Optimal Taxation in Infinitely-Lived Agent and Overlapping Generations Models: A Review

Andrés Erosa and Martin Gervais

T he literature concerned with dynamic fiscal policy has evolved in two main directions over the last 20 years or so. On the one hand, there is a large literature on optimal taxation. In the context of a standard neoclassical growth model with infinitely-lived individuals, Chamley (1986) and Judd (1985) establish that an optimal income-tax policy entails taxing capital at confiscatory rates in the short run and setting capital income taxes equal to zero in the long run. Only labor income should be taxed in the long run. On the other hand, most applied work concerned with the dynamic impact of fiscal policy uses the life-cycle framework (Auerbach, Kotlikoff, and Skinner [1983], Auerbach and Kotlikoff [1987], and many others, surveyed in Kotlikoff [1998]). Unfortunately, the prescriptions that emanate from the former framework do not necessarily generalize to the latter.

This article reviews the basic results obtained under both the infinitely-lived agent model and the life-cycle model. The first section, which presents a nontechnical introduction to the optimal taxation literature, discusses the optimal taxation problem and the intuition behind the results obtained from both types of models. Section 2 more formally presents the results for the infinitely-lived agent model and Section 3 presents results for the life-cycle economy.

The review of the literature presented here complements that of Chari and Kehoe (1999) and Atkeson, Chari, and Kehoe (1999). The main focus of their

papers is on the infinitely-lived agent model. As the title of Atkeson, Chari, and Kehoe's article indicates, they emphasize that taxing capital is a bad idea. Their conclusion, especially for life-cycle economies, is based on very special cases. Our review suggests instead that there is no real consensus regarding the optimal tax on capital income. Rather, we demonstrate several empirically relevant settings in which optimal capital taxes are non-zero, both in the short run and in the long run.

1.   A REVIEW OF OPTIMAL TAXATION

Statement of the Ramsey Problem

The problem of finding the optimal manner in which to finance a given stream of expenditures has a long tradition in public finance. The statement of the optimal taxation problem given here follows Ramsey's 1927 seminal paper, which formally recognized that individuals and firms react to changes in fiscal policy. When considering alternative fiscal policies, the government has to take into account that individuals and firms will behave in their own best interest, taking as given whichever fiscal policy the government has chosen. Each fiscal policy implies a feasible allocation of goods and factor services, along with prices, that fully reflects the optimal reaction of individuals and firms; that is, each fiscal policy implies a competitive equilibrium allocation.[1] Given a welfare criterion, which the government uses to evaluate different allocations, the Ramsey problem for the government is to pick the fiscal policy[2] that generates the competitive equilibrium allocation giving the highest value of the welfare criterion.

An equivalent way of formulating the Ramsey problem is to let the government pick an *allocation* directly—rather than a set of tax rates—but to restrict the set of allocations from which the government can choose. This set of allocations can be constructed as follows. Pick an arbitrary fiscal policy. Under this fiscal policy, the optimal behavior of individuals and firms generates a competitive equilibrium allocation. This allocation is one element in the set of allocations from which the government can choose. We refer to such an allocation as an *implementable allocation*: to implement this particular allocation as a competitive equilibrium, the government simply needs to choose the fiscal policy that generated it. By repeating this process for all possible fiscal policies, we can construct the set of all possible allocations that the government can implement. The resulting Ramsey problem then consists of picking, among all implementable allocations, the one that maximizes a

---

[1] Note, however, that many fiscal policies may generate the same competitive equilibrium. This is the case, in particular, when the government has more tax instruments than are needed to generate a particular allocation.

[2] Or one of them if there are many.

welfare criterion. In many situations, this alternative way of stating the Ramsey problem, referred to as the *primal approach* in the literature, turns out to be much more convenient than the dual problem of choosing tax rates.[3]

The Ramsey problem poses some additional challenges when its focus is on dynamic fiscal policies. Implicit in the statement of the problem is the following sequence of actions by the government, individuals, and firms. First, at initial date zero, the government announces a time path for the fiscal policy instruments. Taking this path of tax rates as given, individuals and firms then choose their paths of consumption, savings, leisure, and inputs in order to maximize utility and profits. When we get to period one, however, it is quite possible that the government will choose to revise its path of tax rates if given the opportunity to do so. Furthermore, individuals and firms will behave differently in period zero if they know that the government has an incentive to modify the path of tax rates in the future. This problem, known as the *time inconsistency* of policies, is particularly severe in infinitely-lived agent models, but it is also present in life-cycle economies.[4]

Infinitely-Lived Agent Models

Two central prescriptions emerge from the solution to the Ramsey problem in representative, infinitely-lived agent models. The first is that capital income should not be taxed in the long run. This result makes sense if we understand that a positive tax on the return from today's savings effectively makes consumption next period more expensive than consumption in the current period. In infinitely-lived agent models, then, a positive (and constant) tax on capital income in the steady state implies that the implicit tax rate of consumption in future periods increases without bound. On the other hand, the relevant elasticity of demand for consumption at all dates is constant.[5] Taxing dated consumption at different rates thus violates the general public finance principle that tax rates should be inversely proportional to demand elasticities. It follows that the capital income tax should be zero.

The second important aspect of optimal taxation in infinitely-lived agent models stems directly from the time inconsistency of optimal policies discussed above. Prior to date zero, which is the date when the government chooses the path of fiscal instruments, individuals presumably operate under

---

[3] On the primal approach, see Atkinson and Stiglitz (1980) and Lucas and Stokey (1983).

[4] The classic reference on time inconsistency of optimal plans is Kydland and Prescott (1977).

[5] In general equilibrium, the relevant elasticity does not have a readily recognizable representation. If the individuals' utility function is additively separable over time, then this elasticity depends on the intertemporal elasticity of substitution as well as some cross elasticity between consumption and leisure. In any case, the fact that both consumption and leisure are constant in steady state is sufficient to make this elasticity constant.

the assumption that the old fiscal policy will last forever. As far as the government is concerned, individuals' previous actions translate into the economy's initial conditions at date zero, as summarized by individuals' initial asset holdings (capital and government debt). Since these assets were accumulated in the past, at date zero individuals will supply their capital to firms regardless of the fiscal policy: this factor is inelastically supplied. As a result, taxing the return on these assets perfectly imitates a (nondistortionary) lump-sum tax. Without any restrictions on the size of that tax, it is efficient for the government to tax initial asset holdings at confiscatory rates.[6] In this way, the government can finance its stream of expenditures through the return on the levied capital and avoid distortionary taxes in the future. Indeed, if the return on that capital is sufficiently large to finance all future government expenditures, a Pareto optimal allocation is achieved because there is no need to ever use distortionary taxes.

