TAMPERE ECONOMIC WORKING PAPERS

NET SERIES

ON OPTIMAL INCOME TAXATION WITH HETEROGENOUS WORK PREFERENCES

Ritva Tarkiainen
Matti Tuomala

Working Paper 32
August 2004
http://tampub.uta.fi/econet/wp32-2004.pdf

DEPARTMENT OF ECONOMICS AND ACCOUNTING

FI-33014 UNIVERSITY OF TAMPERE, FINLAND

ISSN 1458-1191
ISBN 951-44-6081-2

On Optimal Income Taxation with Heterogenous Work Preferences*

Ritva Tarkiainen
Department of Mathematics,
University of Jyväskylä,
Finland

Matti Tuomala
Department of Economics
University of Tampere
P.O.Box 607
Fin-33101 Tampere
Finland
e-mail: ktmatuo@uta.fi

17.8.2004

Abstract

This paper considers the problem of optimal income taxation when individuals are assumed to differ with respect to their earnings potential and work preferences. A numerical method for solving this two-dimensional problem has been developed. We assume an additive utility function, and utilitarian social objectives. Rather than solve the first order conditions associated with the problem, we directly compute the best tax function which can be written in terms of a second order B-spline function. Among our findings are that marginal tax rates are higher than one might anticipate, and that very little bunching occurs at the optimum.

Key words: Non-linear income taxation, heterogenous work preferences, two-dimensional population, numerical simulation.
JEL classification: C63, H21, H24.

*We thank Vidar Christiansen, Sören Bo Nielsen, Jukka Pirttilä, seminar participants at EPRU, Copenhagen Business School, and two anonymous referees for very helpful comments on an earlier version. We are also grateful to the Yrjö Jahnsson Foundation for financial support.

1 Introduction

There is a large literature, initiated by Mirrlees (1971), that addresses the optimal design of nonlinear income taxes. In this literature, the issue of incentives for supplying labour is tackled directly by modelling individuals as choosing between work and leisure given the tax-transfer schedule they face. There are assumed to be a large number of individuals, differing only in a single parameter, the pre-tax wage they can earn. In other words people have identical work preferences but face unequal opportunities in the labour market. The government then chooses a schedule that maximizes a social welfare function defined on individuals' welfare, that is, on the utility they derive from their consumption-leisure bundles.

Although the models of the one-dimensional population have been useful for computations and examinations of optimal income tax problem, they are not in all respects accurate pictures of reality (see Tuomala (1990) for a survey of this literature). To analyze redistributional policies more fully, it would be useful to consider situations where individuals are characterized by more than just one parameter. In the simple model with identical preferences, all workers earning the same level of income also have the same wage rates (marginal productivities). When preferences vary in unspecified ways across individuals, this will not be the case any more. Some people will be earning a given income because they are more productive, while others because they are more hard working. The assumption that differences in earnings are completely to be explained by differences in preferences (working or leisure preferences) is obviously an unrealistic one, just as the alternative explanation in terms of ability differences is also a simplification of reality.[1] A more realistic model should take into account of both.[2]

How might these factors including to the optimal income tax model affect our views as to the optimal level of income taxation and transfer payments?

[1] There are still other ways to relax the homogeneity assumption. One way is to differentiate the population by easily observable indicators that are correlated with the unobservable characteristic of interest. An individual's labour market status or demographic attributes, for instance, may convey information on underlying ability. The theory of the optimal use of such information was first considered by Akerlof (1978) and developed further by several authors (for a recent contribution to this tradition see Immonen-Kanbur-Keen-Tuomala, 1998).

[2] Although economists typically think of preferences in terms of psychological characteristics of the individual in question, those characteristics may also have non-psychological dimensions. Sandmo (1993) noted that there is a parallel here to Sen's discussion of capabilities (see Sen, 1992). Namely there is an important general problem of interpersonal variations in converting incomes into the actual ability of an individual to do this or be that. This could be rephrased to refer to the work preferences in our model.

Should this, as many seem to believe, be an argument for a more regressive tax/transfer systems. In fact considerations of that kind can be seen to lay behind income tax reforms implemented in several countries during last decades.[3] One of our objects is to attempt to answer this question.

This paper builds on a model in which individuals are assumed to differ in abilities and work preferences. It is clear that the theory of optimal nonlinear income taxation becomes more complicated since we allow for individuals to differ in more than one dimension. Papers in this area are few. Mirrlees (1976, 1986) sets out the framework of a multidimensional optimal tax problem. He derives first order conditions for a tax schedule be optimal. It is well-known that formal results about the optimal (one-dimensional) income tax are limited. Hence it is not surprising that Mirrlees (1976, 1986) has no general results concerning the shape of the optimal tax schedule in the multidimensional case.[4]

In the context of nonlinear pricing Armstrong (1996) and Wilson (1993) have analyzed the problem of multidimensional types.[5] Armstrong (1996) shows a class of cases that allow explicit solutions. Wilson (1993) also computes numerical solutions. These contributions are to some extent important in the context of optimal income taxation. As is typically the case with nonlinear pricing literature, Armstrong (1996), Rochet-Choné (1998) and Wilson (1993) exclude income effects. The motivation for ignoring income effects when constructing nonlinear tariffs for services and goods offered to household customers is that their income elasticities are small and/or their residual incomes are large in relation to their expenditures on the nonlinearly priced goods and services. This assumption cannot be easily justified in the context of an optimal income taxation model. In particular, in conjunction with the utilitarian objective it eliminates the equity considerations that motive the income tax problem in the first place. Therefore we face a more complex problem here.

