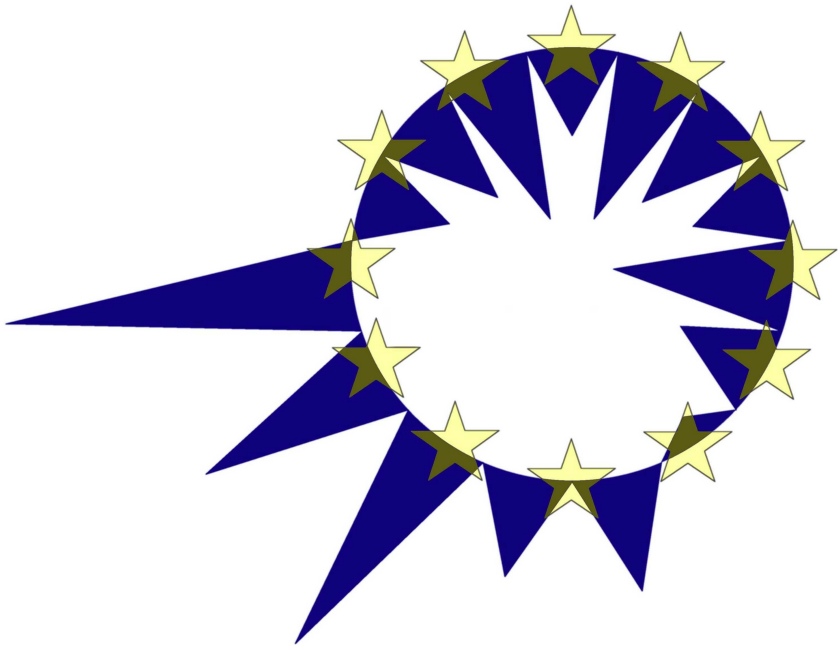
EUROMOD

WORKING PAPER SERIES



EUROMOD Working Paper No. EM9/08

TAX-BENEFIT REVEALED SOCIAL PREFERENCES

François Bourguignon
Amedeo Spadaro

October 2008

Tax-Benefit Revealed Social Preferences*

François Bourguignon

Paris School of Economics

Amedeo Spadaro

Paris School of Economics and Universitat de les Illes Balears

* This paper uses EUROMOD version 15a. EUROMOD is continually being improved and updated and the results presented here represent the best available at the time of writing. Any remaining errors, results produced, interpretations or views presented are the authors' responsibility. EUROMOD relies on micro-data from twelve different sources for fifteen countries. This paper uses data from the Enquête sur les Budgets Familiaux (EBF) made available by INSEE. INSEE does not bear any responsibility for the analysis or interpretation of the data reported here.

TAX-BENEFIT REVEALED SOCIAL PREFERENCES

François Bourguignon
Amedeo Spadaro[1]

Abstract

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 may appear, under some plausible scenario concerning the size of the labor supply behavioral reactions, non Paretian (e.g. giving negative marginal social weights to the richest class of tax payers).

JEL Classification: H11, H21, D63, C63

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

Corresponding author:

Amedeo Spadaro
Paris School of Economics
48 Boulevard Jourdan
75014 Paris
France
Email: spadaro@pse.ens.fr

[1] This is a complete revised version of the paper "Tax-Benefit Revealed Social Preferences". PSE Working Paper nº 2005-22. We thank Tony Atkinson, Salvador Balle, Roger Guesnerie, Jim Mirrlees, Emmanuel Saez, two anonymous referees and the participants to seminars in Barcelona, Madrid, Berlin, Paris, Venezia and Formentera for useful comments. We also thank Pascal Chevalier and Alexandre Baclet from INSEE to help us with the French Fiscal Data and are indebted to all past and current members of the EUROMOD consortium for the construction and development of EUROMOD. We are solely responsible for any remaining error. Amedeo Spadaro acknowledges financial support of Spanish Government (SEJ2005-08783-C04-03/ECON) and of French Government (ANR BLAN06-2_139446).

1. 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 seem more relevant now 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 approach in this paper is the opposite of this standard approach. The focus is on the social welfare function that makes optimal the actual 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', and in particular whether the marginal social welfare is everywhere decreasing (ensuring the concavity of the social objective) and positive. If the first condition does not hold, then it is the whole optimization concept behind Mirrlees' framework that would become doubtful. It would indeed be very difficult to assume that the redistribution authority attempts to maximize a non-concave welfare function if other than trivial redistributions policies are observed. If, on the contrary, the second condition fails, then the revealed social welfare function may not be deemed to be Paretian.

The method proposed here provides a kind of new 'reading' of the effective average and marginal tax curves that are commonly used to describe a redistribution system. It translates the observed shape of these curves into a social welfare function. Comparing two redistribution systems is cast in terms of the social welfare which would make them optimal. Instead of analyzing who is getting more out of redistribution and who is getting less, or the way work incentives are distorted, the marginal tax rate schedule can be made to inform directly on the differential implicit marginal social welfare weight given to one part of the distribution versus another.

These 'revealed social preferences' necessarily rely on auxiliary assumptions about labor supply behavior and the distribution of individual abilities. With the direct or standard approach to optimal taxation, the optimal tax schedule is known to be very sensitive to these assumptions. The same is true of the social preferences revealed by a given marginal tax schedule. If revealed preferences are really odd, this may be because some common assumptions on labor supply behavior or on the distribution of abilities are inconsistent, which should be 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). 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.

In a more recent theoretical paper, Choné and Laroque (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 labor supply behavior - that makes an observed 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. Unlike 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 Mirrlees approach builds on a labor supply model which only focuses on hours-of-work responses. A labor supply model, incorporating labor market participation responses as well as the choice of hours, may provide more realistic results in optimal income taxation, as first shown by Saez (2002). Accordingly, the present paper also shows the results of the inversion of an optimal tax problem à la Saez (2002) in which extensive and intensive labor supply behaviors are explicitly taken into account.

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 labor supply elasticities, which allow deriving the distribution of individual abilities from observed labor incomes. Section 4 extends the analysis to the case of non-zero income-elasticity of labor supply. Section 5 analyzes the case where labor supply is discrete (as in Saez 2002). Section 6 concludes.

[2] See Sutherland (2001).

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, under some plausible assumption on the labor supply elasticity, the redistribution authority appears to be non Paretian (e.g. giving negative social weights to some class of tax payers). The inclusion of labor market participation responses as well as the choice of hours behaviors confirms that high marginal tax rates are compatible with the maximization of a Paretian social welfare function only if the labor supply elasticities are low.

2. 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 (*c*) /labor (*L*) combination that maximizes their preferences, *U(c, L)*, given the budget constraint imposed by the government: $c = wL - T(wL)$, where *w* is the productivity, taken to also be the wage, 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:

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

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

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

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

where $\overline{T}$ is the budget constraint of the government and *G[.]* is the function that transforms individual utility, *V(.)*, into social welfare. Somewhat improperly, this function will be referred to as the 'social welfare function' in what follows. The main argument in this paper is based on the special case where the function *U(c, L)* is quasi-linear with respect to *c* and isoelastic with respect to *L*, a case extensively used in both the theoretical and applied optimal

tax literature[4]. Formally, the utility function writes: $U(c, L) = c - B(L)$ with

$$B(L) = \frac{\varepsilon}{1+\varepsilon} L^{\frac{1+\varepsilon}{\varepsilon}} \quad (2)$$

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

It can be shown – see for instance Atkinson and Stiglitz (1980) or Atkinson (1995) - that this particular case leads to the following simple characterization of the optimal tax schedule:

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

In that expression, *t(y)* is the (optimal) marginal tax rate faced by an agent with productivity, *w*, and therefore with (gross) earnings *y=wL** -i.e. *t(y) = 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:

$$S(w) = \frac{1}{\left[1-F(w)\right]} \int_w^Z \frac{G'\left[V(x, T(xL)\right]}{\lambda} f(x)dx \quad (5)$$

where λ is the Lagrange multiplier associated with the constraint 1.4.

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.

---

[3] See for instance Atkinson and Stiglitz (1980) or Tuomala (1990).

[4] See in particular Atkinson (1995) or Diamond (1998).

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(Y)*, the elasticity of labor supply, ε, the 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]: $TR = \frac{t(Y)}{1-t(Y)} \frac{W.f(W)}{1+1/\varepsilon} dT$. 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 $S(w_0)(TI - TR)$. 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.

---

[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.