The time inconsistency problem exists because when the government switches to a new fiscal policy, individuals are "surprised" and cannot react to the government's action. As such, the time inconsistency problem and the optimality of the front-loading policy are directly related. The former is not, however, confined to the initial switch in fiscal policy. As long as the path of taxes initially announced by the government involves distortionary taxes at some future date, the government has an incentive to redesign its original plan in order to take advantage of whatever lump-sum tax (capital levy) becomes available in the future. Economists have dealt with the general time inconsistency problem by assuming that the government has access to some commitment device, or a *commitment technology*, that allows the government to commit itself once and for all to the sequence of tax rates announced at date zero. In other words, the commitment technology prevents the government from revising the path of fiscal instruments over time. The optimality of confiscating initial holdings of financial assets, however, still remains an integral part of the solution to the Ramsey problem.

To avoid this arguably trivial solution to the optimal taxation problem, it is usually assumed that the government faces exogenous bounds on the size of feasible tax rates. For example, Chamley (1986) assumes that tax rates have to lie between zero and one. Chamley shows that the optimal policy under some assumptions, with respect to preferences, entails taxing capital income at the highest possible rate for a finite amount of time—while the lump-sum aspect of this tax outweighs its distortionary cost—and setting the capital income tax equal to zero thereafter.[7] Although this exogenous upper bound assumption

---

[6] Similarly, it is efficient for the government to renege on all government debt outstanding at date zero.

[7] In discrete time models, there is a period of transition during which the tax on capital income is strictly between zero and one. There is no such transition period in Chamley's original result since it was derived in a continuous time model.

may seem realistic, it is also completely arbitrary and has a pronounced impact on the solution to the optimal taxation problem: The higher the bound is, the more capital the government accumulates during the first few periods after the switch in fiscal policy and the lower the tax rate on labor income individuals have to face in the future, including the steady state. More generally, the size of the bound determines the magnitude of the welfare gains achieved by switching to the taxes prescribed by the Ramsey problem.

Life-Cycle Economies

The result developed above holds because the elasticity of consumption expenditures exhibits steady state constancy. In turn, this elasticity is constant precisely because consumption and leisure are themselves constant in steady state, which need not be the case in life-cycle economies.[8] In fact, one of the main reasons why economists use the life-cycle model is precisely because observed lifetime consumption and leisure profiles are not flat. Because the behavior of individuals has this life-cycle pattern, there is no reason to expect the relevant consumption elasticities to be constant.

It follows from this reasoning that consumption at different ages should be taxed at different rates. Alternatively, capital income should be taxed or subsidized at rates that depend on the age of the individual supplying the capital.[9] Through a similar argument, optimal labor income tax rates also vary with the age of the individual supplying labor. Although these arguments indicate that the relative capital and labor income tax rates should vary over the lifetime of individuals, they leave open the question of how to determine the level at which these tax rates should be set. We will return to this question below.

The choice of a welfare function to evaluate different implementable allocations is not as straightforward in life-cycle economies as it is in infinitely-lived agent models. The fact that standard infinitely-lived agent models are populated by a single representative individual dictates that the benevolent government or planner would use the representative agent's utility function as the welfare function. A life-cycle economy, however, involves many heterogeneous agents: each generation has (at least) a representative member, and a new generation is born every period. At a minimum, relative weights need to be assigned to each individual. It is usually assumed that these weights take

---

[8] This remains true even in infinitely-lived agent economies with heterogeneous consumers. See Judd (1985) for details.

[9] Recall that taxing consumption tomorrow more (less) than today is equivalent to taxing (subsidizing) capital income tomorrow.

the form of a discount factor, so that the planner places an ever decreasing weight on future generations.[10]

Irrespective of the precise form of these welfare weights, the impact of a capital levy is very different in a life-cycle environment than in an infinitely-lived agent economy, simply because a capital levy explicitly involves a redistribution between generations: The individuals on whom the burden of a capital levy falls are different from those who benefit from lower distortionary taxation in the future. For example, consider the impact of confiscating the assets of a (possibly retired) individual who, at date zero, is in his last period of life. Under this front loading policy, this individual's consumption would be very low (it may be zero) and so would his utility; his utility is not affected by the lower tax rates that future generations would face. Since this individual's utility has a positive weight in the welfare function, the value of the government's objective would be driven down considerably by the front loading policy. This is not to say that the government would not tax initial assets at all, but rather that the extent to which the government will do so is limited, at least relative to what is optimal in infinitely-lived agent models. Accordingly, there is no need to impose arbitrary bounds on feasible tax rates in life-cycle economies.

Recall that the level of the long-run labor income tax in the infinitely-lived agent model is a function of these exogenous bounds on feasible tax rates. The size of these bounds determines how much capital the government accumulates during the transition, and thus the tax revenue that needs to be collected in the long run and the tax rate on labor income. Since there is no need to impose such bounds in life-cycle models, what, then, determines the tax revenue that needs to be collected in the long run? The answer lies in the weights that the planner puts on different generations. In the usual case, where these weights are represented by a discount factor, the steady-state amount of government debt and the amount of tax collection are entirely driven by the size of the discount factor. A relatively low discount factor indicates that the government puts low weights on future generations relative to current generations. In such cases, the government will tend to have a relatively high amount of accumulated debt in the long run and will need to collect a relatively high amount of taxes. Similarly, a high discount factor implies low (or even negative) government debt and that a small amount of taxes is to be collected.

---

[10] Note that these weights have to get smaller over time—they have to converge to zero—for the welfare function to be well defined. For example, if all generations had the same weight, all (positive) allocations would give the same welfare value (infinity).