As is often the case with the one-dimensional optimal income tax model we have to rely on numerical simulations. In a two-dimensional model this

[3]It is not obvious that one would want to redistribute income from those who prefer to work much to those who prefer to work little. Many people, politically on the right, seemed to believe in that way.

[4]In Ebert (1988) the problem is considered with some strict restrictions on the utility function which allows to transform the two-dimensional problem into the one-dimensional one. Sandmo (1993) (section 6) discusses the difficulty to sign marginal tax rate in a linear income tax model when individuals differ both in tastes and productivities.

[5]See also Rochet (1985), Laffont, Maskin and Rochet (1987) and McAfee and MacMillan (1988). For a recent important contribution to the nonlinear pricing of multidimensional types see Rochet-Choné (1998).

is even more the case. Moreover now the computation is much demanding.

The structure of the paper is as follows. Section 2 presents the formulation of two-dimensional optimal nonlinear income tax problem introduced by Mirrlees (1976, 1986) and derives necessary conditions for a solution. In section 3 we adopt a computational approach to the problem and the appendix presents the basic outlines of our computational method. Section 4 reports a range of numerical calculations. Section 5 concludes the paper.

2 Mirrlees' formulation

The economy consists of many different types of individuals who are distinguished both by earnings abilities, denoted by t and work (or leisure) preferences, denoted by s. Thus each individual is characterized by a vector $[t, s]'$ of type parameters that varies among individuals. Both characteristics, t and s, are assumed to be private information that is not available to the government. The distribution of these type parameters in the population is given by a density function f such that $f(t, s) \geq 0$ on a rectangular domain $\Omega = [t_0, t_1] \times [s_0, s_1]$. There are two commodities in an economy, namely a consumption good x and labour supply y. The economy is competitive so that pre-tax pay in each job is the workers marginal product in that job. Hence his gross income z is given by $z = t\,y$. The government knows that when it provides a nonlinear income tax schedule $T : \mathbf{R}^+ \to \mathbf{R}$, where $T(z)$ is the tax paid on gross income z, each individual maximizes his or her concave utility function of the following additive form[6]

$$u(x, y) = g(x) + s\,h(y), \tag{1}$$

subject to

$$x + T(z) = z, \;\; z = t\,y \tag{2}$$

in choosing his or her labour supply. We assume that $u \in C^2$, $g_x > 0$ and $h_y < 0$ for all x, $y \geq 0, y < 1$. The higher is the value of s, the stronger is the preference for leisure or the weaker is the preference for work (or the greater is the disutility of work). The additivity assumption in (1) has very strong implications. Namely the marginal utility of consumption is independent of work preferences.

The labour supply-behaviour can be modelled by first-order partial differential equations. These are so called incentive compatibility conditions.

[6]This assumption permits interpersonal comparability.

Next we introduced formally these conditions. For that we define the maximum value function $v : \Omega \to \mathbf{R}$ by

$$v(t,s) = \max_{x,y}\{g(x) + s\,h(y) : x + T(z) = z,\ z = t\,y\}. \tag{3}$$

For each (t,s) we denote the optimum of (1)–(2) by $(x(t,s),\ y(t,s))$, and then we have

$$v(t,s) = g(x(t,s)) + s\,h(y(t,s)). \tag{4}$$

It is assumed that x and y are differentiable with respect to t and s. Then by differentiating (4) with respect to t and s and making use of necessary first order conditions of individual's maximization problem (1)–(2) we obtain an envelope (or incentive compatibility) conditions

$$v_t = -s\,y\,h_y/t \tag{5}$$

and

$$v_s = h \tag{6}$$

for all $(t,s) \in \Omega$.

The conditions (5) and (6) are only necessary for the individuals' choices to be optimal. Sufficient conditions for a global maximum are considered in Mirrlees (1976). We only assume that the conditions (5) and (6) are also sufficient for a global maximum, so that we can substitute the individuals' utility maximization conditions by the weaker conditions (5) and (6).

Given the tax schedule T, the government can calculate the gross income $z(t,s)$ and the consumption (net income) $x(t,s)$ for an individual with characteristics (t,s). The problem of the utilitarian government in choosing the optimal income tax schedule can be described as follows

$$\max_{x,y}\{W = \int\!\!\int_{\Omega} v(t,s)\,f(t,s)\,dt\,ds\} \tag{7}$$

subject to the revenue constraint

$$\int\!\!\int_{\Omega} T(z(t,s))\,f(t,s)\,dt\,ds = \int\!\!\int_{\Omega} [t\,y(t,s)) - x(t,s)]\,f(t,s)\,dt\,ds = R, \tag{8}$$

where R is revenue requirement.

