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

Hisahiro Naito*
Institute of Social Economic Research
Osaka University

First version January 2001
Current Version May 1, 2002

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 heterogeneous agents and quasi-linear preference under asymmetric information environment we characterizes 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 higher level of asset is lower than the expected tax liability of those who have lower level of asset. Thus, the optimal inheritance tax is regressive.

Keywords: Capital income taxation, heterogenous agents, redistribution.

*Tel:1-6-6879-8581; fax: 1-6-6878-2766.
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 is 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, top 1 percent families (as ranked by marketable wealth) own 45 percent of total household financial wealth and the top 20 percent own 92 percent 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 inheritance whose average value is $800,000 while 20% of the total population receive 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 affects 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.  

Given those arguments, it is interesting to know 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, by using a framework developed by a seminal work by Mirrlees (1971), Atkinson and Stiglitz (1975) showed that if leisure is weakly

\footnote{For other statistics, the top 1% of families (as ranked by marketable wealth) owned 36 percent of total household wealth and the top 20% percent of households owned 84 percent. Total wealth is financial wealth plus net equity of owner occupied housing and automobile.}
separable with consumptions in different periods, then an intertemporal distortion is not optimal. 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 for the analysis of capital income taxation using two period overlapping generation model is misleading because two period overlapping generation model does not capture the income effect that is caused by a change of discount rate for the future income. Fourth, the literature ignore the important fact that transfer of assert between generation is also accompanied by intangible assets such as knowledge on management. If it is so, discouraging to continue business might implies 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 exist in reality. It seems difficult to explain those observed inheritance among the rich by accidental bequest model. Also using heterogenous model is needed to discuss redistribution of the unequal wealth distribution. We use quasi preference assumption to make the analysis under the environment of asymmetric information tractable.²

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

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. a. analyzed capital income taxation with nonlinear income tax and accidental 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 tax in future.

Methodologically this paper borrows an approach used in a dynamic contract theory such

²In the mechanism design literature, quasi-linear preference is assumed in many cases.
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 contract 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 higher level of capital $K^H$ and type L has 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 mid of each period, there are four types of worker, $Hs, Hu, Ls, Lu$. Throughout this paper, we use index $i$ to indicate capital types and index $j$ and $m$ ($m \neq j$) to indicate labor types ($i = H, L$ and $j, m = s, u$). We assume that the probability that an agent becomes skilled or unskilled is i.i.d. and it does not depends on his past history or past labor supply. We also assume that skilled worker is $\theta^s / \theta^u$ times productive than unskilled workers ($\theta^s > \theta^u$). We denote the population of 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 $u''(c) < 0$. $\alpha^j$ 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 $\gamma$ as a parameter exhibiting a degree of altruism. At the end of each period, each agent makes 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 $x^i_t$ 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 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. Given type \( ij \)'s investment \( I_{ij}^t \), the probability of his child receiving \( K^H \) at period \( t+1 \) is \( p_i(I_{ij}^t) \). \( p_i(I_{ij}^t) \) is the probability that type \( ij \)'s children becomes type \( H \) given his investment \( I_{ij}^t \) and we assume that \( p_H(I) \geq p_L(I) \) for the same level of investment \( I \). The assumption that the probability of having \( K^H \) level of capital depends on the index \( i \) comes from the idea that the probability of successfully leave the business to his child depend not only on the amount of the investment but also the amount of capital that the parent holds because of the effect of intangible asset and reputation.\(^3\) After receiving capital at the beginning of period \( t+1 \), he again turns out to be a skilled or 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 \( \{n_{1H}, n_{1L}\} \), the social planner solves the following problem:

**Primary program**

\[
\max_{\{v_{1H}, v_{1L}\}} \Psi(v_{1H}, v_{1L})
\]

\[
s.t. \quad E(v_{1H}, v_{1L}; n_{1H}, n_{1L}) \leq \bar{A}
\]

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

\[
W_{ij}(x, v_{t+1}^{H}, v_{t+1}^{L}) \equiv \max_{\{i\}} u(x - I) + \gamma p_i(I)v_{t+1}^{H} + \gamma(1 - p_i(I))v_{t+1}^{L}
\]

