Journal of Public Economics 28 (1985) 59–83. North-Holland

REDISTRIBUTIVE TAXATION IN A SIMPLE PERFECT FORESIGHT MODEL

Kenneth L. JUDD*

Department of Managerial Economics and Decision Sciences, J.L. Kellogg Graduate School of Management, Northwestern University, Evanston, IL 60201, USA

Received September 1983, revised version received May 1985

We investigate the redistributive potential of capital taxation in an intertemporal maximizing model of capital formation. First, even unanticipated redistributive capital taxation is severely limited in its effectiveness since it depresses wages. Second, under any convergent redistributive tax policy which maximizes a Paretian social objective, the capital income tax will converge to zero, independent of the factor supply elasticities. These results are independent of workers' holdings of capital.

1. Introduction

An important question in public finance is the ultimate incidence of a tax. One particularly interesting aspect of this question is the redistributive potential of capital income taxation: How much will the disincentive effects of capital income taxation on capital accumulation and the resulting loss in wages reduce the net redistribution to the workers, the presumed recipients? In this paper we examine the redistributive potential of capital income taxation in general equilibrium growth models.

Dynamic general equilibrium incidence of capital income taxation has been studied in various versions of the neoclassical growth model by Feldstein (1974), Grieson (1975), Stiglitz (1978), Boadway (1979), Bernheim (1982), and Homma (1981). These studies demonstrated that the incidence of a capital income tax may be significantly shifted to labor in the long run, reducing the redistributive potential of capital income taxation, but generally not eliminating it. The major shortcoming of these studies was their concentration on long-run effects, often ignoring the adjustment process to the steady state, which is only realized in the limit. The true incidence of any tax includes the incidence along the adjustment path as well as in the new steady state. In the

*The author would like to thank W.A. Brock for helpful conversations and access to unpublished notes, and D. Brito, L. Kotlikoff and anonymous referees for their helpful comments. Any remaining errors are the author's. Financial support from the NSF and J.L. Kellogg Graduate School of Management is gratefully acknowledged.

0047-2727/85/$3.30

absence of a utility function, as in neoclassical growth models, it is not clear how this dynamic incidence is to be valued. When the adjustment path is modeled in such models (as in Boadway and Bernheim) the results are sensitive to the discount rate, a parameter of intertemporal preferences which, in their models, does not affect savings behavior. In contrast, we examine dynamic general equilibrium incidence of capital taxation in a perfect foresight model of growth where capital accumulation is determined by the maximization of a dynamic utility functional for the owners of capital. In such a model we can both examine the anticipation effects absent in neoclassical growth models and calculate the dynamic marginal value of a tax change, taking into account the adjustment process.

We find two recurring themes, one expected and the other surprising. First, the inelastic short-run supply of capital makes both temporary and permanent unexpected increases in the redistributive tax on capital income attractive to the agents who possess less capital than the average holding. This makes it tempting for a relatively poor, but politically powerful, majority to impose unanticipated capital income taxes for the purposes of redistribution. Second, the long-run incentives are quite different since, if both workers and capitalists have the same rate of time preference in the steady state, *the optimal redistributive tax on capital income from the point of view of any agent is asymptotically zero, independent of long-run factor supply elasticities.* This last result stands in stark contrast with neoclassical growth analyses, such as Hamada (1967), Homma (1981), Pestieau and Possen (1978), and Stiglitz (1978), which seem to argue that some redistribution generally benefits workers even in the long run. Third, when we weigh the relative importance of the short- and long-run impacts, we find that the long-run effects are substantial, whereas Boadway showed that in neoclassical growth models the long-run regressive impacts are of lesser importance because the adjustment process is slow. Together, these results indicate that redistribution of income through capital income taxation is effective only if it is unanticipated and will persist only if policy-makers cannot commit themselves to low taxation in the long run. More generally, these results indicate that the true long-run burden of a factor income tax is not correctly represented by the long-run impact of the tax on the net-of-tax factor price, the usual index considered in the study of tax shifting.

Section 2 presents the equilibrium characterization of the inelastic labor supply model we initially examine. Section 3 then determines the short-run and long-run redistributive effects of capital income taxation when the workers do not save. Section 4 repeats this analysis when workers do participate in the capital market. In both cases we show that no redistributive taxation is desired by the workers if the classes have the same rate of time preference. Section 5 generalizes this result to the case of elastic labor supply and flexible rates of time preference. Section 6 concludes this study.