But if the government chooses a tax schedule T, the individuals react to this schedule and modify their labour supply behaviour. These reactions are presented by the conditions (5), and (6).

Adding utilities in (7) may face some ethical objections. Namely if individuals are identical in work preferences, equal marginal utilities of all coincides with equal total utilities. Sen (1973) pointed out that with diversity of human beings (eg. different work preferences in our case) the two can pull in opposite directions.[7] For those who prefer to think of the justification for redistribution as being based on inequality of opportunity, differences in preferences may provide a suitable basis for distinguishing economic rewards but differences in abilities in turn do not. This point of view raises questions on the nature of the parameter s. Namely it may be argued that both attributes (working preferences and productivity) are "circumstances of birth". Or as Sandmo (1993) pointed out there is a very fine dividing line between differences in preferences that are due on the one hand to physiological characteristics and on the other hand to psychological attitudes to work. Therefore, it is far from clear how unsuitable is the utilitarian criterion in our case. Moreover, it has been the dominant framework in the optimal income tax literature.

We may eliminate x by inverting (4) so that $x(t,s) = \Gamma(v(t,s), y(t,s), t, s)$ for all $(t,s) \in \Omega$. Now by choosing y as a control and v as a state function, we can formulate the problem as an optimal control problem as follows:

$$\max_{y,v}\{W = \iint_\Omega v(t,s)\, f(t,s)\, dt\, ds\} \tag{9}$$

subject to the conditions (5), (6) and the revenue constraint

$$\iint_\Omega [t\, y(t,s)) - \Gamma(v(t,s), y(t,s), t, s)]\, f(t,s)\, dt\, ds = R. \tag{10}$$

Then we can construct a Lagrangean[8] by defining multipliers λ for (10) and $\nu(t,s)$ and $\alpha(t,s)$ for (5) and (6)

$$\begin{aligned} L(v,y,\lambda,\nu,\alpha) \quad = \quad & \iint_\Omega \{[v(t,s) - \lambda(\Gamma(v(t,s), y(t,s), t, s) - t\, y(t,s))]\, f(t,s) \\ & + \nu(t,s)[v_t + s\, y\, h_y/t] \\ & + \alpha(t,s)[v_s - h]\}\, dt\, ds \end{aligned} \tag{11}$$

Using Green's formula, taking into account that $\nu(t_0,s) = \nu(t_1,s) = 0, s \in [s_0,s_1]$, and $\alpha(t,s_0) = \alpha(t,s_1) = 0, t \in [t_0,t_1]$ (transversality conditions),

[7] In fact this led Sen to propose his weak equity axiom which says that redistribution should be directed to those with lower levels of utility.

[8] Following Wilson (1993) this formulation can be called as a relaxed version of the problem in which some of the possibly relevant constraints are omitted. Namely we know in a one-dimensional model that relatively weak conditions suffice to ensure that the solution the relaxed problem is the solution of the complete problem.

where t_0, s_0, t_1, s_1 are lower and upper bounds for t and s, and setting the derivatives of L with respect to y and v equal to zero we obtain

$$T'(z) = 1 + \frac{s\,h_y}{t\,g_x} = -\frac{\nu(t,s)\,s[h_{yy}\,y/t + h_y/t]}{\lambda\,t\,f(t,s)} + \frac{\alpha(t,s)\,h_y}{\lambda\,t\,f(t,s)} \tag{12}$$

$$(1 - \frac{\lambda}{g_x})\,f(t,s) = \nu_t + \alpha_s \tag{13}$$

where we have used the facts that $\Gamma_v\,g_x = 1, \Gamma_y = -\Gamma_v\,s\,h_y = -s\,h_y/g_x$. The equation (12) is the marginal tax rate formula. To say something about the properties of the tax schedule we should be able to deduce from (12) and (13) the sign of multipliers. Unfortunately, this is not possible. Namely, if we solve y and v from (12) and (13) as a function of $\lambda, \alpha, \nu, \alpha_s, \nu_t, t$ and s, substitution in the state equations yields a system of partial differential equations for α and ν. This system seems to be rather complicated even to be solved numerically.[9] Therefore we will adopt a direct method to solve this problem numerically. Wilson (1993,1995) describes various computational methods and solves some examples of nonlinear pricing without income effects.

It is still a necessary condition for an optimum that the marginal tax rate faced by the highest income earner should be zero.[10] This can be deduced from (12) and the transversality conditions. Thus, the optimal income tax schedule must have a zero marginal tax rate at the top, even when the preference structures underlying the work-behaviour of different consumers differ in any number of ways. The top-income person will now not necessarily be the one with highest wage rate.[11] This result is interesting in the sense that this property is not purely the result of an assumption that taxpayers differ only by a scalar parameter. (The additivity assumption in (1) is not crucial for the result here). Unfortunately these end-point results offer us little concrete guidance for tax policy purposes.

