Tax-Benefit Revealed Social Preferences: Are Tax Authorities Non-Paretian?

François Bourguignon  
DELTA (Joint Research Unit CNRS-ENS-EHESS)  
48 Bd Jourdan 75014 Paris  
email: fb@delta.ens.fr  
and World Bank

Amedeo Spadaro  
DELTA (Joint Research Unit CNRS-ENS-EHESS)  
and  
GREP - Department of Economics, Universitat de les Illes Balears,  
Ctra Valldemossa Km 7,5 07071 Palma de Mallorca, Spain  
email: amedeo.spadaro@uib.es

Abstract

This paper inverts the usual logic of applied optimal income taxation. Standard practice investigates the shape of the optimal tax schedule that is consistent with a given social welfare function, a statistical distribution of individual productivities that fits available data on labour incomes and given preferences between consumption and leisure. The argument in this paper goes in the opposite direction. It starts from the observed distribution of gross and disposable income within a population and from observed marginal tax rates. Under a set of simplifying assumptions, it is then possible to recover the social welfare function that would make the observed marginal tax rate schedule optimal. This provides an alternative way of reading marginal tax rates calculations routinely provided by tax-benefit models. In this framework, the issue of the optimality of an existing tax-benefit system may be analysed by considering whether the social welfare function associated with that system satisfies elementary properties. A detailed application is given in the case of France, Spain and UK.

This version  
September 2002

---

1 This is a revised version of a previous paper “Social Preferences Revealed through Effective Marginal Tax Rates” initially prepared for the Workshop on “Fighting Poverty and Inequality through Tax Benefit Reform: Empirical Approaches”, Universitat Autonoma de Barcelona, November 2000. The authors are indebted to the EUROMOD team. They also thank participants to seminars in DELTA, EEA and ESEM 2002 Venezia conference, IEF Madrid for useful comments. But they remain solely responsible for any remaining error. We acknowledge financial support of EC-DGXII (Targeted Socio-Economic Research Program contract nº ERBFMBICT960879), Spanish Government- MEC, (Programa de cooperación Franco-Española PFE2001-0010), Spanish Government- MCYT (Programa nacional de promoción general del conocimiento BEC2000-0415 and SEC2002-02606).
Introduction

Several attempts were recently made at analysing existing redistribution systems in several countries within the framework of optimal taxation theory. The basic question asked in that literature is whether it is possible to justify the most salient features of existing systems by some optimal tax argument. For instance, under what condition would it be optimal for the marginal tax rate curve to be U-shaped - see Diamond (1998) and Saez (2001) for the US and Salanié (1998) for France? Or could it be optimal to have 100 per cent effective marginal tax rates at the bottom of the distribution as in some minimum income programs - see d'Autume (2000), Choné and Laroque (2001) Bourguignon and Spadaro (2000), Piketty (1997) in the case of France and other European countries. Such questions were already addressed in the early optimal taxation literature and in particular in Mirrlees (1971) but the exercise may now seem more relevant because of the possibility of relying on large and well documented micro data sets rather than on hypothetical statistical distributions.

The results obtained when applying the standard optimal taxation framework to actual data depend very much on several key ingredients of the model. The shape of the social welfare function may be the most important one. As already pointed out by Atkinson and Stiglitz (1980) in their comments of Mirrlees' original work, using a Rawlsian social objective or a utilitarian framework on a hypothetical but distribution of abilities make a big difference. The first would lead to very high effective marginal rates for low individual abilities, whereas the second would be closer to a linear tax system with a constant marginal tax rate. As the sensitivity with estimated distribution of abilities is likely to be the same as with hypothetical but reasonably realistic ones, what should be done? Should one use a Rawlsian objective and conclude that part of actual redistribution systems are sub-optimal, or should one use a less extreme social welfare function and conclude that another part of the redistribution schedule in non-optimal? The point of view taken in this paper is the opposite of the previous one. Instead of taking the social welfare function as given and deriving the optimal schedule of effective marginal tax rates along the income or ability scale, it is the reverse that is being done. The focus is here on the social welfare function that makes optimal the effective marginal tax rates schedule that corresponds to the redistribution system actually in place. This approach is the dual of the previous one. In the first case, wondering about the optimality of an actual redistribution system consists of comparing an optimal effective marginal tax rate schedule derived from some 'reasonable' social welfare function with the actual one. In the second case, it consists of checking whether the social welfare function implied by the actual redistribution
schedule is in some sense 'reasonable', that is whether the marginal social welfare is everywhere positive and decreasing along the horizontal axis.

In effect, the approach that is proposed here is simply a way of 'reading' the redistribution schedule, or that is the average and marginal net tax curves that are commonly used to describe redistribution. This reading simply translates the observed shape of these curves into social welfare language. Comparing two redistribution systems or analysing the reform of an existing system can thus be made directly in terms of social welfare. Instead of determining who is getting more out of redistribution and who is getting less, this reading of the marginal tax rate schedule informs directly on the differential implicit marginal social welfare weight given to one part of the distribution versus another.

Of course, this revelation of social preferences relies on several auxiliary assumptions about labour supply behaviour and about the distribution of individual abilities. It is well known that the optimal tax schedule depends crucially on these assumptions. The same is true of the social preferences revealed by a given marginal tax schedule. It is even conceivable that apparent anomalies in these preferences may be due to these assumptions being unsatisfactory. The observation of the effective average or marginal tax rate schedule may thus reveal more than social preferences. In some cases, it may suggest either that the tax schedule is inconsistent with optimality. But in others it may also reveal that some common assumptions on labour supply behaviour or on the distribution of abilities are inconsistent. This seems equally useful information.

