bond interest costs associated with financing the purchase of the asset plus the opportunity cost of tying up equity capital in the asset (this is the cost RP). However, these costs are partially offset by the fact that at the end of the period, we have an asset that could be sold for the amount $(1–\delta)(1+\alpha)P$. Thus, the net cost of using the asset during the period (including the opportunity costs of the equity capital that is tied up in the purchase of the asset) is the right hand side of (32). Note that the end of the period value of the asset, $(1–\delta)(1+\alpha)P$, becomes next period's beginning of the period value of the asset.

Formula (33), which is simply a rearrangement of (32), also has a nice intuitive interpretation. It says that the user cost of an asset has an interest/opportunity cost of capital equal to RP less a capital gains component αP plus a depreciation component that is indexed for asset inflation $\delta(1+\alpha)P$. The first two components of this formula can be combined into the term $(R - \alpha)P$ which can be interpreted as a real interest rate term.

The user cost of capital plays a fundamental role in any economic approach to modeling producer behavior. It plays the role of a period specific price for a durable capital input and is analogous to a wage rate (as the price for a unit of labour) or an output price (as the price for a unit of output).

We now bring business income taxes into the picture.

Unfortunately, the business income tax does not treat capital costs in the manner indicated above by formulae (32) or (33), which is an economic approach based on current opportunity costs. Most systems of business income taxation use the conventions of historical cost accounting to define period by period capital costs that are to be used in defining income for tax purposes. Thus, for tax purposes, an accounting user cost A for the asset described in the previous section might be defined by something like the following formula:

(34) $$A = [fR + d]P$$

where f is the fraction of interest and equity cost that is tax deductible (typically interest costs are deductible but equity opportunity costs are not so that f would depend on the firm's debt-equity ratio) and d is the depreciation rate for the asset that is prescribed by the tax code. The actual formula for A is typically a lot more complicated than the right hand side of (34) due to various incentives and exceptions that are invariably written into the tax code and due to the lack of indexation of depreciation allowances for inflation.

We continue to suppose the firm uses only one capital input, say the one described in the previous section. Suppose further that the firm purchases K units of this durable input at the beginning of the accounting period and has a cash flow (the value of outputs produced during the period minus the value of nondurable inputs used during the period) of CF during the period. Then the firm's profits before income taxes would be $CF - UK$ where U is the user cost described by (32) or (33) above. The firm's profits after business income taxes, π, are equal to before tax profits, $CF - UK$, minus the business tax rate, τ, times taxable income as defined by the tax code, $CF - AK$; ie we have the following definition for after business tax income:

(35) $\pi = [CF - UK] - \tau[CF - AK]$

(36) $= (1 - \tau)[CF - UK] + \tau[A - U]K$ rearranging terms

(37) $= (1 - \tau)[CF - \{U + [\tau/(1 - \tau)](U - A)\}K]$ rearranging terms

(38) $= (1 - \tau)[CF - \{U + W\}\mathrm{K}]$ using (39)

where the business tax wedge W is defined as:

(39) $W = [\tau/(1 - \tau)](U - A)$.

Note that if either the business income tax rate τ equals 0 or if the no tax economic user cost U equals the tax accounting user cost A, then the tax wedge W is 0 and the tax distorted user cost of capital $U + W$ is equal to the undistorted user cost of capital U.

The tax distorted user cost of capital $U + W$ was first derived by Jorgenson (1963) and it is widely used by economists when they model the effects of the business income tax. In order for us to use it in our study, we would need a time series of business tax rates τ and a time series of accounting user costs A for the various types of asset in our model.

However, the (business) tax distorted user cost of capital, $U + W$, is not the end of the story when we want to calculate the deadweight cost of taxes on capital, because the above material neglects the fact that capital is not only taxed at the business level, it is also taxed at the personal level. In addition, there are various specific commodity taxes that fall on various capital stock components, like property taxes and sales taxes on purchases of machinery and equipment and structures.

The return to financial capital (i.e., interest and dividends and realized capital gains) is taxed at the personal level. The user costs developed above assumed that the firm faced the average cost of capital R. However, the individual investor does not receive this entire return to capital: the dividends and interest received by the domestic investor are subject to personal taxation at the rate t say (foreign investors are typically subject to interest or dividend withholding taxes at rates which approximate the personal tax rate). Thus, the individual investor receives only the after personal tax rate of return r where r is related to R as follows:

(40) $r = (1 - t)R$.

Thus, removing the tax distortions on capital would involve eliminating both the business income tax (or setting A equal to U) as well as the personal tax on the return to capital. To find the completely undistorted user cost of capital, we need only replace R in (32) or (33) by r. Thus, define the undistorted user cost of capital U^* by:

(41) $U^* = [r - \alpha + \delta(1 + \alpha)]P.$

Using (40), it can be seen that the relationship of our initial economic user cost of capital U defined by (33) and our new undistorted user cost of capital U^* defined by (41) is:

(42) $U = U^* + [t/(1 - t)]rP$

(43) $\quad = U^* + W^*$

where the personal taxation wedge W^* is defined as:

(44) $W^* = [t/(1 - t)]rP.$

In words, taxation of capital income at the personal level causes the undistorted user cost of capital U^* to increase by the personal taxation wedge, W^*. This analysis has neglected the effects of the business income tax. However, the analysis presented in the previous section is still valid but in order to determine the distortions due to both the business and personal taxation of capital, we need to replace U in (38) by $U^* + W^*$. The resulting formula for after business tax income becomes:

(45) $\pi = (1 - \tau)[CF - \{U^* + W^* + W\}K].$

Thus, the total tax wedge created by taxation at both the business and personal levels is:

(46) $T = W^* + W$

where the personal tax wedge W^* is defined by (44) and the business tax wedge W is defined by (39).

Unfortunately, we did not have an accurate information base that would have allowed us to construct the above personal and business tax wedges defined above. However, there is little evidence that businesses actually rearrange terms as was done in (35) to (38) above to obtain the Jorgenson tax adjusted user costs $U + W$. There is, however, some evidence that firms do use the user costs defined by (32) or (33) (see Diewert and Fox 1999). Typically, the business income tax is ignored entirely in cost allocation models and simply treated as a charge against earnings. In this case, a sophisticated business might treat the amount of business tax that it pays during the accounting period as a capital charge that should be spread evenly on its assets. Following this line of thought, turn to our one asset example again. Define the business asset tax rate τ^* as:

(47) $\tau^* =$ (business income taxes paid during the period)$/PK$

and the tax adjusted user cost becomes:

(48) $U' = U + \tau^*P$

(49) $= [R - \alpha + \tau^* + \delta(1+\alpha)]P$ using (33)

(50) $= U^* + W^* + \tau^*P$ using (43)

(51) $= U^* + t^*P + \tau^*P$

where we have converted the personal tax wedge W^* into a tax wedge as a fraction of the asset value; ie t^* can be defined as:

(52) t^* = (personal capital taxes paid during the period)/PK.

The user cost formula (51) is the user cost formula that we use in this study. From the viewpoint of real life firm accounting practices, the case for using this user cost formula is just as strong as the case for using the Jorgensonian user cost formula $U^* + W^* + W$. Essentially, we have replaced the Jorgensonian tax wedge W by the aggregate business and personal tax wedge $(t^*+\tau^*)P$.

The approach we adopt in this report to calculating the capital tax rate is to assume that the tax rate is the same across the various asset categories. It is calculated by taking the ratio of actual capital tax payments to the value of assets. This produces an estimate of the average capital tax rate for each year.

Using the average capital tax rate has the advantage of being based on relatively 'hard', observable data and, for most countries, it will provide a reasonable approximation of the capital tax burden faced by producers. However, an argument can also be mounted that deadweight loss studies should use estimates of the effective marginal tax rate (EMTR) producers face as this will be a closer approximation to the rate producers respond to in making investment and other production decisions. While the use of EMTRs may be desirable, their construction is informationally very demanding. We hope to implement this approach in the future. In the meantime, we use the average tax rate on capital to proxy the tax burden producers face and to test our producer model.

To summarize: we assume that the undistorted user cost of capital for each of our three assets has the following form:

(53) $U^* = uP \equiv [R - \alpha + \delta(1+\alpha)]P$

where P is the beginning of the period asset price, R is a tax free opportunity cost of capital[8], α is an anticipated asset inflation rate[9] and δ is our assumed geometric or declining balance depreciation rate. The assumed depreciation rate for nonresidential structures was 3.5% (i.e., $\delta_{NR} \equiv .035$), for machinery and equipment was 12.5% (i.e., $\delta_{ME} \equiv .125$) and for inventory stocks was 0 (i.e., $\delta_{IS} \equiv 0$). The nominal interest rates or opportunity costs of capital R are tabled below in Table 4 along with our estimated asset inflation rates, α_{NR}, α_{ME} and α_{IS} for nonresidential structures, machinery and equipment and inventory stocks respectively. The beginning of the period asset prices, P_{NR}, P_{ME} and P_{IS} are tabled in Table 5 below. The undistorted user costs in percentage form, the terms u in (53), are tabled in Table 4 below for each of our three assets.

To obtain the final tax distorted user cost for each asset, we need to add the total distortion wedge, wP, to the undistorted user costs defined by (53). These wedge terms w for each asset are defined as follows:

(54) $w_{NR} \equiv \tau^* + t^* + \tau_{PNR} + \tau_{CNR}$;

(55) $w_{ME} \equiv \tau^* + t^* + \tau_{CME}$;

(56) $w_{IS} \equiv \tau^* + t^*$

where $\tau^* + t^*$ is the combined (asset) business and personal tax rate on capital, τ_{PNR} is the property tax rate on nonresidential capital, and τ_{CNR} and τ_{CME} are the commodity tax rates times the undistorted user costs u_{NR} and u_{ME} respectively.[10] These tax rates and are listed in Table 3 below along with R, the nominal opportunity cost of capital, and the three (anticipated) asset inflation rates, α_{NR}, α_{ME} and α_{IS}. The capital tax rates are also graphed in Figure 4.

[8] We assume that $1+R = (1 + .05)(1 + \alpha_C)$ where r = .05 is an assumed required real rate of return and α_C is a smoothed inflation rate for Canadian consumer prices. We used smoothed ex post actual consumer price inflation rates to proxy these anticipated inflation rates. We used the Lowess nonparametric smoothing option in Shazam with the smoothing parameter f = .2.

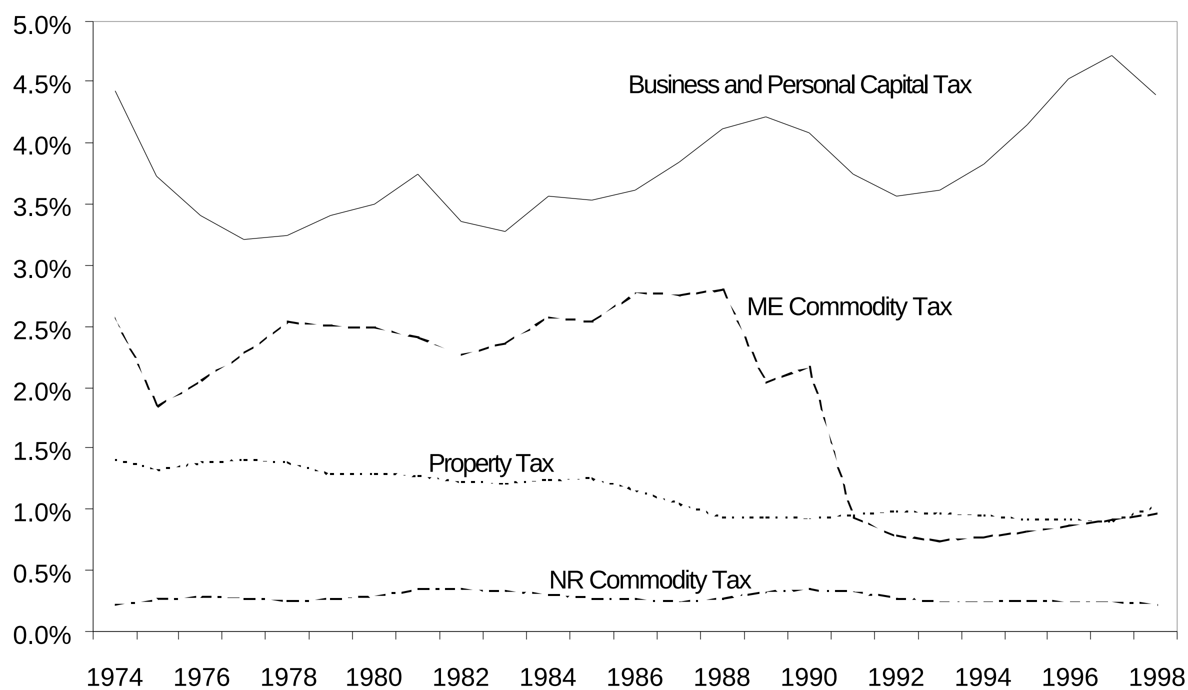
[9] We used smoothed ex post asset inflation rates to proxy these anticipated asset inflation rates. We used the Lowess nonparametric smoothing option in Shazam with the smoothing parameter f = .2.

