THE UNIVERSITY OF
MELBOURNE

Department of Economics

Working Paper Series

Optimal Marginal Income Tax Reforms: A Microsimulation Analysis

John Creedy & Nicolas Herault

August 2009

Research Paper Number 1080

ISSN: 0819-2642
ISBN: 978 0 7340 4044 2

Department of Economics
The University of Melbourne
Parkville VIC 3010
www.economics.unimelb.edu.au

Optimal Marginal Income Tax Reforms: A Microsimulation Analysis*

John Creedy and Nicolas Hérault†

Abstract

Extensive research has shown that few robust results regarding the optimal tax structure are available. Moreover, the stylised models used in optimal tax analyses are not appropriate for practical policy advice. This paper proposes a method of examining optimal marginal income tax reforms using 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. The approach is illustrated using the Australian microsimulation model MITTS. The results show that the marginal welfare changes for the Australian income tax structure are not symmetric with respect to increases and decreases in tax rates, largely because of the asymmetry in tax revenue changes arising from differential labour supply effects in different ranges of the income distribution. In addition, the extent of inequality aversion was found to play a much larger role in the determination of the optimal direction of rate changes than the form of the welfare metric or the specification of adult equivalence scales.

JEL Classification: D63; H21; H31; I31; J22.

Keywords: Optimal taxation; marginal reforms; behavioural microsimulation; social welfare function; money metric utility.

*We are grateful to Hielke Buddelmeyer and Guyonne Kalb for helpful comments on an earlier version of this paper.

†Department of Economics and Melbourne Institute of Applied Economic and Social Research respectively, University of Melbourne, Victoria 3010, Australia.

1 Introduction

The optimal tax problem, which in general asks what tax structure maximises a specified evaluation, or social welfare function, subject to a government budget constraint, is known from Mirrlees (1971) to give rise to intractable problems so that numerical simulation methods are needed, unless strong simplifying assumptions are made.[1] Extensive research has shown that few robust results are available: the optimal structure depends on the nature of the social welfare function examined as well as the wage rate distribution and the nature of individuals' preferences.[2]

Optimal tax modelling relies on small simulation models in which there is negligible population heterogeneity. However, practical policy advice requires 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).[3] It does not seem practicable to use such models to produce optimal nonlinear structures, but this is not a severe limitation because practical policy concerns involve changes in an existing tax structure rather than a complete redesign of an existing tax and transfer system.

The contribution of this paper is to propose a method of examining the social welfare effects of marginal income tax reforms using a behavioural microsimulation model. The approach involves the computation of social welfare changes per dollar of revenue, for explicit forms of social welfare function, for each existing marginal income tax rate. The welfare metric is the money metric utility per adult equivalent person. The approach is illustrated using the Melbourne Institute Tax and Transfer Simulator (MITTS).[4] Although marginal tax reforms have been examined in the context of indirect taxes,

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

[2]For reviews of policy implications of optimal tax theory see, for example, Stern (1984), Heady (1993), Tuomala (1995) and Bradbury (1999). For a critique of the approach, see Slemrod (1990).

[3]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 (2006).

[4]MITTS is described briefly in the Appendix and in detail in Creedy *et al.* (2002). Creedy and Kalb (2006) describe some further features of MITTS, and Kalb and Lee (2007, 2008) report updated wage and labour supply estimates underlying the labour supply responses. For a comparison of MITTS results with experimental evaluation results of the same policy change, see Cai *et al.* (2008).

they have not previously been computed for income taxation in view of the difficulties of dealing with population heterogeneity, obtaining an appropriate welfare metric for each individual (given the complexity of budget constraints) and the need to deal with the relevant probability distributions when using a discrete hours random utility approach to labour supply modelling.

Section 2 first briefly reviews the main results of optimal tax analyses and explains the approach used to examine marginal income tax reforms. Implementation of the method in a microsimulation model is discussed in Section 3. Unlike the case of marginal indirect tax reforms, the method is complicated by the existence of highly nonlinear budget constraints facing individuals. This affects the computation of welfare changes and also leads to an asymmetry in the effects of marginal tax rate increases and decreases. Furthermore, labour supply in the MITTS model is based on a discrete hours structural approach. Hence the method must also deal with the fact that simulations produce, for each individual, a probability distribution over the discrete hours levels, rather than a specific point. Section 4 presents numerical results relating to marginal reforms to the Australian income tax structure. Actual tax structure changes in Australia are then examined in Section 5. Brief conclusions are in section 6.

2 Optimal Taxation and Tax Reform

This section briefly reviews the main lessons regarding tax design, from optimal tax theory. It then considers the problem of obtaining information about optimal reforms, involving small changes to income tax rates which lead to increases in a specified social welfare, or evaluation, function.

2.1 Optimal Structures

Earlier optimal tax 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.[5] 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 utilities; that is, it is expressed in terms of variables which matter to

[5] A small selection of examples includes those discussed by Diamond (1998), Chang (1994), Hashimzade and Myles (2004), Myles (1999), Saez (2001) and Tuomala (2006).

individuals. But 'non-welfarist' forms are sometimes used. For example, social welfare may be based solely on an income-based measure of poverty, which can give quite different results.[6] Non-welfarist objectives 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).[7]

One approach has been to consider piecewise-linear tax functions with just two or three rates. This allows 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. 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.[8] The 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).

In practice, governments seldom have the 'luxury' of completely redesigning tax structures, so that from a practical point of view it is useful to consider optimal marginal reforms rather than attempting the much more complex task of evaluating an optimal structure.

