Tax-benefit revealed social preferences

François Bourguignon, Amadéo Spadaro

PARIS-JOURDAN SCIENCES ECONOMIQUES
48, BD JOURDAN – E.N.S. – 75014 PARIS
TEL. : 33(0) 1 43 13 63 00 – FAX : 33 (0) 1 43 13 63 10
www.pse.ens.fr

WORKING PAPER N° 2005 - 22

Tax-Benefit Revealed Social Preferences

François Bourguignon

Amedeo Spadaro

Codes JEL : H11, H21, D63, C63

Mots clés : Social Welfare Function, Optimal Income Tax, Microsimulation, Optimal Inverse Problem

Tax-Benefit Revealed Social Preferences

François Bourguignon
The World Bank
and
PSE -Paris-Jourdan Sciences Economiques-
(Joint research unit 8545 CNRS -EHESS-ENPC-ENS), Paris

and

Amedeo Spadaro
PSE -Paris-Jourdan Sciences Economiques-
(Joint research unit 8545 CNRS -EHESS-ENPC-ENS), Paris
and
Universitat de les Illes Balears, Palma de Mallorca

Abstract[1]

This paper inverts the usual logic of applied optimal income taxation. It starts from the observed distribution of income before and after redistribution and corresponding 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. In this framework, the issue of the optimality of an existing tax-benefit system is transformed into the issue of the shape of the social welfare function associated with that system and whether it satisfies elementary properties. This method is applied to the French redistribution system with the interesting implication that the French redistribution authority either has a rather low estimate of the labor supply elasticity or does not give positive social weights to the richest tax payers.

Keywords: Social Welfare Function, Optimal Income Tax, Microsimulation, Optimal Inverse Problem.

JEL Classifications: H11, H21, D63, C63.

Résumé
Cet article renverse la logique classique de la littérature de la fiscalité optimale. On part de la distribution observée de revenus bruts et de revenus disponibles d'une population et des taux d'imposition marginaux observés calculés par un modèle de microsimulation. On montre alors que, sous certaines conditions simplificatrices, il est possible d'identifier la fonction de bien-être social qui rendrait optimal le schéma des taux marginaux observés sous certaines hypothèses sur les préférences consommation-loisir. Dans ce cadre, la question de l'optimalité d'un système tax-benefit concret peut-être analysée en vérifiant si la fonction de bien-être sociale associée satisfait certaines propriétés. Cette méthode est appliquée au système de redistribution français. On observe que, ou l'autorité fiscale assigne des valeurs peu élevées aux élasticités de l'offre de travail, ou bien elle assigne des poids sociaux négatifs aux agents les plus riches.

[1] This is a completely revised version of the DELTA WP 2000-29. We thank Tony Atkinson, Roger Guesnerie, Jim Mirrlees and participants to seminars in Barcelona, Madrid, Berlin, Paris, Venezia for useful comments. We also thank Pascal Chevalier and Alexandre Baclet from INSEE to help us with the French Fiscal Data. We are solely responsible for any remaining error. Amedeo Spadaro acknowledges financial support of Spanish Government (Programa Nacional de Promoción General del Conocimiento SEC2002-02606) and Fundación BBVA.
Correspondence: Amedeo Spadaro, Universitat de les Illes Balears, Department of Economics, Ctra Valldemossa Km 7,5 07122 Palma de Mallorca, Spain, email : amedeo.spadaro@uib.es.

Introduction

Several attempts were recently made at analyzing existing redistribution systems in several countries within the framework of optimal income 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 implied by some minimum income programs - see Piketty (1997), d'Autume (2001), Choné and Laroque (2005) and Bourguignon and Spadaro (2000) 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) on the basis of arbitrary parametric representations of the distribution of individual abilities. The exercise may now seem more relevant because of the possibility of relying on large and well documented micro data sets giving some indication on the 'true' distribution of abilities. The results obtained when applying the standard optimal taxation calculation 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 distribution of abilities meant to approximate real world distributions makes 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 yield the same range of results, what should one conclude? Should one refer to a Rawlsian objective and conclude that some part of observed redistribution systems are clearly sub-optimal, or should one use a less extreme assumption for the social welfare function and then conclude that another part of the redistribution schedule in non-optimal? The point of view taken in this paper is the opposite of this standard approach. The focus is here on the social welfare function that makes optimal the effective marginal tax rate schedule that corresponds to the redistribution system actually in place. This approach may thus be considered as the dual of the previous one. In the standard approach, 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 present case, it consists of checking whether the social welfare function implied by the actual redistribution schedule is in some sense 'reasonable', that is in particular whether the marginal social welfare is everywhere positive and decreasing. The approach proposed here provides a way of 'reading' effective average and marginal tax curves that are commonly used to describe a redistribution system. This reading consists of translating the observed shape of these curves into social welfare language. The issue of comparing two redistribution systems, or analyzing the reform of an existing system, can thus be cast in terms of social welfare rather than the distribution of abilities and/or labor supply behavior. Instead of determining who is getting more out of redistribution and who is getting less, both overall and at the margin, 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, these 'revealed social preferences' rely on several auxiliary assumptions about labor supply behavior and about the distribution of individual abilities. With the direct or standard approach to optimal taxation, 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. The observation of the effective average or marginal tax rate schedule may thus reveal more than social preferences. It may suggest that the tax schedule is inconsistent with optimality. But, in some instances, it may also reveal that some common assumptions on labor supply behavior or on the distribution of abilities are inconsistent. This seems equally useful information.

To our knowledge, this paper represents the first attempts to 'reveal' the implicit social welfare preferences by applying an 'optimal inverse' technique to direct taxation, within the framework of Mirrlees' optimal labor income tax model. A similar approach has been used in the field of indirect taxation by Ahmad and Stern (1984). In that paper, they apply the optimal inverse method to the indirect taxation system in India and conclude that tax authorities are not Paretian in the sense that some agents have a negative marginal weight in the revealed social welfare function. They then derive a set of tax reforms that are Pareto improving over the status quo situation. A more recent theoretical paper by Laroque and Choné (2005) use the optimum inverse within Mirrlees' optimal direct redistribution framework but focus on the distribution of individual abilities rather than the social welfare function. More precisely they show that there always exists a distribution of abilities - conditional on individual preferences - that make an observed income redistribution schedule marginal tax rate schedule optimal with a Rawlsian welfare function. However, they do not apply empirically their inversion method so that it is difficult to know how 'reasonable' would be the 'revealed' distribution of abilities under the assumption of Rawlsian social welfare. In comparison with Laroque and Choné (2005), the present paper inverts the optimal taxation model with respect to social welfare rather than the distribution of abilities and provides an empirical application of this method. 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 France, taking advantage of the easy identification of marginal tax rate schedule with the EUROMOD model[2]. In each case, the social welfare function is characterized under a set of simple alternative assumptions about the labor supply elasticity, the distribution of individual abilities being itself derived from observed labor incomes. Section 4 analyzes the case in which income effects are considered and, finally, Section 5 concludes.

This paper is both methodological and factual. On the methodological side, it shows how the characteristics of any given redistribution system may be expressed in social welfare terms. On the factual side, the main lesson drawn from the practical applications handled in the paper is essentially that revealed social preferences satisfy the usual regularity assumption – positive and decreasing marginal social welfare – as long as the wage elasticity of labor supply is below some threshold. In the case of France it cannot be ruled out that revealed social preferences are non-Paretian beyond some income level if labor elasticity is high enough, but unreasonably so.