In sum we can conclude that analytically we have no results concerning the shape of the optimal schedule. Therefore computer simulation is the only

[9] In principle, the problem could be solved indirectly using necessary conditions, which must hold for a solution of the problem, considered in Cesari (1969) and Kazemi-Dehkordi (1984). However, in practice this turned out to be very difficult even in very simple cases.

[10] This is the counterpart of a result known as "no distortion at the top" in nonlinear pricing theory discussed by Armstrong (1996) and Rochet-Choné (1998).

[11] This does not affect our conclusions here, even when the preference structures underlying the work-behaviour of different individuals differ in any number of ways. An intuition behind this result is that the only reason to have a marginal tax rate differing from zero is to raise an average tax rate above that point but at the top is no one to take from.

way to gain further insights. In fact the most interesting results obtainable in the optimal income tax theory are numerical calculations for specific examples. It can be said that the very basic nature of income tax problems requires quantitative results. It is also good to remember that in modern physics and applied mechanics, where the many partial differential equations encountered are almost invariably solved numerically.

3 A computational approach

Tarkiainen–Tuomala (1998) developed a numerical method to solve a two-dimensional nonlinear income tax problem (the Lagrangean (12)).[12] Using this method we managed to solve the income tax problem only when the characteristics (t, s) are uniformly distributed. This experience and the complexity of the multidimensional nonlinear income tax problem suggest that an entirely different formulation and computational method might be useful in practice.

In this section we introduce a new numerical method for solving the optimal nonlinear income tax problem with two-dimensional population. The problem is stated alternatively in terms of taxes. In other words we attempt to find directly the tax function $T(z)$ that maximizes the social welfare function with constraints. Instead of using the incentive compatibility constraints (5)–(6) we take into account the individual's optimization problem directly.

Now we are looking for T^* which solves the following optimization problem (P):

$$W(T^*) = \max_T \, \{W(T) = \int\!\!\int_\Omega [g(z - T(z)) + s\, h(z/t)]\, f(t,s)\, dt\, ds\} \tag{14}$$

subject to

$$\int\!\!\int_\Omega T(z)\, f(t,s)\, dt\, ds = R, \tag{15}$$

where $z(t,s)$ given T, (denoted by $z(T)$), is a solution of the following state

[12] This method is based on the expansion of state and control variables in Lagrange series and on a spectral collocation method for approximating state equations. In this way, the optimal control problem can be reduced to a finite-dimensional nonlinear programming problem in expansion coefficients. This problem can be solved by using standard nonlinear programming methods. We used NAG library subroutine E04UCF, which is based on sequential quadratic programming method, to solve the nonlinear programming problem (see Tarkiainen–Tuomala, 1998, for a more detailed exposition of this method).

problem (S):

$$z(T) = \underset{z \geq 0}{\operatorname{argmax}}\, g(z - T(z)) + s\, h(z/t) \text{ for all } (t, s) \in \Omega.^{13} \tag{16}$$

In order to solve the problem (14)–(15) we need the values of state mapping $T \to z(T)$ and thus at each evaluation of z we have to solve the state system (16). The state constraint $z \geq 0$ causes the mapping $T \to z(T)$ to be nonsmooth.

In order to solve the problem (14)–(16) in practice some numerical method have to be used. The idea is to expand the tax function in terms of second order B-spline functions. In this way, the problem is replaced by a nonsmooth parameter optimization problem, that can be solved e.g. by a Broximal Bundle method. We shall describe briefly a numerical method for solving the optimal income tax problem (P) in the appendix.

4 Numerical results

This section reports numerical results in a Cobb-Douglas case. It means that computations were carried out for the Cobb-Douglas utility function

$$u = ln\, x + s\, ln(1 - y) \tag{17}$$

the time endowment being normalized at unity.

The elasticity of substitution between leisure and consumption is 1. Social welfare is taken to be the simple sum of utilities cardinalised as in (17). The revenue requirement, R, is set at zero so that the sole purpose of taxation is redistributive. It is natural to make comparisons between the one-dimensional case and the two-dimensional case. The solutions obtained are displayed in Figures 1, 2 and 3. In Figure 1 we assumed that the characteristics are uniformly distributed. The domain $\Omega = [0.03,1.0] \times [0.03,1.0]$. In Figures 2 and 3 we in turn assumed that the characteristics (t, s) are distributed in the population according to a bivariate lognormal distribution with parameters (μ, σ, ρ) representing the means, variances and correlation of the characteristic parameters (t, s) in the population. The distributional coefficient are therefore the five parameters $(\mu_t, \mu_s, \sigma_t, \sigma_s, \rho)$. In Figures 2 and 3 we have the following cases $(\mu_t = \mu_s = -1, \sigma_t = \sigma_s = 0.5, \rho = -0.3$ and $\rho = 0.3)$ and $(\mu_t = -1, \mu_s = -1.2, \sigma_t = 0.5, \sigma_s = 0.3, \rho = -0.3)$. The

[13] In other words the incentive compatibility constraint is now written as the selected income supply, $z(t, s)$, maximizes utility, given the tax function.

domain $\Omega = [0.05,1.4]\times[0.05,1.4]$. In the one-dimensional case $\mu_t = -1$ and $\sigma_t = 0.5$.

