On income and capital taxation in a life cycle model with extensive labor supply

Guy Laroque\textsuperscript{1}
INSEE-CREST and University College London

Essay in honor of Angus Deaton

November 6, 2009

\textsuperscript{1}First draft, very preliminary. I have benefited from the comments of Philippe Choné, Stéphane Gauthier and Anne Laferrère and from discussions with Nicola Pavoni.
Abstract

The paper studies redistributive taxation in a stationary life cycle model with extensive labor supply. Two forms of taxation are investigated in some depth: nonlinear income taxation and linear capital taxation.

The optimal income taxation programs of the life cycle and of the static model have similar properties. The life cycle model differs in that the social weights of the dynasties depend on their permanent incomes, not on the observed taxable current income, which is an imperfect signal of the variable of interest.

A tax on saving therefore appears as a potentially useful complement to the income tax. The formula for the derivative of social welfare with respect to the tax rate is derived when financial markets are perfect. This derivative, evaluated at the point of zero capital tax, is equal to the opposite of the correlation between the social weight of the dynasties and their aggregate lifetime savings. When high permanent incomes induce high savings, a tax on capital may have redistributive value, as illustrated on a simple example.

JEL classification numbers: H21, H31.

Keywords: optimal taxation, extensive model, welfare.

Please address correspondence to Guy Laroque, laroque@ensae.fr, INSEE-CREST J360, 15 boulevard Gabriel Péri, 92245 Malakoff Cedex, France.
1 Introduction

One function of the modern state is to make more equal the lifetime welfare of individuals. While private insurance may go a long way into smoothing out the hazards of existence, preexisting differences between the members of the economy are not insurable and can only be reduced by government interventions, provided it has information and power to detect and correct these differences (Varian (1980)). With this background, optimal taxation has developed a set of tools that allow to derive the properties of tax schemes given a specific information structure. A large part of the literature has focused on the static Mirrlees intensive labor supply setup, while some dynamic features, along the lines followed by Werning (2007), are required to discuss the kind of information issue just mentioned.

The present paper considers a stationary model in continuous time with generations of consumer-workers reproducing identically. At any point in time, the typical agent has an instantaneous productivity, observable and taxable by the government if she works, and a privately known cost of going to work. She decides whether to work or not, comparing her after tax income net of cost when working with the subsistence allowance paid to the unemployed. Given her intertemporal resources and perfect financial markets, she chooses a smooth consumption profile over her lifetime.

The paper studies some of the tax arrangements that a utilitarian government might want to implement in this context. The first one is a (time invariant) nonlinear income tax. To design such a tax, it is notable that the tax authority only needs to know the joint distribution of agents’ types, productivities and opportunity costs of work, without being informed on the time path of income and consumption as people age. Indeed, it is shown that the optimal tax program has the same shape as in a static model. A major difference however is in the social weights of the agents, which here depend on their lifetime permanent incomes, while in the static model they are linked to the current incomes. We describe a set of assumptions on the fundamentals of the economy that lead to ‘well behaved’ tax schedules. Nevertheless, the redistributive scope of the tax authority is limited by the fact that its information, current income, is an imperfect signal of its ultimate objective, lifetime income. This also drives the results of Gorry and Oberfield (2008) who study the properties of the optimal tax schedule in a similar life cycle model with an intensive labor supply à la Mirrlees.

It is therefore natural to seek other information and tax tools that might complement income tax to facilitate redistribution. There is a large literature on the topic, recently surveyed in Banks and Diamond (2008). It has been suggested and experimented to base taxes on a smoothed income (the average of the past three years, say, instead of only the current one). But this does not seem to work in practice (Vickrey (1939)). Recent works, Blomquist and Micheletto (2008) and
Weinzierl (2008), have looked at the effect of having taxes depend both on income and on age. Following a different track, keeping in mind the textbook property of the life cycle model that accumulated savings serve to finance consumption during retirement, a tax on capital is a natural instrument. Note that the motivation is different from the one put forward by Saez (2002) and Diamond and Spinnnewijn (2009): these authors advocate a tax on capital based on the idea of heterogeneity in discount rates, the rich being more patient and therefore saving more than the poor. Here the consumption-income profiles of the life cycle model are the relevant underlying feature, not impatience. Indeed, as shown by Erosa and Gervais (2002) in a nonstationary Ramsey framework, the celebrated Chamley-Judd optimality of non taxation of capital does not extend to the life cycle model. In the stationary framework studied here, one can easily compute the derivative of the government social welfare function when financial markets are perfect: at the point of zero tax on capital, this derivative is equal to the opposite of the correlation between the dynastic social weights and their total savings, summed up over their life times. If the high productivity types have wage and consumption profiles who make them high savers, capital taxation is a valuable redistributive tool. An illustrative example shows that this may be the case in the presence of borrowing constraints, and that a channel for redistribution is the taxation of retirement accounts.

Capital taxation is a controversial and lively topic. The point made here is theoretical and it would be of interest, but difficult, to supplement it with a fully fledged calibration exercise, using some of the tools advocated by Judd and Su (2006). It would also be good to know whether the mechanism underlined here is at work in the empirical life cycle models of İmrohoroğlu (1998) and Conesa, Kitao, and Krueger (2009).

