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

WORKING PAPER N° 2006 - 01

Optimal nonlinear labor income taxation in dynamic economies

Salvador Balle

Amedeo Spadaro

JEL Codes : H21, C63

Keywords : Dynamic optimal income taxation, private information, learning by doing.

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE – ÉCOLE DES HAUTES ÉTUDES EN SCIENCES SOCIALES
ÉCOLE NATIONALE DES PONTS ET CHAUSSÉES – ÉCOLE NORMALE SUPÉRIEURE

OPTIMAL NONLINEAR LABOR INCOME TAXATION IN DYNAMIC ECONOMIES*

Salvador Balle and Amedeo Spadaro

The aim of this paper is to explore the characteristics of the optimal nonlinear labor income tax in dynamic economies with information asymmetries and human capital accumulation. We develop a dynamic optimal income tax model in which agent's productivity evolves over time according to two different factors: an exogenous component and a learning by doing process endogenous to the fiscal policy. The latter is determined by the government, maximizing in the initial period a social welfare function capturing some level of aversion to inequality. We characterize analytically the first order condition driving the optimal tax schedule in a model in which agents choose the consumption and labor supply patterns that maximize their lifetime utility function. We show that the inclusion of the endogenous evolution of productivities into the tax problem changes the results with respect to the static framework à la Mirrlees (1971). We find that the optimal tax strategy balances social marginal costs of increasing marginal tax rates with social marginal benefits of doing so. The costs are related with the reduction of both past and future human capital accumulation, with the negative impact on aggregate social welfare due to the reduction of the individual utility of all the agents paying more taxes and with the increase of the necessity to redistribute more in the future (given that the spread among social marginal weights of each agents will be higher). The benefits derive from the increase of the instantaneous tax receipts with the consequent reduction of the overall tax burden all along the time horizon and from the reduction of present inequality of incomes and future inequality of the productivities. As a particular extreme result we find that it can be optimal to subsidize (instead of taxing) at the margin high productivity agents.

Keywords: Dynamic Optimal Income Taxation, Private Information, Learning by Doing.
JEL classification numbers: H21, C63

* Spadaro acknowledges the support of Spanish Government-MEC (SEJ2005-08783-C04-03) and Fundación FBBVA. Versions of this paper were presented at a large number of seminars; we thank the seminar participants for their comments. We also thank Rolf Aaberge, Nuria Badenes, Lluis Bru, Daniel Cardona, Joan Esteban, Roger Guesnerie, Pierre Pestieau and Thierry Verdier for helpful comments. The usual caveat applies.

I. INTRODUCTION

The role of income taxes in achieving equity objectives has been highlighted in most of the economic analysis. However, there is a controversy about the degree of progressivity the income tax should have. High rates may affect incentives to work and may therefore imply efficiency losses.

The modern setup for analyzing the equity-efficiency tradeoff using a general nonlinear income tax was built by Mirrlees (1971). It is a static optimal income tax model that defines the tax schedule as a function of three key ingredients: the distribution of productivity in the population[1], individual preferences for consumption and leisure, and the shape of the social welfare function. The optimal tax schedule then maximizes this social welfare functions under some government budget constraint and the constraints born from the labor supply behavior of agents. From this model (and its extensions) we learned that the lower is the elasticity of labor supply and the higher is the aversion to inequality of the government, the higher is the optimal marginal tax rate to be imposed on a given class of abilities/productivities. We know also that, if the distribution is bounded, the optimal marginal tax rate on the highest-productivity agent is zero [Seade (1982)]. The general intuition is the following: the nonlinear taxation of income (implemented for redistributive purposes) has a negative effect on the incentives to carry out some activities that, while contributing to global welfare, also involve a disutility for the individual. The optimal tax schedule balances the efficiency loss of a marginal increase in taxation with the equity gains obtained by a high level of redistribution financed with the taxes paid by the most productive agents.

Since the work of Mirrlees, the theory of optimal income taxation has been considerably developed (see Tuomala 1990). Most of these contributions, however, do not take into account the dynamic evolution of the economy and its effects on redistribution policies. Economies evolve due to the interplay of multiple factors. In particular, the productivity of the agents may evolve over time due to mechanisms that can be considered endogenous to individual decisions (e.g. education decisions or learning by doing mechanisms) or exogenous (e.g. aggregate productivity increases or salary indexation to inflation).

The effect of income taxes on human capital formation has been extensively analyzed [see, for instance, Heckman et al. (1998)]. This literature attempts to assess the effects of a given tax code on human capital accumulation, but does not provide the normative prescription of how an optimal income tax schedule should look like. On the other hand, several papers have adopted the normative approach of Mirrlees (1971) to incorporate education in the analysis, but have remained static. They typically consider a one shot investment decision in education [for a review, see Tuomala(1990) or Brett and Weymark (2003)]. More recently, Kapicka (2004) has attempted a dynamic description of an optimal labor income tax problem with education decisions, but the analysis is limited to the steady state environments, leaving aside the characterization of the optimal fiscal policy all along the transition.

Labor economics have a strong tradition in the analysis of learning by doing as a mechanism of human capital accumulation. Studies by Willis (1986), Altug and Miller (1998) or Cossa, Heckman and Lance (2000) find a significant effect of past work experience on current wage earnings. It is also well documented that displaced workers suffer important wage losses (Jacobson, LaLonde and Sullivan (1993)) and that wage profiles are affected by job tenure (Topel (1991)). These findings from microeconomic data suggest that workers experience systematic changes in labor productivity that are related with their labor experience. It is reasonable to assume that, as labor supply depends on the tax schedule, this evolution is likely to be endogenous to the redistribution policy.

The bulk of the dynamic taxation literature has studied the optimal taxation of labor income and wealth under the assumption that the government is constrained to use linear taxes (see Chari and Kehoe (1999) for a review). Non-linear taxation (with asymmetric information) has been considered in Diamond and Mirrlees (1978), Roberts (1984), Brito et al. (1991), Werning (2002), Golosov, Kocherlakota and Tsyvinsky (2003), Golosov and Tsyvinski (2003), Battaglini and Coates (2003), Berliant and Ledyard (2005), Kocherlakota (2005), Albanesi and Sleet (2005) and Conesa and Krueger (2005). All these papers make different assumptions about the dynamic evolution of individuals' unobservable types. They assume that either the abilities do not evolve over time [Roberts (1984), Brito et al. (1991) and Berliant and Ledyard (2005)], or that abilities evolve following a stochastic process completely exogenous to the fiscal policy [Diamond and Mirrlees (1978), Werning (2002), Golosov, Kocherlakota and Tsyvinsky (2003), Golosov and Tsyvinski (2003), Battaglini and Coates (2003), Kocherlakota (2005), Albanesi and Sleet (2005) and Conesa and Krueger (2005)].

To our knowledge, there are no papers trying to extend the optimal nonlinear labor income tax problem to a dynamic setting with evolving productivities where the evolution is due to learning by doing. This is different from education decision mechanisms, because the accumulation of human capital is given by the labor supply of each agent instead of being given by his/her schooling time.

In this paper, we tackle this specific problem following the mechanism design approach of Mirrlees (1971) where the informational asymmetries preclude the use of non-distortionary taxation. Thus, we make no a priori restrictions on the type of taxes the government may use. The objective of our work is to shed light on the dynamic aspect of the trade off equity-efficiency, leaving aside the insurance aspect related to taxation; for this reason we do not model stochastic environments.