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-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[V(w)]*, implied by a tax-benefit system under the assumption that it is optimal requires some additional steps. Equations (4) or (6) can be simply rewritten as:

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

Identifying the marginal social welfare functions, *G'(.)*, itself requires an additional step. Differentiating (7) and using the definition of *S(w)* in (5) yields:

$$\frac{G'[V(w,T(y)]}{\lambda} = 1 + \left(\frac{\varepsilon}{1+\varepsilon}\right)\left(\frac{t(y)}{1-t(y)}\right)\left[1 + \eta(w) + \nu(y)\frac{1+\varepsilon}{1-t(y)+\varepsilon\nu(y)t(y)}\right] \qquad (8)$$

where $\eta(w) = wf'(w)/f(w)$ is the elasticity of the density and $\nu(y) = yt'(y)/t(y)$ that of the marginal tax rates with respect to labor income *y*. Putting (7) and (8) together, it can be seen that the function *S(w)*, is the "upper average marginal social welfare" (UAMSW) of people with productivity equal or greater than *w* may thus be recovered from the knowledge of

[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(y)*, *f(w)* and ε. Choné and Laroque (2005) solve a

primary data, that is the marginal tax rate schedule, *t(y)* the elasticity ε and the distribution of abilities, i.e. *f(w)* and *F(w).* Recovering *G'(.)* itself requires information on the derivatives of *t(y)* and *f(w).* Because of this, the estimate that can be empirically obtained of *S(w)* is likely to be much more robust than that of *G'(.).* Most of the empirical application in this paper will thus mostly be based on UAMSW rather than marginal welfare.

All the previous results are based on the hypothesis that the observed marginal tax rate is the result of maximization of a social welfare function under the budget and the incentive compatibility constraints. This assumption imposes several restrictions on the shape of the observed marginal tax rate. If they are not satisfied, then the whole inversion procedure becomes inconsistent. In Appendix 1 we analyze them in details. If one of the conditions in the appendix does not hold, then it is the whole optimization concept behind Mirrlees framework that would become doubtful. For instance, it would be difficult to assume that the redistribution authority attempts to maximize a non-concave welfare function if other than trivial redistribution policies are observed[7].

Let us now derive a few consequences of the optimal inverse framework for the Paretianity of the revealed social preferences.

Definition of Paretian Social Welfare Function: *A SWF is sad to be Paretian if G'(V)≥ 0 everywhere. It is Non Paretian otherwise.*

Proposition 1*. A necessary condition for the social welfare function, G(.), that makes the observed effective marginal tax rate schedule, t(y), optimal with respect to the observed distribution of productivities, f(w) to be Paretian is that:*

$$t(y) \le \frac{\frac{1+\varepsilon}{\varepsilon}\frac{1-F(w)}{w.f(w)}}{1+\frac{1+\varepsilon}{\varepsilon}\frac{1-F(w)}{w.f(w)}} \quad \text{for all } w \in [w_0, Z] \qquad (9)$$

symmetric problem by identifying the pair (*f(w), ε*) knowing *t(y)* and *S(w).*

[7] Of course, from a mathematical point of view we cannot completely rule out a maximizing behavior. The point is that we are not able to characterize it.

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 (9) 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 (8) of marginal social welfare.

The Paretian condition given in proposition 1 can be also reinterpreted as a test on the relative position of the tax schedules with respect to the "Laffer bound". This bound is defined as the revenue maximizing or efficiency cost minimizing tax system [see Canto, Joines and Laffer (1982) and Laroque (2005)], and it is precisely the right hand side of (9). If marginal tax rates are not below the Laffer bound, then observed the tax system can be optimal only with non Paretian social preferences.

Interestingly, where the ability distribution *f(w)* 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 $\frac{w.f(w)}{1-F(w)} = a$, (9) is equivalent to: $t(y) \leq \frac{1+1/\varepsilon}{1+1/\varepsilon+a}$

(10)

For instance, with not unreasonable values like $a = 3$ and $\varepsilon = 0.5$, this condition states that a redistribution system where the effective 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, G(.), is necessarily non-decreasing everywhere.*

This proposition follows directly from (8). If indeed $\eta(w)$ and $\nu(y)$ take only finite values, the second term on the RHS of (8) can be made as small as desired in absolute value 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 and the density of the ability distribution is decreasing, a sufficient condition for the social welfare function G(.) to be locally non-decreasing is:* $$t(y) \leq \frac{1+\varepsilon}{1-\eta(w)\varepsilon} \qquad (11)$$

Again, this proposition is directly derived from (8). 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, (11) 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 (10) in the case where the productivity distribution may be approximated by a Pareto. On the other hand, condition (11) becomes a necessary condition for marginal social welfare to be non-negative when the marginal tax rate is decreasing - for instance for low incomes.

A last 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) [8]. As this is not observed in actual tax-benefit systems, there are two alternative interpretations: the first is to say that the tax-authority knows the highest wage rate but it is not pursuing the maximization of some well behaved social welfare function. The second is that tax authority is unable to identify the top wage and gives a non zero probability for the top wage to be above any arbitrary bound. In this second case, zero marginal taxation at the top becomes irrelevant. This second alternative is, in our opinion, the most plausible from an empirical point of view (as noted also by Atkinson 1985:57, Mirrlees 1976:340 and Diamond 1998). In what follows the support of *f(w)* is assumed to be large enough (with *z* tending to infinity). This is obtained by computing adaptive kernel densities for extreme high class of productivities or, as an alternative, by making the hypothesis that for the upper tails, the distribution follows a Pareto of parameter *a*.

[8] A simple proof of that property is obtained by considering the limiting case *1-F(w)=0* in the intuitive argument justifying (4) above. In that limiting case, optimization requires that *TR=0*, and therefore that *t=0*.

3. Empirical implementation issues

The previous methodology requires estimates of the elasticity of labor supply, ε, the distribution *f(w)* and the marginal rate of taxation, *t(y),* to be available. Practically, what is observed in a typical household survey? Essentially total labor income, $y=wL$, and disposable income, *c*, or by difference, total taxes net of benefits, $T(wL)$[9]. 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 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 previous work on applied optimal income tax, see Diamond, 1998, Salanié, 1998, or d'Autume, 2001. This is the first approach pursued below.

Although simple, this 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 (Blundell et al., 1998). On the other hand, relying on simpler alternative estimates based on standard 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

[9] To keep with the logic of the optimal taxation model, non-labor taxable income is ignored in all what follows.

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? 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"? 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 extremely complex econometric estimation procedure that would deal with the previous points 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:

$$Y = wL^* = k.w^{1+\varepsilon}.[1-t(wL^*)]^{\varepsilon} \qquad (12)$$

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

$$w = Y^{\frac{1}{1+\varepsilon}}[k(1-t(Y))]^{\frac{-\varepsilon}{1+\varepsilon}} \qquad (13)$$

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 (13). 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, $t(Y_i)$ and all the derivatives from which the social marginal welfare function may be inferred - see (7) and (8) above. Of course, from the household point of view, the elasticity of the household labor supply is not simply the average of the spouses' elasticies. A more appropriate measure should take into account the activity status of the household components. For example, in the typical case of a one earner household in which the spouse is potentially active, the key labor supply parameter is the participation elasticity of the spouse. In the preceding framework, that extensive elasticity actually becomes the relevant household intensive labor supply elasticity.

4. Application to the French redistribution system

As mentioned above, there is considerable imprecision about the value of labor supply elasticity, 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) and Eissa and Hoynes (2006) give 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 (2005), 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 (0.6 to 1) if they have children (Piketty 1998, Bargain, 2005, Choné et al. 2003). This second result is mainly driven by participation effects. Similar results have been obtained on the basis of the relationship between taxable incomes and changes in tax rates. In the case of France, Piketty (1999) found average elasticities of taxable income around 0.1 with participation elasticities around 0.2. 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 2 and 3 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. 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. Nevertheless, indirect taxes can be easily introduced. They would simply increase the marginal tax faced by every household.[10]. The corresponding effective 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 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 essentially due to 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] This is what justifies the distinction made here between the two

---

[10] Note, however, that the increase depends on the initial marginal tax rate. If θ is the indirect tax rate, the overall effective marginal tax rate becomes $[t(y) + \theta]/(1+\theta)$ rather than $t(y) + \theta$. See Atkinson and Stiglitz (1980, chapter 9).

[11] In effect, the health insurance system and the way it is financed may be 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.

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 (Sutherland et al., 2001). 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 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 size of these benefits that depend on areas of residence. Once smoothened through kernel techniques, the net effective marginal tax rate function, *t(y)*, 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 roughly 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 depending on whether the hourly wage is defined as net or gross of the health insurance contribution (kernel smoothing has been used). Next figures show the results of the optimal inverse procedure.