2. AN INFINITELY-LIVED AGENT ECONOMY

Consider an economy populated by a large number of identical individuals with infinite lives.[11] Each period, individuals are endowed with one unit of productive time. The representative individuals' preferences are ordered according to the following utility function

$$\sum_{t=0}^{\infty} \beta^t U(c_t, 1 - l_t). \tag{1}$$

Equation (1) expresses that in each period of their infinite lives, individuals care about consumption, denoted c_t, and leisure, $(1-l_t)$; the latter corresponds to the total endowment of time minus time devoted to work (l_t). The discount factor $1 > \beta > 0$ is used by individuals to discount utility in future periods to utility in the current period. We assume that the utility function U is strictly increasing in both arguments, is strictly concave, and satisfies the standard Inada conditions.[12] Two commonly used algebraic forms for the utility function will be considered in the text. The first is a utility function which is *separable* between consumption and leisure:

$$U(c, 1 - l) = \frac{c^{1-\sigma}}{1 - \sigma} + V(1 - l), \tag{2}$$

where the function V satisfies the above stated conditions and $1/\sigma$ is the intertemporal elasticity of substitution, which measures the degree to which individuals are willing to substitute consumption over time. The second functional form we consider is a *Cobb-Douglas* utility function:

$$U(c, 1 - l) = \frac{c^{1-\sigma}(1 - l)^{\eta}}{1 - \sigma}, \tag{3}$$

where $\eta = \theta(1 - \sigma)$. In equation (3), $1/\sigma$ has the same interpretation as it did under the separable utility function and θ measures the intensity of leisure in individuals' preferences.

Each period, individuals face the budget constraint

$$c_t + a_{t+1} = w_t l_t + (1 + r_t)a_t, \tag{4}$$

where w_t is the after-tax wage rate, r_t is the after-tax interest rate, and a_{t+1} is the amount of resources carried over from period t to period $t + 1$. Letting p_t be the Lagrange multiplier on the time-t budget constraint, the first order

---

[11] This section draws from Atkeson, Chari, and Kehoe (1999).

[12] The Inada conditions state that the marginal utility of consumption or leisure is very high (low) at very low (high) consumption levels, that is, $\lim_{c \to 0} U(c, 1 - l) = \lim_{l \to 1} U(c, 1 - l) = \infty$ and $\lim_{c \to \infty} U(c, 1 - l) = 0$. Note that leisure time cannot exceed one since working time cannot be negative.

conditions for individuals are

$$\beta^t U_{c_t} - p_t = 0, \tag{5}$$

$$\beta^t U_{l_t} + p_t w_t = 0, \tag{6}$$

$$-p_t + (1 + r_{t+1}) p_{t+1} = 0, \tag{7}$$

where U_{c_t} and U_{l_t} denote the derivative of U with respect to c_t and l_t respectively, that is, the marginal utility of consumption and leisure. One could use these conditions to obtain the optimal consumption and leisure demands of individuals. Naturally, these demand functions would depend on the fiscal policy chosen by the government, and they would represent the reaction functions that the government takes into account when formulating a Ramsey problem. We show below that these first order conditions can be used not only to construct the budget constraint but also to construct a constraint that can be imposed on the government when formulating a Ramsey problem where the government chooses allocations rather than tax rates.

There is a single produced good in our economy that can be used as consumption (private or public) or capital. For the goods-producing sector of our economy, we assume that the input-output technology is represented by a neoclassical production function with constant returns to scale, $y_t = f(k_t, l_t)$, where y_t, k_t, and l_t denote the aggregate (or per capita) levels of output, capital, and labor, respectively. Profit maximization by firms implies that capital and labor services are paid their marginal products: before-tax prices of capital and labor in period t are given by $\hat{r}_t = f_{k_t} - \delta$, where $0 < \delta < 1$ is the depreciation rate of capital, and $\hat{w}_t = f_{l_t}$.

Feasibility requires that total (private and public) consumption plus investment be less than or equal to aggregate output

$$c_t + k_{t+1} - (1 - \delta)k_t + g_t \leq y_t, \tag{8}$$

where g_t stands for date-t government consumption and all aggregate quantities are expressed in per capita terms.

To finance its given stream of expenditures, the government has access to a set of fiscal policy instruments and to a commitment technology to implement its fiscal policy. The set of instruments available to the government consists of government debt and proportional taxes on labor income and capital income.[13] The date-t tax rates on capital and labor services are denoted by τ_t^k and τ_t^w, respectively. In per capita terms, the government budget constraint at date t is given by

$$(1 + \hat{r}_t)b_t + g_t = b_{t+1} + (\hat{r}_t - r_t)a_t + (\hat{w}_t - w_t)l_t, \tag{9}$$

---

[13] In this framework, consumption taxes need to be ruled out to make the problem interesting since they can be used in conjunction with labor income taxes to perfectly imitate a levy of initial holdings of assets. See Chari and Kehoe (1999) for details.

where b_t represents government debt issued at date t, $w_t \equiv (1 - \tau_t^w)\hat{w}_t$, and $r_t \equiv (1 - \tau_t^k)\hat{r}_t$. Equation (9) expresses that the government pays its expenditures, which are composed of outstanding government debt payments (principal plus interest) and other government outlays, either by issuing new debt, by taxing interest income, or by taxing wage income.

In the spirit of Ramsey, the government takes individuals' optimizing behavior as given and chooses a fiscal policy to maximize a given welfare criterion. Since there is but a single representative agent in this economy, a natural way for the government to evaluate different allocations is to use the representative individual's utility function. If we let π denote a fiscal policy and denote $c_t(\pi)$ and $l_t(\pi)$ the solution to the consumer problem as a function of the fiscal policy, then the Ramsey problem is

$$\max_{\{\pi\}} \sum_{t=0}^{\infty} \beta^t U(c_t(\pi), 1 - l_t(\pi)), \tag{10}$$

subject to feasibility (8) and the government budget constraint (9) for all $t \geq 0$. This Ramsey problem corresponds to the dual. Note that the problem is fairly difficult to analyze because any tax instrument enters all demand functions. Given this difficulty, the primal approach is much more tractable.