There are several features of numerical results that we would like to emphasize. First, in the two-dimensional cases the marginal tax rates are not monotone functions of gross income. Secondly, in terms of levels of marginal tax rate there is a substantial difference between the one- and two-dimensional cases. We found that optimal marginal tax rates are higher for almost all income levels in the two-dimensional case compared to those obtained from one-dimensional model. In fact we have two possibilities in the case of one-dimensional population.

Namely if people have identical preferences but differ in abilities, we are back in the Mirrlees model. The opposite case of the Mirrlees model is that one in which diversity of preferences is the sole source of inequality. Thus we consider an economy in which everyone has exactly the same level of productivity. In Figure 1 we also show the marginal tax rate in this case. We assume that everyone has a wage rate of 1 i.e. t=1. They turn out to be nondecreasing over the majority of the population. Thus there is redistributive taxation in this case. Some people would, however, say that if individuals have the same opportunities, then while their choices may differ, there is no ethical basis for redistributive taxation.

Our results may be obvious to some, and surprising to others. At least they are surprising to those who believe that taking into account work preferences in the population we have an argument for a less redistributive tax/benefit system. Given our specifications numerical results suggest that this is not necessarily so. Note also that a higher correlation between t and s lowers marginal tax rates uniformly. Or put it another way marginal tax rates are higher when there is positive (negative) correlation between work(leisure)preferences and productivity. From Figures 2 and 3 we can also see that the marginal tax rates are increasing around average z.

What might be an intuitive explanation of these results?[14] We know from the numerical results in Kanbur–Tuomala (1994) that a greater inherent in equality leads to the higher marginal tax rates at each income level in the one dimensional income tax model. We might reasonable expect in the two-dimensional world that there is a greater dispersion of income. At least it seems to us that this is so when there is a positive correlation between work preferences and productivity. In conjunction with the maximization of the concave social welfare in the economy with a greater dispersion of income

[14]It is difficult to find intuition because several factors are going on at the same time.

will imply to higher marginal tax rates.[15]

Table 1 also shows that the levels of z and x are considerably higher with heterogenous work preferences than with identical preferences. In other words the economy with heterogenous work preferences seems to be richer than that of identical work preferences. How to explain this? For that it may useful to consider how things are in a no-tax economy in which diversity of work preferences is the sole source of inequality. Thus people face equal opportunities in the labour market. To simplify further this economy consists of two types of individuals who differ in their work preferences s_1 and s_2, where $s_2 > s_1$. Given the utility function (17) in a no tax economy the demand for leisure is $(1 - y) = s/(1 + s)$. If we now increase dispersion of s in this economy, the average leisure decreases. Or put it another way the average labour supply increases and consequently the average gross income increases so that the economy becomes richer.

One has to be careful in making comparisons between the one- and two dimensional problem. It is simply so that we have different populations in different cases. Therefore we should not take literally our comparisons in the Table 1 and Figures 2, 3. They just show the solutions in the two dimensional and the case, when $t = 1$.

Moreover, one interesting thing of numerical results is that there were quite little “bunching”. In the circumstances under which ”bunching” occurs each individual faces the same pre-tax income and consumption. This means that $\partial z(t, s)/\partial t = \partial z(t, s)/\partial s$=0 in a subdomain of Ω. In two-dimensional problems one might expect more “bunching” simply because some taxpayers with different work preferences will end up to supply the same amount of labour. In the case of Figure 1 (uniform distribution) the amount of “bunching” was about 2 percent. For example in Figure 2 (ρ = -0.3) the amount of “bunching” was about 3 percent. In the one-dimensional case the amount of “bunching” turns out to be practically speaking zero.

[15] The referee has encouraged us to take this view.