2.2 Marginal Tax Reform

The problem of how to move towards an optimal structure by a process of marginal tax reform has previously been examined in the context of indirect taxation, involving fixed

[6] For further discussion see Kanbur, Keen and Tuomala (1994) and for a broader treatment of non-welfarist objectives, see Kanbur, Pirttilä and Tuomala (2004).

[7] See, for example, Besley and Coate (1992).

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

incomes but differential taxation on a range of goods.[9] Let x_{ji} denote the consumption of the ith taxable good by the jth household, and t_i the unit tax on good i. A standard result shows that an increase in social welfare, W, resulting from a change in the ith tax rate is:

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

where v_j is the the social value of additional consumption by household j. Furthermore, aggregate tax revenue, R, from indirect taxes on K goods is given by:

$$R = \sum_{j=1}^{n} \sum_{k=1}^{K} t_k x_{jk} \tag{2}$$

and $\frac{\partial R}{\partial t_i}$ can be expressed in terms of (Marshallian) demand elasticities, at observed consumption levels, and expenditures. Multiplying $\frac{\partial R}{\partial t_i}$ by p_i allows the ratio to be expressed in terms of expenditures (rather than quantities) and cross-price elasticities, since, for example:

$$p_i \frac{\partial R}{\partial t_i} = \sum_{j=1}^{n} p_i x_{ji} + \sum_{j=1}^{n} \sum_{k=1}^{K} \tau_k \eta_{jki} p_k x_{jk} \tag{3}$$

where η_{jki} is household j's elasticity of demand for good k with respect to the price of good i, and τ_k is the ratio of the tax to the tax-inclusive price. Hence τ_i is the tax-inclusive *ad valorem* rate. This simplifies further if households are assumed to have equal elasticities. 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. An optimal system is characterised by an equi-marginal condition whereby all these marginal welfare costs are equal. Hence they can be used to determine the required direction of tax changes.

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 that if t_i now represents the ith marginal rate, and y_j is person j's earnings, a similar result

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

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:

$$T(y_j) = t_k (y_j - a'_k) \tag{4}$$

where y_j falls into the kth 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 kth bracket, T_k, is thus:

$$T_k = t_k \sum_{j=1}^{n} (y_j - a'_k) \tag{5}$$

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 particular rate, to consider whether (for a given set of tax thresholds) that marginal rate should be reduced or increased. For an extensive discussion of the use of elasticities to derive optimal tax rates, see Saez (2001).

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.

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, can 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 marginal welfare and revenue changes, denoted ΔW and ΔR respectively, for each tax parameter in turn. The marginal welfare costs and benefits, that is the changes in welfare for small increases and decreases in 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. However, there is a further complication arising from the nonlinearity of the budget constraints facing individuals. The effects of small tax rate changes are not symmetrical with respect to increases and decreases.[10] Hence the investigation of optimal marginal reforms must consider explicitly the resulting changes for rate changes in each direction. Unlike the case of marginal indirect tax reform, the present context requires the explicit evaluation of a welfare metric for each individual, and thence the social welfare function under each tax regime considered.

3 Computation of the Welfare Metric

The computation of the welfare metric for each individual must allow for the complexity of budget constraints. The method used here of computing welfare changes for each individual in a discrete hours context follows that proposed in Creedy and Kalb (2005), and implemented and extended to the calculation of money metric utility in a microsimulation context in Creedy, Hérault and Kalb (2008). MITTS uses a discrete hours labour supply model. Hence, the behavioural simulations produce a frequency distribution of post-reform hours for each individual, conditional on the individual's optimal pre-reform hours being equal to observed (discretised) hours.[11] This flows on to the welfare calculation; that is, a frequency distribution of welfare and net income changes is obtained for each individual. A method of dealing with the distributions of net income and welfare metrics, when each person faces a probability distribution, was proposed by Creedy *et al.* (2006), involving the production of a 'pseudo' distribution. The probability distributions of welfare and net income changes by hours worked are used for social welfare evaluations. This provides more accurate results than taking,

[10] In addition, the effects on welfare and revenue of further increasing (or reducing) the marginal tax rates are not expected to be linear.

[11] The conditional distribution is obtained using the method of generating conditional draws described in Bourguignon, Fournier and Gurgand (1998). The error terms are drawn from the set of error terms which all result in observed labour supply.

say, the arithmetic mean values for each person.

Population-level evaluations of social welfare, and changes arising from tax reforms, necessarily involve value judgements, so that a decision must be made regarding the social evaluation method. Value judgements concern three aspects: the welfare metric, the definition of the unit of analysis and the form of the social welfare function to be used. As mentioned above, money metric utility is used here as the welfare metric, but is also compared below with the use of a 'non-welfarist' metric based on net income. A feature of optimal tax models is that results depend on the cardinalisaton of utility functions used. The use of money metric utility, with current 'prices' as reference prices, is invariant with respect to monotonic transformations of utility.[12] The choice of welfare function is closely related to value judgements regarding inequality aversion and the implied inequality measure. Different values of inequality aversion are used in the analyses in this paper. Any evaluation for a broad group of income units involves comparisons of units of different size and composition. Therefore, the reported results are based on the use of money metric utility per adult equivalent, using parametric equivalence scales described below.

