Discussion Paper No. 595

OPTIMAL NONLINEAR INCOME AND INHERITANCE TAXATION IN AN INFINITE HORIZON MODEL WITH QUASI-LINEAR PREFERENCE

Hisahiro Naito

September 2003

The Institute of Social and Economic Research
Osaka University
6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan

Optimal Nonlinear Income and Inheritance Taxation in an Infinite Horizon model with quasi-linear preference

Hisahiro Naito*
Institute of Social Economic Research
Osaka University
and
Department of Economics
University of California Irvine

Current Version March 31, 2003

Abstract

This paper analyzes optimal nonlinear income and inheritance taxation by incorporating two types of models that were developed independently in the public finance literature: an infinite horizon representative agent model such as Judd (1995), Chamley (1986) and Lucas (1992), and asymmetric information model analyzed by Mirrlees (1971) and Stiglitz (1982). In this paper, by using an infinite horizon model with heterogenous agents and quasi-linear preference under an asymmetric information environment we characterize optimal income and inheritance taxation. This paper shows that, contrary to the general perception that inheritance taxation should be progressive to some extent, the expected tax liability of those who have a higher level of assets is lower than the expected tax liability of those who have a lower level of assets. Thus, the optimal inheritance tax is regressive.

Keywords: Capital income taxation, heterogenous agents, redistribution, inheritance taxation.

*Institute of Social Economic Research,Mihogaoka 6-1, Ibaraki City, Osaka, Japan, Postal-Code 567-0047
Tel:81-6-6879-8581; fax: 81-6-6878-2766.
E-mail address: naito@iser.osaka-u.ac.jp

1 Introduction

Inheritance taxation is one of the most controversial issues in modern economic policy. Conservatives often argue that inheritance taxes discourage saving, entrepreneurial activity and labor supply and that they have large negative effects on the size of economic output and its growth. Liberals argue that those taxes are necessary to put people on the same start line, to reduce income inequality and to decrease the concentration of wealth. In addition, sometimes it is argued that such redistributive taxes work as social insurance if we consider being rich or poor as idiosyncratic shocks (Varian 1980). Furthermore, a recent increase of inequality of wealth distribution in the US accelerated the policy debate on the effect of inheritance taxes (Gale et al. 2000).

According to the Survey of Consumer Finance in 1995 (Wolf 2000), the top 1% of families (as ranked by marketable wealth) own 45% of total household financial wealth and the top 20% own 92 % of the marketable financial wealth. The average financial wealth is $7,000,000 for the top 1% while the average financial wealth of the population is $155,000. In addition, half of the top 1% receive an inheritance whose average value is $800,000 while 20% of the total population receive an inheritance whose value is $96,000. This implies that the rich are already richer at the beginning than the average population. Thus, on wealth inequality and the transmission of wealth, the argument of the liberals is not groundless. However, those top 1% rich are also very productive people. According to the same data, among the top 1% rich, 69% are self-employed while only 17% are so among the total population. 40% of the top 1% go to graduate school while among the total population only 11% do so. Hence, it is also quite possible that the top 1% are knowledgeable entrepreneurs and their wealth is the reward for their entrepreneurial activity as the conservative argue. Given the fact that half of the top 1% receive inheritance, it is not surprising that inheritance taxes affect incentive for entrepreneurial activity due to business owner's incentive to leave their business to their family members. In such a case, inheritance taxes can have large negative effect on the economy. [1]

Given those arguments, it is interesting to look at how the previous public finance literature treated the issue of inheritance taxation. There are several papers that should be discussed in such a context. In seminal papers, Chamley (1986) and Judd (1985) showed that in an infinite horizon representative agent model, the constrained Pareto-efficient tax rate on capital income should be zero at the steady state and that there should be no intertemporal distortion at the steady state. In a static analysis with heterogenous agent, Atkinson and Stiglitz (1975) showed

[1] For other statistics, the top 1% of families (as ranked by marketable wealth) owned 36% of the total household wealth and the top 20% percent of households owned 84% of it. The total wealth is the financial wealth plus the net equity of owner occupied housing and automobiles.

that if leisure is weakly separable with consumptions in different periods, then an intertemporal distortion is not optimal by using a framework developed by Mirrlees (1971). Since inheritance taxes can be interpreted as one form of intertemporal distortion, those results suggest that the government should not use inheritance taxation.

