TAMPERE ECONOMIC WORKING PAPERS
NET SERIES

ON OPTIMAL NON-LINEAR INCOME TAXATION: NUMERICAL RESULTS REVISITED

Matti Tuomala

Working Paper 17
September 2003
http://tampub.uta.fi/econet/wp17-2003.pdf

DEPARTMENT OF ECONOMICS

FIN-33014 UNIVERSITY OF TAMPERE, FINLAND

ISSN 1458-1191
ISBN 951-44-5774-9

On optimal non-linear income taxation: numerical results revisited*

by

Matti Tuomala[#]

Abstract: Based on numerical simulations there seems to be a kind of consensus in the optimal tax literature that the marginal tax rate should fall rather than rise with income. Retaining the same formal structure as in Mirrlees (1971) this paper shows that this consensus is sensitive to a choice of the assumed form of utility of consumption. For the utility function quadratic in consumption optimum tax schedules look rather like those traditionally chosen by governments, i.e. the marginal tax rates rise with income.

* I am grateful to Jim Mirrlees from very helpful comments.

Corresponding author. Address: Department of Economics, 33014 University of Tampere, Finland. Email: ktmatuo@uta.fi

1. Introduction

In the numerical calculations that followed Mirrlees's (1971) pioneering contribution marginal tax rates decrease with income for the vast majority of the population. This picture is at odds with observed patterns in most countries. There are, however, in the optimal income tax literature some counter examples. Kanbur-Tuomala (1994) show that increases in inherent inequality can alter the qualitative pattern of optimal marginal tax rates. The optimal graduation can indeed be such that marginal tax rates increase for the majority of the population, but there continues to exist a significant income range at the top where marginal tax rates decline. Diamond (1998) assumes a Pareto distribution of skills rather than the lognormal distribution that Mirrlees and other have assumed. Using quasi-linear preferences he finds rising marginal tax rates on those above the modal skill level to be optimal[1]. [2]

Should we then conclude that the existing optimal non-linear income tax literature does not provide support for the progressive marginal tax rates? We think it is premature to reach this conclusion There are still unexplored questions in the standard non-linear continuum model. Although the numerical simulations that followed Mirrlees (1971) have tried many variations on his basic parametrization, it is slightly surprising to notice that essentially only two utility functions – a Cobb-Douglas and a CES with the elasticity of substitution between consumption and leisure being 0.5- are employed in these simulations. In particular, the role of the assumed form of utility of

[1] Progressive marginal rates may be optimal under non-standard assumptions e.g when people care about relative levels of consumption (Tuomala,1990) , when there is significant wage uncertainty (Tuomala, 1990, Low and Maldoom,2002), and when earnings are allowed to be less than perfectly correlated with productivity (Bevan,2002).

consumption in determining the shape of the tax schedule is not clear. Dahan-Strawczynski (2000) show, however, that a rising marginal tax rates at high incomes presented by Diamond (1998) depend on the assumption of utility of consumption. Their simulations focus only on high levels of income, not the whole schedule. It is our aim to show numerically how sensitive marginal tax rate structure is to the assumed form of utility of consumption. We replace the assumptions of utility of consumption used in previous simulations to quadratic utility of consumption. This form is essentially less curvature than those used in previous simulations.

The plan of the paper is as follows. In section 2 we set up the basic Mirrlees (1971) model and highlight the role of the underlying factors in determining the shape of optimal non-linear tax schedule. However, analytical characterisation does not lead far. In section 4 using numerical simulations we study the role of the assumed utility of consumption in determining the pattern of optimal non-linear taxation. Section 4 concludes.

2. The Mirrlees model

It useful to lay out the basic model, even though it is well-known. There is a continuum of individuals, each having the same preference ordering, which is represented by an additive utility function $u = U(x) - V(y)$ defined over consumption x and hours worked y, with $U_x > 0$ and $V_y < 0$ (subscripts indicating partial derivatives) and where V(.) is convex. Workers differ only in the pre-tax wage n they can earn. There is a distribution of n on the interval (s,h) represented by the density function f(n).Gross income $z = ny$.