Finally a more ambitious theoretical undertaking would be to look for the full dynamic optimal tax schedule in the extensive labor supply model, using all the information potentially available to the government, replicating here the type of analysis that has been carried out for the intensive model à la Mirrlees, see e.g. Golosov, Tsyvinski, and Werning (2006) or Grochulski and Kocherlakota (2007). The government would keep track of the history of employment and incomes and design an optimal tax scheme depending on age, current income and this whole history. This should be the subject of further research.

The paper is organized as follows. The model and notations are set in the next section. Nonlinear income taxation is studied in Section 3, while Section 4 looks at the optimal linear tax on savings.
2 The description of the economy

Consider an economy in continuous time where all the agents have the same fixed length of life $A$. Agents differ through a characteristic $\alpha$ determined at birth. They have a concave increasing instantaneous utility function $u(c, \alpha)$, where $c$ is their non-negative consumption. At each date, labor supply $\ell$ is either zero or one. At age $a$, the typical agent has a productivity $\omega_{\alpha}(a)$ and a private cost of going to work $\delta_{\alpha}(a)$, measured in units of good. Neither labor nor the good are storable. There is no discounting and there are perfect financial markets with a zero interest rate.

For simplicity, one may think of the exogenous trajectory $(\omega_{\alpha}(a), \delta_{\alpha}(a))$, $a$ in $[0, A]$ as deterministic and known from birth. It only depends on the parameter $\alpha$. Many of the results extend to the case where this trajectory is random (so that agents with the same $\alpha$ ex post differ depending on the timing of their shocks), under the assumption of perfect insurance markets. In the absence of government intervention agent $\alpha$ chooses the consumption and labor supply profile solution of the program

$$\max \int_0^A u(c(a), \alpha) \, da$$

$$\int_0^A [c(a) - (\omega_{\alpha}(a) - \delta_{\alpha}(a))\ell(a)] \, da \leq 0. \quad (1)$$

By concavity of the utility function, agents choose a constant consumption flow. They decide to work, $\ell(a) = 1$, whenever net production $\omega_{\alpha}(a) - \delta_{\alpha}(a)$ is positive, not to work $\ell(a) = 0$, whenever $\omega_{\alpha}(a) - \delta_{\alpha}(a)$ is negative, so that at laissez-faire

$$c^*_\alpha = \frac{1}{A} \int_{\omega_{\alpha}(a) - \delta_{\alpha}(a) \geq 0} (\omega_{\alpha}(a) - \delta_{\alpha}(a)) \, da.$$

The economy is stationary. At each instant a new cohort, with the same characteristics as all the previous cohorts, is born. There is an infinitely lived government with powers of taxation that wants to redistribute welfare between the various $\alpha$ members of a cohort. A fundamental information asymmetry limits the government actions. It does not know the type $\alpha$ of the agents and it never observes the cost of going to work $\delta$. But it sees the incomes of the workers, equal to their productivities $\omega$ under suitable incentive conditions. In this setup, the paper focuses on two tax instruments. First taxation of current gross income: the results of the static model, if of any relevance, should be useful in this stationary environment. The attention therefore here will be on the limits of myopic income taxation to redistribute lifetime welfare, linked to the correlation between current and permanent income. A natural extension then is to suppose that the government can observe and tax savings. Conditions under which a positive linear taxation of savings is optimal are given.
3 Income tax

Under the stationarity assumption, the distribution of agents’ characteristics \((\omega, \delta)\) for a given type \(\alpha\) in a cross-section is identical to that faced by a new born \(\alpha\) over her life time. To study the optimal tax schedule, it is therefore simpler to describe the economy through the distributions of \((\omega, \delta)\) conditional on \(\alpha\), together with the marginal distribution of \(\alpha\), rather than with the functions \((\omega_\alpha(a), \delta_\alpha(a)), a \in [0, A]\) introduced above. Indeed we can altogether forget about age for the time being.

While the economy is fully described with a probability distribution on the space \((\omega, \delta, \alpha)\), in the computations some sections of this distribution will only be needed. The letter \(F\) will be used for the cdf of \(\delta\), so that

\[ F(\delta|\omega, \alpha) \]

is the fraction of her lifetime an agent of type \(\alpha\) has work opportunity cost at most equal to \(\delta\) when her productivity is equal to \(\omega\). Similarly \(f(\delta|\omega, \alpha)\) is the associated pdf. The letter \(G\) is associated with distributions of productivity \(\omega\). The quantity

\[ G(\omega|\alpha) \]

is the fraction of her lifetime an agent of type \(\alpha\) has productivity smaller than or equal to \(\omega\). The marginal distribution of productivities will be denoted \(G(\omega)\). The letter \(g\) is used for the corresponding pdf’s. Similarly

\[ H(\alpha|\omega) \]

is the cdf of characteristic \(\alpha\), given a productivity level \(\omega\): it reflects the information that the tax authority can infer on the value of \(\alpha\) of someone of current income \(\omega\). The marginal distribution of \(\alpha\) is \(H(\alpha)\). The letter \(h\) is used to denote the associated pdf’s.