However, the previous literature is limited in several ways. First, the representative agent model does not seem to be appealing when we need to discuss redistribution of wealth from those who have to those who do not have. Second, the assumption on the weak separability between consumption and labor supply is useful only when labor is supplied in the first period. The result of Atkinson and Stiglitz (1975,1980) does not hold when labor is supplied in multiple periods. Third, the two period overlapping generation model is limited in the analysis of capital income taxation. In the public finance literature, Summers (1981) showed that using the two period overlapping generation model is misleading for the analysis of capital income taxation because the two period overlapping generation model does not capture the income effect for the future income that is caused by a change of discount rate. Fourth, the literature ignores the important fact that the transfer of assets between generations is also accompanied by intangible assets such as knowledge on management. If it is so, discouraging to continue business might imply a large social welfare loss.

In this paper, we examine the issue of inheritance taxation in an infinite horizon model with heterogenous agents, asymmetric information and quasi-linear preference (linear in labor supply). We assume an infinite horizon model given that the substantial size of inheritance among the rich exists in reality. It seems difficult to explain those observed inheritances among the rich by the accidental bequest model. Also using a heterogenous model is needed to discuss redistribution of the unequal wealth distribution. We use quasi-linear preference assumption to make the analysis under the environment of asymmetric information tractable. [2]

In this paper, we first characterize the optimal income and inheritance tax system and show that in the environment of asymmetric information and quasi-linear preference the tax liability of those who have a higher level of capital is on average lower than those who have a lower level of capital, contrary to the general perception that the inheritance should be progressive to some extent.

In the public finance literature, to our knowledge there are two papers that are close to our analysis: Pirttila and Tuomala (2001) Boadway, Merchand and Pestieau (2000). Pirttila and Tuomala analyzed capital income taxation under nonlinear income tax in an OLG model. Boadway et al. analyzed capital income taxation with nonlinear income tax and accidental

[2] In the mechanism design literature, the quasi-linear preference is assumed in many cases for tractability. See Fudenberg and Tirole (1993).

bequest. But neither of the papers analyzed the capital income taxation in a dynastic framework where the amount of investment depends on the structure of income taxes in the future.

Methodologically, this paper borrows an approach used in a dynamic contract theory such as Spear and Srivastava (1987) and Abreau, Pearce, and Stacchetti (1990), which showed that by using the future's expected discounted utility as state variables many interesting dynamic contracts can be analyzed in a tractable way.

2 The model

2.1 Set up

In this economy, time is infinite. Before period 1, there are 2 types of agents denoted by index H and index L. Type H has a higher level of capital K^H and type L has a lower level of capital K^L. Types of capital is observable and verifiable to a social planner. Then, at the beginning of period 1, type H and type L agents turn out to be skilled workers or unskilled workers. This implies that at the middle of each period, there are four types of workers, Hs, Hu, Ls, Lu. Throughout this paper, we use an index i to indicate capital types($i = H, L$) and indices j and m ($m \neq j$) to indicate labor types ($j, m = s, u$). We assume that the probability that an agent becomes skilled or unskilled is i.i.d. and it does not depend on his past history or past labor supply. We also assume that the skilled worker is θ^s/θ^u times productive than unskilled workers ($\theta^s > \theta^u$). We denote the population of type $ij(i = H, L$ and $j = s, u)$ at period t by n_t^{ij}. The type ij agent supplies l_t^{ij} units of labor and earns ${z^{ij}}_t = \theta^j w l_t^{ij}$ units of labor income and $F_k K^i$ units of capital income where w and F_k is the marginal product of labor and the marginal product of capital, respectively. A production function $F(\cdot)$ has two arguments (capital and labor) and exhibits constant returns to scale. Given type ij's capital income and labor income, the social planner determines his tax liability. We assume that the tax liability of each agent does not depend on information of their parents due to the social planner's concern about equity. Once after tax income is determined, the agent decides how much to invest for future and how much to consume for today. Each agent has a dynastic utility function:

$$v_1 = u(c_1) - l_1 + E_1[\sum_{t=1}^{\infty} \gamma^t \{u(c_{t+1}) - l_{t+1}]\}$$

where c_t and l_t are consumption and labor supply at period t and $u'(c) > 0$ and $u''(c) < 0$. α^j is the probability to become type j ($j = s, u$) worker and $\alpha^s + \alpha^u = 1$. We assume that c_t and l_t take any value on the real number. This implies that there is no corner solution for l_t and c_t at the optimum. We assume that one generation lives for only one period. Thus, we interpret γ as a parameter exhibiting a degree of altruism. At the end of each period, each agent makes a consumption and investment decision. Because each agent lives for one period, we assume that the agent cannot borrow money using children's future income as collateral. At the end of period t, given after tax income x_t^i for agent i he decides how much to invest out of his after tax income. At the beginning of the period $t+1$, some investment will succeed and some investments will fail. As a result, some agents will receive a high level of capital K^H becoming type H and other agents will receive K^L becoming type L. We assume that the amount of investment is not observable to the social planner. Thus, there is a moral hazard problem regarding investment. We assume that given type ij's investment I_t^{ij}, the probability of his child receiving K^H at period $t+1$ is $p_i(I_t^{ij})$ and