1. The duality between optimal marginal tax rates and the social welfare function

The basic optimal taxation framework is well known.[3] Agents are assumed to choose the consumption (y) /labor (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( )* the net tax schedule. If the distribution of agents' productivity in a population of size unity is represented by the density function *f(w)* defined on the support $[w_0, Z]$, the optimal taxation problem may be written as:

$$Max_{T()} \int_{w_0}^{Z} G[V[w,T(\ )]]\, f(w)dw \qquad (1.1)$$

$$s.t: \quad (y^*,L^*) = Argmax[U(y,L);\ y = wL - T(wL),\ L \geq 0] \qquad (1.2)$$

$$V[w,T(\ )] = U(y^*,L^*) \qquad (1.3)$$

$$\int_{w_0}^{Z} T(wL^*)f(w)dw \geq B \qquad (1.4)$$

where *G[ ]* is the social welfare function that transforms individual utility, *V( ),* 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*, a case extensively used in both the theoretical and applied optimal tax literature[4]:

[2] See Sutherland (2001).
[3] See for instance Atkinson and Stiglitz (1980) or Tuomala (1990).
[4] See in particular Atkinson (1995) or Diamond (1998).

$$U(y,L) = y - k^{-\frac{1}{\varepsilon}} \frac{L^{1+\frac{1}{\varepsilon}}}{1+\frac{1}{\varepsilon}} \tag{2}$$

where k is a constant and ε is the elasticity of labor supply, L^*, with respect to the marginal return of labor. Indeed, together with (2) the solution of (1.2) above yields the labor supply function given by the solution of the following equation: $L^* = kw^{\varepsilon}\left[1 - T'(wL^*)\right]^{\varepsilon}$ (3)

It can be shown – see for instance Atkinson (1995) - that this particular case leads to the following simple characterization of the optimal tax schedule:

$$\frac{t(w)}{1-t(w)} = (1+\frac{1}{\varepsilon})\frac{1-F(w)}{w.f(w)}(1-S(w)/S(w_0)] \tag{4}$$

In that expression, *t(w)* is the (optimal) marginal tax rate faced by an agent with productivity, *w*, and therefore with (gross) earnings *wL** -i.e. *t(w) = T'(wL*)*. *F(w)* and *f(w)* are respectively the cumulative and the density functions associated to the distribution of productivity in the population. Finally, *S(w)* stands for the average marginal social utility of all agents with productivity no smaller than *w*, which is given by:

$$S(w) = \frac{1}{1-F(w)} \int_{w}^{Z} G'\left[V(w,T())\right] f(w) dw \tag{5}$$

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 arbitrary tax system, the government decides to increase the tax payment by a small increment *dT* for each agent whose labor income is equal or higher than *Y* and labor productivity equal or higher than *W*, leaving the rest of the tax schedule unchanged. Such a measure has three effects: a) it reduces the labor supply of people with income in the neighborhood of *Y* because the marginal return to their labor 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). With the optimal tax system, the total effect of these changes on social welfare must be equal to zero for all *Y*.

The tax reduction effect a) depends on the marginal rate of taxation, *t(W)*, the elasticity of labor supply, ε, 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]

$$TR = \frac{t(W)}{1-t(W)} \frac{W.f(W)}{1+1/\varepsilon} dT$$

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*:

$$TI = [1 - F(W)]dT$$

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 the term $S(w_0)(TI - TR)$. 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 in the neighborhood of *W* – are not affected because they compensate the drop in the effective price of their labor and its negative effect on consumption by a reduction in the labor 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*:

$$[1 - F(W)]\frac{S(W)}{S(w_0)} = \frac{TI - TR}{dT} \qquad (6)$$

which, after rearranging, leads to (4) above.

What is attractive in the preceding expression is that the right-hand side is essentially of a positive nature whereas the left-hand side is 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 also exhibits 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-

---

[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 labor 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.

benefit system in a given economy and its implied effective marginal tax rate schedule, provided that some estimate of the labor supply elasticity is available. Then the left-hand side of (6) yields information on the social welfare function that is consistent with the observed 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) corresponds to the "optimum inverse method"[6].

Characterizing precisely the social welfare function, *G(w)*, implied by a tax-benefit system under the assumption that it is indeed optimal requires some additional steps beyond (6). First, normalize the welfare function *G(w)* so that the mean marginal social welfare for the whole population is equal to unity, $S(w_0) = 1$. Equation (6) now writes simply:

$$S(w) = 1 - \frac{t(w)}{1-t(w)} \frac{1}{1+1/\varepsilon} \frac{w.f(w)}{1-F(w)} \qquad (7)$$

The function S(w), which stands for the 'average upper marginal social welfare' of people with productiviy equal or greater than w may this be recovered from the knowledge of the marginal tax rate schedule, *t(w)* and the distribution of abilities, i.e. *f( )* and *F( )*.

Identifying the marginal social welfare functions, *G'( )* , itself requires an additional differentiation step. Differentiating the definition of *S(w)* in (5) it comes easily that :

$$G'[V(w,T()] = -\Theta'(w)/f(w) \qquad (8)$$

where $\Theta(w) = [1-F(w)] - \frac{t(w)}{1-t(w)} . \frac{w.f(w)}{1+1/\varepsilon}$

In other words, the knowledge of the actual marginal tax schedule *t(w)* and the distribution of individual productivities, *f()* and *F( )* , are sufficient to derive the marginal social welfare associated with each productivity level, *w*, provided that a continuous and differentiable approximation of the 'upper average marginal social welfare function', *S(w)*, is available. However, it must be kept in mind that, through the differentiation of that function S(w), the identification of the marginal social welfare function, *G'()*, is much more exigent. Indeed, it requires the knowledge not only of the tax schedule, *t(w)*, and the density function, *f(w)*, but also of their elasticity with respect to individual productivity, *w*. More precisely, it can be seen by taking the derivative of (7) that the analytical expression of *G'( )* is given by:

[6] See, for instance, Kurz (1968). Going back to expression (4) above the optimum inverse problem considered in this paper consists of identifying S(w) given the knowledge of t(w), f(w) and ε. Choné and Laroque (2005) solve a symmetric problem by identifying the pair (f(w), ε) knowing t(w) and S(w).

$$G'[V(w,T())] = 1 + \left(\frac{1}{1+1/\varepsilon}\right)\left(\frac{t(w)}{1-t(w)}\right)\left(1+\eta(w)+\frac{\nu(w)}{1-t(w)}\right) \qquad (9)$$

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

Recovering the upper average marginal social welfare function, *S(w)*, , thus only requires primary data, that is marginal tax rates and the frequency and cumulative distribution of productivities. Because of this, the estimate that can be obtained of *S(w)* is likely to be much more robust than that of *G'(w)*. Most of the empirical application in this paper will thus mostly be based on upper average marginal social welfare rather than marginal welfare.

Overall, the way in which the marginal tax rate schedule of a tax-benefit system maps into the marginal social welfare is far from simple. It indeed combines the marginal tax rate schedule itself, its derivative, the density of the distribution of abilities and its derivatives. Yet, based on the 'upper average marginal social welfare' function, *S(w)*, some simple rules allowing for a test of the consistency of an observed tax-benefit system and an observed distribution of productivities with the optimal taxation framework may be obtained.

Proposition 1. *A necessary condition for the social welfare function making the observed effective marginal tax rate schedule, t(w), optimal with respect to the observed distribution of productivities, f(w) to be Paretian - e.g. non-decreasing everywhere- is that :*

$$t(w) \le \frac{\dfrac{1+\varepsilon}{\varepsilon}\dfrac{1-F(w)}{w.f(w)}}{1+\dfrac{1+\varepsilon}{\varepsilon}\dfrac{1-F(w)}{w.f(w)}} \quad \text{for all } w \in [w_0, Z] \qquad (10)$$

The proof of that proposition is easily established. If the social welfare function is Paretian, the derivative of G( ) is positive everywhere and S(w), as defined by (5) too. Inequality (10) then follows from (7). This is only a necessary condition, but its interest is that it relies only on the knowledge of the marginal tax rate schedule and the distribution of productivities and should therefore be more robust that dealing directly with expression (9) of marginal social welfare. Where the distribution may be approximated by a Pareto with parameter *a*, the preceding condition may be simply expressed as a ceiling on the marginal tax rate. Given that $\frac{w.f(w)}{[1-F(w)]} = a$, it comes that:

$$t(w) \le \frac{1+1/\varepsilon}{1+1/\varepsilon + a} \qquad (11)$$

For instance, with not unreasonable figures like a = 3 and ε =0.5, this condition states that a redistribution system where the marginal tax rate would exceed 50 per cent could be deemed 'optimal' only on the basis of a non-Paretian social welfare function.

Proposition 2. *If the elasticity of the marginal tax rate and the density function are bounded, then there exists a threshold for the wage elasticity of labor supply below which the social welfare function is necessarily non-decreasing everywhere.*

This proposition follows directly from (9). If indeed $\eta(w)$ and $\nu(w)$ take only finite values, the second term on the RHS of (9) can be made small enough by allowing ε to tend towards zero. Thus there always exists a value of ε small enough so that marginal social welfare is positive for all values of w.

