# The Personal Income Tax Structure: Theory and Policy[*]

John Creedy
The University of Melbourne

**Abstract**

There is now a large and complex literature on optimal income taxation, within the context of second-best welfare economics. This paper considers the potential role of this analysis in the practical design of direct tax and transfer structures. It is stressed that few results are robust, even in simple models, in view of the important role played by alternative social welfare functions, the nature of the distribution of abilities and the preferences of individuals. In view of these negative results, it is suggested that a range of empirical tax analyses, capturing particular issues, can provide helpful guidance for policy analysts. Numerical illustrations are provided, paying attention to the role of a 'top' marginal tax rate applied to higher-income groups. In particular, behavioural microsimulation models can be used to examine marginal direct tax reform. Such models have the advantages of capturing the full extent of population heterogeneity and the complexity of the tax structure.

# 1  Introduction

If asked for practical advice about taxation, economists for many years would have referred to Adam Smith's (1776) famous four maxims (contribution according to ability to pay[1], certainty, convenience, and 'efficiency' – the latter including administrative costs, distortions to activity, and the 'vexation and oppression' involved). While the list of such criteria was extended and clarified,[2] in the discussion of ability to pay it is clear that there was no acceptance of a redistributive role.[3] An explicit preference for equality of treatment via proportional taxation was made most clear in the often quoted remark by McCulloch, the author of the most extensive and systematic treatment of public finance in the classical literature. McCulloch (1845) argued that 'The moment you abandon the cardinal principle of exacting from all individuals the same proportion of their income or of their property, you are at sea without rudder or compass, and there is no amount of injustice and folly you may not commit'. The appropriate tax rate is thus determined by the independently given revenue requirement.[4]

The main change in the approach to taxation came from the later integration of public finance into the general area of welfare economics, which was itself a major concomitant of the successful introduction of a utility maximising approach to exchange in the 1870s. However, the most systematic early developments came from Cohen-Stuart (1889) and Edgeworth (1897) in investigating the broad implications for progressivity of the utility maximising principle, in the context of the minimisation of the total disutility from taxation – ignoring any possible benefits.[5] Even here, taking a 'classical utilitarian' perspective, there was no explicit independent role for redistribution: the maximand was strictly considered to be *total* utility.

This movement reached its ultimate conclusion in the 'optimal tax' literature, be-

---

[1] There was some ambiguity here as to whether Smith held a benefit view or a sacrifice view.

[2] In particular it was extended by Lord Overstone, who suggested that a tax should be productive, computable, divisible, frugal, non-interferent, unannoyant, equal, popular, and uncorruptive. This classification influenced Norman, who added 'unevasibility'. Norman's approach was dominated by his utilitarian view of sacrifice in an ability to pay context; see O'Brien (2009).

[3] Later explorations of the concept of 'equal sacrifice', as discussed for example by J.S. Mill (1848), left some ambiguity here, but see below.

[4] The income tax structure at the time had a tax-free threshold (and there were no transfer payments), but the 'degressive' rate structure meant that it was proportional for higher-income recipients.

[5] With the criterion of minimising total sacrifice, progression arises from decreasing marginal utility, but with equal absolute sacrifice it depends on the precise behaviour of the marginal utility of income.

ginning virtually a century after the initial introduction of a utility analysis into economics,[6] whereby the mathematical analyses of Edgeworth were extended by allowing, in particular, for labour supply incentive effects of taxes and transfers, and including a range of specifications of the objectives of taxation, thus introducing an explicit redistributive role. In this final development, most of the important criteria suggested by Smith and others were quietly ignored. The relevant branch of welfare economics into which optimal tax theory falls is the theory of the 'second best', in view of the fact that the government is unable to tax individuals' ability 'endowments' and instead taxes their incomes.[7]

In drawing on this branch of modern economic theory to provide policy advice regarding tax structures, a number of serious difficulties immediately arise. Tax models have a way of getting very complicated very quickly.[8] Many interdependencies are usually involved and while economists are specially trained to identify such inter-relations, clear views can only be obtained by abstracting from many realistic features. Indeed, many of the strong results from tax analyses (for example certain equivalence results concerning uniform direct and indirect taxes in general equilibrium models) are best interpreted as demonstrating that in fact they are most unlikely to apply in practice.[9] Furthermore, value judgements are inevitably involved because virtually every tax and transfer has distributional implications, and ultimately what matters is not the equity effect of a single tax considered in isolation, but the overall impact of a wide range of taxes and transfer payments (the latter being a relatively modern development). Many economic analyses are specially designed to give insights into the nature of the interdependencies involved.

The possible effects of taxes are often expressed in terms of measures for which it

---

[6] And indeed that analysis, at its inception, included the utility treatment of labour supply behaviour, by Jevons (1871). It arises from the emphasis on exchange, that is, selling labour in return for consumer goods.

[7] Sometimes stress is placed on asymmetric information aspects, in that the government cannot observe ability levels. However, this seems rather artificial as the maximisation of a social welfare function itself presupposes a huge amount of information (including net incomes, hours levels, and resulting utilities, of all individuals).

[8] For example, for the simplest possible tax-transfer structure – the linear income tax – with individuals differing only by productivity, the government budget constraint is nonlinear in the tax parameters if there are some non-workers.

[9] A further example is the result that indirect taxes are not necessary if preferences are separable between goods and leisure (marginal rates of substitution between goods do not vary with the wage rate).

is extremely difficult to obtain empirical counterparts. Insofar as these models provide policy advice, this is often expressed in broad terms and is of a negative nature. As Edgeworth suggested when discussing income taxation and the concept of minimum sacrifice:

> Yet the premises, however inadequate to the deduction of a definite formula, may suffice for a certain negative conclusion. The ground which will not serve as the foundation of the elaborate edifice designed may yet be solid enough to support a battering-ram capable of being directed against simpler edifices in the neighbourhood. (1925, ii, p. 261)

It is important to understand why clear results may not be achieved, or why intuitively appealing results may not be reliable. But patience in the face of such limited results requires a taxonomic turn of mind that is more congenial to economists than those who need to provide direct policy advice or support. The more realistic the model, the more it has to be restricted to highly specific questions.[10]

The theory of optimal income taxation provides an interesting case study. It can be argued that analyses in this tradition have generated valuable insights into the highly complex relationships involved. They have clarified what early investigators referred to as 'the grammar of arguments' and there was no pretence that they were designed as guidance for practical policy advice. The results are largely – as in so many other cases – of a negative nature. Nevertheless, it is possible that the substantial changes in the personal income tax structures of many countries over the last thirty years – in particular, reductions in the number of marginal tax rates and the degree of rate progression – have been influenced by the optimal tax literature.[11] Care must be taken in making such statements. Establishing a clear rationale for each policy action is of course far from straightforward, and the social welfare functions which play a fundamental role in optimal tax theory seldom represent the varied objectives of politicians.[12]

---

[10]For a general discussion of types of tax model, see Creedy (2001a).

[11]For an empirical study of changes in income tax structures in many countries, see Peter *et al.* (2007). They show how the flattening of rate structures spread from higher to lower income countries, and document the use of flat-rate systems in post-communist counties since the mid-1990s.

[12]Indeed, there are familiar theories about the ways in which politicians may wish to exploit the lack of transparency of complex tax structures.

In view of these problems, the aims of this paper are extremely modest. It begins in section 2 by briefly discussing the role of 'rules of thumb' in policy advice, in particular the rule concerning broad bases and low rates, and the adherence to it in New Zealand. Section 3 then considers the simplest possible tax and transfer system, the linear income tax, as a way of illustrating why optimal tax models quickly become intractable. Insights regarding the tax rate schedule from optimal tax modelling are then discussed briefly in section 4. Optimal tax modelling typically relies on small simulation models in which there is negligible population heterogeneity. However, practical policy advice regarding the effects of alternative tax structures can be provided with the use of behavioural microsimulation models in which the full extent of population heterogeneity is represented along with all the details of highly complex tax and transfer systems (compared with the simple stylised forms used in optimal tax analyses).[13] While it does not seem practicable to use such models to produce optimal nonlinear structures, a method of examining marginal reforms is proposed and illustrated using a simple model in section 5.

In the absence of direct policy advice from economic theory, more partial arguments are discussed in subsequent sections. These regard things to keep in mind with nonlinear personal tax structures. In view of recent and planned changes to the New Zealand income tax structure, special attention is given to the role and effectiveness of a top marginal tax rate at higher income levels, in revenue raising and generating redistribution and progressivity. Section 6 examines the effect of a top rate on selected progressivity measures. The complications relating to potential labour supply effects, caused by increasing marginal income tax rates and means-tested transfer payments, are then considered in section 7. Examples are given in section 8 of the variety of welfare effects arising from the introduction of a top marginal tax rate. Brief conclusions are in section 9.

---

[13]Nevertheless, when using such models it is important to be aware of their limitations. In particular, they deal only with the supply side of the labour market and, despite modelling labour supply, have no genuine dynamic element. Furthermore, they deal only with financial incentive effects rather than administrative behaviour and monitoring features designed to reduce moral hazard. On behavioural modelling, see Creedy and Kalb (2007).

# 2 A Rule of Thumb

There are some basic principles which are worth keeping in mind in thinking about tax structures.[14] These include points such that: there is a difference between legal and economic incidence and taxes can be shifted in various ways (including tax capitalisation); large efficiency costs (in the form of excess burdens) can arise even when taxes appear to have little effect on behaviour; incentives matter. This list could easily be extended, but still does not provide strong positive advice. However, many economists (particularly those working in treasury departments) take as a starting point the basic principle that the best taxes are those having a broad base and low tax rate. As a simple 'rule of thumb', this is not a statement derived from a set of fundamental or universal principles, or axioms. It is meant only as a guiding aim which, though not precise, does have valuable content. Departures from the rule require a special case to be made. A broad base, which is obtained by allowing few exemptions and deductions, is of course required in order to achieve a low tax rate, for a given revenue objective. In turn, the need for low rates is generally seen in terms of the efficiency costs of taxation, where appeal is made to the long-established result that the excess burden of a tax is approximately proportional to the square of the tax rate.[15]

Of course, it is not straightforward to explain concepts such as excess burden to politicians, or to persuade them of the importance of the principle of tax capitalisation or of incentive effects. These latter effects influence another term in the excess burden approximation, the compensated demand elasticity – whether of consumption goods or leisure. The confusion between Marshallian and Hicksian (compensated) responses is also common in this context. Nevertheless it is important to continue to stress these concepts and to argue that particular policies require a trade-off to be made in terms of a balance of the perceived benefits of a tax policy against the estimated efficiency costs. It is necessary to make explicit the value judgements involved and the nature of the trade-off between gains and losses. This leads to the concept of the social welfare function which is central to optimal tax theory.

The New Zealand tax system typically scores relatively well when applying the

---

[14]These points differ from the kind of maxim proposed by Smith (1776), in that here the emphasis is on the link from theoretical insights to specific policy advice.