The solid curves in Figure 3 show the UAMSW 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(y)*, shown in figure 1. The horizontal axis is defined in net hourly 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'(V(w))*

(divided by the constant λ). It is derived from the solid curve through expression (8) above. The fact that the *G'(.)* curves are decreasing and above the *S(.)* curves is consistent with the hypothesis that *G(.)* is concave[13].

Focusing on the UAMSW function, *S(w)*, Figure 3 shows that it is consistent with marginal social welfare a) being everywhere positive, and b) declining with income. It can also be seen that the UAMSW curve is everywhere lower and with a higher slope when the elasticity of labor supply takes the high value, 0.5. 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 seems to be 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 it values redistribution more since it is willing to accept that the same redistribution schedule lead to and a bigger loss in total income.

If Figure 3 is consistent with a net redistribution system that would maximize a Paretian 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 UAMSW curve in Figure 4 becomes negative for high levels of wage and for the high value of the labor supply elasticity. This phenomenon is statistically significant because it occurs much before the range of wages where the scarcity of observations makes any conclusion somewhat fragile because depending on the smoothing technique being used. 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 finding is interesting and somewhat surprising. It can be enunciated in the following way. *"If the French redistribution authority anticipates an elasticity of labor supply around 0.5 or higher, then it is non-Paretian and imputes a <u>negative</u> marginal social welfare to people at the upper end of the distribution of wages."* Practically, the thin bottom curve in figure 4 shows that marginal social welfare becomes negative for the top vintile of the population. In other words, social welfare would be directly increased by <u>reducing</u> the income of the richest 5 per cent of the population. The only reason why it would not be

[13] See the discussion on condition D) in appendix 1.

optimal to reduce it further than what is presently done is the loss of tax receipts and therefore the drop in transfers to the bottom part of the distribution that this would entail. Including also indirect taxes would reinforce this non Paretianity results given that in France the VAT is levied on consumption of goods mostly at a rate of 19.6%[14].

The role of the anticipations of the redistribution authority on the 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 anticipated 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 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, or it is non Paretian"*. 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.

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 (13) 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

[14] Certain types of goods are taxed a different rates (2.1% and 5.5%) but their contribution to the total receipts is really small.

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 (13), 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 UAMSW 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.

Obtaining the same features for singles and 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, the issue of what elasticity is the most reasonable one arises with much more strength. In particular, as discussed at the end of section 2, 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 individual participation elasticity of the secondary earner becomes then the household intensive elasticity (given that there is always a first earner working full time). The value $\varepsilon = 0.5$ would thus be more likely than $\varepsilon = 0.1$ (see Piketty 1998 and, more recently, Choné et al. 2003). 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[15].

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 – 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 and $\varepsilon = 0.5$, the condition for the Paretianity of revealed social preferences as given by (10) is that the marginal tax rate is below 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 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[16]. With such a value of the Pareto parameter non Paretianity holds.

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

[15] These results are available upon request.

[16] As in the case of the Kernel, we interpolate the density using an unbounded Pareto distribution.

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 larger than some threshold. We asked the French National Statistic Institute (INSEE) to perform the estimations and to provide us with the estimates of the Pareto parameter using a more numerous and more precise survey based on income tax returns: the "Survey on the Fiscal Incomes 1996" [17]. The estimations yielded values of *a* in the interval [2.9-3.2] reinforcing our non Paretianity result when the elasticity of labor supply is high enough.

5. Income Effects

The inclusion of income effects influences the Non Paretianity results. With preferences represented by a function of the following type: $U(c, L) = A(c) - B(L)$ (where $A(c)$ is not supposed to be linear anymore) it may be shown that the optimal taxation formula (4) becomes:

$$\frac{t(y)}{1-t(y)}\psi[c(w)] = (1+\frac{1}{\varepsilon}).\frac{1-F(w)}{w.f(w)}.[\bar{\psi}[c(w)] - S(w)] \qquad (14)$$

where $\psi[c(w)] = \frac{1}{A'(c)}$ is the inverse of the marginal utility of income, and $\bar{\psi}[c(w)] = \frac{1}{1-F(w)}\int_w^z \psi[c(w)]f(x)dx$ i.e. the mean value of that inverted marginal utility for people with productivity above *w*. From equation (14) we can recover the equivalent of proposition 1 that is:

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

[17] 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.

$$t(y) \leq \frac{\frac{1+\varepsilon}{\varepsilon}\frac{1-F(w)}{w.f(w)}\frac{\overline{\psi}[c(w)]}{\psi[c(w)]}}{1+\frac{1+\varepsilon}{\varepsilon}\frac{1-F(w)}{w.f(w)}\frac{\overline{\psi}[c(w)]}{\psi[c(w)]}} \quad for\ all\ w \in [w_0, Z] \qquad (15)$$

This is the equivalent of (9) when income effects are considered. By comparing (9) and (15) we can easily see that, as $\frac{\overline{\psi}[c(w)]}{\psi[c(w)]} \geq 1$, the right hand side of (9) is always smaller that the corresponding term in (15). This implies that the inclusion of income effect mitigates the possibility to be Non Paretian. Results of the computation of *S(w)* obtained using $U(c,L) = \frac{c^{1-\frac{1}{\alpha}}}{1-\frac{1}{\alpha}} - \frac{L^{1+\frac{1}{\beta}}}{1+\frac{1}{\beta}}$ (which leads to the constant labor supply elasticity $\varepsilon = \frac{\beta(\alpha-1)}{\alpha+\beta}$) 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[18]. It may be seen that, in both cases, the upper incomplete mean marginal social welfare, *S(w)*, became negative only beyond the 95th centile. 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.

6. Taking into account explicitly participation decisions: the Saez model

The Mirrlees approach builds on a labor supply model which only focuses on hours-of-work responses. More realistic labor supply model, incorporating labor market participation responses, can provide some quite different results in optimal income taxation[19]. To check to what extent this is an important assumption and to contrast welfare weights in the standard intensive elasticity scenario with a scenario with participation effects we now present the results of the inversion of an optimal tax problem à la Saez (2002) where extensive and

[18] To compute the terms $\psi[c(w)]$ and $\overline{\psi}[c(w)]$ in formula (14) we used the observed disposable income as proxy for the optimal consumption *c*.

[19] In the original Mirrlees (1971) formulation, there is a threshold skill level under which individuals do not work. This implies that the intensive elasticity at the bottom is infinite. Therefore, there is an element of labor force participation in the intensive model of Mirrlees (1971). But the labor force participation choice is only between unemployment and an infinitesimal amount of work. This feature is not empirically realistic, because of fixed costs of work.

intensive labor supply behaviors are explicitly taken into account in an optimal labor income taxation model.

Saez (2002) sets up a discrete optimal tax problem conceptually very similar to the one described above with the particularity of differentiating explicitly the labor supply decisions (how much to work) from the participation decisions (working or not). In his model there are *I+1* groups in the labor market: *I* groups of individuals who do work (ranked by increasing earnings from 1 to I) plus one group consisting of those who do not work (group 0). Individuals choose whether or not to participate (the extensive margin), and which group to choose (the intensive margin). In this framework, optimal taxation has the following form [see Saez (2002) for a formal derivation]:

$$\frac{T_i - T_{i-1}}{C_i - C_{i-1}} = \frac{1}{\mu_i h_i} \sum_{j \geq i}^{I} h_j \left[1 - g_j - \chi_j \frac{T_j - T_0}{C_j - C_0} \right] \qquad (16a)$$

and

$$T_0 = \Phi - \sum_{j \geq i}^{I} h_j T_j \qquad (16b)$$

In this expression, Φ is the exogenous government financial constraint, T_i is net tax paid by group *i* and C_i is the net household income of this group. The term on the left-hand side is the discrete equivalent of the marginal tax rate, i.e. the extra tax paid when moving from group *i-1* to *i* divided by the gain in net income. Non-workers receive benefits $-T_0$, by definition identical to C_0. Gross earnings within group *i*, Y_i, equal to $C_i + T_i$, are supposed to be fixed. h_i measures the share of group *i* in the population. The social welfare function is summarized by g_i, the marginal weight the government assigns to group *i*. This weight represents the value (expressed in terms of public funds) of giving an additional euro to an individual in group *i*. Alternatively, one can say that the government is indifferent between giving one more euro to an individual in occupation *i* and getting g_i more euros of public funds. It is equivalent to $\frac{G(.)}{\lambda}$ (eq. 8) in the standard Mirrlees setup. The intensive elasticity, μ_i, is defined as:

$$\mu_i = \frac{C_i - C_{i-1}}{h_i} \frac{dh_i}{d(C_i - C_{i-1})} \qquad (17)$$

This mobility elasticity captures the percentage increase in the number of agents in group *i* when C_i-C_{i-1} is increased by 1%, and is defined under the assumption that individuals are restricted to adjust their labor supply to the neighboring group. Note that, as shown in Saez (2002), this intensive elasticity is related with the classical labor supply elasticity ε_i in the Mirrlees model by the following relationship: $\mu_i = \frac{Y_i}{Y_i - Y_{i-1}} \varepsilon_i$ (18)

Finally, χ_i is a measure of the extensive elasticity, and is defined as the percentage of individuals in group *i* who stops working when the difference between the net household income out of work and at earnings point *i* is reduced by 1%:

$$\chi_i = \frac{C_i - C_0}{h_i} \frac{dh_i}{d(C_i - C_0)} \quad (19)$$

The main implication of the optimal tax rule above is that the optimal tax system depends heavily on whether labor supply responses are concentrated at the intensive or extensive margin.