\(^3\)Theoretically, we can assume that the probability also depends on agent’s labor type. However, addition this assumption will destroy a single crossing property of indirect utility function for different labor types. To avoid a complication, we assume that the probability only depends on capital types.
Then, the expenditure function, \( E(\cdot) \) is defined recursively as follows:

Sub-program

\[
E(v^L_t, v^H_t; n^L_t, n^H_t) = \min_{\{c_{ij}, t_{ij}; j=1,2,\ldots,J\}} Q(b_t) + \gamma E(v^L_{t+1}, v^H_{t+1}; n^L_{t+1}, n^H_{t+1})
\]

s.t. \[
\sum_{j=s,u} \alpha^j \{W^{ij}(x^{ij}_t, v^H_t, v^L_t) - l^{ij}_t\} - v^i_t = 0 \quad \text{for} \quad i = H, L
\]

(PUi)

\[
W^{ij}(x^{ij}_t, v^H_t, v^L_t) - l^{ij}_t \geq W^{ij}(x^{im}_t, v^H_t, v^L_t) - \frac{\theta^m}{\theta^j} i^{jm}_t \quad \text{for} \quad i = H, L; j, m = s, u \text{ and } m \neq j
\]

(ICij)

\[
b_t + \sum_{i=H,L} n^i_t K^j + \sum_{i=H,L,j=s,u} \sum_{i=H,L,j=s,u} n^i_t \alpha^j \frac{x^{ij}_t}{l^{ij}_t} \geq \sum_{i=H,L,j=s,u} \sum_{i=H,L,j=s,u} n^i_t \alpha^j x^{ij}_t
\]

(RC)

\[
n^H_{t+1} = \sum_{i=H,L,j=s,u} \sum_{i=H,L,j=s,u} \alpha^j p_i (I^{ij}(x^{ij}_t, v^H_{t+1}, v^L_{t+1})) N^i_t
\]

(TRNH)

\[
n^L_{t+1} = \sum_{i=H,L,j=s,u} \sum_{i=H,L,j=s,u} \alpha^j [1 - p_i (I^{ij}(x^{ij}_t, v^H_{t+1}, v^L_{t+1}))] n^i_t
\]

(TRNL)

\[
I^{ij}(x^{ij}_t, v^h_{t+1}, v^l_{t+1}) = \arg \max_u u(x^{ij}_t - I) + \gamma p(I)x^{h}_t + (1 - p(I))x^{l}_t
\]

\[
t = 0, 1, 2, \ldots;
\]

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 \( v^i_t \) for each \( i \) to maximize his social welfare function with the total discounted resource constraint. Then, given those chosen \( v^i_t \) for each \( i \), the social planner will choose \( x^{ij}_t \) for all \( i \) and \( t \) to minimize the discounted resource cost to achieve \( v^i_t \) 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 \( \delta \) a quite 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 \( v^i_t \) for each \( i \). By summarizing the effect of all policies in future in \( v^i_{t+1} \), we can design today’s tax policies when agent behavior also depends on tax policies in future. (ICis) is the incentive compatibility constraint for those who have \( K^i \) level of capital and who are skilled workers. Because of hidden type 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-select an allocation that the social planner intended. The constraint says that for type \( ij \) worker has an incentive to announce that he is type \( j \) worker, to
work $t^{ij}_t$ hours, earn $z^{ij}_t \equiv w^{ij} t^{ij}_t$ dollars and receive $x^{ij}_t$ unit of income rather than announcing that he is type $m$ worker, $w^m t^m_l$ hours, earn $z^{im}_t \equiv w^m t^m_l$ units of income. (RC) is the resource constraint for the social planner. Definition of $I(x^{ij}_t, v^H_{t+1}, v^L_{t+1})$ requires that the investment is consistent with intertemporal maximization.

2.2 Analysis

Before analyzing the problem with incentive problems, it would be useful to know the first best case where the social planner can control $x^{ij}_t$ and $I^{ij}_t$ perfectly. In this case, it is straightforward to show that at the steady state (i) consumption level is equal for all types of agent (ii) investment level of all types of agent 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 intrinsic type but observe earned income alone, 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 worker from unskilled worker, the social planner will give and require higher consumption level and higher labor supply to those who announced that they are skilled workers and give lower consumption level and lower labor supply to those who announce that they are unskilled.