We consider a dynamic economy where the productivities of the agents evolve over time, and we model the problem of a government averse to inequality seeking to find the tax schedule that optimally solves the trade off among equity (i.e. a reduced income inequality) and efficiency (a higher income per capita) in a finite horizon model in which agents have private information on their initial endowment of productivity. In addition, the government must satisfy an aggregate intertemporal budget constraint. The optimal tax code is announced at the beginning of the game and the government commits himself to respect it all along the periods (from Roberts (1984) and Berliant and Ledyard (2005), we know that in such case, separation of types will occur in each period). Agents choose in the first period, for a given tax schedule, the consumption and labor supply patterns that maximize their lifetime utility function. In the model, we consider that the evolution of the productivity depends on two factors: an exogenous component and the effort of the agent. The inclusion of the exogenous component allows us to explore the implications on optimal tax policy of aggregate changes of labor force productivity due to technological progress or to institutional mechanisms, often observed in reality, like indexation of wages to inflation, etc. The endogeneity of the second mechanism to the fiscal policy derives from the fact that the optimal individual effort in each period depends on the tax schedule. For the sake of simplicity, we choose the simplest form (linear) of wage evolution that incorporates wage increases due to learning-by-doing.

In such a setting, we determine the first order condition to be satisfied by the optimal tax schedule. The analysis of this first order condition reveals that the introduction of the endogenous evolution of the productivities adds several important insights to the optimal labor income tax problem with respect to the static Mirrlees framework. In our case, the classical static equity-efficiency trade off is modified by the dynamic effects (past and future) of taxation. In particular we find that the optimal tax strategy balances the social marginal costs of increasing marginal tax rates in a given instant t on a given class of productivities with the social marginal benefits of doing it. The costs are related with the reduction of both past and future human capital accumulation, with the negative impact on aggregate social welfare due to the reduction of the individual utility of all the agents paying more taxes at t and with the increase of the necessity to redistribute more in the future (given that the spread among social marginal weights of each agent will be higher). This last negative effect is due to concavity of the social welfare function. The benefits derive from the increase of the instantaneous tax receipts and then the reduction of the overall tax burden all along the τ periods (evaluated in terms of social welfare) and from the reduction of present inequality of incomes and future inequality of the productivities.

The structure of the paper is the following. First we state the problem of interest. Second we characterize the optimal behavior of agents. Third we study the general solution of the social planner problem. In sections 5, 6, 7 and 8 we analyze particular cases allowing to better understand the nature of the dynamic equity-efficiency trade off. Last section presents our conclusions.

II. STATEMENT OF THE PROBLEM

We consider a dynamical model where the productivity of each individual is evolving over time. A particular individual is endowed with a productivity $w_o \epsilon [0, \infty[$ at the initial stage, and from there on his productivity evolves over time according to

$$\partial_t w(w_o, t) = [a(t) + bL(w_o, t)]\, w(w_o, t) \qquad \text{subject to } w(w_o, 0) = w_o \ , \tag{1}$$

where $a(t)$ is the exogenous contribution to the wage increase, $L(w_o, t)$ is the work (supposed to take values in the interval $[0, 1]$) supplied by this individual at time t and $bL(w_o, t)$ is the rate of increase of the individual's wage due to the learning-by-doing process.

We consider that there is a continuum of individuals and we call $f(w_o) \geq 0$ the initial distribution of productivities. We assume that $\int_0^\infty dw_o f(w_o) = 1$ and that the moments of any order n of the distribution are finite - i.e. $\int_0^\infty dw_o w_o^n f(w_o)$ finite. Moreover we consider that both agents and government are aware of the rule of evolution of the wages and, in particular, the learning by doing mechanism; however, the initial productivity w_o of each agent is private information not available to the government, who can only know $f(w_o)$

In this framework, a social planner seeks to maximize the finite-horizon social welfare function

$$Max_{T()}\ S = \int_0^\tau \ dt e^{\alpha t} \int_0^{+\infty} dw_o f(w_o) G(U(w_o, t)) \ , \tag{2}$$

where α is the time preference parameter of the government. It can be positive, null or negative [2]. A positive value means that the social planner gives a higher weight to the future social welfare. On the contrary, a negative α implies a government more prone to the present welfare.

$G()$ is the weight function that transforms individual instantaneous utility, $U()$, into social welfare. In order to model a government averse to inequality we assume that $G'() \geq 0$ and $G''() \leq 0$, where primes denote derivatives of the function with respect to its argument, i. e., $h'(z) = dh/dz$. The more negative is $G''()$, the higher is the aversion of the social planner to inequality[3].

The above maximization is constrained by an exogenous budgetary restriction

$$R = \int_0^{\tau} dt e^{\beta t} \int_0^{+\infty} dw_o f(w_o) T(w(w_o,t)L(w_o,t),t) \tag{3}$$

that represents the non redistributive public spending to be financed all along the τ periods. β is the opposite of the exogenous interest rate at which the government can lend or borrow money in the financial market.

The budgetary restriction affects the social welfare because $U(w_o,t)$ is affected by the tax schedule $T(wL,t)$. We consider that individual instantaneous utilities depend only on consumption and effort, being increasing and concave with respect to consumption and decreasing, strictly concave and twice continuously differentiable with respect to effort. For simplicity (as it is often the case in this type of literature[4]), we will restrain our analysis to the class of individual instantaneous utility functions which are quasi-linear in consumption,

$$U(w_o,t) = C(w_o,t) - B(L(w_o,t)) \;, \tag{4}$$

where $C(w_o,t)$ is consumption and $B(L)$ denotes the loss of utility associated to the supply of L labor units. We also impose that individuals (as in Mirrlees) do not accumulate any wealth and, unlike the government, they do not have access to the credit market. To avoid having to worry about corner solutions, we assume that $B'(0) = 0$ and that $B'(1) = B(1) = \infty$. Since agents consume all their net income, $C(w_o,t) = w(w_o,t)L(w_o,t) - T[w(w_o,t)L(w_o,t),t]$ thus

$$U(w_o,t) = w(w_o,t)L(w_o,t) - T[w(w_o,t)L(w_o,t),t] - B(L(w_o,t)) \;, \tag{5}$$

Therefore, the social planner must appropriately choose (at time 0) the profile of instantaneous income tax function $T(x,t)$ that maximizes the social welfare S while respecting the budgetary restriction.

Hereafter, for alleviating the notation we shall not explicitly write the dependencies on w_o and t for every term in a given function, but instead we shall use the shorthand $[x]_{w_o,t}$ to indicate for which individual and period the function should be computed, e. g.,

$$U(w_o,t) = [wL - B(L) - T(wL,t)]_{w_o,t} \equiv w(w_o,t)L(w_o,t) - B[L(w_0,t)] - T[w(w_0,t)L(w_o,t),t] \;. \tag{6}$$

Before proceeding further with the development of the model, we would like to remark that we have modelled a social planner caring about the evolution of the inequality all along the transition path. This is different with respect to a social planner caring only on the distribution of the intertemporal utilities of the agents, in which case the social welfare function should be expressed as

$$S = \int_0^{+\infty} dw_o f(w_o) G\left[\int_0^{\tau} dt U(w_o,t)\right] . \tag{7}$$

This specification for the social welfare function coincides with ours (2) only for the utilitarian case (i.e. $G()$ is linear) when the time preference parameters for the agents' and social planner are the same. In our specification, the concavity of $G()$ imposes limits to the level of inequality socially admissible in any instant, thus describing a social planner more averse to inequality.