The Primal Approach

To construct a Ramsey problem where the government chooses allocations rather than tax rates, we must restrict the set of allocations from which the government can choose. This set should include only allocations that are competitive equilibrium under some fiscal policy. To construct this set, we use the fact that for any given fiscal policy, the competitive equilibrium must satisfy the consumer's optimality conditions, including the budget constraint, as well as those of the firm. Using these optimality conditions, we can derive a condition that competitive equilibria must satisfy.

Our first step is to iterate on the budget constraint (4) to express this sequence of constraints as a single, present-value budget constraint

$$\sum_{t=0}^{\infty} \left[\prod_{j=1}^{t} \frac{1}{1+r_j} \right] (c_t - w_t l_t) = (1 + r_0)a_0. \tag{11}$$

The term inside the square brackets is a shorthand to express the multiplication of many terms, $\frac{1}{1+r}$ in this case. Next, the consumer's first order conditions (5)

through (7) imply

$$p_t/p_0 = \prod_{j=1}^{t} \frac{1}{1+r_j}, \tag{12}$$

$$w_t = \frac{U_{l_t}}{U_{c_t}}. \tag{13}$$

Using equations (5), (12), and (13) we can rewrite the present value budget constraint (11) as

$$\sum_{t=0}^{\infty} \beta^t (U_{c_t} c_t + U_{l_t} l_t) = A_0, \tag{14}$$

where $A_0 \equiv (1+r_0)U_{c_0}a_0$, $r_0 = (1-\tau_0^k)\hat{r}_0$ and τ_0^k is taken to be fixed to make the problem interesting.

Equation (14) is referred to as the *implementability constraint*. It can be shown that any competitive equilibrium allocation has to satisfy this constraint, and that any feasible allocation satisfying the implementability constraint is a competitive equilibrium (see Chari and Kehoe [1999] for details). Imposing this constraint on the government's problem accomplishes exactly what we wanted: It ensures that any allocation picked by the government can be implemented as a competitive equilibrium.

We can now state the Ramsey problem in terms of allocations. This problem consists of maximizing welfare, given by the representative consumer's utility function (1), subject to feasibility (8) and the implementability constraint (14). Note that by Walras's law, if the individual's present value budget constraint (11) holds under a feasible allocation, then the government's budget constraint (9) is also satisfied. Let λ be the Lagrange multiplier on the implementability constraint (14) and define the pseudo-welfare function W to include the implementability constraint

$$W_t = U(c_t, 1 - l_t) + \lambda(U_{c_t} c_t + U_{l_t} l_t). \tag{15}$$

The multiplier λ will be strictly positive if it is necessary for the government to use distortionary taxation. The term multiplying λ essentially gives a bonus to date-t allocations that bring in extra government revenues, thereby relieving other periods from distortionary taxation, and the same term imposes a penalty in the opposite situation. The Ramsey problem, in terms of allocations, is

$$\max_{\{c_t, l_t, k_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t W_t - \lambda A_0, \tag{16}$$

subject to feasibility (8). The form the above problem takes, the *primal*, is very similar to a first-best planning problem except that the pseudo-welfare function replaces the utility function.

Prescriptions

Using the primal allows us to characterize optimal fiscal policies. With few exceptions, our focus will be on the capital income tax. The first order conditions for an optimum imply[14]

$$-\frac{W_{l_t}}{W_{c_t}} = -\frac{U_{l_t}[1 + \lambda(1 + H_t^l)]}{U_{c_t}[1 + \lambda(1 + H_t^c)]} = f_{l_t}, \tag{17}$$

$$\frac{W_{c_t}}{W_{c_{t+1}}} = \frac{U_{c_t}[1 + \lambda(1 + H_t^c)]}{U_{c_{t+1}}[1 + \lambda(1 + H_{t+1}^c)]} = \beta(1 + f_{k_{t+1}} - \delta), \tag{18}$$

for $t = 1, 2, \ldots$, where

$$H_t^c \equiv \frac{U_{c_t,c_t}c_t + U_{l_t,c_t}l_t}{U_{c_t}}, \tag{19}$$

$$H_t^l \equiv \frac{U_{c_t,l_t}c_t + U_{l_t,l_t}l_t}{U_{l_t}}. \tag{20}$$

Equation (17) equates the government's marginal rate of substitution between consumption and leisure at date t to the marginal product of labor of that date. Similarly, equation (18) sets the government's marginal rate of substitution between consumption today and consumption tomorrow equal to the return on capital (net of depreciation). Note that the government's marginal rate of substitution, unlike that of an individual, takes the implementability constraint into account. Also, the government cares about before-tax prices, whereas individuals face after-tax prices.

Atkeson, Chari, and Kehoe (1999) refer to the terms H^c and H^l as *general equilibrium elasticities* since they capture the relevant distortions for setting the capital and labor income tax rates in general equilibrium. Notice that if $H_t^c = H_{t+1}^c$, then equation (18) implies that

$$\frac{U_{c_t}}{U_{c_{t+1}}} = \beta(1 + f_{k_{t+1}} - \delta) = \beta(1 + \hat{r}_{t+1}). \tag{21}$$

This condition is of particular interest, for when $H_t^c = H_{t+1}^c$, the tax on capital income in period $t+1$ is zero. To observe this result, notice that the consumer's first order conditions (5) and (7), which the Ramsey allocation has to satisfy, imply that

$$\frac{U_{c_t}}{U_{c_{t+1}}} = \beta(1 + r_{t+1}) = \beta(1 + (1 - \tau_{t+1}^k)\hat{r}_{t+1}). \tag{22}$$

For both equations (21) and (22) to hold, τ_{t+1}^k must equal zero. Similarly, if $H_t^c = H_t^l$, the tax rate on labor income has to be equal to zero for the optimality

---

[14] Note that the first-order conditions at time zero are different from the above equations since consumption at date zero appears inside the term A_0.

condition in order for the labor decision of individuals ($-U_{l_t}/U_{c_t} = w_t$) to be compatible with that of the government (equation (17)). We have just demonstrated the following proposition:

Proposition 1 *In our infinitely-lived agent model, (i) the optimal tax rate on labor income at date t is different from zero unless $H_t^c = H_t^l$, and (ii) the optimal tax rate on capital income at date $t + 1$ is different from zero unless $H_t^c = H_{t+1}^c$.*

Chamley's (1986) celebrated result on the optimality of not taxing capital in the long run follows directly from Proposition 1. Suppose that the Ramsey allocation converges to a steady state where consumption and leisure are constant by definition. It follows immediately that in such a steady state, the function H_t^c is also constant, which implies that the optimal capital income tax is zero in the long run.