2. The model

Assume that we have an economy of a large fixed number of identical, infinitively-lived individuals. The common utility functional is initially assumed to be additively separable in time:

$$U = \int_0^\infty e^{-\rho t} u(c(t))\,dt,$$

where $c(t)$ is consumption of the single good at time t. To abstract from differences in taste and construct a model where equilibrium is determined by the evolution of aggregate capital and consumption, we will initially assume that workers and capitalists have the same *constant* elasticity of marginal utility, $\beta \equiv -u''(c)c/u'(c)$, and the same pure rate of time preference, ρ.

We will assume initially that labor is supplied inelastically by all. This is in keeping with the previous studies and is appropriate since our concern is with the income inequality due to wealth inequality. Also, our results indicate that capital income taxation is a poor instrument for redistribution because of the induced capital decumulation. Adding an elastic labor supply would reinforce our results, since capital decumulation would be reinforced by labor supply movements [see Judd (1984)]. At all times t, L^c units of labor are supplied inelastically by each capitalist and L^n units of labor are inelastically supplied by agents who do not participate in the capital market, consuming their wages at each moment. All are paid at a wage rate of $w(t)$. We normalize so that the total labor supply, L, is unity.

Physical capital will be the only asset in this economy. Let $F(k)$ be the concave production function giving output per unit of labor as a function of the aggregate capital–labor ratio, k. $k_i(t)$ will represent the ith capitalist's ownership of capital and k_{i0} is his initial endowment of capital at $t=0$. Therefore, $k=\sum_i k_i$ since labor supply is unity. Capital depreciates at a rate of $\delta>0$ and $f(k)$ is the national product net of depreciation. σ is the elasticity of substitution between capital and labor in the *net* production function.

We shall keep the institutional structure simple. Think of each capitalist as owning his own firm, hiring labor and paying himself a rental of $r(t)$ per unit of capital at t, gross of taxes and depreciation. It is straightforward to show that the alternative assumption of value-maximizing firms would be equivalent [e.g. see Brock and Turnovsky (1981)].

The government's role will be standard: at time t, it taxes capital income net of depreciation at a proportional rate $\tau(t)$, makes a lump-sum transfer of T_i^c to the ith capitalist, transfers totalling T^n to noninvestors, and consumes G units of the good, such consumption not affecting the demand of any agent for private consumption goods. For technical reasons, we assume that

$\tau(t)$ is constant for sufficiently large t. This assumption will ensure the existence of a steady state and convergence to that steady state.

The ith capitalist will choose his consumption path, $c_i(t)$, to maximize lifetime utility, subject to the instantaneous budget constraint, taking the wage, rental, and tax rates as given:

$$\underset{c_i(t),k_i(t)}{\text{maximize}} \int_0^\infty e^{-\rho t} u(c_i(t))\,dt \tag{1}$$

$$\text{s.t. } c_i + \dot{k}_i = wL^c + (r-\delta)(1-\tau)k_i + T_i^c, \quad k_i(0) = k_{i0}$$

(Time arguments are suppressed when no ambiguity results.) The basic arbitrage condition which must hold is

$$u'(c_i) \equiv \int_t^\infty e^{\rho(t-s)}(r-\delta)(1-\tau)u'(c_i)\,ds. \tag{2}$$

This states that along an optimum path, each capitalist is indifferent between an extra unit of consumption and the extra future consumption that would result from an extra unit of investment. Upon differentiation, (2) yields:

$$\dot{\alpha}_i = \alpha_i(\rho - (r-\delta)(1-\tau)), \tag{3a}$$

where $\alpha_i = u'(c_i)$. Since $u(c)$ is isoelastic, this implies that consumption follows:

$$\dot{c}_i = -c_i(\rho - (r-\delta)(1-\tau))/\beta. \tag{3b}$$

We shall also assume that the transversality condition at infinity holds:

$$(TVC_\infty) \lim_{t\to\infty} u'(c_i(t))\,e^{-\rho t} = 0. \tag{4}$$

This condition will be used below to insure that c_i and k_i remain bounded as $t\to\infty$ and is a necessary condition for the agent's problem if $u(\cdot)$ is bounded, which applies here since the net production function is bounded [see Benveniste and Scheinkman (1982)].