3.1 The optimal taxation program

The tax authority at any date is faced with a stationary distribution of productivities and work opportunity costs \((\omega, \delta)\). Typically it wants to redistribute from the highly productive low work opportunity costs persons to the low productive high opportunity costs types. Since it does not observe the \(\delta\)’s, and only knows the productivities of the workers, it announces a nondecreasing after-tax income schedule \(R(\cdot)\) and a subsistence income level \(s\) common to all those out of work.

\(^1\)We suppose that once she has paid the fix cost of work a worker of productivity \(\omega\) is indifferent between producing any before tax income less than \(\omega\), in the interval \([0, \omega]\). Incentive compatibility then obtains when the after-tax income schedule is a non-decreasing function of before tax income \(\omega\).
Facing a time invariant tax subsidy schedule \((R(\cdot), s)\), the typical agent maximizes her consumption flow which is given by her intertemporal budget constraint
\[
c = \int_{\omega} \int_{\delta} [R(\omega) - \delta] \ell(\omega, \delta) f(\delta|\omega, \alpha) g(\omega|\alpha) \, d\delta \, d\omega + s \int_{\omega} \int_{\delta} [1 - \ell(\omega, \delta)] f(\delta|\omega, \alpha) g(\omega|\alpha) \, d\delta \, d\omega,
\]
with respect to her labor supply \(\ell(\omega, \delta)\). The optimal labor supply therefore is to work \((\ell = 1)\) whenever the financial incentive to work \(R(\omega) - s\) is larger than the work opportunity cost \(\delta\). The government intervention, setting a positive \(s\) and often announcing a value of \(R(\omega)\) smaller than laissez-faire \(\omega\), reduces labor supply. The consumption of consumer \(\alpha\), as a function of the tax schedule, is given by
\[
C(R, s; \alpha) = \int_{\omega} \int_{\delta} [R(\omega) - \delta] f(\delta|\omega, \alpha) g(\omega|\alpha) \, d\delta \, d\omega + s \int_{\omega} [1 - F(R(\omega) - s|\omega, \alpha)] g(\omega|\alpha) \, d\omega.
\]
(2)

It is useful for further reference to compute the derivative of consumption both with respect to a change in \(R(\omega)\) in an interval \([\omega, \omega + d\omega]\) and to a change in \(s\):
\[
\frac{\partial C}{\partial R(\omega)} = F(R(\omega) - s|\omega, \alpha) g(\omega|\alpha)
\]
\[
\frac{\partial C}{\partial s} = \int_{\omega} [1 - F(R(\omega) - s|\omega, \alpha)] g(\omega|\alpha) \, d\omega.
\]
(3)

The optimal taxation problem of the government is then to find the couple \((R(\cdot), s)\) that maximizes
\[
\int_{\alpha} u(C(R, s; \alpha), \alpha) \, dH(\alpha)
\]
where lifetime consumption of type \(\alpha\) is given by [2], subject to the budget constraint
\[
\int_{\alpha} \int_{\omega} [\omega - R(\omega) + s] F(R(\omega) - s|\omega, \alpha) g(\omega|\alpha) \, d\omega \, dH(\alpha) \geq s.
\]
(4)

A marginal increase of the consumption of a person of type \(\alpha\) increases the social objective by \(u_{c}(C(R, s; \alpha), \alpha)\). This quantity, which we shall also note \(u_{c}(\alpha)\) to alleviate notations, is the social weight of person \(\alpha\).

### 3.2 First order conditions

Let \(\lambda\) be the Lagrange multiplier associated with the government budget constraint, which measures the marginal cost of public funds. The Lagrangian of the program is
\[
L(R, s) = \int_{\alpha} u(C(R, s; \alpha), \alpha) \, dH(\alpha) + \lambda \int_{\omega} \int_{\alpha} \left\{[\omega - R(\omega) + s] F(R(\omega) - s|\omega, \alpha) - s \right\} g(\omega|\alpha) \, d\omega \, dH(\alpha).
\]
(5)
A small change in income affecting equally everyone in the economy (dR(\omega) = ds for all \omega) must leave the Lagrangian unchanged at the margin, which using (2) yields the first order condition

\int_\alpha u_c(\alpha) dH(\alpha) = \lambda \tag{6}

The marginal cost of public funds is equal to the average of the social weights.

Denote \partial L/\partial R(\omega) the change in the Lagrangian induced by a change dR(\omega) in a small interval [\omega, \omega + d\omega], normalized by dR(\omega) d\omega. We have

\frac{\partial L}{\partial R(\omega)} d\omega = \int_\alpha u_c \frac{\partial C}{\partial R(\omega)} h(\alpha) d\omega d\alpha

+ \lambda \int_\alpha \{\omega - R(\omega) + s\} f(R(\omega) - s|\omega, \alpha) - F(R(\omega) - s|\omega, \alpha)\} g(\alpha|\omega) d\omega h(\alpha) d\alpha.

