



How to Use the Working Paper Series

John Creedy

WORKING PAPER 0(/2013
June 2013

The Working Papers in Public Finance series is published by the Victoria Business School to disseminate initial research on public finance topics, from economists, accountants, finance, law and tax specialists, to a wider audience. Any opinions and views expressed in these papers are those of the author(s). They should not be attributed to Victoria University of Wellington or the sponsors of the Chair in Public Finance.

Further enquiries to:
The Administrator
Chair in Public Finance
Victoria University of Wellington
PO Box 600
Wellington 6041
New Zealand

Phone: +64-4-463-9656
Email: cpf-info@vuw.ac.nz

Papers in the series can be downloaded from the following website:
<http://www.victoria.ac.nz/cpf/working-papers>

The Elasticity of Taxable Income, Welfare Changes and Optimal Tax Rates*

John Creedy[†]

Abstract

This paper provides a technical introduction to the use of the elasticity of taxable income in welfare comparisons and optimal tax discussions. It draws together, using a consistent framework and notation, a number of established results concerning marginal welfare changes and optimal taxes, in addition to presenting some new results, particularly in terms of non-marginal tax changes.

JEL Classification: H21; H24; H31

Keywords: Income taxation; Taxable income; Elasticity of taxable income; Excess burden of taxation; marginal welfare cost; optimum tax.

*In preparing this paper I have benefited from discussions with Norman Gemmill. I am very grateful to Martin Keene for pointed out an error in an earlier draft.

[†]New Zealand Treasury and Victoria University of Wellington.

1 Introduction

The concept of the elasticity of taxable income has, following Feldstein (1995), become widespread in both the positive literature on the behavioural incentive effects of income taxation and in the normative literature on welfare effects and optimal taxation.¹ This elasticity is defined as the elasticity of taxable income with respect to the net-of-tax rate (one minus the marginal tax rate), and is therefore positive. For example, much use was made of this elasticity, following the contribution of Saez (2001), in the report chaired by Sir James Mirrlees, which consisted of two substantial volumes (Mirrlees, 2010 and 2011) produced under the aegis of the Institute for Fiscal Studies in London. The attraction is obvious: the use of a reduced-form approach eliminates the need to construct and estimate a fully specified structural model of taxpayers' behaviour, and optimal tax rates can be readily discussed and expressed explicitly in terms of the elasticity of taxable income. Such elasticities are 'bread and butter' to economists and may be regarded as being more 'concrete' than the elements which enter into the determinants of optimal tax rates in the types of structural labour supply model which followed the earlier work of Mirrlees (1971).² These advantages nevertheless come at a cost, in terms of the difficulties of empirical estimation and the strong underlying assumptions required to generate some of the results.

The aim of the present paper is to provide a technical introduction to the use of the elasticity of taxable income in welfare comparisons and optimal tax discussions. It draws together, using a consistent framework and notation, a number of established results concerning marginal welfare changes and optimal taxes, in addition to presenting some new results, particularly in terms of non-marginal tax changes.

Section 2 introduces the quasi-linear utility function that is implicit in all studies which use a constant elasticity specification in which there are no income effects on taxable income of marginal tax rate changes. Section 3 discusses the revenue and welfare effects, measured in terms of the excess burden and marginal welfare cost of small

¹Saez *et al.* (2012) survey a vast literature on the elasticity of taxable income, and an introduction to some of the basic analytics can be found in Creedy (2010).

²However, in the case of the linear income tax, Tuomala (1985) gives some elegant results which can also be expressed in terms of easily interpreted elasticities: for extensions, including comparisons with majority-voting outcomes, see also Creedy (2008).

increases in the marginal tax rate. For simplicity the tax function is assumed to have a single marginal rate applied above an income threshold. The results are extended in Section 4 to allow for the situation in which some income is shifted into an alternative form which faces a lower marginal tax rate. Section 5 introduces the more realistic multi-rate income tax function and shows that the results of the previous sections apply directly to the top marginal tax rate in such a structure. Section 5 then extends those results to deal with tax rates in any of the income brackets in the multi-rate function. Expressions for optimal rates are examined in section 6. The discussion emphasises the treatment of value judgements in this context.

All the results in sections 3 to 6 apply to small changes: this is appropriate when considering optimal rates, for which an equi-marginal condition applies whereby the marginal benefits of a small increase in a tax rate (arising from the extra expenditure financed by the revenue increase) must be equal to the marginal cost (in terms of the weight attached to welfare changes), as perceived by an independent judge or decision maker. The marginal excess burden per dollar of revenue raised (the marginal welfare cost) clearly becomes infinitely large as the tax rate (within any income band) approaches its revenue-maximising rate, in view of the fact that the change in revenue (the denominator) becomes zero. However, the total excess burden as a proportion of the total tax raised remains finite for rates beyond this point. In some policy contexts the total efficiency loss arising from the tax is of primary interest rather than the marginal excess burden or, more importantly, efficiency losses from discrete changes in tax rates are relevant. Section 7 thus extends results to deal with the welfare changes arising from non-marginal tax rate changes. Brief conclusions are in section 8.

2 The Basic Specification

In the literature on the elasticity of taxable income, a constant-elasticity reduced-form specification is ubiquitous, yet its derivation is seldom discussed explicitly. It is therefore useful to begin by stating clearly the nature of the required assumptions concerning individuals' utility functions and budget constraints. Let c denote net income and z gross taxable income, with z_0 the value of income in the absence of taxation. Consider

the quasi-linear form:

$$U = c - \left(\frac{1}{1 + \frac{1}{\eta}} \right) z_0^{-\frac{1}{\eta}} z^{1 + \frac{1}{\eta}} \quad (1)$$

with budget constraint:

$$c = \mu + z(1 - \tau) \quad (2)$$

Here μ is virtual income³ τ is the marginal tax rate. This is associated with a tax function of the form, for $z > a$:

$$T(z) = \tau(z - a) \quad (3)$$

and $T(z) = 0$ for $z < a$. Hence, $\mu = a\tau$, and a is the income threshold, in a tax structure having a constant rate above the tax-free threshold. Substituting for c into U and differentiating gives:

$$\frac{\partial U}{\partial z} = (1 - \tau) - \left(\frac{z}{z_0} \right)^{\frac{1}{\eta}} \quad (4)$$

Setting this equal to zero and solving for z gives:

$$z = z_0(1 - \tau)^\eta \quad (5)$$

and the elasticity of taxable income, $\eta_{z,1-\tau} = \frac{(1-\tau)}{z} \frac{dz}{d(1-\tau)}$, is constant at η . It is the linear term in c , and the absence of μ from the term involving z , which ensures that income effects are zero. It must be acknowledged that the (largely untested) assumption of no income effects is made mainly for pragmatic reasons, as it considerably simplifies the analysis.⁴ Income effects would mean, for example, that the behaviour of higher-rate taxpayers in a multi-rate structure changes in response to tax rate changes in lower-tax brackets.