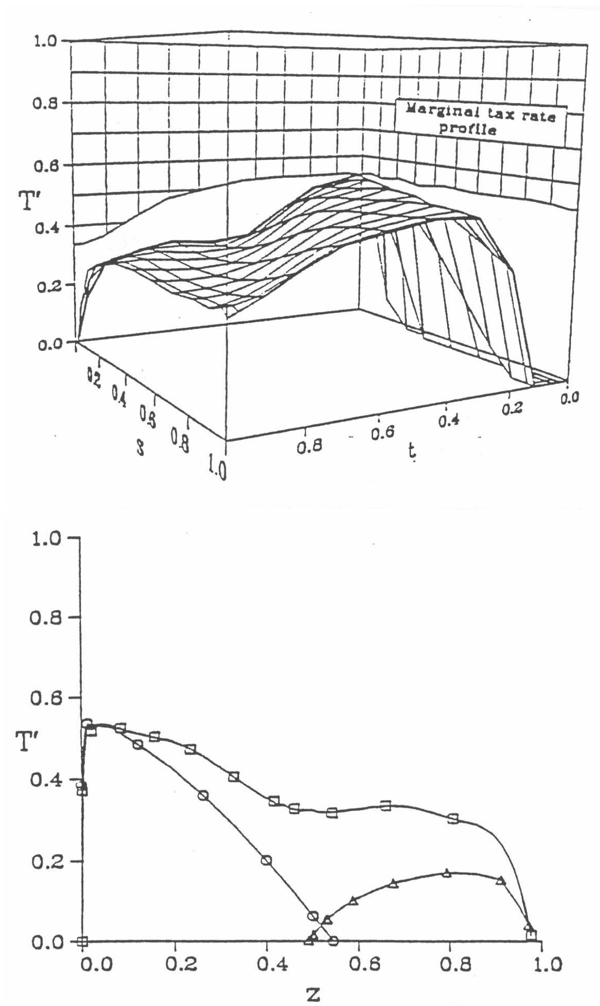

Figure 1: Optimal marginal income tax rates and tax schedule

$(t, s)(\square); t(\circ)$ uniformly distributed (s=1) and s(Δ) uniformly distributed (t=1).

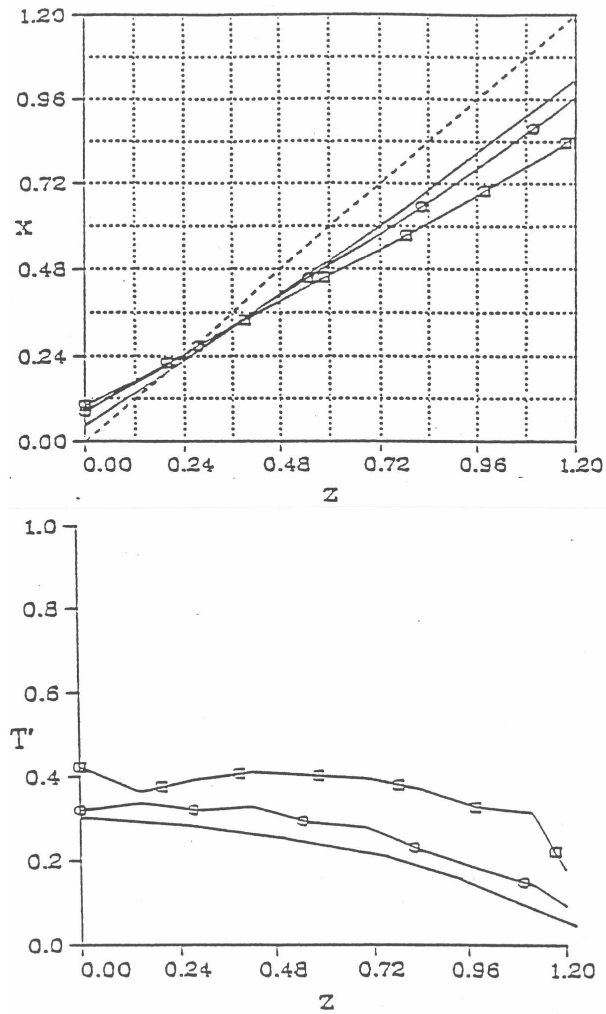

Figure 2: Optimal marginal income tax rates and tax schedules

(t, s) and t lognormally distributed
□ (μ_t=μ_s=-1, σ_t=σ_s=0.5, ρ=-0.3)
○ (μ_t=μ_s=-1, σ_t=σ_s=0.5, ρ= 0.3)
– (μ_t=-1, σ_t=0.5)