This property shows the importance of the assumption made on the wage sensitivity of labor supply to judge the optimality of a given redistribution system. *Any redistribution system may be said to optimize a Paretian social welfare function, provided that the redistribution authority has a low enough estimate of the wage elasticity of labor supply.*

Proposition 3. *Wherever the marginal tax rate is increasing with income, a sufficient condition for the social welfare to be everywhere non-decreasing is:*

$$t(w) \leq \frac{1+\varepsilon}{1-\eta(w)\varepsilon} \qquad (12)$$

Again, this proposition is directly derived from (9). It is of relevance in connection with the discussion on whether the marginal tax rate curve must be U-shaped – see Diamond (1998) and Saez (2001). In that part where the marginal tax rate is increasing, that is for high incomes, (12) gives an upper limit for the marginal tax rate – in the reasonable case where $\eta(w)$ is negative of course. It can be checked that this condition is the same as (11) in the case where the productivity's distribution may be approximated by a Pareto. On the other hand, condition (12) becomes a necessary condition for marginal social welfare to be non-negative when the marginal tax rate is decreasing - for instance for low incomes.

Going further in checking the consistency of a redistribution system with optimizing a well-behaved social welfare function requires studying the concavity of the revealed social welfare function. Conditions for concavity are analyzed in appendix 3. It can be seen there that they are based on the 'curvature' of both the tax schedule t(w) and the density function f( ). As not enough precision can be obtained on these functions, the empirical application that follows mostly focuses on the shape of function S(w), and whether it is decreasing, an obvious implication of the concavity of the social welfare function G( ).

The rest of this paper focuses on the sign and the slope of the function S(w) and *G'(w)* that may be associated with the tax-benefit systems and the distribution of productivities observed on actual data provided some basic assumptions on ε. The key test is whether S(w) is everywhere positive and decreasing. If the first condition is not satisfied, then the revealed social welfare function may not be deemed to be Paretian. If the second condition does not hold, then it is the whole optimization concept behind Mirrlees framework that would become irrelevant. It would indeed make no sense to assume that the redistribution authority attempts to maximize a non-concave welfare function if other than trivial redistributions policies are observed.

Two final remarks are necessary at this stage. The first relates to the previous statements and is to stress that the optimal tax conditions (4) do not require that social preferences be Paretian but only concave. Indeed, these conditions are necessary and sufficient for all *concave* social welfare functions and not only those which are monotonically increasing. This may be seen by rewriting the original optimization 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:

$$Max \int_{w_0}^{Z} G[u].f(w).dw$$

whereas the motion equation is defined by a condition of the type $\dot{u} = du/dw = g(Y,u)$, 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 this is not observed 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 or simply large enough. This latter assumption is maintained in the rest of this paper[7].

2. Empirical implementation issues

[7] On this, see also the argument in Diamond (1998).

The previous methodology requires estimates of the elasticity of labor supply, ε, 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 labor income, $Y=wL$, and disposable income, y, or by difference, total taxes and benefits, $T(wL)$[8]. 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 labor income by a small amount. To be in the situation to apply the optimum inverse method analyzed above, it is thus necessary to impute a value of the productivity parameter, w, to the households being observed with total income Y and then to estimate the statistical distribution of individual productivities, *f(w).*

When labor supply, L, is observed, the simplest way to proceed would consist of assimilating productivity with observed hourly wage rates, and then using an econometrically estimated values for the labor supply elasticity, ε, which, without loss of generality, might even be specified as a function of productivity, w (as it has been done in the previous work on applied optimal income tax). This is the first approach pursued below.

Although simple, the preceding approach can be inappropriate for several reasons. First, the distribution of hourly wages may be an imperfect proxy for the distribution of productivities because actual labor supply may differ quite significantly from observed working hours when unobserved efforts are taken into account. Second, econometric estimates of the labor supply elasticity are extremely imprecise, and ambiguous. Econometric estimation requires taking into account the non-linearity inherent to most tax-benefit systems and the endogeneity of marginal tax rates that it entails. Moreover, econometric estimates derived from these non-linear models are known to be little robust[9]. On the other hand, relying on simpler alternative estimates based on simple linear specifications introduces some arbitrariness in the estimation procedure. Third, econometric estimates of the elasticity of labor supply, whether they are obtained from models with endogenous or exogenous marginal tax rates, 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 ε ? Fourth, and more fundamentally, it seems natural that a welfare analysis of taxes and benefits focus on households rather than individuals. 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 an "household productivity"?

[8] To keep with the logic of the optimal taxation model, non-labour taxable income is ignored in all what follows.
[9] See Blundell, Duncan and Meghir (1998).

Likewise, if the elasticity of labor supply has been estimated at the individual level and is different across various types of individuals, how should it be averaged within the household?

An alternative approach to the complex econometric estimation problems just described is the following. Instead of assuming that observed hourly wages and hours of work are good proxies for individual productivities and labor supply, and deriving from them an estimate of labor supply elasticity, the whole procedure is inverted. An arbitrary value of the elasticity of labor supply is chosen within the range of values found in the literature. Then, this value is used to derive the implicit productivity and labor supply of households or individuals from observed labor incomes.

The latter operation is a simple inversion of the labor supply equation (3). Multiply both sides of that equation by *w* so that the gross labor income, Y, appears on the left hand side:

$$Y = wL^* = k.w^{1+\varepsilon}.[1-T'(wL^*)]^{\varepsilon} \qquad (13)$$

After inversion, one gets for a given value of ε:

$$w = Y^{\frac{1}{1+\varepsilon}}[k(1-T'(Y))]^{\frac{-\varepsilon}{1+\varepsilon}} \qquad (14)$$

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

The preceding inversion procedure allows for a consistent definition of all the variables of which observation is necessary for recovering the social welfare function from the optimal taxation formula. Moreover, this procedure may be applied to individual agents as well as households comprising various potential earners. For household *i*, observed with gross labor income, Y_i, and marginal tax rate, t_i, a value of the implicit productivity characteristic, w_i, may be imputed through (14). Then all households may 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.

3. Application to the French redistribution system

A key parameter of the methodology for recovering the social welfare function from marginal tax rates is the labor supply elasticity. As mentioned above, there is considerable imprecision about the value of that parameter, which moreover is likely to depend on individual characteristics like gender, age, marital status or household composition.

A recent survey of estimation techniques and results obtained in studies of labor supply in UK and US by Blundell and MaCurdy (1999) gives a range of values mostly concentrated in the interval (0,1). In the case of France, Bourguignon and Magnac (1991), Piketty (1998), Donni (2000), Bargain (2004), Choné et al. (2003) and Laroque and Salanié (2002) found labor supply elasticity estimates in the same interval. Values between 0.1-0.2 are found for men and an average of 0.5 is found for married women - and slightly more if they have children (Bargain, 2004).

Similar results have been obtained on the basis of the relationship between taxable incomes and changes in tax rates. The difference in difference estimation performed on tax returns panel data by Feldstein (1995) using the 1986 US tax reform as a natural experiment yields estimates on the high side, in the interval [1, 3]. However, more recent work by Auten and Carroll (1999) and Gruber and Saez (2002) have questioned such high values. In particular, they have shown that Feldstein's estimates were probably affected by a bias coming from the mean reversion tendency present in panel income data. Improving on Feldstein's methodology lowered the estimates of the elasticity of US taxable income to values around 0.7 in Auten and Carroll (1999), and 0.4 in Gruber and Saez (2002). In the case of France, Piketty (1998) found elasticies of taxable income even lower, around 0.1.

In line with the empirical findings for France, we shall be working in what follows with two extreme values of the labor supply elasticity, a low-value equal to 0.1 and a high value equal to 0.5. It turns out that these two values are sufficient to illustrate the various conclusions that may be drawn from the analysis. Appendix 1 and 2 give more technical detail about the implementation of the preceding methodology to French data as well as about the datasets and the micro-simulation model being used.

Several calculations have been performed. They differ depending on the definition of the redistribution system, the definition of individual productivities and the sample being used. The first definition of the redistribution system includes income taxes and assimilated contributions like the 'Cotisation Sociale Généralisée' and all non-contributory benefits. In other words, this definition includes all taxes and benefits with an 'explicit' redistributive role.[10] The corresponding marginal tax rate is referred to as '*net*' in what follows, in the sense that it does not incorporate social contributions paid by employers or workers. The second definition of the redistribution system adds contribution to health insurance on the 'tax' side. In France, that contribution is levied on all labor

[10] This is equivalent to considering that other taxes, including indirect taxes, which are mostly neutral with respect to consumption, are essentially aimed at covering non-redistributive public expenditures. The implications of this hypothesis are discussed in Atkinson and Stiglitz (1980), chapter 9.

incomes at a virtually uniform rate whereas the corresponding benefits - that is health insurance - may be considered, as a first approximation, as being the same for the whole population and, in any case, very imperfectly related to income and therefore to the contribution itself. Thus, the redistributive role of the health insurance system is quite substantial and is generated essentially by the quasi proportionality of contributions with respect to income.[11] By contrast, most other contributions, for instance contributions to pensions or unemployment insurance give rise to a delayed benefit that, in actuarial terms and as a first approximation, is not very different from the value of contributions. Even though actuarial neutrality does not really hold for these contributions, their redistributive role may be considered of much lesser importance than that of the health insurance contribution.[12] Hence the distinction made here between the two types of contribution. The marginal tax rate associated with this second definition of the redistribution system that includes health insurance will be referred to as 'gross' below.