3 Marginal Revenue and Welfare Changes

This section examines the revenue and welfare effects of small changes in the marginal tax rate for the simple linear tax function. It is shown in section 5 that these results

³In a diagram of the budget constraint, with net income (consumption) on the vertical axis and gross income on the horizontal axis, virtual income is the intercept on the vertical axis. In a multi-rate tax schedule virtual income refers to the extension to the vertical axis of the particular segment under consideration.

⁴The Cobb-Douglas case, which is usually so convenient, produces a much more awkward expression for z : see Creedy (2010, p. 564).

can be directly applied to the top rate in a multi-rate structure.

3.1 Marginal Revenue Changes

The tax paid by an individual is given, for the simple tax function discussed above and for $z > a$, by $\tau(z - a)$. Total revenue collected is thus:

$$R = \tau \sum_{z_i > a} (z_i - a) \quad (6)$$

Let \bar{z}_T denote the arithmetic mean of those above the threshold, and N_T the number of people above the threshold. Then total revenue becomes:

$$R = N_T \tau (\bar{z}_T - a) \quad (7)$$

The effect on R of a small change in τ , denoted MR , is:

$$MR \equiv \frac{dR}{d\tau} = \frac{\partial R}{\partial \tau} + \frac{\partial R}{\partial \bar{z}_T} \frac{d\bar{z}_T}{d\tau} \quad (8)$$

The first term is a pure ‘tax rate’ effect while the second term is a ‘tax base’ effect of the tax rate change. These effects are also sometimes referred to as ‘mechanical’ and ‘behavioural’ effects respectively. This terminology is used, for example, by Saez *et al.* (2012). In Creedy and Gemmell (2013) the tax base effect is decomposed further and shown to depend on the elasticity of taxable income and the revenue elasticity. Writing (8) in elasticity form gives:

$$\eta_{R,\tau} = \eta'_{R,\tau} + (\eta_{R,\bar{z}_T}) (\eta_{\bar{z}_T,\tau}) \quad (9)$$

where $\eta'_{R,\tau}$, the mechanical effect, is a partial elasticity and $\eta_{R,\bar{z}_T} = \eta'_{R,\bar{z}_T}$. The term η_{R,\bar{z}_T} is the tax revenue elasticity at \bar{z}_T . In terms of the elasticity of taxable income, and using for this tax structure, $\eta_{R,\bar{z}_T} = \bar{z}_T / (\bar{z}_T - a)$ and $\eta'_{R,\tau} = 1$, this becomes:

$$\eta_{R,\tau} = 1 - \left(\frac{\bar{z}_T}{\bar{z}_T - a} \right) \left(\frac{\tau}{1 - \tau} \right) \eta_{\bar{z}_T, 1-\tau} \quad (10)$$

Differentiating (6) with respect to τ and \bar{z}_T gives:

$$\frac{\partial R}{\partial \tau} = N_T (\bar{z}_T - a) \quad (11)$$

$$\frac{\partial R}{\partial \bar{z}_T} = N_T \tau \quad (12)$$

and:

$$\frac{dz}{d\tau} = \frac{-\eta z}{1 - \tau} \quad (13)$$

Hence, marginal revenue becomes:

$$\frac{dR}{d\tau} = N_T (\bar{z}_T - a) \left\{ 1 - \eta \left(\frac{\bar{z}_T}{\bar{z}_T - a} \right) \left(\frac{\tau}{1 - \tau} \right) \right\} \quad (14)$$

Define the term α as the ratio of average income, \bar{z}_T , obtained by those above the threshold, a , divided by the average income measured in excess of the threshold, so that:

$$\alpha = \frac{\bar{z}_T}{\bar{z}_T - a} \quad (15)$$

This is also the total income of those above the threshold divided by the total income measured in excess of the threshold. From above, it is also known that this is the same as the revenue elasticity, η_{R, \bar{z}_T} , at \bar{z}_T . Hence (14) is more succinctly written as:

$$\frac{dR}{d\tau} = N_T (\bar{z}_T - a) \left\{ 1 - \eta \alpha \left(\frac{\tau}{1 - \tau} \right) \right\} \quad (16)$$

The elasticity of tax revenue with respect to the tax rate, $\eta_{R, \tau}$, is thus:

$$\eta_{R, \tau} = \frac{1 - \tau (1 + \alpha \eta)}{1 - \tau} \quad (17)$$

The tax rate, τ^* , which maximises revenue, obtained by setting $dR/d\tau = 0$, is thus a simple function of α and the elasticity, η , whereby:

$$\tau^* = (1 + \alpha \eta)^{-1} \quad (18)$$

Thus the revenue change in (14) depends on the precise form of the distribution of declared income and the income threshold above which the tax rate, τ , applies.

3.2 Marginal Welfare Changes

Consider the marginal welfare change arising from a small change in the marginal tax rate, τ . Let $E(\tau, U)$ denote the expenditure function, expressed in terms of virtual income, μ , where individual subscripts have been omitted. For the equivalent variation,

the welfare change resulting from a change in the tax rate from τ_1 to τ_2 , where the change in the tax rate has a dual effect of changing the ‘price’ and the virtual income, is defined, using subscripts to denote appropriate values of U and μ and omitting individual subscripts, as:⁵

$$EV = \{E(\tau_2, U_2) - E(\tau_1, U_2)\} - \{\mu_2 - \mu_1\} \quad (19)$$

The first term is the ‘price effect’ and the second term is the ‘income effect’ of the tax change, and $E(\tau_2, U_2) = \mu_2$. For small changes this can be written as:

$$EV = \frac{\partial E(\tau, U)}{\partial \tau} - \frac{d\mu}{d\tau} \quad (20)$$

Using Shephard’s Lemma (the Envelope theorem), it is known that $\partial E(\tau, U) / \partial \tau = z^H$, where the superscript indicates that it is the Hicksian, or compensated, ‘demand’. In the present context, income effects are absent so that Marshallian and Hicksian demands are equal for each individual. Hence, $\partial E(\tau, U) / \partial \tau = z$.