[10] The taxation of purchases of capital inputs by the use of Provincial sales taxes (and some Federal sales taxes before the imposition of the GST) leads to a tax distortion at the moment of purchase. We essentially distribute this distortion over the useful life of the asset so the tax rates listed in Table 3 are lower than the instantaneous rate of taxation.

Table 3: Tax Rates and Asset Inflation Rates (percentages)

Year	$\tau^{*}+t^{*}$	τ_{PNR}	τ_{CNR}	τ_{CME}	R	α_{NR}	α_{ME}	α_{IS}
1974	4.43	1.41	0.22	2.57	14.57	11.22	9.43	12.90
1975	3.73	1.33	0.27	1.85	13.71	8.60	8.61	9.11
1976	3.41	1.39	0.29	2.06	12.78	6.94	6.47	5.66
1977	3.21	1.41	0.28	2.28	13.05	7.57	4.58	6.52
1978	3.24	1.39	0.26	2.54	14.17	9.25	2.54	10.11
1979	3.40	1.29	0.27	2.51	15.25	10.18	1.85	12.38
1980	3.50	1.30	0.30	2.49	15.43	9.11	1.54	11.69
1981	3.75	1.28	0.35	2.41	14.31	6.45	1.37	8.49
1982	3.36	1.23	0.35	2.27	12.33	4.11	0.83	5.25
1983	3.28	1.22	0.34	2.37	10.48	2.56	-0.98	2.94
1984	3.56	1.25	0.31	2.58	9.46	2.60	-1.95	2.01
1985	3.54	1.26	0.28	2.54	9.19	3.34	-2.02	1.83
1986	3.62	1.16	0.27	2.78	9.24	3.93	-1.48	2.02
1987	3.84	1.05	0.25	2.76	9.32	4.43	-0.72	2.40
1988	4.11	0.94	0.28	2.80	9.53	3.78	-0.76	2.21
1989	4.21	0.94	0.33	2.04	9.46	1.97	-1.37	1.14
1990	4.09	0.93	0.35	2.17	8.80	0.56	-1.96	0.04
1991	3.74	0.96	0.33	0.94	7.91	0.22	-1.30	0.39
1992	3.57	0.99	0.28	0.79	6.78	0.97	0.18	1.31
1993	3.61	0.97	0.25	0.75	6.35	1.57	0.72	2.73
1994	3.82	0.96	0.25	0.78	6.39	1.56	0.00	2.47
1995	4.15	0.92	0.26	0.82	6.53	1.37	-0.97	-0.41
1996	4.53	0.92	0.25	0.87	6.58	1.64	-2.23	-2.10
1997	4.71	0.91	0.25	0.92	6.65	1.99	-3.63	-3.50
1998	4.40	1.03	0.23	0.97	6.67	2.42	-5.03	-4.81

Figure 4: Capital Tax Rates



It can be seen that the combined personal and business tax rates on the return to capital are in the 3.5% to 4.5 % range for our sample period, 1974-1998. Our estimated property tax rate which falls on the use of structures is in the .9% to 1.4% per year range. Our estimated commodity tax rate on the use of structures is in the .2% to .35% range and on the use of machinery and equipment is in the 2.5% to .75% range. The sum of all of these tax rates leads to an overall wedge tax rate in the 5% to 6% range for nonresidential structures, in the 4% to 7% range for machinery and equipment and in the 3.5% to 4.7% range for inventory stocks.

The undistorted user costs for nonresidential structures, u_{NR}, machinery and equipment, u_{ME}, and for inventory stocks, u_{IS}, (in proportional form; see (53) above) are listed in Table 4 below, along with the tax wedges. The undistorted user costs and tax wedges are then graphed in Figures 5 and 6, respectively.

Table 4: Undistorted User Costs (deflated by Asset Prices) and Tax Wedges

Year	u_{NR}	u_{ME}	u_{IS}	w_{NR}	w_{ME}	w_{IS}
1974	0.0724	0.1882	0.0166	0.0605	0.0700	0.0443
1975	0.0891	0.1868	0.0460	0.0533	0.0559	0.0373
1976	0.0958	0.1962	0.0712	0.0508	0.0547	0.0341
1977	0.0925	0.2154	0.0652	0.0490	0.0549	0.0321
1978	0.0875	0.2445	0.0406	0.0489	0.0578	0.0324
1979	0.0892	0.2614	0.0287	0.0496	0.0591	0.0340
1980	0.1014	0.2658	0.0373	0.0510	0.0598	0.0350
1981	0.1158	0.2561	0.0582	0.0538	0.0616	0.0375
1982	0.1187	0.2411	0.0708	0.0495	0.0563	0.0336
1983	0.1150	0.2383	0.0754	0.0485	0.0566	0.0328
1984	0.1045	0.2366	0.0745	0.0512	0.0614	0.0356
1985	0.0947	0.2346	0.0736	0.0508	0.0608	0.0354
1986	0.0894	0.2304	0.0722	0.0504	0.0639	0.0362
1987	0.0854	0.2245	0.0692	0.0514	0.0660	0.0384
1988	0.0938	0.2269	0.0731	0.0532	0.0691	0.0411
1989	0.1106	0.2316	0.0832	0.0548	0.0625	0.0421
1990	0.1176	0.2301	0.0875	0.0537	0.0626	0.0409
1991	0.1120	0.2155	0.0752	0.0503	0.0469	0.0374
1992	0.0934	0.1912	0.0547	0.0484	0.0436	0.0357
1993	0.0833	0.1822	0.0362	0.0483	0.0436	0.0361
1994	0.0838	0.1888	0.0392	0.0503	0.0460	0.0382
1995	0.0871	0.1987	0.0694	0.0534	0.0497	0.0415
1996	0.0849	0.2103	0.0868	0.0570	0.0540	0.0453
1997	0.0823	0.2232	0.1015	0.0587	0.0563	0.0471
1998	0.0784	0.2357	0.1148	0.0566	0.0537	0.0440

Figure 5: Undistorted User Costs

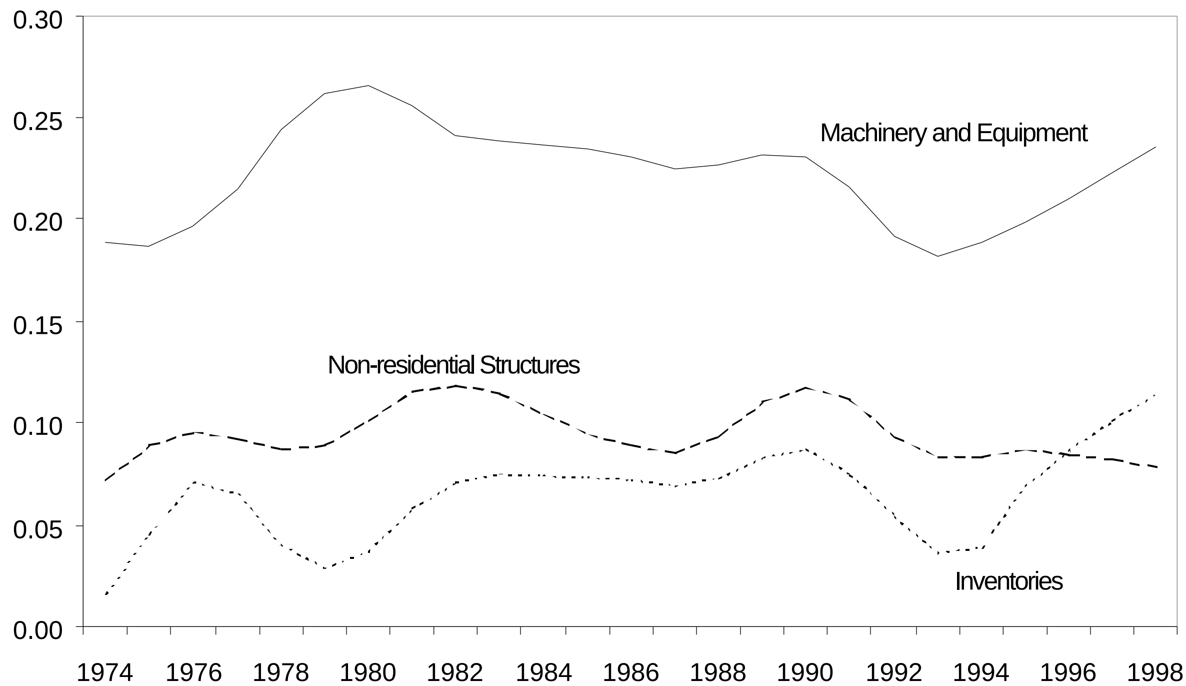
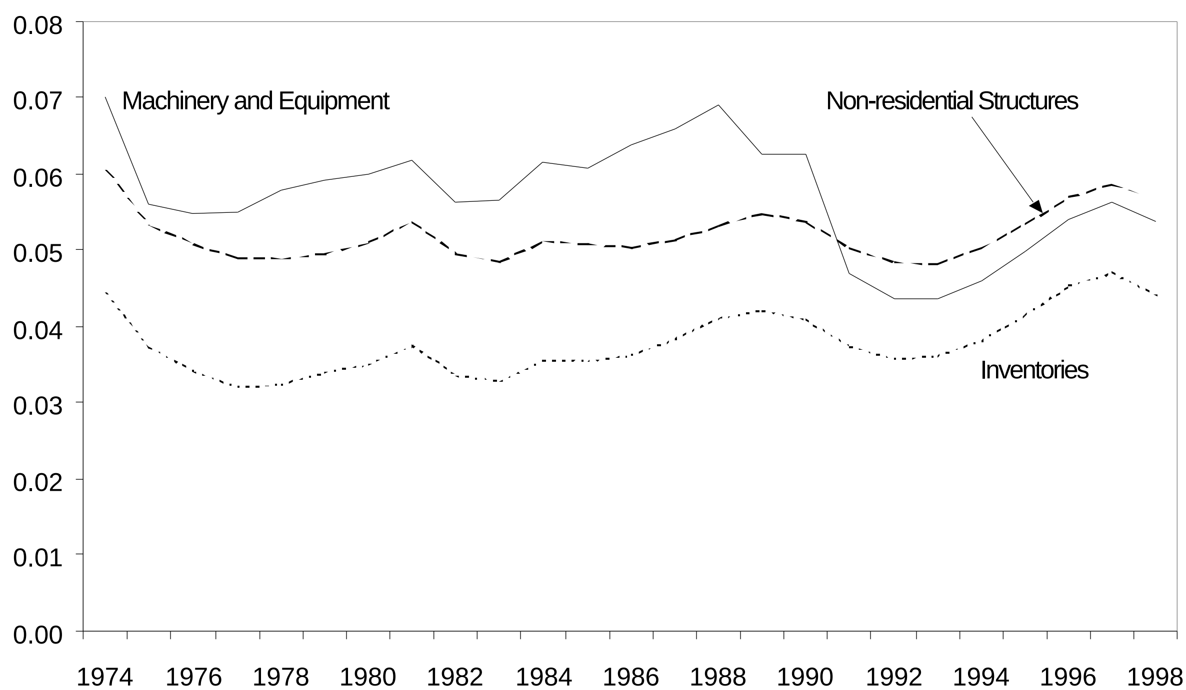


Figure 6: Tax Wedges



Finally, the beginning of the period asset prices (before commodity taxes), P_{NR}, P_{ME} and P_{IS} are listed in Table 5 along with the corresponding beginning of the period capital stocks, K_{NR}, K_{ME} and K_{IS} (in billions of 1974 Canadian dollars). The capital stocks were constructed using the perpetual inventory method using investment data back to 1926. Starting stocks for 1926 for nonresidential structures and machinery and equipment were constructed using the assumed depreciation rates plus the assumption that investment had been growing at a 2% rate for all years prior to 1926.

Table 5: Capital Prices and Quantities (in billions of 1974 dollars)

Year	P_{NR}	P_{ME}	P_{IS}	K_{NR}	K_{ME}	K_{IS}
1974	4.43	1.41	0.22	2.57	6.05	7.00
1975	3.73	1.33	0.27	1.85	5.33	5.59
1976	3.41	1.39	0.29	2.06	5.08	5.47
1977	3.21	1.41	0.28	2.28	4.90	5.49
1978	3.24	1.39	0.26	2.54	4.89	5.78
1979	3.40	1.29	0.27	2.51	4.96	5.91
1980	3.50	1.30	0.30	2.49	5.10	5.98
1981	3.75	1.28	0.35	2.41	5.38	6.16
1982	3.36	1.23	0.35	2.27	4.95	5.63
1983	3.28	1.22	0.34	2.37	4.85	5.66
1984	3.56	1.25	0.31	2.58	5.12	6.14
1985	3.54	1.26	0.28	2.54	5.08	6.08
1986	3.62	1.16	0.27	2.78	5.04	6.39
1987	3.84	1.05	0.25	2.76	5.14	6.60
1988	4.11	0.94	0.28	2.80	5.32	6.91
1989	4.21	0.94	0.33	2.04	5.48	6.25
1990	4.09	0.93	0.35	2.17	5.37	6.26
1991	3.74	0.96	0.33	0.94	5.03	4.69
1992	3.57	0.99	0.28	0.79	4.84	4.36
1993	3.61	0.97	0.25	0.75	4.83	4.36
1994	3.82	0.96	0.25	0.78	5.03	4.60
1995	4.15	0.92	0.26	0.82	5.34	4.97
1996	4.53	0.92	0.25	0.87	5.70	5.40
1997	4.71	0.91	0.25	0.92	5.87	5.63
1998	4.40	1.03	0.23	0.97	5.66	5.37

In the next section, we show how the simple excess burden model explained in section 2 above can be generalized to cover the case of many (noncapital) outputs and inputs and 3 types of reproducible capital (compared to the single reproducible capital stock model of section 2).