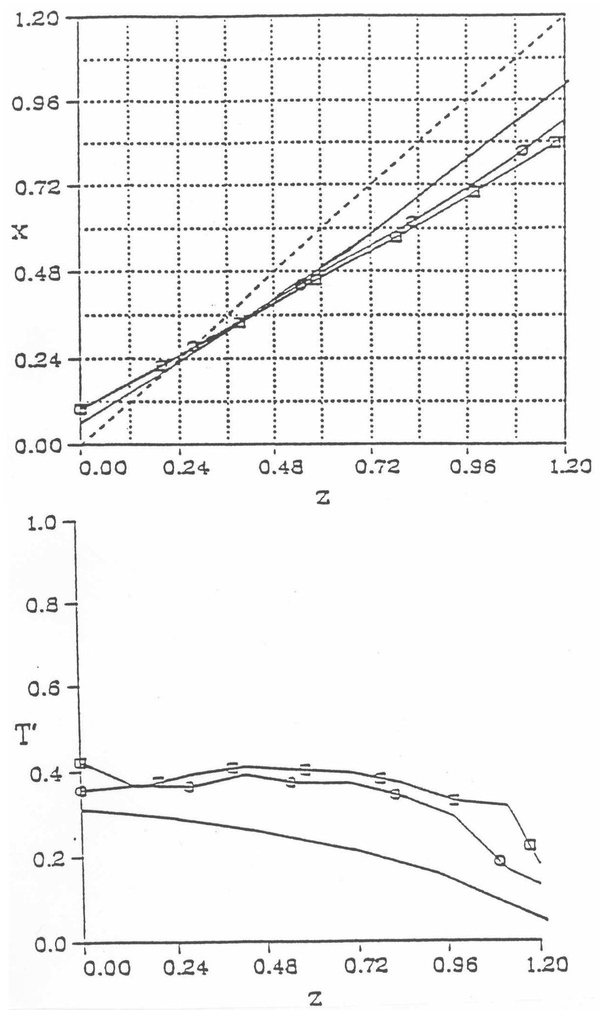

Figure 3: Optimal marginal income tax rates and tax schedules.

(t, s) and t lognormally distributed
□ (μ_t=μ_s=-1, σ_t=σ_s=0.5, ρ=-0.3)
○ (μ_t=μ_s=-1.2, σ_t=0.5, σ_s=0.3, ρ=-0.3)
– (μ_t=-1, σ_t=0.5)

Table 1

Case 1 (Figure 1) ((t, s) uniformly distributed)	
Average z and x	0.29
Marginal tax rate at average z	39 %
The maximun value of social welfare	-1.53
Case 2 (Figure 2) (μ_t=μ_s=-1, σ_t= σ_s=0.5, ρ=-0.3 and 0.3)	
Average z and x	0.26
Marginal tax rate at average z	32 % (ρ=0.3), 38 % (ρ=-0.3)
Average tax rate at z=1.2	20 % (ρ=0.3), 30 % (ρ=-1.3)
The maximum value of social welfare	-1.74 (ρ=0.3), -1.71 (ρ=-0.3)
Case 3 (Figure 3) (μ_t=-1, μ_s=-1.2, σ_t=0.5, σ_s=0.3, ρ=-0.3)	
Average z and x	0.28
Marginal tax rate at average z	38 % and 37 %
Average tax rate at z=1.2	26 %
The maximum value of social welfare	-1.62
One-dimensional population (cases 2 and 3)	
Average z and x	0.18
Marginal tax rate at average z	28 %
Average tax rate at z=1.2	16 %

5 Conclusions

This paper has examined the utilitarian case for a redistributive nonlinear income tax under the assumption that individual differ in their work preferences and abilities. A numerical method for solving this problem has been developed. The problem is difficult to solve, but we managed to solve some examples. The numerical results obtained give us a direction to influences that richer picture of population would have on the optimal income tax schedule. On the basis of our numerical results we conclude that the tax system is more redistributive compared to those obtained from the one-dimensional case. This may be surprising to those who believe that taking into account different work preferences is an argument for having less redistribution and hence lower levels of income taxation and social security payments.

Appendix: Computational procedure

Instead of trying to approximate the tax function $T(z)$ over the entire interval by one polynomial of high degree, we partition the interval into smaller intervals and approximate the tax function on each of these smaller intervals by a low degree polynomial. We approximate the tax function T using a quadratic (second order) B-spline basis.[16] B-splines have small support, i.e., they are zero on a large set. For computational convenience it is desirable to have basis functions with this property. Let $[z_0, z_n]$ be a finite interval, and let $z_0 < z_1 < \cdots < z_{n-1} < z_n$ be the partition points (nodes) for this interval. Moreover, we introduce additional nodes z_{-2}, z_{-1}, z_{n+1} and z_{n+2} such that $z_{-2} < z_{-1} < z_0$ and $z_n < z_{n+1} < z_{n+2}$. Now we write the tax function T as a sum of basis functions with real parameters $\{k^i, i = 1, ..., N\}$ as

$$T^h(z) = \sum_{i=1}^{N} k^i \Phi^i(z) \text{ for all } z \in [z_0, z_n] \tag{18}$$

where Φ^i is a second order B-spline function with support $[z_{i-3}, z_i]$ and $N = n+2$. Second order B-spline basis implies C^1 continuity for the tax function. Now the tax function T^h is uniquely determined by N-dimensional parameter vector $k = [k^1, k^2, \cdots, k^N]'$. Thus we denote $z(k)$ as the solution of the state problem (S) corresponding to parameter vector k.

The discretized problem corresponding to (P) is (P^h):
Find $k^* \in \mathbf{R}^N$ such that

$$W(k^*) = \max_k \, \{W(k) = \iint_\Omega [g(z(k) - T^h(z(k))) \, + \, s\, h(z(k)/t)]\, f(t,s)\, dt\, ds\} \tag{19}$$

subject to

$$G(k) = \iint_\Omega T^h(z(k))\, f(t,s)\, dt\, ds = R. \tag{20}$$

The discretized problem (P^h) can be regarded as a constrained nonsmooth optimization problem and hence the problem can be solved e.g. by Broximal Bundle method. We apply the code MPBNGC from a sub-routine library NSOLIB to solve the nonsmooth optimization problem (19)–(20) (see Mäkelä, 1990 and 1993).