Furthermore, from the budget constraint defined above, $\mu = a\tau$, and so $d\mu/d\tau = a$ for all individuals above the threshold. Hence the welfare change is simply:

$$EV = z - a \quad (21)$$

This welfare change is equivalent to $\partial R / \partial \tau$ for each individual taxpayer, which is the tax-rate, or mechanical, effect on revenue of a change in τ . Adding these changes over all those above a gives the aggregate welfare change per taxpayer as:

$$EV = \bar{z}_T - a \quad (22)$$

The aggregate marginal excess burden per taxpayer, $MEB = EV - MR$, arising from the tax is, from (16):

$$MEB = (\bar{z}_T - a) \eta \alpha \left(\frac{\tau}{1 - \tau} \right) \quad (23)$$

The MEB is thus equal to the absolute value of the tax-base, or behavioural, effect on tax revenue of a rate change.⁶ The total marginal welfare cost per dollar of extra

⁵On welfare changes and associated concepts, see Creedy (1998a).

⁶This is probably the source of a misunderstanding, regarding the comment by Brewer *et al.* (2011, p. 61) that, ‘A tax change that would have been revenue neutral in the absence of a reduction in work effort will instead produce a revenue loss. It is the size of this revenue loss that determines the [marginal] “excess burden” of taxation’. In his review, Feldstein (2012, p. 782) criticised this comment, interpreting the revenue change, to which Brewer *et al.* alluded, as the total change in revenue, rather than only the behavioural component.

revenue, MWC , is defined as the aggregate marginal excess burden divided by the change in aggregate tax revenue. This is:

$$MWC = \frac{\eta\alpha\tau}{1 - \tau(1 + \eta\alpha)} \quad (24)$$

This expression is relevant only when the marginal tax rate is below the revenue-maximising rate given in (18), so that $dR/d\tau > 0$. The MWC initially rises slowly as τ increases, for low values of τ . Then as τ approaches the value for which $dR/d\tau = 0$, the MWC increases rapidly for further tax rate increases. At the tax rate for which $dR/d\tau = 0$, no extra revenue can be obtained from a small increase in the tax rate and so the marginal welfare cost per dollar of extra revenue is clearly infinitely large. Sometimes this expression for MWC is used to compute its value for increasing values of τ , holding η and α constant. The latter obviously relies on the assumption that the ratio $\alpha = \frac{\bar{z}_T}{\bar{z}_T - a}$ remains constant; that is, it is independent of the tax structure. This property holds only for Pareto distributions. For alternative income distributions, the value of α is likely to change as \bar{z}_T – itself a function of τ – changes. The MWC can also be expressed in terms of the two elasticities – the elasticity of taxable income and the revenue elasticity, as follows:

$$MWC = \alpha \left(\frac{\tau}{1 - \tau} \right) \frac{\eta_{z,1-\tau}}{\eta_{R,\tau}} \quad (25)$$

4 The Effect of Income Shifting

The previous discussion has assumed that the disincentive effect of taxation involves a reduction in taxable income that is also the same as gross income. However, a proportion, s , of income that would otherwise be obtained, or reported, may be shifted into another source, where it is taxed at a lower rate, $t < \tau$. For example, individuals may form trusts for which income is taxed at a lower rate. As in the previous sections, consider the simple income tax function where a single rate is applied to income measured in excess of a threshold, a . The effective tax function is thus:

$$T(z) = \tau(z - a) + st(z_0 - z) \quad (26)$$

Thus, the imposition of income tax at the rate, τ , means that a proportion, s , of the income reduction, $z_0 - z$, is taxed at the rate, t . The individual's optimisation problem is

thus to maximise utility, as in (1), subject to the budget constraint whereby net income, c , is given by:

$$c = z - \tau(z - a) - st(z_0 - z) \quad (27)$$

This can be written as:

$$c = (at - stz_0) + z(1 - (\tau - st)) \quad (28)$$

Virtual income thus becomes $(at - stz_0)$ and the tax rate becomes $(\tau - st)$. As before, substituting for c in the utility function, setting $dU/dz = 0$ and solving for z gives:

$$z = z_0 \{1 - (\tau - st)\}^\eta \quad (29)$$

The solution for z therefore takes the constant elasticity form, as above, but with the rate, τ , replaced by the effective tax rate $\tau - st$.⁷

The tax-rate, or mechanical, effect on revenue of a marginal increase in τ is, as before, $\partial R/\partial\tau = z - a$, while $dz/d\tau = z_0\eta\{1 - (\tau - st)\}^{\eta-1}$ and $\partial R/\partial z = \tau - st$. Hence the tax-base, or behavioural, effect of an increase in τ is given by:

$$\frac{\partial R}{\partial z} \frac{dz}{d\tau} = \left(\frac{\tau - st}{1 - (\tau - st)} \right) \eta z \quad (30)$$

This is, as shown above, the same as the excess burden, so that the marginal welfare cost of a small increase in τ , following the same steps as before, becomes:

$$MWC = \frac{\alpha\eta(\tau - st)}{1 - (\tau - st)(1 + \alpha\eta)} \quad (31)$$

This is clearly the same as the earlier result for $s = 0$, but with τ replaced by $\tau - st$.⁸

5 A Multi-Rate Tax Structure

The previous sections have considered the case of a tax structure having a single rate applied to income measured above a tax-free threshold. The present section extends the

⁷This clearly raises problems for the estimation of η , since s cannot be observed. Estimation is beyond the scope of the present paper.

⁸Saez *et al.* (2012, p. 11) give an incorrect expression, by not recognising that in this case τ must be replaced by $\tau - st$ in the solution for z . An incorrect form is also given in Creedy (2010, p. 572), which also contains a printing error, and Claus *et al.* (2012, p. 301), who follow Saez *et al.*

results to the more realistic multi-rate structure that is widely used in practice. Subsection 5.1 describes the tax structure and shows that the results in previous sections can be interpreted as simply applying to the top rate in a multi-rate structure. Subsection 5.2 considers marginal revenue and welfare changes for intermediate rates.