4. Excess Burdens in a Single Capital Model Using Cash Flow Profit Functions

The model presented in section 2 above was adequate to introduce the reader to the basic concepts involved in measuring the deadweight loss due to the taxation of capital. However, this model suffered from a number of defects including: (i) it was highly aggregated; (ii) the investment good was assumed to be perfectly substitutable with the consumption good and (iii) the economy was closed, ie there was no international trade in goods and services. In this section, we shall relax the above restrictions except we continue to assume that there is only one reproducible capital stock in the economy. In the following section, we shall deal with the multiple stock case.

In order to simplify our derivation of excess burden formulae, we shall make use of the producer's profit function. The profit function simply provides an alternative method for representing the production function, or more generally, the producer's production possibilities set.[11] The use of the profit function not only facilitates the derivation of deadweight loss formulae but it is also very convenient from the viewpoint of the econometric estimation of production functions or technology sets. For additional material on profit functions and references to the literature, see the section on our producer model methodology.

We assume that there are M non-capital variable outputs and inputs that are produced and utilised in the private production sector. The positive prices that producers face in period t for these M variable commodities are denoted by $(p_1^t,\ldots,p_M^t) \equiv p^t$. The corresponding variable outputs and inputs produced and used during period t are denoted by the quantity vector $y^t \equiv (y_1^t,..,y_M^t)$. The list of outputs includes consumption goods and services, government purchases of goods and services from the private sector, an investment good that corresponds to the single reproducible capital stock in our model, exports, imports and labour inputs. If commodity m is an input, then y_m^t has a negative sign. The price of one unit of the reproducible capital stock is P^t in period t. The private business sector of the economy utilises the beginning of period t capital stock K^t and the fixed factor input F^t. The period t set of feasible net output vectors y, conditional on a beginning of the period capital stock K^t and fixed factor input F^t is denoted by the set S^t. The private sector's period *t cash flow profit function* π^t is defined as follows:

$$(57) \qquad \pi^t(p^t, K^t, F^t) \equiv \max_y \{p^t \bullet y : (y, K^t, F^t) \quad S^t\}$$

where $p^t \bullet y$ denotes the inner product of the vectors p^t and y. In words, $\pi^t(p^t, K^t, F^t)$ is the maximum value added less the value of labour inputs that the private sector can produce given that producers face the prices p^t for these variable outputs and inputs and given that producers have the fixed stocks K^t and F^t of reproducible and non-reproducible capital available to them at the beginning of period t. In other words, $\pi^t(p^t, K^t, F^t)$ is the maximum *cash flow* that the

[11] For additional material on profit functions and duality theory, see Diewert (1974a) (1993).

economy can earn in period t, given that it faces the price vector p^t for variable inputs and outputs and has the use of K^t units of reproducible capital and F^t units of fixed factors.

The counterpart to our earlier tax distorted profit maximisation problem (5) is now the following problem:

(58) $\max_K \{\pi^t(p^t, K, F^t) - (u^t + w^t)P^t K\} \equiv \Pi^t(p^t, (u^t + w^t)P^t, F^t)$

where u^t is the period t (deflated) undistorted user cost of capital defined in (53) in the previous section and w^t is the period t (deflated) total tax distortions wedge defined by one of (54)-(56) in the previous section. Note that the optimised objective function in (58) defines another profit function Π^t, which we call the *rents profit function*. We will use the rents profit function in the next section. We assume that the observed period t capital stock K^t satisfies the first order necessary conditions for the unconstrained maximisation problem in (57):

(59) $\partial\pi^t(p^t, K, F^t)/\partial K = (u^t + w^t)P^t$.

Equation (59) is the counterpart to our old equation (6). For an arbitrary capital tax wedge w, we denote the capital solution to (59), where w replaces w^t, by $K(w)$. Substituting $K(w)$ into (59) and differentiating with respect to w yields the following equation for $K'(wt)$:

(60) $K'(w^t) = [\partial^2\pi^t(p^t, K^t, F^t)/\partial K^2]^{-1} P^t$.

Recall the definition of the producer surplus function $S(w)$, defined by (8) above. This function evaluated the outputs produced and the inputs used by the private sector at undistorted prices. Using the cash flow profit function and the capital demand function $K(w)$, which solves (59) when the observed period t distortion term w^t is replaced by an arbitrary distortion term w, we redefine the *surplus function* for period t[12], $S(w)$ as follows:

(61) $S(w) \equiv \pi^t(p^t, K(w), F^t) - u^t P^t K(w)$.

Differentiating $S(w)$ with respect to w yields the following equations:

(62) $S'(w) = [\partial\pi^t(p^t, K(w), F^t)/\partial K]\, K'(w) - u^t P^t K'(w)$

$= [(u^t + w)P^t]\, K'(w) - u^t P^t K'(w)$ using (59) with w replacing w^t

$= w\, P^t K'(w)$.

Evaluating (61) at $w = 0$ and $w = w^t$ leads to the following equalities:

[12] The capital demand function, K(w), and the surplus function, S(w), should actually be denoted as $K^t(w)$ and $S^t(w)$ to indicate that these functions depend on the period t price data and on the period t cash flow function π^t. The function π^t depends on time t due to technical change.

(63) $S'(0) = 0$;

(64) $S'(w^t) = w^t P^t K'(w^t)$

$$= w^t P^t [\partial^2\pi^t(p^t, K^t, F^t)/\partial K^2]^{-1} P^t \qquad \text{using (60).}$$

It is possible to approximate $S(w^t) - S(0)$ by a second order Taylor series expansion, as we did in section 2 above. Of course, if $S(w)$ is (locally) a quadratic function, this approximation will be (locally) exact. Another approximation for $S(w^t) - S(0)$ that is exact if $S(w)$ is quadratic is the following one[13]:

(65) $S(w^t) - S(0) \approx (1/2)\,[S'(w^t) + S'(0)]\,[w^t - 0]$

$$= (1/2)\, S'(w^t)\, w^t \qquad \text{using (63)}$$

$$= (1/2)\,(w^t)^2 P^t [\partial^2\pi^t(p^t, K^t, F^t)/\partial K^2]^{-1} P^t \qquad \text{using (64)}$$

Thus, we have the following quadratic approximation for $S(0) - S(w^t)$:

(66) $S(0) - S(w^t) \approx -(1/2)\,(w^t)^2 P^t [\partial^2\pi^t(p^t, K^t, F^t)/\partial K^2]^{-1} P^t$.

Recall our earlier definition (13) of the loss of surplus due to the distortion wedge w^* as a fraction of undistorted GDP, $[S(0) - S(w^*)]/Y(0)$, and the second order approximation to this loss, $A(w^*)$ defined by (14). We will now express the loss as a fraction of the tax distorted level of GDP, which we denote by $Y(w^t)$ for period t. Thus, using (58), our new *second order approximation to the loss of output in period t* is:

(67) $A(w^t) \equiv -(1/2)\,(w^t)^2 P^t [\partial^2\pi^t(p^t, K^t, F^t)/\partial K^2]^{-1} P^t / Y(w^t)$.

Given an econometrically estimated cash flow function π^t, we can readily calculate $A(w^t)$ defined by (67).

We turn now to the problem of generalising our old marginal cost function defined earlier by (23). We now define the *cost of the system of capital taxation function* $C(w)$ as the difference between the optimal value of output $S(0)$ and the tax distorted value of output $S(w^t)$:

(68) $C(w^t) \equiv S(0) - S(w^t)$

where $S(w)$ is now defined by (61) above. Differentiating (68) with respect to w and evaluating w at the period t tax distortion rate w^t leads to the following period *t marginal cost of the system of capital taxation*:

(69) $MC(w^t) \equiv C'(w^t)$

[13] See Diewert's (1976; 118) quadratic approximation lemma.

$$= -\, S'(w^t) \qquad \text{using (68)}$$

$$= -\, w^t\, P^t\, K'(w^t) \qquad \text{using (64)}$$

(70) $$= -\, w^t\, P^t\, [\partial^2 \pi^t(p^t, K^t, F^t)/\partial K^2]^{-1}\, P^t \qquad \text{using (64).}$$

Thus, (70) defines the marginal cost of increasing the period t distortion wedge w^t by a small amount. We turn now to the problem of defining the corresponding marginal benefit function for our new model.

Recall equation (24) above, which expressed the total revenue T from all sources of capital taxation. The counterpart is now:

(71) $T^t = w^t\, P^t\, K^t$.

We replace the period t distortion rate w^t by a general distortion rate w and let $K(w)$ be the solution to (59) when w^t is replaced by w. Making these substitutions into (71) leads to the following definition for the *period t total tax revenues* as a function of the distortion rate w:

(72) $T(w) \equiv w\, P^t\, K(w)$.

We can now obtain the marginal increase in capital tax revenues by differentiating $T(w)$ defined by (72) with respect to w and evaluating the resulting derivative at the observed period t distortion rate w^t. Thus, define the period *t marginal benefit of an increase in the rate of capital taxation* as:

(73) $MB(w^t) \equiv T'(w^t)$

$$= P^t\, K(w^t) + w^t\, P^t\, K'(w^t) \qquad \text{differentiating (72)}$$

(74) $$= P^t\, K(w^t) + w^t\, P^t\, [\partial^2 \pi^t(p^t, K^t, F^t)/\partial K^2]^{-1}\, P^t \qquad \text{using (64).}$$

Thus, given an econometric estimate for the period t cash flow function, π^t, the right hand side of (74) can be evaluated using observable data. The first term on the right hand side of (74) will be positive and the second term will be negative. If there is a high degree of substitutability of capital for other inputs and outputs in period t, then the second term can make the overall marginal benefits of increasing capital taxes negative. In this case, the government will achieve both increased productive efficiency and higher tax revenues by reducing capital tax distortions.

The *marginal excess burden of increasing the period t capital tax distortion rate* w^t by a small amount is simply the ratio of the marginal cost $MC(w^t)$ defined by (70) above divided by the marginal tax revenue $MB(w^t)$ defined by (74) above:

(75) $MEB(w^t) \equiv MC(w^t) / MB(w^t) = -\, w^t\, K'(w^t) / \{P^t\, K(w^t) + w^t\, P^t\, K'(w^t)\}$.

Thus, with the use of duality theory, it proved to be quite straightforward to generalise the very simple one output model of the previous section to an open economy with many outputs and inputs. However, the model presented in this section still only had a single reproducible capital input. Before we deal with the case of many capital inputs, we rework the analysis presented in this section using the pure rents profit function Π^t (see (58) above) in place of the cash flow profit function π^t, since we used the former function in our econometric work.

5. Excess Burdens in a Single Capital Model Using Pure Rents Profit Functions

Recall that the period *t pure rents* profit function $\Pi^t(p^t, (u^t + w^t)P^t, F^t)$ was defined by (58) above. Recall also the capital demand function $K(w)$ that was the solution to the first order condition (59), where w^t was replaced by a general distortion wedge w. It can be shown[14] that the capital demand function $K(w)$ can be obtained directly from the pure rents profit function by differentiating $\Pi^t(p^t, (u^t + w)P^t, F^t)$ with respect to the tax distorted user cost, which we denote by $V \equiv (u^t + w)P^t$:

(76) $K(w) = -\,\partial\,\Pi^t(p^t, (u^t + w)P^t, F^t)/\partial V.$

We can differentiate $K(w)$ with respect to w and evaluate the resulting derivative at $w = w^t$. We obtain the following counterpart to our old formula (60) in the previous section:

(77) $K'(w^t) = -\,[\partial^2\Pi^t(p^t, (u^t + w)P^t, F^t)/\partial V^2]\, P^t.$

Recall our old definition (61) of the *surplus function* $S(w)$ using the cash flow profit function π^t. Using the pure rents profit function Π^t, we can redefine $S(w)$ as follows:

(78) $S(w) \equiv \Pi^t(p^t, (u^t + w)P^t, F^t) + w\,P^t\,K(w).$

Differentiating $S(w)$ with respect to w yields the following equations:

(79) $S'(w) = [\partial\Pi^t(p^t, (u^t + w)P^t, F^t)/\partial V]\,P^t + P^t\,K(w) + w\,P^t\,K'(w)$

$= -\,K(w)\,P^t + P^t\,K(w) + w\,P^t\,K'(w)$ using (76)

$= w\,P^t\,K'(w).$

Evaluating (79) at $w = 0$ and $w = w^t$ leads to the following equalities:

(80) $S'(0) = 0;$

[14] This is known as Hotelling's Lemma; see Diewert (1993; 166).