As with the Mirrlees model, it is possible to invert the model in order to reveal the social preferences about inequality (i.e. the term g_i). From equation (16) we obtain that:

$$g_i = 1 - \chi_i \frac{T_i - T_0}{C_i - C_0} - \mu_i \frac{T_i - T_{i-1}}{C_i - C_{i-1}} + \frac{1}{h_i} \sum_{j=i+1}^{I} h_j \left[1 - g_j - \chi_j \frac{T_j - T_0}{C_j - C_0} \right] \quad (20)$$

and, given that the group *I* is the last one in the population of workers:

$$g_I = 1 - \chi_I \frac{T_I - T_0}{C_I - C_0} - \mu_I \frac{T_I - T_{I-1}}{C_I - C_{I-1}} \quad (21)$$

Equations (20) and (21), jointly with the normalizing condition $\sum_{i=0}^{I} h_i g_i = 1$[20], allows us to compute recursively, from the observation of T_i, C_i, h_i, μ_i and χ_i for each $i = I, I-1, \ldots 0$, the marginal weights g_i that the government assigns to each class of agents.

[20] Note that, as shown in Saez (2002) this condition holds only if income effects are ruled out.

It is easy to see (Saez 2002) that, when the elasticity of participation χ_i tends to zero, then equations (20) and (21) reduce to a discrete version of equation (8) and then all the results previously obtained still hold. It is also immediate to see that, in the classical intensive labor supply framework (i.e. $\chi_i = 0$), the condition of Paretianity of the social welfare function (i.e. the equivalent of proposition 1) for the last group of agent (the group I) is, by equation (21):

$$\frac{T_I - T_{I-1}}{C_I - C_{I-1}} \leq \frac{1}{\mu_I} \qquad (22)$$

and then that there exists a threshold for the intensive elasticity of labor supply below which the social welfare function is necessarily non-decreasing everywhere (proposition 2).

When the elasticity of participation χ_i is positive, the possibility to be non Paretian (i.e. $g_i \leq 0$) increases given that the terms $-\chi_j \frac{T_j - T_0}{C_j - C_0}$ are negative (see equation 20).

For comparability with Saez (2002), we first report the results on a sample of singles aged 18 to 65, in which students and individuals with non-labor income above 10 per cent of total income are eliminated from the sample. The final sample used in this exercise contains 1028 singles (963 working). The rate of nonlabor force participation (zero yearly earnings reported) for this group is around 9 percent. We present only the case in which the redistribution system is the "gross" (as defined in the previous section. We have defined a discrete grid of eleven income levels Y_i trying to obtain the same frequency h_i in each class. Table 1 gives the statistics of our sample. Following the discussion in section 3, about the elasticity parameters summarizing the behavioral responses, we present simulations using 3 ranges of parameter values (see Table 2). Three main groups of scenarios are simulated: no participation effects (scenarios A and B), medium participation elasticity (scenarios C, D, E and F) and high participation elasticity (scenarios G, H, I and L). In particular, the values for the participation elasticity χ is taken as constant and equal to 0, 0.5 or 1 for incomes below 75000 francs (i.e. until group 2) per year and equal to 0 for the rest of the population because it is certainly small for middle and higher income earners.

In order to compare these new results with the previous one, all simulations are presented in terms of the intensive labor supply elasticity from the standard model which is denoted by ε.

The intensive elasticity ε for incomes below 75000 francs per year is taken as constant and equal to 0, 0.1 or 0.5. The middle and high income (above 75000 francs) elasticity is taken as constant and equal to 0.1 or 0.5. All simulations have been carried out assuming no income effects.

Figure 11 shows the social marginal weights computed in a complete intensive labor supply framework (i.e. $\chi = 0$), with average ε = 0.1 (scenario A) and 0.5 (scenario B) (the values of g_i are reported in columns A and B of table 3). As expected, the same qualitative results are found as in the standard Mirrlees framework (Figure 4). In particular, with high intensive elasticities (scenario B), we obtain negative social marginal weights for the upper part of the income distribution (last decile).

Table 3 reports the social marginal weights associated to each income group under different scenarios (negative weights are in bold). The inclusion of medium participation effects (scenarios C, D, E, and F) does not change qualitatively the results obtained in a classical intensive labor supply model à la Mirrlees. The non-Paretianity results are limited to the upper class of incomes and under the hypothesis of high intensive labor supply reactions. On the contrary, including high participation effects ($\chi_i = 1$), implies revealing negative marginal weights not only for the upper part of the workers population but also for the first category of workers (i =1) (see scenarios G, H, I, L). Again it must be underscored that this conclusion depends on the prior that the redistribution authority has about the participation elasticity. In practice, participation elasticity bigger than 0.5 are extremely implausible for sample of singles (there is no empirical evidence, at the moment supporting such a scenario neither in France nor in other countries)[21]. On the contrary, looking at the results of the empirical literature on discrete choice model of labor supply (see in particular Piketty 1998) values of participation elasticity higher than 0.5 are observed for the so-called secondary household members (in particular women with children). Probably the best way to incorporate the big size participation elasticity of the 2nd earner in a household income tax model (where household are considered as an agent) is to treat it as an intensive elasticity given that there is always a first earner working full time. This is what we have done in the section 3.

[21] It must be also stressed that revealing non decreasing social marginal weights implies the violation of the concavity conditions on the social welfare function and then the impossibility to ensure that the observed redistribution policy is consistent with a maximizing behavior.

Alternatively it would be necessary to write down a model of household income taxation explicitly accounting for both spouses participation but this is out of the scope of this paper (see Kleven et al. 2006).

The results of this section are in line with the ones obtained by Saez (2002), Laroque (2005) and Blundell et al. (2006). In his simulations on US data, Saez (2002) shows that a Negative Income Tax program with a large guaranteed income level which is taxed away at high rates (as the French Minimum Income Guaranteed Scheme –RMI-) it is never optimal when participation elasticities are high and the social welfare function is increasing and concave everywhere. In that case, an Earning Tax Credit implying negative marginal tax rates on the low ability agents (class 1) is optimal. In our simulation, we show that, with high participation elasticities, a redistribution system with positive and very high marginal tax rate at the bottom of the income distribution (as it is the case for the RMI) can be candidate to be optimal only if the social planner is Non Paretian.

Blundell et al. (2006) perform analogous simulations on a sample of lone mothers in Germany and UK, using elasticities (both intensive and of participation) estimated econometrically with values corresponding to our low-medium scenarios (from 0 to 0.2 in Germany and from 0 to 0.5 in UK). They find patterns of social marginal weights very similar to ours.

Laroque (2005) computes the Laffer bound in France (e.g. and the revenue maximizing and efficiency cost minimizing tax system under Paretian specification of social preferences) and compare it with the 1999 French tax system. He finds that the French system is on the left of the Laffer bounds. This implies that French tax policy is Rawlsian, so that Laroque's results also become a statement about the social preferences implied by observed tax policy. In terms of our results, this means that, among the scenarios we simulated, the most plausible are the low elasticity ones.

7. 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 paper offers a set of optimality conditions that the observed marginal tax rate must respects in order to be compatible with the maximization of a concave social welfare function

and derives conditions for the revealed social welfare function be Paretian (i.e. increasing everywhere).