5.1 The Tax Function

Consider the multi-step tax function, which is defined by a set of income thresholds, a_k , for $k = 1, \dots, K$, and marginal income tax rates, τ_k , applying in tax brackets, that is between adjacent thresholds a_k and a_{k+1} . The function can be written as:

$$\begin{aligned} T(z) &= \tau_1(z - a_1) & a_1 < z \leq a_2 \\ &= \tau_1(a_2 - a_1) + \tau_2(z - a_2) & a_2 < z \leq a_3 \end{aligned} \quad (32)$$

and so on. If z falls into the k th tax bracket, so that $a_k < z \leq a_{k+1}$, $T(z)$ can be written for $k \geq 2$ as:

$$T(z) = \tau_k(z - a_k) + \sum_{j=1}^{k-1} \tau_j(a_{j+1} - a_j) \quad (33)$$

Letting $b_k = \sum_{j=1}^{k-1} \tau_j(a_{j+1} - a_j)$ this becomes $T(z) = \tau_k(z - a_k) + b_k$.⁹ Hence for an individual whose income falls into the k th tax bracket, the budget constraint in (2) becomes:

$$c = (\tau_k a_k - b_k) + z(1 - \tau_k) \quad (34)$$

and the virtual income, μ , is simply reduced by the term b_k . This means that all the above results can be applied directly to the top rate in a multi-rate structure. Importantly, references to tax revenue must all refer to revenue collected at the top marginal rate only. The assumption that the top tail of the distribution can be approximated by the Pareto distribution is clearly more reasonable in this context. The above results can easily be extended to the case of any tax rate in a multi-rate structure, as follows.

⁹This expression for $T(z)$ can be rewritten as $T(z) = \tau_k(z - a_k^*)$ where $a_k^* = \frac{1}{\tau_k} \sum_{j=1}^k a_j (\tau_j - \tau_{j-1})$ and $\tau_0 = 0$. Thus the tax function facing any individual taxpayer in the k th bracket is equivalent to a tax function with a single marginal tax rate, τ_k , applied to income measured in excess of a single threshold, a_k^* . Therefore, unlike a_j , a_k^* differs across individuals depending on the marginal income tax bracket into which they fall.

5.2 Changes in Intermediate Tax Rates

In order to consider changes in lower tax rates, rather than the top rate, it is sufficient here to consider a two-rate structure, where the rate τ_L applies to incomes between the income thresholds a_L and a_H . Let N_L denote the number of people in the first tax bracket and N_H the number in the top bracket.¹⁰ Let R_{τ_L} denote the total tax revenue raised at the rate τ_L , that is only from income that falls into the lower bracket, for which $a_L < z < a_H$. The higher-rate payers must pay τ_L on an amount, $a_H - a_L$, of their income, so that:

$$R_{\tau_L} = N_H \tau_L (a_H - a_L) + N_L \tau_L (\bar{z}_L - a_L) \quad (35)$$

where \bar{z}_L is the arithmetic mean income of those who fall into the tax bracket with the marginal of rate τ_L . The corresponding marginal revenue, using $\frac{dR_{\tau_L}}{d\tau_L} = \frac{\partial R_{\tau_L}}{\partial \tau_L} + \frac{\partial R_{\tau_L}}{\partial \bar{z}_L} \frac{d\bar{z}_L}{d\tau_L}$, is:

$$\frac{dR_{\tau_L}}{d\tau_L} = N_H (a_H - a_L) + N_L (\bar{z}_L - a_L) \left\{ 1 - \eta \alpha \left(\frac{\tau_L}{1 - \tau_L} \right) \right\} \quad (36)$$

where in this case:

$$\alpha = \frac{\bar{z}_L}{\bar{z}_L - a_L} \quad (37)$$

From earlier results the aggregate marginal excess burden is:

$$MEB_{\tau_L} = N_L (\bar{z}_L - a_L) \eta \alpha \left(\frac{\tau_L}{1 - \tau_L} \right) \quad (38)$$

and the marginal welfare cost is thus found to be:

$$MWC_{\tau_L} = \frac{\alpha \eta \tau_L}{1 - \tau_L (1 + \alpha \eta \tau_L) + D} \quad (39)$$

where:

$$D = (1 - \tau_L) \frac{(a_H - a_L) N_H}{(\bar{z}_L - a_L) N_L} \quad (40)$$

The expression for the marginal welfare cost of raising the lower tax rate is thus the same as for the top tax rate, with the addition of the term D in the denominator.

The rate that maximises revenue from the rate τ_L is given by:

$$\tau^* = \left[1 + \eta \left\{ \frac{N_L \bar{z}_L}{N_H (a_H - a_L) + N_L (\bar{z}_L - a_L)} \right\} \right]^{-1} \quad (41)$$

¹⁰In general, of course, N_H can refer to all those in higher-rate brackets than the one being considered.

which clearly reduces to $\tau^* = (1 + \eta\alpha)^{-1}$ for the top tax rate, as obtained above in (18). Furthermore, the interpretation of the term in curly brackets in (41) is the same as that of α for the earlier result: it is the ratio of total income of those who fall into the relevant bracket to the total income that is taxed at the relevant marginal rate.

6 Optimal Tax Rates

The use of a reduced-form expression for taxable income in terms of the marginal tax rate means that it is also possible to express optimal tax rates in terms of the elasticity of taxable income, using the above results. In the ‘standard’ optimal tax literature stemming from Mirrlees (1971), in the context of a tax and transfer system where labour supply is endogenous, a starting point is a social welfare, or evaluation, function expressed in terms of individuals’ utilities. This welfare function reflects the value judgements of an independent decision maker, in particular regarding the judge’s aversion to inequality. The decision maker selects the tax rate (for example, the single rate in a linear tax function) to maximise the welfare function, while individuals select their labour supply to maximise utility. The value of a transfer payment is determined by the need to satisfy a government budget constraint. This budget constraint may involve a requirement to raise a given amount of non-transfer expenditure per person (rather than considering a ‘pure’ transfer system), but the optimal tax models usually consider this as involving a ‘black hole’, in that the benefits of the resulting expenditure do not enter either individuals’ utility functions or the welfare function of the decision maker. It is well known that in general numerical simulation methods must be used to obtain results.¹¹ However, for this structural model, the value judgements of the decision maker, the nature of the tax and transfer system, and the government’s budget constraint are entirely transparent.

In the present context a social welfare function is not fully specified but a decision maker is assumed to take a view about the value of additional government tax-financed expenditure resulting from the extra revenue from a small tax increase.¹² The additional

¹¹For references to special cases where explicit solutions are available, and an approximation in the case of the linear income tax, see Creedy (2009).