Figure 1 shows the 'net' and 'gross' effective net marginal tax rates for the sub-sample of single workers in the 1995 French Household Survey, ranked by increasing hourly wage level. Only those individuals with labor income representing 90 per cent or more of total income have been selected, to be consistent with the fact that the optimal income tax model being used refers only to labor income. Focusing on singles avoids the ambiguity mentioned before in defining productivity and labor supply for households with multiple potential earners. Marginal tax rates are computed on the basis of official rules for the calculation of taxes, health insurance contributions and non-contributory benefits, as modeled by the EUROMOD micro-simulation package. The figure also shows a continuous approximation to the relationship between net or gross marginal tax rates and individual hourly wage obtained through adaptive kernel techniques. Details on the calculation of the marginal effective tax rates and the application of kernel techniques can be found in Appendix 1 and 2. It is important to observe that there is some heterogeneity of marginal tax rates for low levels of the hourly wage rate. This heterogeneity reflects differences in non-wage characteristics of workers that affect the benefits they are entitled to - for instance their right to housing benefit and the rate of these benefits that depend on areas of residence. Once smoothened through kernel techniques, the net effective marginal tax rate function, $t(w)$, raises from 18 per cent at the lower end of the distribution to 36 per cent at the upper end, whereas the gross rate lies 15 per cent above the net marginal tax rate curve. Figure 2 shows the estimate of the density function of the distribution of hourly wage rates, $f(w)$, among single households. Two distributions are shown

[11] In effect, the health insurance system and the way it is financed maybe seen as one of the most important channel for redistribution in France - see Rochet (1996).

[12] Another reason to ignore these contributions is that the redistribution they actually achieve is technically difficult to assess, mostly because of its inter-temporal nature.

depending on whether the hourly wage is defined as net or gross of the health insurance contribution.

The solid curves in Figure 3 show the upper average marginal social welfare function *S(w)* derived from the density function, *f(w)*, shown in figure 2, its primitive, *F(w)*, and the continuous approximation of the 'net' marginal tax rate function, *t(w)*, shown in figure 1. The horizontal axis is defined in net wage percentiles. The top curve has been obtained under the assumption of a low labor supply elasticity, $\varepsilon = 0.1$, whereas the bottom one corresponds to the high elasticity value, $\varepsilon = 0.5$. Thin curves show marginal social welfare by percentile of productivity, *G'(w)*. It is derived from the solid curve through expression (9) above. As clear in this expression, the curves *G'(w)* are everywhere above *S(w)*.

It can be seen that all curves in figure 3 lack very much precision at the very top of the distribution of productivities. This is due to the fact that not enough observations are available in that part of the distribution of productivities for Kernel estimates to be reliable, despite using variable bands. Of course, the problem is more serious for the identification of the marginal social welfare since it relies on derivatives of Kernel estimates. In any case, it must be kept in mind that strong conclusions derived from the shape of the curves in the last two or three percentiles of the distribution of productivities may not be warranted.

Focusing on the upper average marginal welfare function, *S(w)*, Figure 3 shows that it is consistent with: a) marginal social welfare being everywhere positive, b) declining with income (or productivity levels); c) and declining faster with the high elasticity of labor supply. These features are fully consistent with the idea of a French redistribution authority that would be maximizing a well-behaved - i.e. increasing and concave - social welfare function. That the function is more concave when the labor supply elasticity is assumed to be high is easy to understand. If the redistribution authority believes the labor supply elasticity is high and yet applies the same redistribution schedule as when it believes it is low, it means that its preferences have changed. It values more redistribution than before since it is willing to accept that the same redistribution schedule lead to a bigger loss of output. All these features are indeed observed on the marginal social welfare curves, G'(w) shown in Figure 3. Note, however, that the preceding conclusions do not necessarily apply at the very top of the distribution where marginal social welfare appears to become negative. As mentioned earlier, however, this result has little significance because referring to that part of the distribution where Kernel estimates have limited validity.

If Figure 3 is consistent with a net redistribution system that would maximize a well behaved social welfare function, Figure 4 suggests that this is not the case any more when introducing health insurance in the redistribution system ('gross' marginal tax rates). The upper average marginal social welfare in Figure 4 is declining again and its slope is more pronounced with the high than the low elasticity of labor supply. The new feature is that upper average marginal social welfare becomes negative for high levels of wage and for the high value of the labor supply elasticity. Moreover, this phenomenon is significant since it occurs much before the range where imprecision is making any conclusion somewhat fragile. Indeed, it can be seen the upper average marginal social welfare S(w) becomes negative around the 92th percentile whereas imprecision affects the top 2 or 3 percentiles.

The interpretation of this conclusion is interesting. It can be enunciated in the following way. *"If the French redistribution authority believes that the elasticity of labor supply is around 0.5, then it is non-Paretian and imputes a negative marginal social welfare to people at the upper end of the distribution of productivities."* Practically, the thin bottom curve shows that marginal social welfare becomes negative for the top vintile of the population. In other words, social welfare would be directly increased by reducing the income of the richest 5 per cent of the population, the only reason why it is not optimal to reduce it further than what is presently done being the loss of tax receipts and therefore transfers to the bottom part of the distribution that this would entail.

That this conclusion depends on the prior that the redistribution authority has about a reasonably high elasticity of labor supply must be underscored. The upper curves in Figure 4 show that the redistribution authority would behave in a fully Paretian way if it believed that the elasticity of the labor supply would be as low as 0.1, rather than 0.5. The preceding conclusion might thus be reformulated as follows: *"the French redistribution authority is either non-Paretian or persuaded that the elasticity of labor supply is low enough for relatively high marginal tax rates to be optimal in the upper range of the distribution"*. In other words, in conformity with Proposition 2, we see that there exists a threshold for the elasticity of labor supply such that the redistribution authority is Paretian at all levels below that threshold. An interesting feature of the inversion methodology shown in the present paper is that it permits identifying that threshold. In the present case, a trial and error procedure showed the threshold was around 0.35 when using gross marginal tax rates and 0.75 when using net marginal tax rates. Conversely, if the elasticity is above that threshold, then the redistribution authority imputes a negative marginal social welfare to the upper end of the distribution of productivities. The higher is the elasticity the broader is the range where marginal social welfare is negative.

Figures 5 to 9 may be used to check whether the preceding conclusions still hold when modifying the way in which the distribution of productivities is being estimated and when the universe of income recipients is modified. Figure 5 shows the distribution of productivities obtained on single workers by inverting the basic labor supply model used throughout this paper with the appropriate wage elasticity - see (14) above. The interest of this procedure is to yield a distribution of productivities which is fully consistent with the method used to recover the social welfare function that makes the observed marginal tax rate schedule optimal, rather than the distribution of hourly wage rates. Of course, the distribution of productivities consistent with the observed distribution of total labor incomes depends on the labor supply elasticity being used. Productivities are distributed less equally when the elasticity is low. Figure 6 shows the resulting estimates of the upper average marginal social welfare and marginal social welfare for low and high elasticity. The shape of these curves is the same as before with the upper average marginal social welfare becoming negative still around the 92th percentile when the elasticity of labor supply is high.

Figures 7-9 apply the same technique to all households whose labor income represents 90 per cent or more of total income. Household of different size are being made comparable by deflating gross labor income by the number of adults at working age in the household. This makes the implicit productivity, w, derived from the inversion formula (14), a sort of average productivity among household individual members. Figure 7 shows the distribution of marginal tax rates among households ranked by productivity whereas figure 8 show the productivity distribution under the two same arbitrary assumptions about the elasticity of (household) labor supply as before. Finally, figure 9 shows the upper average marginal social welfare curve (solid curves) and the corresponding marginal social welfare curves (thin curves). All these operations are done using the 'gross' definition of marginal tax rates.

The same features as in the case of singles may be observed. Marginal social welfare is positive and declining everywhere for the low elasticity of labor supply. It is decreasing, with a steeper slope, for the high elasticity, but it is also negative in the upper part of the distribution. Moreover, the upper average marginal social welfare becomes negative practically at the limit of the 9^{th} decile, slightly sooner than for singles.