Let $\mu_t, \lambda^i_t$ and $\phi^{ij}_t$ be the Lagrangian multipliers of RC, PU$i$ and ICij of the sub-program, respectively. Then, the first-order conditions for $x^{ij}_t, l^{ij}_t$ are

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

\[ x^{ij}_t : \lambda^i_t \alpha^j \frac{\partial W^{ij}}{\partial x^{ij}} + \phi^{ij} \frac{\partial W^{ij}}{\partial x^{ij}} - \phi^{im} \frac{\partial W^{im}}{\partial x^{ij}} - \mu_t n^i \alpha^j + \frac{\partial I}{\partial x^{ij}_t} p^i_t (I^{ij}) n^i t \alpha^j \left\{ \frac{\partial E}{\partial n^H_{t+1}} - \frac{\partial E}{\partial n^L_{t+1}} \right\} = 0 \]

\[ l^{ij}_t : -\alpha^j \lambda^i_t - \phi^{ij} + \phi^{im} \frac{\theta^i}{\theta^m} + \mu_t n^i \alpha^j \theta^j w = 0 \]

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

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

\[
\begin{pmatrix}
\lambda_{t+1}^H \\
\lambda_{t+1}^L
\end{pmatrix} = \begin{pmatrix}
\sum_j \alpha_j p_H(I_t^H) & \sum_j \alpha_j p_L(I_t^L) \\
\sum_j \alpha_j (1 - p_H(I_t^H)) & \sum_j \alpha_j (1 - p_L(I_t^L))
\end{pmatrix}
\begin{pmatrix}
\lambda_{t}^H \\
\lambda_{t}^L
\end{pmatrix}
+ \left( \begin{array}{c}
\varphi_t \\
-\varphi_t
\end{array} \right)
\]

(1)

\[
\begin{pmatrix}
\nu_{t+1}^H \\
\nu_{t+1}^L
\end{pmatrix} = \begin{pmatrix}
\sum_j \alpha_j p_H(I_t^H) & \sum_j \alpha_j p_L(I_t^L) \\
\sum_j \alpha_j (1 - p_H(I_t^H)) & \sum_j \alpha_j (1 - p_L(I_t^L))
\end{pmatrix}
\begin{pmatrix}
\nu_{t}^H \\
\nu_{t}^L
\end{pmatrix}
\]

(2)

where \( \varphi_t = \sum_{j=s,u} \sum_{i=H,L} \phi_{ij}(p_i(I_t^ij) - p_i(I_t^{im})) + \sum_{i=H,L} \sum_{j=s,u} \frac{\partial I^ij}{\partial n^H_{t+1}} p_i^j(I_t^ij) n^i_t \alpha_j \left\{ \frac{\partial E}{\partial n^H_{t+1}} - \frac{\partial E}{\partial n^L_{t+1}} \right\} \)

Note that 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:

\[
\frac{\partial E}{\partial n^H_{t+1}} - \frac{\partial E}{\partial n^L_{t+1}} = \mu_{t+1} \left\{ \left[ F_k K^H + w \sum_{j=s,u} \alpha_j I^{Hj} - \sum \alpha_j x^{Hj} \right] - \left[ F_k K^L + w \sum_{j=s,u} \alpha_j I^{Lj} - \sum \alpha_j x^{Lj} \right] \right\}
\]

Let \( F_k K^i + w_l^{Hj} - x_i^{ij} \) be \( T_{t}^{ij} \). Since \( F_k K^i + w_l^{Hj} \) is the total income of type \( ij \) agent and \( x_i^{ij} \) is after tax income, \( T_{t}^{ij} \) can be interpreted as the tax liability of the type \( ij \) agent at period \( t \). Thus, \( \sum_{j=s,u} \alpha_j T_{t}^{Hj} - \sum_{j=s,u} \alpha_j T_{t}^{Lj} \) is the difference of the expected tax liability of type \( H \) and type \( L \) agent. In addition, in the steady state, \( \frac{\partial E}{\partial n^H_{t+1}} - \frac{\partial E}{\partial n^L_{t+1}} = \mu_{*} \left\{ \sum_{j=s,u} \alpha_j T_{s}^{Hj} - \sum_{j=s,u} \alpha_j T_{s}^{Lj} \right\} \) where \( * \) indicates variables at the steady state and \( \Delta_{*} = [1 - \gamma\alpha(s(p(I_s^H) - p(I_s^L)) + \alpha(u(p(I_s^H) - p(I_s^L))))] > 0 \). Thus, the sign of \( \frac{\partial E}{\partial n^H_{t+2}} - \frac{\partial E}{\partial n^L_{t+2}} \) will determine the sign of the difference of the expected tax liability between type \( H \) and type \( L \) agent 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 tax liability of type \( H \) and type \( L \) agent for the same labor type. We characterize the structure of inheritance tax in the following steps:

**Claim 1** \( \mu_t \) and \( \lambda_t^i \) are strictly negative.

Proof. Suppose that RC constraint is not binding. Then, by decreasing \( b_t \), the social planner can decrease the cost. Thus, \( \mu_t \) is strictly negative. Next we prove that \( \lambda_t^i \) is strictly negative. Suppose that at the optimum, one of (PUi) is not binding. Then, by increasing \( l_t^s \) and \( l_t^u \) 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** Bunching of type \( s \) and \( u \) does not occur for any type \( i = H, L \).
Proof. See Appendix.

Claim 3 \( \phi^{is}_t \) is strictly negative and \( \phi^{iu}_t \) is equal to zero.

Proof. First, note that from the assumption of \( p_i(I) \), 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 \( i = H, L \). Now we show that both ICHs and ICLs bind. Suppose that either of ICHs or ICLs are not binding. Then increase in a dimension of \( x_t \). But this implies that \( z_t^{is} \geq z_t^{iu} \). Suppose that ICis is binding. Then, \( \lambda_t^{is} = 0 \) for \( i = H, L \). Now we will show that if ICis is binding, then ICiu is automatically satisfied as long as \( z_t^{is} \geq z_t^{iu} \). Suppose that ICis is binding. Then,

\[
W^{is}(x_t^{is}) - \frac{z_t^{is}}{w^s} = W^{is}(x_t^{iu}) - \frac{z_t^{iu}}{w^s}
\]

This implies that \( W^{is}(x_t^{is}) - W^{is}(x_t^{iu}) = (1/\theta^s)\{z_t^{is} - z_t^{iu}\} \). Under the assumption that \( z_t^{is} \geq z_t^{iu} \),

\[
W^{is}(x_t^{is}) - W^{is}(x_t^{iu}) \leq (1/\theta^s)\{z_t^{is} - z_t^{iu}\}
\]

But this implies that \( W^{is}(x_t^{is}) - (1/\theta^s)z_t^{is} \leq W^{is}(x_t^{iu}) - (1/\theta^s)z_t^{iu} \). Note that from our assumption on \( p(I) \), \( W^{is}(x) = W^{iu}(x) \) for the same \( x \). Thus, \( W^{iu}(x_t^{is}) - (1/\theta^u)z_t^{is} \leq W^{iu}(x_t^{iu}) - (1/\theta^u)z_t^{iu} \). This implies that ICiu is satisfied. Finally, derive the first order condition for \( x_t^{is} \) and \( x_t^{iu} \) by assuming that ICHs and ICLs bind and ICHu and ICLu do not exist. We claim that the solution of this relaxed problem is the solution of the original problem Because, from the FOC of \( x_t^{ij} \), \( x_t^{is} > x_t^{iu} \). But since ICis bind, \( z_t^{is} > z_t^{iu} \). In this case, ICHu and ICLu are also satisfied from the above argument. Thus, this is the solution. 

Claim 4 \( \lambda^H_t/n^H_t = \lambda^L_t/n^L_t \).

Since only ICHs and ICLs bind, from the first order condition of \( l^{is} \) and \( l^{iu} \),

\[
-\lambda_t^i - \phi^{is}_t + \mu_i n^i \alpha^s w = 0 \text{ for } i = H, L
\]

\[
-\lambda_t^i + \phi^{iu}_t \theta^u + \mu_i n^i \alpha^u w = 0 \text{ for } i = H, L
\]

Those four equations imply that \( \lambda^H_t/n^H_t + \phi^H/s/n^H_t = \lambda^L_t/n^L_t + \phi^L/s/n^L_t \) and \( \lambda^H_t/n^H_t + (\theta^u \phi^H)/((\theta^s n^H_t)) = \lambda^L_t/n^L_t + (\theta^u \phi^L)/((\theta^s n^L_t)) \). This implies that \( \lambda^H_t/n^H_t = \lambda^L_t/n^L_t \).