[15]For an indirect tax, the approximation is one half of the product of expenditure on the good, the compensated demand elasticity and the square of the tax rate.

basic 'broad base–low rate' rule of thumb. For example, the goods and services tax, GST, has few exemptions and a single rate, whereas many other countries exempt (or zero rate) goods such as food, and apply a higher rate to 'luxury' goods. Attempts to justify exemptions are usually made on equity grounds: household budget shares for those goods are typically higher for households with low total expenditure. But this kind of selectivity has poor 'target efficiency' qualities, the rich after all do spend more on the same goods, and there are also good administrative arguments for uniformity.[16] Furthermore, the information needed for the imposition of an optimal non-uniform indirect structure are unlikely to be available and equity objectives can be achieved using direct taxes and transfers (which relate directly to the characteristics of the individuals involved). Exceptions are of course made in the case of excises, such as those imposed on tobacco, alcohol and petrol, which are imposed for a variety of paternalistic and externality arguments (along with revenue raising qualities). But these appear to have little overall redistributive effect.[17]

Regarding the income tax, New Zealand has no tax-free threshold and has not (for the employed) a wide range of deductions relating to various expenses, as in many countries, such as Australia where base erosion seems endemic. This all produces in New Zealand a higher tax base and helps to keep income tax rates lower than otherwise.[18] However, in viewing the base, New Zealand cannot be said to have a 'comprehensive' income tax: in particular, most capital gains are not taxed.

The broad base–low rate rule of thumb suggests, through the excess burden relationship with the tax rate, a preference for a flattish income tax rate structure, although of course this does not allow for redistribution objectives and associated trade-offs which may lead to modifications. In New Zealand there are means-tested transfer payments which produce very high effective marginal tax rates for lower-income earners. Furthermore, the abatement (that is, taper, or benefit withdrawal) range for some transfers extends far up the income ranges, so that effective tax rates are significantly above the marginal income tax rates. Furthermore, a higher 'top' marginal income tax rate

---

[16]These administrative arguments have less force in the context of developing economies. The spurious argument is often used that indirect taxes are 'regressive' because high-income households save more, but this is easily dismissed. On the role of exemptions, see Creedy (2001b).

[17]However, excess burdens on some household types can be quite large; on excise taxes in New Zealand, see Creedy and Sleeman (2006).

[18]When making comparisons it is necessary to keep in mind that there are no separate social insurance contributions in New Zealand, as in many other countries.

was reintroduced, after being eliminated in the 1980s reforms which had substantially flattened the income tax rate structure.

A top marginal tax rate is typically introduced on the grounds that it is needed for reducing inequality. However, it is well-known that tax rate progression is not needed for progressivity – indeed a linear tax (basic income–flat tax, or BI–FT) is highly redistributive, such that many individuals face a negative average rate, and the average rate is increasing over the whole income range. Furthermore, the top rate may in practice produce little extra revenue: section 6 discusses this aspect further. The argument is also sometimes made that 'the labour supply elasticity', especially for higher-income 'prime-age' males, is low so that the incentive effects of a top rate are negligible. However, labour supply behaviour is much too complicated to be described by a single elasticity. Further, even if the observed elasticity were low or zero, this could still coexist with large welfare costs and high excess burdens because it is the Hicksian elasticity that matters. Section 8 below returns to this issue.

## 3   The Linear Income Tax

It seems useful, though at first sight perhaps perverse, to begin by ignoring the fundamental question of the 'optimal' *form* of an income tax and transfer system, and to concentrate instead on the narrower question of determining the optimal tax rate in the simplest possible kind of modelling framework, namely a linear (or basic income–flat tax) structure. There is thus only one policy variable, the tax rate, that can be chosen independently, while the other decision variable, the universal or basic income, is determined via the government's budget constraint (discussed in more detail below). Indeed, this type of system is the 'basic workhorse' of optimal tax analysis, on which a number of variants have been built. It serves to demonstrate that the simplest system – it would indeed be impossible to imagine a simpler redistributive tax and economic environment to examine – is very far from being straightforward.

The framework is entirely static, involving a single period partial equilibrium model. Individuals can vary their hours of work continuously without constraint, at a fixed gross wage rate. Each individual's ability level, which determines the wage rate, is fixed and there are no other sources of income other than the untaxed transfer payment and earnings. Apart from the exogenous endowment of ability, there is an endowment of

time (normalised to unity) which is divided between work and leisure. Individuals are considered to maximise utility, which depends only on an index of the consumption of marketed goods and services (where the price index is normalised to unity) and the proportion of time spent working. Apart from differences in abilities, individuals are assumed to be identical: they have the same utility functions and have no non-income characteristics which influence judgements regarding, for example, special 'needs' (hence the basic income can be the same for all individuals).[19]

The framework is thus designed to be the simplest possible model for considering the elements involved in the trade-off between equity and efficiency considerations in setting the parameters of the simplest possible tax and transfer system. It is indeed trivially simple – and yet it turns out to give rise to substantial complexities. One simplification – the concentration on one form of government expenditure – is perhaps worth stressing here, particularly in view of the fact that one important lesson from a welfare economics perspective is that what matters is the overall effect of a multi-tax, multi-expenditure system, rather than the effects of each component taken in isolation. Any revenue collected for non-transfer purposes is considered simply to disappear into a 'black hole', rather than being devoted to activities such as the production of public goods, or education and health services, which would in principle affect individuals' utilities.[20] The extension of an optimal tax approach to the analysis of the *composition* of government expenditure, as well as its magnitude, in this framework presents further complexities which cannot be discussed here, but which of course cannot be ignored in practical discussions.

The cornerstone of the analysis is the government's budget constraint, showing the relationship between the basic income, $b$, and the tax rate, $t$, needed to finance it. This is illustrated in Figure 1 where a minimum tax rate, $t_{\min}$, is needed to finance the non-transfer expenditure. If the tax is below this rate, the system requires the use of a negative basic income, that is a poll tax, to raise the non-transfer revenue. The constraint reflects a 'battle' between 'rate' and 'base' effects of an increase in $t$:

---

[19]Of course, population heterogeneity is a fundamental feature of the 'real world' and plays a crucial role in actual tax policy design (and is discussed further below). As mentioned earlier, the aim here is to demonstrate how awkward even the simplest unrealistic model is. The introduction of taste differences raises the potential for 'incentive compatibility' problems.

[20]Furthermore, such expenditure would be expected to affect indivduals' productivities and thus their wage rates.
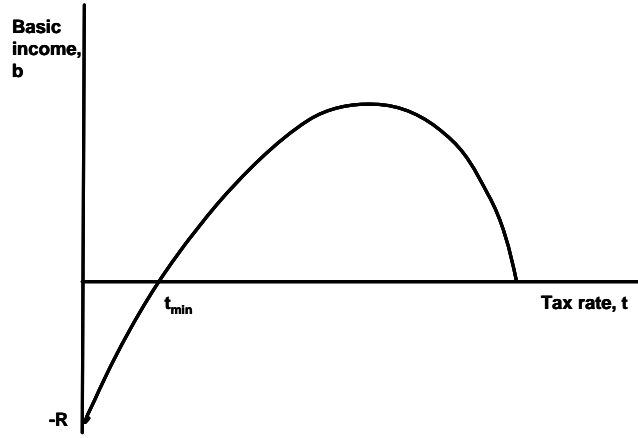
Figure 1: The Government's Budget Constraint

an increase in the rate produces a higher revenue with an unchanged total income (the base), but the base falls because of adverse incentive effects.[21] Initially the 'tax rate' effect dominates the 'tax base' effect, but at some point, where the percentage increase in the tax rate is exactly matched by the percentage reduction in total income, total revenue (and thus $b$) reaches a maximum. If $\bar{y}$ denotes average income and $R$ is non-transfer revenue per head, the budget constraint is:

$$b = t\bar{y} - R \qquad (1)$$

Hence, $\frac{d(b+R)}{dt} = \bar{y} + t\frac{d\bar{y}}{dt}$ and:

$$\frac{t}{(b+R)}\frac{d(b+R)}{dt} = 1 + \frac{t}{\bar{y}}\frac{d\bar{y}}{dt} \qquad (2)$$

Thus the budget constraint becomes flat when the elasticity, $\frac{t}{\bar{y}}\frac{d\bar{y}}{dt} = -1$. This point is clearly reached before $t$ reaches unity.

Although the equation of the budget constraint looks superficially simple, to move from (1) to a relationship between $b$ and $t$ is not simple because $\bar{y}$ is a function of the tax parameters, the preferences of individuals and the form of the distribution of wage rates, $w$. Unless strong assumptions are made regarding utility functions and the wage

---

[21]This supposes that the substitution effect – arising from the cheaper price of leisure – outweighs the income effect of the change in the net wage.

rate distribution, (1) is nonlinear.[22]

The choice of $t$ (and consequently of $b$) must obviously be somewhere along the government's constraint. The issue thus arises of the choice of decision mechanism. The crucial point about the optimal tax framework is that tax choices are considered to be made by a single individual – variously referred to as a policy maker or decision-maker or independent judge – who is entirely disinterested, that is, does not have a personal interest in the outcome. Hence a social welfare, or social evaluation, function is maximised, representing the value judgements of the decision-maker. The welfare function is *not* regarded as representing any kind of aggregation of the views of members of the population.

The role of optimal tax analysis is therefore, from the very outset, to examine the implications of adopting particular value judgements. Even if the model were not so unrealistic, no professional economist could state, on the basis of an optimal tax analysis, that, for example, 'the tax rate should be set at $t = x$'[23] The approach can say only, for example, that 'for the value judgements made explicit in the evaluation function used, the optimal rate turns out to be $t = x$'.

Clearly, a vast range of welfare functions could potentially be examined. In the majority of analyses, the implications of adopting a 'welfarist' evaluation function are examined: this means that welfare, $W$, is considered (where it is also individualistic) to depend on the variables that are of importance for the individuals themselves, that is, their utility levels. Hence, with $n$ individuals $W = W(U_1, ... U_n)$.[24] Each individual is maximising utility, subject to the tax parameters chosen, so indirect utility for each person, $V_i$, can be written as $V_i(t, b)$ where the other parameters (such as those of the direct utility function) have been suppressed. Hence, in principle, it is possible to think

---

[22]Suppose hours worked, expressed as a function of the gross wage, are $h(w) = 0$ for $w \leq w_m$ and $h > 0$ for $w > w_m$. Then if $F(w)$ is the distribution function of wage rates, arithmetic mean earnings are $\bar{y} = \int_{w_m} wh(w) dF(w)$. Even in the special case where $h(w)$ is linear in $w$ (which arises if preferences are Cobb-Douglas), the threshold wage, $w_m$, is itself a function of preferences and the tax parameters.

[23]Furthermore, the addition of the words 'in order to maximise the welfare of society' at the end of that sentence would be meaningless, given the concept of the welfare function used.