Proposition 1 also implies that for utility functions that are separable in consumption and leisure (so that $U_{c,l} = 0$) and have a constant intertemporal elasticity of substitution (so that $U_{c,c}c/U_c$ is constant), the capital income tax should be zero in all but the first period. Of course, capital should be taxed at a confiscatory rate in the first period regardless of preferences since this tax perfectly imitates a lump-sum tax.

Proposition 2 *Under utility functions of the form given by (2), the optimal capital income tax in our infinitely-lived agent model is zero for $t > 1$.*

Chamley also shows, under the same separable utility function, that imposing a bound on feasible tax rates implies the following for the optimal tax on capital income: The tax rate should be equal to the upper bound for a finite amount of time, after which it should be equal to zero. In discrete time models, there is a period between the two regimes where the tax rate is strictly between zero and the upper bound. The intuition for this result is that taxing capital income has two effects. While the capital income tax partially imitates a lump-sum tax because the initial stock of capital is given, it also introduces a distortion on the savings decision. As a result, the lump-sum aspect of the tax dominates for periods sufficiently close to date zero, and the distortionary aspect of the tax dominates thereafter.

The intuition for the above result remains intact under more general utility functions, in the sense that early taxation of capital income is preferred to later taxation. It is less clear, however, whether capital income should be taxed throughout the transition. In particular, for the Cobb-Douglas utility function (3), the optimal capital income tax is zero only in the long run.

3. A LIFE-CYCLE ECONOMY

We now consider a life-cycle economy.[15] This economy is similar to the infinitely-lived agent economy considered in Section 2, except it is populated by overlapping generations of individuals with finite lives. Individuals still make consumption and labor/leisure choices in each period so as to maximize their lifetime utility, and firms still operate a neoclassical production technology. The payments received by individuals on their factors (capital and labor) are subject to proportional taxes, which we now assume can be conditioned on age.

Individuals live $(J+1)$ periods, from age 0 to age J. At each time period, a new generation is born and is indexed by date of birth. At date zero, when the change in fiscal policy occurs, the generations alive are $-J, -J+1, \ldots, 0$. In order to take these initial generations into account in the following analysis, it will prove convenient to denote the age of individuals alive at date zero by $j_0(t)$. For all other generations we set $j_0(t) = 0$, so that for any generation t, $j_0(t) = \max\{-t, 0\}$. One can thus think of $j_0(t)$ as the first period of an individual's life affected by the date zero switch in fiscal policy. We let $\mu_j = 1/(J+1)$ represent the share of age-j individuals in the population. The labor productivity level of an age-j individual is denoted z_j.

We let $c_{t,j}$ and $l_{t,j}$, respectively, denote consumption and time devoted to work by an age-j individual who was born in period t. Note that $c_{t,j}$ and $l_{t,j}$ actually occur in period $(t + j)$. Similarly, the after-tax prices of labor and capital services are denoted $w_{t,j}$ and $r_{t,j}$, respectively. The problem faced by an individual born in period $t \geq -J$ is to maximize lifetime utility subject to a sequence of budget constraints:

$$U^t(\pi) \equiv \max \sum_{j=j_0(t)}^{J} \beta^{j-j_0(t)} U(c_{t,j}, 1 - l_{t,j}), \qquad (23)$$

$$\text{s.t.} \quad c_{t,j} + a_{t,j+1} = w_{t,j} z_j l_{t,j} + (1 + r_{t,j}) a_{t,j}, \quad j = j_0(t), \ldots, J. \quad (24)$$

This problem mimics the consumer's problem from Section 2, except for the need to index all variables by age. We denote $U^t(\pi)$ the indirect utility function, that is, the maximum lifetime utility obtained by an individual from generation t under fiscal policy π.

Let $p_{t,j}$ denote the Lagrange multiplier associated with the budget constraint (24) faced by an age-j individual born in period t. The necessary and sufficient conditions for a solution to the consumer's problem are given by (24)

---

[15] This section draws from Erosa and Gervais (2000).

and

$$\beta^{j-j_0(t)}U_{c_{t,j}} - p_{t,j} = 0, \tag{25}$$

$$\beta^{j-j_0(t)}U_{l_{t,j}} + p_{t,j}w_{t,j}z_j \le 0, \quad \text{with equality if } l_{t,j} > 0, \tag{26}$$

$$-p_{t,j} + p_{t,j+1}(1 + r_{t,j+1}) = 0, \tag{27}$$

$$a_{t,J+1} = 0, \tag{28}$$

$j = j_0(t), \ldots, J$, where $U_{c_{t,j}}$ and $U_{l_{t,j}}$ denote the derivative of U with respect to $c_{t,j}$ and $l_{t,j}$, respectively.[16]

The feasibility constraint is still given by (8). However, the date-t aggregate levels of consumption and labor input—the latter being expressed in efficiency units—are now obtained by adding up the weighted consumption (or effective labor supply) of all individuals alive at date t, where the weights are given by the fraction of the population that each individual represents:

$$c_t = \sum_{j=0}^{J} \mu_j c_{t-j,j},$$

$$l_t = \sum_{j=0}^{J} \mu_j z_j l_{t-j,j}.$$

The set of instruments available to the government consists of government debt and proportional, age-dependent taxes on labor income and capital income.[17] The date-t tax rates on capital and labor services supplied by an age-j individual (born in period $(t - j)$) are denoted by $\tau_{t-j,j}^{k}$ and $\tau_{t-j,j}^{w}$, respectively. In per capita terms, the government budget constraint at date $t \ge 0$ is given by

$$(1 + \hat{r}_t)b_t + g_t =$$

$$b_{t+1} + \sum_{j=0}^{J} (\hat{r}_t - r_{t-j,j})\mu_j a_{t-j,j} + \sum_{j=0}^{J} (\hat{w}_t - w_{t-j,j})\mu_j z_j l_{t-j,j}, \tag{29}$$

where $w_{t,j} \equiv (1 - \tau_{t,j}^{w})\hat{w}_{t+j}$ and $r_{t,j} \equiv (1 - \tau_{t,j}^{k})\hat{r}_{t+j}$. Equation (29) has exactly the same interpretation as equation (9) in the infinitely-lived agent model.