$$(A1) \qquad p_H(I_t^{ij}) \geq p_L(I_t^{ij})$$

This implies that the probability of successfully leaving capital to his/her child for those who have a high level of capital is higher than for those who have a low level of capital. This can be so because of the effect of intangible assets and reputation of conducting business. The assumption (A1) means that whether the agent can leave his business successfully to his child does not depend on the agent's genetic ability once it is conditioned by the agent's investment. In other words, other than physical investment, there is no factor that causes intergenerational transmission of wealth inequality. After receiving capital at the beginning of period $t+1$, the child again turns out to be a skilled worker or an unskilled worker and the same game will be repeated.

The social planner's objective is to maximize the social welfare function evaluated at $t=1$ with the intertemporal government budget constraint. Thus, given an initial population distribution $\{n_1^H, n_1^L\}$, the social planner solves the following problem:

Primary program

$$\max_{\{v_1^H,\, v_1^L\}} \Psi(v_1^H, v_1^L)$$
$$\text{s.t.} \qquad E(v_1^L, v_1^H; n_1^L, n_1^H) \leq \overline{A}$$

where $\Psi(\cdot)$ is a social welfare function for the social planner. It is concave and strictly increasing with respect to its arguments. $E(v_1^H, v_1^L;, n_1^H, n_1^L)$ is the expenditure function and it is the additionally necessary resource to achieve a lifetime utility v_1^i for each i agent when the population of type i is n_1^i. The above Primary Program says that the social planner will choose v_1^i for each i to maximize his social welfare with the constraint that the additionally necessary cost to achieve a life time utility v_1^i for each i is less than $\overline{A}$. We normally set $\overline{A}$ to zero. For the analysis, define $W^{ij}(x, v_{t+1}^H, v_{t+1}^L)$ as follows:

$$W^{ij}(x, v_{t+1}^H, v_{t+1}^L) \equiv \max_{\{I\}} \; u(x-I) + \gamma p_i(I) v_{t+1}^H + \gamma(1-p_i(I)) v_{t+1}^L$$

$W^{ij}(x, v_{t+1}^H, v_{t+1}^L)$ is the sub-indirect utility when the type ij agent receives after tax income x and the optimal investment is chosen. Given this individual optimization problem, since investment I is a function of x, v_{t+1}^H and v_{t+1}^L, we can write it as $I(x, v_{t+1}^H, v_{t+1}^L)$.

Then, the expenditure function, $E(\cdot)$ is defined recursively as follows:

Sub-program

$$E(v_t^L, v_t^H; n_t^L, n_t^H) = \min_{\{c^{ij},, I_t^{ij}, v_t^j : j=1,2,\ldots,J\}} Q(b_t) + \gamma E(v_{t+1}^L, v_{t+1}^H; n_{t+1}^L, n_{t+1}^H)$$

$$\text{s.t.} \qquad \sum_{j=s,u} \alpha^j \{W^{ij}(x_t^{ij}, v_{t+1}^H, v_{t+1}^L) - l_t^{ij}\} - v_t^i = 0 \ \text{ for } i = H, L \tag{PUi}$$

$$W^{ij}(x_t^{ij}, v_{t+1}^H, v_{t+1}^L) - l_t^{ij} \geq W^{ij}(x_t^{im}, v_{t+1}^H, v_{t+1}^L) - \frac{\theta^m}{\theta^j} l_t^{im} \text{ for } i = H, L \ ; j, m = s, u \text{ and } m \neq j \tag{ICij}$$

$$b_t + \sum_{i=H,L;} \sum_{j=s,u} \alpha^j n_t^{ij} K^i (1+R) + \sum_{i=H,L} \sum_{j=s,u} n_t^i \alpha^j l_t^{ij} w \theta^j \geq \sum_{i=H,L} \sum_{j=s,u} n_t^i \alpha^j x_t^{ij} \tag{RC}$$

$$n_{t+1}^H = \sum_{i=H,L} \sum_{j=s,u} \alpha^j p_i(I^{ij}(x_t^{ij}, v_{t+1}^H, v_{t+1}^L)) N_t^i \tag{TRNH}$$