Using (3) and the conditional densities property \(g(\omega|\alpha)h(\alpha) = h(\alpha|\omega)g(\omega),\)

this yields

\frac{\partial L}{\partial R(\omega)} = \int_\alpha (u_c - \lambda) F(R(\omega) - s|\omega, \alpha) h(\alpha|\omega) d\alpha g(\omega)

+ \lambda[\omega - R(\omega) + s] \int_\alpha \{f(R(\omega) - s|\omega, \alpha)\} h(\alpha|\omega) d\alpha g(\omega).

It is useful to define the (average) social weight of the workers of productivity \omega, i.e. the average of \(u_c(\alpha)\) on all the agents of productivity \omega and opportunity costs of work smaller than \(R(\omega) - s\). It is given by

\(p_E(R, s|\omega) = E[u_c(\alpha)|\omega, \delta \leq R(\omega) - s] = \int_\alpha u_c(\alpha) dH(\alpha|\delta \leq R(\omega) - s, \omega)

= \frac{\int_\alpha u_c(\alpha)F(R(\omega) - s|\omega, \alpha) h(\alpha|\omega) d\alpha}{\int_\alpha F(R(\omega) - s|\omega, \alpha) h(\alpha|\omega) d\alpha}.

Noting that \(\int F(\delta|\omega, \alpha) h(\alpha|\omega) d\alpha\) is simply the marginal \(F(\delta|\omega)\) and similarly with the pdf \(f\), the first order condition can be rewritten

\frac{1}{g(\omega)} \frac{\partial L}{\partial R(\omega)} = [p_E(R, s|\omega) - \lambda] F(R(\omega) - s|\omega) + \lambda[\omega - R(\omega) + s] f(R(\omega) - s|\omega). \tag{7}

At any point where \(R(\omega)\) is strictly increasing, a small enough variation \(dR\) is admissible without violating the monotonicity constraint bearing on the function \(R\). At any such point, \(\partial L/\partial R(\omega)\) is equal to zero along the optimal solution. This expression is identical to that of the static model (see Choné and Laroque (2008)). The crucial difference comes from the social weights, which are derived here from the underlying life cycle model. The game is to see whether the assumptions made for the static model in Choné and Laroque (2008) on the social weights \(p_E\) have natural counterparts in the life cycle setup.
3.3 Optimal participation tax rates

A worker of productivity $\omega$ gets a net income $R(\omega) - s$ from working, so that she faces a participation tax rate

$$\tau(\omega) = \frac{\omega - R(\omega) + s}{\omega}.$$

The elasticity of the labor supply $F(R - s|\omega)$ of the agents of productivity $\omega$ with respect to after tax income $R$ is

$$\varepsilon_R = R \frac{f(R - s|\omega)}{F(R - s|\omega)}.$$

With these two definitions, (7) can be rewritten as in Choné and Laroque (2008)

$$\omega \tau(\omega) = \frac{R}{\varepsilon_R} \left[ 1 - \frac{p_E(R, s|\omega)}{\lambda} \right].$$

At a productivity level where the after tax schedule is strictly increasing and there are both workers and non-workers ($0 < F(R - s|\omega) < 1$), the financial incentive to work is larger (resp. smaller) than productivity and the participation tax rate is negative (resp. positive) whenever the social weight of the workers is larger (resp. smaller) than the marginal cost of public funds.

3.4 A simple utilitarian case

The overall shape of the optimal tax schedule depends on the social weights of the employed agents $p_E(R, s|\omega)$. Without restrictions on these weights, (7) indicates that there is little scope of deriving properties satisfied by after tax income.

However a natural case, stemming from utilitarianism with concave utility indices, is one where the function $p_E$ is nonincreasing both in $R$ and in $\omega$. Under these two assumptions the after tax schedule has a simple shape described in Choné and Laroque (2008), which is recalled below. First a set of restrictive assumptions bearing on the model’s fundamentals is presented under which the utilitarian properties hold.

3.4.1 Deriving utilitarian properties from the fundamentals

A first task is to pin down how consumption varies with $\alpha$. Consumption is an increasing function of $\alpha$ provided that the distribution of $(\omega, -\delta)$ is first order stochastically increasing with the (unidimensional) parameter $\alpha$. This is a strong restriction. It is not difficult to imagine situations where education implies a low market productivity when young, to get high returns later: then the distributions of productivities of educated and non-educated persons would not be comparable according to the stochastic dominance criterion.
Assumption 1. The distribution of $\delta$ conditional on $(\omega, \alpha)$ is first order stochastically nonincreasing in $(\omega, \alpha)$. The distribution of $\omega$ conditional on $\alpha$ is first order stochastically increasing in $\alpha$.

Lemma 1. Under Assumption 1, for any nondecreasing after-tax income schedule $R(\cdot)$, consumption $C(R, s; \alpha)$ is a nondecreasing function of $\alpha$.