[24]The optimal rate depends also on the cardinalisation of utility used in the social welfare function, which is not surprising as any monotonic transformation is equivalent to attaching different weights to utility at different levels. When individuals have relevant non-income characteristics (concerning the size and composition of their household) further value judgements are required in relation to the use of adult equivalent scales and the income unit.

of the welfare function being abbreviated into an expression in terms of the decision variables $t$ and $b$, and this in turn gives rise to social indifference curves. These can be represented in Figure 1 as upward sloping convex indifference curves the nature of which depend on, among other things, the judge's attitude towards inequality.[25] If the judge has extreme aversion to inequality, concern is only to maximise the welfare of the poorest person and indifference curves are horizontal. Such a judge would select the tax rate at which the budget constraint reaches a peak, that is for which the elasticity, $\frac{t}{\bar{y}}\frac{d\bar{y}}{dt} = -1$. The downward sloping section of the constraint is thus irrelevant. As inequality aversion falls, and the judge places relatively greater emphasis on efficiency considerations, the indifference curves become steeper (a given increase in $t$ must be accompanied by a higher increase in $b$ to compensate) and the optimal tax rate falls.[26]

Any information that an increase in an existing tax rate is likely to produce a reduction in total revenue clearly suggests that, whatever the inequality aversion of the judge, the rate is too high. Non-transfer government revenue also plays a role. With zero net revenue, the optimal linear rate is expected to fall as the elasticity of substitution between net income (consumption) and leisure rises, because the substitution effect of a tax rise is likely to be larger (so the budget constraint is flatter). But with positive net revenue, the fall in the optimal rate is eventually reversed as the elasticity of substitution rises. This is because the minimum rate needed to finance the non-transfer revenue increases as the elasticity rises.

So far, the results are quite general. Anyone asking for specific advice about the optimal rate to impose in a linear tax structure would not welcome the information simply that the first-order conditions imply a tangency between the budget constraint and the highest social indifference curve. More structure must be imposed, but even with simple forms for the social welfare function (such as the widely used iso-elastic case of constant relative inequality aversion), the common utility functions, and the wage rate distribution, closed-form solutions are not available.[27] Numerical simulation results, involving iterative methods to solve for the basic income which can be financed

---

[25]In fact, they may not be convex, but the discussion here is not affected.

[26]However, an indifference to inequality – as in the case of classical utilitarian judges – does not necessarily produce vertical indifference curves. The optimal rate is likely to produce a positive $b$ and may also involve the existence of some non-workers.

[27]Special cases of explicit solutions have been produced, some of them involving quasi-linear utility functions.

with any given tax rate, are therefore ubiquitous in the optimal tax literature.

The tangency solution view of first-order conditions for social welfare maximisation therefore provides limited insights. Starting from the same conditions, an alternative approach, which exploits duality theory, can be used to show that the optimal linear tax must satisfy the following condition, first given by Tuomala (1985):[28]

$$1 - \frac{\tilde{y}}{\bar{y}} = \left| \eta_{\bar{y},t} \right| \tag{3}$$

where $\left| \eta_{\bar{y},t} \right|$ represents the absolute value of the elasticity of average earnings with respect to the tax rate, and has already been discussed in the context of the shape of the government budget constraint. The term $\tilde{y}$ is a welfare-weighted average of earnings, where the weights are equal to $v_i / \sum_{i=1}^{n} v_i$, with $v_i = \frac{\partial W}{\partial V_i} \frac{\partial V_i}{\partial b}$. The latter is the weight attached by the social welfare function to an addition to person $i$'s income (from an increase in the basic income).

The left-hand side represents the proportional difference between the welfare-weighted mean of earnings and the arithmetic mean, which can be interpreted as a measure of the inequality of gross earnings. For example, extreme inequality aversion on the part of the judge would attach the highest value of 1 to this inequality measure, whatever the statistical earnings differences, and as seen above $\left| \eta_{\bar{y},t} \right| = 1$ represents the highest point of the government budget constraint. This might be thought to provide further insights in view of the fact that both sides of the equality in (3) deal with earnings, which are clear 'empirically relevant counterparts', unlike some other components of the model. Discussion can proceed in term of a diagram showing the profiles of each side of the equation as $t$ varies, and the way in which they, and hence their point of intersection, are likely to shift as basic assumptions are changed. However, this expression does not provide a closed-form solution. Even the welfare weights, $v_i$, depend in general on the tax parameters.

The optimal policy perspective of choice by an independent decision-maker clearly differs from public choice models which consider particular aggregation mechanisms, such as majority voting.[29] However, some interesting comparisons can be made between the different approaches. For example, in the simple majority voting framework, a

---

[28] See Appendix B for further details.

[29] The first problem in considering voting schemes is that indivduals' preferences over $t$ are not in general single peaked.

voting equilibrium is known to exist, despite the existence of double-peaked preferences regarding the tax rate, if there is 'hierarchical adherence', such that the ordering of individuals by their income is independent of the tax rate. In this case the median voter theorem can be invoked and the median here is the person with median wage. Equation (3) applies also in this case, except that the welfare-weighted mean, $\tilde{y}$, is replaced by the median. In the case of stochastic voting, the voting equilibrium is generated by maximisation of a function that may appear to look something like a social welfare function, involving the weighted arithmetic mean of utilities, or the weighted geometric mean, depending on the precise details. Again, unless special assumptions are made, interior solutions are not available.

The absence of closed form solutions for this, the simplest possible model of a population group and a tax structure, is not of course the property that limits its value in providing policy advice. A wide range of simulation analyses can easily be carried out. The problem is that it is so far removed from reality, in order to illustrate as simply as possible the nature of the trade-offs involved and the underlying structure of such models, as to be useless as a practical tool. It was never intended to provide such a tool. However, its very simplicity serves to demonstrate that more realistic models are likely to be completely intractable, which raises the question of how to proceed. Before considering this question, the following section turns to the initial question raised and ignored in starting from a linear tax – that of the form of optimal tax function.

# 4   Optimal Tax Structures

In view of the difficulties of examining even the simplest structure, it cannot be expected that a more general analysis of the optimal rate structure would generate clear results. The optimal tax problem, which in general asks what tax structure maximises a specified evaluation, or social welfare function, is known from the work of Mirrlees (1971) to give rise to a problem in the calculus of variations. In view of the nonlinearities involved, solutions generally require numerical simulation methods, unless strong simplifying assumptions are made.[30] Unfortunately, it turns out that there are few general results available – the optimal structure depends on the nature of the social

---

[30]Explicit solutions include Atkinson and Stiglitz (1980), Deaton (1983) and examples given in Hindriks and Myles (2006).

welfare function examined as well as the wage rate distribution and the nature of preferences. Earlier work suggested that, for a range of assumptions, the optimal structure is approximately linear.

But in fact it is not difficult to produce models in which different tax structures are optimal.[31] A basic issue relates to the social welfare function itself. Much of the literature is 'welfarist' in that the welfare function is specified as some function of (indirect) utilities, that is, terms which matter to individuals.[32] But since the fundamental aim is to consider the implications of adopting alternative value judgements, other 'non-welfarist' forms cannot be ruled out. For example, social welfare may be based solely on an income-based measure of poverty.[33] These can give quite different results.

Hence the general treatment of optimal tax structures yields very few clear results. The most unambiguous results can in fact easily be established.[34] First consider Figure 2 which shows a tax function in a diagram with net income on the vertical and gross income on the horizontal axis. This displays a range AB where the marginal tax rate (equal to 1 minus the slope of the tax schedule) is greater than 100 per cent. This range is clearly irrelevant, since indifference curves relating net income and gross earnings are upward sloping and convex: an increase in gross earnings involves an increase in hours worked, which must be compensated by an increase in net income (consumption). Hence, without loss AB can be replaced by a marginal tax rate of 100 per cent. It can also be shown that negative marginal tax rates (in contrast with average tax rates) can be ruled out.[35]

A further result states that the marginal tax rate on the highest income should be zero. In Figure 3 consider the tax function AB, where the person with the highest wage rate reaches a tangency position at C. If the tax function is changed to ACD, where

---

[31] A small selection of examples includes those discussed by Diamond (1988), Chang (1994), Hashimzade and Myles (2004), Myles (1999), Saez (2001) and Tuomala (2006).

[32] Here there is a problem relating to the cardinalisation of utility functions, discussed further below.

[33] For further discussion see Kanbur, Keen and Tuomala (1994) and for a broader treatment of non-welfarist objectives, see Kanbur, Pirttilä and Tuomala (2004). Non-welfarist objects may go further than simply attaching no value to leisure, in that they may prefer to encourage labour supply (whereas in a welfarist approach the existence of non-workers is acceptable in an optimal structure): see, for example, Besley and Coate (1992).

[34] For other reviews of policy implications of optimal tax theory see, for example, Stern (1984), Heady (1993), Tuomala (1995) and Bradbury (1999). For a more critical discussion, see Slemrod (1990).

[35] However, this is not true of a social welfare function based on income-poverty, where a negative marginal rate can result for the lowest earners.
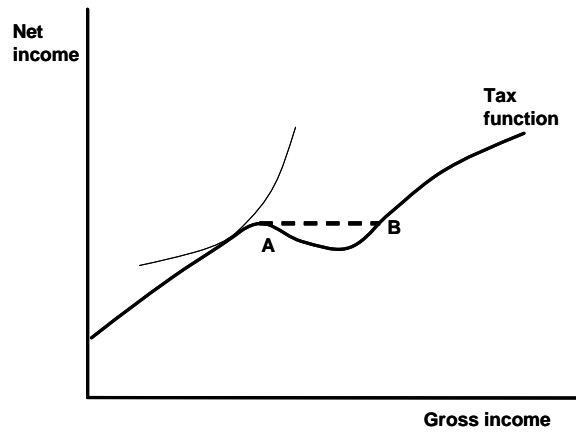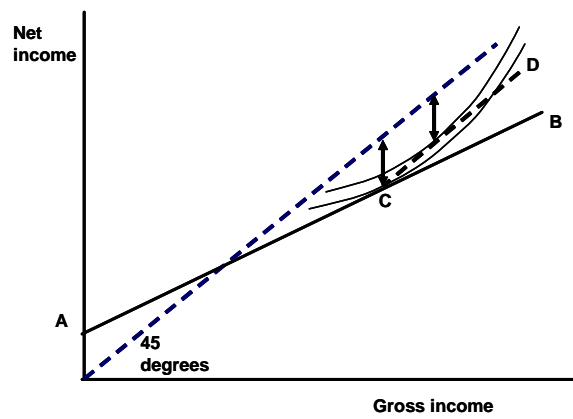
Figure 2: Maximum Tax Rate



Figure 3: Tax Rate on Top Income

16

CD is parallel to the 45 degree line, the individual then faces a zero marginal rate on any extra income earned. This induces a movement to a new tangency on a higher indifference curve. The total tax revenue is unchanged and no one else is affected: hence the richest person is better off and the non-zero top marginal rate cannot have been optimal, for any Paretian welfare function. However, this result is of no practical relevance as there is no way to determine just where the rate should become zero.