$$n_{t+1}^L = \sum_{i=H,L} \sum_{j=s,u} \alpha^j [1 - p_i(I^{ij}(x_t^{ij}, v_{t+1}^H, v_{t+1}^L))] n_t^i \tag{TRNL}$$

$$I^{ij}(x, v_{t+1}^h, v_{t+1}^l) = \arg\max \ u(x - I) + \gamma p_i(I) v_{t+1}^h + \gamma(1 - p_i(I)) v_{t+1}^l$$

$$t = 0, 1, 2, \ldots;$$

where $Q(b_t)$ is a penalty function from lending. We assume that $Q(b_t) = b_t$ for $b_t < 0$ and $Q(b_t) = \delta G(b_t)$ for b_t where $\delta > 0$,$G(b_t) \geq B_t, G'(b_t) > 1$ and $G''(b_t) > 0$.

The above two programming problems deserve several comments. First, we consider the problem of the social planner in two steps. First the social planner chooses v_1^i for each i to maximize his social welfare function with the total discounted resource constraint. Then, given those chosen v_1^i for each i, the social planner will choose x_t^{ij} for all i and t to minimize the discounted resource cost to achieve v_1^i for each i. Second, $Q(b_t)$ is a function that captures the social planner's accessibility to international capital market (openness of the economy). For example, if this economy is closed, we can obtain the solution by setting δ at quite a large number. If $Q(b_t)$ is equal to b_t for all b_t, it means that the economy is open for the social planner and the social planner can lend and borrow at the same price. Third, (PUi) is the promised utility constraint. It says that for those who have K^i level of capital, the expected utility from today must be equal to v_t^i for each i. Summarizing the effect of all policies in the future in v_{t+1}^j, we can design today's tax policies when the agents' behavior also depends on tax policies in the future. (ICis) is the incentive compatibility constraint for those who have K^i level of capital and who are skilled workers. Because of the hidden types assumption, at each period t, the social planner cannot know whether an agent is a skilled worker or an unskilled worker. Therefore, the tax system must be designed so that each type self-selects an allocation that the social planner intended. The constraint says that the type ij worker has an incentive to announce that he is type j worker, to work l_t^{ij} hours, to earn $z_t^{ij} \equiv w\theta^j l^{ij}$ dollars and to receive x_t^{ij} units of income rather than to announce that he is type m worker, to work $\frac{w^m}{w^j} l_t^{im}$ hours, to earn $z_t^{im} \equiv w\theta^m l_t^{im}$ units of income and to receive x_t^{im} units of income. (RC) is the resource constraint for the social planner. Definition of $I(x_t^{ij}, v_{t+1}^H, v_{t+1}^L)$ requires that the investment is consistent with intertemporal maximization.

2.2 Analysis

Before analyzing the case with incentive problems, it would be useful to know the first best case where the social planner can control x_t^{ij} and I_t^{ij} perfectly. In this case, it is straightforward to show that at the steady state (i) the consumption levels are equal for all types of agents (ii) the investment levels of all types of agents are equal (iii) the first order condition of the investment implies that $P'(I^{ij}) \times F_k \times (K^H - K^L) = 1/\gamma$ where F_k is the marginal product of capital. Thus, at the first best solution, the consumption is perfectly smoothed for all types of agents and investment is made so that the expected marginal product of investment is equal to the discount rate.

Now consider the second best situation where the social planner cannot observe individual labor types and investment but can observe earned income, which was initially analyzed by Mirrlees (1971) and Stiglitz (1982). In this case, the social planner will use a nonlinear income tax system to distinguish skilled workers from unskilled workers. The social planner will give and require a higher consumption level and higher labor supply to those who announced that they are skilled and will give a lower consumption level and lower labor supply to those who announce that they are unskilled.

Let μ_t, λ_t^i and ϕ_t^{ij}, be the Lagrangian multipliers of RC, PUi and ICij of the sub-program, respectively. Then, the first-order conditions for x_t^{ij}, l_t^{ij} are

$$B_t : Q'(b_t) + \mu_t = 0.$$

$$x_t^{ij} : \lambda_t^i \alpha^j \frac{\partial W^{ij}}{\partial x_t^{ij}} + \phi^{ij} \frac{\partial W^{ij}}{\partial x_t^{ij}} - \phi^{im} \frac{\partial W^{im}}{\partial x_t^{ij}} - \mu_t n^i \alpha^j + \frac{\partial I}{\partial x_t^{ij}} p_i'(I^{ij}) n_t^i \alpha^j \left\{ \frac{\partial E}{\partial n_{t+1}^H} - \frac{\partial E}{\partial n_{t+1}^L} \right\} = 0$$