There are several reasons justifying the adoption of such specification. First, the social welfare function retained allows for a richer analysis of the social planner intertemporal redistributive arbitrage, i.e. the optimal intertemporal allocation of the redistribution charge. Second, recent papers in the field of political economy of taxation and redistribution [see for example Piketty (1995)] show that agents (and then social planner's) attitudes toward redistribution evolves in time as a result of a mix of ideologic concerns and a learning process based on their past histories. In particular, empirical studies on this issue [see, e.g., Fong (2001), Piketty (1995, 1999), Gilens (1999, p.51), Kluegel and Smith (1986) or Konrad and Spadaro (2005)] show that the responses given by individuals, when asked for his opinions about the role of the government in reducing inequality of incomes and opportunities, have a strong correlation with the level of redistribution implemented by the government in the previous period. In our opinion this evidence is better captured by the specification we use. Finally, note also that, as compared to (7), our choice allows us to model a social planner with a time preference different from that of the individuals [who, as demonstrated in many empirical studies - see Frederick, Loewenstein and O'Donoghue (2002) - are impatient]. In our opinion this is an important issue given that nothing prevents the social planner to care more about the final periods (for instance for electoral reasons).

III. INDIVIDUAL OPTIMUM

In response to a tax schedule $T()$, agents decide how many labor units to supply at each time by optimizing their own welfare. We consider that these long-sighted agents decide their instantaneous labor supply, $L(w_o,t)$, by maximizing their accumulated utility over the same temporal horizon as the social planner,

$$Max_{L(w_o,t)} \mathcal{U}(w_o) = \int_0^{\tau} dt e^{\gamma t} U(w_o,t) . \tag{8}$$

where γ is the time preference parameter as defined in Frederick, Loewenstein and O'Donoghue (2002). An impatient individual is characterized by a large, negative γ.

Given (5), the individual optimum is thus determined by

$$Max_{L(w_o,t)}\mathcal{U}(w_o) = \int_0^\tau dte^{\gamma t}[wL - B(L) - T(wL,t)]_{w_o,t} \tag{9}$$

$$\text{subject to } \partial_t w_{w_o,t} = [(a+bL)w]_{w_o,t} \; . \tag{10}$$

This is an standard optimal control problem for each individual which can be solved by applying Pontryagin method. Defining the conjugate moment to $w(w_o,t)$, $\mu(w_o,t)$ and introducing the hamiltonian of the individual,

$$H(w_o,t) = e^{\gamma t}[wL - B(L) - T(wL,t)]_{w_o,t} + [\mu(a+bL)w]_{w_o,t} \; , \tag{11}$$

the (interior) optimal solution is determined by the system of ordinary differential equations

$$\partial_t w = (a+bL)w \tag{12}$$

$$\partial_t \mu = -e^{\gamma t} L \left[1 - \partial_x T(x,t)\right]_{x=wL} - (a+bL)\mu \tag{13}$$

$$0 = e^{\gamma t}\left\{w\left[1 - \partial_x T(x,t)\right]_{x=wL} - B'(L)\right\} + b\mu w \; , \tag{14}$$

where primes denote the total derivative with respect to the function's argument, i. e., $B'(L) = dB/dL$.

From these equations we have that

$$w(w_o,t) = w_o e^{\int_0^t ds[a(s)+bL(w_o,s)]} \; , \tag{15}$$

$$\mu(w_o,t) = e^{-\int_0^t dsa(s)} \int_t^\tau dse^{\gamma s} e^{\int_0^s dra(r)} \left[\frac{B'(L)L}{w}\right]_{w_o,s} \; , \tag{16}$$

$$w(w_o,t)\;\left[1 - \partial_{wL}T(wL,t)\right]_{w_o,t} + [bQ]_{w_o,t} = \left[B'(L)\right]_{w_o,t} \; , \tag{17}$$

where we have defined

$$Q(w_o,t) = \left[\mu w e^{-\gamma t}\right]_{w_o,t} \; . \tag{18}$$

It is interesting to note that this term (divided by $w(w_o,t)$) is the current value Lagrange multiplier associated to the human capital accumulation constraint. It represents the shadow value, expressed in term of current utility, of an increase of the stock of productivity in the future. It is decreasing in time and, at the last period, it is null for each class of productivity. Note also that very impatient individuals (γ tending to $-\infty$) have vanishing μ and Q, expressing the negligible current value of future stock of productivity.

The first order condition (17) clearly shows that the arbitrage condition driving the optimal agent's behavior depends on the comparison of the marginal cost of an increase of the instantaneous labor supply (i.e the term $B'(L)$) with the marginal net benefit of doing it. This benefit is the sum of an instantaneous benefit (i.e the net remuneration of a supplementary unit of labor supply $w(w_o,t)\,[1-\partial_{wL}T(wL,t)]_{w_o,t}$) and the marginal return of investment in human capital bQ (i.e. the shadow value, expressed in term of current utility, of the marginal increase of the stock of productivity in the future, due to instantaneous labor supply increase). It is worth noting that, given a tax schedule, the labor supply of the agent is higher when $b \neq 0$, expressing that the future wage increases incentivate the agent to make larger efforts.

IV. SOCIAL OPTIMUM

From eqs.(5) and (17), we find that the change in indirect utility among individuals at any time is

$$\frac{\partial U(w_o,t)}{\partial w_o} = \left[\frac{L}{w}\frac{\partial w}{\partial w_o}(B'(L)-bQ) - bQ\frac{\partial L}{\partial w_o}\right]_{w_o,t} , \tag{19}$$

which upon integration yields

$$U(w_o,t) = U(0,t) + \int_0^{w_o} dx \left[\frac{L}{w}\frac{\partial w}{\partial x}(B'(L)-bQ) - bQ\frac{\partial L}{\partial x}\right]_{x,t} . \tag{20}$$

The relationship between tax and utility given in eq.(5) together with eq.(20) allow us to reformulate the problem for the social planner in terms of the individual labor output only, namely

$$Max_{L(w_o,t)}\ S = \int_0^{\tau} ds e^{\alpha s} \int_0^{+\infty} dx f(x) G(U)_{x,s} \tag{21}$$

$$\text{subject to } R = \int_0^{\tau} ds e^{\beta s} \int_0^{+\infty} dx f(x)[wL - B(L) - U]_{x,s} , \tag{22}$$

where w and U are given by eqs.(15) and (20), respectively. By introducing the (constant) Lagrange multiplier λ, we can define the restricted objective functional

$$\mathcal{S} = S + \lambda R , \tag{23}$$

and then the first order condition for the social optimum can be found by direct functional derivation,

$$\frac{\delta \mathcal{S}}{\delta L(w_o,t)} = 0 . \tag{24}$$

Since we have assumed that $G'' \leq 0$, a Taylor expansion of the integrand in (21) evidences that the first order condition corresponds to a maximum.