The paper is organized as follows. Section 1 recalls the optimal taxation model and derives the key duality relationship between the effective marginal tax rate schedule and the marginal social welfare function in the simple case where individual preferences between consumption and leisure are assumed to be quasi-linear. The second section discusses the empirical implementation of the preceding principles. The third section applies them to three EU countries - France, Spain and the UK- taking advantage of preliminary results obtained with the EUROMOD model. In each case, the social welfare function is characterized under a set of simple alternative assumptions about the labour supply elasticity and some characteristics of the distribution of individual abilities. The fourth section deals with a more general specification of individual preferences allowing for the introduction of income effects. The main lessons drawn from the whole exercise are gathered in the concluding section.

---

2 See Immervoll, O'Donoghue and Sutherland (2000).
1. The duality between optimal marginal tax rates and the social welfare function

The basic optimal taxation framework is well known. Agents are assumed to choose the consumption \( y \) /labour \( L \) combination that maximizes their preferences, \( U(y, L) \), given the budget constraint imposed by the government: \( y = wL - T(wL) \), where \( w \) is the productivity of the agent and \( T(\cdot) \) the net tax schedule. If the distribution of agents' productivity is represented by the density function \( f(w) \) defined on the support \([w_0, A] \), the optimal taxation problem may be written as:

\[
\text{Max}_{T(w)} \int_{w_0}^{A} G[V[w, T(w)]] f(w) \, dw
\]

\[
\text{s.t.} \quad (y^*, L^*) = \text{Argmax} \{ U(y, L); y = wL - T(wL), L \geq 0 \}
\]

\[
V[w, T(w)] = U(y^*, L^*)
\]

\[
\int_{w_0}^{A} T(wL^*), f(w) \, dw \geq B
\]

where \( G[\cdot] \) is the social welfare function that transforms individual utility, \( V(\cdot) \), into social welfare and \( B \) is the budget constraint of the government.

The main argument in this paper is based on the special case where the function \( U(y, L) \) is quasi-linear with respect to \( y \) and iso-elastic with respect to \( L \):

\[
U(y, L) = y - k \frac{L^{1 + \frac{1}{\varepsilon}}}{1^{1 + \frac{1}{\varepsilon}}}
\]

In that case, the labour supply function that is solution of (1.2) above is given by:

\[
L^* = k \cdot w^{\varepsilon} \cdot [1 - T'(wL^*)]^\varepsilon
\]

where \( k \) is a constant. Given that expression, \( \varepsilon \), may be interpreted as the elasticity of labour supply, \( L^* \), with respect to the marginal return to the labour of the agent, the latter being his/her productivity corrected by the marginal rate of taxation \( T'(\cdot) \).

This particular case has been studied in some detail in the literature. It leads to the following simple characterization of the optimal tax schedule:

---

In that expression, \( t(w) \) is the (optimal) marginal tax rate faced by an agent with productivity, \( w \), and therefore with earnings \( wL^* \). \( F(w) \) is the cumulative distribution function associated with the density \( f(w) \), whereas \( S(w) \) stands for the average marginal social utility of all agents with productivity above \( w \), which is given by:

\[
S(w) = \frac{1}{1-F(w)} \int_w^A G[V(w, T())] f(w) dw
\]

The duality between the marginal rate of taxation and the social welfare function, which is exploited in the rest of this paper, lies in the two preceding relationships. It is thus important to have a good intuition of what they actually mean. Consider the following thought experiment. Starting from an existing tax system, the government decides to increase the tax payment by a small increment \( dT \) for people whose labour income is \( Y \) and labour productivity \( W \), leaving the rest of the tax schedule unchanged. Such a measure has three effects: a) it reduces the labour supply of people at \( Y \) and just below that level because the marginal return to their labour falls by \( dT \); b) it increases the tax payment of all people whose earnings is above \( Y \) by \( dT \); c) it increases total tax receipts by the difference between effects b and a. At an optimum, the total effect of these changes on social welfare must be equal to zero.

The tax reduction effect a) depends on the marginal rate of taxation, \( t(W) \), the elasticity of labour supply, \( \varepsilon \), productivity itself, \( W \), and the density of people around that level of productivity, \( W \). This tax reduction effect (\( TR \)) may be shown to be equal to: \(^5\)

\[
TR = \frac{t(W) W f(W)}{1-t(W)} \frac{dW}{1+\varepsilon}
\]

The tax increase effect (\( TI \)) is simply equal to the proportion of people above the productivity level \( W \) times the infra-marginal increase in their tax payment, \( dT \):