The steps in the social evaluation are as follows. For each income unit, the initial money metric utility, M_0, is obtained, using pre-reform taxes as 'reference prices'; this is equal to 'full income' under the pre-reform system (defined as the net income which could be obtained if all the endowment of time were devoted to work at the going wage rate). For each income unit, the net income at 80 hours of work by all adult members of the income unit under pre-reform taxes is calculated. Assuming that 80 hours is the maximum number of hours that can be worked per week, this net income represents full income for the income unit. Following the approach introduced by Creedy, Hérault and Kalb (2008), the equivalent variation from a tax policy change, EV, is obtained by searching all discrete labour supply points for the minimum compensation for each conditional draw, while taking into account the non-linearity and non-convexity of the budget constraint. Then, given the equivalent variation, EV, resulting from the reform for each of the discrete hours levels, money metric utility is computed as $M_1 = M_0 - EV$. A probability distribution of EVs and thus money metric utilities is obtained for each unit.

For each income unit, the adult equivalent size, s, is obtained using equivalence

[12] The use of money metric utility in optimal tax models was investigated by Creedy (1998)

scales. Following Banks and Johnson (1994) and Jenkins and Cowell (1994), the following parametric scales are used:

$$s = (n_a + \theta n_c)^{\alpha} \tag{6}$$

where n_a and n_c are respectively the number of adults and children in the unit, θ is the weight attached to children and α represents the extent of economies of scale. The size, s_i, is used to compute money metric utility per adult equivalent, m_{ji}, where j refers to the tax structure and i refers to the income unit. The distributions of pre-reform and post-reform money metric utility can be used to calculate social evaluation measures.

In computing inequality and welfare measures with the individual as the unit of analysis, each value of m_{ji} is weighted by the actual number of persons in the income unit, n_i. This paper uses Atkinson's inequality measure, $A(\varepsilon)$, where ε is the degree of relative inequality aversion. The inequality measure is expressed as 1 minus the ratio of the equally distributed equivalent value to the arithmetic mean. The equally distributed equivalent value is the value which, if obtained by everyone, gives the same social welfare as the actual distribution. Using an additive welfare function based on constant relative inequality aversion, ε, of the form:

$$W = \frac{1}{1-\varepsilon} \sum_{i=1}^{n} y_i^{1-\varepsilon} \tag{7}$$

the equally distributed equivalent value, y_{ede}, is in general, for a set of values y_i, for $i = 1, \ldots, N$, equal to:

$$y_{ede} = \left(\frac{1}{N} \sum_{i} y_i^{1-\varepsilon} \right)^{1/(1-\varepsilon)} \tag{8}$$

Results can be obtained for a range of inequality aversion parameters, ε. Finally, social welfare in each system is most conveniently obtained using the abbreviated form of the welfare function in (7). This is:

$$W_j = \bar{m}_j \left(1 - A(\varepsilon)\right) \tag{9}$$

where $\bar{m}_j$ is the arithmetic mean value of the money metric utility per adult equivalent. It is possible to compare results based on money metric utility with those obtained using net incomes in the social welfare function. The use of the abbreviated form, showing the trade-off between 'equity and efficiency', is convenient because it ensures that W is positive for all values of ε.

The orders of magnitude of ε can be interpreted by considering the 'leaky bucket' experiment of making a transfer from a richer to a poorer person. Consider two individuals, a transfer from person 2 to person 1, in the context of incomes, $y_2 > y_1$, which leaves social welfare unchanged is given by:

$$\left.\frac{dy_1}{dy_2}\right|_W = -\left(\frac{y_2}{y_1}\right)^{-\varepsilon} \tag{10}$$

Hence if, for discrete changes and a transfer of 1 from the richest person, $\Delta y_2 = -1$, and if incomes are such that $y_2 = 2y_1$, then the amount which must be given to the poorest person is $\Delta y_1 = 2^{-\varepsilon}$. For example, if $\varepsilon = 0.5$, $\Delta y_1 = 0.71$ and a 'leaky bucket' involving a loss of 29 cents from the dollar is tolerated. Also, if $\varepsilon = 1$, $\Delta y_1 = 0.5$ and the judge is prepared to lose 50 cents from the dollar taken from the richest person. If ε is as high as 1.4, it can be seen that $\Delta y_1 = 0.39$ and the judge is prepared to lose 61 cents from the dollar. By contrast, if the judge has no aversion to relative inequality, $\varepsilon = 0$ and no loss is tolerated. Using questionnaires based on this experiment, Amiel *et al.* (1999) reported an average value of around 0.2 for economics students in three countries.

4 Simulation Results

The following simulation results are based on the Australian tax and transfer system for the financial year of 2003/2004, using the MITTS model with the 2003/2004 Survey of Income and Housing Costs.[13] Population values were obtained by using the sample weights provided. The income tax structure is given in Table 1, where t_i is the marginal rate applied above an annual income of a_i, for $i = 1, ..., 5$. Each of these marginal rates was in turn decreased by one percentage point, and then increased by one percentage point, and the resulting changes in total revenue and social welfare were obtained.

[13] It is desirable to use the actual tax structure in operation at the time of the survey, so that calibration places individuals at their labour supplies corresponding to actual incentives faced.