After some algebra and using the individual first order condition (17), we obtain an interesting intermediate step that provides us with a first intuition on the mechanisms driving the optimal tax policy. At this stage, the first order condition can be written as

$$\left[\frac{\partial T(wL,t)}{\partial(wL)}\right]_{w_o,t} = Z_1 + Z_2 - Z_3 \ , \tag{25}$$

where

$$Z_1 = b\left[\mu e^{-\gamma t} - \frac{\int_t^\tau ds\ e^{\beta(s-t)}[wL]_{w_o,s}}{w(w_o,t)}\right] \ , \tag{26}$$

$$Z_2 = \frac{e^{-\beta t}}{w(w_o,t)f(w(w_o,t))}\left[\frac{\delta \int_0^\tau dse^{\beta s}\int_0^\infty dx f(x)U(x,s)}{\delta L(w_o,t)}\right] \ , \tag{27}$$

and

$$Z_3 = \frac{e^{-\beta t}}{w(w_o,t)f(w(w_o,t))}\left[\frac{\delta \int_0^\tau ds\frac{e^{\alpha s}}{\lambda}\int_0^\infty dx f(x)G[U(x,s)]}{\delta L(w_o,t)}\right] \ . \tag{28}$$

The term Z_1 captures the effect of the productivity dynamics. In particular, the term $\mu e^{-\gamma t}$ is the current value Lagrange multiplier associated to the agent's human capital accumulation constraint. It represents the benefit of loosing utility today (by working more) in terms of future stock of productivity. If this term is high, there will be a strong incentive for the individual to invest (by working more today) in human capital accumulation in order to have a higher productivity in the future. It implies that the efficiency distortion of an increase of the marginal tax rate at time t will be low (i.e. the individual will still work hard even if the net wage is reduced). On the contrary, if this term is low (as e.g. in the case of very impatient individuals) the efficiency distortions of taxation will be higher. Of course, if the endogenous part of the learning by doing process generates a strong impact of the present effort on the future productivity, the problems of dynamic efficiency will be more important. This intuition is behind the observation that the term $\frac{\int_t^\tau ds[bwL]_{w_o,s}}{w(w_o,t)}$ impacts negatively on the optimal marginal tax rate (this term corresponds exactly to future endogenous productivity increase divided by the productivity present value). It is interesting to note that, in the extreme case in which the human capital accumulation process presents strongly increasing returns to scale (in that case the term $\frac{\int_t^\tau ds\ e^{\beta(s-t)}[bwL]_{w_o,s}}{w(w_o,t)}$ is extremely high), it can be the case that the productivity accumulation constraint prevails over the equity concerns up to the point where it can be optimal to set negative marginal tax rates on the agents with high levels of productivity.

The term Z_2 is the product of the sum over time of the instantaneous average sensibility of individual indirect utility to variations of the labor supply of agents in the class w_0 and the factor $\frac{e^{-\beta t}}{w(w_o,t)f(w(w_o,t))}$. It is related with efficiency aspects of tax problem. It captures the behavioral effects induced by an arbitrary increase of the marginal tax rate on a given class of productivities in a given instant t (evaluated by the government at $t=0$). Such a measure has several efficiency effects. First, it reduces the labor supply of people in the neighborhood of $w(w_o,t)$ because the instantaneous marginal return to their labor falls. Second, it reduces the future value of their stocks of human capital. Third, it reduces the utility of all the agents with productivity higher than $w(w_o,t)$ because they also pay additional taxes (even if both the marginal tax rate and the shadow cost are unchanged - i.e their labor supply remain unchanged). It can be easily seen that the sum of this negative efficiency effects is higher the higher is the instantaneous elasticity of labor supply and the higher is the impact of the present effort on the future productivity. In this case the term $\left[\frac{\delta\int_0^\tau dse^{\beta s}\int_0^\infty dxf(x)U(x,s)}{\delta L(w_o,t)}\right]$ is lower.

The net increase of the tax receipts —i.e. the loss on the "$w(w_o,t)$" agents and the subsequent increase on the "$1-F(w_o)$" agents with higher productivities— is then redistributed via a lump sum subsidy to everybody in any instant because the intertemporal budget constraint must be maintained unchanged. This operation has a positive social welfare effects if the weight of the poorest agents is higher than the richest (i.e. the government has redistributive tastes). The term Z_3 captures these effects. It is the product of the sum over time of the instantaneous average social welfare sensibility to variations of the labor supply of agents in the class w_0 and, again, the factor $\frac{e^{-\beta t}}{w(w_o,t)f(w(w_o,t))}$. With an increasing and concave social welfare function, a low productivity agent weights more (at the margin) than an high productivity agent, then, it can be easily shown that the term Z_3 decreases with the productivities, and, as it impacts negatively on the marginal tax rate, the lower is the productivity at time t of an agent, the higher is the contribution to the social welfare of a change of his labor supply, the lower should be his optimal marginal tax rate.

In both Z_2 and Z_3, the term $\frac{e^{-\beta t}}{w(w_o,t)f(w(w_o,t))}$ takes into account the number of agents in each class of productivities $w(w_0,t)$ and discounts to time t the effects previously described.

It is interesting to note that, in a purely static framework we have that (see appendix)

$$\frac{\delta U(x,s)}{\delta L(w_o,t)} = \frac{(LB')'}{w_o}\theta(x-w_o)\delta(s-t) \tag{29}$$

and therefore,

$$Z_2 - Z_3 = \frac{B' + LB''}{(w_o)^2 f(w_o)} \int_{w_o}^{\infty} dxf(x)\left[1 - \frac{G'(U)}{\lambda}\right] . \tag{30}$$

Since $Z_1 = 0$, the first order condition (25) then reduces to

$$\frac{\partial_{w_o L} T(w_o L)}{1 - \partial_{w_o L} T(w_o L)} = \left[\frac{B'(L) + LB''(L)}{B'(L)}\right]\left[\frac{1 - F(w_o)}{w_o f(w_o)}\right]\left[1 - \frac{\int_{w_o}^{\infty} dx f(x) \left[\frac{G'(U)}{\lambda}\right]}{1 - F(w_o)}\right] , \qquad (31)$$

hence recovering what has been found by Mirrlees in his seminal paper. The intuition in the static framework is that the higher is the aversion to inequality of the social planner the higher should be the marginal tax rate on a given class of income. The second message concerns the term $\frac{B'(L)+LB''(L)}{B'(L)}$. It is a measure of the inverse of the sensibility of labour/effort supply to changes in net wages. The higher is the sensibility of labour/effort supply to changes in net wages term the lower is the marginal tax rates can be implemented. The difference with our first order condition is that the efficiency and equity concerns are simply static, meanwhile, in our case, the government should also consider the impact of the instantaneous taxation both on the past and the future evolution of the wage inequality.

The complete characterization of the solution of this general problem is not easy to identify. As we can see, the complete development of the first order condition (see Appendix), give us a very complex relation:

$$\begin{aligned}
\left[\frac{\partial T(wL,t)}{\partial (wL)}\right]_{w_o,t} &= b\left[\mu e^{-\gamma t} - \frac{\int_t^{\tau} ds\ e^{\beta(s-t)}[wL]_{w_o,s}}{w(w_o,t)}\right] + \\
&+ \left[\frac{(B'(L)L)' - bQ}{w}\right]_{w_o,t} \frac{1 - F(w_o)}{w(w_o,t) f(w_o)}[1 - S(w_o,t)]\left[\frac{\partial w}{\partial w_o} - W\right]_{w_o,t} \\
&+ \frac{b}{w(w_o,t)} \int_t^{\tau} ds\ e^{\beta(s-t)} \times \left\{Q\left[\beta - \gamma - \frac{\alpha-\gamma}{\lambda} e^{(\alpha-\beta)s} G'(U) - \frac{e^{(\alpha-\beta)s}}{\lambda} G''(U)\partial_s U\right]\right\}_{w_o,s} \\
&+ b\frac{1 - F(w_o)}{w(w_o,t) f(w_o)} \int_t^{\tau} ds\ e^{\beta(s-t)} \left\{\frac{\partial Q}{\partial w_o}[\gamma - \beta + (\alpha - \gamma + P)S]\right\}_{w_o,s} \qquad (32)
\end{aligned}$$

where we have defined

$$F(w_o) = \int_0^{w_o} dx f(x) \qquad (33)$$

$$S(w_o,t) = \frac{e^{(\alpha-\beta)t}}{\lambda} \frac{1}{1 - F(w_o)} \int_{w_o}^{\infty} dx f(x) G'(U(x,t)) \qquad (34)$$

$$P(w_o,t) = \frac{\int_{w_o}^{\infty} dx f(x) G''(U(x,t))\partial_t U(x,t)}{\int_{w_o}^{\infty} dx f(x) G'(U(x,t))} \qquad (35)$$

$$W(w_o,t) = \frac{1}{1 - S(w_o,t)} \int_0^t ds\ e^{A(t-s)}[1 - S(w_o,s)]\left[\frac{\partial (bwL)}{\partial w_o}\right]_{w_o,s} \qquad (36)$$

$$A(t-s) = \int_s^t dr\, [\gamma - \beta + a(r)] \ . \qquad (37)$$

More useful insights about the nature of the solution can be found analyzing particular cases.