\(^5\) The change in the tax receipt is given by \( T'(Y).dY/dT g(Y) \), where \( g(Y) \) is the density of people at the gross labour income \( Y \). Given (3), it is easily shown that \( dY/dT = \varepsilon Y/(1-T'(Y)) \) and that \( g(Y) = f(W)W/[Y.(1+\varepsilon)] \). The expression of \( TR \) follows.
In order for the government’s budget constraint to keep holding, the resulting net increment in tax receipts, $TI – TR$, is to be redistributed. Since net effective marginal tax rates are not to be changed, except at $Y$, this requires redistributing a lump sum $TI – TR$ to all individuals in the population. The marginal gain in social welfare of doing so is given by $(TI – TR).S(w_0)$. The loss of social welfare comes from people above $W$ whose disposable income is reduced by $dT$. People whose marginal tax rate is actually modified – i.e. people at $W$ and just below – are not affected because they compensate the drop in the effective price of their labour and its negative effect on consumption by a reduction in the labour they supply and an increase in their leisure. This is the familiar envelope theorem. Under these conditions the loss of social welfare is simply equal to the proportion of people above $W$ times their average social marginal welfare, $S(W)$. The optimality condition may thus be written as:

$$[1 – F(W)].S(W).dT = (TI – TR).S(w_0)$$

and after dividing through by $S(w_0)$ and $dT$:

$$[1 – F(W)] \frac{S(W)}{S(w_0)} = \frac{TI – TR}{dT}$$

which, after rearranging, leads to (5) above.

What is attractive in that expression is that the right-hand side is essentially of a positive nature whereas the left-hand side is essentially normative. The right hand side measures the net tax gain by Euro confiscated from people at and above $W$. The left hand side measures the relative marginal social loss of doing so.

The preceding expression illustrates the duality that is used in the rest of this paper. For a given distribution of productivities, $f(w)$, the right-hand side may be easily evaluated by observing the tax-benefit system in a given economy and its implied effective marginal tax rate schedule, provided that some estimate of the labour supply elasticity is available. Then (6) yields information on the social welfare function that is consistent with the tax-benefit system. When read in the reverse direction, (6) shows the tax-benefit system that is optimal for a given social welfare function. The latter is the usual approach in the applied optimal taxation literature. The former approach that ‘reveals’ the social welfare function consistent with an
existing tax-system, under the assumption that this system is indeed optimal in the sense of model (1) is less conventional, although it has been known for some time in the dynamic optimisation literature as the “optimal inverse problem”\(^6\).

Characterizing precisely the social welfare function, \(G(w)\), optimally implied by a tax-benefit system requires some additional steps after (6). Define the right-hand side of (6) as the following function of the observed marginal tax rates and productivity distribution:

\[
\Theta(w) = \left[1 - F(w)\right] - \frac{t(w)}{1-t(w)} \cdot w f(w) \frac{1}{1+1/\varepsilon}
\]  

(7)

Normalize the welfare function \(G(w)\) so that the mean marginal social welfare is equal to unity, \(S(w_0) = 1\). Finally take derivatives on both sides of (5) and (6). It comes that the marginal social welfare, associated with the productivity level, \(w\), is given by:

\[
G'[V(w, T)] = -\Theta'(w) / f(w)
\]  

(8)

It may readily be seen in this expression that the revealed marginal social welfare function depends not only on the effective marginal tax rate and the density of the distribution of productivities, but more importantly on the derivatives of these two functions with respect to productivity. Taking the derivative of (7) leads to the following expression of the marginal social welfare:

\[
G'[V(w, T)] = 1 + \left(\frac{1}{1+1/\varepsilon}\right) \cdot \left(\frac{t(w)}{1-t(w)} \cdot \frac{V(w)}{1-t(w)} + \eta(w) + 1\right)
\]  

(9)

where \(\eta(w) = w f''(w) / f(w)\) is the elasticity of the density and \(V(w) = w t'(w) / t(w)\) that of the marginal tax rates with respect to individual productivity.

The preceding expression suggests two basic tests of the consistency of an observed tax-benefit system and an observed distribution of productivities within the optimal taxation framework. First, the social welfare function is generally assumed to be increasing so as to ensure consistency with the Pareto principle. Thus, expression (9) must be positive everywhere for the redistribution authority to be considered as Paretian. Second, the derivation of the

---

\(^6\) See, for instance, Kurz (1968).
optimal redistribution schedule (4) requires the social welfare function to be concave\(^7\). Taking the derivative of (9) leads to the following condition:

\[
\phi(g + wg') + \phi'(2 + wg) + w\phi'' \leq 0
\]  

(10)

where \( g(w) = f'(w) / f(w) \) and \( \phi(w) = t(w) / [1 - t(w)] \).

Note, however, that this condition relies on the second derivatives of both the density function and the marginal tax rate schedule. It is thus much stronger than the Pareto condition.

The rest of this paper focuses on the sign and the slope of the function \( G'(w) \) that may be associated with the tax-benefit systems observed in several countries and with some basic assumptions on \( \varepsilon \), and possibly on \( f(w) \) for extreme ranges of productivity, and on the preceding conditions. As far as the sign test is concerned, note that a more robust test that does not rely on the derivatives of the density function and marginal tax rate schedule may be obtained. \( S(w) \) in equation (5) is the average of marginal social welfare for people above \( w \). The sign condition on \( G''(w) \) implies that this function must be non-negative and non-increasing. It is easily shown that:

\[
S(w) = 1 - \frac{t(w)}{1 - t(w)} \frac{w f'(w) \varepsilon}{1 - F(w) 1 + \varepsilon}
\]  

(11)

The sign test provides a test of the Pareto consistency of the redistribution authority that depends only on marginal tax rate and the density function, rather than their derivatives. If \( S(w) \) as given by the preceding expression is negative, then it may be concluded that the redistribution authority is non-Paretian.