Table 1: The Australian Income Tax 2003/2004

No.	Income threshold (in AU$)	Marginal tax rate
1	$a_1 = 0$	$t_1 = 0$
2	$a_2 = 6,000$	$t_2 = 0.17$
3	$a_3 = 21,600$	$t_3 = 0.30$
4	$a_4 = 52,000$	$t_4 = 0.42$
5	$a_5 = 62,500$	$t_5 = 0.47$

4.1 Aggregate Effects

Before presenting the main results of the marginal reforms, Table 2 presents summary information about simple aggregate effects, for increases and decreases in each marginal tax rate by one percentage point, obtained by adding all expected EV and net incomes across all income units without equivalising the amounts. Changes in net government revenue account for changes in income tax revenue and also for consequent changes in expenditures on social security. Indeed, government spending on pensions, allowances and rebates is affected by changes in income taxes and labour supply. Table 2 also presents labour supply responses by demographic group. For couples, the first amount relates to the male partner while the second amount is for the female partner.

Table 2: Aggregate Effects of Marginal Increases and Decreases in Income Tax Rates

	Net rev change ($m)	Agg net inc change ($m)	EV ($m)	Average hours change in hours per week Couples	Single males	Single females	Single parents	ALL
Increase in marginal tax rate t_i								
t_1	542	-654	600	-0.01, -0.02	-0.01	0.00	-0.05	-0.01
t_2	1,181	-1,548	1,367	-0.03, -0.05	-0.03	-0.01	-0.10	-0.04
t_3	1,256	-1,834	1,522	-0.04, -0.04	-0.04	-0.02	-0.11	-0.04
t_4	182	-255	216	0.00, 0.00	-0.01	0.00	0.00	0.00
t_5	429	-626	532	-0.02, 0.00	-0.01	0.00	0.00	-0.01
Reduction in marginal tax rate t_i								
t_2	-1,258	1,501	-1,369	0.03, 0.03	0.03	0.02	0.10	0.03
t_3	-1,349	1,781	-1,524	0.03, 0.03	0.04	0.03	0.10	0.04
t_4	-177	263	-216	0.01, 0.00	0.00	0.00	0.00	0.00
t_5	-306	764	-534	0.02, 0.01	0.01	0.00	0.00	0.01

Changes in t_2 and t_3 have the largest effects. Changes in other tax rates have smaller effects either because they apply over a smaller income range (as in the case of t_1) or because they concern fewer households (as for rates t_4 and t_5). Increases in tax rates typically lead to a net decrease in labour supply. The latter implies a further reduction in household income and thus a smaller increase in government revenue than would be expected under fixed labour supply. Therefore, the reduction in household net income is larger than the increase in government revenue. By contrast, reductions in taxes typically increase labour supply, leading to reductions in government revenue that are smaller than the increases in household income.

For tax increases, reductions in welfare (as measured by aggregate EV) are smaller than reductions in household income because welfare measures account for the increase in leisure and home production time, following the decrease in labour supply. Similarly, changes in welfare are smaller than changes in net income for tax reductions because welfare measures value the decrease in leisure and home production time. Overall, increases and decreases in tax rates have fairly symmetric effects on aggregate welfare and average hours worked.

4.2 Marginal Reforms

Table 3 shows the resulting absolute values, $\left|\frac{\Delta W}{\Delta R}\right|$, of the marginal welfare cost and benefit per dollar of revenue for increases and decreases in each marginal tax rate by one percentage point. The social welfare function is based on money metric utilities as described above. Results are shown for two values of the economies of scale parameter, α, in the parametric adult equivalence scales, and three values of inequality aversion, ε. The weight attached to children was set at 0.6.

Marginal welfare gains and losses are not symmetric. Given ε and α, the marginal welfare gain associated with a decrease in t_2 and t_3 is always smaller than the corresponding marginal welfare loss associated with a tax increase. The opposite result is found for t_4 and t_5, with marginal welfare gains from tax rate reductions always larger than the marginal welfare losses from tax rate increases. Since changes in aggregate EV are largely symmetric, this can only be explained by an asymmetric behaviour of inequality changes or net government revenue, ΔR. Table 4 reports inequality measures of money metric utility for increases and decreases in the marginal income tax rates. This shows that inequality changes are indeed also highly symmetric. Therefore, the

Table 3: Values of $\left|\frac{\Delta W}{\Delta R}\right|$ Using Money Metric Utility

	Increase in t			Reduction in t		
		ε			ε	
rate	0.2	0.8	1.4	0.2	0.8	1.4
Scale parameter: $\alpha = 0.8$						
t_1	1.35	1.30	1.25			
t_2	1.40	1.33	1.24	1.32	1.25	1.17
t_3	1.46	1.31	1.15	1.36	1.22	1.07
t_4	1.42	1.16	0.94	1.46	1.19	0.97
t_5	1.43	0.99	0.69	2.01	1.38	0.96
Scale parameter: $\alpha = 0.4$						
t_1	1.85	1.77	1.69			
t_2	1.93	1.81	1.68	1.81	1.70	1.58
t_3	2.03	1.79	1.54	1.89	1.67	1.43
t_4	2.01	1.60	1.24	2.07	1.65	1.28
t_5	2.08	1.37	0.90	2.92	1.91	1.25

asymmetric behaviour of the marginal welfare gains and losses is driven by differences in ΔR. For increases in t_1 to t_5 by one percentage point, the increases in net revenue are respectively \$542m, \$1,181m, \$1,256m, \$182m and \$429m. For decreases in t_2 to t_5, the reductions in net revenue are respectively \$1,258m, \$1349m, \$177m and \$306m.

Marginal welfare gains for reductions in t_2 and t_3 are smaller than the corresponding marginal welfare losses associated with increases in these tax rates because ΔR is larger for reductions than for increases in t_2 and t_3. The opposite is true for t_4 and t_5, with ΔR larger for tax increases than for tax reductions.