¹²Perhaps understandably, the report in Mirrlees (2011) often conflates the two approaches, suggesting that the use of reduced-form elasticities, allowing income adjustment in addition to labour supply incentive effects, is in the Mirrlees tradition. The common ground is of course a concept of an optimum,

expenditure is not explicitly divided into transfer and other expenditure. Given that a reduced-form model of individual behaviour is used, neither component of this expenditure is considered to enter the utility functions. The independent judge also forms a view about the weight attached to the loss of welfare resulting from the small tax increase. The loss of welfare is expressed as in previous sections above.

6.1 First-order Conditions in a Multi-rate Structure

The approach involves considering each tax bracket in turn; hence decisions regarding income thresholds are supposed already to have been made.¹³ The value judgements of the decision maker are reflected in two terms. The social marginal valuation, SMV , reflects the weight attached to the loss of welfare suffered by those in the relevant tax bracket as a result of a small tax increase. The $MVPF$ is the value attributed by the judge to the extra tax-financed expenditure resulting from the small tax increase. The optimal tax rate in the bracket is that rate for which (in the view of the judge) the marginal benefit of a further tax increase just matches the marginal cost. Hence the first-order condition for each tax bracket is:

$$(EV)(SMV) = (MR)(MVPF) \quad (42)$$

The left-hand side of (42) is the marginal cost, while the right-hand side is the marginal benefit of the tax increase. The previous sections have expressed the efficiency cost of a marginal tax increase in terms of the marginal excess burden per dollar of extra revenue, the MWC . Thus it is useful to convert this ‘equi-marginal’ condition into one that involves the MWC . First, rewrite (42) as:

$$\left. \frac{EV}{MR} \right|_{\tau_{opt}} = \left. \frac{MVPF}{SMV} \right|_{\tau_{opt}} \quad (43)$$

and since by definition:

$$MWC = \frac{EV}{MR} - 1 \quad (44)$$

based on value judgements, an allowance for incentive effects, and the ability to express the optimum in terms of an equi-marginal condition.

¹³A more general approach in which the tax rate can vary continuously over the whole income range is discussed in Saez (2001) and in Brewer *et al.* (2010). The present approach is adopted for simplicity.

this first-order condition becomes:

$$MWC|_{\tau_{opt}} = \frac{MVPF}{SMV}|_{\tau_{opt}} - 1 \quad (45)$$

Public tax-financed projects may be subject to decreasing marginal valuation by the judge, and the valuation, SMV , may well depend on the tax bracket being considered.

For example, consider the simplest case above, where income is not shifted to lower-taxed sources (so that $s = 0$) and the rate being examined is the top rate in a multi-tax structure. Let g denote the reciprocal of $\frac{MVPF}{SMV}|_{\tau_{opt}}$. Substituting for $MWC = \frac{\eta\alpha\tau}{1-\tau(1+\eta\alpha)}$ from (24) and re-arranging (45) gives the optimal rate as:

$$\tau_{opt} = \frac{1-g}{1+\eta\alpha-g} = \left(1 + \frac{\alpha\eta}{1-g}\right)^{-1} \quad (46)$$

Furthermore, substituting for α gives the alternative expression:

$$\tau_{opt} = \left\{1 + \left(\frac{\eta}{1-g}\right) \left(\frac{N_T \bar{z}_T}{N_T(\bar{z}_T - a)}\right)\right\}^{-1} \quad (47)$$

The term $\frac{N_T \bar{z}_T}{N_T(\bar{z}_T - a)}$ measures the ratio of the total income of those in the top tax bracket to the total income that is subject to the top tax rate. In the extreme case where the judge does not care about top-rate taxpayers, $g = 0$ and the optimal rate is the same as the rate which maximises revenue from those taxpayers. However, this is a closed-form solution only in the case where g is considered to be constant (that is, independent of the tax rate), otherwise the precise form of $g(\tau)$ must be known.

Consider the optimal value for a lower marginal tax rate, τ_L , in the two-rate structure considered earlier (and which is easily extended to the multi-rate form). From above, this must satisfy:

$$MWC_{\tau_L} = \frac{\alpha\eta\tau_L}{1 - \tau_L(1 + \alpha\eta\tau_L) + (1 - \tau_L) \frac{(a_H - a_L) N_H}{(\bar{z}_L - a_L) N_L}} = \frac{1}{g} - 1 \quad (48)$$

which can be solved to give:

$$\tau_{L,opt} = \left\{1 + \frac{\alpha\eta}{(1-g) \left(1 + \frac{(a_H - a_L) N_H}{(\bar{z}_L - a_L) N_L}\right)}\right\}^{-1} \quad (49)$$

Another way to express this is:

$$\tau_{L,opt} = \left\{1 + \left(\frac{\eta}{1-g}\right) \left(\frac{N_L \bar{z}_L}{N_H(a_H - a_L) + N_L(\bar{z}_L - a_L)}\right)\right\}^{-1} \quad (50)$$

The term $N_L \bar{z}_L$ is the income of those in the relevant tax bracket, while the term $N_H (a_H - a_L) + N_L (\bar{z}_L - a_L)$ measures the income to which the rate τ_L is applied. Hence the expression for the optimal rate corresponds precisely with that given in (47) for the optimal top marginal rate.

6.2 Comparison with Earlier Results

Instead of writing the optimal condition in terms of the marginal welfare cost, Saez (2001) expressed the condition for the optimal rate using the decomposition of marginal revenue into mechanical and behavioural terms, M and B respectively (where of course B is negative). Write (8) as $MR = M + B$ and from (22), it is known that $EV = M$. Hence rearranging (42) as $Mg = M + B$ gives the first-order condition as $M(1 - g) + B = 0$, which is the form given in Saez (2001, p. 210), who does not state (42) explicitly. For the revenue-maximising rate, $MR = 0$ and $M = -B$: Brewer *et al.* (2010, p. 102) thus refer to this rate as ‘balancing mechanical and behavioural effects’.

When discussing optimal rates Brewer *et al.* (2010) write the condition, using current notation, as $M + B - gM = 0$. In their discussion, the value of $MVPPF$ is implicitly set at 1.¹⁴ Hence the term gM is effectively $(M)(SMV)$ and as $EV = M$ this is the change in ‘social welfare’ resulting from a small tax rate change (that is, the left hand side of the optimal rate condition in (42)). In their own notation, Brewer *et al.* write mechanical and behavioural effects on revenue as dM and dB respectively, and they write the social welfare change, $-gM$, as dW . Thus their condition is written as $dM + dB + dW = 0$. In their discussion of appropriate settings for the value of g , they therefore consider only the variation in the welfare loss.