Proof: From (2),

$$ C(R, s; \alpha) = \int_\omega \int_\delta \max[R(\omega) - \delta, s] f(\delta|\omega, \alpha) g(\omega|\alpha) \, d\delta \, d\omega. $$

Let

$$ \tilde{A}(r, \omega, \alpha) = \int_\delta \max[r - \delta, s] f(\delta|\omega, \alpha) \, d\delta, $$

$$ A(\omega, \alpha) = \tilde{A}(R(\omega), \omega, \alpha), $$

and

$$ B(a, \alpha) = \int_\omega A(\omega, a) g(\omega|\alpha) \, d\omega. $$

By definition $\tilde{A}$ is nondecreasing in $r$, and by stochastic dominance it is nondecreasing in $(\omega, \alpha)$. Then $A$ is nondecreasing in $(\omega, \alpha)$ by the monotonicity of $R(\cdot)$. Finally $B$ is nondecreasing in both its arguments, the first one by the monotonicity of $A$, the second by monotonicity of $A$ and first order stochastic dominance. The result follows from the remark that $C(R, s; \alpha) = B(\alpha, \alpha)$.

We now turn to the properties of the function $p_E$.

Assumption 2. The distribution of $\alpha$ conditional on $(\omega, \delta \leq R - s)$ is first order stochastically increasing in $\omega$.

Lemma 2. Under Assumption 2, the function $p_E(R, s|\omega)$ is nonincreasing in $\omega$ provided that consumption is nondecreasing with $\alpha$.

Proof: This is a direct consequence of the formula

$$ p_E(R, s|\omega) = \int_\alpha u_c(\alpha) \, dH(\alpha|\delta \leq R(\omega) - s, \omega), $$

since the marginal utility of consumption $u_c(\alpha)$ decreases with $\alpha$.

Increasing the current income of the employed increases their permanent income, and therefore lowers their social weights. But it also brings into the labor force some previously unemployed persons. The next assumption warrants that the arrival of the newcomers does not prevent the function $p_E$ to be decreasing in $R$. 

8
**Assumption 3.** The distribution of $\alpha$ conditional on $(\omega, \delta = R - s)$ first order stochastically dominates the distribution of $\alpha$ conditional on $(\omega, \delta \leq R - s)$.

**Lemma 3.** Under Assumption 3, the function $p_E(R, s|\omega)$ is nonincreasing in $R$ provided that consumption is nondecreasing with $\alpha$.

**Proof:** We use the expression

$$p_E(R, s|\omega) = \frac{\int_{\alpha} u_c(\alpha) F(R - s|\omega, \alpha) h(\alpha|\omega) d\alpha}{\int_{\alpha} F(R - s|\omega, \alpha) h(\alpha|\omega) d\alpha}.$$

Differentiating with respect to $R$ gives

$$\frac{dp_E}{dR} = \frac{\int_{\alpha} u_c(\alpha) \frac{dc}{dR} F(R - s|\omega, \alpha) h(\alpha|\omega) d\alpha}{\int_{\alpha} F(R - s|\omega, \alpha) h(\alpha|\omega) d\alpha} + \frac{\int_{\alpha} u_c(\alpha) f(R - s|\omega, \alpha) h(\alpha|\omega) d\alpha}{\int_{\alpha} F(R - s|\omega, \alpha) h(\alpha|\omega) d\alpha} - \frac{\int_{\alpha} u_c(\alpha) F(R - s|\omega, \alpha) h(\alpha|\omega) d\alpha}{\int_{\alpha} F(R - s|\omega, \alpha) h(\alpha|\omega) d\alpha} \frac{\int_{\alpha} f(R - s|\omega, \alpha) h(\alpha|\omega) d\alpha}{\int_{\alpha} F(R - s|\omega, \alpha) h(\alpha|\omega) d\alpha}.$$

The first term is negative by concavity of the utility function, The second line of the formula has the same sign as

$$\int_{\alpha} u_c(\alpha) h(\alpha|\delta = R - s, \omega) d\alpha - \int_{\alpha} u_c(\alpha) h(\alpha|\delta \leq R - s, \omega) d\alpha.$$

The desired result follows from Assumption 3 and the fact that $u_c$ is decreasing in $\alpha$. ■

**3.4.2 The shape of the after tax income schedule under utilitarianism**

Under the Assumptions of the previous section together with some regularity properties, the optimal after tax schedules have a specific shape, illustrated on Figure 1. The Figure represents the plan $(\omega, \delta)$ of agents’ characteristics. The curve $R(\omega) - s$ delimits the set of workers, who lie below $(\delta < R(\omega) - s)$, from the unemployed who are above $(\delta > R(\omega) - s)$. Laissez-faire corresponds to the 45 degree line, $R(\omega) - s = \omega$. Under Lemmas 2 and 3 the social weights of the workers decrease when one travels up along the 45 degree line. Ignoring pooling, the optimal tax schedule satisfies the first order condition (7), so that $R(\omega) - s - \omega$ has the same sign as $p_E(R, s|\omega) - \lambda$. As a consequence, the after tax income schedule $R(\omega)$ crosses at most once, and from above, the $\omega + s$ line.