Under a fixed labour supply scenario, ΔR would be symmetric but after allowing for labour supply changes, the reduction in the extra government revenue generated by increases in t_2 or t_3 is larger than the reduction in the extra government expenditure generated by a tax rate reduction. The explanation is that although aggregate labour supply responses are quite symmetric, those reducing their labour supply in the face of an increase in t_2 or t_3 on average have higher incomes than those increasing their labour supply when the same tax rate is reduced, so that the impact of labour supply responses on government revenue is larger for a tax increase. By contrast, those experiencing the largest increases in labour supply following a reduction in t_4 or t_5 have higher incomes than those with the largest reductions in labour supply after an increase in t_4 or t_5,

Table 4: Atkinson's Index based on Money Metric Uility and Net Income (changes in per cent)

	Increase in t			Reduction in t		
		ε			ε	
Tax rate	0.2	0.8	1.4	0.2	0.8	1.4
Based on money metric utility						
Scale parameter: $\alpha = 0.8$						
t_1	0.03	0.01	0.01			
t_2	0.01	-0.04	-0.06	0.00	0.04	0.06
t_3	-0.20	-0.27	-0.30	0.21	0.27	0.30
t_4	-0.08	-0.08	-0.09	0.08	0.09	0.09
t_5	-0.45	-0.42	-0.37	0.46	0.42	0.37
Scale parameter: $\alpha = 0.4$						
t_1	0.03	0.02	0.01			
t_2	0.02	-0.02	-0.05	-0.02	0.02	0.05
t_3	-0.17	-0.23	-0.27	0.17	0.23	0.27
t_4	-0.07	-0.08	-0.08	0.07	0.08	0.08
t_5	-0.45	-0.40	-0.34	0.45	0.40	0.35
Based on net income						
Scale parameter: $\alpha = 0.8$						
t_1	0.11	0.08	0.05			
t_2	0.15	0.05	-0.03	-0.10	-0.01	0.06
t_3	-0.32	-0.41	-0.48	0.39	0.48	0.54
t_4	-0.15	-0.15	-0.14	0.15	0.15	0.15
t_5	-0.79	-0.65	-0.54	0.93	0.77	0.65

which means that changes in government revenue are more limited for reductions than for increases in t_4 or t_5.

The last segment of Table 4 gives inequality measures when net income is used as the welfare metric. These are still approximately symmetric, but to a lesser extent than inequality changes based on money metric utility. In addition, they are substantially larger in magnitude compared with the corresponding results obtained using money metric utility (with two exceptions for a reduction in t_2 with $\varepsilon = 0.4$ or $\varepsilon = 1.4$). This arises because of the failure to value leisure time in measures based on net income only. By contrast, measures based on money metric utility account for the fact that increases in income are due to some extent to an increase in labour supply. In some cases, for changes in lower tax rates, the inequality measures move in different directions.

Welfare changes per dollar for reductions and increases in tax rates indicate the direction of optimal marginal changes to the tax system. The optimal single rate change is the reduction in the tax rate associated with the largest marginal welfare gain. If it is only required to obtain a marginal increase in revenue, the optimal single rate change is a tax increase resulting in the smallest marginal welfare cost. However, marginal adjustments to the tax system that keep total tax revenue unchanged by combining an increase in a tax rate to fund a decrease in another tax rate may also improve social welfare. Such marginal adjustments are of particular interest because they contribute to a movement towards an optimal tax system, at no cost for the government. The optimal marginal change keeping total tax revenue unchanged is the one involving an increase in the tax rate with the lowest marginal welfare cost combined with a decrease in the tax rate with the highest marginal welfare gain.

Consider the marginal welfare gains and losses for an inequality aversion of $\varepsilon = 0.2$, along with the use of $\alpha = 0.8$ for the scale economy parameter in the adult equivalence scales. The largest welfare gain per dollar, is of 2.01 for a reduction in the top marginal income tax rate. Then moving to the column relating to tax increases, the smallest welfare loss per dollar of extra revenue is for the imposition of a 1 per cent tax rate in the tax-free income range. Thus the optimal revenue-neutral marginal change involves a flattening of the rate structure from a combination of reducing the top rate and increasing (from zero) the lowest rate. A revenue-neutral change involving a reduction in t_5 and an increase in t_1 can be obtained as:

$$\left.\frac{dt_1}{dt_5}\right| = -\frac{\partial R/\partial t_5}{\partial R/\partial t_1} \tag{11}$$

and the resulting welfare increase from such a revenue-neutral change is:

$$\frac{dW}{dt_1} = \frac{\partial W}{\partial t_1} + \left(\frac{\partial W}{\partial t_5}\right)\left.\frac{dt_5}{dt_1}\right| \tag{12}$$

For this value of $\varepsilon = 0.2$, the next biggest marginal welfare gain from reducing a marginal tax rate is for a reduction in t_4, and the next smallest marginal welfare loss is for an increase in t_2. This means that a further flattening of the rate structure will also improve social welfare.

This optimal flattening of the rate structure contrasts with the implications of adopting the very high degree of inequality aversion of $\varepsilon = 1.4$. In this case the highest marginal welfare gain arises from a reduction in t_2, and the smallest welfare loss is for an increase in t_5. Hence the optimal marginal tax reform keeping total tax revenue unchanged is for a small increase in the degree of rate progression. Further welfare improving adjustments to other rates involve a reduction in t_3 along with an increase in t_4.