That the same features as for singles are obtained for all households is interesting for various reasons. As labor supply is certainly much more elastic at the level of the household than for single individuals, the issue of what elasticity is the most reasonable one arises with much more strength. In particular, it could make sense to assimilate the household elasticity of labor supply to the individual elasticity of so-called secondary household members – spouses, young children, heads close to retirement. The value $\varepsilon = 0.5$ would thus be more likely than $\varepsilon = 0.1$. On the other hand, it must be stressed that the treatment of household size in the optimal redistribution model is totally ignored, even though it is certainly responsible for differing marginal tax rates of households with the same total labor income per member at working age. To circumvent this problem, an alternative would be to run the inverse optimal taxation model on samples of households with comparable composition – i.e. couples without children, couples with 1 child, etc… When doing so, it is reassuring that the same result obtains, namely negative upper average marginal social welfare in the upper range of productivities[13].

To conclude, it may be worth comparing the preceding conclusions to previous empirical application of the optimal income taxation model to French data. In those direct applications of Mirrlees model using individual wage rates as a measure of productivity – Saez (2001), d'Autume (2001), Salanié (1998), Piketty (1997), - it was found that optimal marginal tax rates had a U-shape (as in figure 7) with the right-hand end marginal tax rate comparable to rates actually observed in the French redistribution system. In those models, the redistributive authority was maximizing a well-behaved social welfare functions. Under these conditions, why is it found here that observed marginal tax-rates for the top of the income distribution may not always be consistent with a Paretian social welfare function?

The answer to the preceding question relies essentially on the assumptions that are made about the distribution of productivities at the upper end of the distribution. Because of the lack of observations in that part of the income range – or more exactly the distance at which top observations are from each other – it is extremely difficult to obtain satisfactory continuous approximations of the distribution. A very common assumption consists of assuming that the distribution can be approximated there by a Pareto. For example, Piketty (1997, 2001) makes that assumption for the very top incomes of the French distribution and finds that the best fit is offered by a Pareto with coefficient $a = 2.1$. With such a value, the condition for the Paretianity of revealed social preferences as given by (11) is that:

$$t(w) \leq \frac{1 + 1/\varepsilon}{1 + 1/\varepsilon + a} = t^*$$

With ε = 0.5, the maximum value of the marginal tax rate compatible with Paretianity is $3/(3+a)$. With $a = 2.1$, the maximum marginal tax rate thus is 58.8%, a value that is indeed slightly above the maximum gross marginal tax rate observed in the case of France which in our sample turns out to be 57 per cent.

It is thus sufficient that the top of the distribution roughly corresponds to a Pareto distribution with coefficient equal or larger than 2.263, for the social welfare function revealed by existing marginal tax rates be non-Paretian. It turns out that the Pareto coefficient estimated for the top part of the distribution in our sample is superior to 2.5 whatever the percentile at which the original distribution is replaced by a Pareto. The same result was obtained using a more numerous and more

[13] These results are available upon request.

precise survey on French incomes: fitting a Pareto distribution to the top of the income distribution observed in the "Survey on the Fiscal Incomes 1996" yielded values of *a* in the interval [2.9-3.2][14].

Estimating the shape of the distribution of the productivity will always be difficult and imprecise at the very top of the distribution. However the important feature of the previous results is that the negativity of S(w) occurs much below the range in which the density and the cumulative of the productivity distribution are imperfectly known. This is either because the Pareto shape does not fit that part of the distribution, or because the parameter *a* is much larger than 2.263 there.

4. Income Effects

All the preceding results are based on the assumption that individual preferences were quasi-linear in consumption and iso-elastic in labor supply. With this specification the income effect on labor supply is simply nil. A slightly less restrictive form consists of assuming separability of consumption and labor 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:

$$\frac{t(w)}{1-t(w)}\psi[y(w)] = (1+\frac{1}{\varepsilon}).\frac{1-F(w)}{w.f(w)}.[\overline{\psi}[y(w)] - S(w)/S(w_0)] \qquad (15)$$

where $\psi[y(w)] = \dfrac{1}{A'(y)}$ i.e the inverse of the marginal utility of income, and $\overline{\psi}[y(w)] = \int_w^{w\max} \dfrac{1}{A'(y)} f(x)dx$ i.e. the mean value of that inverted marginal utility for people with productivity above w.

From equation (15) we can recover the equivalent of proposition 1 that is:

Proposition 4. *A necessary condition for the social welfare function (making the observed effective marginal tax rate schedule, t(w), optimal with respect to the observed distribution of productivities, f(w)-when the individual utility function is separable in consumption and labor) to be Paretian is that :*

[14] In French: "Enquête sur les Revenus Fiscaux 1996". We thank Pascal Chevalier and Alexandre Baclet from the French National Statistic Institute (INSEE) for accepting to perform the estimations and for providing us with these figures.

$$t(w) \leq \frac{\frac{1+\varepsilon}{\varepsilon}\frac{1-F(w)}{w.f(w)}\frac{\overline{\psi}[y(w)]}{\psi[y(w)]}}{1+\frac{1+\varepsilon}{\varepsilon}\frac{1-F(w)}{w.f(w)}\frac{\overline{\psi}[y(w)]}{\psi[y(w)]}} \quad for\ all\ w \in [w_0, Z] \qquad (16)$$

that is the equivalent of (10) when income effects are considered.

Recall that condition (10) is

$$t(w) \leq \frac{\frac{1+\varepsilon}{\varepsilon}\frac{1-F(w)}{w.f(w)}}{1+\frac{1+\varepsilon}{\varepsilon}\frac{1-F(w)}{w.f(w)}} \quad for\ all\ w \in [w_0, Z]$$

By comparing this last two conditions we can easily see that, as $\frac{\overline{\psi}[y(w)]}{\psi[y(w)]} \geq 1$, the right hand side of (10) is always smaller that the correspondent term in (16). This implies that the inclusion of income effect mitigate the possibility to be Non Paretian.

To see it empirically, we have computed the values of S(w) correspondent to the same inversion technique as above using the following analytical specification for *U(y, L)*:

$$U(y,L) = \frac{y^{1-\frac{1}{\alpha}}}{1-\frac{1}{\alpha}} - k\frac{L^{1+\frac{1}{\beta}}}{1+\frac{1}{\beta}} \qquad (17)$$

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

Results obtained with this specification are presented in Figure 10 for the sample of French singles, using the gross wages as a proxy of the productivities, with the following two sets of parameters values ($\alpha = 2$, $\beta = 2$) and ($\alpha = 5$, $\beta = 5/7$) both leading to $\varepsilon = 0.5$ but with different marginal utilities of income[15].

It may be seen that, in both cases, the upper incomplete mean marginal social welfare, *S(w)*, became negative only beyond centile 95.

These empirical results show that the non-Paretian nature of the social welfare function in presence of a medium value for the elasticity of labor supply is influenced by the presence of income effects.

5. 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

[15] To compute the terms $\psi(y)$ and $\overline{\psi}(y)$ in formula (19) we used the observed disposable income as proxy for the optimal consumption *y*.

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 labor supply.

The detailed analysis performed on France shows that revealed marginal social welfare curves are in agreement with standard theory when the elasticity of labor supply is assumed to be low and when the redistribution system excludes the health insurance contributions. Marginal social welfare then is both positive and decreasing throughout the range of individual productivities. However, marginal social welfare turns out to be negative at the very top of the distribution when the labor supply elasticity is assumed to be around the average of estimates available for secondary workers in the literature, and the health insurance contribution is included in the redistribution system.

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 Paretian social welfare function. It was found in this paper that its behavior could be of three different types. Either the redistribution authority under-estimates the labor supply response to taxation, or it has non-Paretian social preferences, or it does not optimize at all (a Paretian social welfare function) and there is space for a Pareto improving tax reform. 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 might result more from political economy forces than from 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.

Appendix 1. Technicalities

Equation (7) yields the basic principle of the inversion methodology. Its actual implementation raises additional complications, however. They are listed below together with the choices made to overcome them.

a) Continuity and differentiability

The application of the inverted optimal taxation formulae, (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[16]. 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[17].

(ii) Estimate the derivatives of $t(w)$ and $f(w)$ using again a kernel approximation computed over the whole sample.[18]

(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 (7).

b) Households with zero income and households with apparently irrational behavior

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 pats 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

---

[16] 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).

[17] Occupational status and home ownership are other sources of heterogeneity with respect to the tax system.

[18] For technical details about the computations of kernel derivatives see Pagan and Ullah (1999, pag. 164).

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: $y = RMI + .25* wL$ if this person qualifies for the RMI –i.e. $wL<RMI$. But it is simply: $y = wL$ if $wL > RMI$[19]. This budget constraint is clearly convex. Therefore, there should be a range of labor incomes around the RMI where it would be irrational to be[20]. 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, 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).