Two final remarks are necessary at this stage. The first one is to stress that the optimal tax conditions (4) do not require that social preferences be Pareian. They are necessary and sufficient for all concave social welfare functions and not only monotonically increasing one. This may be seen by rewriting the original optimisation problem (1) as an optimal control problem as originally done by Mirrlees (1971) – see also (Atkinson and Stiglitz, 1980, p. 415). In this optimal control problem, the utility, \( u \), of an individual with productivity \( w \) is the state variable and the objective function simply writes:

\[
\text{Max } \int_{w_0}^A G[u] f(w) dw
\]
whereas the motion equation is defined by a condition of the type \( \dot{u} = du/\dot{w} = g(Y,u) \), where disposable income, \( Y \), is the control variable and \( g( ) \) may be derived from first order optimization conditions of agents. It is well-known that the maximum principle that leads to the optimality conditions (4) is necessary and sufficient provided that the two functions \( G( ) \) and \( g( ) \) are concave. This requirement is consistent with \( G( ) \) being inverted-U shaped.

A second remark has to do with the well known results of the optimal income tax theory that the optimal marginal tax rate on the most productive agent must be zero when the support of \( f(w) \) is finite (Seade 1977, 1982). As it is not the case in actual tax-benefit systems, this would seem sufficient to rule out that the tax-authority is pursuing the maximization of some well behaved social welfare function. Thus, no well-behaved social welfare function should be obtained from the preceding inversion procedure. But of course, zero marginal taxation at the top does not necessary hold if the support of \( f(w) \) is assumed to be infinite. This latter assumption is maintained in the rest of this paper. 8

2. Basic principles for empirical implementation

The previous methodology requires estimates of the elasticity of labour supply, \( \varepsilon \), the distribution \( f(w) \) and the marginal rate of taxation, \( t(w) \), to be available. Practically, what is observed in a typical household survey? Essentially total labour income, \( wL \), and disposable income, \( y \), or equivalently, total taxes and benefits, \( T(wL) \). When the household survey is connected with a full tax-benefit model, it is possible to compute the latter on the basis of the observed characteristics of the household and the official rule for the calculation of taxes and benefits. With such a model, it is also possible to evaluate the effective marginal tax rate by simulating the effects of changing observed labour income by a small amount. To be in the situation to invert the marginal tax rates into marginal social welfare, it is thus necessary to impute a value of the productivity parameter, \( w \), to the households being observed and to estimate the statistical distribution of individual productivities, \( f(w) \).

The simplest way to proceed would consist of assimilating productivity with observed hourly wage rates, and then estimating econometrically the labour-supply elasticity, \( \varepsilon \), which, without loss of generality, might even be specified as a function of productivity, \( w \). This approach was not followed for several fundamental reasons. First, labour supply may differ quite significantly from working hours when unobserved efforts are taken into account. Second,

---

8 On this, see also the argument in Diamond (1998).

9 To keep with the logic of the optimal taxation model, non-labour taxable income are ignored in all what follows.
the econometric estimation of a labour supply model requires taking into account the non-linearity introduced by the tax-benefit system actually faced by individuals, and in particular the endogeneity of marginal tax rates. Econometric estimations of this type are now known to be little robust. But relying on simpler alternative estimates based on simple linear specifications is like introducing some arbitrariness in the analysis. Third, econometric estimates of the elasticity of labour supply are known to differ substantially across various types of individuals. In particular, it is small for household heads and larger for spouses, young people and people close to retirement age. Under these conditions, what value should be chosen for \( \varepsilon \)? Fourth, and more fundamentally, it seems natural to choose the household as the economic unit in a welfare analysis of taxes and benefits. But, then, the problem arises of aggregating at the household level concepts or measures that are valid essentially at the individual level. In particular, how should individual productivities be aggregated so as to define 'household productivity'? Likewise, if the elasticity of labour supply has been estimated at the individual level and is different across various types of individuals, how should it be averaged within the household?

The approach followed in this paper is opposite to the previous one. Instead of starting from observed productivity and deriving an estimate of labour supply from observed labour incomes, the elasticity of labour supply is taken to be arbitrarily given. Then, it is used to derive the implicit productivity of households from observed labour incomes. The latter operation is a simple inversion of the labour supply equation (3). Multiply both sides of that equation by \( w \) so that the gross labour income, \( Y \), appears on the left hand side:

\[
Y = wL^* = k w^{1+\varepsilon} [1 - T'(wL^*)]^\varepsilon
\]

After inversion, one gets for a given value of \( \varepsilon \):

\[
1 = \frac{1}{w} k Y^{1+\varepsilon} [1 - T'(Y)]^{-\frac{\varepsilon}{1+\varepsilon}}
\]

Thus, implicit productivity turns out to be an iso-elastic function of observed gross labour income corrected by a term that depends positively on the marginal tax rate. This correction is easily understood. For a given gross labour income, the higher the marginal tax rate, the lower is the labour supply as given by (3), and therefore the higher the implicit productivity.

The preceding inversion equation allows for a satisfactory definition of all functions necessary for recovering the social welfare function from the optimal taxation formula. For household, \( i \), observed with gross labour income, \( Y_i \), and marginal tax rate, \( t_i \), a value of the implicit productivity characteristic, \( w_i \), may be imputed through (13). Then all households may

---

10 See Blundell, Duncan and Meghir (1998).
be ranked by increasing value of that productivity. It is then possible to identify the distribution function $F(w)$, the marginal tax rate function, $t(w)$ and all the derivatives from which the social marginal welfare function may be inferred - see (7), (8) and (9) above.