Suppose that the aim of policy can be expressed as maximizing the following social welfare criterion

[2] Boadway et al (2000) provide a full characterization of the solution when preferences are quasi-linear in leisure.

$$S = \int_{s}^{h} W(u(n)) f(n) dn \ , \qquad (1)$$

where W(.) is an increasing and concave function of utility. The government cannot observe individuals' productivities and thus is restricted to setting taxes and transfers as a function only of earnings, T[z(n)]. The government maximizes S subject to the revenue constraint

$$\int_{s}^{h} T(z(n)) f(n) dn = R \qquad (2)$$

where in the Mirrlees tradition R is interpreted as the required revenue for essential public goods. The more non-tax revenue a government receives from external sources, the lower is R. In addition to the revenue constraint, the government faces incentive compatibility constraints. These in turn state that each n individual maximizes utility by choice of hour.

Totally differentiating utility with respect to n, and making use of workers utility maximization condition, we obtain the incentive compatibility constraints,

$$\frac{du}{dn} = -\frac{yV_y}{n} . \qquad (3)$$ [3]

Since T = ny-x, we can think of government as choosing schedules y(n) and x(n). In fact it is easier to think of it choosing a pair of functions, u(n) and y(n), which maximize welfare index (1) subject to the incentive compatibility condition (3) and the revenue requirement (2). Omitting details (for an

[3] The 1.order condition of individual's optimisation problem is only a necessary condition for the individual's choice to be optimal, but we assume here that it is sufficient as well. Assumptions that assure sufficiency are provided by Mirrlees (1976). Note also that while we here presume an internal solution for y, (3) remains valid even if individuals were bunched at y=0 since, for them, du/dn=0.

exposition see Tuomala ,1990), the first order conditions of this problem imply a pattern of marginal rates[4], t(z) = T'(z), satisfying

$$\frac{t}{1-t} = \frac{(e^{-1}+1)U_x}{\lambda n f(n)}\mu(n) \tag{4}$$

where λ is the multiplier on the revenue constraint and

$$\mu(n) = \int_s^n ((W'U_x - \lambda)(1/U_x)f(p)dp. \tag{5}$$

is the multiplier on the incentive compatibility constraint. This latter satisfies the transversality conditions

$$\mu(s) = \mu(h) = 0. \qquad (6).$$

As in Atkinson-Stiglitz (1980) $e = V'/yV''$. It is the elasticity of labour supply with respect to net wage, holding marginal utility of income constant, i.e. e is “compensated” wage elasticity in rather unusual sense.

It should be clear from (4) and (5) that the variation of the optimal marginal tax rate with the level of income is a complex matter.[5] [6] To see the complications that arise, consider equations (4) and (5).

---

[4] There are other papers that have looked at alternative derivations and formulae for non-linear taxation, see Revesz (1989), Roberts (2000) and Saez (2001)

[5] Equations (4) - (6) lead to the few qualitative conclusions available in this framework (see Tuomala, 1990). It can be shown that the marginal tax rate on income is nonnegative. This is more striking than it at first looks. It may very well be optimal to have the average tax rate less than zero, but it is never optimal to subsidize earnings at margin. An intuition is that it is cheaper to get people to given indifference curve by reducing average rate rather than by exacerbating deadweight loss through distorting their labour supply decisions. It can also be shown that the marginal tax rate is less than one. We also have the famous "end point" results. If wage distribution is bounded above, then the marginal tax rates at the top is zero. If it is optimal for least able individual to work then the marginal tax rate on least able is zero. An intuition behind these endpoint results is that only reason to have a marginal tax rate differing from zero is to raise an average tax rate above that point and lower it below i.e. equity considerations. But at the top is no one to take from and at the bottom there is no one to give to. So at the end points only efficiency considerations matter. Numerical solutions (Tuomala, 1984) have shown, however, that these results have very little practical relevance.