---

[16] We assumed in Section 2 that the labor supply was always between zero and one. In the context of a life-cycle economy, however, the labor supply can realistically hit a corner solution if labor productivity gets sufficiently small. For example, individuals may become less productive as they age and choose to retire.

[17] Recall that consumption taxes had to be ruled out in the infinitely-lived agent model because they could perfectly imitate a levy of initial holdings of assets. Because the government's incentive to confiscate initial holdings of assets is endogenously limited in life-cycle economies, we could allow for consumption taxes. However, it can be shown that a consumption tax would be a redundant instrument. See Erosa and Gervais (2000) for details.

Because of the presence of many heterogeneous individuals in this economy, the choice of a welfare function is not as straightforward here as it is in the infinitely-lived agent model. Below, we assume that social welfare is defined as the discounted sum of individual lifetime welfares (as in Samuelson [1968] and Atkinson and Sandmo [1980]). In other words, the government chooses a sequence of tax rates in order to maximize

$$\sum_{t=-J}^{\infty} \gamma^t U^t(\pi),\tag{30}$$

where $0 < \gamma < 1$ is the intergenerational discount factor and $U^t(\pi)$, as was defined earlier, is the indirect utility function of generation t as a function of the government tax policy. As was the case with the infinitely-lived agent model, it is much easier to characterize optimal fiscal policies using the primal approach.

The Primal Approach

As before, we need to impose restrictions so that any allocation chosen by the government can be decentralized as a competitive equilibrium. In life-cycle models, each generation has its own implementability constraint. The implementability constraints are obtained by using the consumers' optimality conditions (25) through (27) and acknowledging the fact that factors are paid their marginal products to substitute out prices from consumers' budget constraints (24). After adding up these budget constraints, the resulting implementability constraint associated with the cohort born in period t is given by

$$\sum_{j=j_0(t)}^{J} \beta^{j-j_0(t)}(U_{c_{t,j}}c_{t,j} + U_{l_{t,j}}l_{t,j}) = A_{t,j_0(t)},\tag{31}$$

where $A_{t,j_0(t)} \equiv U_{c_{t,j_0(t)}}(1 + r_{t,j_0(t)})a_{t,j_0(t)}$. It should be emphasized that implicit in this implementability constraint is the existence of age-dependent tax rates. Additional restrictions need to be imposed for an allocation to be implementable with age-independent taxes. In other words, the set of allocations from which the government can pick depends crucially on the instruments available to the government.

The Ramsey problem in this life-cycle economy consists of choosing an allocation to maximize the discounted sum of successive generations' utility, subject to each generation's implementability constraint (31) as well as the feasibility constraint (8) for $t = 0, \dots$:

$$\max_{\left\{\{c_{t,j}, l_{t,j}\}_{j=j_0(t)}^{J}, k_{t+J+1}\right\}_{t=-J}^{\infty}} \sum_{t=-J}^{\infty} \gamma^t W_t.$$

The pseudo-welfare function W_t is defined as in Section 2 to include generation t's implementability constraint in addition to its lifetime utility. If we let $\gamma^t \lambda_t$ be the Lagrange multiplier associated with generation t's implementability constraint (31), then the function W_t is defined as

$$W_t = \sum_{j=j_0(t)}^{J} \beta^{j-j_0(t)} \left[U\left(c_{t,j}, 1 - l_{t,j}\right) + \lambda_t (U_{c_{t,j}} c_{t,j} + U_{l_{t,j}} l_{t,j}) \right]$$

$$- \lambda_t A_{t,j_0(t)}. \quad (32)$$

As was the case in the infinitely-lived agent model, the government budget constraint (29) has been omitted from the Ramsey problem since it has to hold by Walras's law.

Prescriptions

In this section we show that the solution to the Ramsey problem generally features non-zero tax rates on labor and capital income and demonstrate how these rates vary with age. In particular, and in contrast with infinitely-lived agent models, if the Ramsey allocation converges to a steady state solution, optimal capital income taxes will in general differ from zero even in that steady state. Although the main results of this section hold more generally, we will restrict attention to steady states for ease of exposition.

Let $\gamma^t \phi_t$ denote the Lagrange multiplier associated with the time-t feasibility constraint (8). The steady state solution is characterized by the following equations:

$$1/\gamma = 1 - \delta + f_k, \quad (33)$$

$$-\frac{W_{l_j}}{W_{c_j}} = -\frac{U_{l_j}[1 + \lambda(1 + H_j^l)]}{U_{c_j}[1 + \lambda(1 + H_j^c)]} \leq z_j f_l, \quad \text{with equality if } l_j > 0, \quad (34)$$

$$\frac{W_{c_j}}{W_{c_{j+1}}} = \frac{U_{c_j}[1 + \lambda(1 + H_j^c)]}{\beta U_{c_{j+1}}[1 + \lambda(1 + H_{j+1}^c)]} = f_k + 1 - \delta, \quad (35)$$

where

$$H_j^c = \frac{U_{c_j,c_j} c_j + U_{l_j,c_j} l_j}{U_{c_j}}, \quad (36)$$

$$H_j^l = \frac{U_{c_j,l_j} c_j + U_{l_j,l_j} l_j}{U_{l_j}}, \quad (37)$$

as well as the feasibility and implementability constraints (8) and (31).

The first order condition with respect to capital, equation (33), implies that the solution to this Ramsey problem has the modified golden rule property: The marginal product of capital (net of depreciation) equals the discount rate

applied to successive generations $(1/\gamma - 1)$ (see Samuelson [1968]). Equation (34) equates the government's marginal rate of substitution between consumption and leisure of an age-j individual to the effective marginal product of labor of that same individual. Similarly, equation (35) sets the government's marginal rate of substitution between consumption today and consumption tomorrow equal to the return on capital (net of depreciation).

We now derive necessary conditions under which the Ramsey allocation features zero taxation of either labor or capital income. Any optimal fiscal policy has to satisfy the consumer's optimality conditions, and we derive the necessary conditions by comparing the optimality conditions from the consumer's problem to those of the Ramsey problem.