To describe equilibrium, we impose the equilibrium conditions,

$$r = f'(k) + \delta, \tag{5a}$$

$$w = f(k) - kf'(k), \tag{5b}$$

on (3) and the budget constraint, sum the resulting individual arbitrage conditions and capital accumulation equations, thereby yielding the equilibrium equations:

$$\dot{C} = -C(\rho - (1-\tau)f'(k))/\beta, \tag{6a}$$

$$\dot{k} = f(k) - C - L^n(f(k) - kf'(k)) - T^n - G, \tag{6b}$$

where C is total consumption by capitalists per worker and G is government consumption per worker. (6a) follows from substituting (5a) into (3) and summing over i. (6b) states that net aggregate investment equals net output less capitalist consumption, wages and transfers paid to nonsaving workers, and government consumption.

The pair of equations, (6), describes the equilibrium of our economy at any t such that C and k are differentiable. To determine the system's behavior at points where C may not be differentiable, we impose the equilibrium conditions on (2), yielding:

$$u'(c_i(t)) = \int_t^{\infty} e^{-\rho(s-t)} u'(c_i(s)) f'(k(s))(1-\tau(s))\, ds, \tag{7}$$

showing that $c_i(t)$ and $C(t)$ are continuous functions of time.

Since the equilibrium system has a saddlepoint structure, for any fixed collections of values for T^n, L^n, G, and τ, there is a unique steady state and for any k there is a unique stable path leading the economy to that steady state from k. We can also prove that that convergent path is the unique equilibrium. If the economy does not follow the unique path converging to the steady state, then, due to the saddlepoint stability nature of (6), either k becomes zero at a finite time, T_0, or k converges to the maximum sustainable capital stock, $\hat{k}$, where $f(\hat{k}) = 0$ defines $\hat{k}$. Integrating (3a) shows that

$$\alpha(t) = \alpha(0) \exp\left\{\int_0^t (\rho - \bar{r}(z))\, dz\right\},$$

where $\bar{r}(t) \equiv (1-\tau)\ f'(k(t))$. If k disappears at T_0, then consumption must be zero and u' infinite thereafter; however, our solution for α implies that u' must be zero after T_0 since $f'(0) = \infty$, a contradiction.

On the other hand, if k approaches $\hat{k}$, then $\bar{r}(t)$ is negative for all sufficiently large t, since $f'(\hat{k}) < 0$. However, $\ln(e^{-\rho t}\alpha)$ converges to $\lim_{t\to 0} (-\int_0^t (1-\tau) f'(k))$, which cannot be $-\infty$, as required by TVC_∞, if $f'(k)$ is asymptotically negative, implying a violation of TVC_∞. More intuitively, at some finite time f' is driven negative and consumption is driven arbitrarily near to zero. At such a point, a rational individual will realize that if he would stop accumulating capital he would be able to achieve greater consumption at every future time. Since capital can neither become zero at some finite time nor converge to the maximal sustainable level, the unique equilibrium once the policy parameters are constant is that path which converges to the steady state. Hence,

$$0<\lim_{t\to\infty} C(t), \qquad \lim_{t\to\infty} k(t)<\infty. \tag{8}$$

Since we assume that the policy parameters become constant after some time, T^f, this argument gives the locus of (C,k) points at T^f that are consistent with long-run equilibrium. The fact that C and k must be continuous functions of time implies that (6) must put the economy on that locus at T^f. Since the initial capital stock is fixed, this condition determines the initial aggregate capitalist consumption level [see the appendix in Judd (1985) for the proof of what is needed for our analysis]. The system of relations given by eqs. (6) and (7) and the inequality (8) therefore describes the general equilibrium of our economy at all times t.

The differences between this model and the neoclassical growth models previously used by Grieson, Hamada, Homma, Pestieau–Possen, and Stiglitz, are substantial. In neoclassical growth models current savings is a function solely of the current interest rate and current income, whereas in optimal growth models, current savings is a function solely of expected future returns and lifetime income. While it is not realistic to assume that people ignore future returns in making current savings choices since many investments, such as education, have predominately long-run returns, some may argue that this perfect foresight optimal growth model goes too far in the other direction. The debate over appropriate modelling of the formation of expectations applies to these problems as well as those of macroeconomics.