In the case of $\varepsilon = 0.8$, determining the optimal revenue-neutral marginal change is slightly more difficult since t_5 is associated with both the largest marginal welfare gain for tax reductions and the smallest marginal welfare loss for tax increases. In this particular case, it is useful to consider a more general rule. That is that the optimal revenue-neutral marginal change to the tax system is the one involving the largest difference between the marginal welfare cost of the tax increase and the marginal welfare gain of the tax decrease. Accordingly, the use of $\varepsilon = 0.8$ implies a reduction in t_2 along with an increase in t_5 because it is the combination leading to the largest welfare gain per dollar of revenue shifted from one tax bracket to another. Thus, this invovles an increase in the degree of rate progression, as with $\varepsilon = 1.4$. This is despite the fact that, with this degree of inequality aversion, the single rate change leading to the largest welfare gain per dollar results from a reduction in the top tax rate t_5. Examination of the lower half of Table 3 confirms that the marginal policy reform for each value of inequality aversion is not affected by a reduction in α, involving a greater degree of scale economies within households.

4.3 The Use of Net Income

Table 5 presents marginal tax reform results when using net income per adult equivalent person in the social welfare function. This shows that the results, in terms of the optimal single marginal tax rate adjustments, are the same when the welfare metric is net income instead of money metric utility. That is, the largest marginal welfare gains are obtained by reducing the top tax rate for $\varepsilon = 0.2$ or $\varepsilon = 0.8$ and by reducing the second tax rate for $\varepsilon = 1.4$. The smallest marginal welfare costs are obtained by increasing the first tax rate for $\varepsilon = 0.2$, and by increasing the top tax rate for $\varepsilon = 0.8$ or $\varepsilon = 1.4$. Table 5 also shows that, for $\varepsilon = 0.2$ and $\varepsilon = 1.4$, the results, in terms of the optimal direction of revenue-neutral marginal adjustments, are not affected when the welfare metric is net income instead of money metric utility.

Table 5: Values of $\left|\frac{\Delta W}{\Delta R}\right|$ Using Net Income

	Increase in t			Reduction in t		
		ε			ε	
rate	0.2	0.8	1.4	0.2	0.8	1.4
Scale parameter: $\alpha = 0.8$						
t_1	1.48	1.42	1.32			
t_2	1.60	1.49	1.34	1.45	1.34	1.20
t_3	1.75	1.49	1.24	1.56	1.31	1.07
t_4	1.62	1.23	0.90	1.73	1.31	0.95
t_5	1.62	1.01	0.63	2.78	1.75	1.09
Scale parameter: $\alpha = 0.4$						
t_1	2.05	1.94	1.77			
t_2	2.21	2.03	1.79	2.00	1.83	1.61
t_3	2.43	2.05	1.64	2.17	1.80	1.41
t_4	2.30	1.70	1.17	2.46	1.81	1.24
t_5	2.35	1.40	0.81	4.02	2.43	1.41

However, in the case of $\varepsilon = 0.8$, the direction of the optimal revenue-neutral change is modified by the use of net income instead of money metric utility in the welfare metric. Based on net income measures, the optimal change now involves a reduction in t_5 along with an increase in t_4. This is not, as before, a flattening of the rate structure from the bottom to the top of the income scale, but a flattening at the top end only. This result is essentially due to the substantially higher value of the marginal welfare gains for a reduction in the top tax rate when net income is used instead of money metric utility in the welfare metric.

4.4 Changes to Income Thresholds

The previous discussion has concentrated on marginal adjustments to income tax rates. However, it is of interest to consider the effects of changes in income thresholds. Clearly, any change in tax thresholds which do not move an individual into a different tax bracket do not affect the marginal rate faced, but there is a change in net income if the threshold is below gross income (since a portion of lower income is taxed at a different rate). Individuals who are moved across thresholds experience changes in both marginal and average rates (at the pre-change gross income level).

Table 6: Values of $\left|\frac{\Delta W}{\Delta R}\right|$ Using Money Metric Utility: Income Threshold Changes (Equivalence Scale Parameter $\alpha = 0.8$)

	Increase			Decrease		
Threshold	$\varepsilon = 0.2$	$\varepsilon = 0.8$	$\varepsilon = 1.4$	$\varepsilon = 0.2$	$\varepsilon = 0.8$	$\varepsilon = 1.4$
a_2	1.31	1.24	1.17	1.39	1.32	1.24
a_3	1.36	1.26	1.16	1.46	1.36	1.25
a_4	1.42	1.18	0.98	1.41	1.18	0.98
a_5	1.58	1.25	1.00	1.45	1.17	0.93

Table 6 shows absolute values of marginal welfare changes arising from increases and decreases of \$1,000 in annual income thresholds, using a value of $\alpha = 0.8$ for the adult equivalence scale parameter (along with $\theta = 0.6$). Increases in thresholds a_2 to a_5 involve reductions in aggregate revenue of \$1,449m, \$897m, \$256m, and \$59m respectively, and give rise to welfare increases. Reductions in the thresholds involve increases in revenue of \$1,374m, \$853m, \$271m and \$67m respectively, and are associated with welfare reductions. As for changes in tax rates, given ε and α, marginal welfare gains are smaller than the corresponding marginal welfare losses for the lowest thresholds a_2 and a_3, but larger for the top threshold a_5. However, the marginal welfare changes per dollar of revenue are generally smaller than when the tax rates are modified, and display smaller ranges. In addition, differences between marginal welfare gains and losses are more limited, which implies a lower potential for social welfare improvements following marginal revenue-neutral adjustments.