The two typical cases are shown on the Figure in a case where the support of the distribution of work opportunity costs has a positive lower bound $\delta$. The solid line represents a situation where the participation tax rate is negative for

---

2This property holds in general, allowing for pooling; the argument follows the lines of the proof of Proposition 4 in Choné and Laroque (2008).
low productivities in the interval $[\delta, \omega_m]$, the quantity $p_E(\omega + s, s|\omega)$ being larger than $\lambda$ in this interval. The dashed line has positive participation tax rates everywhere, and social weights of the employees always smaller than the marginal cost of public funds. From equation (6) which states that the overall average of the social weights is equal to the marginal cost of public funds, this can only occur when the average social weight of the unemployed agents is larger than $\lambda$, the marginal cost of public funds.

3.5 Discussion and possible extensions

The above analysis shows that, by and large, taxation of income in the stationary life cycle model follows similar principles to that of the static model. The main difference comes from the fact that current income in the static model is a direct measure of welfare while here it is a noisy signal of the permanent income which the government wants to redistribute. It is therefore likely that redistribution based on taxation of current income is less efficient here, because of the reduced information contained in the signal.

There is no scope for improvement if, as postulated above, the only information available to the government is the joint distribution of $(\omega, \delta, \alpha)$. But better information, together with a social willingness to condition the tax schedule on this information, will enlarge the scope for redistribution.

A widely discussed possibility in the recent years is to condition taxes on
age, see e.g. Banks and Diamond (2008), Blomquist and Micheletto (2008), or Weinzierl (2008). Since age is an exogenous variable, under the assumption that the distributions of characteristics \((\omega, \delta)\) conditional on \(\alpha\) and age are given and known to the government, the derivation of the optimal tax schedules by age is similar to what has been described above, with a first order condition (7) for each age, or age bracket.

The derivation of the properties of the optimal tax schedule has been facilitated by the assumption of perfect financial markets and perfect insurance across the life cycle. While the notations would be heavier and the computations more involved, a similar first order condition as (7) holds in an environment with imperfect insurance and/or borrowing constraints. The complication is in figuring out the social weight associated with a marginal increase in income: compared with the case studied here, the marginal utility of income would vary over the life cycle instead of being constant, being higher in periods where the credit constraint binds. Intuitively it does not seem to jeopardize the properties proved in Section 3.4.1 but this should be the subject of future research.

The similarity of the static and life cycle models noted above largely comes from a property of the models, left implicit up to now, which is worth making explicit. The determinants \((\omega, \delta)\) of the labor supply decision and of the associated income tax are exogenous, independent of the government chosen tax schedule. An agent cannot modify her productivity profile when taxes become more progressive. Only consumption and welfare depend on taxes. This completely ignores the role of taxes on human capital accumulation, as recently studied for instance by Bohacek and Kapicka (2008).

4 The taxation of capital

Taxing savings or capital to redistribute welfare is natural provided that the agents with the largest permanent incomes are also those with the largest savings. Contrary to permanent income which only depends on the distribution of \((\omega, \delta)\) over the lifetime, savings is a function of the precise trajectory: indeed, for a given income distribution, the positive savings that obtains if income continuously decreases over the lifetime changes sign with increasing incomes. Taxation of capital for a redistributive motive relies on a different type of information than current income taxation. I therefore go back to the formalism of the beginning of the paper.
4.1 How welfare depends on the capital tax rate

In the absence of a tax on wealth, with a zero interest rate, accumulated wealth or savings at age \( a \) is

\[
W_{\alpha a} = \int_{t=0}^{a} [(R(\omega_{\alpha}(t)) - \delta_{\alpha}(t))\ell(t) + s(1 - \ell(t)) - c(t)] \, dt.
\]

Suppose now that saving bears a capital tax at constant rate \( \tau_k \) (this becomes a subsidy in case of borrowing, negative saving). The proceeds of the tax are distributed equally in a lump sum fashion, everyone in the economy receiving \( s_k \) at each instant. The formula becomes

\[
W_{\alpha a} = \int_{t=0}^{a} \exp(-\tau_k(a-t))[(R(\omega_{\alpha}(t)) - \delta_{\alpha}(t))\ell(t) + s(1 - \ell(t)) + s_k - c(t)] \, dt.
\]

The typical consumer chooses a consumption labor profile that maximizes her lifetime utility

\[
\int_{0}^{A} u(c(a), \alpha) \, da
\]

subject to the intertemporal budget constraint \( W_{\alpha A} = 0 \). With this separable formulation, labor supply is chosen so as to maximize income, and follows the rules of the previous section. In particular it does not depend on capital taxation: we can denote \( i_{\alpha t} = (R(\omega_{\alpha}(t)) - \delta_{\alpha}(t))\ell(t) + s(1 - \ell(t)) \) the income at age \( t \). On the other hand, the consumption profile does depend on the capital taxation policy, i.e. positive taxes induce agents to tilt consumption towards younger ages. The computation is reminiscent of that of the lifetime excess burden of an interest tax in Levhari and Sheshinski (1972). Let \( C_{\alpha a}(\tau_k, s_k) \) be the corresponding consumption function. The government optimal taxation program consists in maximizing

\[
U = \int_{\alpha} \int_{a} u(C_{\alpha a}(\tau_k, s_k), \alpha) \, da \, dH(\alpha)
\]

subject to the budget constraint

\[
A s_k - \tau_k \int_{\alpha} \int_{a} W_{\alpha a} \, da \, dH(\alpha) = 0.
\]