6.3 Imposing Value Judgements

The role of professional economists, following the famous statement by Robbins (1935), is to examine the implications of adopting alternative value judgements. In the present context this means examining the effects on optimal tax rates of alternative values of the ratio $g = SMV/MVPPF$. The question therefore arises of how to interpret different orders of magnitude.

¹⁴An allusion to this is later made in Brewer *et al.* (2010, p. 166, n. 75).

In the branch of optimal tax literature that follows the structural modelling approach of Mirrlees (1971), it is usual to consider the independent judge as selecting tax rates which maximise the value of a particular social welfare function, expressed in terms of individuals' utilities. The objective is thus entirely transparent and it is clear that interpersonal comparisons of utility are explicitly being made by the judge. Although this allows for a range of types of welfare function, the most common form to be examined is the additive, individualistic, Paretean, and Utilitarian form with constant relative inequality aversion, ε , so that $W = \left(\frac{1}{1-\varepsilon}\right) \sum_{i=1}^N U_i^{1-\varepsilon}$.¹⁵ The exercise then becomes one of examining the effects of using different values of ε , and in interpreting orders of magnitude it is useful to consider the well-known 'leaky bucket' experiment.¹⁶ Within this framework, the optimal tax rate depends not only on the form of the social welfare function but also on the cardinalisation used for individuals' utility functions, although this is usually given less attention.¹⁷

The general structural approach can also deal with 'non-welfarist' social welfare functions, in which the judge does not evaluate outcomes in terms of things that matter directly to the individuals involved (such as their utility) but in terms of, for example, some aggregate poverty measure, or the number of non-workers.

In the present context of using the elasticity of taxable income in a reduced-form model, the choice of alternative values of g is less straightforward. Little guidance is given by Saez (2001), Brewer *et al.* (2010) and Mirrlees (2011). As mentioned earlier, Brewer *et al.* (2010) set the value of $MVPF$ equal to one, and concentrate on discussing values of SMV .¹⁸ The Mirrlees (2011) report gives most emphasis to the revenue maximising rate in the top tax bracket, which (2011, p. 65) is 'equivalent to placing a zero value on

¹⁵The 'classical utilitarian' form – which was in fact the one considered in Mirrlees's original paper – is of course simply the sum of individuals' utilities (that is, inequality aversion is zero).

¹⁶This involves considering taking a \$1 from one person and deciding how much one is prepared to lose (from the leaky bucket) in making a transfer to a poorer person. With this welfare function, and incomes of the rich and poor individuals as z_R and z_P respectively, a judge would tolerate a leak of $1 - (z_P/z_R)^\varepsilon$ from the initial \$1 taken from z_R .

¹⁷See Creedy (1998), where the use of money metric utility is explored; this is of course a particular cardinalisation which is invariant with respect to monotonic transformations of utility but does depend on the choice of 'reference prices'.

¹⁸However, they persistently refer to society's or government's views about inequality. In view of well-known problems relating to the aggregation of preferences, the 'social welfare function' instead must be interpreted as representing the value judgements of a single independent person.

their (marginal) welfare’.¹⁹ Different judges may be concerned more explicitly with the question of how the tax revenue is spent: in the structural approach there is an explicit transfer payment and some non-transfer expenditure, the amount of which is considered to have previously been determined and which does not enter individuals’ utilities.²⁰

One approach might be to suppose that the judge envisages a tax and transfer system and applies an evaluation function of the form, $(\frac{1}{1-\varepsilon}) \sum_{k=1}^K \bar{z}_k^{1-\varepsilon}$, that is a weighted sum over the K tax brackets of average taxable income in each bracket. Consider the tax rate in the k th bracket, where transfers are assumed to go to those in the 1st bracket (and with no income effects there are no consequences for the behaviour of those in the 1st bracket). The (absolute) slope of the ‘social indifference curve’ relating \bar{z}_k and \bar{z}_1 values for which social welfare is unchanged is thus $(\bar{z}_1/\bar{z}_k)^\varepsilon$. For example, if the tax rate is being considered in a bracket for which the average income is twice that in the lowest tax bracket, and $\varepsilon = 1$, then $g = 0.5$. A lower value of $\varepsilon = 0.5$ gives $g = 0.71$ while a higher value of $\varepsilon = 2$ gives $g = 0.25$. Of course, a difficulty here is that the incomes are themselves endogenous and the link with utility is not straightforward.

Table 1: Examples of Optimal Rates: New Zealand Thresholds and Income Distribution 2010

a_L	N_L	\bar{z}_L	η	g	τ_{opt}
0	807.04	6.948	0.2	0.996	0.126
14	1687.21	27.079	0.5	0.880	0.207
48	462.37	57.546	0.5	0.455	0.331
70	347.84	115.419	0.6	0.075	0.378

To illustrate the use of the results, Table 1 is based on the New Zealand taxable income distribution for 2010, with the 2009-1010 income thresholds. As the exercise is purely illustrative, only income taxation is considered and hence no account of taxable welfare benefits or indirect taxes, and their effect on overall effective tax rates, is taken.

¹⁹This led to Feldstein’s (2012, p. 783) question, ‘what kind of nation places no value on the welfare of those with income in the top bracket, treating them as the revenue producing property of the state?’ and comment that ‘many noneconomists would find the Review’s suggestion that a society could disregard the welfare of any group of taxpayers repugnant’. Here it is of course important to distinguish between the value attached to the *total welfare* of those in the top bracket and the value attached to a *marginal reduction* in welfare. It is the latter to which *SMV*, and hence g , applies.

²⁰However, some authors have investigated the implications of allowing public good expenditure to influence individuals’ labour supply decisions.