(81) $S'(w^t) = w^t P^t K'(w^t)$

(82) $= -w^t P^t [\partial^2\Pi^t(p^t, (u^t + w)P^t, F^t)/\partial V^2] P^t$. using (77).

Using (80) and (82) and making use of the quadratic approximation (65) again to approximate $S(0)$, we have the following quadratic approximation for $S(0) - S(w^t)$:

(83) $S(0) - S(w^t) \approx (1/2)\,(w^t)^2 P^t [\partial^2\Pi^t(p^t, (u^t + w)P^t, F^t)/\partial V^2]\; P^t$.

Recall our earlier definition of the loss of surplus due to the distortion wedge w^* as a fraction of undistorted GDP, $[S(0) - S(w^*)]/Y(0)$, and the second order approximation to this loss, $A(w^*)$ defined by (14). As in the previous section, we will again express the loss as a fraction of the tax distorted level of GDP, which we again denote by $Y(w^t)$ for period t. Thus, using (83), we obtain the following counterpart to (67) in the previous section; i.e., our new *second order approximation to the loss of output in period t* is:

(84) $A(w^t) \equiv (1/2)\,(w^t)^2 P^t [\partial^2\Pi^t(p^t, (u^t + w)P^t, F^t)/\partial V^2]\; P^t / Y(w^t)$.

Given an econometrically estimated pure rents function Π^t, we can readily calculate $A(w^t)$ defined by (84).

We turn now to the problem of generalising our old marginal cost function $MC(w^*)$ defined earlier by (69). As in the previous section, we define the *cost of the system of capital taxation function* $C(w)$ as the difference between the optimal value of output $S(0)$ and the tax distorted value of output $S(w^t)$:

(85) $C(w^t) \equiv S(0) - S(w^t)$

where $S(w)$ is now defined by (78) above. Differentiating (85) with respect to w and evaluating w at the period t tax distortion rate w^t leads to the following period *t marginal cost of the system of capital taxation*:

(86) $MC(w^t) \equiv C'(w^t)$

$= - S'(w^t)$ using (85)

(87) $= - w^t P^t K'(w^t)$ using (81)

(88) $= w^t P^t [\partial^2\Pi^t(p^t, (u^t + w)P^t, F^t)/\partial V^2] P^t$ using (82).

Thus, (86) defines the marginal cost of increasing the period t distortion wedge w^t by a small amount and (87) and (88) are formulae which can be used to evaluate this marginal cost. We turn now to the problem of defining the corresponding marginal benefit function for our new model.

Recall equations (24) and (71) above, which expressed the total revenue T from all sources of capital taxation in terms the distortion wedge $W = wP$. We can rewrite (24) for period t total capital tax revenue T^t as a function of the period t distortion wedge w^t as follows:

(89) $T^t = w^t P^t K^t (w^t).$

We replace the period t distortion rate w^t by a general distortion rate w and let $K(w)$ be defined by (76). Making these substitutions into (89) leads to the following definition for the *period t total tax revenues* as a function of the distortion rate w:

(90) $T(w) \equiv w P^t K(w).$

Now differentiate T(w) defined by (90) with respect to w and evaluate the resulting derivative at the observed period t distortion rate w^t. This defines the period t *marginal benefit of an increase in the rate of capital taxation* as:

(91) $MB(w^t) \equiv T'(w^t)$

(92) $= P^t K(w^t) + w^t P^t K'(w^t)$ differentiating (90)

(93) $= P^t K(w^t) - w^t P^t [\partial^2 \Pi^2(p^t, (u^t + w)P^t, F^t)/\partial V^2] P^t$ using (77).

Thus, given an econometric estimate for the period t cash flow function, Π^t, the right hand side of (93) can be evaluated using observable data. As was the case with our earlier formulae (74), the first term on the right hand side of (93) will be positive and the second term will be negative. If there is a high degree of substitutability of capital for other inputs and outputs in period t, then the second term can make the overall marginal benefits of increasing capital taxes negative. In this case, the government will achieve both increased productive efficiency and higher tax revenues by reducing capital tax distortions.

The *marginal excess burden of increasing the period t capital tax distortion rate* w^t by a small amount is simply the ratio of the marginal cost $MC(w^t)$ defined by (87) above divided by the marginal tax revenue $MB(w^t)$ defined by (92) above:

(94) $MEB(w^t) \equiv MC(w^t) / MB(w^t) = - w^t K'(w^t) / \{P^t K(w^t) + w^t P^t K'(w^t)\}.$

Note that our new formula for the $MEB(w^t)$, (94), coincides with our formula for the marginal excess burden of an increase in the wedge w^t in the previous section, (75). However, in this section, we obtain an estimate of the capital demand derivative $K'(w^t)$ using equation (77), which involves the second order partial derivative of the period t pure profits function Π^t with respect to the tax distorted user cost $V \equiv (u + w)P$, whereas in the previous section, we obtained an estimate of the capital demand derivative $K'(w^t)$ using equation (60), which involved the second order partial derivative of the period t cash flow function π^t with respect to capital K.

The profit function models presented in this section and the previous section had only a single reproducible capital input. In the following section, we relax this restriction.

6. A Multiple Capital Stock Model

In this section, we assume that $K \equiv (K_1, K_2, K_3)$ is now a three dimensional vector of reproducible capital stocks rather than being the scalar capital stock assumed in the previous sections of this paper. We now adapt the analysis presented in the previous section to this multiple capital stock case.

The period *t cash flow* profit function π^t can still be defined by (57) above but we now interpret K as a vector. Recall that the period *t pure rents* profit function $\Pi^t(p^t, (u^t + w^t)P^t, F^t)$ was defined by (58) above. This same definition is still applicable with K being interpreted as a vector but now $(u^t + w^t)P^t \equiv [(u_1^t + w_1^t)P_1^t, (u_2^t + w_2^t)P_2^t, (u_3^t + w_3^t)P_3^t]$ is interpreted as a vector of period t tax distorted user costs for the reproducible capital stocks. We define the period t vector of stock prices for the reproducible capital stock components as $P^t \equiv [P_1^t, P_2^t, P_3^t]$. The vector of period t undistorted (deflated) user costs of capital is defined as $u^t \equiv [u_1^t, u_2^t, u_3^t]$ and the vector of deflated period t distortion terms is defined as $w^t \equiv [w_1^t, w_2^t, w_3^t]$. Each distortion term is defined as in equations (54)-(56) in section 3 above.[15] The undistorted (deflated) user costs of capital, u_n, are defined as in section 3; ie we have:

$$(95) \qquad u_n \equiv R - \alpha_n + \delta_n (1 + \alpha_n) ; \qquad\qquad n = 1,2,3.$$

It can be shown that the vector of capital demand functions $K(w) \equiv [K_1(w_1, w_2, w_3), K_2(w_1, w_2, w_3), K_3(w_1, w_2, w_3)]$ can be obtained directly from the pure rents profit function by differentiating $\Pi^t(p^t, (u^t + w)P^t, F^t)$ with respect to the components of the tax distorted user costs, which we denote by the vector $V \equiv (u^t + w)P^t \equiv [(u_1^t + w_1)P_1^t, (u_2^t + w_2)P_2^t, [(u_3^t + w_3)P_3^t]$:

$$(96) \quad K(w) = - \nabla_V \Pi^t(p^t, (u^t + w)\mathrm{P}^t, F^t)$$

where $\nabla_V \Pi^t(p^t, (u^t + w)P^t, F^t) = \nabla_V \Pi^t(p^t, V, F^t)$ denotes the vector of first order partial derivatives of $\Pi^t(p^t, V, F^t)$ with respect to the components of V; ie we have

$$(97) \qquad \nabla_V \Pi^t(p^t, V, F^t) \equiv [\, \partial \Pi^t(p^t, V, F^t)/\partial V_1, \partial \Pi^t(p^t, V, F^t)/\partial V_2, \partial \Pi^t(p^t, V, F^t)/\partial V_3] \,.$$

[15] Note that we have changed our notation for the deflated wedges.

We can differentiate the vector $K(w)$ defined by (96) with respect to the components of w and evaluate the resulting three by three matrix of derivatives at $w = w^t$. We obtain the following counterpart to our old formula (77) in the previous section:

(98) $\nabla_w K(w^t) = -[\nabla^2_{VV} \Pi^t(p^t, (u^t + w)P^t, F^t)]\,\mathrm{Diag}(P^t)$

where $\nabla^2_{VV} \Pi^t$ denotes the three by three matrix of second order partial derivatives of Π^t with respect to the three user costs and $\mathrm{Diag}(P^t)$ is a three by three diagonal matrix with the elements of the three dimensional vector of period t capital stock prices, $P^t \equiv [P_1^t, P_2^t, P_3^t]$, running down the main diagonal.

Recall our old definition (78) of the *surplus function* $S(w)$. We can redefine the *surplus function* $S(w)$ using the cash flow profit function Π^t as follows:

(99) $S(w) \equiv \Pi^t(p^t, (u^t + w)P^t, F^t) + \sum_{n=1}^{3} w_n P_n^t K_n(w).$

Differentiating $S(w)$ with respect to the components of w yields the following equations:

(100) $$\begin{aligned}\partial S(w)/\partial w_n &= [\partial\Pi^t(p^t, (u^t + w)P^t, F^t)/\partial V_n]\, P_n^t + P_n^t K_n(w) \\ &\quad + \sum_{j=1}^{3} w_j P_j^t\, \partial K_j(w)/\partial w_n \\ &= -K_n(w) P_n^t + P_n^t K_n(w) + \sum_{j=1}^{3} w_j P_j^t\, \partial K_j(w)/\partial w_n \qquad \text{using (96)} \\ &= \sum_{j=1}^{3} w_j P_j^t\, \partial K_j(w)/\partial w_n \qquad \text{for } n = 1,2,3.\end{aligned}$$

Evaluating (100) at $w = 0_3$ and $w = w^t$ leads to the following equalities:

(101) $\partial S(0_3)/\partial w_n = 0$ for $n = 1,2,3$;

(102) $\partial S(w^t)/\partial w_n = \sum_{j=1}^{3} w_j^t P_j^t\, \partial K_j(w^t)/\partial w_n$ for $n = 1,2,3$

(103) $\qquad = -\sum_{j=1}^{3} w_j^t P_j^t [\partial^2\Pi^t(p^t, (u^t + w^t)P^t, F^t)/\partial V_j \partial V_n]\, P_n^t.$ using (98).

Using (101) and (103) and making use of Diewert's (1976; 118) quadratic approximation lemma, we have the following quadratic approximation for $S(0_3) - S(w^t)$:

(104) $S(0_3) - S(w^t) \approx (1/2) \sum_{j=1}^{3} \sum_{n=1}^{3} w_j^t P_j^t [\partial^2\Pi^t(p^t, (u^t + w^t)P^t, F^t)/\partial V_j \partial V_n] w_n^t P_n^t\,.$

As in the previous section, we will again express the loss (104) as a fraction of the tax distorted level of GDP, which we again denote by $Y(w^t)$ for period t. Thus, using (104), we obtain the following counterpart to (84) in the previous section; ie our new *second order approximation to the loss of output in period t* is[16]:

(105) $A(w^t) \equiv (1/2) \sum_{j=1}^{3} \sum_{n=1}^{3} w_j^t P_j^t [\partial^2 \Pi^t(p^t, (u^t + w^t)P^t, F^t)/\partial V_j \partial V_n] w_n^t P_n^t / Y(w^t)$.

Given an econometrically estimated pure rents function Π^t, we can readily calculate $A(w^t)$ defined by (105).

We turn now to the problem of defining the marginal cost function. As in the previous section, we define the *cost of the system of capital taxation tax function* $C(w)$ as the difference between the optimal value of output $S(0_3)$ and the tax distorted value of output $S(w^t)$:

(106) $C(w^t) \equiv S(0_3) - S(w^t)$

where $S(w)$ is now defined by (99) above. Differentiating (106) with respect to the components of w and evaluating w at the period t tax distortion vector w^t leads to the following period t *marginal costs with respect to a small increase in the n*th *deflated distortion wedge* w_n:

(107) $MC_n(w^t) \equiv \partial C(w^t)/\partial w_n$ for $n = 1,2,3$

$= - \partial S(w^t)/\partial w_n$ using (106)

(108) $= - \sum_{j=1}^{3} w_j^t P_j^t \partial K_j(w^t)/\partial w_n$ using (102)

(109) $= \sum_{j=1}^{3} w_j^t P_j^t [\partial^2 \Pi^t(p^t, (u^t + w^t)P^t, F^t)/\partial V_j \partial V_n] P_n^t$ using (103).