$$l_t^{ij} : -\alpha^j \lambda_t^i - \phi_t^{ij} + \phi^{im} \frac{\theta^i}{\theta^m} + \mu_t n^i \alpha^j \theta^j w = 0$$

$$i = H, L \; ; j, m = s, u \text{ and } j \neq m$$

As for the FOC of v_{t+1}^i and the population transition equation, it would be useful to write in matrix form:

$$
\begin{aligned}
&\begin{pmatrix} \lambda^H_{t+1} \\ \lambda^L_{t+1} \end{pmatrix} = \begin{pmatrix} \sum_j \alpha^j p_H(I^{Hj}_t) & \sum_j \alpha^j p_L(I^{Lj}_t) \\ \sum_j \alpha^j (1-p_H(I^{Hj}_t)) & \sum_j \alpha^j (1-p_L(I^{Lj}_t)) \end{pmatrix} \begin{pmatrix} \lambda^H_t \\ \lambda^L_t \end{pmatrix} \\
&+\begin{pmatrix} \varphi_t \\ -\varphi_t \end{pmatrix} \qquad (1) \\
&\begin{pmatrix} n^H_{t+1} \\ n^L_{t+1} \end{pmatrix} = \begin{pmatrix} \sum_j \alpha^j p_H(I^{Hj}_t) & \sum_j \alpha^j p_L(I^{Lj}_t) \\ \sum_j \alpha^j (1-p_H(I^{Hj}_t)) & \sum_j \alpha^j (1-p_L(I^{Lj}_t)) \end{pmatrix} \begin{pmatrix} n^H_t \\ n^L_t \end{pmatrix} \qquad (2) \\
&\text{where } \varphi_t = \sum_{j=s,u}\sum_{i=H,L} \phi^{ij}(p_i(I^{ij}_t) - p_i(\widehat{I}^{ij,m}_t)) + \sum_{i=H,L}\sum_{j=s,u} \frac{\partial I^{ij}}{\partial v^H_{t+1}} p'_i(I^{ij}_t) n^i_t \alpha^j \left\{ \frac{\partial E}{\partial n^H_{t+1}} - \frac{\partial E}{\partial n^L_{t+1}} \right\}, \\
&I^{ij}_t = I^{ij}(x^{ij}, v^h_{t+1}, v^l_{t+1}), \text{ and } \widehat{I}^{ij,m}_t = I^{ij}(x^{im}, v^h_{t+1}, v^l_{t+1}) \qquad (3)
\end{aligned}
$$

Note that the matrix of the RHS of (2) is the Markov matrix. For the property of the Markov matrix, see Simon and Blume (1994). From the envelope theorem, $\frac{\partial E}{\partial n^H_{t+1}} - \frac{\partial E}{\partial n^L_{t+1}}$ can be calculated as follows:

$$
\begin{aligned}
\frac{\partial E}{\partial n^H_{t+1}} - \frac{\partial E}{\partial n^L_{t+1}} &= \mu_{t+1}\{[F_k K^H + w\sum_{j=s,u}\alpha^j l^{Hj} - \sum \alpha^j x^{Hj}] - [F_k K^L + w\sum_{j=s,u}\alpha^j l^{Lj} - \sum \alpha^j x^{Lj}]\} \\
&+ \gamma\{\alpha^s(p(I^{Hs}_{t+1}) - p_{Ls}(I^{Ls}_{t+1})) + \alpha^u(p(I^{Hu}_{t+1}) - p(I^{Lu}_{t+1}))\}\{\frac{\partial E}{\partial n^H_{t+2}} - \frac{\partial E}{\partial n^L_{t+2}}\}.
\end{aligned}
$$

Let $F_k K^i + w l^{Hj}_t - x^{ij}_t$ be T^{ij}_t. Since $F_k K^i + w l^{Hj}_t$ is the total income of type ij agent and x^{ij}_t is after tax income, T^{ij}_t can be interpreted as the tax liability of the type ij agent at period t. Thus, $\sum_{j=s,u}\alpha^j T^{Hj}_t - \sum_{j=s,u}\alpha^j T^{Lj}_t$ is the difference of the expected tax liability of type H and type L agents. In addition, in the steady state, $\frac{\partial E}{\partial n^H_t} - \frac{\partial E}{\partial n^L_t} = \frac{\mu_*}{\Delta_*}\{\sum_{j=s,u}\alpha^j T^{Hj}_* - \sum_{j=s,u}\alpha^j T^{Lj}_*\}$ where $*$ indicates variables at the steady state and $\Delta_* = [1-\gamma\{\alpha^s(p_H(I^{Hs}_*) - p_L(I^{Ls}_*)) + \alpha^u(p_H(I^{Hu}_*) - p_L(I^{Lu}_*))\}] > 0$. Thus, the sign of $\frac{\partial E}{\partial n^H_t} - \frac{\partial E}{\partial n^L_t}$ will determine the sign of the difference of the expected tax liability between type H and type L agents at the steady state.