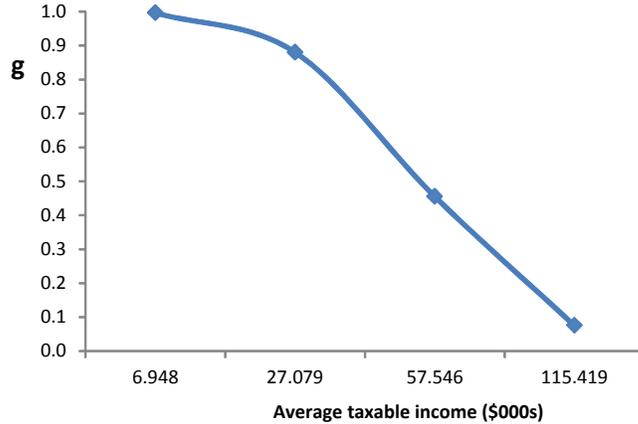


Figure 1: Variation in g with \bar{z}

The first column gives the lower income threshold (in thousands of dollars) for each tax band, while columns headed N_L and \bar{z}_L show, again in units of thousands, the number of individuals in each bracket and the arithmetic mean income respectively. The column headed η gives assumed values of the elasticity of taxable income within each bracket: these are hypothetical but are based on the results of Claus *et al.* (2012). Imposed values of g and the implied optimal rates are shown in the final two columns of the table. The final column may be compared with the actual New Zealand rates of, respectively: 0.125; 0.21; 0.33 and 0.38. Hence, the values of g were obtained in each case following a trial and error search process such that the optimal rate matches the actual rate very closely.

It can be seen that the values of g , required for the optimal rate schedule to replicate the actual tax rates, fall rapidly. These are plotted in Figure 1. The top marginal rate of 0.38 could be said to be consistent with the value judgement that places very little value on the marginal welfare loss of those in the top bracket: indeed, the revenue-maximising rate in this bracket is 0.396.²¹ For the other tax brackets, the rates that maximise revenue are (from brackets one to three respectively) equal to 0.973, 0.685 and 0.476. These are in each case substantially higher than the actual rates.

²¹Using the constant inequality aversion approach described above, $(\bar{z}_1/\bar{z}_4)^2 = 0.0036$, which gives an optimal rate of 0.395. A parameter of 2 implies a very high value of inequality aversion: the judge would be prepared to take \$1 from a person on average income in the top bracket, give less than half a cent to an average person in the lower bracket and throw away the rest.

7 Welfare and Non-marginal Tax Changes

Instead of considering small changes, it is useful to be able to evaluate the welfare changes associated with a given tax rate compared with a no-tax situation or, more often, to evaluate the effect of a significant (non marginal) change. In practice, many tax reforms cannot be considered to involve ‘marginal’ changes in tax rates. An example involves the introduction of a new top marginal rate in a multi-rate structure. In order to evaluate such changes, it is necessary to consider the precise form of the expenditure function in (19).

7.1 The Expenditure Function

First obtain the indirect utility function, V , by substituting (5) and (2) into (1) to get:

$$\begin{aligned} V &= \mu + z_0(1 - \tau)^{1+\eta} - \left(\frac{\eta}{1 + \eta}\right) z_0^{-\frac{1}{\eta}} \{z_0(1 - \tau)^\eta\}^{1+\frac{1}{\eta}} \\ &= \mu + z_0(1 - \tau)^{1+\eta} \left(\frac{1}{1 + \eta}\right) \end{aligned} \quad (51)$$

Hence solving for μ in terms of U and τ gives:²²

$$E(\tau, U) = U - z_0(1 - \tau)^{1+\eta} \left(\frac{1}{1 + \eta}\right) \quad (52)$$

Substituting into (19) gives, for an increase in τ from τ_1 to τ_2 :

$$EV = \left(U_2 - (1 - \tau_2)^{1+\eta} \frac{z_0}{1 + \eta} \right) - \left(U_2 - (1 - \tau_1)^{1+\eta} \frac{z_0}{1 + \eta} \right) - \{\mu_2 - \mu_1\} \quad (53)$$

and:

$$EV = \frac{z_0}{1 + \eta} \{(1 - \tau_1)^{1+\eta} - (1 - \tau_2)^{1+\eta}\} - \{\mu_2 - \mu_1\} \quad (54)$$

The change in revenue from a non-marginal tax rate change is:

$$\begin{aligned} \Delta R &= \tau_2(z_2 - a) - \tau_1(z_1 - a) \\ &= \tau_2 z_2 - \tau_1 z_1 - a(\tau_2 - \tau_1) \end{aligned} \quad (55)$$

and:

$$\Delta R = \tau_2 z_0 (1 - \tau_2)^\eta - \tau_1 z_0 (1 - \tau_1)^\eta - a(\tau_2 - \tau_1) \quad (56)$$

Using $\mu = at$, the term $a(\tau_2 - \tau_1)$ is equal to $\mu_2 - \mu_1$.

²²Using this result, Shephard’s Lemma referred to above is easily confirmed, whereby $\partial E/\partial \tau = z$.

7.2 The case where $t_1 = 0$

Consider first the case where $\tau_1 = 0$ and $\tau_2 = \tau$. That is, consider the welfare change from the introduction of a tax, rather than a change in the tax rate. Then setting $\tau_1 = 0$ in (54):

$$EV = \frac{z_0}{1 + \eta} \{1 - (1 - \tau)^{1+\eta}\} - \mu_2 \quad (57)$$

and:

$$\Delta R = \tau z_0 (1 - \tau)^\eta - a\tau \quad (58)$$

Hence the excess burden, EB , is:

$$EB = \frac{z_0}{1 + \eta} \{1 - (1 - \tau)^{1+\eta}\} - \tau z_0 (1 - \tau)^\eta \quad (59)$$

Using $z_0 = z(1 - \tau)^{-\eta}$ this becomes:

$$\begin{aligned} EB &= \frac{z}{(1 - \tau)^\eta (1 + \eta)} \{1 - (1 - \tau)^{1+\eta}\} - \tau z \\ &= \frac{z}{1 + \eta} \{(1 - \tau)^{-\eta} - (1 - \tau)\} - \tau z \end{aligned} \quad (60)$$

The excess burden per taxpayer (that is for those N_T people with $z > a$) is thus obtained from (60) by replacing z with \bar{z}_T . The tax revenue per taxpayer is $\tau(\bar{z}_T - a)$. Hence in aggregate the welfare cost, the excess burden per dollar of revenue, now denoted simply as WC , is:

$$WC = \frac{\bar{z}_T}{\bar{z}_T - a} \left[\frac{(1 - \tau)^{-\eta} - (1 - \tau)}{(1 + \eta)\tau} - 1 \right] \quad (61)$$

As before, in the Pareto case $\alpha = \bar{z}_T / (\bar{z}_T - a)$ is constant.