Thus, (107) defines the marginal cost of increasing the period t distortion wedge for capital input n, w_n^t, by a small amount, and (108) and (109) are formulae which can be used to evaluate this marginal cost. However, now we encounter a difference in the multiple capital stock model of this section compared with the single capital stock model in the previous section. In the previous sections, we did not have to consider in detail the effects of changes in each tax policy parameter; all we had to know is whether the change in tax policy increased or decreased the single deflated wedge. Now we have to consider changes in each tax parameter separately. The changes in tax policy that we will consider in section 8 are:

- An *increase* in the business income tax rate τ^*;
- An *increase* in the rate of property tax τ_{PNR} or indirect sales taxation τ_{CNR} (an increase in either of these tax rates has the same effect); or

[16] Since the pure rents profit function Π^t must be a convex function in its price arguments, the matrix of second order partial derivatives in (105) must be positive semidefinite. This means that the approximate loss defined by (105) must be nonnegative.

- An *increase* in the rate of indirect sales taxation on purchases of machinery and equipment, τ_{CME}.

To indicate how we can work out the marginal cost of each type of tax increase, we show how to do this for the first case above, an increase in the rate of business income taxation. Recall definitions (54) to (56) which defined each deflated wedge w_n as a function of all tax parameters. Now regard each w_n as a function of the business income tax rate,τ^*; ie we have:

(110) $w_n(\tau^*) = g_n(\tau^*)$ for $n = 1,2,3$.

Now we define the *cost of the business income tax function* $C(\tau^*)$ as the difference between the optimal value of output $S(0_3)$ and the tax distorted value of output $S(w^t)$, but where the wedges w_n^t are regarded as functions of τ^*:

(111) $C(\tau^*) \equiv S(0_3) - S[g_1(\tau^*),g_2(\tau^*),g_3(\tau^*)]$.

Now differentiate (111) with respect to τ^* and evaluate the resulting derivatives at the observed period t tax rate τ^t. Using (111) leads to the following period *t marginal cost of an increase in the business income tax rate*:

(112) $MC(\tau^t) \equiv \partial C(\tau^t)/\partial\tau$

$= -\sum_{n=1}^{3} [\partial S(w^t)/\partial w_n]\, \partial g_n(\tau^t)/\partial\tau$ using (110)

(113) $= -\sum_{n=1}^{3} \sum_{j=1}^{3} w_j^t P_j^t [\partial K_j(w^t)/\partial w_n]\, \partial g_n(\tau^t)/\partial\tau$ using (102).

Equations (109) can be substituted into (113) in order to obtain a formula for $MC(\tau^t)$ that can be evaluated empirically, given an econometrically estimated pure rents function Π^t.

We turn now to the problem of defining the corresponding marginal benefit function for our new model. We continue to focus on changes in the business income tax rate. The treatment of changes in other tax parameters is similar.

Recall equation (90) above, which expressed the total revenue T from all sources of capital taxation in terms of the distortion wedge $W = wP$. The multiple capital stock generalisation of equation (90) in the previous section is (114) below; ie we can write total capital tax revenue T^t in period t as a function of the distortion wedges $w^t = [w_1^t, w_2^t, w_3^t]$ as follows:

(114) $T^t = \sum_{n=1}^{3} w_n^t P_n^t K_n^t$.

We replace the period t vector of distortion rate w^t by a general vector of distortion rates w and let $K(w)$ be defined by (96). Finally, we regard each wedge w_n as a function of the business income tax rate τ, as in equations (110) above. Making these substitutions into (114) leads to the following definition for the *period t total tax revenues* as a function of the business income tax rate τ*:

$$(115)\quad T(\tau^*) \equiv \sum_{n=1}^{3} w_n(\tau^*)\, P_n^{\,t}\, K_n[w_1(\tau^*), w_2(\tau^*), w_3(\tau^*)].$$

We can obtain the marginal increase in capital tax revenues by differentiating $T(\tau^*)$ defined by (115) with respect to τ* and evaluating the resulting derivative at the observed period t tax rate τ^t. Thus, we define the period *t marginal benefit of an increase in the business income tax rate* as:

$$(116)\quad MB(\tau^t) \equiv T'(\tau^t)$$

$$= \sum_{n=1}^{3} [\partial g_n(\tau^t)/\partial\tau]\, P_n^{\,t}\, K_n(w^t) + \sum_{n=1}^{3} w_n^{\,t}\, P_n^{\,t}\, [\sum_{j=1}^{3} \partial K_n(w^t)/\partial w_j]\, \partial g_j(\tau^t)/\partial\tau$$

$$= \sum_{n=1}^{3} K_n(w^t) P_n^{\,t}\, \partial g_n(\tau^t)/\partial\tau]$$

$$+ \sum_{n=1}^{3} \sum_{j=1}^{3} w_n^{\,t}\, P_n^{\,t}\, [\partial^2\, \Pi^{\,t}(p^t, (u^t + w^t)P^t, F^t)/\partial V_n \partial V_j]\, \partial g_j(\tau^t)/\partial\tau$$

where the last equality follows using (98).

Thus, given an econometric estimate for the period t cash flow function, $\Pi^{\,t}$, the right hand side of (116) can be evaluated using observable data. As was the case with our earlier formulae, (74) and (93), the first term on the right hand side of (116) will be positive but we can no longer guarantee that the second term will be negative. If there is a high degree of substitutability of capital for other inputs and outputs in period t, then as in the previous sections, the second term can make the overall marginal benefits of increasing capital taxes negative. In this case, the government will achieve both increased productive efficiency and higher tax revenues by reducing capital tax distortions.

The *marginal excess burden of increasing the period t business income tax rate* τ^t by a small amount is simply the ratio of the marginal cost $MC(\tau^t)$ defined by (113) above divided by the marginal tax revenue $MB(\tau^t)$ defined by (116) above:

$$(117)\quad MEB(\tau^t) \equiv MC(\tau^t) / MB(\tau^t).$$

The calculations of marginal excess burdens for the other tax parameters is similar.

We turn now to our econometric model.

7. The Producer Model

As we saw in the previous section, the key determinants of the size of the deadweight loss or loss of efficiency in the economy due to the taxation of capital are the size of the capital tax distortion wedges and the magnitudes of various elasticities of demand and supply for private sector producers. This section and the next one will focus on the empirical estimation of these producer elasticities.

In this paper, we use the data pertaining to the Canadian economy that is developed in Appendix A below to estimate a system of private producer supply and demand equations. Flexible functional form techniques are used: ie the functional form we use to model the technology does not impose unwarranted *a priori* restrictions on elasticities of substitution between the outputs and inputs. In the present section, we lay out a preliminary version of our model.

In the next section, we note that there is a potential problem with our preliminary model: the elasticities that it generates may trend significantly over the sample period in a manner that is not warranted. Given the importance of determining accurate elasticities in order to calculate excess burdens, we discuss how this problem can be remedied.

In the data appendix, we describe our disaggregated database covering the private production sector of the Canadian economy for the period 1974 to 1998. Our disaggregated producer model database contains price and quantity information for a total of 39 inputs and outputs. These were aggregated to 8 commodities to permit econometric estimation.

When the number of commodities in an applied general equilibrium model is large, it becomes difficult or impossible to estimate flexible functional forms. When there are $N+1$ commodities and T observations on prices and quantities for each commodity, there are $(N+1)T$ degrees of freedom available for econometric estimation and this number is an upper bound to the number of unknown parameters characterising technology that can be estimated. A bare bones basic flexible functional form for a production function (or the dual cost or profit functions) in the constant returns to scale case has $N(N+1)/2$ unknown parameters. Hence, as soon as N (the number of commodities less one) is equal to or greater than $2T$, it becomes impossible to estimate a flexible functional form using time series data.

Given that we could not accurately estimate a flexible producer model with all 39 commodities in our database, in this preliminary work, we aggregated the 39 commodities to obtain the following 8 commodities: (1) consumption plus government purchases of intermediate inputs plus investment; (2) exports; (3) imports; (4) labour input; (5) nonresidential stocks; (6) stocks of machinery and equipment; (7) inventory stocks and (8) inputs of land and other fixed factors.

Flexibility is a desirable property for a functional form since an inflexible functional form will restrict elasticities of substitution between commodities in some arbitrary *a priori* fashion. A way of dealing with this inflexibility problem when the number of commodities is large relative to the number of observations was suggested by Diewert and Wales (1988) in the consumer theory context: instead of estimating a general N by N symmetric substitution matrix A of full rank, they restricted A to be a symmetric substitution matrix of rank J where J is smaller than N. Diewert and Wales (1988) termed functional forms of this type *semiflexible*. In the present section, we shall adapt their technique to the producer context.

The technology of the private production sector could be described by a production, cost or variable profit function. In this study, we will describe technology by means of a pure profits variable profit function of the type defined in the previous section.[17]

Recall the definition of the period *t cash flow profit function* π^t, (57) above, which we rewrite as (118) below:

(118) $\pi^t(p^t, K^t, F^t) \equiv \max_y \{p^t \bullet y : (y, K^t, F^t) \in S^t\}$

where $(p_1^t,\ldots, p_{11}^t) \equiv p^t$ is the vector of positive prices that producers face in period t for the 4 noncapital variable inputs and outputs in our model and where $p^t \bullet y$ denotes the inner product of the of the vectors p^t and y. The corresponding variable outputs and inputs produced and used during period t are denoted by the quantity vector $y^t \equiv (y_1^t,\ldots y_4^t)$. Recall that if commodity m is an input, then y_m^t has a negative sign. The private business sector of the economy utilises the beginning of period t capital stock vector $K^t \equiv (K_1^t, K_2^t, K_3^t)$ and the fixed factor input F^t. The period t set of feasible net output vectors y, conditional on a beginning of the period capital stock K^t and fixed factor input F^t is denoted by the set S^t. In words, $\pi^t(p^t, K^t, F^t)$ is the maximum value added less the value of labour inputs that the private sector can produce given that producers face the prices p^t for these variable outputs and inputs and given that producers have the vector of

[17] In this approach, we treat the user costs of our three types of reproducible capital as exogenous variables and the corresponding capital input demands are treated as endogenous variables. Thus, in our present econometric approach, the demand for capital is treated in a symmetric manner with the demand for labour. In contrast, in our earlier study of the New Zealand economy, Diewert and Lawrence (1994), we treated the stocks of reproducible capital as exogenous variables and the corresponding rental prices as endogenous variables. We feel that our present approach is more appropriate in the context of determining the excess burdens of capital taxation, a topic that our earlier study did not address.

fixed stocks of reproducible capital K^t and the quantity F^t of non-reproducible capital available to them at the beginning of period t.

As in the previous sections, we use the period t cash flow profit function π^t defined by (118) above in order to define the period *t pure rent profit function* Π^t as follows:

$$(119)\quad \Pi^t(p^t, (u^t + w^t)P^t, F^t) \equiv \max_K \{\pi^t(p^t, K^t, F^t) - \sum_{n=1}^{3} (u_n^t + w_n^t)P_n^t K_n\}$$

where the price of one unit of the ith type of reproducible capital stock is P_i^t in period t and $u^t \equiv (u_1^t, u_2, u_3^t)$ is the vector of period t (deflated) undistorted user costs of capital defined by (95) in the previous section and $w^t \equiv (w_1^t, w_2^t, w_3^t)$ is the period t vector of (deflated) total tax distortion wedges defined by (54)-(56) in section 3.

In our econometric work, we hold the input of land and other fixed factors that are used by the Canadian private production sector fixed throughout our sample period. Hence, in what follows, we will omit F^t from $\Pi^t(p^t, (u^t + w^t)P^t, F^t)$. We will also absorb the three user costs of capital, $(u^t + w^t)P^t$, into the p^t vector; ie we define $p_5^t \equiv (u_1^t + w_1^t)P_1^t$, $p_6^t \equiv (u_2^t + w_2^t)P_2^t$ and $p_7^t \equiv (u_3^t + w_3^t)P_3^t$. Finally, the notation $\Pi^t(p)$ indicates that the pure rents profit function depends on the period t as well as on the price vector $p \equiv (p_1, p_2,\ldots,p_7)$. We will rewrite this dependence as $\Pi(p, t)$.

Once a functional form for Π has been chosen, estimating equations can be obtained by differentiating the profit function with respect to the prices p_m, see Diewert (1974a; 137 and 140), (1993; 166 and 168):

$$(120)\quad y_m(p, t) \equiv \partial\Pi(p, t)/\partial p_m\ ; \qquad m = 1,\ldots,7.$$

The functional form for the pure rents function Π that we chose was a variant of the normalised quadratic functional form,[18] since this functional form allows us to impose the appropriate curvature conditions without destroying its flexibility properties. Using matrix notation, the function can be defined as follows:

$$(121)\quad \Pi(p, t) \equiv p\bullet b + p\bullet d(t - 1) + (1/2)\, p\bullet Ap / p\bullet g\ ; \qquad t = 1, 2,\ldots, 25$$

where $b \equiv [b_1,\ldots,b_7]$ and $d \equiv [d_1,\ldots, d_7]$ are parameter vectors to be estimated and $A \equiv [a_{mn}]$ is a 7 by 7 symmetric matrix of parameters to be estimated. The vector $g \equiv [g_1, \ldots, g_7]$ is a vector of exogenously determined parameters. The components of g were chosen to be the absolute values of the sample means of the observed net output vectors $y^t \equiv [y_1^t,\ldots,y_7^t]$ normalised so that:

$$(122)\quad p^*\cdot g = 1$$

[18] See Diewert and Wales (1987) (1992) and Lawrence (1988) (1989) (1990).

where p^* was a fixed vector.[19] The variable t which appears in (121) is a scalar time variable which serves as a proxy for technological change.