It turns out that the derivative of the social objective at the no-tax point \( \tau_k = s_k = 0 \) takes a very simple form. Note that at this point, consumption is constant over the life time, so that the marginal utility of consumption \( u_c(\alpha) \) is independent of age. Then:

\[
\frac{dU}{d\tau_k} = -\text{cov} \left( u_c(\alpha), \int_{0}^{A} W_{\alpha a} \, da \right). \tag{8}
\]
Lemma 4. With perfect financial markets, the derivative of social welfare with respect to the tax rate on capital at the no tax point is equal to the opposite of the covariance of the social weights with the aggregate savings of the cohorts over their life times.

Proof: We first make explicit the conditions associated with the individual behavior of the agents of type $\alpha$. To alleviate notations, let $u_{\alpha a}$ denote the marginal utility of consumption of type $\alpha$ at age $a$. The first order conditions from consumer optimization over the intertemporal budget constraint, marginal utility proportional to the tax factor, are

$$u_{\alpha a} \exp(\tau_k (A - a)) = u_{\alpha 0} \exp(\tau_k A) = u_{\alpha A}.$$  

Also, differentiating the intertemporal budget constraint $W_{\alpha A} = 0$ with respect to $\tau_k$ and $s_k$ gives

$$\int_{a=0}^{A} \left\{ -\exp(-\tau_k (A - a)) \frac{\partial C_{\alpha a}}{\partial \tau_k} - (A - a) \exp(-\tau_k (A - a)) [i_{\alpha a} + s_k - c_a] \right\} da = 0$$  

(9)

and

$$\int_{a=0}^{A} \exp(-\tau_k (A - a)) \left[ \frac{\partial C_{\alpha a}}{\partial s_k} - 1 \right] da = 0.$$  

(10)

We first work on (9). Through an integration by parts:

$$\int_{a=0}^{A} W_{\alpha a} da = [a W_{\alpha a}]^A_{a=0} - \int_{0}^{A} a \exp(-\tau_k (A - a)) [i_{\alpha a} + s_k - c_a] da,$$

or

$$\int_{a=0}^{A} W_{\alpha a} da = - \int_{0}^{A} a \exp(-\tau_k (A - a)) [i_{\alpha a} + s_k - c_a] da.$$

Therefore, using the fact that $W_{\alpha A} = 0$, (9) can be rewritten

$$\int_{a=0}^{A} \exp(-\tau_k (A - a)) \frac{\partial C_{\alpha a}}{\partial \tau_k} da + \int_{a=0}^{A} W_{\alpha a} da = 0$$

Multiplying through by $u_{\alpha A}$ yields

$$\int_{a=0}^{A} u_{\alpha a} \frac{\partial C_{\alpha a}}{\partial \tau_k} da + u_{\alpha A} \int_{a=0}^{A} W_{\alpha a} da = 0$$  

(11)

The same multiplication applied to (10) gives

$$\int_{a=0}^{A} u_{\alpha a} \frac{\partial C_{\alpha a}}{\partial s_k} da = Au_{\alpha A}.$$  

(12)
Using (11) and (12), the derivative with respect to the tax rate on capital of the government objective

$$
\frac{dU}{d\tau_k} = \int_\alpha \int_a u_{\alpha a} \left( \frac{\partial C_{\alpha a}}{\partial \tau_k} + \frac{\partial C_{\alpha a}}{\partial s_k} \frac{ds_k}{d\tau_k} \right) da \, dH(\alpha),
$$
simplifies into

$$
\frac{dU}{d\tau_k} = -\int_\alpha u_{\alpha A} \int_{a=0}^A W_{aa} da \, dH(\alpha) + A \int u_{\alpha A} dH(\alpha) \frac{ds_k}{d\tau_k}.
$$

There remains to evaluate the derivative of government income with respect to the rate of tax on capital, i.e. $ds_k/d\tau_k$. It is well defined at the no-tax point $\tau_k = s_k = 0$. Indeed, at this point the government budget constraint yields through the implicit function theorem

$$
\frac{ds_k}{d\tau_k} = \frac{1}{A} \int_\alpha \int_a W_{aa} da \, dH(\alpha).
$$

Substituting yields the desired result.