3. Case I: Workers do not save

In this section we turn our attention to a special case similar to the neoclassical savings models used, for example, by Bernheim, Boadway, Hamada, and Grieson — only capitalists save and only workers work. In terms of our notation, this means $L^c=0$ and $L^n=L=1$. Both assumptions are consistent with basic economic theory: we may assume that capitalists are on a corner of their labor supply decision due to their wealth, leisure being a normal good, and workers find neither saving nor borrowing valuable because of the transactions costs associated with small transactions. The crucial difference between our model and these neoclassical savings models is that capitalists' behavior is governed by the maximization of an intertemporal utility function. The case of all agents saving will be analyzed separately since the results are substantially different.

We will analyze only the case where all capital income tax receipts are redistributed uniformly among the workers. This is the only interesting case since there is no point in this model to taxing capital income and returning it to capitalists. Therefore, the equilibrium of this economy is described by

$$\dot{C} = -C(\rho - f'(k)(1-\tau))/\beta, \tag{9a}$$

$$\dot{k} = (1-\tau)kf'(k) - C. \tag{9b}$$

We first should note that the steady-state capital stock, k^{∞}, is a function of the constant steady-state tax rate τ, that relation being

$$f'(k^{\infty}) = \rho/(1-\tau). \tag{10}$$

In particular, the long-run capital supply curve is perfectly elastic, the net-of-tax return being ρ. We shall see below, however, that this is not the basis of our results. We examine this case because the essential points may be easily illustrated, and where a result appears not to be robust to more general utility functionals, we shall prove the desired general result.

3.1. Impact effects of tax changes

To study the total incidence of taxation, we examine the desirability of a tax on capital income from the point of view of the workers who will receive the revenue in the form of lump-sum transfers. We first address this question by assuming that the economy is in the steady state associated with a constant tax rate of τ on capital income and ask if the workers want to increase the current tax, increase the tax in the near future, and/or increase the tax rate in the distant future by a small increment. More precisely, if we consider $t=0$ to be the present, we want to compute the net impact on worker welfare of increasing the capital income tax rate at time $t>0$ by $\varepsilon h(t)$ for small ε. [For technical reasons we assume $h(t)$ is eventually constant.] That policy change is announced and enacted immediately. The new equilibrium is the solution to

$$\dot{C} = -C(\rho - f'(k)(1-\tau-\varepsilon h(t)))/\beta, \tag{11a}$$

$$\dot{k} = (1-\tau-\varepsilon h(t))\, kf'(k) - C. \tag{11b}$$

For any ε, the solutions of C and k in the system (11) can be expressed as $C(t,\varepsilon)$ and $k(t,\varepsilon)$, respectively. We are interested in determining the impact on investment and consumption of a marginal policy change, modelled as an increase in ε from an initial value of zero. These initial impacts of the change in ε are denoted:

$$k_\varepsilon(t) \equiv \frac{\partial k}{\partial \varepsilon}(t,0), \qquad C_\varepsilon(t) \equiv \frac{\partial C}{\partial \varepsilon}(t,0),$$

$$\dot{k}_\varepsilon(t) \equiv \frac{\partial}{\partial \varepsilon}\left(\frac{\partial k}{\partial t}\right)(t,0), \qquad \dot{C}_\varepsilon(t) \equiv \frac{\partial}{\partial \varepsilon}\left(\frac{\partial C}{\partial t}\right)(t,0).$$

To determine the impact of changing $\tau(t)$ by $\varepsilon h(t)$, we differentiate the system (11) with respect to ε, and evaluate the derivative at $\varepsilon=0$ and at the initial steady-state level of capital. [For a general treatment of this procedure, see Judd (1982, 1985).] The result is a system of linear differential equations:

$$\begin{pmatrix} \dot{C}_\varepsilon \\ \dot{k}_\varepsilon \end{pmatrix} = J \begin{pmatrix} C_\varepsilon \\ k_\varepsilon \end{pmatrix} - \begin{pmatrix} Cf'h(t)/\beta \\ h(t)k'f \end{pmatrix}, \tag{12}$$

where

$$J = \begin{pmatrix} 0 & (1-\tau)f''(k)C/\beta \\ -1 & (1-\tau)(f'(k)+kf''(k)) \end{pmatrix}.$$