The detailed empirical analysis performed on France shows that the observed marginal tax rate is in agreement with standard optimal tax theory and that the revealed social welfare function is increasing and concave 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. Taking explicitly into account participation decisions confirms the result that high marginal tax rates are compatible with the maximization of a Paretian social welfare function only if the labor supply elasticities are low.

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 takes it that labor supply responses to taxation are low, or it has non-Paretian social preferences (and there is then space for a Pareto improving tax changes), or it does not optimize 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 might result more from political economy forces than from the pursuit of some well defined social objective. They may also reflect various other constraints that policy makers face (e.g., that real-world tax systems for practical reasons have to be piecewise linear).

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. The optimum inverse problem

All the results in section 1 are based on the hypothesis that the observed marginal tax rate is the result of maximization of a social welfare function with the budget and the incentive compatibility constraints. This assumption imposes several restrictions on the shape of the observed marginal tax rate that, if not satisfied make the whole inversion procedure inconsistent. Let see them in details.

First we have to ensure that an observed marginal tax rate *t(y)* is consistent with an agent maximizing behavior and that the individual utility function chosen fulfills the Spence-Mirrlees condition (this condition ensure that the first order approach to the incentive compatibility constraint is sufficient, see Ebert 1992). The conditions to be checked are:

A) $t(y) < 1$ for any *w* (from the f.o.c. of problem 1.2);

B) $t'(y) > \frac{U_{LL} + U_{cc}[w(1-t(y))]^2}{w^2 U_c}$ for any *w* (from the s.o.c. of problem 1.2);

C) $\frac{\partial C}{\partial w} > 0$ without taxes; this is the Spence-Mirrlees condition.

Second, we have to ensure that the observed *t(y)* is consistent with the solution of an optimization problem à la Mirrlees. Let start by rewriting the original optimization problem (1) as an optimal control problem– see also Atkinson and Stiglitz (1980), p. 415. Using the utility function (2), the correspondent Hamiltonian is:

$$H[L(w), V(w), \mu(w), \lambda] = \left[G(V) + \lambda\left(wL - V(w) - B(L) - \bar{T}\right)\right] f(w) + \mu(w)\frac{L}{w} B_L(L)$$

where *L(w)* is the control variable, *V(w)* is the state variable, λ is the Lagrange multiplier associated to constraint (1.4) and $\mu(w)$ is the co-state variable associated to the first order incentive compatibility constraint $\frac{\partial V(w)}{\partial w} = \frac{L}{w} B_L(L)$ (an alternative way to rewrite constraints 1.2 and 1.3)

The Pontryagin Maximum principle states that the following first order conditions are necessary:

(p. foc 1) $\frac{\partial H()}{\partial L} = \lambda(w - B_L) f(w) + \mu(w) \frac{(LB_L)_L}{w} = 0$

(p. foc 2) $\frac{\partial H()}{\partial V} = -\frac{\partial \mu(w)}{\partial w} \Rightarrow [G'(.) - \lambda] f(x) = -\frac{\partial \mu(w)}{\partial w}$ that, after integration and making use of the transversality condition $\mu(Z) = 0$ implies that $\int_w^Z \left(1 - \frac{G'(.)}{\lambda}\right) f(x)dx = -\frac{\mu(w)}{\lambda}$.

Consolidating the two and making use of the f.o.c of problem (1.2) we obtain the condition (4) on the marginal tax rate *t(y)*.

It is well-known (for the Mangasarian theorem) that the Pontryagin Maximum Principle that leads to the optimality conditions (p. foc 1) and (p.foc 2) are necessary and sufficient provided that *H(.)* is differentiable and concave in the variables *(L,V)* jointly. Given that in our case *H* is separable in *(L,V)*, the Mangasarian theorem implies that:

D) $\frac{\partial^2 G(.)}{\partial V^2} < 0$ (e.g. the concavity of social welfare function. It ensures the concavity of the Hamiltonian with respect to *V*).

and

E) $\frac{\mu(w)}{\lambda} < w f(w) \frac{B_{LL}}{[LB_L]_{LL}}$ (from $\frac{\partial^2 H()}{\partial L^2} < 0$; it ensures the concavity of the Hamiltonian with respect to the control variable *L*).

The empirical tests of conditions A, B and C are immediate. Note that with *U(c,L)* isoelastic and quasi-linear in consumption $U_{cc}=0$ and $U_c=1$ then condition B) can be rewritten as:

$t'(y) > \frac{U_{LL}}{w^2} = -\frac{(1 - t(y))^{1-\varepsilon}}{\varepsilon w^{1+\varepsilon}}$ or $t'(y) < \frac{1 - t(y)}{\varepsilon y}$.

The Spence-Mirrlees condition C) is always satisfied with the *U(.)* chosen.

The empirical test of condition D) is based on the 'curvature' of the density function *f(w)* and on the derivatives of the marginal tax rate *t(y)* (see appendix 4 for a formal derivation). As only a limited precision can be empirically obtained on these functions this approach will not

be pursued in this paper. A much simpler test on the concavity of the social welfare function can be easily implemented by an inspection of the shape of the *G'(V(w))* computed as in (8) when plotted on *w*. If it is everywhere decreasing then we can easily prove that $\frac{\partial^2 G(.)}{\partial V^2} < 0$ given that $\frac{\partial G'(.)}{\partial w} = \frac{\partial^2 G(.)}{\partial V^2}\frac{\partial V(.)}{\partial w}$ and that $\frac{\partial V(.)}{\partial w}$ is always positive. This is the basic concavity test we will follows.

With *U(c,L)* iso-elastic and quasi-linear in consumption, the condition E) becomes $\frac{\mu(w)}{\lambda} < \frac{\varepsilon}{1+\varepsilon} wf(w)$. Using (p.foc 2), it can be written as $S(w) < 1 + \frac{\varepsilon}{1+\varepsilon}\frac{wf(w)}{1-F(w)}$. Then, using (7), he reduces to: $-1 < \frac{t(y)}{1-t(y)}$ that is always satisfied if $0 \leq t(y) \leq 1$. Note that if conditions A), B) and C) are satisfied, condition E is automatically satisfied given that ensuring the compatibility of the marginal tax rate with the maximizing behavior of any agent means that the tax scheme is incentive compatible.

Appendix 2. 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), requires the knowledge of the continuous functions *f(w)*, *t(y)* 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, $t(y_i)$ The following operations permit to smooth those functions and to get estimates of their derivatives.

(i) For any arbitrary value of productivity, *w*, obtain an estimate of the density function *f(W)* and the effective marginal tax rate *t(wL)* by kernel techniques defined over the whole sample of observations - using a Gaussian kernel with an adaptive window. 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[22].

(ii) Estimate the derivatives of *t(y)* and *f(w)* using again a kernel approximation computed over the whole sample.[23]

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

(iv) Compute the function *G'(w)* as in (8) and the function *S(w)* as in (7).

b) *The Adaptive Nadaraya- Watson kernel*

For a theoretical explanation see Silverman (1986) and Pagan and Ullah (1999). The adaptive kernel techniques allow to compute, starting from a set of pair wise observations x_i and $y_i(x_i)$, a one-dimensional kernel density $f(x_i)$ and a kernel estimation $\hat{y}(x_i)$ of $y_i(x_i)$.

In our case $x_i = w_i$ and $y_i(x_i) = t(w_i L_i)$.

Let $\hat{y}(x) = \dfrac{\sum_{i=1}^{n} K_h(x - x_i) y_i n^{-1}}{\sum_{j=1}^{n} K_h(x - x_i) n^{-1}}$ be the normal kernel that will be used as pilot kernel (following Silverman 1989). The kernel actually used is a modified Gaussian $K_h(x - x_i) = \exp\left(-h(x - x_i)^2\right)$

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

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

where $h = \frac{1}{optimalwindow^2} = \frac{1}{\left(\frac{1.06 * \text{standard deviation(x)}}{n^{0.2}}\right)^2}$ as in Hardle (1991) pag 91.

Define now the local bandwidth factor:

$$\lambda(x) = \left(\frac{g}{\hat{y}(x)}\right)^{\alpha} \quad \text{where } g = \textit{geometric means of } \hat{y}(x): \; g = \exp\left[\frac{\sum_{i=1}^{n}\log(\hat{y}(x))}{n}\right]$$

and α is a sensitivity parameter for the adaptive bandwidth taking values from 0 to 1. Near 0 it reproduces the normal kernel. Near 1 is high adaptive. The value used in our simulations is 0.5 (following Silverman 1986). The *Adaptive Nadaraya – Watson* kernel estimator is:

$$\hat{y}(x) = \frac{\sum_{i=1}^{n} \lambda K_h(x - x_i) y_i n^{-1}}{\sum_{j=1}^{n} \lambda K_h(x - x_i) n^{-1}}$$

where $\sum_{j=1}^{n} \lambda K_h(x - x_i) n^{-1}$ is the Adaptive Rosenblatt-Parzen density estimation.