Combining the consumer's first order conditions for consumption (25) and labor (26), and applying them to the nontrivial case of positive labor supply, we obtain

$$-\frac{U_{l_j}}{U_{c_j}} = z_j w_j = z_j \hat{w}(1 - \tau_j^w), \tag{38}$$

which corresponds to the usual optimality condition that the marginal rate of substitution between labor and consumption be equal to the relative price of labor faced by the consumer. Compare equation (38) to its analogue from the Ramsey problem (34). Using the fact that $\hat{w} = f_l$, the tax rate on labor income for an age-j individual is given by

$$\tau_j^w = \frac{\lambda(H_j^l - H_j^c)}{1 + \lambda + \lambda H_j^l}. \tag{39}$$

Since λ is generally different from zero, this tax rate on labor income will be equal to zero only if $H_j^l = H_j^c$.

The same logic applies to the tax rate on capital income. For this case, consider the consumer's first order condition for consumption (25) at age j and $j + 1$. Using the consumer's first order condition for asset holdings (27), we get

$$\frac{U_{c_j}}{\beta U_{c_{j+1}}} = 1 + r_{j+1} = 1 + (1 - \tau_{j+1}^k)\hat{r}. \tag{40}$$

Equation (40) corresponds to the usual intertemporal condition that sets the marginal rate of substitution between consumption today and consumption tomorrow equal to the relative price of the same commodities, which is equal to the gross interest rate. The government's counterpart of (40) is given by equation (35). Using these two equations and the fact that $\hat{r} = f_k - \delta$, we obtain

$$\frac{1 + \hat{r}}{1 + (1 - \tau^k_{j+1})\hat{r}} = \frac{1 + \lambda + \lambda H^c_j}{1 + \lambda + \lambda H^c_{j+1}}, \qquad (41)$$

which implies that the tax rate on capital income is different from zero unless $H^c_j = H^c_{j+1}$. These results are summarized in the following proposition.

Proposition 3 *In our life-cycle economy, (i) the optimal tax rate on labor income at age j is different from zero unless $H^l_j = H^c_j$, and (ii) the optimal tax rate on capital income at age $j + 1$ is different from zero unless $H^c_j = H^c_{j+1}$.*

Proposition 3 is very similar to Proposition 1, but with age of individuals replacing the time period. Essentially, prescriptions that hold in the transition of infinitely-lived agent models hold in the steady state of life-cycle economies. Since zero-capital income tax is merely a special case in the transition of infinitely-lived agent models, we should expect the same prescription in the steady state of life-cycle economies.

The intuition as to why the celebrated Chamley-Judd zero-capital tax result does not extend to life-cycle economies should be clear. Since consumption and leisure are constant in the steady state of infinitely-lived agent models, H^c is constant; thus, zero-capital income taxation is optimal regardless of the form of the utility function. In contrast, consumption and leisure are generally not constant over an individual's lifetime in life-cycle models, even in steady state. There is in fact no reason to expect $H^c_j = H^c_{j+1}$, and consequently capital income taxes will generally not be equal to zero in the long run. Obviously, if the economy is specified so that individuals' behavior features no life-cycle elements, i.e., labor supply and consumption are independent of age, then optimal taxation works as in infinitely-lived agent models in the sense that capital income is not taxed.

Proposition 4 *In our life-cycle model, if (i) $z_j = z$, $j = 0, \ldots, J$ and (ii) $\gamma = \beta$, then it is not optimal to tax capital income in the long run.*

The proof of Proposition 4 is fairly intuitive. From the first order condition with respect to capital (33), when $\gamma = \beta$, the steady state return on capital coincides with individuals' rate of time preference, i.e., $f_k - \delta = \hat{r} = 1/\gamma - 1 = 1/\beta - 1$. The consumer's optimization conditions (25) through (27) then imply that, given a constant productivity profile, consumption and leisure do not depend on age in steady state. In turn, the absence of life-cycle behavior implies that the function H^c_j does not depend on age, which, following Proposition 3, is sufficient for the optimal capital income tax rate to be zero in steady state.

Proposition 4 is a generalization of a result in Atkeson, Chari, and Kehoe (1999) that they use to prescribe zero-capital income taxation in overlapping generations economies. It should be noted, however, that the conditions stated

in Proposition 4 have empirically unappealing implications. In particular, they imply that individuals consume their labor earnings period by period. As a result, individuals do not accumulate any assets and the entire stock of capital is owned by the government.

Optimal capital income taxes are also zero when preferences are such that uniform commodity taxation over the lifetime of individuals is optimal. The separable utility function (2) is one form under which the capital income tax is zero. Because this utility function is separable in consumption and leisure (so that $U_{c,l} = 0$) and has a constant intertemporal elasticity of substitution (so that $U_{c,c}c/U_c$ is constant), the general equilibrium elasticity (the function H^c) is constant. Relative to Proposition 2, the general equilibrium elasticity here is not only independent of time but of age as well.

Proposition 5 *Under utility functions of the form given by* (2)*, the optimal capital income tax in our life-cycle model is zero for $t > 1$.*

Proposition 5, combined with Proposition 2, sounds like good news. If we were confident that individuals' preferences were reasonably well represented by utility functions that are separable in consumption and leisure, then we would be reasonably confident that prescribing zero-capital income tax was the right thing to do. There is, however, an important caveat: The result in Proposition 5 relies heavily on the government's ability to age-condition tax rates. If the government were constrained to use tax rates that are independent of age, then the optimal capital income tax would no longer be zero.[18] Furthermore, applied work in public finance is usually conducted with utility functions that are not separable in consumption and leisure.[19]

Under the Cobb-Douglas utility function, it is straightforward to show that optimal capital income taxes in this case are zero in the long run only under very restrictive conditions, as stated in Proposition 4. The principles guiding the optimal manner in which to tax capital over the lifetime of individuals are stated in the following proposition:

Proposition 6 *For utility function of the form given by* (3)*, the tax rate on capital income at age $j + 1$ is positive if and only if $l_{j+1} < l_j$.*

---

[18] Although showing this result is beyond the scope of this paper, the intuition is that the government will use a non-zero capital income tax to imitate the optimal age-dependent labor income tax (see Erosa and Gervais [2000] for details). Alvarez, Burbidge, Farrell, and Palmer (1992) derive a similar result in a partial equilibrium setting. This type of finding is reminiscent of results in Stiglitz (1987) and Jones, Manuelli, and Rossi (1997), where the government taxes capital income when it is constrained to use tax rates that are independent of individuals' skill levels.