First, the term $(1+e^{-1})$, reflecting also conventional wisdom, (4) says that, other things equal, the marginal tax rate should be lower the larger is the compensated elasticity of labour supply, e. Governments fearing that disincentive effects are large will tend to set lower marginal rates. The term, U_x, denotes the marginal utility of consumption representing income effects. It also appears in the third term, $\mu(n)$, that incorporates distributional concerns. Hence it is obvious that the functional form of U_x has the important role in determining the shape of the schedule. The fourth term, nf(n), in turn is an indicator of the extent of earnings at the wage level n. Equation (4) also suggests that, other things equal, the marginal tax rate should be lower the denser the population at that point, i.e. higher f(n). In other words the more people affected, the higher is deadweight loss. On the other hand for the typical distribution the density weighted by n, (eg. lognormal distribution) is likely to decline with n above some point suggesting a higher marginal tax rate on high earners. The term $\mu(n)$ in turn increases with n for low n and decreases with n for high n. The turning point depends on λ[7]. The lower is λ, the higher is the n at which the turning point occurs. Thus as the revenue requirement falls, and hence λ falls, the range over which $\mu(n)$ is increasing stretches further. Since $\mu(n)$ affects the marginal tax rate positively, this means that the range over which the latter increases also stretches further - at least for this reason. In this sense, therefore, more tax revenue leads to a less progressive tax structure.

But this is about as far as we can get at this level of generality[8]. Of course, following studies e.g Atkinson (1995), Diamond (1998) and Dahan and Strawczynski (2000) we can assume income effects away. Then it is possible to deduce on the basis of (4) that with Pareto and lognormal distributions marginal tax rates rise with income at high levels of income. Once we assume income effects this not any more possible. A concave utility of consumption means that income effects are weaker for high income earners which pushes marginal tax rates down at high levels of income. On the other hand when U_x decreases with n its impact on inequality aversion pushes in an opposite

---

[7] It can be shown that the left hand side of (5) is decreasing in n* (n* is the skill level at which $W'(u(n^*))U_x[x(n^*)]=\lambda$) so long as W(u) is concave and leisure is normal.

direction. The interaction of these two forces determines the optimal tax rate schedule. In the tradition of the non-linear taxation literature, we can provide better understanding of the form of optimal policy through numerical simulations. We can compute post tax income at each level of n, and thus calculate inequality of pre and post tax income as well as total income, for different values of key parameters. The next section takes up this task.

3. Model calibration and numerical results

Following the lead of Mirrlees (1971) numerical calculations have proved useful in generating useful results[9]. We follow this route here. We assume n to be distributed lognormally with parameters m and σ (see Atichison and Brown, 1957). This assumption is common in the literature, following Mirrlees (1971). For numerical simulations we choose σ = 0.39, 0.5 and 0.7 as a standard deviation of n and mean n = 0.4. Preferences over consumption and working time are given by the following utility function that is quadratic in consumption (quadratic approximation)[10]

$$u = (x-1) - 5(x-1)^2 - \frac{1}{(1-y)} \quad ; \; x \in (0,1) \qquad (7)$$

The previous simulations (see Tuomala, 1990, Kanbur-Tuomala 1994)) have used either the logarithmic utility of consumption or U(x)=-1/x. The utility of consumption in (7) is a kind of “mixture” of these two forms. As displayed in Figure 1 the utility of consumption in (7) is essentially less curvature than those used in the previous simulations.

---

[8] There are also a number of asymptotic results for the unbounded case in Mirrlees (1971)
[9]Tuomala (1984, 1990) gives details of the computational procedure.