Equation (13) yields the basic principle of the methodology. Its actual implementation raises additional complications, however. They are listed below together with the choices made to overcome them.

\[ \text{a) Continuity and differentiability} \]

The application of the inverted optimal taxation formula, (7)-(8)-(9), requires the knowledge of the continuous functions $f(w)$, $t(w)$ and their derivatives. As just discussed, above, however, what may be obtained from households data bases is a set of discrete observations of the imputed productivity characteristic, $w_i$, the associated cumulative distribution function, $F(w_i)$ and the marginal tax rate function, $t(w_i)$. The following operations permit to get an estimate of the derivatives of the function $f(w)$ and $t(w)$ and therefore of the marginal social welfare function.

(i) For any arbitrary value of productivity, $W$, obtain an estimate of the density function $f(W)$ and the effective marginal tax rate $t(W)$ by kernel techniques defined over the whole sample of observations - using a Gaussian kernel with an adaptive window\(^{11}\). These Kernel approximations are made necessary first by the need to switch from a discrete to a continuous representation of the distribution and the tax schedule and second by the heterogeneity of the population with respect to some characteristics that may influence marginal tax rates and productivity estimates - household composition, for instance\(^ {12}\).

(ii) Estimate the derivatives of $t(w)$ and $f(w)$ using again a kernel approximation computed over the whole sample.\(^ {13}\)

(iii) Compute the elasticity of $t(w)$ and $f(w)$ (i.e. the terms $\eta(w) = wf''(w)/f(w)$ and $\nu(w) = wt'(w)/t(w)$).

(iv) Compute the function $G'(w)$ as in (9) and the function $S(w)$ as in (11).

\[ \text{b) Household size} \]

It was assumed in the preceding section that all households had identical preferences and indirect utility functions. Practically, actual tax-benefit systems discriminate households

\(^{11}\) This choice was justified by the lack of observations and the increasing distance between them in the upper tail of the distribution. For technical details, see Hardle (1990).

\(^{12}\) Occupational status and home ownership are other sources of heterogeneity with respect to the tax system.

\(^{13}\) The function $\Theta(\cdot)$ itself may be approximated by Kernel techniques and then differentiated numerically. For technical details about the computations of kernel derivatives see Pagan and Ullah (1999, pag. 164).
according to various characteristics. Size and household composition are the main dimensions along which this discrimination is taking place. The issue thus arises of the way in which these characteristics can be implicitly or explicitly incorporated in the imputation of the social welfare function.

The results shown in the next sections are based on two extreme views. In the first one, the size of households is simply ignored in both the imputation of productivity and in tax optimisation. The implications of this choice are somewhat ambiguous. It may be seen in (13) that size affects productivity through two channels. On the one hand, a larger family - in terms of the number of potentially active adults - will generally have a greater gross labour income, which will contribute to a larger estimate of productivity. On the other hand, it will also face a different marginal tax rate. If the marginal tax rate is a decreasing function of household size for a given household income, as in most tax-benefit systems, then the preceding bias in the estimation of productivity will be attenuated.

The other extreme assumption consists of considering groups of households with the same size or the same composition as populations, which the redistribution authority seeks to maximize social welfare independently of each other. In other words, the optimal taxation problem involves finding an optimal tax-benefit schedule separately for each household group. This is implicitly done under some exogenous budget constraint, which makes the aggregate redistribution of income across the various groups of households exogenous. Thus, the analysis does not seek to recover any information about the social preferences of the redistribution authority across the various types of households.

Other choices would probably be possible, which would permit inferring social preferences about household composition. However two issues are intertwined here. One has to do with the autonomous labour supply behaviour of households with different composition - i.e. how the preferences \( U(y, L) \) should depend on size. The other is concerned with the weight of households with different composition in social preferences. Clearly, some arbitrary assumption is necessary on the former to be able to recover some information on the latter. It seemed simpler to limit the present analysis to the two extreme cases described above. \(^{14}\)

\[ c) \textit{Households with zero income and households with apparently irrational behaviour} \]

\(^{14}\) The only attempt made in that direction consisted of defining the preferences \( U(y, L) \) on a per (adult) capita basis. This had for consequence that the optimal tax-benefit system had to satisfy a homogeneity property close to the ‘quotient familial’ principle in the French income tax. According to that property, taxes and benefits per capita should depend only on gross income per capita. As this property is seldom satisfied by the whole tax-benefit system neither in France nor in other EU countries considered here, this route was abandoned. .
In presence of a guaranteed minimum income in a tax-benefit system, some households may find it optimal not to work at all. In the simple labor supply model above, this would correspond to a situation where the marginal tax rate is 100 percent. However, there is some ambiguity about these situations. Practically, some households are observed in parts of their budget constraint where the marginal tax rate is indeed 100 percent. There are two possible reasons for this. First, transitory situations may be observed where households have not yet converged towards their preferred consumption-labor combination. Second, transition periods are allowed by tax-benefit systems where beneficiaries of minimum income schemes may cumulate that transfer and labor income for some time so as to smoothen out the income path on return to activity.