Appendix 2. Microsimulation model, dataset and data treatment.

The samples and the micro simulation model were taken from EUROMOD, a project whose objective is to build an integrated micro simulation model for the 15 countries of the European Community. A complete and detailed description of the EUROMOD micro simulation model as well as the datasets is contained in Sutherland (2001).

The version of the model used in this paper is a prototype replicating the laws enforced in 1995 in France. All the modules replicate social contributions levied on wages (for employers and employees) and on self-employed workers; social contributions on other types of income (unemployment benefits, income from pensions and capital return); income taxes; family benefits and social assistance mechanisms. The datasets used for France are the 1995 Households Budget Survey of INSEE.

The micro simulation model has been used in order to compute the effective marginal tax rate for each household. This variable was obviously not present in the survey and it was therefore necessary to compute it. The definition of effective marginal tax rate used was the derivative, in each point, of the budget constraint. A possible method of calculation consists of the assignment of a lump-sum amount of gross income to each household (in our case the equivalent of 5000 French

[19] All other benefits that may complement the RMI are ignored in this argument, but they are taken into account in the calculations made below.

[20] 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.

francs per household per year) and, in the computation with the micro-simulation model, of a new distribution of disposable incomes. The effective marginal rate of taxation is thus obtained from the formula:

$$emtr = \frac{\Delta Taxes + \Delta Benefits}{\Delta Gross\ \mathrm{I}ncome} = 1 - \frac{\Delta Yd}{\Delta y} \qquad (18)$$

To keep with the logic of the optimal taxation model, all households with zero income and with non-labor income, including pension and unemployment benefits above 10 per cent of total income were eliminated from the sample. The final samples used in the paper contain 5527 households (on a total of 10214), 963 of which are singles.

Appendix 3. More on the consistency of a redistribution system with optimizing a well-behaved social welfare function

Going further in checking the consistency of a redistribution system with optimizing a well-behaved social welfare function requires studying the concavity of the revealed social welfare function. We have that:

Proposition 5*. A necessary and sufficient condition for the concavity of the social welfare function is that:*

$$v(w)*\left[\eta(w)+2+\sigma_t+\frac{2t(w)v(w)}{1-t(w)}\right]+\left[1-t(w)\right]*\left[\eta(w)(1+\sigma_f)-\eta(w)^2\right]\leq 0 \qquad (19)$$

where $w\dfrac{f''(w)}{f'(w)}=\sigma_f(w)$ is the curvature of *f(w)* and $w\dfrac{t''(w)}{t'(w)}=\sigma_t(w)$ the curvature of *t(w).*

The proof of that proposition is established by taking the derivative of (9).

The expression (19) it is not easy to interpret because of its complexity and the fact that it relies on the second derivatives of both the density function and the marginal tax rate schedule. In the particular case in which *f(w)* is distributed as a Pareto with parameter *a*, it is easily shown that $\eta(w)=-1-a$ and $\sigma_f=-2-a$. It follows that the term $\eta(w)(1+\sigma_f)-\eta(w)^2$ vanishes, leading to the following property.

Proposition 6*. Wherever the productivity distribution function f(w) is a Pareto distribution of parameter a, a necessary and sufficient condition for social welfare to be concave is that the marginal tax rate function follow a path such that:*

$$v(w)^2 \frac{2t(w)}{1-t(w)} + v(w)\left[1 - a + \sigma_t\right] \leq 0 \qquad (20)$$

Again, propositions 5 and 6 are of relevance in connection with the discussion on whether the marginal tax rate curve must be U-shaped – see Diamond (1998) and Saez (2001). Notice that *v(w)* is positive when the marginal tax rate is increasing, that is for high incomes. Then, equation (20) reduces to:

$$v(w)\frac{2t(w)}{1-t(w)} + \sigma_t \leq a - 1 \qquad (21)$$

The preceding equation tells us that if the curvature of the observed marginal tax rate is highly positive, the revealed social welfare function would not be concave, which makes the whole optimum inverse procedure irrelevant. This result extends the one obtained by Diamond (1998) because it restraints the class of U shaped marginal tax rates (consistent with an increasing average tax rate) consistent with a concave social welfare function.

Figure 1. Kernel smoothed marginal tax rates for singles: net and gross scenarios

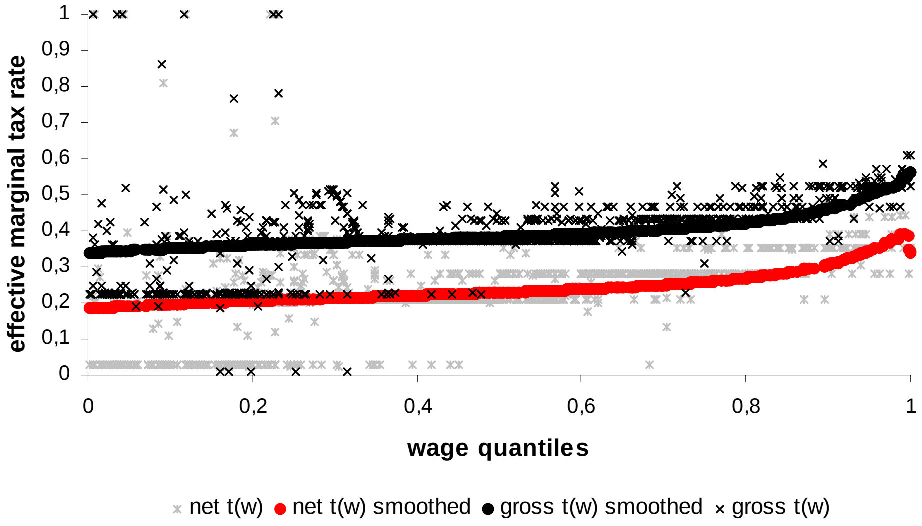


Figure 2. Kernel wage densities for singles: net and gross scenario

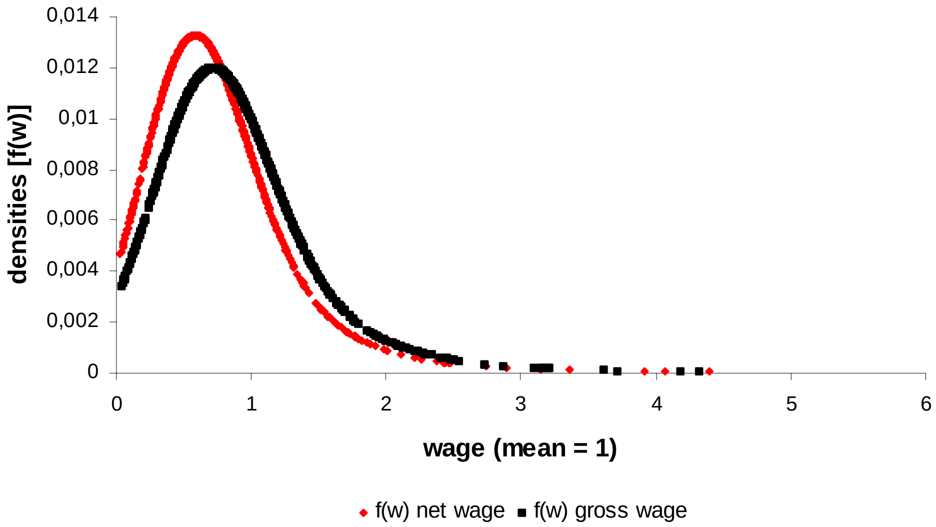


Figure 3. Social marginal welfare for singles (on net wages)

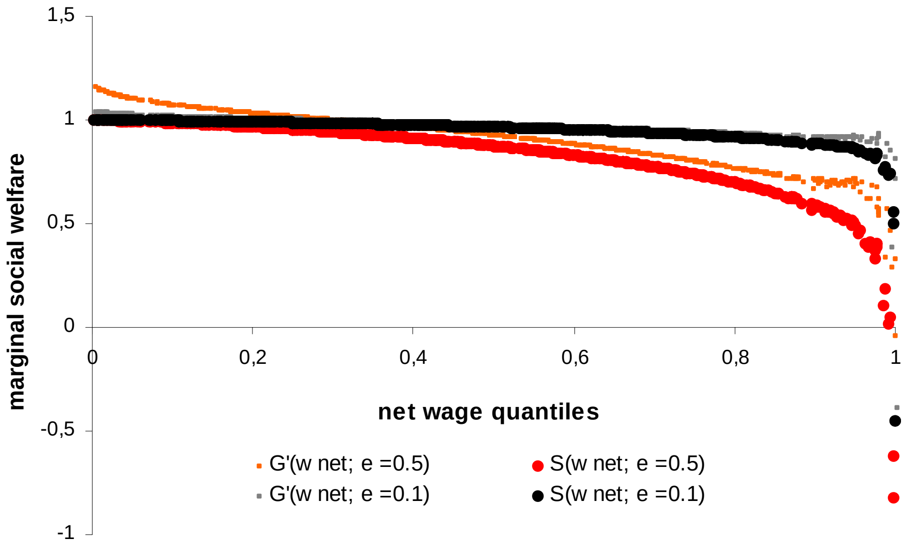


Figure 4. Social marginal welfare for singles (on gross wages)