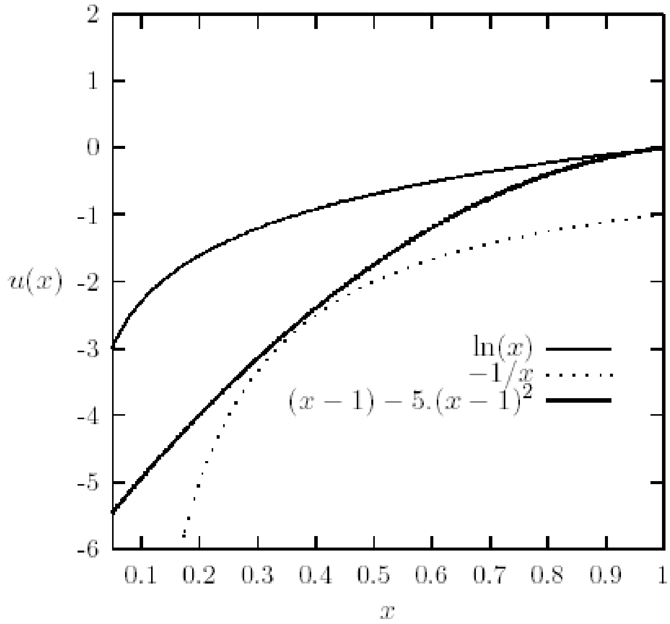

Figure 1

The social welfare function of the recipient government is specified as $S(u) = -\frac{1}{\beta}e^{-\beta u}$ so that ß measures the degree of inequality aversion in the social welfare function of the government (in the case of ß = 0, we define W = u). R=$1 - \int x(n)f(n)dn / \int y(n)f(n)dn$ is specified as a fraction of national income, and is assumed to be 0.0 (pure redistributive) and 0.1.

The Tables 1-4 give labour supply, y, gross income, z, net income, x and optimal average (ATR) and marginal tax rates (MTR) at various percentiles of the ability distribution. Tables also provide the decile ratio (P90/P10) for net income and gross income. Unlike the scalar inequality measures the use of fractile measures such as the decile ratio allows us to consider changes in inequality at various different points in the distribution. Since marginal tax rates may be a poor indication of the redistribution powers of an optimal tax structure we measure the extent of redistribution, denoted by

[10] Stern (1986) discusses the quadratic specification and other alternatives. In fact in the recent empirical labour supply studies, see Blundell et al (1999) and Keane and Moffitt (2001), preferences over working time and net income are

RD, as the proportional reduction between the decile ratio for market income, z, and the decile ratio for disposable income, x.

Table 1

β = 0	σ=0.39	R= 0.0			
		S=-5.98			
F(n)	y	z	x	ATR%	MTR%
0.10	0.32	0.07	0.08	-14	6
0.50	0.44	0.16	0.16	-2	9
0.90	0.51	0.31	0.30	4	13
0.99	0.55	0.46	0.46	9	17
P90/P10		4.27	3.63		
RD(%)			15.1		

Table 2

β = 1	σ=0.39	R= 0.0			
		S=-6.01			
F(n)	y	z	x	ATR%	MTR%
0.10	0.26	0.06	0.09	-46	20
0.50	0.38	0.14	0.15	-5	27
0.90	0.47	0.29	0.13	13	31
0.99	0.53	0.48	0.39	20	30
P90/P10		4.93	2.94		
RD(%)			40		

given by the utility function that is quadratic in hours and net income.

Table 3

β = 0	σ=0.5	R= 0.1			
		S=-6.04			
F(n)	y	z	x	ATR%	MTR%
0.10	0.27	0.05	0.05	0	8
0.50	0.43	0.16	0.15	7	14
0.90	0.52	0.36	0.32	13	21
0.99	0.55	0.64	0.53	18	27
P90/P10		6.72	5.85		
RD(%)			13		

Table 4

β = 0	σ=0.7	R= 0.0			
		S=-5.47			
F(n)	y	z	x	ATR%	MTR%
0.10	0.15	0.02	0.07	-100	10
0.50	0.39	0.14	0.17	-16	22
0.90	0.51	0.46	0.38	16	37
0.99	0.52	0.96	0.68	30	45
P90/P10		6.72	5.85		
RD(%)			75		

The striking thing in the numerical results shown in Tables 1-4 is that once we assume that preferences are given by the utility function that is quadratic in consumption, the shape of optimum tax schedules may be altered drastically. The marginal tax rates rise with income over the whole range. In all four cases, the optimal income tax rate schedule features marginal tax rate progressivity, except that in case 1 (Table 1) rates decline slightly at the very top of wage distribution. In fact the marginal rate rises after the ninety-ninth percentile point. In the case of Table 3 the marginal tax rate increases with income up to the 99.7 percentile point. Of course it is still true that the marginal tax rate at the top is zero. Hence our results also confirm how local is no distortion at the top-result. In particular, when the standard deviation of n is 0.5 or 0.7 marginal tax rate structures look like those observed in most developed countries.