7.3 An Increase from τ_1 to τ_2

For a non-marginal change in the tax rate, (54) and (55) give, where ΔEB is written instead of MEB to indicate that discrete changes are considered:

$$\Delta EB = \frac{z_0}{1 + \eta} \{(1 - \tau_1)^{1+\eta} - (1 - \tau_2)^{1+\eta}\} - \{\tau_2 z_2 - \tau_1 z_1\} \quad (62)$$

$$= \frac{z_2}{1 + \eta} \left\{ \frac{(1 - \tau_1)^{1+\eta}}{(1 - \tau_2)^\eta} - (1 - \tau_2) \right\} - z_2 \left\{ \tau_2 - \tau_1 \frac{z_1}{z_2} \right\} \quad (63)$$

Using the earlier results, the change in the total excess burden per dollar of extra revenue, $\tau_2(\bar{z}_2 - a) - \tau_1(\bar{z}_1 - a)$, denoted ΔWC , becomes:

$$\Delta WC = \frac{\frac{1}{1+\eta} \left\{ \frac{(1-\tau_1)^{1+\eta}}{(1-\tau_2)^\eta} - (1-\tau_2) \right\} - \left\{ \tau_2 - \tau_1 \frac{\bar{z}_1}{\bar{z}_2} \right\}}{\left\{ \tau_2 - \tau_1 \frac{\bar{z}_1}{\bar{z}_2} \right\} - \frac{a}{\bar{z}_2} (\tau_2 - \tau_1)} \quad (64)$$

Furthermore:

$$\frac{\bar{z}_1}{\bar{z}_2} = \left(\frac{1-\tau_1}{1-\tau_2} \right)^\eta \quad (65)$$

and writing $\alpha = \bar{z}_2 / ((\bar{z}_2 - a))$, then $a/\bar{z}_2 = 1 - \frac{1}{\alpha}$. It can be seen that by letting $\tau_1 = 0$, this result reduces to (61).

8 Conclusions

The aim of this paper has been to provide a technical introduction to the use of the elasticity of taxable income in welfare comparisons and optimal tax discussions. This concept is now widely used in discussions of income tax policy, although a number of the results and assumptions are not entirely transparent in the literature. Using a consistent framework and notation, a number of established results concerning marginal welfare changes and optimal tax rates are derived. In addition, some new results relating to non-marginal tax changes, which are often relevant in practice, are presented.

It is particularly important to be able to consider the relevant value judgements used, so that the sources of policy disagreements can more easily be identified. Attention was given to the way value judgements enter into the calculation of optimal tax rates using the elasticity of taxable income measure, where they are somewhat less explicit than in the context of structural models which maximise a specified social welfare, or evaluation, function.

It was stated in the introduction that the use of the reduced-form concept of the elasticity of taxable income allows some strong results to be obtained in terms that are perhaps ‘more concrete’ than the results from structural models of optimal income taxation. However, it has also been seen that the results come at a cost of some strong simplifying assumptions. One point that perhaps needs stressing here is that the elasticity of taxable income, even within a model having a constant elasticity, is not in fact

a fixed parameter but depends on many elements of the tax structure including, in particular, the ease of shifting income between sources. This leads to a distinction between optimal tax *rates* and optimal tax *structures*, where the latter includes things like the ease of becoming incorporated, establishing family trusts, and so on.

Both structural and reduced-form models are clearly highly simplified, both in terms of the economic environment and the behaviour of individuals. Neither approach can of course be expected to provide detailed practical policy advice. However, they can both be used, in their different ways, to illuminate and clarify different aspects of the complex relationships involved in choosing a tax rate structure.

References

- [1] Brewer, M., Saez, E. and Shephard, A. (2010) Means-testing and tax rates on Earnings. In *Dimensions of Tax Design* (chair: Sir James Mirrlees), pp. 90-173. Oxford: Oxford University Press for the Institute for Fiscal Studies.
- [2] Claus, I., Creedy, J. and Teng, J. (2012) The elasticity of taxable income in New Zealand. *Fiscal Studies*, 33, pp. 287-303.
- [3] Creedy, J. (1998a) *Measuring Welfare Changes and Tax Burdens*. Cheltenham: Edward Elgar.
- [4] Creedy, J. (1998b) The optimal linear income tax model: utility or equivalent income? *Scottish Journal of Political Economy*, 45, pp. 99-110.
- [5] Creedy, J. (2008) Choosing the tax rate in a linear income tax structure. *Australian Journal of Economics*, 11, pp. 257-276.
- [6] Creedy, J. (2009) Explicit solutions for the optimal linear income tax rate. *Australian Economic Papers*, 48, pp. 224-236.
- [7] Creedy, J. (2010) Elasticity of taxable income: an introduction and some basic analytics. *Public Finance and Management*, 10, pp. 556-589.
- [8] Creedy, J. and Gemmell, N. (2013) Measuring revenue responses to tax rate changes in multi-rate income tax systems: behavioural and structural factors. *International Tax and Public Finance* (forthcoming).
- [9] Feldstein, M. (1995) The effect of marginal tax rates on taxable income: a panel study of the 1986 Tax Reform Act. *Journal of Political Economy*, 103, pp. 551-572.
- [10] Mirrlees, J.A. (1971) An exploration in the theory of optimal income taxation. *Review of Economic Studies*, 38, pp. 175-208.
- [11] Mirrlees, J.A. (Chair) (2011) *Tax by Design*. Oxford: Oxford University Press for the Institute for Fiscal Studies.

- [12] Robbins, L.R. (1935, 2nd edition) *An Essay on the Nature and Significance of Economic Science*. London: Macmillan.
- [13] Saez, E. (2001) Using elasticities to derive optimal income tax rates. *Review of Economic Studies*, 68, pp. 205-229.
- [14] Saez, E., Slemrod, J.B. and Giertz, S.H. (2012) The elasticity of taxable income with respect to marginal tax rates: a critical review. *Journal of Economic Literature*, 50, pp. 3-50.
- [15] Tuomala, M. (1985) Simplified formulae for optimal linear income taxation. *Scandinavian Journal of Economics*, 87, pp. 668-672.

About the Author

John Creedy is Visiting Professorial Fellow at Victoria Business School, Victoria University of Wellington, New Zealand, and a Principal Advisor at the New Zealand Treasury. He is on leave from the University of Melbourne where he is The Truby Williams Professor of Economics.
Email: john.creedy@vuw.ac.nz