The example of the French minimum income program (RMI) suggests the following way of handling the 100 marginal tax rate issue. People receiving the minimum income RMI and taking up a job lose only 50 percent of additional labor income during a so-called 'intéressement' period – 18 months. At the end of that period, however, they would lose all of it if they wanted to keep benefiting from the RMI. Discounting over time, this means that the actual marginal tax rate on the labor income of a 'RMIste' is between 50 and 100 percent. Taking the middle of that interval, the budget constraint of that person thus writes:

\[ y = \text{RMI} + 0.25 \times wL \]

if this person qualifies for the RMI – i.e. \( wL < \text{RMI} \). But it is simply:

\[ y = wL \quad \text{if} \quad wL > \text{RMI} \]

This budget constraint is clearly convex. Therefore, there should be a range of labour incomes around the RMI where it would be irrational to be. But, of course, some households are actually observed in that range, which is inconsistent with the model being used and/or the assumption made on the marginal tax rate associated with the RMI. One way of dealing with this inconsistency is to assume that all gross labor incomes are observed with some measurement error drawn from some arbitrary distribution. The measurement error is such that,

15 The Income Support mechanism in UK is a real 100% effective marginal tax rate. For each pound earned by working the Income Support is reduced by an equal amount.
16 All other benefits that may complement the RMI are ignored in this argument, but they are taken into account in the calculations made below.
17 This interval may easily be computed using the preference function of households and the budget constraint described by the preceding conditional system. Note that it depends on the size and the socio-demographics characteristics of each household.
without it, households would be rational and supply a quantity of labor outside the preceding range. This treatment of the data is analogous to the original econometric model describing the labor supply behavior of households facing a non-linear and possibly discontinuous budget constraint by Hausman (1985).

This methodology has been also applied in the case of UK but not to Spain because of the absence of a minimum income scheme in that country.

3. Application to three EU countries

The methodology, which has just been presented, is now applied to data from France, Spain and UK. This application draws on a prototype version of the European tax-benefit model EUROMOD. This model simulates the tax-benefit systems of EU countries using representative household samples in each country. To keep with the logic of the optimal taxation model, all households with zero income and with non-labour income, including pension and unemployment benefits above 10 per cent of total income were eliminated from the sample. Disposable income is computed with official rules for taxes and benefits instead of being taken directly from the data. Effective net marginal tax rates are calculated using the same rules. Following the last set of remarks in the preceding section, several applications have been run. They differ with respect to the value selected for the elasticity of labour supply and the choices made for handling household size.

Results are summarized by curves showing the marginal social welfare $G'( )$ for by percentiles of the population of households ranked by level of productivity as well as the upper incomplete means curve $S()$.

Two figures are shown for each country. The first figure comprises three panels (except for France where a fourth panel appears). The top left hand panel shows the effective net marginal tax rates, as computed on the basis of official rules modelled in Euromod, for the various percentiles of the population ranked by productivity level. The bottom left panel shows the distribution of productivity consistent with the distribution of gross labour incomes and observed marginal tax rates under two alternative assumptions about the elasticity of labour-supply, a 'very low' elasticity case with $\varepsilon=.1$ and a 'moderate' elasticity case with $\varepsilon=.5^{19}$. In both cases, the mean productivity is normalized to unity. Finally, the right hand panel shows the

---

18 See Immervoll et al. (2000).
19 The latter may be considered as a rough average estimate obtained in the labour supply econometric literature (see Blundell and MaCurdy 1999).
marginal social welfare the function $G'(w)$ and its upper incomplete mean $S(w)$ as derived from (9) and (11) respectively by productivity percentile. Here again, the mean marginal social welfare is normalized to unity. Both curves are drawn for the two values selected for the elasticity of labour supply. On the same panel, with the same scale, is represented the function $S(w)$ [equation (11)] that allows, as had before, for a robust test of the sign of the marginal social welfare $G'(w)$. In the case of France the bottom right panel shows (panel 1d) the function $S(w)$ computed under an alternative assumption about individual preferences. The second figure for each country shows the marginal social welfare curves when the methodology is applied separately to different household types.

The first set of figures calls for several remarks. In the case of France and UK, the original marginal tax rate curve has a U-shape. It is extremely high at the bottom of the distribution because of households facing marginal tax rates equal to 100 per cent due to the minimum guarantee (RMI in France and Income Support in UK). Then, the marginal tax rate falls until a little after the median and then increases slowly with the progressivity of the income tax. In the Spanish case, where no minimum income guarantee does exist, the effective marginal tax rate start from zero and increases regularly reaching, at the end of the distribution, values around 50%. In panels showing the productivity density functions associated with the distribution of gross incomes and effective marginal tax rates, less inequality is obtained for productivities when the elasticity of labour supply is moderate than when it is very low. This is because the distribution of gross income tends to amplify more the inequality of productivity in the first case. Thus, for a given distribution of gross incomes and marginal tax rates, there is less inequality in productivities when labour supply is more elastic. Turning now to the marginal social welfare curves, three features are readily apparent: a) overall, marginal social welfare is declining; b) its slope is more pronounced with the medium than the low elasticity of labour supply; c) the marginal social welfare is not always positive at the top of the distribution of productivities.

That revealed marginal social welfare is declining with the level of household productivity, - except for the very top of the distribution in UK, which may be due to the instability of Kernel in a range with very distant observations - suggests that the redistribution system in these three countries is consistent with the hypothesis of an optimising redistribution authority. Clearly, the basic optimisation problem (1) would not make sense if the objective function were not concave. This is an interesting result, which was certainly not guaranteed by the inversion methodology used in this paper. That the slope of the marginal social welfare

---

20 The meaning of it will be explained in section 4.
curve increases significantly with the elasticity of labour supply was to be expected. It was seen above that the observed distribution of gross incomes led to more inequality in the distribution of productivity when the elasticity of labour supply was low – in other words, the elasticity of labour supply contributes to amplifying the inequality of productivity. Since redistribution remains the same, it must be the case that revealed social preferences are less averse toward inequality when the elasticity of labour supply is low. The difference appears clearly for all countries.