Claim 5 \( \varphi_t = 0 \)
Note from the FOC of \( t^j_1 \), \( \lambda^H_t/n^H_t = \lambda^L_t/n^L_t \) at period 1. On the other hand, once \( n^H_t \) is determined, \( n^L_t \) is determined by equation (TRNH) and (TRNL), and \( \lambda^L_t \) is determined by equation (16). Thus, when \( \lambda^H_t/n^H_t = \lambda^L_t/n^L_t \) for \( t = 2, \varphi_t = 0 \). Extending this logic, it is obvious that it must be true for all \( t \geq 1 \).

The above observation has an important implication for the sign of \( \frac{\partial E}{\partial n^H_{t+1}} - \frac{\partial E}{\partial n^L_{t+1}} \). Because ICIs binds, \( x^i_t > x^i_{1,t} \) from the FOC of \( x^i_j \). This implies that \( I^i > I^u \). Therefore, in order that Claim 4 is true, \( \frac{\partial E}{\partial n^H_{t+1}} - \frac{\partial E}{\partial n^L_{t+1}} > 0 \) for all \( t \). This implies that at the steady state \( \lambda^H_t \{ \sum_{j=s,u} \alpha^j T^H_{s,t} - \sum_{j=s,u} \alpha^j T^L_{s,t} \} \) is positive. Since \( \mu_s \) is the Lagrangian multiplier of the resource constraint at the steady state, it 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 higher level of asset is smaller than the tax liability of those who have lower level of asset!

**Proposition 1** The average tax liability of those who have high level of asset is lower than the average tax liability of those who have low level of asset.

### 3 Implication and Discussion

There is a general perception that inheritance taxes should be redistributive to some extent, or at least not regressive. In contrast, Proposition 1 shows that such a perception is not necessarily correct in particular assumptions on preference and technology. To understand policy implications and the limit of the above proposition, it would be useful to understand the basic logic behind the above proposition. For that purpose, consider a utility possibility frontier between \( v^H_t \) and \( v^L_t \) assuming that the expenditure of the social planner constant. When the preference is quasi-linear, since the marginal disutility of labor is constant, the slope of the utility possibility frontier is \( \frac{dv^H_{t+1}}{dv^L_{t+1}} = n^H_{t+1}/n^L_{t+1} \). By the definition of \( \lambda^H_t, \frac{dv^H_t}{dv^L_t} \) is also equal to \( \lambda^H_t/\lambda^L_t \) for all \( t \). On the other hand, changing \( v^L_{t+1} \) on the utility possibility frontier implies \( dv^H_t \) and \( dv^L_t \) are changed by \( dv^H_t = \sum_j \alpha^j p_H(I^H_{t,j}) \frac{dv^H_{t+1}}{dv^L_{t+1}} + \sum_j \alpha^j p_L(I^L_{t,j}) \) and \( dv^L_t = \sum_j \alpha^j (1 - p_H(I^H_{t,j})) \frac{dv^H_{t+1}}{dv^L_{t+1}} + \sum_j \alpha^j (1 - p_L(I^L_{t,j})) \).

Using the definition of \( n^H_t \) and \( n^L_t \) and \( \frac{dv^H_{t+1}}{dv^L_{t+1}} = n^H_{t+1}/n^L_{t+1} \), we obtain \( \frac{dv^H_{t+1}}{dv^L_{t+1}} = n^H_{t+1}/n^L_{t+1} \). However, there are additional two effects that we need to take into consideration. First, when \( v^H_{t+1} \) and \( v^L_{t+1} \) are changed, it will change the amount of investment and changes the distribution of high level capital owners and low level capital owners. Whether such a change of composition of the population generates a positive effect or a negative effect depends on whether tax liability of high capital owners is larger than the one of low capital owners. If the tax liability of high
capital owners is larger than low level capital owners, then it is generate a positive effect. In addition, there is the second effect of from changing $v_{t+1}^H$ and $v_{t+1}^L$ in the presence of asymmetric information. When the social planner cannot identify the labor type of agents, the incentive compatibility of skilled type binds. As a result, the after tax income of the skilled is higher than the after tax income of the unskilled and the higher after tax income implies higher investment for skilled type than for unskilled type. In this case, increasing $v_{t+1}^H$ and decreasing $v_{t+1}^L$ will relax the incentive compatibility constant and generates additional benefit for the social planner. Note that the slope of the utility possibility frontier is always $n_t^H / n_t^L$. This implies that those two additional effects must offset each other. Hence, the effect through changing population structure and changing tax liability must be strictly positive.