V. NO WAGE EVOLUTION

This is the simplest possible case, the only evolution of the optimal tax rate being dictated by the discount rates of the social planner, α, and the financial market, β. With respect to a purely static problem, the possibility to finance redistribution a time t with loans (the government budget constraint holds for the whole sequence of periods) influences the optimal tax rate at period t. In this framework, we have that $a(t) = 0 = b$, thus $w(w_o, t) = w_o$ and we find that

$$\frac{\partial_{w_o L} T(w_o L, t)}{1 - \partial_{w_o L} T(w_o L, t)} = \frac{1 - F(w_o)}{w_o f(w_o)} [1 - S(w_o, t)] \left[\frac{B'(L) + L B''(L)}{B'(L)} \right]_{w_o, t} . \tag{38}$$

This result is formally equivalent to that by Mirrlees, although in our case there is temporal evolution of $S(w_o, t)$. This implies that the average redistribution is not constant over time. In particular, if the discount rate of the social planner is higher than the interest rate (in such a case $\alpha > \beta$) then it is optimal to finance the present redistribution with debt and to postpone its payment by rising taxes in the future. On the contrary, if the discount rate of the social planner is lower than the interest rate ($\alpha < \beta$) then it is optimal to increase the fiscal pressure at early stages and finance future redistribution with the interests generated by the excess of tax receipts collected. Even though we have $a = b = 0$, in both cases —see eq. (17)— the labor supply of the agents will become time-dependent due to the time-dependent tax schedule: in the first case, L will decrease in time in response to the increasing tax rate, and viceversa. When $\alpha = \beta$, there is no possibility to play strategically with the debt in order to finance redistribution, thus the problem becomes purely static and the optimal tax schedule corresponds to a ”repeated Mirrlees”.

VI. EXOGENOUS WAGE EVOLUTION

In this case, $b = 0$, and we have a marginal tax rate which is formally equivalent to that found before,

$$\left[\frac{\partial_{wL} T(wL, t)}{1 - \partial_{wL} T(wL, t)} \right]_{w_o, t} = \frac{1 - F(w_o)}{w(w_o, t) f(w_o)} [1 - S(w_o, t)] \left[\frac{B'(L) + L B''(L)}{B'(L)} \right]_{w_o, t} , \tag{39}$$

although in this case it must be recalled that, besides the time dependence of $S(w_o, t)$, we now have a time dependence of $w(w_o, t)$. The labor supply $L(w_o, t)$ depends on these two factors. The optimal tax schedule depends on the same key elements as in the static Mirrlees case: the higher the sensibility of labor supply to changes in net wages, the lower the optimal tax rate in any period. At the same time, the higher the aversion to inequality, the higher the marginal tax rate on a given level of productivity, independently on the evolution they suffer.

A first interesting remark is that the optimal tax rate in this case is lower than in the previous one (no wage evolution). To see it, it is sufficient to replace $w(w_o, t)$ in (39) with the definition given in (15). We can then rewrite (39) as

$$\left[\frac{\partial_{wL} T(wL,t)}{1-\partial_{wL} T(wL,t)}\right]_{w_o,t} = \frac{1-F(w_o)}{w_o f(w_o) e^{\int_0^t ds[a(s)]}}[1-S(w_o,t)]\left[\frac{B'(L)+LB''(L)}{B'(L)}\right]_{w_o,t} . \quad (40)$$

The intuition is straightforward: the presence of the exogenous evolution determine an increase of the total output of the economy. With a fixed intertemporal budget constraint, it allows the social planner to reduce at any instant (with respect to a situation with no wage evolution) the average fiscal pressure. This effect is higher the higher is the exogenous component $a(t)$. Another immediate implication of this result is that the optimal marginal tax rate on a given class of agent is (in this case) decreasing over time.

Another important feature is that - as shown in this section and in the previous one - in a dynamic economy where productivities do not evolve or where they evolve exogenously, the government will commit on a sequence of optimal tax rates $T[(wL)_{w_o,t}, t]$ that follows the same rules that in Mirrlees. In particular, in these cases we also recover the very well known result of Seade (1982) about the optimality of a zero marginal tax rate on the agent with the highest productivity.

The only difference is the possibility to "play" strategically with the debt (i.e the average level of redistribution in each period): in economies with high interest rates, the social planner will prefer to increase the fiscal pressure at the beginning of the "game" and to relax it near the end, financing the redistribution with the public wealth accumulated, and viceversa.

On this subject, Bassetto and Kocherlakota (2004) show that, in a dynamic stochastic optimal tax problem, when taxes can depend on past incomes the intertemporal structure of the debt is irrelevant for the optimal allocation of resources in the economy. We have shown that if optimal taxes are levied on instantaneous income but depend on the entire intertemporal horizon this statement does not hold.

VII. ENDOGENOUS EVOLUTION: EQUAL DISCOUNTS FACTORS

Suppose that the time preference parameter of the government is equal to the discount factor of the agents and to the interest rate (i.e. $\alpha = \beta = \gamma$) then we can rewrite the first order condition (32) as:

$$\left[\frac{\partial T(wL,t)}{\partial (wL)}\right]_{w_o,t} = Z1 + M1 + M2 + M3 + M4 \quad (41)$$

with

$$M1 = \left[\frac{(B'(L)L)' - bQ}{w}\right]_{w_o,t} \frac{1 - F(w_o)}{w(w_o,t)f(w_o)}[1 - S(w_o,t)] \left[\frac{\partial w}{\partial w_o}\right]_{w_o,t} \tag{42}$$

$$M2 = -\left[\frac{(B'(L)L)' - bQ}{w}\right]_{w_o,t} \frac{1 - F(w_o)}{w(w_o,t)f(w_o)} \int_0^t ds\; e^{A(t-s)}[1 - S(w_o,s)] \left[\frac{\partial(bwL)}{\partial w_o}\right]_{w_o,s} \tag{43}$$

$$M3 = \frac{b}{w(w_o,t)} \int_t^\tau ds\; e^{\beta(s-t)} \left[-Q\frac{\partial_s G'(U)}{\lambda}\right]_{w_o,s} \tag{44}$$

$$M4 = \frac{b}{w(w_o,t)f(w_o)} \int_t^\tau ds e^{\beta(s-t)} \left[\frac{\partial Q(w_o,s)}{\partial w_o} \int_{w_o}^\infty dx f(x) \partial_s \frac{G'(U(x,s))}{\lambda}\right] \tag{45}$$

In this case it is possible to have a clearer intuition of the elements that define the optimal marginal tax rate on a given class in a given instant.