One approach has been to consider piecewise-linear tax functions with just two or three rates. This allow for consideration of the question of whether marginal tax rates should be higher for those with relatively low earnings. Such higher marginal rates arise from the means-testing of transfer payments. The resulting non-convexity of budget sets facing individuals can give rise to complex labour supply behaviour, as discussed below. Means-testing is preferred by those who advocate 'target efficiency' as the criterion by which schemes should be judged. Numerical analyses using 'welfarist' social welfare functions show that in a very wide range of situations, the evaluation function is increased by a shift to lower taper rates and a flatter rate schedule.[36] But such results inevitably involve special cases in highly simplified models with little of the considerable population heterogeneity that is observed in practice. They therefore cannot provide a strong basis for policy advice. Indeed, assumptions giving rise to means-testing in an optimal structure are given in Diamond (1998). An emphasis on 'workfare', designed largely to encourage positive labour supply, rather than 'welfare', can also lead to high marginal rates imposed on low earners, as shown by Besley and Coate (1992).

Despite the dearth of general results, the optimal tax literature does support a broad argument for relatively low rates (the second part of the broad base–low rate rule of thumb), even with high inequality aversion, especially compared with the top rates operating in many countries in the 1970s. A small degree of substitution between leisure and net income imposes strong constraints on the government's ability to redistribute income.[37] As suggested earlier, the finding that, even in simplified models, optimal

---

[36] See, for examples, simulation results reported in Creedy (1998a). For wide-ranging discussions of means-testing, see also Atkinson (1995) and Bradbury (1999).

[37] In the linear tax case, non-transfer government revenue also plays a role. With zero net revenue, the optimal linear rate falls as the elasticity of substitution (between consumption and leisure) rises. With positive net revenue, the optimal rate falls and then rises as the elasticity rises (because the minimum rate needed to finance the non-transfer revenue rises as the elasticity rises).

rates are not high may perhaps have had some influence on policy since the late 1970s when the very high top marginal rates existing in some countries started to be reduced.

A behavioural tax microsimulation model, which encapsulates the actual degree of population heterogeneity and most of the detail of actual tax and transfer systems, does not provide a convenient vehicle for producing an optimal tax structure. However, its advantage of providing information about changes in taxes may be exploited to examine the optimal direction of reforms. That is, such a model may be used to explore the directions of small changes in tax parameters which improve a specified social welfare function. An illustration of the kind of analysis which would be possible, but using a much simpler model, is provided in the following section.

# 5  Marginal Income Tax Reform

Consider the more realistic problem of how to move towards an optimal structure by a process of marginal tax reform. This has previously been examined in the context of indirect taxation, involving fixed incomes but differential taxation on a range of goods, where $x_{ji}$ is the consumption of the $i$th taxable good by the $j$th household, and $t_i$ is the unit tax on good $i$.[38] A standard result shows that an increase in social welfare, $W$, resulting from a change in the $i$th tax rate is:

$$\frac{\partial W}{\partial t_i} = -\sum_{j=1}^{n} v_j x_{ji} \tag{4}$$

where $v_j$ is the the social value of additional consumtpion by household $j$ (see Appendix B). Furthermore, aggregate tax revenue, $R$, from indirect taxes on $K$ goods is given by:

$$R = \sum_{j=1}^{n} \sum_{k=1}^{K} t_k x_{jk} \tag{5}$$

and $\frac{\partial R}{\partial t_i}$ can be expressed in terms of (Marshallian) demand elasticities, at observed consumption levels, and expenditures. It is then possible to obtain values of the ratio, $-\frac{\partial W}{\partial t_i} / \frac{\partial R}{\partial t_i}$, the marginal welfare cost of raising tax rate $i$, for each good and thus to

---

[38]On marginal indirect tax reform see, for example, Ahmed and Stern (1984), Madden (1996) and Creedy (1999)).

determine the required direction of changes, since an optimal system is characterised by an equi-marginal condition.

In considering income tax structures, a similar type of approach could be adopted to examine marginal adjustments to a piecewise linear system. The social welfare function $W$ is in this case a function of a set of marginal income tax rates and income thresholds (as well as summarising the value judgements involved). Similarly, total net revenue, $R$, is a function of the same set of tax parameters. It can be shown (see Appendix B) that if $t_i$ now represents the $i$ marginal rate, and $y_j$ is person $j$'s earnings, a similar result to that given above holds, whereby $\frac{\partial W}{\partial t_i} = -\sum_{j=1}^{n} v_j y_j$. For any piecewise linear tax function, the tax paid by household $j$, $T(y_j)$, can be expressed as:

$$T(y_j) = t_k(y_j - a'_k) \tag{6}$$

where $y_j$ falls into the $k$th tax bracket and $a'_k$ is a function of all the thresholds (denoted $a_1, ..., a_K$) and marginal rates. The total tax revenue raised by those in the $k$th bracket, $T_k$, is thus:

$$T_k = t_k \sum_{j=1}^{n} (y_j - a'_k) \tag{7}$$

It would be possible to combine this kind of approach with a number of simplifying assumptions regarding elasticities of earnings with respect to tax rates and the proportions of earnings and people above a top rate, to consider, say, whether (for a given set of tax thresholds) a top marginal rate should be reduced or increased.[39] This appears to be an attractive route because, unlike the usual approach to optimal tax modelling, the conditions can be expressed in terms of (what appear to be) empirically observable counterparts such as elasticities. However, given the considerable complexities introduced by nonlinear budget constraints – unlike the consumption tax case where linear pricing is a reasonable assumption – any clear results need strong assumptions and could only be regarded as illustrative rather than of practical relevance. Nonlinear budget sets make it difficult to generalise regarding labour supply responses and welfare changes even for workers with similar preferences and with relatively simple tax structures. In practice, populations display considerable heterogeneity in preferences and household circumstances, and tax and transfer structures are extremely complex.

---

[39]For an extensive discussion on the use of elasticities to derive optimal tax rates, with emphasis on the choice of a top tax rate, and references to earlier literature, see Saez (2001).

However, given a behavioural tax microsimulation model, there is no need to make simplifying assumptions about elasticities of earnings with respect to tax rates. The full extent of heterogeneity can be captured in behavioural microsimulation models. Such models could not realistically be used to produce an optimal tax structure, even for a clearly specified social welfare function. But for marginal reforms, the required changes, from an initial actual tax structure, could be obtained numerically. Starting from the actual tax structure, and considering small changes in a range of tax parameters, it is possible to use a microsimulation model to obtain values of welfare and revenue changes, denoted $\Delta W$ and $\Delta R$ respectively, for each tax parameter in turn. The marginal welfare costs, that is the change in welfare per dollar of extra revenue, $\Delta W / \Delta R$, should be similar for all tax parameters in an optimal system and so the direction of an optimal reform is indicated by relative orders of magnitude of these ratios.

In the absence of a New Zealand behavioural microsimulation model, the present section illustrates – nevertheless using a very simple model – the way in which such a model could be used to examine optimal marginal income tax reforms. The first stage in illustrating the approach is to specify a hypothetical population. Two sets of populations were examined – single individuals with no dependents and workers in couple households with children (each couple is in fact modelled as if there is just one potential worker). The tax and transfer structure for each group is fully specified by a set of income thresholds and marginal rates, $t_i$, which apply above those thresholds, giving rise to a piecewise-linear budget line relating net income and hours of work. Each linear section, $i$, has a slope equal to the net wage $w = w_g (1 - t_i)$, where $w_g$ is the gross wage rate, and when extended to the intercept where hours of work are zero, a non-wage virtual income of $\mu_i$.

Each group is considered to face the tax and transfer structures shown in Table 1. In the table, the tax rates and income thresholds (expressed in terms of weekly income) are given, along with the $\mu_i$. Only $\mu_1$ is in fact specified and the other information is used to produce the other virtual incomes.[40] These effective tax schedules are chosen to be similar to those in New Zealand (see IRD, 2009) although the marginal rate of 0.85 applied after \$200 is less than the very high rate of almost 100 per cent which actually applies as a result of means-testing. The extension of means-tested benefits to higher-income ranges means that for couples with children, the marginal tax rate

---

[40]It can be shown that, for $i = 2, ..., n$, $\mu_i = \mu_{i-1} + y_i (t_i - t_{i-1})$.

Table 1: Hypothetical Tax Structure

|   | Singles | | | Couples with children | | |
|---|---|---|---|---|---|---|
|   | Threshold | $t_i$ | $\mu_i$ | Threshold | $t_i$ | $\mu_i$ |
| 1 | 0 | 0.15 | 150 | 0 | 0.15 | 300 |
| 2 | 200 | 0.25 | 107 | 200 | 0.85 | 440 |
| 3 | 650 | 0.35 | 235 | 450 | 0.25 | 170 |
| 4 | 1000 | 0.40 | 285 | 650 | 0.55 | 365 |
| 5 | – | – | | 1000 | 0.60 | 415 |
| 6 | – | – | | 1400 | 0.40 | 135 |

actually drops from 0.60 to 0.40 at a weekly income of $1,400. Calculation of labour supply must allow for the possibility of multiple local optima, of the kind disucussed in section 7 below.

The population heterogeneity in each case is such that the joint distribution of wage rates and preference parameters ($\alpha$, the coefficient on net income) in Cobb-Douglas utility functions is jointly lognormal, with a correlation of -0.75. The arithmetic mean of $\alpha$ is set at 0.75 for the case of both singles and couple workers, while the mean of log-wages is 2.7 and 3.0 for singles and couples respectively. The variance of logarithms of $\alpha$ is 0.10 for singles and 0.15 for couple workers. These values were chosen following experimentation in which the distributions of earnings (for simulated populations of 10,000 in each case) were similar to those for New Zealand in terms of the pattern of modes and antimodes. Suppose that the social welfare function, $W$, for each group takes the familiar additive individualistic Paretean form and is expressed as a function of utilities, $U_j$, with an inequality aversion coefficient of $\varepsilon = 0.5$:[41]

$$W = \frac{1}{1 - \varepsilon} \sum_{j=1}^{n} U_j^{1-\varepsilon} \qquad (8)$$

Tables 2 and 3 show marginal changes for each marginal tax rate, for single individuals and workers in couple households. The change in each marginal tax rate involved a reduction of 0.02, that is 2 percentage points, with thresholds unchanged.

---

[41]A feature of optimal tax models, mentioned briefly above, is that results depend on the cardinalisaton of utility functions used. The use of money metric utility, with current 'prices' as reference prices, was suggested by Creedy (1998b). The calculation of a social welfare function using money metric utility in a behavioural microsimulation model based on discrete hours random utility modelling of labour supply is examined by Creedy, Herault and Kalb (2008).