When solving the nonsmooth optimization problem (P^h) one needs the (sub)gradient and the function values $W(k)$ and $G(k)$ at each iteration point. The numerical procedure for computation of the values of the social welfare

[16] Mortenson (1985) provides a detailed exposition on the B-splines.

function, constraints and the subgradients is first to compute values $z(k)$ at a set of quadrature points $(t_i, s_j), i, j = 1, \cdots, M$. At each evaluation of $z(t_i, s_j; k)$ we have to solve the state problem (S). This problem is solved by NAG-library subroutine E04KBF. The values of the social welfare function, constraints and their subgradients can be evaluated by the quadrature. The quadratures are computed by the Romberg integration. The interval $[z_0, z_n]$ was partitioned into $n = 9$ and $n = 10$ equal subintervals. Thus the number of unknowns was $N = 11$ or $N = 12$. If the number of unknowns was increased over that the value of the objective function didn't increase significantly. The number of discretization points for t and s variables is 33 $(=M)$.

References:

Akerlof, G., (1978), 'The economics of 'tagging' as applied to the optimal income tax, welfare programs and manpower planning', American Economic Review, 68, 8-19.

Armstrong, M., (1996), 'Multiproduct nonlinear pricing', Econometrica, 64, 51-75.

Cesari, L., (1969), 'Optimization with partial differential equations in Dieudonne-Rashevsky form and conjugate problems', Arch. Rational Mech. Anal., 33, 339-357.

Ebert, U., (1988), 'Optimal income tax problem: On the case of two-dimensional population', Discussion paper A-169, Department of Economics, University of Bonn.

Immonen, R., R. Kanbur, M. Keen and M. Tuomala, (1998), 'Tagging and taxing: The optimal use of categorical and income information in designing tax/transfer schemes', Economica, 65, 179-192.

Kanbur, R., and M. Tuomala, (1994), 'Inherent inequality and the optimal graduation of marginal tax rates', Scandinavian Journal of Economics, vol. 96, 275-82.

Kazemi-Dehkordi, M. A., (1984), 'Optimal control of systems governed by partial differential equations with integral inequality constraints', Nonlinear Analysis, 8, 1409-1425.

Laffont, J.-J., E. Maskin, and J.-C. Rochet, (1987), 'Optimal nonlinear pricing with two-dimensional characteristics'. In information, incentives, and economic mechanisms, eds. T. Groves, R. Radner, and S. Reiter, University of Minnesota Press.

McAfee, P., and J. McMillan, (1988), 'Multidimensional incentive compatability and mechanism design', Journal of Economic Theory, 46, 335-354.

Mirrlees, J. A., (1971), 'An exploration in the theory of optimal taxation', Review of Economic Studies, 38, 175-208.

Mirrlees, J. A., (1976), 'Optimal tax theory, a synthesis', Journal of Public Economics, 6, 327-358.

Mirrlees, J. A., (1986), 'The theory of optimal taxation', Chapter 24 in Arrow and Intriligator (eds.), Handbook of Mathematical Economics, Vol III, 1197-1249, Amsterdam: Elsevier Science Publishers.

Mortenson, M. E., (1985), Geometric modelling, John Wiley&Sons, New York.

Mäkelä, M., (1990), 'Nonsmooth Optimization. Theory and Algorithms with Application to Optimal Control', Doctoral Thesis, Report 47/1990, University of Jyväskylä, Department of Mathematics.

Mäkelä, M., (1993), 'Issues of implementing a Fortran subroutine package NSOLIB for nonsmooth optimization', University of Jyväskylä, Department of Mathematics, Laboratory of Scientific Computing, Report 5/1993.

Rochet, J.-C., (1985), 'The taxation principle and multi-time Hamilton-Jacobi equations', Journal of Mathematical Economics, 14, 113-128.

Rochet, J.-C. and P. Choné, (1998), 'Ironing, sweeping and multidimensional screening', Econometrica, 66, 783-826.

Sandmo, A., (1993), 'Optimal redistribution when tastes differ', Finanz Archiv, 50, 149-163.

Sen, A., (1973), On economic inequality, Clarendon Press, Oxford.

Sen, A, (1992), Inequality reexamined, Clarendon Press, Oxford.

Tarkiainen, R. and M. Tuomala, (1998), 'Optimal nonlinear income taxation with two-dimensional population; a computational approach', Computational Economics, 13, 1-16.

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

Wilson, R., (1993), Nonlinear pricing, New York: Oxford University Press.

Wilson, R., (1995), 'Nonlinear pricing and mechanism design', in Amman, Kendrick and Rust (eds.) Handbook of Computational Economics, Volume 1, Amsterdam, Elsevier Science Publishers.