The term $Z1$ is the same as in (26).

The term $\frac{\partial w}{\partial w_o}$ in (42), gives us a local measure (in the distribution) of the degree of inequality of the productivities at time t compared with his initial level. The higher it is, the higher will be the social benefit of an increase of the marginal tax rate on $w(w_o,t)$. This effect is reinforced if the number of agents that, as a consequence of that increase, will pay more taxes (i.e. $1 - F(w)$) is high. Of course, if this part of the population has a high weight (at the margin) in the social welfare (the term S), the social marginal benefits of an increase of the marginal tax rate in a given instant will be low. This last effect implies a lower optimal marginal tax rate.

The term (43) opposes the previous term (42) by considering how much of the present degree of inequality in the productivities (and its impact on the social welfare) is due to the past individual effort. In fact, the term $d(bwL)/dw_o$ corresponds to the local variation in the contribution of the learning-by-doing process to the evolution of the individual wage. An increase of the marginal tax rate at the instant t will induce agents in the neighborhood of w_o to decrease their accumulation of human capital in the period from 0 to t (such an effect arises from the long-sightedness of the agents). If the evolution of the wages has been driven mainly by the individual effort trajectories (i. e., the term $d(bwL)/dw_o$ is high), then this effects is much more important and it is optimal for the government to set a lower marginal tax rate. Of course, this negative efficiency effect is mitigated (as in (42)) by the marginal social benefits of an increase of the tax burden on the $1 - F$ richest part of the population.

The term (44) gives us the impact of the future evolution of the productivity of agents in the class $w(w_o,t)$ on their social marginal weights. The higher are the actual wage w and the actual marginal return of investment in human capital μ, the higher will be the increase of the productivity in the future. As a consequence, there will be an increase of the utility in the future and, then, a decrease of the social marginal weight ($\frac{\partial_s G'[U]}{\lambda}$) of agents in the class $w(w_o, t)$. The social planner anticipates this effects by increasing (at the margin) the marginal tax rate on this class of productivity. Of course, the importance of such effect will diminish when t approaches τ.

If individual utilities increase over time, the concavity of $G(U)$ implies that it is more efficient to redistribute more (at the margin) tomorrow instead of today given that the spread among social marginal weights decreases over time. The social planner can enhance this effect by taxing less at time t, which allows to increase the accumulation of human capital. This effect is higher the higher is the differential evolution of the marginal returns with respect to the initial productivity (i.e. the term $\frac{\partial Q(w_o,s)}{\partial w_o}$). This is the mechanism captured by (45). Indeed, it impacts negatively on the marginal tax rate (note that $\partial_s G'(U) = G''(U)\partial_s U < 0$). Note also that, as in (44), the importance of this term is decreasing in time.

As we have highlighted in the general case, the optimal tax strategy balances social marginal costs of increasing marginal tax rates in a given instant on a given class of productivities with social marginal benefits of doing it. The costs are: the reduction of both past and future human capital accumulation, the reduction of social welfare due to the reduction of the individual utility of all the agents paying more taxes at t (i.e. $1 - F(w_o)$) and the increase of the necessity to redistribute more in the future because the effect in (45). The benefits are: the increase of the tax receipts at t and then the reduction of the overall tax burden all along the τ periods (evaluated in terms of social welfare), the reduction of present (at t) inequality of incomes and future inequality of the productivities. The analysis performed in this section clearly shows that the introduction of the endogenous evolution of the productivities adds several important insights to the optimal labor income tax problem with respect to the static Mirrlees framework. In what follow we analyze the implications for the optimal taxation on the class of individuals with highest productivity.

VIII. ENDOGENOUS EVOLUTION: THE HIGHEST PRODUCTIVITY INDIVIDUAL

The last (if any) individual in the distribution is characterized by an initial wage Ω_o such that $F(\Omega_o) = 1$. Then, his/her marginal tax rate at every time period is determined as

$$\left[\frac{\partial T(wL,t)}{\partial wL}\right]_{\Omega_o,t} = \frac{b}{w(\Omega_o,t)}Q(\Omega_o,t) - \frac{b}{w(\Omega_o,t)}\int_t^\tau dse^{\beta(s-t)}[wL]_{\Omega_o,s}$$
$$+ \frac{b}{w(\Omega_o,t)}\int_t^\tau dse^{\beta(s-t)} \times Q(\Omega_o,s)\left[\beta-\gamma-\frac{\alpha-\gamma}{\lambda}e^{(\alpha-\beta)s}G'(U)-\frac{e^{(\alpha-\beta)s}}{\lambda}G''(U)\partial_s U\right]_{\Omega_o,s} \quad (46)$$

Interestingly, we see that the marginal tax rate of the last individual of the wage distribution does not necessarily be null, as in the static Mirrlees case. This holds only in the last period, $t = \tau$, where we have $Q(x,\tau) \equiv 0$ for any wage class.

Instead, during the time interval considered the marginal tax rate over the last individual will in general evolve in time and it can be either null, positive or negative depending on the balance of the different contributions.

In the simplest case where $\alpha = \beta = \gamma$, from eq.(46) we can see that the marginal tax rate will be positive, null or negative depending on the following condition:

$$-\int_t^\tau ds\left[\mu w\frac{\partial_s G'(U)}{\lambda}\right]_{\Omega_o,s} \gtreqless \int_t^\tau dse^{\beta s}[wL]_{\Omega_o,s} - [\mu w]_{\Omega_o,t}\,. \quad (47)$$

The meaning of this condition is the following: a positive marginal tax rate reduces the accumulation of human capital and then the future income production of the richest agent (i.e. the term $\left[\int_t^\tau dse^{\beta s}[wL]_{\Omega_o,s}\right]$). This reduction is smaller the higher is the marginal return to the human capital (i.e. the term $[\mu w]_{\Omega_o,t}$). On the other side, it reduces his future contribution to productivity inequality and then his social marginal weight (i.e. the term $\int_t^\tau ds\left[\mu w\frac{\partial_s G'(U)}{\lambda}\right]_{\Omega_o,s}$). If this last effect is higher than the first, it will be optimal for the government to set a positive marginal tax rate. On the contrary, if the efficiency effect dominates, it will be optimal to subsidize the highest productivity individual in order to allow for a faster increase of his stock of human capital and then his total output.

IX. CONCLUSIONS

In this paper, we have extended the optimal nonlinear labor income tax problem [Mirrlees (1971)] to a dynamic setting with evolving productivities where the evolution depends both on an exogenous mechanism and on the effort of the agent. The endogeneity is given by the dependence of the optimal individual effort in each period on the tax schedule. In particular, we have analyzed the problem of a government averse to inequality seeking to find the tax schedule that optimally solves the trade off among equity and efficiency in a finite horizon model in which agents have private information on their initial endowment of productivity. In addition, the government must satisfy an aggregate intertemporal budget constraint. The optimal tax code is announced at the beginning of the game and the government commits himself to respect it all along the periods. Agents choose in the first period, for a given tax schedule, the consumption and labor supply patterns that maximize their lifetime utility function.