Each table shows the welfare and revenue changes, along with the ratios $\Delta W/\Delta R$, and the elasticities of welfare with respect to revenue, $\eta_{W,R}$. In the case of couples, it can be seen that a reduction in the 5th marginal rate of two percentage points actually leads to an increase in tax revenue: this marginal rate is in practice affected by both means testing of transfers and the top marginal income tax rate.

Table 2: Marginal Tax Reforms: Singles

| $k$ | $\Delta W$ | $\Delta R$ | $\Delta W/\Delta R$ | $\eta_{W,R}$ |
|---|---|---|---|---|
| 1 | 1.1974 | -4.5569 | -0.2628 | -0.0646 |
| 2 | 1.9103 | -7.9664 | -0.2398 | -0.0589 |
| 3 | 0.5746 | -2.4185 | -0.2376 | -0.0584 |
| 4 | 0.3560 | -0.6511 | -0.5467 | -0.1343 |

Table 3: Marginal Tax Reforms: Couples

| $k$ | $\Delta W$ | $\Delta R$ | $\Delta W/\Delta R$ | $\eta_{W,R}$ |
|---|---|---|---|---|
| 1 | 1.1489 | -5.6348 | -0.2039 | -0.0930 |
| 2 | 0.9349 | -2.6391 | -0.3543 | -0.1616 |
| 3 | 0.7277 | -2.3056 | -0.3156 | -0.1440 |
| 4 | 0.7248 | -1.7888 | -0.4052 | -0.1848 |
| 5 | 0.3486 | 0.2484 | 1.4032 | 0.6400 |
| 6 | 0.395 | -0.6345 | -0.6225 | -0.2839 |

In these illustrative examples, the optimal reform is clear: a reduction in the top income tax rate produces the greatest 'payoff' in terms of social welfare. Such illustrative exercises cannot be used as a basis for policy advice in practice, but they do highlight the way in which a behavioural microsimulation model could be exploited.

# 6    Redistribution and Progressivity

It has been mentioned that a common argument for applying a higher marginal tax rate to 'top incomes' is that it is needed for redistribution and progressivity. But the effects of a tax structure cannot be evaluated by looking at the marginal rate structure alone, as they depend fundamentally on the nature of the income distribution. This

section provides some illustrative calculations of the effects of income tax structures under different income distributions.[42]

Consider the effect of income taxation alone. Table 4 shows thresholds and rates in New Zealand for the years 2006-7 and 2008-9. The rate structures are similar (except for the lowest rate) but thresholds are higher in 2008-9, reflecting the growth of incomes.

Table 4: New Zealand Income Tax Structures

| 2006-7 | | 2008-9 | |
|---|---|---|---|
| Threshold | Tax rate | Threshold | Tax rate |
| 0 | 0.15 | 0 | 0.125 |
| 9,500 | 0.21 | 14,000 | 0.21 |
| 38,000 | 0.33 | 40,000 | 0.33 |
| 60,000 | 0.39 | 70,000 | 0.39 |

On the assumption that pre-tax incomes are lognormally distributed as $\Lambda\left(\mu, \sigma^2\right)$ were $\mu$ and $\sigma^2$ are the mean and variance of logarithms, suppose $\sigma^2 = 0.8$.[43] This gives a Gini inequality measure of 0.4589, and an Atkinson inequality measure (with inequality aversion of 0.5) of 0.1775.[44] Tables 5 and 6 report, for alternative values of $\mu$, summary measures of the distribution of pre-and post-tax income. The upper part of each table shows the results for the actual structure and the lower part gives results obtained without the top marginal rate. The variable $\bar{x}$ is arithmetic mean pre-tax income, while $G$, $K$, $L$ and $g$ measure respectively the Gini inequality of post-tax income, the Kakwani measure of tax progressivity, the reduction in the Gini measure when moving from pre- to post-tax incomes (that is, the Reynolds-Smolensky measure of redistribution), and the overall effective tax ratio.[45]

As expected, the tax structure (with or without a top rate) is more progressive and more redistributive as $\bar{x}$ increases, since more people are in the higher marginal rate ranges (though clearly this effect is reduced when the top rate of 0.39 is eliminated).

---

[42]It is therefore concerned more with statistical issues than economic theory insights.

[43]This is obviously a simplifying assumption for illustrative purposes. In view of the New Zealand tax and benefit structure, a further (small) mode is generated at the lower end of the distribution of taxable income, and other small modes and antimodes appear as a result of labour supply effects discussed above.

[44]Refer to values given in Tsy working paper p.10 gives average gross taxable income 2006/7 of $33,503 with $G = 0.464$ and $g = 0.23$. IRD paper, average 'wage' of $46k.

[45]On these measures see, for example, Creedy (1996).

However, for any given value of $\bar{x}$, the changes resulting from the elimination of the top marginal rate are very small. For example, considering the changes in the Gini measures of post-tax income, these are only affected at the third decimal place. The progressivity measures change by less than 0.01, and the effects on revenue, measured by $g$, involve reductions of less than one percentage point.

Table 5: 2006-7 Tax Structure

| Income tax only | | | | | | Income tax with MIG | | | |
|---|---|---|---|---|---|---|---|---|---|
| $\mu$ | $\bar{x}$ | $G$ | $K$ | $L$ | $g$ | $y_m = 12k$ | | $y_m = 15k$ | |
| | | | | | | $G$ | $g$ | $G$ | $g$ |
| With top marginal tax rate | | | | | | | | | |
| 9.8 | 26903 | .4389 | .1031 | .0300 | .226 | .3373 | .152 | .2875 | .101 |
| 10.0 | 32860 | .4351 | .1079 | .0338 | .239 | .3669 | .194 | .3269 | .160 |
| 10.2 | 40134 | .4318 | .1101 | .0372 | .252 | .3885 | .226 | .3582 | .204 |
| 10.3 | 44356 | .4304 | .1101 | .0386 | .259 | .3966 | .240 | .3710 | .222 |
| Without top marginal tax rate | | | | | | | | | |
| 9.8 | 26903 | .4436 | .0906 | .0253 | .218 | .3425 | .145 | .2927 | .094 |
| 10.0 | 32860 | .4415 | .0924 | .0274 | .229 | .3737 | .184 | .3340 | .150 |
| 10.2 | 40134 | .4400 | .0919 | .0289 | .239 | .3971 | .213 | .3672 | .191 |
| 10.3 | 44356 | .4395 | .0908 | .0294 | .245 | .4062 | .225 | .3809 | .207 |

Table 6: 2008-9 Tax Structure

| Income tax only | | | | | | Income tax with MIG | | | |
|---|---|---|---|---|---|---|---|---|---|
| $\mu$ | $\bar{x}$ | $G$ | $K$ | $L$ | $g$ | $y_m = 12k$ | | $y_m = 15k$ | |
| | | | | | | $G$ | $g$ | $G$ | $g$ |
| With top marginal tax rate | | | | | | | | | |
| 9.8 | 26903 | .4355 | .1302 | .0334 | .204 | .3402 | .134 | .2923 | .086 |
| 10.0 | 32860 | .4316 | .1332 | .0373 | .219 | .3682 | .177 | .3303 | .145 |
| 10.2 | 40134 | .4281 | .1333 | .0408 | .234 | .3884 | .210 | .3600 | .189 |
| 10.3 | 44356 | .4267 | .1321 | .0422 | .242 | .3958 | .224 | .3721 | .207 |
| Without top marginal tax rate | | | | | | | | | |
| 9.8 | 26903 | .4391 | .1201 | .0298 | .199 | .3441 | .129 | .2963 | .081 |
| 10.0 | 32860 | .4367 | .1204 | .0322 | .212 | .3736 | .169 | .3358 | .137 |
| 10.2 | 40134 | .4350 | .1177 | .0339 | .224 | .3955 | .199 | .3673 | .179 |
| 10.3 | 44356 | .4345 | .1153 | .0344 | .230 | .4039 | .212 | .3803 | .195 |

For comparison, the right hand side of each table shows the effect on the Gini inequality measure of post-tax income and the overall average tax rate of introducing

a minimum income guarantee (MIG), whereby those with after tax income below $y_m$ have their income brought up to $y_m$. Such transfer payments clearly have a much larger effect than marginal rate progression applied to the higher income ranges.

# 7   Labour Supply and Marginal Rate Changes

This section turns to particular aspects of nonlinear tax structures and their implications for individual labour supply behaviour. These features are well established, but are sometimes forgotten when discussing piecewise-linear tax schedules, particularly the implied labour supply elasticities. They are also relevant when considering welfare effects of tax changes, as in the next section. The discussion applies to a single individual who is able continuously to vary the number of hours worked in one job, and who faces a fixed gross hourly wage rate. The net wage depends on the chosen position on the budget constraint and is therefore, like the number of hours worked, endogenous.

With a piecewise-linear budget constraint any interior (or tangency) solution and corner solution can be regarded as being generated by a simple linear constraint of the form:

$$c = wh + \mu \tag{9}$$

In the case of tangency solutions, $w$ and $\mu$ represent the appropriate net wage rate and 'virtual' income respectively. Virtual income is the intercept (where $h = 0$) corresponding to the relevant segment of the budget constraint and associated net wage; it is therefore distinct from actual non-wage income. In the case of a corner solution, the appropriate virtual income is defined as the value generated by a linear constraint having a net wage, the virtual wage, equal to the slope of the indifference curve at the kink.[46]

The effect of an increase in the gross wage is shown in Figure 4. The constraint ABC has a kink at B, reflecting the presence of an earnings threshold where the marginal effective tax rate increases. The budget set is convex: a straight line joining any two points is associated with a feasible position. The hours level at which the earnings threshold is reached depends on the wage rate. For a higher wage rate, the budget constraint pivots to $AB'C'$, and the kink point $B'$ moves to the left of B. A lower hours

---

[46]The concept of the virtual wage is the same as that of the virtual price used in the theory of rationing.
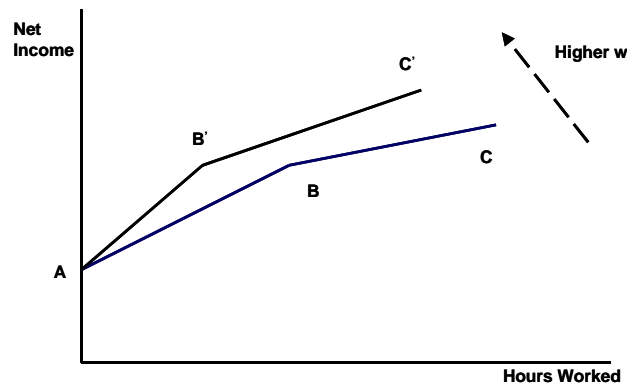
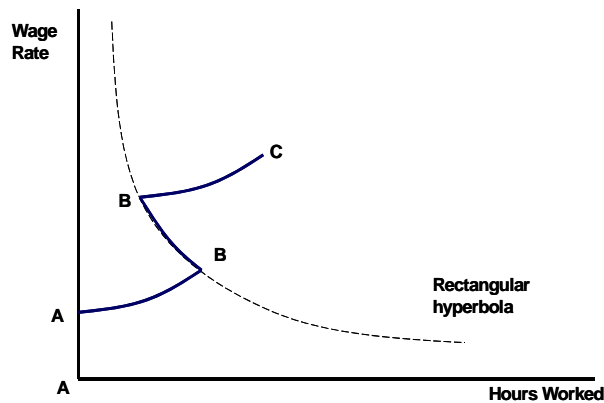Figure 4: A Convex Budget Set



Figure 5: Labour Supply Curve

level is required, at the higher wage, to reach the earnings threshold where the marginal rate increases; gross and thus net income remain constant at the kink.