and k and C are evaluated at their steady-state values corresponding to τ. The most convenient method of solving the linear differential equation system in (12) is by Laplace transforms.[1] Taking Laplace transforms of (12) yields an algebraic system in the transform variable, s, the solution of which is

$$\begin{pmatrix} \Psi_\varepsilon(s) \\ K_\varepsilon(s) \end{pmatrix} = (sI - J)^{-1} \begin{pmatrix} -Cf'H(s)/\beta + C_\varepsilon(0) \\ -H(s)\,kf' \end{pmatrix}, \tag{13}$$

where Ψ_ε, K_ε and H are Laplace transforms of C_ε, k_ε, and h, respectively. $C_\varepsilon(0)$ is the initial change in C per unit of ε made to ensure stability of the system. Our analysis extensively uses the eigenvalues of J, which are

$$\mu, \lambda = \frac{\rho}{2}\left(1 - \frac{\theta_L}{\sigma} \pm \sqrt{\left(1 - \frac{\theta_L}{\sigma}\right)^2 + \frac{4}{\beta}\frac{\theta_L}{\sigma}}\right), \tag{14}$$

where θ_L is labor's share of net output. Both eigenvalues are important to our analysis. λ is the rate of convergence to the steady state along the stable

[1]The Laplace transform of $g(t)$ is $G(s)$, where $G(s)=\int_0^\infty e^{-st}g(t)\,dt$. Intuitively, the Laplace transform of $g(t)$ is the present value of the stream $g(t)$ discounted at s.

manifold corresponding to a particular value of τ, and μ is the rate of divergence from the steady state off that stable manifold. Note that they are independent of τ and that $\mu > 0 > \lambda$. Also note that $\mu \gtrless \rho$ as $\beta \lessgtr 1$.[2]

Using Judd (1982, 1985) we find that TVC_∞ implies the stability of capital accumulation and the boundedness of $k_\varepsilon(\cdot)$ and $K_\varepsilon(\mu)$, implying after some manipulation[3] that

$$C_\varepsilon(0) = H(\mu)\frac{\rho}{1-\tau}\left(1 - \frac{\mu}{\rho}\beta\right)\frac{C}{\beta}. \tag{15}$$

Combining (13) with this solution of $C_\varepsilon(0)$ yields a complete solution for $\Psi_\varepsilon(s)$ and $K_\varepsilon(s)$.

It is straightforward to show that $\mu b/\rho \lessgtr 1$ as $\beta \lessgtr 1$. Therefore, (15) shows that capitalists may increase or decrease their consumption in response to a tax increase. This is not surprising since the income effect of lower future income on demand for goods today and the substitution effect due to today's goods becoming cheaper relative to tomorrow's goods act in different directions. If $\beta > 1$, capitalists have a strong preference for a smooth consumption path due to the high curvature of the utility function and the income effect dominates, resulting in less consumption today. If $\beta < 1$, the price effect dominates and consumption jumps up. In either case, the change in consumption is proportional to $H(\mu)$, the tax rate change discounted at the rate μ. Since $H(\mu)$ is greater as the tax increase continues for a longer time, we see that the *magnitude* of the change in consumption is greater for tax increases of greater duration, whereas the *sign* depends only on β.

Finally, we should note that an announcement of future higher taxes will reduce consumption if $\beta > 1$, thereby increasing investment at $t = 0$ if there is no tax change at $t = 0$. This possibility of current investment being stimulated by future taxation will be important in determining the desirability of future tax increases.

The impact on the workers' utility of this tax change can be calculated from the solution for $K_\varepsilon(s)$. Let c^w denote consumption of the representative worker. Workers consume their wages and the lump-sum transfer:

$$c^w(t,\varepsilon) = f(k) - kf'(k) + (\tau + \varepsilon h(t))kf'(k). \tag{16}$$

[2]Straightforward calculations also show that the following useful identities hold for the model of section 3:

$$\Delta \equiv (\rho - \mu)(\rho - \lambda) = \frac{\rho^2\theta_L}{\sigma}(1 - \beta^{-1}),$$

$$\frac{f'^2(1-\tau)}{\Delta\beta} = \frac{-\sigma}{(1-\tau)\theta_L(1-\beta)}.$$

[3]For details, see Judd (1985). The key detail is that $K_\varepsilon(\mu)$ must be finite since $\mu > 0$, but from (13) this is possible only if $c_\varepsilon(0)$ is that value which offsets the singularity of $\mu I - J$.