In order for $\prod(p,t)$ to be a well behaved profit function which is a convex function in its price variables p, we set A to equal the following product of two matrices, U and its transpose U^T :[20]

(123) $A = UU^T$

where U is a lower triangular matrix and U^T is an upper triangular matrix which satisfies the following restrictions[21]:

(124) $U^T p^* = 0_7$

where 0_7 is a vector of zeros of dimension 7.

Differentiating the profit function (121) with respect to the components of p leads to the following system of 7 estimating equations:

(125) $y^t = b + d(t-1) + Ap^t / p^t \cdot g - (1/2)\, p^t \cdot Ap^t\, g / (p^t \cdot g)^2 + e^t$; $\quad t = 1,2,\ldots,25$

where $e^t \equiv [e_1{}^t,\ldots,e_7{}^t]$ is a vector of independently distributed normal residuals where each of the residuals $e_m{}^t$ has mean 0 and variance $\sigma_m{}^2$ for $m = 1,\ldots,7$ and $t = 1,\ldots,25$.

The vector of parameters d essentially adds a linear trend to each estimating equation in order to allow for the effects of technical progress in the Canadian economy over our sample period.

Unfortunately, (123) and (125) did not represent our final model because there is a problem with the profit function defined by (121). The problem is that the elasticities of demand and supply derived from the profit function defined by (121) can have substantial trends built into them. This is a problem in the present context due to the importance of elasticities in determining marginal excess burdens. We deal with this problem in the following section.

8. The Problem of Trending Elasticities

If we differentiate the pure rents profit function defined by (121) above with respect to the mth component of the price vector p, we obtain the following equation that describes the net supply of commodity m as a function of the price vector p in period t:

[19] We chose p* to be a vector of ones.
[20] See Diewert and Wales (1987; 52-53) for further explanation.
[21] Restrictions like (7) are required in order to identify the components of the b vector. Alternatively, restrictions (7) could be dropped but then the b vector would have to be dropped as well.

(128) $y_m(p,t) = b_m + d_m\,(t-1) + \sum_{j=1}^{7} a_{mj}\,(p_j / p\bullet g) - (1/2)\, g_m\, p\cdot Ap / (p\cdot g)^2.$

Now differentiate (128) with respect to p_n, the *n*th component of the price vector *p*:

(129) $\partial y_m(p,t)/\partial p_n = a_{mn} / p\bullet g - \sum_{j=1}^{7} a_{mj}\, p_j\, g_n / (p\bullet g)^2 - \sum_{j=1}^{7} \mathrm{a}_{nj}\, \mathrm{p}_j\, \mathrm{g}_m / (p\bullet g)^2$

$$+ g_m\, g_n\, p\cdot Ap / (p\bullet g)^3 .$$

Now turn (129) into the cross elasticity of net supply of commodity m with respect to a change in the price of commodity *n*, e_{mn}:

(130) $e_{mn}\,(p,t) \equiv [p_n / y_m]\, \partial y_m(p,t)/\partial p_n$

$$= a_{mn}\,(p_n / y_m\, p\cdot g) - \sum_{j=1}^{7} a_{mj}\, p_j\, g_n\, p_n / y_m\,(p\bullet g)^2$$

$$- \sum_{j=1}^{7} a_{nj}\, p_j\, g_m\, p_n / y_m\,(p\bullet g)^2 + g_m\, g_n\, p\cdot Ap\, p_n / y_m\,(p\bullet g)^3 .$$

The last three terms on the right hand side of (130) will be zero when $p = p^*$ and in our empirical work, these last three terms were typically small in magnitude. Thus, the key determinant of the magnitude of the elasticity e_{mn} will typically be the first term on the right hand side of (130), namely, $a_{mn}\,(p_n / y_m\, p\bullet g)$. Of course, the parameter a_{mn} will be constant over time but the other terms, p_n (the price of commodity *n*), y_m (the net output of commodity *m*) and $p\bullet g \equiv \sum_{j=1}^{7} p_n\, g_n$ (a fixed basket price index of all 7 variable input and output prices) can all have substantial trends over our sample period. Thus, our chosen functional form has built in these possible trends in elasticities.

A solution to this problem is readily at hand but at a cost in terms of using up degrees of freedom. We have followed the example of most applied production function researchers and allowed technical progress to affect the constant terms in the system of net supply functions (128) but we have left the substitution matrix *A* unchanged over time. To solve the problem of trending elasticities, all we have to do is allow *A* to change over time as well. Thus, in our empirical work, we set the *A* matrix in (128) above equal to weighted average of a matrix *B* (which characterises substitution possibilities in 1974) and a matrix *C* (which characterises substitution possibilities in 1998); ie, we define *A* as follows in terms of *B* and *C* and the time variable *t*:

(131) $A^t = (1 - [(t-1)/24])\, B + [(t-1)/24]\, C\,; \qquad t = 1,2,\ldots,25.$

Essentially, we now let technical progress affect not only the constant terms in (121) but we also allow it to affect substitution possibilities as well. Another way of viewing our new functional form is that *we allow the functional form to be flexible at two points* (the first sample point and the last) *instead of the usual one point*.

In order to impose the correct curvature conditions (globally), we need to set the 7 by 7 symmetric matrices B and C equal to the product of UU^T and VV^T respectively, where U and V are lower triangular matrices; ie we set:

(132) $B = UU^T$; $C = VV^T$; U and V lower triangular.

We also impose the following normalisations on the matrices U and V:

(133) $U^T p^* = 0_7$; $V^T p^* = 0_7$.

Now we are ready to describe our empirical results.

9. Empirical Results for the Production Model

The unknown parameters which appear in (125) (where A is replaced by A^t defined by (131) above) can in theory be estimated using nonlinear systems maximum likelihood estimation commands in econometric packages such as TSP or SHAZAM (see White (1978)). However, due to the very large number of parameters in our model, these econometric programs failed to converge. Thus, we decided to run all 7 of our estimating equations in (125) as one big nonlinear regression with only one variance parameter σ^2. This approach proved to be quite successful using SHAZAM. Once we obtained satisfactory parameter estimates using this one big regression approach, we calculated the inverse of the square root of the squared residuals for each equation and then multiplied both dependent and independent variables for that equation by this variance stabilising factor and then we performed a second single big regression using these transformed dependent and independent variables. This two stage procedure controls for equation by equation variance heteroskedasticity in the original regression model.

It should be noted that equations (125) are linear in the unknown vectors of parameters, b and d, and linear in the unknown components of the matrices of parameters, B and C. However, when we impose the correct curvature conditions on our estimated profit function by setting $B = \mathrm{UU}^\mathrm{T}$ and $C = VV^T$, the resulting estimating equations (125) turn out to be nonlinear in the components of the matrices U and V. When we attempted to estimate the parameters in b, d, U and V by running one big regression, we found that it was difficult to achieve convergence if we attempted to estimate *all* of the parameters in an initial regression. Thus, we used the following strategy: (i) the parameters in the vectors b, and d were estimated in an initial linear regression (with U and V being set equal to zero matrices initially); (ii) we ran nonlinear regressions, using equations (125), introducing one column of the U matrix and one column of the V matrix into our nonlinear regression; (iii) the final parameter values from stage (ii) above were used as starting values in a new nonlinear regression where an additional column of *U* and *V* were added with starting values close to zero; (iv) step (iii) was repeated until all columns of the *U* and *V* matrices were entered into the big nonlinear regression, with at least one nonzero component. This algorithm lead us to introduce 5 columns of the *U* matrix and 5 columns of the *V* matrix.[22] In view of the restrictions (133), this means that we should have 20 u_{mn} parameters and 20 v_{mn} parameters in our final regression. We also have 14 b_m and d_m parameters or an additional 14 parameters to estimate. This means we have a total of 54 parameters to estimate with 7 times 25 or 175 degrees of freedom. This is a large number of parameters for the available degrees of freedom but as we shall see, most of them appear to be necessary to describe substitution elasticities for the Candadian economy over our sample time period.

Table 6: Estimated Coefficients for the Producer Model

Variable	*Estimate*	*St. Error*	*t-statistic*	*Variable*	*Estimate*	*St. Error*	*t-statistic*
b_1	124.820	1.958	63.767	u_{22}	-3.648	1.669	-2.186
b_2	34.037	1.193	28.530	v_{32}	-13.216	1.581	-8.358
b_3	-37.804	1.705	-22.175	u_{32}	8.460	1.382	6.122
b_4	-83.097	1.164	-71.363	v_{42}	2.776	0.775	3.584
b_5	-15.242	0.660	-23.099	u_{42}	-7.975	1.095	-7.286
b_6	-10.829	0.571	-18.955	v_{52}	-1.916	0.754	-2.539
b_7	-2.540	0.249	-10.193	u_{52}	0.085	1.134	0.075
d_1	7.008	0.463	15.129	v_{62}	-0.887	0.793	-1.119
d_2	2.088	0.190	10.965	u_{62}	1.953	0.909	2.150
d_3	-2.513	0.285	-8.820	v_{33}	-4.822	1.055	-4.571
d_4	-3.666	0.287	-12.764	u_{33}	-1.261	1.517	-0.831
d_5	-1.184	0.106	-11.199	v_{43}	6.680	0.865	7.721
d_6	-0.806	0.125	-6.435	u_{43}	3.041	1.308	2.325
d_7	0.010	0.034	0.301	v_{53}	0.947	0.909	1.042
v_{11}	23.267	2.496	9.323	u_{53}	-2.330	0.799	-2.916
u_{11}	14.383	1.745	8.243	v_{63}	-2.560	1.064	-2.407
v_{21}	-10.447	0.958	-10.899	u_{63}	0.042	1.074	0.039
u_{21}	8.355	1.142	7.314	v_{44}	-0.078	1.566	-0.050
v_{31}	-5.441	1.819	-2.991	u_{44}	-0.791	1.956	-0.404
u_{31}	-8.817	1.600	-5.512	v_{54}	-3.804	0.855	-4.447
v_{41}	-9.405	0.863	-10.892	u_{54}	1.306	1.206	1.082
u_{41}	-12.194	1.157	-10.544	v_{64}	4.203	1.188	3.538
v_{51}	-0.503	0.946	-0.532	u_{64}	-0.522	1.221	-0.427
u_{51}	-3.590	0.815	-4.403	v_{55}	-0.170	2.441	-0.070
v_{61}	1.979	0.848	2.333	u_{55}	-0.279	1.204	-0.231
u_{61}	1.548	0.606	2.556	v_{65}	0.179	2.630	0.068
v_{22}	13.209	1.258	10.497	u_{65}	0.546	1.105	0.494

[22] Adding an extra *U* or *V* column led to a negligible increase in the log likelihood after 4 columns were added so we stopped at 5 columns for the *U* and *V* matrices.

The one big nonlinear regression with 54 parameters generated the following 7 variance weighting factors (the inverses of the square roots of the sum of the squared residuals for each of our 14 estimating equations): 0.04474143; 0.09917493; 0.06528360; 0.1064402; 0.1990786; 0.2345888; 0.5263991. The resulting parameter estiamtes are presented in Table 6.

It should be noted that the elements in the last row of the lower triangular U matrix are defined in terms of the other elements of the U matrix as follows:

(134) $u_{7,\mathrm{m}} \equiv - \sum_{\mathrm{j}=1}^{6} u_{jm}$; $m = 1,\ldots,5.$

Similarly, the elements in the last row of the lower triangular V matrix are defined in terms of the other elements of the V matrix as follows:

(135) $v_{7,m} \equiv - \sum_{\mathrm{j}=1}^{6} v_{\mathrm{j}m}$; $m = 1,\ldots,5.$

Of the 40 price substitution coefficients u_{mn} and v_{mn}, 26 had t statistics greater than two.