For inequality aversion of $\varepsilon = 0.2$, the optimal revenue-neutral marginal change implies an increase in a_5 along with a reduction in a_2. The effect is a yet longer income range between thresholds two and five. The use of a very high degree of inequality aversion of $\varepsilon = 1.4$ leads to changes which are the reverse of those arising from $\varepsilon = 0.2$. The value of $\varepsilon = 0.8$ implies similar changes to the use of $\varepsilon = 1.4$, except that the increase in a_3 takes first priority over a_2.

5 Actual Tax Policy Changes

The Australian income tax structure has been changed in several ways since 2003/4. In particular there have been annual changes in the thresholds, except for a_2 which has remained constant in nominal terms (and a_1 which remains at zero). Furthermore, all the marginal tax rates have been reduced slightly, except for t_3 which has remained unchanged. The 2008/9 thresholds and marginal rates are shown in Table 7. This table also shows the thresholds adjusted to 2003/4 prices using the consumer price index. Hence in real terms a_2 has been reduced, while the biggest changes have been the increase in a_5, which more than doubled, followed by the increases in a_3 and a_5. While the rate changes are quite small, the threshold changes could not be considered as 'marginal'. These changes effectively reflect two types of policy change, since they involve changes in the income tax structure which are revenue reducing.

Table 7: The Australian Income Tax 2009/2010

No.	Income Threshold (AU$ per year)	Inflation Adjusted (to 2003/4)	Marginal tax rate
1	$a_1 = 0$	0	$t_1 = 0$
2	$a_2 = 6,000$	5, 155	$t_2 = 0.15$
3	$a_3 = 35,000$	30, 072	$t_3 = 0.30$
4	$a_4 = 80,000$	68, 736	$t_4 = 0.38$
5	$a_5 = 180,000$	154, 657	$t_5 = 0.45$

From results reported above, the adoption of inequality aversion of $\varepsilon = 0.2$, would lead, given a policy of revenue reduction, to reductions in t_5, t_4 and t_3 along with increases in income thresholds a_5, a_4 and a_3: these rates and thresholds are listed in order, with the first mentioned giving the highest marginal welfare increase per dollar of reduced revenue. This pattern of changes is not precisely reflected in the actual policy changes, since the latter include a reduction in a_2 along with a reduction in t_2, with t_3 unchanged. Nevertheless, they are much closer to the tax reforms implied by the high values of $\varepsilon = 0.8$ and $\varepsilon = 1.4$.

This suggests that the adoption of a value of 0.2 for ε would imply a larger increase in the social welfare function, with either net income or money metric utility per adult equivalent person as the welfare metric, compared with the value judgements of higher aversion. Indeed, using the inflation adjusted thresholds and the new rates, compared with the actual 2003/4 structure, and using money metric utility (with $\alpha = 0.8$ and $\theta = 0.6$), the value of social welfare, as defined by (9), increases by 2.27 per cent for $\varepsilon = 0.2$, compared to 2.03 and 1.81 percent for inequality aversion of, respectively, 0.8 and 1.4. When net income is used as the welfare metric, the percentage increases in (abbreviated) social welfare are 5.58, 4.83 and 4.08 for inequality aversion parameters of 0.2, 0.8 and 1.4 respectively.[14]

Of course it would be extremely foolish to suggest that actual policy could be described as being decided by a single judge with sufficient information to plan policy reforms based on an evaluation function of the kind considered here. However, it is not unreasonable, as an exercise in ex-post rationalisation, to ask what value judgements might be found to be implicit in actual policy changes. In this case a relative inequality aversion of 0.2 implies changes which are closer to actual reforms than the higher values examined here.

6 Conclusions

This paper began with the observation that the extensive optimal tax literature does not provide, and was never expected to provide, clear guidance about tax structures. Instead, the literature 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. In looking for practical advice, it was argued that a behavioural microsimulation model can provide the kind of detail needed, and

[14] In each case social welfare increases despite the increase in inequality arising from the threshold increases and lower tax rates.

can deal with the considerable complexity of actual tax systems and the large degree of population heterogeneity found in practice. While such models are not well-suited to generating optimal tax structures, an approach to the analysis of optimal marginal reforms was proposed and implemented using the Australian microsimulation model MITTS. Marginal welfare changes per dollar of revenue were obtained using additive Paretian welfare functions displaying constant relative inequality aversion. Two alternative welfare metrics were used in the social welfare function. One, for 'welfarist' functions, involves the use of money metric utility per adult equivalent person. A 'non-welfarist' alternative, that of net income, was also investigated.

It was found that the marginal welfare changes for the Australian income tax structure were not symmetric with respect to increases and decreases in tax rates, largely because of the asymmetry in tax revenue changes arising from differential labour supply effects in different ranges of the income distribution. It was also found that the degree of relative inequality aversion plays a substantial role in the determination of the optimal direction of rate changes. For moderate levels of aversion, a flattening of the rate structure is recommended. However for high degrees of aversion, a further increase in marginal rate progression arises as the preferred direction of change leading towards an optimal structure. The extent of inequality aversion was found to play a much larger role than the form of the welfare metric or the specification of adult equivalence scales. However, although the specification of adult equivalence scales did not affect the direction of optimal tax reforms, the form of the welfare metric was found to modify, for one particular value of inequality aversion, the direction of the optimal revenue-neutral tax reform.