Kernels Derivatives (see: Pagan and Ullah 1999)

Let $\hat{y}(x) = \frac{g(x)}{f(x)}$ be the kernel estimator. The analytical expression for its derivative is:

$$\hat{y}'(x) = \frac{g'(x)}{f(x)} - \frac{g(x) f'(x)}{f(x)^2} \quad \text{and} \quad \hat{y}''(x) = \frac{g''(x)}{f(x)} - 2\frac{g'(x) f'(x)}{f(x)^2} - \hat{y}(x)\left[\frac{f''(x)}{f(x)} + 2\left(\frac{f'(x)}{f(x)}\right)^2\right]$$

In order to adjust the windows for the kernel estimation of derivatives (e.g. the optimal window for derivatives is not the optimal window for the primitive kernel), we use the following formula for the optimal window of *s-order* derivative (on the subject see Pagan and Ullah 1999 pag 56-57):

optimal window for $\hat{y}^{(s)} = \frac{1.06 * \text{standard deviation(x)}}{\frac{1}{n^{5+2s}}}$ from which (after some algebra) we obtain:

$$\hat{h}_s = h * n^{\left(\frac{-4s}{25+10s}\right)}$$

c) 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 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 $wL > RMI$[24]. This budget constraint is clearly convex. Therefore, there should be a range of labor

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

incomes around the RMI where it would be irrational to be[25]. 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 3. Microsimulation model, dataset and data treatment

The sample and the microsimulation 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 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:

[25] 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.

$$emtr = \frac{\Delta Taxes + \Delta Benefits}{\Delta Gross\ Income} = 1 - \frac{\Delta Yd}{\Delta y}$$

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 4. On the concavity of the social welfare function

We can show differentiating (8) with respect to w that:

$$\frac{\partial G'(.)}{\lambda \partial w} = \frac{\varepsilon}{1+\varepsilon} z(y)\phi(y)\frac{1+\hat{\varepsilon}}{w}\left[1+\eta(w)\left(1+\frac{\hat{\varepsilon}_w}{1+\hat{\varepsilon}}\left(1+\sigma_f-\eta(w)\right)\right)+(1+\hat{\varepsilon})(z(y)+y)\right]$$

where $w\frac{f''(w)}{f'(w)} = \sigma_f(w)$ is the curvature of *f(w)* and where we have defined the following functions:

$$\frac{v(y)}{1-t(y)} = z(y);$$

$$\frac{t(y)}{1-t(y)} = \phi(y);$$

$$\hat{\varepsilon} = \frac{w}{l}\frac{\partial l(w)}{\partial w} = \varepsilon\left[\frac{1-t(y)(1+v(y))}{1-t(y)(1+v(y))\varepsilon}\right] \text{ and}$$

$$\hat{\varepsilon}_w = \frac{\partial \hat{\varepsilon}(w)}{\partial w} = \varepsilon t(y)\left[\frac{\left(\varepsilon z(y) - v(y)(1+\hat{\varepsilon})\frac{\varepsilon z(y)}{w}\right)[1-t(y)z(y)] - \left(\frac{1+\hat{\varepsilon}}{w}[v(y)+z(y)y]\right)}{[1+t(y)z(y)\varepsilon]^2}\right].$$

Given that: $\frac{\partial^2 G(.)}{\partial V^2} = \frac{\frac{\partial G'(.)}{\partial w}}{\frac{\partial V}{\partial w}}$ and $\frac{\partial V(.)}{\partial w} = \frac{l}{w}B'(l) = l(1-t(y)) > 0$, we have that

$$\frac{\partial^2 G(.)}{\lambda \partial V^2} = \frac{\varepsilon}{1+\varepsilon} z(y)\phi(y) \frac{1+\hat{\varepsilon}}{y(1-t(y))}\left[1+\eta(w)\left(1+\frac{\hat{\varepsilon}_w}{1+\hat{\varepsilon}}\left(1+\sigma_f-\eta(w)\right)\right)+(1+\hat{\varepsilon})(z(y)+y)\right]$$

•

The concavity of *G(V)* implies that

$$\frac{\partial^2 G(.)}{\lambda \partial V^2} = \frac{\varepsilon}{1+\varepsilon} z(y)\phi(y) \frac{1+\hat{\varepsilon}}{y(1-t(y))}\left[1+\eta(w)\left(1+\frac{\hat{\varepsilon}_w}{1+\hat{\varepsilon}}\left(1+\sigma_f-\eta(w)\right)\right)+(1+\hat{\varepsilon})(z(y)+y)\right] < 0$$

(23)

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 (23) becomes:

$$\frac{\partial^2 G(.)}{\lambda \partial V^2} = \frac{\varepsilon}{1+\varepsilon} z(y)\phi(y) \frac{1+\hat{\varepsilon}}{y(1-t(y))}\left[-a+(1+\hat{\varepsilon})(z(y)+y)\right] < 0 \qquad (24)$$

Notice that when the marginal tax rate is increasing, that is for high incomes, *z(y)* is positive. Then, in that case, condition (24) is verified if:

$$a > \frac{(1+\varepsilon)(y+z(y))}{1+t(y)z(y)\varepsilon} \qquad (25)$$

If the preceding conditions (23) or (24) are not satisfied, the revealed social welfare function would be convex, 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 non convex social welfare function.

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., 2005 "*On Modelling Household Labor Supply With Taxation*", IZA Discussion Paper 1455. Bonn

Blundell R., M. Brewer, P. Haan, A. Shephard, 2006, "*Optimal income taxation of lone mothers: an empirical comparison for Germany and the UK*", 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, "Labor 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.

Canto, V. A., D.H. Joines, A. B. Laffer (eds.) 1982. "*Foundations of Supply-Side Economics, Theory and Evidence"*. New York: Academic Press.

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 Labor 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.

Ebert U., 1992, “A Reexamination of the Optimal Nonlinear Income Tax”, *Journal of Public Economics*, 49: 47-73

Eissa N., H. Hoines, 2006, " Behavioral Reponses to Taxes: Lessons from the EITC and Labor Supply" in *Tax Policy and the Economy*, edited by J. Poterba, MIT Press (1968); vol.20, 74-110.

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.

Kleven H.J., Kreiner C.T., E. Saez, 2006 “*The Optimal Income Taxation of Couples*”, NBER Working Paper 12685. Cambridge MA.

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., 2005, “Income Maintenance and Labor Force Participation”, *Econometrica*, 73: 2 341-376

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

Piketty T., 1999, “Les hauts revenus face aux modifications des taux marginaux supérieurs de l’impôt sur le revenu en France, 1970-1996*”*, *Economie et Prévision*, 138-139: 25-60.

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., 1997, “La redistribution fiscale contre le chômage*”*, *Revue Française d’Economie*, 12(1) : 157-203.

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., 2002, “Optimal Income Transfer Programs: Intensive Versus Extensive Labor Supply Responses”, *Quarterly Journal of Economics,* 117: 1039-1072

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

Salanié 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.

Silverman B., 1986. *“Density Estimation for Statistical Data Analysis”* Chapman and Hall, New-York.

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.