Table 7: Own Price Elasticities of Net Supply for Canada

Year	e_{11}	e_{22}	e_{33}	e_{44}	e_{55}	e_{66}	e_{77}
1974	1.657	2.442	-3.991	-2.674	-1.319	-0.626	-0.668
1975	1.724	2.048	-3.825	-2.489	-1.221	-0.518	-1.104
1976	1.714	1.430	-2.854	-2.429	-0.990	-0.440	-1.163
1977	1.811	1.116	-2.696	-2.380	-0.833	-0.397	-0.848
1978	1.766	0.754	-2.340	-2.081	-0.673	-0.401	-0.559
1979	1.703	0.517	-1.974	-1.850	-0.564	-0.390	-0.436
1980	1.764	0.374	-1.708	-1.768	-0.549	-0.321	-0.456
1981	1.882	0.260	-1.466	-1.754	-0.540	-0.275	-0.553
1982	2.006	0.202	-1.270	-1.804	-0.468	-0.251	-0.528
1983	1.954	0.189	-1.056	-1.711	-0.372	-0.246	-0.520
1984	2.006	0.224	-1.012	-1.640	-0.321	-0.243	-0.480
1985	2.088	0.293	-0.992	-1.613	-0.278	-0.236	-0.421
1986	2.107	0.383	-0.982	-1.545	-0.251	-0.240	-0.374
1987	2.089	0.497	-0.931	-1.486	-0.240	-0.238	-0.354
1988	2.115	0.612	-0.903	-1.467	-0.265	-0.238	-0.349
1989	2.112	0.736	-0.891	-1.420	-0.307	-0.241	-0.369
1990	2.135	0.872	-0.928	-1.396	-0.338	-0.239	-0.344
1991	2.230	0.982	-0.944	-1.461	-0.335	-0.207	-0.273
1992	2.422	1.076	-1.032	-1.510	-0.311	-0.186	-0.184
1993	2.445	1.216	-1.134	-1.452	-0.309	-0.191	-0.143
1994	2.431	1.381	-1.258	-1.363	-0.346	-0.212	-0.153
1995	2.467	1.466	-1.302	-1.333	-0.392	-0.223	-0.218
1996	2.372	1.569	-1.278	-1.316	-0.425	-0.235	-0.287
1997	2.403	1.633	-1.309	-1.358	-0.463	-0.250	-0.280
1998	2.412	1.764	-1.421	-1.359	-0.493	-0.266	-0.274
Average	2.073	0.961	-1.580	-1.706	-0.504	-0.292	-0.454

Recall definition (130) above, which defined the cross elasticity of net supply of commodity *m* with respect to a change in the price of commodity *n*, e_{mn}. There are too many elasticities in the full 7 by 7 matrix of elasticities for us to list them all but we do list the own price elasticities of net supply e_{mm} in Table 7.

It can be seen that many of the elasticities are fairly large, especially near the beginning of the sample period. This may be partially due to the rather large number of parameters in our model. However, it is likely that the large elasticities are simply due to the fact that we are using a very flexible functional form and we have disaggregated inputs and outputs to a greater degree than many previous econometric studies. What is very surprising to us is the generally low elasticities we get for the three capital stock components: the average own elasticity of demand for nonresidential structures averaged around −.5; the price elasticity of own demand for machinery and equipment averaged only about −.3 and the price elasticity of own demand for inventory stocks averaged only about −.45 (which is perhaps not surprising).

Consistent with economic theory[23], we suspect that a more disaggregated model would yield bigger (in magnitude) elasticities. This is what we found in our studies of the New Zealand economy; moving from a highly aggregated model to one that is relatively disaggregated has increased the scope for substitution and, consequently, led to substantially larger elasticity estimates. Thus, it is likely that further disaggregation would lead to even *higher* elasticities of demand for capital and this would feed into *higher* estimates of deadweight losses and marginal excess burdens of capital taxation.

Our observation that disaggregation tends to lead to larger estimate of elasticities of supply and demand is one that has not been stressed in the literature a great deal. However, given the importance of elasticity information for a wide variety of policy purposes, we believe that the point is an important one and deserves further research.[24]

We have now assembled all the necessary building blocks for the construction of marginal excess burdens for capital taxation. In the next section, we present our marginal excess burden estimates.

[23] See Diewert (1974b).

[24] One of the last serious discussions about the likely size of elasticities took place 50 years ago in the context of trade elasticities by Orcutt (1950) who argued that elasticities of import demand and export supply were likely to be larger than had been thought. We found that in our work on the New Zealand economy, trade elasticities in a 15 commodity model dropped substantially when we aggregated our two export commodities into a single export aggregate and when we aggregated our three import commodities into a single import aggregate.

10. Marginal Excess Burdens of Capital Taxation for Canada

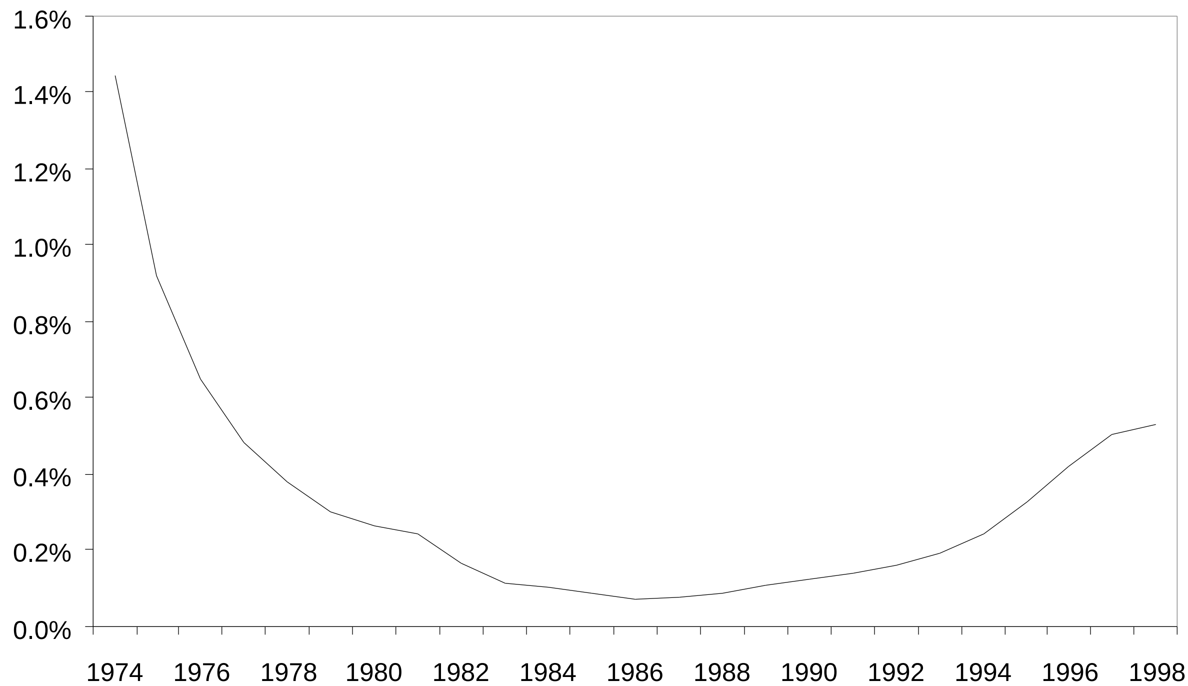
As outlined in section 6 (see formula (105) above), we calculate a second order approximation to the total loss of output that results from the taxation of capital due to both the personal and business income tax, the property tax on structures and sales taxes on the purchases of durable capital equipment. This second order approach is the average of two first order approximations: one around the distorted equilibrium and one around the undistorted equilibrium. These total losses are reported as a fraction of business sector GDP in Table 8 and Figure 7 below.

Table 8: Production Loss as a Proportion of Business Sector Output (percentages)

Year	*Loss*	*Year*	*Loss*	*Year*	*Loss*
1974	1.441	1983	0.115	1992	0.163
1975	0.918	1984	0.103	1993	0.190
1976	0.650	1985	0.086	1994	0.245
1977	0.483	1986	0.073	1995	0.328
1978	0.378	1987	0.075	1996	0.420
1979	0.303	1988	0.090	1997	0.503
1980	0.266	1989	0.111	1998	0.530
1981	0.243	1990	0.125		
1982	0.166	1991	0.140	Average	0.326

After starting from a relatively high proportion of around 1.4 per cent of GDP in 1974, the production loss from capital taxation progressively declines to very low levels of less than 0.1 per cent of GDP in the period 1985-1988. It then increases to finish at around 0.5 per cent of GDP in 1998. We know that these productive efficiency losses grow roughly proportionally to the magnitude of elasticities and increase at a squared rate as the distortion wedges increase. From Table 3, it can be seen that the wedges did not change all that much over the entire sample period. From Table 7, it can be seen that capital elasticities started out at relatively high levels in 1974, generally decreased until they hit their minimum magnitudes in the period 1985-1988 and then these capital elasticities gradually increased again. Thus the pattern of efficiency losses was more or less driven by these fluctuations in capital elasticities rather than by large fluctuations in the burden of capital taxation (i.e., by fluctuations in the wedge rates).

Figure 7: Production Loss as a Proportion of Business Sector Output



As noted in sections 2 and 6 (see formula (117) above), the marginal excess burden of a tax parameter is the loss of output due to a marginal increase in the tax parameter divided by the increase in total tax revenue due to the same marginal change in the tax parameter. In Table 9 and Figure 8 below, we list the MEB for an increase in the business capital tax rate τ^*, the MEB for an increase in the property tax on structures τ_{PNR} (which is equal to the MEB for an increase in the sales tax on structures and materials used on structures, τ_{CNR}) and the MEB for an increase in the sales tax on machinery and equipment τ_{CME}. In all cases, the increase in the tax rate increased tax revenues.

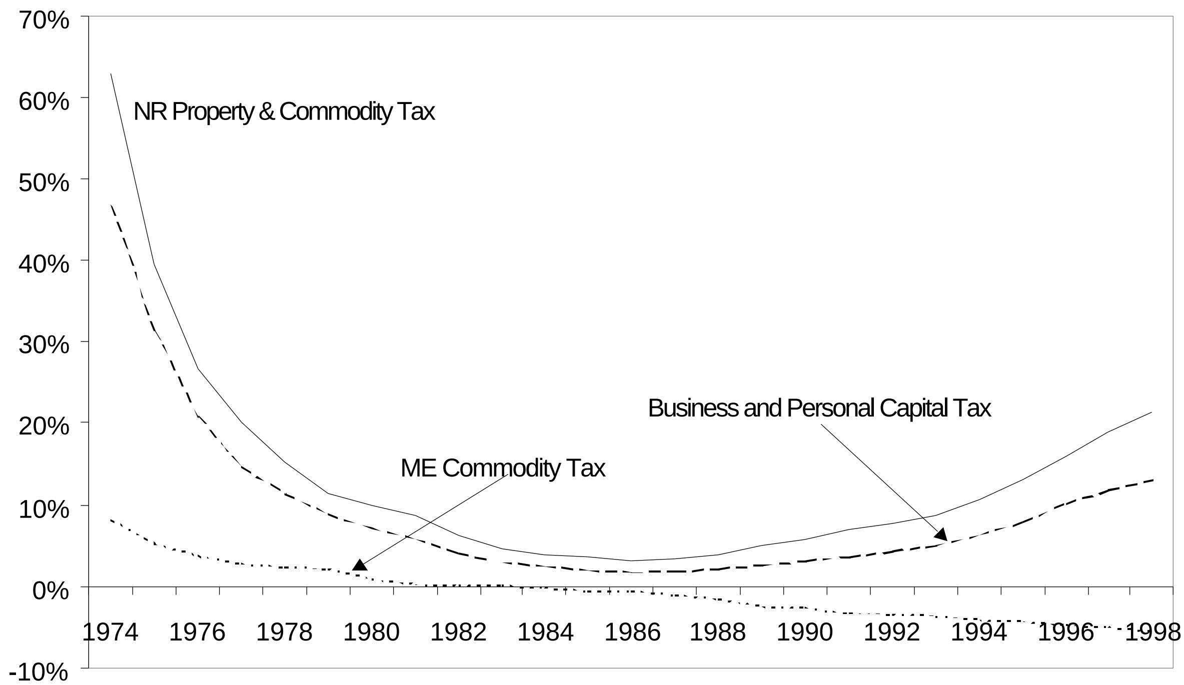
In the case of an increase in the business capital tax rate τ^* or an increase in the property tax rate τ_{PNR}, for all years marginal efficiency declined so the resulting MEBs were positive as we would expect. However, in the case of an increase in the sales tax on machinery and equipment, we found that from 1984 on, this marginal increase actually led to a tiny *increase* in efficiency and thus the resulting MEB's are negative. Thus, according to our model, a small increase in τ_{CME} should lead to a small increase in both tax revenues and in economic efficiency from 1984 on. Economic theory does not rule out this rather strange result but it is difficult to explain intuitively.

Table 9: Marginal Excess Burdens of Various Capital Taxes (percentages)

Year	τ^*	τ_{PNR}	τ_{CME}
1974	46.72	62.93	8.20
1975	31.39	39.49	5.29
1976	20.81	26.57	3.76
1977	14.79	20.24	2.79
1978	11.43	15.25	2.41
1979	8.87	11.26	2.19
1980	7.22	9.93	0.95
1981	5.94	8.65	0.32
1982	4.09	6.29	0.23
1983	3.00	4.55	0.18
1984	2.49	3.96	-0.10
1985	2.04	3.58	-0.51
1986	1.79	3.05	-0.49
1987	1.88	3.31	-1.00
1988	2.20	3.89	-1.46
1989	2.66	4.96	-2.35
1990	3.14	5.71	-2.54
1991	3.70	6.90	-3.27
1992	4.31	7.80	-3.36
1993	5.11	8.80	-3.59
1994	6.39	10.68	-4.05
1995	7.99	13.02	-4.27
1996	10.24	15.97	-4.68
1997	11.88	18.98	-4.97
1998	13.11	21.29	-5.46
Average	9.33	13.48	-0.63

From Table 9, we see that on average the Marginal Excess Burden of a small increase in the sales tax rate on machinery and equipment is negligible. For an increase in the business tax rate, the average MEB is about 9% while a small increase in the property tax rate yields on average an MEB of about 13.5%. These MEB's are much larger for the beginning of our sample period (about 47% for the business tax rate and about 63% for the property tax rate) and they seem to be increasing at the end of our sample period (about 13% for the business tax rate and about 20% for the property tax rate on structures).