At very low wages, utility maximisation gives rise to the corner solution at A. When the wage rate exceeds some level (as the section AB of the constraint pivots about A), the individual moves to a tangency position. Increases in the wage induce higher labour supply until the gross earnings threshold is reached at which the marginal effective tax rate increases. A characteristic of this kind of kink in the budget constraint is that the individual 'sticks' at the corner for a range of wage rates. Gross earnings remain constant as the wage rate rises over a range, while the associated hours level falls. Eventually, for a sufficiently high wage rate, the individual moves to a tangency along

the range BC of the constraint. This implies that, in a graph of hours worked plotted against the wage rate, the hours of work would follow a rectangular hyperbola over the relevant range, as shown in Figure 5. This property is entirely general and applies to any kink in the budget constraint associated with an increase in the marginal effective tax rate at a threshold level of earnings.

This property may suggest that some 'bunching' of individuals around the threshold in the distribution of gross earnings. This kind of phenomenon is nevertheless only observed in particular cases. A tax threshold need not produce a 'spike' in the earnings distribution, and modes may in practice correspond to tangency solutions. Hence, the distribution of earnings need not necessarily provide any information about the extent of the labour supply effects of taxation.[47] Clearly, it makes little sense to attempt to describe the labour supply function in terms of a single elasticity. Even if the ranges AB and BC have a constant elasticity, large variations occur at the kink points, and of course the elasticity changes sign twice.

An example of a budget constraint with a means-tested benefit is given in Figure 6, as ABC. The benefit is withdrawn until it is exhausted at B, when the individual only pays income tax.[48] Here the budget set is non-convex. This raises the possibility of an indifference curve being simultaneously tangential to the two sections of the constraint, for a particular wage rate; this is shown in Figure 6 by the two tangencies at J and K. A small increase in the wage rate would therefore produce a discrete jump in hours worked from J to K. The associated labour supply function is shown in Figure 7. An alternative possibility is that, with a very flat range AB, the individual may jump directly from A to some point on BC. Hence, means-testing is liable to give rise to gaps, or antimodes, in the earnings distribution, though depending on population heterogeneity, these may not necessarily be observed. Again, no single elasticity describes labour supply behaviour. In practice, budget constraints contain several ranges with increasing and decreasing marginal rates, so labour supply cannot be characterised by a convenient smooth schedule. With these complications in mind, the following section considers welfare changes arising from changes to the tax structure.

---

[47]For further treatment of this point, see Creedy (2001c).

[48]It is assumed that integration of the benefit and tax systems avoids a discontinuity, though this is not always achieved in practice.
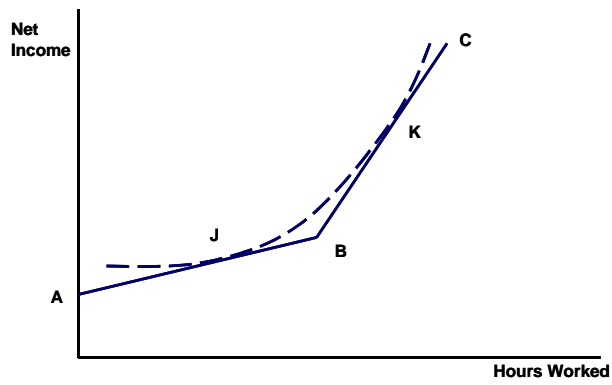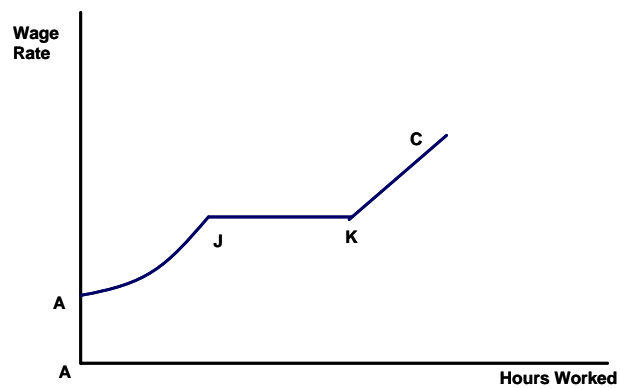
Figure 6: Means-Testing: A Non-Convex Budget Set



Figure 7: Labour Supply with a Non-Convex Budget Set

28

# 8 Welfare Changes and a Top Marginal Rate

This section illustrates the variety of welfare changes arising from the introduction of a top marginal tax rate. It demonstrates that, even for individuals with similar tastes, welfare effects can vary substantially. The welfare changes are measured as follows. Welfare changes are defined in terms of the individual's expenditure function, which is usually written as $E(w, U)$: it represents the minimum income needed to achieve a specified utility level, $U$, at a net wage, $w$. The approach used here is the standard method of obtaining welfare changes, which is sufficient for illustrative purposes.[49] The income involved in the expenditure function is in this context a 'full income' measure, $M$, defined as follows. If $T$ is the maximum number of hours available for work, $M = wT + \mu$, where $\mu$ is virtual non-wage income, that is the value of net income for zero hours of work (for the relevant section of the budget constraint). The expenditure function is thus written as $M(w, U)$ rather than $E(w, U)$ and is obtained by inverting the indirect utility function.[50]

A change to the tax and transfer system may arise from changes in the effective marginal rates, the number of thresholds (and therefore nonlinear segments of the budget constraint), or the gross income thresholds. This may (but need not necessarily) change the individual's optimal labour supply and produce a change in the endogenous net or virtual wage rate and virtual income. The welfare effect of such a tax change is complicated by the fact that a change in either the net or virtual wage affects both the price of leisure and the value of full (or virtual) income. It is therefore useful to decompose the welfare effect into the price and income effects.

For convenience, write the expenditure function in terms of full income, $M = \mu + wT$, and consider a change in the tax system such that the net wage and full income for an individual change from $w^0$ and $M^0$ to $w^1$ and $M^1$. The equivalent variation can be decomposed as follows: and:

$$
\begin{aligned}
EV &= \left\{M^1 - M\left(w^0, U^1\right)\right\} + \left\{M^0 - M^1\right\} \\
&= EV_{\Delta w} + EV_{\Delta M}
\end{aligned} \tag{10}
$$

---

[49] However, Creedy and Kalb (2005) showed how such changes really need to consider the complete range of the nonlinear budget constraint facing each individual.

[50] This inversion may not always be possible analytically. The expenditure function can be expressed alternatively in terms of corresponding virtual incomes, giving the same results for welfare changes.

where the terms $EV_{\Delta w}$ and $EV_{\Delta M}$ relate respectively to the components arising from the reduction in the price of leisure over the range covered by the top rate and the fall in full income. Thus, $EV_{\Delta w}$ measures the welfare gain because leisure hours have become less costly in terms of after-tax income foregone, and $EV_{\Delta M}$ captures the loss in full income via the reduced net wage, $w$, associated with the tax increase. The terms in (10) are defined so that a positive value indicates a reduction in welfare, while a negative value implies a welfare gain. The absolute value of the first term in curly brackets in the above expressions corresponds to an area to the left of a Hicksian (compensated) leisure demand curve between appropriate 'prices' of leisure.

Examples of labour supply and welfare changes for hypothetical single individuals are shown in Table 7. The tax change examined involves a movement from the appropriate structure without a top marginal rate to the one with such a rate, indicated in Table 1. Each individual is assumed to have Cobb-Douglas preferences, with a coefficient on net income of $\alpha = 0.75$: welfare changes for this case are discussed in Appendix A. The only difference between the individuals relates to their gross wage rates. Of course, many individuals – those who are not affected by the tax change – experience no welfare changes.

Table 7: Examples of Single Individuals

| | Initial structure | | With top rate | | Welfare change: $\Delta W$ | | | Tax |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| $w_g$ | $h$ | Position | $h$ | Position | $EV_{\Delta w}$ | $EV_{\Delta M}$ | $EV$ | change: $\Delta R$ |
| 18.4 | 55.09 | s3 | 54.35 | c4 | -11.63 | 11.80 | 0.17 | -4.77 |
| 25.0 | 56.38 | s3 | 55.25 | s4 | -30.02 | 50.00 | 19.98 | 9.14 |
| 30 | 56.99 | s3 | 56.04 | s4 | -34.87 | 70.00 | 35.13 | 24.14 |

Table 8: Examples of Couples with Children

| | Initial structure | | With top rate | | Welfare change: $\Delta W$ | | | Tax |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| $w_g$ | $h$ | Position | $h$ | Position | $EV_{\Delta w}$ | $EV_{\Delta M}$ | $EV$ | change: $\Delta R$ |
| 22.6 | 45.23 | s4 | 44.25 | c5 | -14.11 | 14.31 | 0.20 | -12.25 |
| 25.0 | 51.89 | s4 | 49.62 | s5 | -36.31 | 50.00 | 13.69 | -19.10 |
| 28.2 | 54.61 | s5 | 53.61 | s6 | -36.18 | 62.80 | 26.62 | 15.69 |
| 30.0 | 58.91 | s5 | 58.12 | s6 | -31.83 | 70.00 | 38.17 | 28.94 |

The hypothetical individual with a gross wage of \$18.4 per hour reduces labour

supply from 55.09 hours per week (on the fourth section of the initial relevant budget constraint) to 54.35 hours, which is at the corner solution introduced by the addition of the top marginal rate. The reduction in hours worked is small, at less than one hour, and the welfare change is $EV = 0.17$ is small, though the change in tax paid is actually negative (the individual is on the wrong side of the Laffer curve). However, the person with an hourly wage of $25 also reduces labour supply by less than one hour but has $EV = 19.98$, composed of a marginal excess burden of $10.85, and additional tax of $9.14, implying a marginal welfare cost of $1.19 – that is, an excess burden per dollar of extra revenue of over a dollar. The higher wage of $30 results in a similarly small labour supply response, a much higher excess burden, but a smaller marginal welfare cost.