Table 1. Sample statistics

I	Y_i	C_i	T_i	h_i	*F(Y)*
0	0	12000	-12000	0,09	*9%*
1	48857	35919	12939	0,08	*17%*
2	74340	54398	19942	0,09	*26%*
3	91116	64926	26190	0,09	*35%*
4	105954	73144	32811	0,09	*44%*
5	121247	80750	40497	0,09	*53%*
6	135790	87779	48011	0,09	*63%*
7	152870	95747	57122	0,09	*72%*
8	175352	106173	69179	0,09	*81%*
9	215857	123988	91869	0,09	*90%*
10	408454	217915	190539	0,10	*100%*

Table 2. Elasticity parameters retained in the simulations

Scenarios *(elasticities)*											
A		**B**		**C**		**D**		**E**			
χ	ε	χ	ε	χ	ε	χ	ε	χ	ε		
0	0.1	0	0.5	0.5	0	0.5	0.1	0.5	0.5	***i= 0, 1, 2***	***classes of income***
0	0.1	0	0.5	0	0.1	0	0.1	0	0.1	***i= 3-10***	

Scenarios *(elasticities)*											
F		**G**		**H**		**I**		**L**			
χ	ε	χ	ε	χ	ε	χ	ε	χ	ε		
0.5	0.5	1	0	1	0.1	1	0.5	1	0.5	***i= 0, 1, 2***	***classes of income***
0	0.5	0	0.1	0	0.1	0	0.1	0	0.5	***i= 3-10***	

Table 3. Social marginal weights (g_i) associated to each income group under different scenarios

		No participation effects		**Medium participation elasticity**				**High participation elasticity**			
i	***F(Y)***	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**L**
0	*9%*	1,13	1.63	1,88	2,01	2,52	2,52	2,76	2,89	3,40	3,40
1	*17%*	0,92	1.14	0,48	0,40	0,08	0,08	**-0,04**	**-0,12**	**-0,44**	**-0,44**
2	*26%*	1,28	1.72	0,95	0,90	0,70	1,99	0,57	0,53	0,33	1,61
3	*35%*	1,11	1.72	1,11	1,11	1,11	1,56	1,11	1,11	1,11	1,56
4	*44%*	1,12	1.57	1,12	1,12	1,12	1,57	1,12	1,12	1,12	1,57
5	*53%*	1,03	1.13	1,03	1,03	1,03	1,13	1,03	1,03	1,03	1,13
6	*63%*	1,05	1.22	1,05	1,05	1,05	1,22	1,05	1,05	1,05	1,22
7	*72%*	1,01	1.05	1,01	1,01	1,01	1,05	1,01	1,01	1,01	1,05
8	*81%*	1,06	1.28	1,06	1,06	1,06	1,28	1,06	1,06	1,06	1,28
9	*90%*	0,95	0.73	0,95	0,95	0,95	0,73	0,95	0,95	0,95	0,73
10	*100%*	0,42	**-1.84**	0,42	0,42	0,42	**-1,84**	0,42	0,42	0,42	**-1,84**

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

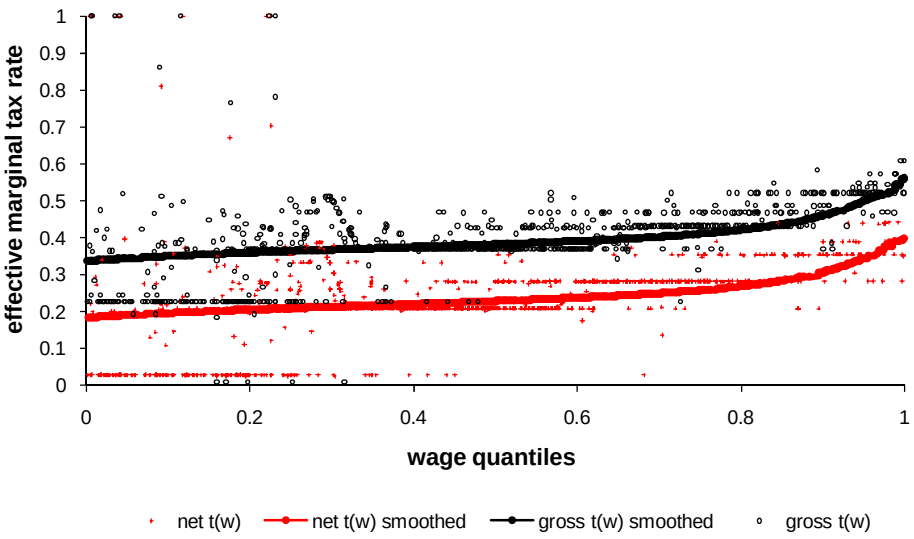


Figure 2. Kernel wage densities for singles: net and gross scenario

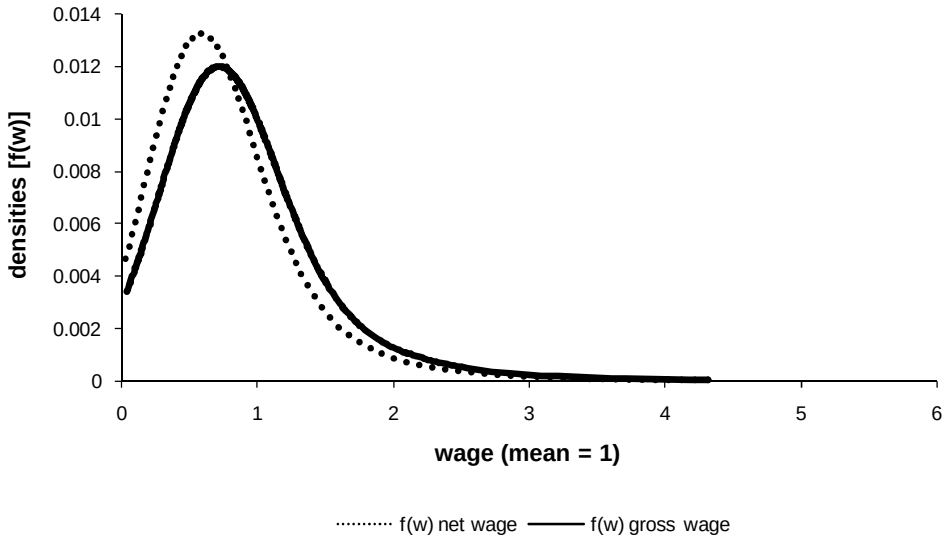


Figure 3. Social marginal welfare for singles (on net wages)

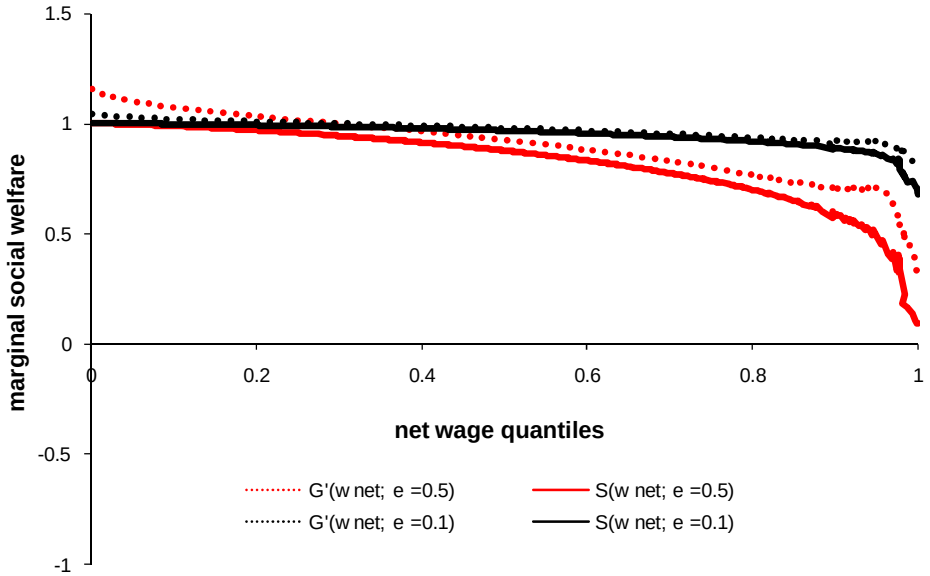


Figure 4. Social marginal welfare for singles (on gross wages)

Figure 5. Kernel productivity densities for singles

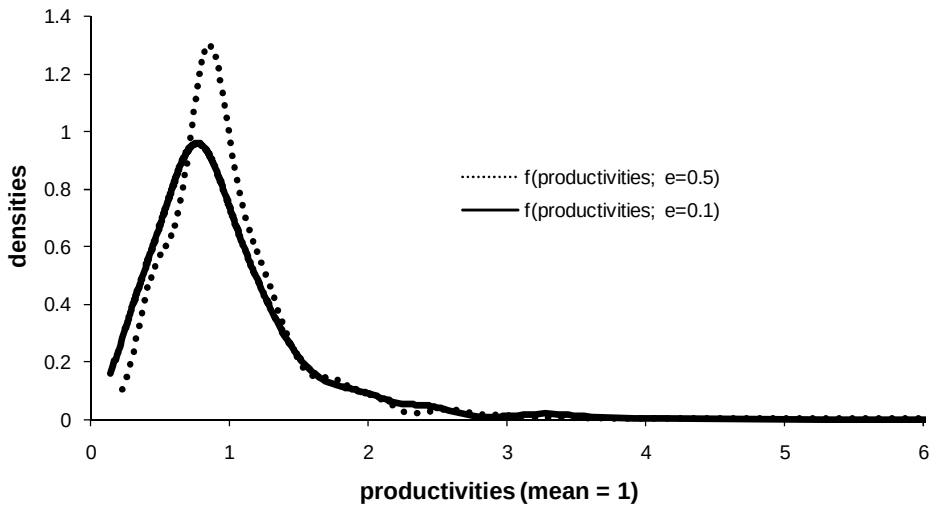


Figure 6. Social marginal welfare for singles (on productivities)