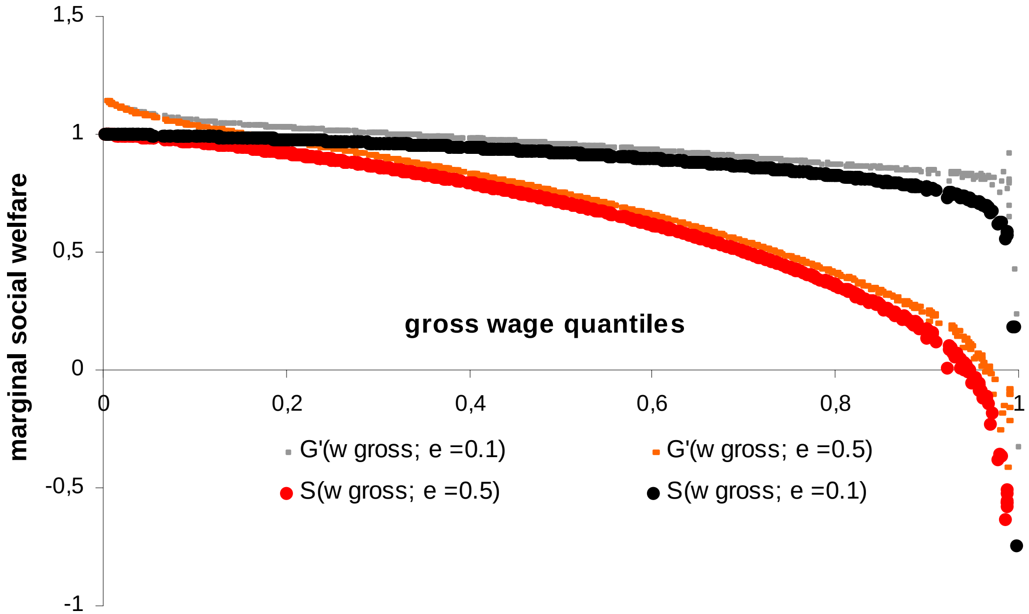


Figure 5. Kernel productivity densities for singles

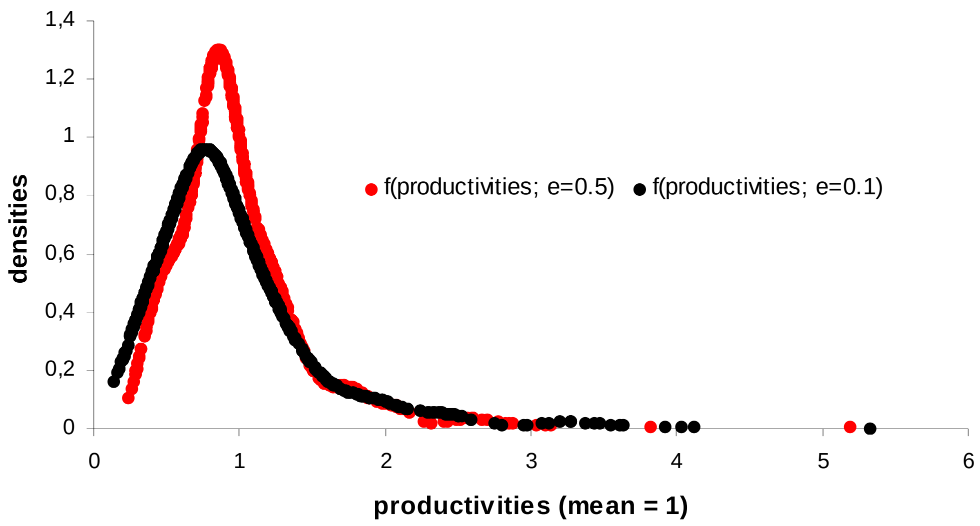


Figure 6. Social marginal welfare for singles (on productivities)

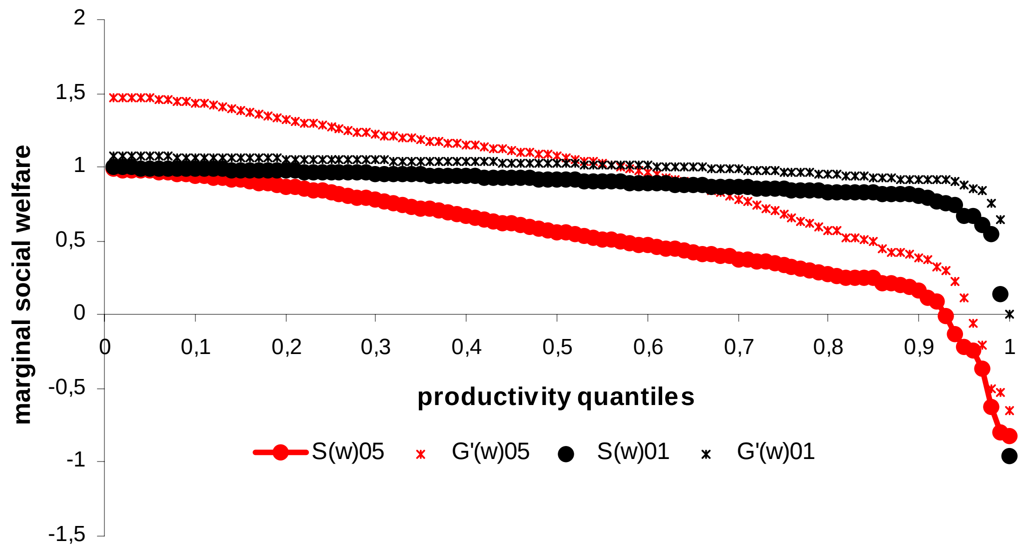


Figure 7. Gross kernel smoothed marginal tax rates for all households

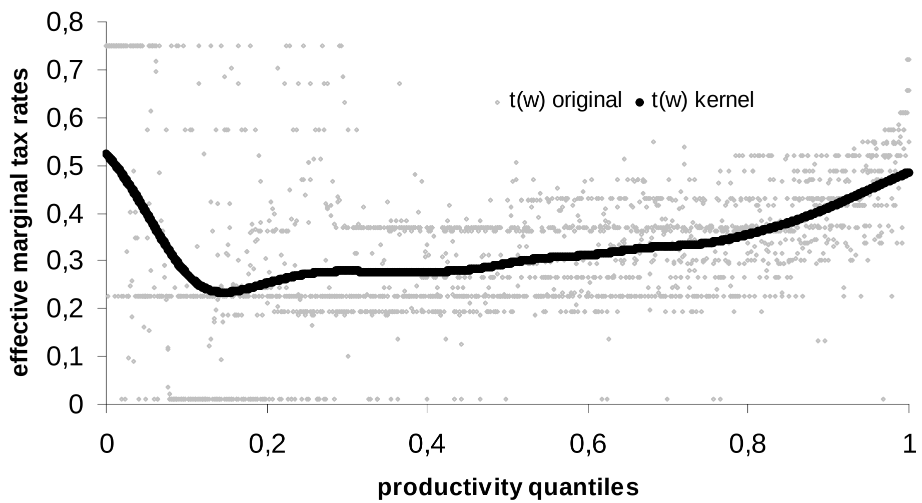


Figure 8. Kernel productivity densities for all households

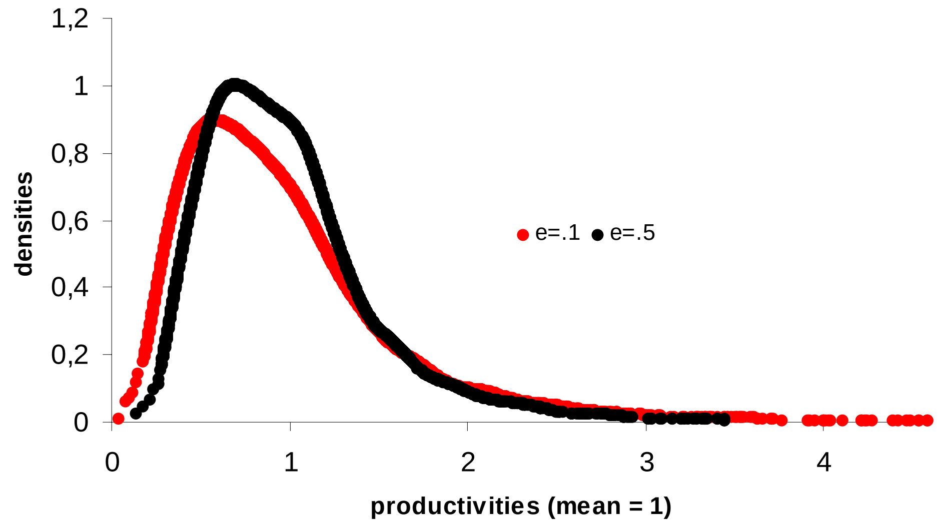


Figure 9. Social marginal welfare for all household (on productivities)