The third noticeable result is that the marginal social welfare becomes negative for high levels of productivity and for the medium value of the labour supply elasticity. This phenomenon is quite strong in the case of France where the marginal social welfare becomes negative around the 9th decile and the upper incomplete mean marginal social welfare turns to negative above the 8th decile. If the French authority indeed believes that the elasticity of labour supply is around .5, this means that it is non-Paretian, social welfare being maximum around the 9th decile of the distribution. In other words, social welfare is directly increased by reducing the income of the richest 10 per cent of the population, the only reason why it is not optimal to further reduce it through taxes being the loss of tax receipts and therefore transfers to the bottom part of the distribution. Of course this conclusion would not hold anymore if the redistribution authority thought the elasticity of labour supply were closer to .1 than .5. Thus the general conclusion must be weakened somewhat. Either the French redistribution believes labour supply is little elastic to the net wage rate and behaves as a Paretian planner, or it believes that the labour supply is somewhere in the range given by the econometricians authority and it is non-Paretian.

In can be seen in figure 3c that the same conclusion is obtained with Spain although the marginal social welfare becomes negative after the 9th decile when the labour supply elasticity is equal to .5. Things are a little less clear in the case of UK because Kernel estimates prove somewhat unstable at the right end of the distribution. But, it is readily apparent in (11) that both functions G(w) and S(w) would become more negative and sooner for slightly higher labour-supply elasticities.

Turning now to the case where households of different size are handled separately - it may be seen in figures 2, 4 and 6 that the inversion of the marginal rates into marginal social welfare curve yields the same type of results. Marginal social welfare is decreasing throughout the whole productivity range for singles couples or couples with children. Some of the
irregularities observed with the whole population even disappeared in the case of the UK.\footnote{This suggests that it was probably due to heterogeneity of the marginal tax rates for high levels of income among various demographic groups.} Marginal social welfare still becomes negative at the top of the distribution for all household groups in France – and for couples with 2 children in UK and Spain. Thus, the conclusion obtained above that redistribution authorities are non-Paretian, under the assumption $\varepsilon = .5$, is not the result of aggregating social preferences across the various household groups. In general, it remains true for particular household groups. As for the whole population, revealed preferences for specific household groups are Paretian when the elasticity of labour supply is low. Together with the remark above that increasing $\varepsilon$ necessarily makes the upper incomplete mean marginal social welfare more negative and sooner, this confirms the conclusion sketched above. If any optimisation is actually taking place, observed marginal tax rates reveal the following. Either redistribution authorities maximise a « well behaved » social welfare function (i.e. increasing and concave) but have a low subjective estimate of labour supply elasticity. Or they expect higher labour supply elasticities, in which case they have non-Paretian preferences, possibly shaped by political economy mechanisms. It is interesting that more heterogeneity appears across countries in that case. In particular preferences are more strongly non-Paretian in France than in UK and Spain, which may reflect different types of political economy equilibria behind the existing tax-benefit schedules. In effect, higher estimates of the elasticity of labour supply would be necessary to get the same of non-Paretian preferences in the last two countries.

Which one of the two explanations is the most plausible? The key element for the answer clearly is the size of the labour supply elasticity. Looking at the results in the literature on the econometrics of labour supply (Blundell and MacCurdy 1999) an elasticity of substitution of .5 appears to be within the range of estimates for households' second workers, which seems the relevant concept to use when households are considered as a single agent. But of course, elasticities slightly higher than .5 are not unlikely, which would then lead to marginal social welfare becoming negative sooner than what appears in the preceding figures. Based on this argument, it is tempting to support the idea that fiscal authorities (of France at least; and to a lesser degree in UK and Spain) may not be Paretian and/or that observed redistribution is the outcome of a collective bargaining process more than the result of the maximisation of a « well behaved » social welfare function.

The robustness of the preceding results must be stressed. It is true that the use of kernel techniques to compute first derivatives of $t(w)$ and $f(w)$ needed to recover the $G'(w)$ function (equation 9) may introduce some bias because of the ambiguity of choosing the appropriate
adaptive windows for the kernel. This bias affects the robustness of the results concerning $G'(w)$. But this is not the case for the upper incomplete mean marginal social welfare curve $S(w)$, which, according to (11) depends only on the marginal tax rate schedule, $t(w)$, and the estimated distribution of productivities, $f(w)$.  

4 Income effects

All the preceding results are based on the assumption that individual preferences were quasi-linear in consumption and iso-elastic in labour supply. With this specification the income effect on labour supply is simply nil. A slightly less restrictive form consists of assuming separability of consumption and labour without assuming quasi-linearity. Preferences are thus represented by a function of the following type:

$$U(y, L) = A(y) - B(L)$$

where $A(y)$ is not supposed to be linear anymore. In that case, it may be shown that the optimal taxation formula (4) becomes:

$$\frac{t(w)}{1-t(w)} a[y(w)] = (1+\frac{1}{\varepsilon}) \frac{1-F(w)}{w f(w)} [\bar{a}[y(w)] - S(w)/S(w_0)]$$  

(14)

where $a(y)$ is the inverse of the marginal utility of income $A(y)$, $\pi(y)$ is the mean value of that inverted marginal utility for people with productivity above $w$. The same inversion technique as above may be implemented but this requires some additional assumption on the functions $A(y)$ and $B(L)$. When $A(y)$ and $B(L)$ are both iso-elastic functions, the elasticity of labour supply becomes constant again, so that it its possible to keep the same approach to estimating the distribution of productivities as above. Yet, it is still necessary to choose arbitrarily the elasticity of one of the two functions $A(y)$ or $B(L)$. The following analytical specification for $U(y, L)$ was tried:

---

22 Basically, the problem is that the optimal adaptive window of the kernel must be adjusted depending of the degree of differentiation of the functions. So, the optimal window for the kernel of the first derivative is not the same of the optimal windows for the second derivative and so on (for a detailed treatment of this problem see Pagan and Ullah 1999).