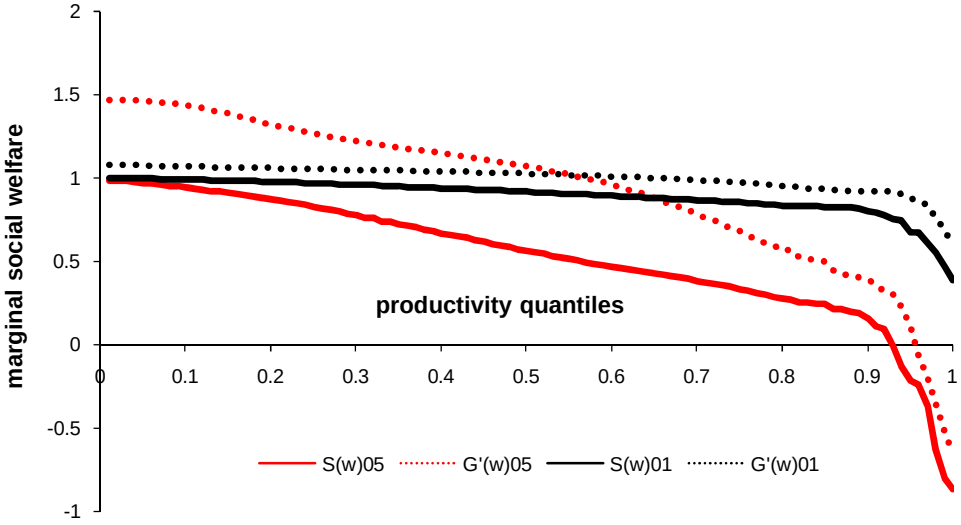


Figure 7. Gross kernel smoothed marginal tax rates for all households

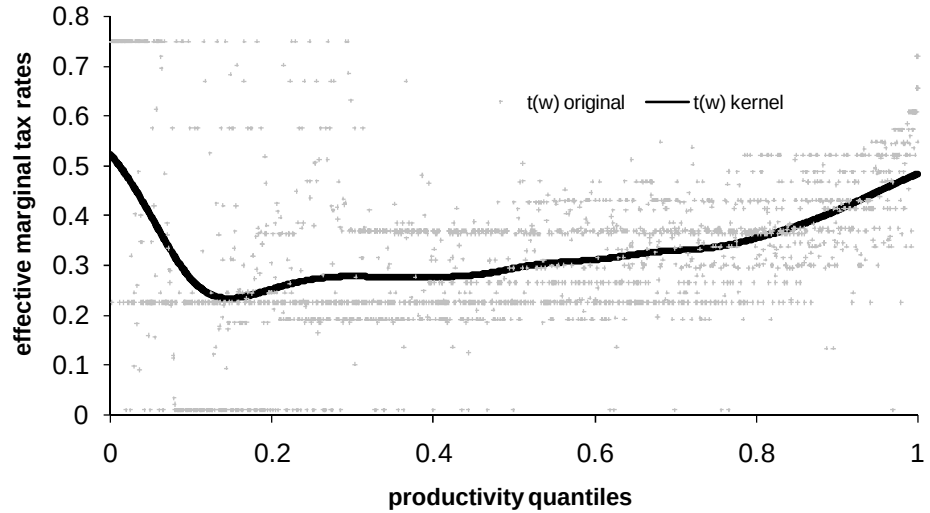


Figure 8. Kernel productivity densities for all households

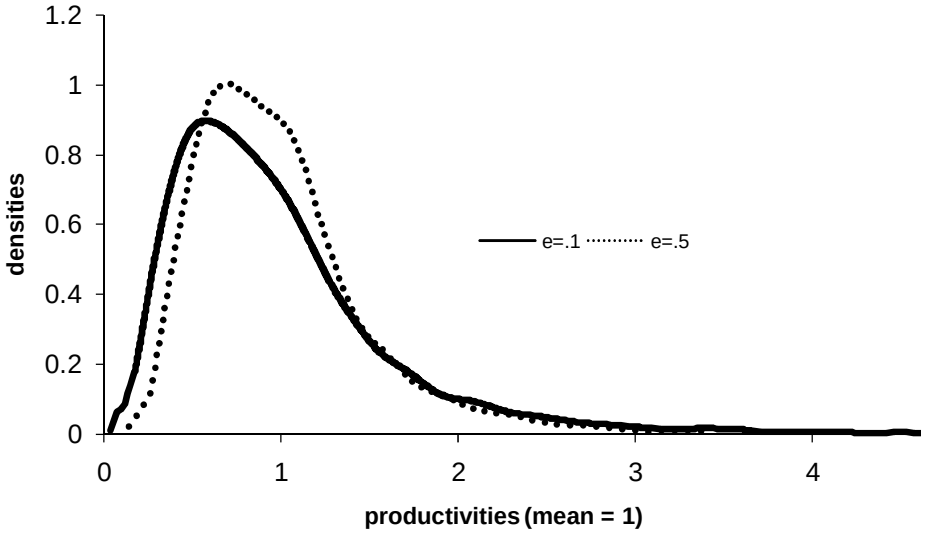


Figure 9. Social marginal welfare for all household (on productivities)

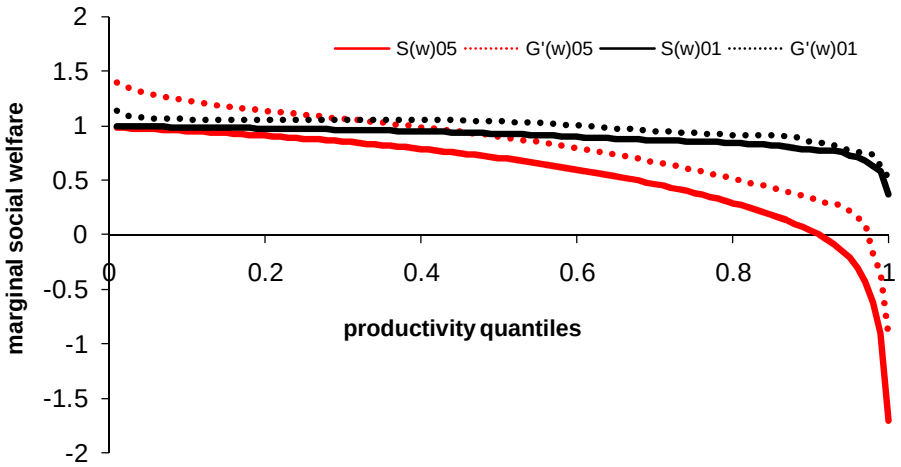


Figure 10. Paretianity test on social marginal welfare for singles (on gross wages) with income effects

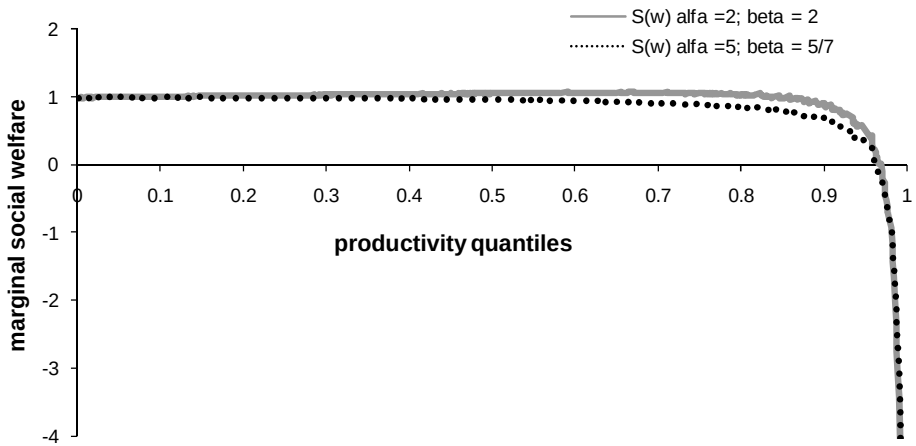


Figure 11. Social marginal weights for singles gross wages (scenario A and scenario B)