Figure 8: Capital Tax Marginal Excess Burdens



11. Conclusions

We conclude with the following observations:

- It is a good idea to try and reduce capital tax distortions because they always involve a loss of productive efficiency. The loss of revenue has to be made up by taxing consumption or labour but with enough tax instruments at its disposal, the tax authority can always design a tax reform strategy that will increase overall welfare.
- Our estimates of the burdens of capital taxation are probably underestimated due to our use of average tax rates. In the real world, the complexities of the tax code lead to a much more dispersed pattern of burdens but since the losses are approximately proportional to the squares of tax distortions, averaging tax distortions will lead to an underestimate of the true efficiency losses.
- Our estimates of the burdens of capital taxation are probably underestimated due to the relatively high degree of aggregation in our model. There is theoretical and empirical evidence that elasticities of substitution increase in magnitude as we disaggregate over

commodities. These higher elasticity estimates would generate proportionately higher estimates of total and marginal efficiency losses.

It may be useful to spell out in more detail why it is not efficient to have a system of capital taxation that generates nonzero distortion wedges. The basic intuition behind the above algebra is explained rather well by Judd (1999):

"One general problem with this literature [on capital taxation] is the lack of economic intuition. … In this paper, we ignore simple dynamic features such as the steady state behavior or long run elasticities, and instead put the zero long run tax results on more economically appealing foundations. To do this, we look to the commodity tax literature. Two results from that literature apply here; first, the optimality of uniform taxation with separable and sufficiently symmetric utility, and, second the prohibition of intermediate good taxation derived in Diamond and Mirrlees (1971). Our methods generalize revious work and tie the results to the commodity tax literature, a change which helps us understand why we often find that the average tax rate on capital income is zero in the optimal policy." Kenneth L. Judd (1999; 2).

It is the second result from optimal tax theory, the prohibition against taxing intermediate inputs in production, that explains our results. Judd goes on to elaborate on this point:

"The second key principle we invoke is the Diamond-Mirrlees argument against the taxation of intermediate goods. This is relevant here since capital goods, physical and human, are intermediate goods. In fact, income taxation is equivalent to sales taxation of intermediate goods. This can be seen by noting, for example, that a 100 % sales tax on capital equipment is equivalent to a 50 % tax on the income flow from capital equipment. Since intermediate good taxation will generally put an economy on the interior of its production possibilities frontier, capital income taxation is likely to produce similar factor distortions, particularly if there are many capital goods. Therefore, an optimal tax structure would tax only final goods". Kenneth L. Judd (1999; 5-6).

Thus a reproducible capital stock component is both *produced* by the production sector (or imported at a fixed world price and thus is produced by an integrated world production sector) and *used* as an input in later periods; ie, it is an *intertemporal intermediate input*. Hence in order for an economy to achieve productive efficiency, it is necessary that all users and producers of an intermediate commodity face the *same* prices.[25] However, the system of business income taxation causes users and producers of reproducible capital to face *different* (intertemporal) prices. Diewert (1988) made the same point as Judd:

"The other major thrust of this paper will be to indicate four major areas where our present tax system is inefficient. Thus in the second, third and fourth sections below, we discuss three different types of deadweight loss induced by our present system of business taxation. In the second section, we discuss the losses due to the fact that the tax

[25] See Diewert (1983b).

system does not treat (nondurable) inputs and outputs in an even handed manner; that is, there are tariffs and sales taxes that fall within the business sector (the manufacturer's sales tax) as well as various output subsidies. In the third section, we discuss the losses due to the uneven tax treatment of durable inputs, such as land, inventories and various types of capital." W. Erwin Diewert (1988; 2).

Thus Diewert noted that both intermediate input taxation and the taxation of reproducible capital inputs led to a loss of productive efficiency. The third type of deadweight loss that leads to a global loss of productive efficiency is transfer pricing. Diewert went on to characterize these three types of loss of productive efficiency as follows:

"The above three types of deadweight loss lead to both a loss of productive efficiency as well as a loss of overall efficiency defined earlier. A tax system is consistent with productive efficiency if the allocation of resources across the entire business sector is such that no output can be increased, holding other aggregate outputs and inputs fixed." W. Erwin Diewert (1988; 2-3).

"In summary: in order to achieve productive efficiency, it is necessary that all producers in the economy face the same relative prices for their outputs and variable inputs." W. Erwin Diewert (1988; 6).

Tax systems that lead to a loss of productive efficiency can always be redesigned so that the inefficiencies are eliminated, the same tax revenues are collected and the utilities of at least some households increase, provided that the government has a sufficient number of tax instruments at its disposal.

REFERENCES

Browning, E.K. (1976), "On the Marginal Welfare Cost of Taxation", *American Economic Review* 71, 11–23.

Browning, E.K. (1987), "The Marginal Cost of Public Funds", *Journal of Political Economy 84*, 283–298.

Christensen, L.R. and D.W. Jorgenson (1969), "The Measurement of U.S. Real Capital Input, 1929–1967," *Review of Income and Wealth* 15(4): 293–320.

Christensen, L.R. and D.W. Jorgenson (1970), "U.S. Real Product and Real Factor Input, 1929–1967," *Review of Income and Wealth* 16(1): 19–50.

Church, A. H. (1901), "The Proper Distribution of Establishment Charges, Part III", *The Engineering Magazine* 21, 904-912.

Diamond, P.A. and J.A. Mirrlees (1971), "Optimal Taxation and Public Production", *American Economic Review* 61, 8–27 and 261–278.

Diewert, W.E. (1974a), “A Note on Aggregation and Elasticities of Substitution”, *Canadian Journal of Economics* 7, 12-20.

Diewert, W.E. (1974b), “Applications of Duality Theory”, pp. 106-171 in *Frontiers of Quantitative Economics*, Vol. II, M.D. Intrilligator and D.A. Kendrick (eds.), Amsterdam: North-Holland Publishing Co.

Diewert, W.E. (1976), “Exact and Superlative Index Numbers”, *Journal of Econometrics* 4, 115–145.

Diewert, W.E. (1980), “Aggregation Problems in the Measurement of Capital”, pp. 433–528 in *The Measurement of Capital*, Dan Usher (ed.), University of Chicago Press, Chicago.

Diewert, W.E. (1983a), “Cost-Benefit Analysis and Project Evaluation”, *Journal of Public Economics* 22, 265–302.

Diewert, W.E. (1983b), “The Measurement of Waste Within the Production Sector of an Open Economy”, *Scandinavian Journal of Economics* 85, 159–179

Diewert, W.E. (1988), “On Tax Reform,” *The Canadian Journal of Economics* 21, 1-40.

Diewert, W.E. (1993), “Duality Approaches to Microeconomic Theory”, pp. 105-175 in *Essays in Index Number Theory*, Vol. 1, W.E. Diewert and A.O. Nakamura (eds.), Amsterdam: North-Holland Publishing Co.

Diewert, W.E. and A.M. Smith (1994), “Productivity Measurement for a Distribution Firm”, *Journal of Productivity Analysis* 5, 335–347.

Diewert, W.E. and D.A. Lawrence (1994), *The Marginal Costs of Taxation in New Zealand*, New Zealand Business Round Table.

Diewert, W.E. and D.A. Lawrence (1996), “The Deadweight Costs of Taxation in New Zealand”, *Canadian Journal of Economics* 29, Special Issue, S658-S673.

Diewert, W.E. and D. Lawrence (1999), “Progress in Measuring the Price and Quantity of Capital”, Discussion Paper 99–17, Department of Economics, University of British Columbia, Vancouver, Canada, V6T 1Z1.

Diewert, W.E. and T.J. Wales (1987), “Flexible Functional Forms and Global Curvature Conditions”, *Econometrica* 55, 43-68.

Diewert, W.E. and T.J. Wales (1988), “A Normalized Quadratic Semiflexible Functional Form”, *Journal of Econometrics* 37, 327-342.

Diewert, W.E. and T.J. Wales (1992), “Quadratic Spline Models for Producer’s Supply and Demand Functions”, *International Economic Review* 33, 704-722.

Findlay, C.C. and R.L. Jones (1982), “The Marginal Cost of Australian Income Taxation”, *Economic Record* 58, 253–262.

Fisher, I. (1896), *Appreciation and Interest*, New York: Macmillan.

Griliches, Z. (1963), "Capital Stock in Investment Functions: Some Problems of Concept and Measurement", in *Measurement in Economics*, C. Christ et al (eds.), Stanford: Stanford University Press, reprinted as pp. 123–143 in Z. Griliches (1988), *Technology, Education and Productivity*, New York: Basil Blackwell.

Hall, R.E. and D.W. Jorgenson, (1967), "Tax Policy and Investment Behavior", *American Economic Review* 57, 391–414.

Harper, M.J., E.R. Berndt and D.O. Wood (1989), "Rates of Return and Capital Aggregation Using Alternative Rental Prices", pp. 331–372 *in Technology and Capital Formation*, D.W. Jorgenson and R. Landau (eds.), The MIT Press, Cambridge, MA.

Hulten, C.R. (1990), "The Measurement of Capital", pp. 119–158 *in Fifty Years of Economic Measurement*, E.R. Berndt and J.D. Triplett (eds.), Chicago: the University of Chicago Press.

Hulten, C.R. (1996), "Capital and Wealth in the Revised SNA", pp. 149–181 *in The New System of National Accounts*, J.W. Kendrick (ed.), Boston: Kluwer Academic Publishers.

Hulten, C.R. and F.C. Wykoff (1981a), "The Estimation of Economic Depreciation Using Vintage *Asset Prices: An Application of the Box-Cox Power Transformation"*, Journal of Econometrics 15, 367–396.

Hulten, C.R. and F.C. Wykoff (1981b), "The Measurement of Economic Depreciation", pp. 81–125 in *Depreciation, Inflation and the Taxation of Income from Capital"*, C.R. Hulten (ed.), Washington D.C.: The Urban Institute Press.

Hulten, C.R. and F.C. Wykoff (1996), "Issues in the Measurement of Economic Depreciation Introductory Remarks", *Economic Inquiry* 34, 10–23.

Jorgenson, D.W. (1963), "Capital Theory and Investment Behavior", *American Economic Review* 53, 366–378.

Jorgenson, D.W. (1973), "The Economic Theory of Replacement and Depreciation", pp. 189–221 in *Econometrics and Economic Theory*, W. Sellekaerts (ed.), New York: Macmillan.

Jorgenson, D.W. (1989), "Capital as a Factor of Production", pp. 1–35 *in Technology and Capital Formation*, D.W. Jorgenson and R. Landau (eds.), Cambridge Mass.: MIT Press.

Jorgenson, D.W. (1996), "Empirical Studies of Depreciation", *Economic Inquiry* 34, 24–42.

Jorgenson, D.W. and Z. Griliches (1967), "The Explanation of Productivity Change", *Review of Economic Studies* 34, 249–283.

Jorgenson, D.W. and Z. Griliches (1972), "Issues in Growth Accounting: A Reply to Edward F. Denison", *Survey of Current Business* 52:5 part 2 (May), 65–94.

Judd, D.L. (1999), "Optimal Taxation and Spending in General Competitive Growth Models", *Journal of Public Economics* 71, 1–26.

Kohli, U.R. (1978), "A Gross National Product Function and the Derived Demand for Imports and Supply of Exports", *Canadian Journal of Economics* 11, 167-82.

Lawrence, D.A. (1988), "Recent Developments in Applying Duality Theory", *Review of Marketing and Agricultural Economics* 56(3), 375-82.

Lawrence, D.A. (1989), "An Aggregator Model of Canadian Export Supply and Import Demand Responsiveness", *Canadian Journal of Economics* 22(3), 503-21.

Lawrence, D.A. (1990), "An Adjustments Costs Model of Export Supply and Import Demand", *Journal of Econometrics* 46, 381-98.

Romer, P.M. (1987), "Crazy Explanations for the Productivity Slowdown," pp. 163-202 in *NBER Macroeconomics Annual* 2, Stanley Fischer (ed.), Cambridge: MIT Press.

Triplett, J.E. (1996), "Depreciation in Production Analysis and in Income and Wealth Accounts: Resolution of an Old Debate", *Economic Inquiry* 34, 93–115.

Walras, L. (1954), Elements of Pure Economics, translated by W. Jaffe (first published in 1874), London: George Allen and Unwin.

White, K.J. (1978), "A General Computer Program for Econometric Methods - SHAZAM", *Econometrica* 46, 239-240.