In such a setting, the classical static equity-efficiency trade off is modified by the dynamic effects (past and future) of taxation. In particular, we find that the optimal tax strategy balances social marginal costs of increasing marginal tax rates in a given instant on a given class of productivities with social marginal benefits of doing so. The costs are related with the reduction of both past and future human capital accumulation, with the negative impact on aggregate social welfare due to the reduction of the individual utility of all the agents paying more taxes and with the increase of the necessity to redistribute more in the future (given that the spread among social marginal weights of each agents will be higher). The benefits derive from the increase of the instantaneous tax receipts with the consequent reduction of the overall tax burden all along the time horizon (evaluated in terms of social welfare) and from the reduction of present inequality of incomes and future inequality of the productivities.

There are several interesting implications of our results. First, the taxation implemented for redistributive purposes should consider the positive effects of labor supply on the human capital accumulation process. On one side, agent's incentive to work hard in order to accumulate future productivity reduces the distortionary effects of taxation on instantaneous labor supply (when agents are impatient this incentive is less effective). On the other side, the need of high levels of per capita output production pushes the social planner to reduce the fiscal pressure in a given instant. In the extreme case in which the mechanism of human capital accumulation presents strongly increasing returns to scale with respect to labor, the productivity accumulation constraint prevails on the equity concerns and it can be optimal to set negative marginal tax rate on agents with high productivity.

Second, the marginal tax rate of the last agent of the wage distribution does not necessarily be null, as in the static Mirrlees case. This holds only in the last period. Instead, during the time interval considered it will in general evolve in time and it can be either null, positive or negative. If the social benefit of the reduction of the future productivity inequality of an increase in the marginal tax today is higher than the net efficiency cost, it will be optimal for the government to set a positive marginal tax rate. On the contrary, if the second effect dominates, it will be optimal to subsidize the highest productivity individual in order to increase faster his stock of human capital and then his total output.

Third, there is a dependence of the intertemporal optimal structure of the redistribution on the time preferences of the social planner and the market (i.e. the interest rate). In economies with high (low) interest rates, the social planner will prefer to increase (decrease) the fiscal pressure at the beginning of the "game" and to relax (increase) it near the end, financing the redistribution with the public wealth (public debt) accumulated.

Fourth, in periods in which the evolution of the wages is strongly driven by exogenous (to individual effort) factors, it is optimal for the social planner to set lower marginal tax rates than in periods in which the wages remain constant.

The analysis performed in this paper clearly shows that the introduction of the endogenous evolution of the productivities adds several important insights to the optimal labor income tax problem with respect to the static Mirrlees framework.

To our knowledge, this is the first attempt to analyze this type of dynamic optimal tax model. In the work presented, there are several limitations and simplifying assumptions of which we are aware. Three of them are, in our opinion, particularly relevant. First, the results are limited to the specific class of individual utility function considered. Second, the complexity of the problem treated makes very difficult to achieve a complete analytical characterization of the solutions; numerical computation of such type of problems is also very complex and demanding, which justify an entire paper devoted to it. Third, more general forms for the learning-by-doing mechanism could be considered which are nonlinear to effort, in order to reflect the plausible assumption that the rate of wage increase saturates as the labor supply increases. Given that the wage increase due to the learning-by-doing mechanisms constitutes an incentive to increase the effort of the agents, concavity on the effort implies that individuals with high productivities will be lesser impelled to increase their labor supply. As a consequence, the optimal marginal tax on these individuals is expected to be lower than in the present case. Along this line, it could be also interesting to consider that different wages increase at different rates as a result of, e.g., syndical agreements.

These more realistic considerations do not qualitatively modify our main conclusions but they add increasing complexity to the problem.

Future work should address these issues jointly with the exploration of more fundamental issues like the introduction of uncertainty (on future productivities) and the modelling of alternative human capital accumulation processes (for example education).

In our opinion, extensions of this model may also be useful in several other fields of economic research.

The more intuitive is the analysis of any principal-agent problem where the agent's type evolves over time endogenously with his effort (and the principal commits himself to respect the contract offered in the first period). In that case, the principal's problem is similar to our social planner problem. He must offer a menu of contracts that are incentive compatible and that consider the positive effects of individual effort on future productivity.

A second interesting issue could be the exploration of the effects on inequality and welfare of international trade. There is a wide theoretical and empirical literature examining the relationship between observed changes in skill premium and international trade. Most of these works [see Acemoglu (2003), Thoenig and Verdier (2003) and Slaughter (2000)] show that countries having opened their economies to trade have experienced increases in wage differentials. In our model, it is relatively easy to incorporate these changes in the wage evolution equation trough the exogenous parameter $a(t)$. It is then possible to evaluate the impact of the trade openness on the optimal redistribution policy and on the final income inequality.

Other possible applications relate to the analysis of optimal redistribution among different regions. Instead of thinking about workers, each characterized by an initial level of productivity, we can analyze models in which there are regions with a different capacity to transform input into output, with imperfect (or impossible) interregional factor mobility (in particular labor) and where there is an aggregate human capital accumulation process. We think that a model of this type may give us a clear picture of the instruments at the social planner's disposition to play optimally with the static and dynamic (positive and negative) effects arising from redistribution among different regions. For example, it would be possible to evaluate the extent to which market distortions and imperfections are worsened or improved by alternative redistribution policies [as in Benabou (2002)].

Institut Mediterrani d'Estudis Avançats, IMEDEA (CSIC-UIB), C/ Miquel Marqués, 21, 07190 Esporles, Spain, e-mail: salvador@imedea.uib.es.

and

PSE Paris-Jourdan Sciences Economiques (Joint Research Unit CNRS-EHESS-ENPC-ENS) Paris and Universitat de les Illes Balears, ctra Valldemossa km 7.5, 07122, Palma de Mallorca, Spain, e-mail: amedeo.spadaro@uib.es

APPENDIX

Here we detail the main steps and intermediate results needed to deduce the first order condition eq.(32). The starting point is eq.(24), so that

$$\begin{aligned} 0 \;=\; & \frac{\delta \mathcal{S}}{\delta L(w_o,t)} = \int_0^\tau ds \int_0^\infty dx f(x) \left[e^{\alpha s} G'(U) - \lambda e^{\beta s} \right]_{x,s} \frac{\delta U(x,s)}{\delta L(w_o,t)} \\ + & \int_0^\tau ds \int_0^\infty dx f(x) \lambda e^{\beta s} \left\{ \left[w - B'(L) \right]_{x,s} \frac{\delta L(x,s)}{\delta L(w_o,t)} + \frac{\delta w(x,s)}{\delta L(w_o,t)} L(x,s) \right\} . \end{aligned} \tag{48}$$

In the first place, we recall that

$$\frac{\delta L(x,s)}{\delta L(w_o,t)} = \delta(x-w_o)\delta(s-t) \;, \tag{49}$$

where $\delta(x)$ denotes Dirac's delta distribution. This distribution is defined as being zero for $x \neq 0$ and $+\infty$ for $x = 0$, such that

$$\int_a^b dx \delta(x) g(x) = g(0) \text{ if } a < 0 < b \;. \tag{50}$$

Hence, from eq.(15) we have that

$$\frac{\delta w(x,s)}{\delta L(w_o,t)} = b w(w_o,s) \delta(x-w_o) \theta(s-t) \;, \tag{51}$$

where $\theta(x)$ denotes Heavyside's distribution, defined as

$$\theta(x) = \int_{-\infty}^x dy \delta(y) \;. \tag{52}$$

Moreover, from eq.(5) we find, by applying the chain rule, that

$$\begin{aligned} \frac{\delta U(x,s)}{\delta L(w_o,t)} \;=\; & \left[\frac{1}{w_o} + b \int_0^t dr \frac{\partial L(w_o,r)}{\partial w_o} \right] \left[B'(L) + L B''(L) - bQ \right]_{w_o,t} \delta(x-w_o)\delta(s-t) \\ - \; & b\theta(s-t) \frac{\partial}{\partial w_o} \left\{ \theta(x-w_o) \left[L B'(L) - bLQ \right]_{w_o,s} \right\} \\ + \; & b\delta(s-t) \frac{\partial}{\partial w_o} \left[\theta(x-w_o) Q(w_o,t) \right] \\ - \; & b \int_0^x dy \left\{ \left\{ y \left[\frac{1}{y} + b \int_0^s dr \frac{\partial L(y,r)}{\partial y} \right] + \frac{\partial L(y,s)}{\partial y} \right\} \right\} \frac{\delta Q(y,s)}{\delta L(w_o,t)} \;. \end{aligned} \tag{53}$$