Table 8 shows hours and welfare changes for hypothetical workers in couple households with children, for the same tax change. In this case the absense of the top marginal rate (of 0.40) means that the initial effective rate structure has rates of 0.15, 0.85, 0.25, 0.55 and 0.35 applying above thresholds of 0, 200, 450, 650 and 1400. The worker with a wage of $22.6 per hour actually has, before the tax change, tangency positions on segments 2 and 4, but the latter is the global maximum. The tax change gives rise to a move to the new corner solution (though the sub-optimal tangency on segment 2 is clearly unchanged). This gives rise to a small labour supply reduction, a small welfare loss, but again a tax reduction, so that again the wrong side of the Laffer curve is relevant. For the two higher wages of 28.2 and 30 per hour, the marginal excess burdens are similar, at 10.93 and 9.23 respectively, but the marginal welfare costs are $0.70 and $0.31 respectively.

These hypothetical examples demonstrate that it is very difficult to generalise about welfare changes, even for individuals with similar preferences. They all display small hours responses but substantial variations in marginal excess burdens and marginal welfare costs.

# 9  Conclusions

This paper has considered, within a limited compass, the extend to which economic theory can offer specific policy advice regarding income tax structures. Ultimately, it was seen that many of the results are negative or too broad to offer direct policy

guidance. Indeed, much of the theory has clarified instead just why it is very difficult to produce clear cut arguments. In providing policy advice, the point inevitably arises that the role of the economist is to examine the implications of adopting alternative value judgements – and there are few results which do not depend in some way on the ultimate objectives of a tax system.

The extensive optimal tax literature does not provide, and was never expected to provide, clear guidance, but instead has clarified the precise way in which the optimal tax system depends on a wide range of factors, some of which relate to value judgements while others concern behavioural responses or basic conditions, such as abilities, which display considerable heterogeneity in practice. Clarifying just how certain conclusions rely on strong, and perhaps unrealistic, assumptions is of course part of the role of economists, yet it is understandable that this can provoke impatience in others.

In looking for practical advice, it appears that a more piecemeal approach must be used. That is, it is necessary to consider in turn a number of features, or their implications for particular specified outcomes, rather than hoping to produce a general rule. This paper has discussed just a few of those factors, mainly in the context of the role of a top marginal income tax rate. Illustrative numerical examples were provided, showing the kind of calculation which may be made. In particular, it was suggested how a behavioural microsimulation model could be used to examine marginal reforms.

It seems almost inevitable that consideration of income tax structures returns yet again to the famous statement made by McCulloch (1845), quoted in the introduction. Modern public finance theorists might argue that the 'cardinal principle' of proportionality involves an unstated value judgement, and that there is now a better understanding of just why a simple 'compass' is not available (in view of the many measurement problems), and a better understanding of constraints imposed on any 'rudder' (or policy instrument) in atttempting to achieve a policy objective. However, his argument does provide a timeless reminder that any policy advice must be extremely tentative and will (given sufficient courage) often take the form of explaining (given the propensity of politicians to tinker with tax structures) just why a stated objective cannot be achieved by a particular policy initiative.

# Appendix A: Welfare Changes and Cobb-Douglas Utility

Consider first the expression for labour supply arising from the Cobb-Douglas utility function. The maximum available number of working hours is $T$. The price index of consumption goods is normalised to 1, so consumption, $c$, is equal to net income, and the direct utility function is:

$$U(c, h) = c^\alpha (T - h)^{1-\alpha} \tag{A.1}$$

Maximisation subject to the linear budget constraint, $c = \mu + wh$, gives the standard interior solutions:

$$\begin{aligned} c &= \alpha M \\ h &= T - (1 - \alpha)\frac{M}{w} \end{aligned} \tag{A.2}$$

where $M$ denotes full income, $M = wT + \mu$. The values of $w$ and $\mu$ are the endogenous net wage and virtual income corresponding to the appropriate linear section.

In the case of corner solutions, it is necessary to find the values of the virtual wage and income, $w_i^K$ and $\mu_i^K$ respectively, which would generate the same position as a tangency solution. Suppose that the corner is at the start of the $i$th linear segment, where the hours of work are $h_i^*$. Net income at that point is thus given by $c_i = \mu_i + w_i h_i^*$. The virtual wage is equal to the slope of the indifference curve at $(c_i, h_i^*)$, so that:

$$w_i^K = \left(\frac{1 - \alpha}{\alpha}\right) \frac{c_i}{T - h_i^*} \tag{A.3}$$

and:

$$\mu_i^K = c_i - w_i^K h_i^* \tag{A.4}$$

Hence corner solutions can be treated in precisely the same way as tangency solutions, so long as the appropriate virtual values are used.

The indirect utility function, $V(w, U)$, is obtained by substituting (A.2) into the direct utility function, giving:

$$V(w, U) = \alpha^\alpha \left(\frac{1 - \alpha}{w}\right)^{1-\alpha} M \tag{A.5}$$

The expenditure function, expressed in terms of full income, is given by inverting (A.5) to give:

$$E(w, U) = U \left(\frac{1}{\alpha}\right)^\alpha \left(\frac{w}{1 - \alpha}\right)^{1-\alpha} \tag{A.6}$$

The expenditure function could be expressed in terms of the virtual income, since $\mu = M - wT$, though the choice is purely one of convenience.

Suppose that the initial tax and transfer system gives rise to an optimum associated with utility, $U^0$, and the values, $M^0$ and $w^0$, with $M^0 = w^0 T + \mu^0$. A change in the tax system gives rise to utility, $U^1$, and, $M^1$ and $w^1$, with $M^1 = w^1 T + \mu^1$. If corner solutions (either before or after the change) are relevant, the values of $w$ and $\mu$ are the

associated virtual values as expressed in (A.3) and (A.4); for tangency solutions they are the net wage and virtual income corresponding to the relevant linear section of the budget constraint.

In the Cobb-Douglas case, appropriate substitution gives:

$$EV = M^1 \left\{ 1 - \left( \frac{w^0}{w^1} \right)^{1-\alpha} \right\} + \left\{ M^0 - M^1 \right\} \tag{A.7}$$

which is used to produce results in section 8.

# Appendix B: Optimal Income Taxation

This appendix first briefly reviews the formal structure of the simplest possible type of optimal tax model, the linear income tax framework with identical individuals except for their earning ability. Second, marginal tax reforms are discussed.

## The Linear Income Tax

The government must select the values of the transfer, $b$, and the tax rate, $t$, which maximise $W = \sum_i G(V_i)$, where $V_i$ is $i$'s indirect utility, subject to the constraint that $b = t\bar{y}$. Where individuals have the same tastes, $\bar{y}$ is in general a complex function of preferences, $b$, $t$, and the wage rate distribution, $F(w)$. The Lagrangean for the optimal tax problem is thus:

$$L = W + \lambda (t\bar{y} - b) \tag{B.1}$$

Simultaneously individuals maximise $U(c_i, h_i)$ subject to the constraint that $c_i = b + w_i (1 - h)(1 - t)$, giving indirect utility $V_i$. The first-order conditions can be written as:

$$\frac{\partial L}{\partial b} = \sum_i \frac{\partial G}{\partial V_i} \frac{\partial V_i}{\partial b} + \lambda \left( t \frac{\partial \bar{y}}{\partial b} - 1 \right) = 0 \tag{B.2}$$

$$\frac{\partial L}{\partial t} = \sum_i \frac{\partial G}{\partial V_i} \frac{\partial V_i}{\partial t} + \lambda \left( \bar{y} + t \frac{\partial \bar{y}}{\partial t} \right) = 0 \tag{B.3}$$

These may initially appear to be quite straightforward, but the general treatment of the first-order conditions is highly complex. Further progress requires more structure to be imposed on the model, along with numerical analysis.

Following Tuomala (1985), turther insight can be obtained by considering the tangency between a social indifference curve and the government budget constraint, obtained by dividing the two first-order conditions to give:

$$\left. \frac{db}{dt} \right|_W = -\frac{\partial W / \partial t}{\partial W / \partial b} = \frac{\left( \bar{y} + t \frac{\partial \bar{y}}{\partial t} \right)}{1 - t \frac{\partial \bar{y}}{\partial b}} = \left. \frac{db}{dt} \right|_R \tag{B.4}$$

where:

$$-\frac{\partial W / \partial t}{\partial W / \partial b} = -\frac{\sum_{i=1}^n \frac{\partial W}{\partial V_i} \frac{\partial V_i}{\partial t}}{\sum_{i=1}^n \frac{\partial W}{\partial V_i} \frac{\partial V_i}{\partial b}} = -\sum_{i=1}^n \left( \frac{v_i}{\sum_{i=1}^n v_i} \right) \left( \frac{\partial V_i / \partial t}{\partial V_i / \partial b} \right) \tag{B.5}$$

where the term $v_i = \frac{\partial W}{\partial V_i} \frac{\partial V_i}{\partial b}$ is the 'welfare weight' attached by the social welfare function to an addition to person $i$'s income (from an increase in the basic income).

At this point, it is useful to employ a standard result from duality theory. In general for an indirect utlity function of the form $V(p, m)$, for goods demanded, $x_i$, at prices, $p_i$, and a budget of $m$, Roy's Identity gives the Marshallian demands as $x_i = -\left(\partial V/dp_i\right)/\left(\partial V/dm\right)$. In the present context, this means that labour supply, $h_i$, can be expressed as:

$$
\begin{aligned}
h_i &= \frac{\partial V/\partial \left(w_i\left(1-t\right)\right)}{\partial V/\partial M_i} \\
&= -\frac{1}{w_i}\frac{\partial V/\partial t}{\partial V/\partial b}
\end{aligned}
\tag{B.6}
$$

In the first line of this expression, the minus sign in the standard form of Roy's Identity has been deleted because the variable in question is the amount supplied, not demanded. Hence:

$$
y_i = -\frac{\partial V/\partial t}{\partial V/\partial b}
\tag{B.7}
$$

and:

$$
\left.\frac{db}{dt}\right|_W = \sum_{i=1}^{n}\left(\frac{v_i}{\sum_{i=1}^{n}v_i}\right)y_i
\tag{B.8}
$$

The right hand side of this expression is a weighted average of the $y_i$s, which can be denoted $\widetilde{y}$. Hence:

$$
\widetilde{y} = \frac{\overline{y} + t\frac{\partial \overline{y}}{\partial t}}{1 - t\frac{\partial \overline{y}}{\partial b}}
\tag{B.9}
$$

so that:

$$
\tilde{y} - \overline{y} = t\left(\frac{\partial \overline{y}}{\partial t} + \tilde{y}\frac{\partial \overline{y}}{\partial b}\right)
\tag{B.10}
$$

The variation in average gross earnings as tax parameters vary can be obtained by totally differentiating, so that $d\overline{y} = \frac{\partial \overline{y}}{\partial t}dt + \frac{\partial \overline{y}}{\partial b}db$ and:

$$
\left.\frac{d\overline{y}}{dt}\right|_R = \frac{\partial \overline{y}}{\partial t} + \left(\frac{\partial \overline{y}}{\partial b}\right)\left.\frac{db}{dt}\right|_R
\tag{B.11}
$$

From the first-order condition, $\tilde{y} = \left.\frac{db}{dt}\right|_R$, and using equation (B.11), the term in brackets on the right hand side of (B.10) can be replaced by $\left.\frac{d\overline{y}}{dt}\right|_R$. Hence, after dividing by $\bar{y}$:

$$
1 - \frac{\tilde{y}}{\overline{y}} = -\frac{t}{\overline{y}}\left.\frac{d\overline{y}}{dt}\right|_R = \left|\eta_{\overline{y},t}\right|
\tag{B.12}
$$

where $\left|\eta_{\overline{y},t}\right|$ represents the absolute value of the elasticity of average earnings with respect to the tax rate. This is the result given in equation (3) above.