Now we are going to ask whether there should be an inheritance tax in this model. Optimality of inheritance tax can be interpreted as the difference of the tax liability of type H and type L agents for the same labor type. We characterize the structure of inheritance tax in the following steps:

Claim 1 *μ_t and λ^i_t are strictly negative.*

Proof. Suppose that RC constraint is not binding. Then, by decreasing b_t, the social planner can decrease the cost. Thus, μ_t is strictly negative. Next we prove that λ^i_t is strictly negative. Suppose that at the optimum, one of (PUi) is not binding. Then, by increasing l^{is}_t and l^{iu} the social planner can decrease b_t. Since (RC) is binding, this will decrease the total cost to the social planner. This is a contradiction. Thus, (PUi) must be binding. ∎

Claim 2 $x_t^{is} \geq x_t^{iu}$ *and* $z_t^{is} \geq z_t^{iu}$

Proof. From the assumption on the utility function, a single crossing property is guaranteed in a dimension of x_t^{ij} and z_t^{ij}. The incentive compatibility constraint and a single crossing property imply that $x_t^{is} \geq x_t^{iu}$ and $z_t^{is} \geq z_t^{iu}$ for each $i = H, L$. ∎

Claim 3 ϕ_t^{is} *is strictly negative.*

Suppose that neither ICHs nor ICLs are binding. Then increase l_t^{is} by dl^{is} and decrease l_t^{iu} by $dl^{iu} = (\alpha^s/\alpha^u)dl^{is}$ so that the expected utility is the same. On the other hand, the change of the total labor supply is $\theta^s\alpha^s dl^{is} + \theta^u\alpha^u((\alpha^s/\alpha^u)dl^{iu} = \alpha^s(\theta^s - \theta^u)dl^{is}$. Thus, as long as $\theta^s > \theta^u$, the social planner can decrease the total cost. This will continue until both ICHs and ICLs bind.
∎

Claim 4 ϕ_t^{iu}*is equal to zero as long as* $x_t^{is} > x_t^{iu}$.

Proof. Now we will show that if ICis is binding, then ICiu is automatically satisfied as long as $z_t^{is} \geq z_t^{iu}$. Note that from a single crossing property and the ICis and ICiu,$x_t^{is} \geq x_t^{iu}$ implies $z_t^{is} \geq z_t^{iu}$. Now suppose that ICis is binding. Then,(x^{is}, z^{is}) and (x^{iu}, z^{iu}) must be on the same indifference curve of type is. Since $z_t^{is} \geq z_t^{iu}$,(x^{is}, z^{is}) is located to the right of (x^{iu}, z^{iu}) or at the same point of (x^{iu}, z^{iu}). Because of a single crossing property and the shape of the indifference curve, the indifference curve of type iu crosses any indifference curve of type is once from the above. This implies that ICiu is also satisfied. ∎

From this point, we consider a problem that ignores ICHu and ICLu but includes a constraint $x_t^{is} \geq x_t^{iu}$. Note that from Claim 2 $x_t^{is} \geq x_t^{iu}$ is a necessary condition from ICis and ICiu under a single crossing property. Thus, we look for the solution with smaller constraints. After finding the solution with smaller constraints, we will check that the solution also satisfies ICHu and ICLu.

Claim 5 $\lambda_t^H/n_t^H = \lambda_t^L/n_t^L$.