[19] Auerbach and Kotlikoff (1987), for example, use a nested CES utility function with intertemporal elasticity of substitution equal to 0.25 and intratemporal elasticity of substitution equal to 0.8. Both elasticities would have to be equal to unity for the CES utility function to be separable in consumption and leisure.

Proof. The proof follows directly from the definitions of H_j^c (equation (36)) under utility function (3). Since $H_j^c = -\sigma - \eta/(1 - l_j)$ we can rewrite equation (41) as

$$\frac{1 + \hat{r}}{1 + r_{j+1}} = \frac{1 + \lambda + \lambda(-\sigma - \eta/(1 - l_j))}{1 + \lambda + \lambda(-\sigma - \eta/(1 - l_{j+1}))}. \tag{42}$$

Notice that $\tau_{j+1}^k > 0$ if and only if

$$\frac{1 + \hat{r}}{1 + r_{j+1}} = \frac{1 + \hat{r}}{1 + (1 - \tau_{j+1}^k)\hat{r}} > 1. \tag{43}$$

From equations (42) and (43), we obtain that $\tau_{j+1}^k > 0$ if and only if $l_{j+1} < lj$.
∎

By taxing (subsidizing) capital, the government makes consumption and leisure in the future more (less) expensive than today. Proposition 6 suggests the government uses capital income taxes to smooth individuals' leisure and consumption profiles over their lifetimes. Under Cobb-Douglas utility, the share of consumption in an individual's total expenditures is constant, so that consumption and leisure always move together over time. If consumption and leisure are high tomorrow relative to today, then the government will tax the return on today's savings at a positive rate tomorrow. By doing so, the government gives individuals an incentive to consume more and save less today, and thus to consume less tomorrow.

An implication of the principle of optimal taxation developed in Proposition 6 is that capital income should not be taxed during retirement. This follows directly from the fact that labor supply is constant during retirement. Notice, however, that leisure time during retirement is taxed indirectly because the return on savings is taxed prior to retirement.

4.   CONCLUSION

We have reviewed optimal taxation in both an infinitely-lived agent model and a life-cycle economy. Our review shows that there is no consensus regarding the optimal tax on capital income. Although the optimal capital income tax is invariably zero in the long-run equilibrium of infinitely-lived agent models, the conditions under which that is the case in life-cycle economies are very stringent. Even under a separable utility function, the capital income tax will only be zero if the government has access to age-dependent labor income taxes. Furthermore, both models suggest that capital income should be taxed at non-zero rates during the transition unless individuals have separable preferences between consumption and leisure. Thus, the strong conclusions and recommendations of much of this literature must be treated with caution.

REFERENCES

Alvarez, Yvette, John Burbidge, Ted Farrell, and Leigh Palmer. "Optimal Taxation in a Life Cycle Model," *The Canadian Journal of Economics*, vol. XXV (February 1992), pp. 111–22.

Atkeson, Andrew, V.V. Chari, and Patrick J. Kehoe. "Taxing Capital Income: A Bad Idea," Federal Reserve Bank of Minneapolis *Quarterly Review,* vol. 23 (Summer 1999), pp. 3–17.

Atkinson, Anthony B., and Agnar Sandmo. "Welfare Implications of the Taxation of Savings." *The Economic Journal*, vol. 90 (September 1980), pp. 529–49.

——————, and Joseph E. Stiglitz. *Lectures on Public Economics*. New York: McGraw-Hill, 1980.

Auerbach, Alan J., and Laurence J. Kotlikoff. *Dynamic Fiscal Policy*. Cambridge: Cambridge University Press, 1987.

——————, and Jonathan Skinner. "The Efficiency Gains from Dynamic Tax Reform," *International Economic Review*, vol. 24 (February 1983), pp. 81–100.

Chamley, Christophe. "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives," *Econometrica*, vol. 54 (May 1986), pp. 607–22.

Chari, V. V., and Patrick J. Kehoe. "Optimal Fiscal and Monetary Policy," in J.B. Taylor and M. Woodford, eds., *Handbook of Macroeconomics, Vol. 1*. Amsterdam: North-Holland, 1999, pp. 1671–745.

Erosa, Andrés, and Martin Gervais. "Optimal Taxation in Life-Cycle Economies," Federal Reserve Bank of Richmond Working Paper 00-2, September 2000.

Jones, Larry E., Rodolfo E. Manuelli, and Peter E. Rossi. "On the Optimal Taxation of Capital Income," *Journal of Economic Theory*, vol. 73 (March 1997), 93–117.

Judd, Kenneth L. "Redistributive Taxation in a Simple Perfect Foresight Model," *Journal of Public Economics*, vol. 28 (October 1985), pp. 59–83.

Kotlikoff, Laurence J. "The A-K Model: Its Past, Present, and Future." NBER Working Paper 6684, August 1998.

Kydland, Finn E., and Edward C. Prescott. "Rules Rather than Discretion: The Inconsistency of Optimal Plans." *Journal of Political Economy*, vol. 85 (June 1977), pp. 473–91.

Lucas, Robert E., Jr., and Nancy L. Stokey. "Optimal Fiscal and Monetary Policy in an Economy without Capital," *Journal of Monetary Economics*, vol. 12 (July 1983), pp. 55–93.

Ramsey, Frank P. "A Contribution to the Theory of Taxation," *Economic Journal*, vol. 37 (1927), pp. 47–61.

Samuelson, Paul A. "The Two-Part Golden Rule Deduced as the Asymptotic Turnpike of Catenary Motions." *Western Economic Journal*, vol. 6 (March 1968), pp. 85–89.

Stiglitz, Joseph E. "Pareto Efficient and Optimal Taxation and the New New Welfare Economics," in A.J. Auerbach and M. Feldstein, eds., *Handbook of Public Economics, Vol. 2*. Amsterdam: North-Holland, 1987, pp. 991–1042.