## Marginal Tax Reforms

In the context of marginal indirect tax reform, the social welfare function is again expressed in general terms as $W = W(V_1, ..., V_n)$, where $V_j$ is the indirect utility of the $j$th household, for $j = 1, ..., n$. If $t_i$ represents the tax imposed on each unit of good $i$, then the change in $W$ resulting from a marginal change in $t_i$ is:

$$\frac{\partial W}{\partial t_i} = \sum_{j=1}^{n} \frac{\partial W / \partial V_j}{\partial V_j / \partial t_i} \tag{B.13}$$

This can be rewritten, using $v_j = \frac{\partial W}{\partial V_j} \frac{\partial V_j}{\partial y_j}$, and where $x_{ji}$ is household $j$'s demand for good $i$, as:

$$\frac{\partial W}{\partial t_i} = \sum_{j=1}^{n} v_j \frac{\partial V_j / \partial t_i}{\partial V_j / \partial y_j} = -\sum_{j=1}^{n} v_j x_{ji} \tag{B.14}$$

which again makes use of the duality property mentioned above, and for small changes $\partial p_i = \partial t_i$. The term $v_j$ is the 'social marginal utility' of income of household $j$.

The aggregate tax revenue, $R$, from indirect taxes on all $K$ goods is given by:

$$R = \sum_{j=1}^{n} \sum_{k=1}^{K} t_k x_{jk} \tag{B.15}$$

The change in revenue arising from a marginal change in $t_i$, $\partial R / \partial t_i$, is the sum of two terms. The first is equal to the initial tax base (that is, total consumption of the $i$th good over all households) and the second depends on the tax rate and the changes in consumption by households. The ratio, $(\partial W / \partial t_i) / (\partial R / \partial t_i)$, measures the reduction in social welfare per dollar of extra tax revenue resulting from a marginal increase in the tax $t_i$. For an optimal tax system, this ratio must be equal for all goods. The direction of marginal tax reform is indicated by the relative magnitudes of this ratio for each commodity group. Multiplying the two terms by $p_i$ allows the ratio to be expressed in terms of expenditures (rather than quantities) and cross-price elasticities, since, for example:

$$p_i \frac{\partial R}{\partial t_i} = \sum_{j=1}^{n} p_i x_{ji} + \sum_{j=1}^{n} \sum_{k=1}^{K} \tau_k \eta_{jki} p_k x_{jk} \tag{B.16}$$

where $\eta_{jki}$ is household $j$'s elasticity of demand for good $k$ with respect to the price of good $i$, and $\tau_k$ is the ratio of the tax to the tax-inclusive price. Hence $\tau_i$ is the tax-inclusive *ad valorem* rate. This simplifies further if households are assumed to have equal elasticities.

In the context of marginal income tax reform, an analytical approach may begin with a multi–step tax function described by a series of marginal tax rates and income thresholds over which the rates apply. Tax paid by household $j$ is (here allowing explicitly for a tax-free zone) $T(y_j) = 0$, for $0 < y_i \leq a_1$; $T(y_j) = t_1(y_i - a_1)$, for $a_1 < y_i \leq a_2$; $T(y_j) = t_1(a_2 - a_1) + t_2(y_i - a_2)$, for $a_2 < y_i \leq a_3$ and so on. If

$a_k < y_j \le a_{k+1}$ so that $y_j$ is in the $k$th tax bracket, and $a_0 = t_0 = 0$, $T(y_j)$ can, as shown by Creedy and Gemmell (2006, p. 25), be written for $k \ge 1$ as:

$$T(y_i) = t_k (y_i - a'_k) \tag{B.17}$$

where:

$$a'_k = \sum_{j=1}^{k} a_j \left( \frac{t_j - t_{j-1}}{t_k} \right) \tag{B.18}$$

Hence the tax function facing any individual taxpayer is equivalent to one with a single marginal tax rate, $t_k$, applied to income measured in excess of a single threshold, $a'_k$. The revenue obtained from any particular threshold, given a density function of $F(y)$, is:

$$T_k = t_k \int_{a_k}^{a_{k+1}} (y_i - a'_k) \, dF(y) \tag{B.19}$$

Total revenue is the sum over all ranges of such terms. Differentiation of total tax revenue with respect to any particular marginal rate is thus considerably more complex than in the case of indirect taxation.

# References

[1] Ahmad, E. and Stern, N.H. (1984) The theory of tax reform and Indian indirect taxes. *Journal of Public Economics,* 25, pp. 259-298

[2] Atkinson, A.B. (1995) *Public Economics in Action: The Basic Income/Flat Tax Proposal.* Oxford: Oxford University Press.

[3] Atkinson, A.B and Stiglitz, J.E. (1980) *Lectures in Public Economics.* London: McGraw-Hill.

[4] Besley, T. and Coate, S. (1992) Workfare versus welfare: incentive arguments for work requirements in poverty-alleviation programs. *American Economic Review*, pp. 249-261.

[5] Bradbury, B. (1999) Optimal taxation theory and the targeting of social assistance. University of New South Wales Social Policy Research Centre.

[6] Chang, C-H. (1994) On the optimum rate structure of an individual income tax. *Southern Economic Journal*, 60, pp. 927-935.

[7] Cohen-Stuart, J. (1889) On progressive taxation. Reprinted in *Classics in Public Finance* (ed. by R.A. Musgrave and A.T. Peacock, 1958). London: Macmillan.

[8] Creedy, J. (1996) *Fiscal Policy and Social Welfare.* Cheltenham: Edward Elgar.

[9] Creedy, J. (1998a) Means-tested versus universal transfers: alternative models and value judgements. *Manchester School,* 66, pp. 100-117.

[10] Creedy, J (1998b) The optimal linear income tax model: utility or equivalent income? *Scottish Journal of Political Economy*, 45, pp. 99-110.

[11] Creedy, J. (1999) Marginal indirect tax reform in Australia. *Economic Analysis and Policy,* 29, pp. 1-14.

[12] Creedy, J. (2001a) Tax modelling. *Economic Record,* 77, pp. 189-202.

[13] Creedy, J. (2001b) Indirect tax reform and the role of exemptions. *Fiscal Studies,* 22, pp. 457-486.

[14] Creedy, J. (2001c) Labour supply, welfare and the earnings distribution. *Australian Journal of Labour Economics,* 4, pp. 134-151.

[15] Creedy, J. and Kalb, G. (2005) Measuring welfare changes in labour supply models. *Manchester School,* 73, pp. 664-685.

[16] Creedy, J. and Kalb, G. (2006) *Labour Supply and Microsimulation: The Evaluation of Tax Policy Reforms.* Cheltenham: Edward Elgar.

[17] Creedy, J. and Gemmell, N. (2006) *Modelling Tax Revenue Growth.* Cheltenham: Edward Elgar.

[18] Creedy, J. and Sleeman, C. (2006) *The Distributional Effects of Indirect Taxes: Models and Applications from New Zealand.* Cheltenham: Edward Elgar.

[19] Creedy, J., Herault, N. and Kalb, G. (2008) Welfare change measures in behavioural microsimulation modelling: accounting for the random utility component. University of Melbourne.

[20] Deaton, A. (1983) An explilcit solution to an optimal tax problem. *Journal of Public Economics*, 20, pp. 333-346.

[21] Diamond, P.A. (1998) Optimal income taxation: an example with a U-shaped pattern of optimal marginal tax rates. *American Economic Review*, 88, pp. 83-95.

[22] Edgeworth, F.Y. (1897) The pure theory of taxation. *Economic Journal*, 7, pp. 46-70, 226-238, 550-571,

[23] Edgeworth, F.Y. (1925) *Papers Relating to Political Economy*, Vol. ii. London: Macmillan.

[24] Hashimzade, N. and Myles, G.D. (2004) The structure of the optimal income tax in the quasi-linear model. *PEUK Working Paper,* no. 2.

[25] Heady, C. (1993) Optimal taxation as a guide to tax policy. *Fiscal Studies*, 14, pp. 15-41.

[26] Hindriks, J. and Myles, G.D. (2006) *Intermediate Public Economics.* Cambridge, Mass.: The MIT Press.

[27] Kanbur, R., Keen, M. and Tuomala, M. (1994) Optimal non-linear income taxation for the alleviation of income-poverty. *European Economic Review*, 38, pp. 1613-1632.

[28] Kanbur, R., Pirttilä, J. and Tuomala, M. (2004) Non-welfarist optimal taxation and behavioural public economics. University of Tampere, Finland.

[29] McCulloch, J.R. (1845) *A Treatise on the Principles and Practical Influence of Taxation and the Funding System.* (Ed. with introduction by D.P. O'Brien, 1975). London: Longman.

[30] Madden, D. (1996) Marginal tax reform and the specification of consumer demand systems. *Oxford Economic Papers,* 48, pp. 556-567.

[31] Mill, J.S. (1848) *Principles of Political Economy.* (Reproduced with editorial material by W.J. Ashley). London: Longmans, Green and Company.

[32] Mirrlees, J.A. (1971) An exploration in the theory of optimum income tax. *Review of Economic Studies*, 38, pp. 175-208.

[33] Myles, G.D. (1999) On the optimal marginal rate of income tax. University of Exeter.

[34] O'Brien, D.P. (2009) Introduction. In *Taxation and the Promotion of Human Happiness: An Essay by George Warde Norman* (Ed. by D.P. O'Brien with J. Creedy). Cheltenham: Edward Elgar.

[35] Saez, E. (2001) Using elasticities to derive optimal tax rates. *Review of Economic Studies*, 68, pp. 205-229.

[36] Slemrod, J. (1990) Optimal taxation and optimal tax systems. *Journal of Economic Perspectives*, 4, pp. 157-178.

[37] Smith, A. (1776) *An Inquiry into the Nature and Causes of the Wealth of Nations.* (Cannan edition reproduced, 1776). Chicago: Chicago University Press.

[38] Stern, N.H. (1984) Optimal taxation and tax policy. *IMF Staff Papers*, 31, pp. 339-378.

[39] Tuomala, M. (1985) Simplified formulae for optimal linear income tax. *Scandinavian Journal of Economics*, 87, pp. 668-672.

[40] Tuomala, M. (1995) *Optimal Taxation and Redistribution*. Oxford: Clarendon Press.

[41] Tuomala, M. (2006) On the shape of optimal non-linear income tax schedules. *Tampere Economic Working Papers*, no. 49.