4 Conclusion

The original motivation of this paper was to examine whether the social planner should emphasize redistributive and social insurance aspect of inheritance taxation or incentive problem of the investment. To do so, I developed a model of infinite horizon heterogenous agents with asymmetric information. By using the framework that I developed, this paper shows that the social planner will emphasize the negative incentive effect of inheritance taxation rather than redistributive effects. This suggests that the argument often done by groups that emphasize the negative disincentive effect of inheritance taxation is not groundless.

5 Appendix

5.1 Proof of Claim 1

Suppose that for some $i$, bunching occurs. Without loss of generality, we can assume that $i = H$. When bunching occur, from a single crossing property $x_t^{Hs} = x_t^{Hu}$ and $w^{θs} I^{Hs} = w^{θu} I^{Hu}$. From the FOC of $x_{i}^{s}$ and $x_{i}^{u}$,

\[
\lambda_t^H \alpha^s \frac{∂W^{Hs}}{∂x_{Hs}} + φ_{Hs} \frac{∂W^{Hs}}{∂x_{Hs}} - φ_{Hu} \frac{∂W^{Hu}}{∂x_{Hu}} - μ_t n^H \alpha^s + \frac{∂I}{∂x_{Hs}} p'_{H}(I^{Hs}) n^H \alpha^s \left\{ \frac{∂E}{∂n_{t+1}^H} - \frac{∂E}{∂n_{t+1}^L} \right\} = 0
\]

\[
\lambda_t^H \alpha^u \frac{∂W^{Hu}}{∂x_{Hu}} + φ_{Hu} \frac{∂W^{Hu}}{∂x_{Hu}} - φ_{Hs} \frac{∂W^{Hs}}{∂x_{Hs}} - μ_t n^H \alpha^u + \frac{∂I}{∂x_{Hu}} p'_{H}(I^{Hu}) n^H \alpha^u \left\{ \frac{∂E}{∂n_{t+1}^H} - \frac{∂E}{∂n_{t+1}^L} \right\} = 0
\]
Note that since \( x_t^{Hs} = x_t^{Hu} \), \( \frac{\partial W^{Hs}}{\partial x_s} = \frac{\partial W^{Hu}}{\partial x_u} = \frac{\partial W^{Hu}}{\partial x_u} = \frac{\partial W^{Hu}}{\partial x_u} \) and \( I^{Hs} = I^{Hu} \). Thus, \( \phi^{si} = \phi^{ui} \).

On the other hand, from the FOC of \( l^{si} \) and \( l^{ui} \) and dividing \( \alpha^s \theta^s \) and \( \alpha^u \theta^u \) respectively,

\[
- \frac{1}{\theta^s} \lambda^H - \frac{1}{\theta^s} \phi^{Hs} + \frac{1}{\theta^u} \phi^{Hu} + \mu w = 0 \\
- \frac{1}{\theta^u} \lambda^H - \frac{1}{\theta^u} \phi^{Hu} + \frac{1}{\theta^s} \phi^{Hs} + \mu w = 0
\]

Thus, we obtain \( \lambda^H (-\frac{1}{\theta^s} + \frac{1}{\theta^u}) + \phi^{Hs} (-\frac{1}{\theta^s} + \frac{1}{\theta^s}) + \phi^{Hu} (-\frac{1}{\theta^u} + \frac{1}{\theta^u}) = 0 \). Since \( \theta^s > \theta^u \), \( \phi^{Hs} + \phi^{Hu} = -\lambda^H \). Because \( \phi^{Hs} \leq 0 \), \( \phi^{Hu} \leq 0 \) and \( \lambda^H < 0 \), this is a contradiction.

\[\blacksquare\]

References