A similar, more complicated, formula holds when there are borrowing constraints.\footnote{Then the marginal utility of consumption $u_c(\alpha)$ typically changes during life, being constant only on the age intervals where the borrowing constraint does not bind. For each type $\alpha$, let $\tilde{A}$ be the subset of ages over which consumption is likely to vary. It includes all the ages such that either the agent is borrowing constrained. Furthermore for each age interval across which she is not constrained and smooths her consumption, the lower bound of this age interval also belongs to $\tilde{A}$. For each age in $\tilde{A}$, define a cumulative wealth $\tilde{W}_{aa}$ as follows. When $\alpha$ is constrained at age $a$, $W_{aa}$ is equal to zero: let $\tilde{W}_{aa} = 0$. At a point $a_0$ where there is consumption smoothing on $[a_0, a_1]$, let

$$
\tilde{W}_{a_0} = \int_{a_0}^{a_1} (a - a_0) W_{aa} da.
$$

Then formula (3) holds with the covariance computed on the variables $(\alpha, \tilde{a})$:

$$
\frac{dU}{d\tau_k} = -\text{cov} \left( u_c(C_{\alpha \tilde{a}}, \alpha), \tilde{W}_{\alpha \tilde{a}} \right).
$$}

\subsection*{4.2 A numerical illustration}

To have a rough idea of the orders of magnitude involved, I consider a stripped down example with two types of agents. I abstract from income taxes, and take as income profiles the deterministic profiles estimated for men on the US PSID, for the two categories, less educated (high school graduate or below) and more educated (above high school), by Meghir and Pistaferri (2004), private
Figure 2: Welfare as a function of the tax rate on capital

Figure 3: The wage and consumption paths of the two types under borrowing constraints
communication from Luigi Pistaferri. A quadratic fit gives a yearly income for ages between 20 and 65,

\[ \omega_h(a) = -3.11 + 0.20a - 0.0006a^2, \]
\[ \omega_l(a) = -2.80 + 0.19a - 0.0014a^2. \]

After 65, I complete the profile with fifteen years of retirement without labor income. Life time utility is the sum of the logarithms of consumption from age 20 to age 80, and social welfare is the sum of the utilities of the two types.

With perfect financial markets, the aggregate savings of the two types are respectively 362 and 309. The difference is small, in the direction of a capital subsidy (!), and the optimum tax rate is zero (see Figure 2). With borrowing constraints, they become respectively 588 and 830. The more educated segment of the population accumulates large savings for retirement, in comparison with the less educated. Taxing these savings, contrary to what is seen in developed countries where savings in view of retirement are often subsidized, is redistributive. The optimum tax rate on capital is .9%, as shown on Figure 2.

Figure 3 shows the wage and consumption time path of the two types under borrowing constraints. Wages are the dashed lines, increasing with a parabolic shape below age 65, zero above 65. Consumption without tax on capital follows the wage schedule until the borrowing constraint becomes lax, and is constant afterwards: it is represented with a line with alternating small and large dashes. Applying the tax of .9% without operating the lump sum transfers yields the thin solid line, increasing the age at which the borrowing constraint ceases binding and inducing a declining consumption level overtime after that. The lump sum transfers move this consumption path upwards, on the bold solid lines. Taxation increases the low educated log life time utility by .69, while that of the high education type barely decreases of .06, for a government budget which amounts to 4% of the aggregate wage bill.

5 On taxation of consumer expenditures

The observation of both current income and the evolution of wealth allows the government to estimate consumer expenditure, \( c + \delta \ell \), through the budget identity. This makes it possible to substitute taxes on income and wealth with a tax on consumers expenditures\(^4\) as advocated in the famous Meade (1978) report. Of course, there is a twist with respect to the simplest interpretations of Meade: the government does not observe consumption, but only expenditure, which prevents it to attain the first best.

To setup the optimal expenditure tax problem, first consider the situation of a general nonlinear tax scheme: to buy a quantity of good \( c \), the consumer must

\(^4\)Richard Blundell urged me to investigate this question.
pay $C[c]$, where $C$ is a nonnegative increasing function. The typical consumer program then is
\[
\begin{align*}
\max & \int_0^A u(c(a), \alpha) \, da \\
\int_0^A \{ C[c(a) + \delta_\alpha(a)\ell(a)] - (\omega_\alpha(a) - \delta_\alpha(a))\ell(a) \} \, da \leq 0.
\end{align*}
\]
(13)

Contrary to what happened with the tax on wealth, we see that the tax on consumer expenditures typically distorts both labor supply and the consumption path. Indeed, comparing the terms in the budget set for $\ell$ equal to 0 or 1, we see that the consumer is willing to work whenever
\[
\omega - \delta \geq C[c + \delta] - C[c].
\]
The consumer works whenever is income is larger than the sum required to finance the pecuniary cost of going to work, given the tax system. The consumer problem looks rather hard to solve in general for a nonlinear tax function!

However a linear tax function, say $C[c] = ac - b$, is easy to handle. The (after tax) cost of going to work becomes $a\delta$, and the constant consumption level of consumer $\alpha$ is given by
\[
aC(\alpha) = b + \int_0^{\omega/(1+a)} [\omega - \delta] f(\delta|\omega, \alpha) g(\omega|\alpha) \, d\delta \, d\omega.
\]
The government then chooses the parameters $a$ and $b$ to maximize welfare
\[
\int_a u(C(\alpha), \alpha) \, dH(\alpha)
\]
subject to the budget constraint
\[
\int_a [aC(\alpha) - b] \, dH(\alpha) \geq \int_a C(\alpha) \, dH(\alpha).
\]

References