Note that we can ignore ICHu and ICLu. Since ICHs and ICLs bind, from the first order condition of l^{is} and l^{iu},

$$-\lambda_t^i - \phi_t^{is} + \mu_t n^i \alpha^s \theta^s w = 0 \text{ for i=H,L}$$
$$-\lambda_t^i + \phi^{is}\frac{\theta^u}{\theta^s} + \mu_t n^i \alpha^u \theta^u w = 0 \text{ for i=H,L}$$

Those four equations imply that $\lambda_t^H/n_t^H + \phi^{Hs}/n_t^H = \lambda_t^L/n_t^L + \phi^{Ls}/n_t^L$ and $\lambda_t^H/n_t^H + (\theta^u\phi^{Hs})/(\theta^s n_t^H) = \lambda_t^L/n_t^L + (\theta^u\phi^{Ls})/(\theta^s n_t^L)$. This implies that $\lambda_t^H/n_t^H = \lambda_t^L/n_t^L$. ■

Claim 6 $\varphi_t = 0$

Proof. Note from the FOC of l_1^{ij}, $\lambda_1^H/n_1^H = \lambda_1^L/n_1^L$ at period 1. On the other hand, once $n_1^i$is determined, n_2^i is determined by equation (TRNH) and (TRNL), and λ_2^i is determined by equation (1). Thus, when $\lambda_t^H/n_t^H = \lambda_t^L/n_t^L$ for $t = 2$, $\varphi_t = 0$. Extending this logic, it is obvious that it must be true for all $t \geq 1$. ■

Claim 7 $\frac{\partial E}{\partial n_{t+1}^H} - \frac{\partial E}{\partial n_{t+1}^L} \geq 0$ *for all t*

Proof. Note that from Claim 4, $\phi_{iu}\{p(I_t^{iu}) - p(\widehat{I}_t^{iu,s})\} = 0$ for $i = H$,L. From Claim 6 it implies that

$$\varphi_t = \sum_{i=H,L} \phi^{ij}(p_i(I_t^{is}) - p_i(\widehat{I}_t^{is,u})) + \sum_{i=H,L}\sum_{j=s,u} \frac{\partial I^{ij}}{\partial v_{t+1}^H} p_i'(I_t^{ij}) n_t^i \alpha^j \left\{ \frac{\partial E}{\partial n_{t+1}^H} - \frac{\partial E}{\partial n_{t+1}^L} \right\} = 0$$

As for the first term, $p_i(I_t^{is}) - p_i(\widehat{I}_t^{is,u}) > 0$ if $x_t^{is} > x_t^{iu}$ and $p_i(I_t^{is}) - p_i(\widehat{I}_t^{is,u}) = 0$ if $x_t^{is} = x_t^{iu}$. Thus, we have $\frac{\partial E}{\partial n_{t+1}^H} - \frac{\partial E}{\partial n_{t+1}^L} \geq 0$ ■

Claim 8 $x_t^{is} > x_t^{iu}$

Proof. Consider the problem with ICHs, ICLs and $x_t^{is} \geq x_t^{iu}$. Let τ_t be the Lagrangian multiplier of the constraint of $x_t^{is} \geq x_t^{iu}$ where $\tau_t \leq 0$. The first order conditions of x^{is} and x^{iu} of this problem are

$$\lambda_t^H \alpha^s \frac{\partial W^{is}}{\partial x^{is}} + \phi_t^{is} \frac{\partial W^{is}}{\partial x^{is}} - \mu_t n^i \alpha^s + \frac{\partial I}{\partial x_t^{is}} p_i'(I^{is}) n_t^i \alpha^s \left\{ \frac{\partial E}{\partial n_{t+1}^H} - \frac{\partial E}{\partial n_{t+1}^L} \right\} + \tau_t = 0$$
$$\lambda_t^i \alpha^u \frac{\partial W^{iu}}{\partial x^{iu}} - \phi_t^{is} \frac{\partial W^{is}}{\partial x^{iu}} - \mu_t n^i \alpha^u + \frac{\partial I}{\partial x_t^{iu}} p_i'(I^{iu}) n_t^i \alpha^u \left\{ \frac{\partial E}{\partial n_{t+1}^H} - \frac{\partial E}{\partial n_{t+1}^L} \right\} - \tau_t = 0$$

It is easy to check that $x_t^{is} = x_t^{iu}$ is not compatible with the above first order conditions. Thus, $x_t^{is} > x_t^{iu}$. ■

Note that from Claim 4, ICHu and ICLu are also satisfied. This means that it was appropriate to ignore ICHu and ICLu. Also, from the argument of Claim 7, we have $\frac{\partial E}{\partial n^H} - \frac{\partial E}{\partial n^L} > 0$. This implies that at the steady state, $\frac{\partial E}{\partial n^H} - \frac{\partial E}{\partial n^L}$ must be positive. On the other hand, at the steady state $\frac{\partial E}{\partial n^H} - \frac{\partial E}{\partial n^L} = \frac{\mu_*}{\Delta_*}\{\sum_{j=s,u} \alpha^j T_*^{Hj} - \sum_{j=s,u} \alpha^j T_*^{Lj}\}$. Since μ_* is the Lagrangian multiplier of the resource constraint at the steady state, μ_* is strictly negative. This means that the inside of the bracket must be strictly negative. In other words, the excepted tax liability of those who have a higher level of assets is smaller than the tax liability of those who have lower level of assets. Thus, we obtain the following proposition:

Proposition 1 *The expected tax liability of those who have a high level of capital is lower than the expected tax liability of those who have a low level of capital.*

The Proposition 1 can be understood as follows. When the investment is not observable to the social planner, but the social planner wants to redistribute income, there is a moral hazard problem and the incentive for investment is reduced. In order to keep the incentive for the investment, there must be some difference between the expected utilities of those who have a high level of capital and those who have a low level of capital. However, the difference between those expected utilities can be generated by the difference between consumption levels or by the difference between labor supply levels and the answer depends on the curvature of the utility function with respect to consumption and labor supply. The idea of the Proposition 1 is that when the marginal disutility of labor is constant, the difference between v^H and v^L should be generated by the difference between labor supply levels. This implies that low level capital owners needs to work more on average than high level capital owners. Other things being equal, this effect will generate a high level of labor income for low capital owners and a low level of labor income for high capital owners. As a result, low level capital owners pays more taxes on average than high level capital owners.

3 Conclusion

The initial motivation of this paper was to examine whether the social planner should emphasize redistributive and social insurance aspects of inheritance taxation or incentive problems to investment by inheritance taxation. To do so, I developed a model of infinite horizon heterogenous agents with asymmetric information. By using this framework, this paper shows that in the case of quasi-linear preference the social planner will emphasize the negative incentive effects of inheritance taxation rather than redistributive effects. This paper suggests that the relative size of the curvature of utility functions with respect to consumption and labor, which are essentially the elasticity of saving and labor supply, are important for designing an efficient intertemporal tax system.

4 Appendix

References

[1] Gale, William, James Hines and Joel Slemrod, *Rethinking Estate and Gift Taxation*, Brooking Institute Press, 2000

[2] Boadway, Robin; Maurice Marchand, Pierre Pestieau, "Redistribution with Unobservable Bequests: A Case for Taxing Capital Income", *Scandinavian Journal of Economics*, June 2000.

[3] Chamley, Christopher, "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives", Econometrica, 1986, Vol 54(3), pp

[4] Judd Kenneth, L. "Redistributive taxation in A Simple Perfect Foresight Model", *Journal of Public Economics*, Vol.28, pp59-83

[5] Atkinson,Anthony and Joseph Stiglitz, 1980, *Lecture on Public Economics*, McGrawhill

[6] Cremer,Helmuth, Pierre Pestieau, "Non-linear Taxation of Bequests, Equal Sharing Rules and the Trade-off between Intra- and Inter-family Inequalities" *Journal-of-Public-Economics* 79(1), January 2001, pp 35-53.

[7] Grossman, Sandy J., Oliver Hart, "An Analysis of the Principal-Agent Problem," *Econometrica*, 1982, 51, pp7-45.

[8] Feldstein, Martin, " College Scholarship Rules and Private Saving, *American Economic Review;* 85(3), June 1995, pages 552-66.

[9] Fudernberg, Drew and Jean Tirole, *Game Theory*, MIT Press, 1993

[10] Mirrlees, James, " The Theory of Moral Hazard and Unobservable Behavior: Part I" Review of Economic Studies, 1999, 66, pp3-21.

[11] Rogerson, Willian P. "The First Order Approach to Principal-Agent Problems", *Econometrica*, 1985, Vol 53(6), pp 1357-1367.

[12] Simon, Carl, Lawrence Blume, *Mathematics for Economics*, 1994, W.W. Norton.

[13] Spear, Stephen E., Sanjay Srivastava, "On Repeated Moral Hazard with Discounting" *Review of Economics Studies*, 1987, Vol54(4), pp599-617.

[14] Summers, Lawrence-H "Capital Taxation and Accumulation in a Life Cycle Growth Model", *American-Economic-Review*, 1981, September 71(4),pp533-544.

[15] Varian, Hal R., "Redistributive Taxation as Social Insurance", *Journal of Public Economics*, 1980, 14(1), pp 49-68.

[16] Wolf, Edward N. "WHo Are the Rich? A Demographic Profile of High-Income and High-Wealth Americans, Chapter 3 in *Does Atlas Shrug? The Ecoomic Consequeces of Taxing the Rich*, edited by Joel B. Slemrod, Harvard University Press, 2000