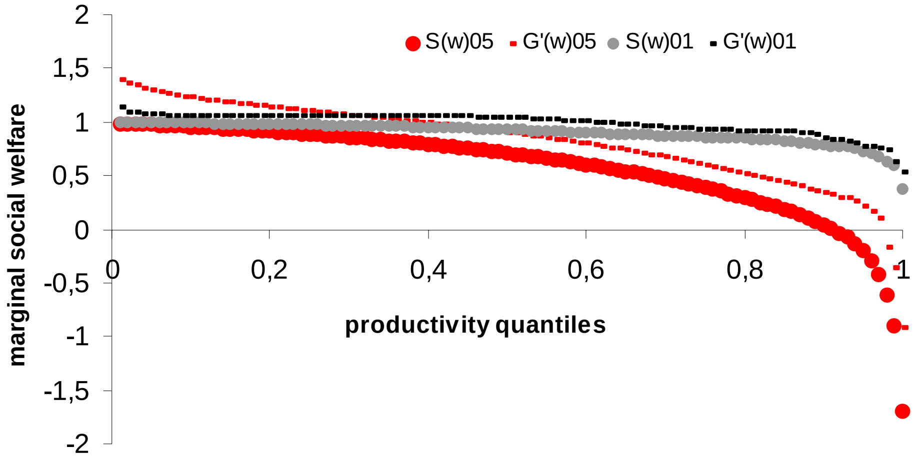
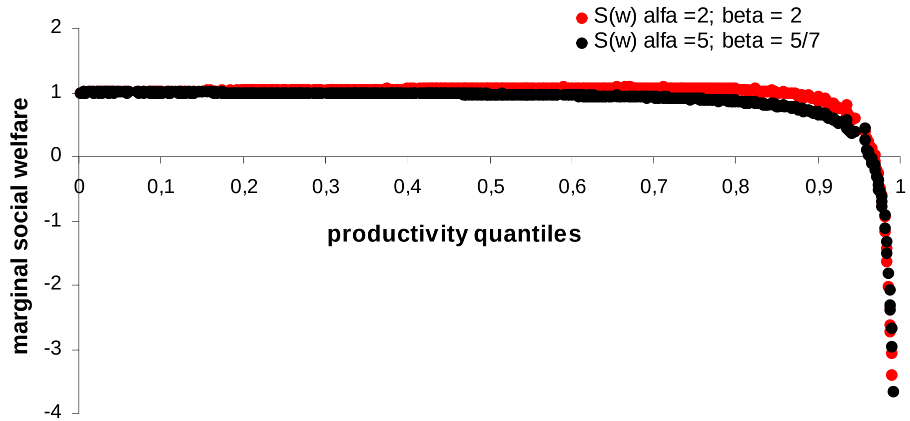


Figure 10. Paretianity test on social marginal welfare for singles (on gross wages) with income effects



References

AHMAD E, N. STERN, 1984, “The Theory of Reform and Indian Direct Taxes”, *Journal of Public Economics* 25: 259-298.

ATKINSON A.B., 1995, "*Public Economics in Action: Basic Income-Flat Tax Proposal*” Clarendon Press Oxford.

ATKINSON A.B., J. STIGLITZ 1980, "*Lectures on Public Economics*” McGraw Hill International Editions.

AUTEN G., R. CARROLL, 1999, “The Effect of Income Taxes on Household Income”, *Review of Economic and Statistic* 81: 681-693.

BARGAIN O., 2004 “*On Modelling Household Labour Supply With Taxation*”, DELTA mimeo.

BLUNDELL R. , A. DUNCAN, C. MEGHIR, 1998, “Estimating Labor Supply Responses to Tax Reform”, *Econometrica* 66: 827-861.

BLUNDELL R., MACURDY, 1999, “Labor Supply: a Review of Alternative Approaches”, *Handbook of Labor Economics vol 3a*. North Holland.

BOURGUIGNON F., T. MAGNAC, 1990, “Labour Supply and Taxation in France”, *Journal of Human Resources*, 25(3).

BOURGUIGNON F, A. SPADARO, 2000, “Redistribution et incitations au travail: une application simple de la théorie de la fiscalité optimale”, *Revue Economique*, 51(3): 473-488.

CHONÉ P., G. LAROQUE, 2005, “Optimal Incentives for Labor Force Participation”, *Journal of Public Economics*, 89(2-3): 395-425

CHONÉ P., D. LE BLANC, I. ROBERT-BOBÉE, 2003, “Female labour Supply and Child Care in France”, *CREST Working Paper* nº 2003-4, INSEE Paris

D’AUTUME A., 2001, “L’imposition optimale du revenu: une application au cas français” *Revue Française d’Economie*, 3: 3-63

DIAMOND P., 1998, “Optimal Income Taxation: An Example with U-Shaped Pattern of Optimal Marginal Tax Rate”, *American Economic Review*, 88(1): 83-95

DONNI O., 2000, “*Essais sur les modèles collectifs de comportement du menage*” Ph.D. Dissertation EHESS, Paris.

FELDSTEIN M., 1995, “The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act”, *Journal of Political Economy*, 103: 551-572.

GRUBER J., E. SAEZ, 2002, “The Elasticity of Taxable Income: Evidence and Implications”, *Journal of Public Economics*, 84: 1-32.

HÄRDLE W., 1990 "*Applied nonparametric regression*" Cambridge University Press.

HAUSMAN J., 1985, “The Econometrics of Nonlinear Budget Set”, *Econometrica*, 53: 1255-1282.

KURZ M., 1968, "On the Inverse Optimal Problem," in *Mathematical Systems Theory and Economics*, edited by H. W. Kuhn and G. P. Szego, Springer-Verlag (1968); 189-202.

LAROQUE G., SALANIE B., (2002) “*Institutions et emploi, le marché du travail des femmes en France*” Economica.

MIRRLEES J.A., 1971, “An Exploration in the Theory of Optimum Income Taxation”, *Review of Economic Studies*, 38: 175-208.

PAGAN A., A. ULLAH, 1999 “*Nonparametric Econometrics”,* Cambridge University Press.

PIKETTY T., 1997, “La redistribution fiscale contre le chômage*”*, *Revue Française d’Economie*, 12(1) : 157-203.

PIKETTY T., 1998, “L’impact des incitations financières au travail sur les comportements individuels : un estimation pour le cas français*”*, *Economie et Prévision*, 132-133: 1-35.

PIKETTY T., 2001, “*Les hauts revenus en France au 20ème siècle – Inégalités et redistributions, 1901-1998*”, éditions Grasset.

ROCHET J.C., 1996, "Les Atouts et les Limites des Systèmes Publics d'Assurance Maladie", *Revue Française d'Economie,* 11(1) : 183-189.

SAEZ E., 2001, “Using Elasticities to Derive Optimal Income Tax Rates”, *Review of Economic Studies,* 68: 205-229

SALANIE B., 1998, "Note sur la Taxation Optimale", *Rapport au Conseil d'Analyse Economique*, La Documentation Française, Paris
SEADE J., 1977, "On the Shape of the Optimal Tax Schedule", *Journal of Public Economics,* 7: 203-236
SEADE J., 1982, "On the Sign of Optimum Marginal Income Taxation", *Review of Economic Studies*, 49: 637-643.
SUTHERLAND H., 2001 "Final Report EUROMOD: An Integrated European Benefit-Tax Model". *EUROMOD Working Paper* nº EM9/01.
TUOMALA M., 1990, "*Optimal Income Tax and Redistribution*", Oxford University Press.