In addition, from eq.(18) we also find

$$\frac{\delta Q(x,s)}{\delta L(w_o,t)} = \delta(x-w_o) \left[\theta(\tau-t) - \theta(s-t) \right] \left[(LB'(L))' - bQ \right]_{w_o,t} e^{-\gamma(s-t) + \int_t^s dr b L(w_o,r)} \;. \tag{54}$$

By substituting eq.(54) into eq.(53) and the result into eq.(24) together with eq.(51), we finally obtain the desired result eq.(32)

REFERENCES

Acemoglu, D.. (2003): "Patterns of Skill Premia", *Review of Economics Studies*, 70, 199-230.

Albanesi, S., and C. Sleet (2005): "Dynamic Optimal Taxation with Private Information", *Review of Economic Studies*, forthcoming.

Altug, S., and R. Miller (1998): "The Effect of Work Experience on Female Wages and Labour Supply", *Review of Economic Studies*, 65(1), 45-85.

Bassetto, M., and N. Kocherlakota (2004): "On the Irrelevance of Government Debt when Taxes Are Distortionary", *Journal of Monetary Economics*, 51, 299-304.

Battaglini, M. and S. Coate (2003): "*Pareto Efficient Income Taxation with Stochastic Abilities*", NBER Working Paper 10119.

Benabou, R. (2002): "Tax and Education Policies in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?", *Econometrica*, 70(2), 481-517.

Berliant, M., and J. Ledyard (2005): "*Optimal Dynamic Nonlinear Income Taxes with No Commitment*", mimeo, California Institute of Technology.

Brett, C., and J. Weymark (2003): "Financing Education Using Optimal Redistributive Taxation", *Journal of Public Economics* 87(11), 2549-2569.

Brito, D., J. Hamilton, S. Slutsky, and J. Stiglitz (1991): "Dynamic Optimal Income Taxation with Government Commitment", *Journal of Public Economics* 44, 15-35.

Chari, V.V, and P. Kehoe (1999): "Optimal Fiscal and Monetary Policy, in *Handbook of Macroeconomics*, ed. by J. Taylor and M. Woodford, Elsevier: New York.

Conesa, J.C., and D. Krueger (2005): "On the Optimal Progressivity of the Income Tax Code", *Journal of Monetary Economics*, forthcoming.

Cossa R., J. Heckman, and L. Lochner (2000): "Wage Subsidies and Skill Formation: A Study of the Earned Income Tax Credit" in *Subsidies to Low Wage Employment*, ed. by E. Phelps, New York, Russell Sage Foundation.

Diamond, P. (2003): "*Taxation, Incomplete Markets, and Social Security*", Cambridge, MIT press.

Diamond, P., and J. Mirrlees (1978): "A Model of Social Insurance with Variable Retirement", *Journal of Public Economics*, 10, 295-336.

Fong, C. (2001): "Social Preferences, Self Interest, and the Demand for Redistribution", *Journal of Public Economics*, 82, 225-246

Frederick S., G. Loewenstein, and T. O'Donoghue (2002): "Time Discounting and

Time Preference: A Critical Review", *Journal of Economic Literature*, 15, 351-401.

Gilens, M. (1999): "*Why Americans Hate Welfare: Race, Media and the Politics of Antipoverty Policy*", University of Chicago Press, Chicago.

Golosov, M., N. Kocherlakota, and A. Tsyvinsky (2003): "Optimal Indirect and Capital Taxation", *Review of Economic Studies*, 70, 569-588.

Golosov, M., and A. Tsyvinsky (2003): "*Designing Optimal Disability Insurance*", Federal Reserve Bank of Minneapolis, Working Paper n. 628.

Heckman, J., J. Lochner, and C. Taber (1998): "Tax Policy and Human Capital Formation", *American Economic Review* 88, 293-297.

Jacobson L., R. LaLonde, and D. Sullivan (1993): "Earnings Losses of Displaced Workers", *American Economic Review*, 83(4), 685-709.

Kapicka, M. (2004): "*Optimal Income Taxation and Human Capital Accumulation*", University of California Santa Barbara, mimeo.

Kluegel, J.R., and E.R. Smith (1986): "*Beliefs About Inequality: Americans' Views of What Is and What Ought to Be*", Aldine de Gruyter. New York.

Kocherlakota, N. (2005): "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation", *Econometrica*, 73(5), 1587-1622.

Konrad, K. and A. Spadaro (2005): "Education, Redistributive Taxation and Confidence", *Journal of Public Economics*, forthcoming.

Mirrlees, J. (1971): "An Exploration in the Theory of Optimal Income Taxation", *Review of Economic Studies*, 38, 175-208.

Piketty, T. (1995): "Social Mobility and Redistributive Politics", *Quarterly Journal of Economics* 110(3), 551-584.

Piketty, T. (1999): "*Attitudes toward inequality in France: Do people really disagree?*", CEPREMAP Working Paper n 9918.

Roberts, K. (1984): "The Theoretical Limits to Redistribution", *Review of Economics Studies*, 51, 177-195.

Seade, J. (1982): "On the Sign of Optimum Marginal Income Taxation", *Review of Economics Studies*, 49, 637-643.

Slaughter, M. (2000): " *What are the Results of Product-Prices Studies and What Can We Learn from Their Differences?*", in Robert Feenstra (ed.), The Impact of International Trade on Wages, NBER Conference Report, Chicago: University of Chicago Press.

Thoenig, M. and T.Verdier (2003): "A Theory of Defensive Skill-Biased Innovation and

Globalization", *American Economic Review*, 93(3):709-28.

TOPEL, R. (1991): "Specific Capital, Mobility and Wages: Wages Rise with Seniority", *Journal of Political Economy*, 99(1), 145-76.

TUOMALA, M. (1990): "*Optimal Income Tax and Redistribution*", Clarendon Press, Oxford.

WERNING, I. (2002): "*Optimal Dynamic Taxation and Social Insurance*" Doctoral Dissertation, University of Chicago.

WILLIS, R. (1986): "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions" in H*andbook of Labor Economics*, ed. by O. Ashenfelter and R. Layard, New York, Elsevier Science Publisher, Vol. 1, 525-602.

---

[1] In this literature, the concepts of productivity, ability and wage are equivalent. They are used to express the capacity of each agent to transform labor (or effort) in income. In our paper we will use them indifferently.

[2] In an infinite horizon model, this parameter has to be negative in order to insure that S is definite.

[3] Note that $G'() \geq 0$ and $G''() \leq 0$ guarantee that S is finite at all times, provided that the moments of the distribution $f(w_o)$ are finite.

[4] The attractiveness and the limits of this specification are discussed in Diamond (2003), pag. 10.