Changes to income tax thresholds were also investigated. Overall, they lead to smaller marginal welfare changes than when the tax rates are modified. Finally, results were compared with the actual income tax structure changes made to the Australian tax structure since 2003/04 regarding income tax rates and thresholds. It was found that, as anticipated from the results for marginal welfare changes, the actual changes lead to the highest increases in social welfare in the case of the low, rather than high, relative inequality aversion parameter investigated.

Appendix A: MITTS - The Melbourne Institute Tax and Transfer Simulator

This appendix provides a brief description of the Melbourne Institute Tax and Transfer Simulator (MITTS), a behavioural microsimulation model of direct tax and transfers in Australia.[15] Since the first version was completed in 2000, it has undergone a range of substantial developments. MITTS is based on the Australian Survey of Income and Housing Costs (SIHC), a representative sample of the Australian population, containing detailed information on labour supply and income from different sources, in addition to a variety of background characteristics of individuals and households. All results are aggregated to the population level using the household weights provided with SIHC. Pre-reform net incomes at alternative hours levels are based on the MITTS calculation of entitlements, not the actual receipt. Furthermore, MITTS applies only income tests, as there is at present no asset imputation in the model. All major social security payments, family payments, rebates and income taxes are included, ensuring a reasonable approximation to net income.

MITTS consists of two components. MITTS-A is the arithmetic tax and benefit modelling component and provides, using the wage rate of each individual, the budget constraints that are crucial for the analysis of behavioural responses to tax changes. For those individuals in the data set who are not working, an imputed wage is obtained. MITTS-B examines the effects of any specified tax reform, allowing individuals to adjust their labour supply. Behaviour is based on quadratic preference functions where the parameters are allowed to vary with individuals' characteristics. Individuals are considered as being constrained to select from a discrete set of hours levels. For singles, 11 discrete points are distinguished. For couples, two sets of discrete labour supply points are used. The female hours distribution covers a wider range of part-time and full-time hours than the male distribution, which is mostly divided between non-participation and full-time work. Therefore, women's labour supply is divided into 11 discrete points, whereas men's labour supply is represented by just 6 points. The joint labour supply of couples is estimated simultaneously, unlike a popular approach in which female labour supply is estimated with the spouse's labour supply taken as

[15] For an overview of refereed publications and books relating to the MITTS model, see: http://www.melbourneinstitute.com/labour/behavioural/mitts_related_pubs.pdf

exogenous. Thus for couples there are 66 possible joint labour supply combinations.

Simulations are probabilistic, as utility at each hours level is the sum of a deterministic component (depending on hours worked and net income) and a random component. Hence MITTS generates a probability distribution over the discrete hours levels. The self-employed, disabled, students and those over 65 have their labour supply fixed at observed hours. Simulations begin by recording the discrete hours level for each individual that is closest to the observed hours level. The deterministic component of utility is obtained using the parameter estimates of the quadratic preference function. To generate the random component, a draw is taken from the distribution of the error term for each hours level (an Extreme Value Type I distribution). The utility-maximising hours level is found by adding the two components of utility for each hours level and choosing the hours with the highest utility. Draws from the error terms are taken conditionally on the observed labour supply; that is, they are taken in such a way that the optimal pre-reform labour supply is equal to the actually observed labour supply. As a result, post-reform labour supply is simulated conditional on the observed pre-reform labour supply. A user-specified number of draws is produced.

For the post-reform analysis, the new net incomes cause the deterministic component of utility at each hours level to change, so using the same set of draws from the calibration stage, a new set of optimal hours of work is produced. This gives rise to a probability distribution over the set of discrete hours for each individual under the new tax and transfer structure.

References

[1] Amiel, Y., Creedy, J. and Hurn, S. (1999) Measuring attitudes towards inequality. *Scandinavian Journal of Economics*, 101, pp. 83-96.

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

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

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

[5] Banks, J. and Johnson, P. (1994) Equivalence scale relativities revisited. *Economic Journal*, 104, pp. 883-890.

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

[7] Bourguignon, F., Fournier, M. and Gurgand, M. (1998) Distribution, development and education: Taiwan, 1979-1994'. (Unpublished paper) Delta (Joint research Unit, ENS-CNRS-EHESS), Crest (Laboratoire de Microéconométrie) and Centre d'études de l'emploi and Lamia (Université Paris I).

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

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

[10] Cai, L., Kalb, G., Tseng,Y.-P. and Vu, H.H. (2008) The effect of financial incentives on labour supply: evidence for sole parents from microsimulation and quasi-experimental evaluation", *Fiscal Studies*, 29, pp. 285-325.

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

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

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

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

[15] Creedy, J., Kalb, G. and Scutella, R. (2006) Income distribution in discrete hours behavioural microsimulation models: an illustration. *Journal of Economic Inequality*, 4, pp. 57-76.

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

[17] Creedy, J., Duncan, A.S., Harris, M., and Scutella, R. (2002) *Microsimulation Modelling of Taxation and The Labour Market: The Melbourne Institute Tax and Transfer Simulator*. Cheltenham: Edward Elgar.

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

[19] 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.

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

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

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

[23] Jenkins, S.P. and Cowell, F.A. (1994) Parametric equivalence scales and scale relativities, *Economic Journal*, 104, pp. 891-900.

[24] Kalb, G. and Lee, W.-S. (2007) Wage and Labour Force Participation in Australia: Accounting for Categorised Observations on Wages. Final Report for Treasury.

[25] Kalb, G. and Lee, W.-S. (2008) Labour Supply in the Australian Population: Using Data from 1999 to 2004. Final Report for Treasury.

[26] 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.

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

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

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

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

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

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

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

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

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