23 The bias on $G'(w)$ explains why, in some case (for example in the case of France, whole sample, $\varepsilon = 0.5$, panel 1c), the $S(w)$ and the $G'(w)$ function becomes negative at different levels of the productivity distribution.
\[ U(y, L) = \frac{y^{\frac{1}{\alpha}}}{\frac{1}{\alpha}} - k \frac{L^{\frac{1}{\beta}}}{\frac{1}{\beta}} \] (15)

which leads to the constant labour supply elasticity \( \varepsilon = \frac{\beta(\alpha - 1)}{\alpha + \beta} \).

Results obtained with this specification are presented in panel (1d) for the whole French sample, with the following two sets of parameters values \((\alpha = 2, \beta = 2)\) and \((\alpha = 5, \beta = 5/7)\) both leading to \( \varepsilon = .5 \) but with different marginal utilities of income\(^24\).

It may be seen that, in both cases, the upper incomplete mean marginal social welfare, \( S(w) \), is decreasing but negative beyond decile 7. These results suggest that the non-Paretian nature of the social welfare function in presence of a medium value for the elasticity of labour supply does not depend on the assumption of quasi-linearity in consumption.

**Conclusion**

This paper has explored an original side of applied optimal taxation. Instead of deriving the optimal marginal tax rate curve associated with some distribution of individual productivities, the analysis consists of retrieving the marginal social welfare functions that makes the observed marginal tax rates optimal under an arbitrary assumption about the wage elasticity of labour supply.

For the three countries analysed in this paper, France, Spain and UK, revealed marginal social welfare curves were found in agreement with standard theory when the elasticity of labour supply was assumed to be low. Marginal social welfare was both positive and decreasing throughout the range of individual productivities, and therefore of individual utilities. However, marginal social welfare turned out to be negative at the very top of the distribution when the labour supply elasticity was assumed to be around the average of estimates available in the literature. This phenomenon was present in the three countries, although more pronounced in the case of France. The same result was also obtained with various specifications of household preferences between labour and consumption.

Two lessons may be drawn from all this exercise. The first sheds some doubt about the idea that the real world is as if a redistribution authority were maximizing some well behaved social welfare function. It was found in this paper that its behaviour could be of three different types. Either the redistribution authority under-estimates the labour supply response to taxation,

\(^{24}\) To compute the terms \( a(y) \) and \( \bar{a}(y) \) in formula (14) we used the observed disposable income as proxy for the optimal consumption \( y \).
or it has non-Paretian social preferences, or it does not optimise at all. This conclusion is not really surprising. To some extent, the last two cases, which seem the most likely, are even reassuring. Indeed, tax-benefit schedules in the real world appear to result more from political economy forces than the pursuit of some well defined social objective.

The second lesson is the practical interest of reading actual tax-benefit systems through the social preferences that they reveal. It is customary to discuss and evaluate reforms in tax-benefit systems in terms of how they would affect some 'typical households' and more rarely what their implications are for the whole distribution, of disposable income. The instrument developed in this paper offers another interesting perspective. By drawing marginal social welfare curves consistent with a tax-benefit system before and after reforms, it is possible to characterize in a more precise way the distributional bias of the reform.
Figure 1. Derivation of the marginal social welfare function: France, all households

Figure 1a. Marginal tax rates by productivity ranking (Kernel Estimates)

Figure 1b. Kernel estimates of productivity density

Figure 1c. Marginal social welfare by productivity ranking with quasi-linear utility function (Kernel estimates)

Figure 1d. Marginal social welfare by productivity ranking with income effects (Kernel estimates)
Figure 2. Marginal social welfare curves: France, household size groups treated separately
(Kernel Estimations)
Figure 3. Derivation of the marginal social welfare function: Spain, all households

Figure 3a. Marginal tax rates by productivity ranking (Kernel Estimates)

Figure 3b. Kernel estimates of productivity density

Figure 3c. Marginal social welfare by productivity ranking (Kernel estimates)
Figure 4. Marginal social welfare curves: Spain, household size groups treated separately
(Kernel Estimations)
Figure 5. Derivation of the marginal social welfare function: UK, all households

Figure 5a. Marginal tax rates by productivity ranking (Kernel Estimates)

Figure 5b. Kernel estimates of productivity density

Figure 5c. Marginal social welfare by productivity ranking (Kernel estimates)
Figure 6. Marginal social welfare curves: UK, household size groups treated separately
(Kernel Estimations)
References

D’Autume, A. (1999), Fiscalité optimale: une application au cas français, Mimeo, EUREQua, Université de Paris I
Salanié, B. (1998), “Note sur la Taxation Optimale”, Rapport au Conseil d'Analyse Economique, La Documentation Française, Paris