Tables 1-4 also confirm well-known earlier results such as that optimal tax/transfer systems become more progressive when inherent inequality increases, $\sigma = 0.7$, $\sigma = 0.5$ and $\sigma = 0.7$. As also expected, greater inequality aversion leads higher marginal rates.

6. Conclusions

Based on numerical simulations there seems to be a kind of consensus in the optimal tax literature that the marginal tax rate should fall rather than rise with income. Retaining the same formal structure as in Mirrlees (1971) this paper shows that this consensus is sensitive to a choice of the assumed form of utility of consumption. For the utility function quadratic in consumption optimum tax schedule look rather like those traditionally chosen by governments, i.e. the marginal tax rates rise with income. Although it is not usual to find support in the optimal income tax literature for progressive marginal tax rates, our study offers some support to defenders of graduated rates[11]. It is also interesting to note that the marginal rates at the bottom are low and the differences between lowest and highest rates are rather big.

[11] Zelenak and Moreland (1999) provide a very interesting discussion on this question.

References

Aitchison, J. & Brown,J.A.C., The lognormal distribution with special reference to its uses in economics, Cambridge University Press,1957.

Atkinson.A.B. and Stiglitz,J , Lectures on Public Economics, McGraw Hill,1980.

Atkinson,A.B, Public economics in Action, The basic income/flat tax proposal, The Lindahl Lectures, Oxford University Press, 1995.

Bevan, D., Optimum income taxation when earnings are imperfectly correlated with productivity, working paper, Oxford University, 2002.

Blundell, R, Duncan,A, McCrae,J and Meghir,C. Evaluating In-Work Reform: The working families' tax credit in the UK, unpublished paper,1999.

Boadway,R.., Cuff,K and M.Marschand, Optimal income taxation with quasi-linear preferences revisited, Journal of Public Economic Theory, 2000.

Diamond, P.: Optimal income taxation: an example with a U-shaped pattern of optimal marginal tax rates, American Economic Review, 1998. 88(1), pp.83-95

Dahan M. and Strawczynski M., The optimal non-linear income tax, American Economic Review.90(3), pp.681-686, 2001

Immonen, R.,Kanbur, R.,Keen, M. & Tuomala, M, Tagging and taxing: The optimal use of categorical and income information in designing tax/transfer schemes, Economica, pp.179-192,1998

Kanbur, R. and Tuomala, M, Inherent inequality and the optimal graduation of marginal tax rates, Scandinavian Journal of Economics 96 (2), pp.275-282, 1994.

Keane, M and Moffitt, R, A structural model of multiple welfare program participation and labour supply, International Economic Review, 39(3),pp.553-589,1998

Mirrlees, J.A., An exploration in the theory of optimum income taxation, Review of Economic Studies 38, pp.175-208,1971.

Mirrlees, J.A., Optimal tax theory, a synthesis, Journal of public Economics 6,327-58, 1976.

Low, H and D.Maldoon, Optimal Taxation, Prudence and Risk-Sharing, Journal of Public Economics, forthcoming, 2003

Revesz, J., The optimal taxation of labour income, Public Finance 44,pp.453-75, 1989.

Roberts, K., A reconsideration of optimal income tax, in Incentives, Organization and public Economics, papers in Honour of Sir James Mirrlees, eds by P.Hammond and G.Myles, Oxford University Press, 2000.

Saez, E. Using elasticities to derive optimal income tax rates, Review of Economic Studies 68, pp.205-229, 2001.

Stern, N. On the specification of labour supply functions, in Unemployment, search and labour supply, edited by R. Blundell and I.Walker, Cambridge University Press,1986.

Tuomala,M., On the optimal income taxation: some further numerical results, Journal of Public Economics 23, 351-66, 1984.

Tuomala, M, Optimal income tax and redistribution, Clarendon Press, Oxford, 1990.

Zelenak,L and K.Moreland Can the Graduated Income Tax Survive Optimal Tax Analysis? Working paper No.149, The Center for Law and Economic Studies, Columbia University School of Law, 1999.