Since we are initially at the steady state associated with $\varepsilon=0$, we may differentiate as before, and find that

$$c_\varepsilon^w(t,0)=h(t)kf'-kf''k_\varepsilon+\tau k_\varepsilon(f'+kf''). \tag{17}$$

Eq. (17) decomposes the impact on worker consumption into its separate components. The first term, $h(t)kf'$, is the increment to tax revenues and resulting rebate to workers. The second term, $-kf''k_\varepsilon$, is the impact on the typical worker's wage of a change of k_ε in capital stock. The last term is the impact of the induced capital accumulation on tax revenues collected.

The change in utility of workers in terms of the good at $t=0$, y_ε^w, is equal to the discounted change in utility divided by the current marginal utility of consumption, $u'(c^w)$, i.e.

$$y_\varepsilon^w=\frac{\int_0^w e^{-\rho t}u'(c^w)c_\varepsilon^w(t,0)\,dt}{u'(c^w)}.$$

By combining (13), (15), and (17),

$$y_\varepsilon^w=kf'H(\rho)\left(1-\left\{1+\frac{H(\mu)}{H(\rho)}\left(\frac{\mu\beta}{\rho}-1\right)-\beta\right\}\left(\frac{\tau}{1-\tau}\frac{\sigma}{\theta_L}+1\right)\bigg/(1-\beta)\right), \tag{18}$$

where θ_K is capital's share of net output.

For tax increases of very short duration, $h(t)$ is one for small t and zero otherwise, and $H(\mu)H(\rho)^{-1}$ is approximately unity. Then

$$y_\varepsilon^w=kf'H(\rho)\left\{\frac{\beta}{1-\beta}\left(1-\frac{\mu}{\rho}\right)\left(\frac{\tau}{1-\tau}\frac{\sigma}{\theta_L}+1\right)+1\right\}. \tag{19}$$

Since $\mu\beta>\rho$ and $\mu>\rho$ if and only if $\beta<1$, utility increases if $\tau=0$, but falls for some positive τ. Hence, a tax increase of short duration will always be desired by the workers if the economy is in the untaxed steady state, but will not be desirable if τ is sufficiently large. This follows from the assumed capital market imperfections: workers are not able to save any of the proceeds from a tax increase, and at high tax rates prefer to keep the capital producing and in the capitalists' hands rather than consume it.

Second, if $h(t)=1$, i.e. a permanent tax increase is enacted, then $H(\mu)H(\rho)^{-1}=\rho/\mu$ and

$$y_\varepsilon^w=\frac{kf'}{\rho}\left(\frac{\rho/\mu-1}{1-\beta}\left(\frac{\tau}{1-\tau}\frac{\sigma}{\theta_L}+1\right)+1\right). \tag{20}$$

Again, utility increases if $\tau=0$, but falls for large τ.

The third case, that of a permanent tax increase which beings at some future time T is more complex. Such a tax increase is represented by $h(t)$ being one for $t > T$ and zero otherwise, where T is understood to be large. Such a tax change is partially anticipated since the change was not anticipated before $t=0$, but is known before it takes effect at $t=T$. The fact that $\mu > \rho$ if and only if $\beta < 1$ is important for our net gain calculation in this case since $H(\mu)H(\rho)^{-1}$ goes to zero as the imposition of the tax is pushed into the future if $\mu > \rho$, whereas if $\mu < \rho$, then $H(\mu)H(\rho)^{-1}$ diverges to infinity as the tax is delayed. These observations immediately lead to the determination of the desirability of imposing a tax which takes effect only in the very distant future. First, if $\beta < 1$, then $\mu > \rho$ and for distant tax increases the $H(\mu)H(\rho)^{-1}$ term becomes negligible. It follows from (18) that for large T, utility is unchanged if $\tau = 0$ initially, and falls if $\tau > 0$. We therefore see that if capitalists have a small elasticity of marginal utility, workers today will not want to impose a partially anticipated tax increase on the capitalists in the distant future, even if the revenues are distributed to the workers. Note that this is also the case where capitalists will increase current consumption in response to an increase in expected future taxation. This capital decumulation in response to future taxation leads to a decline in wages in the near term, offsetting the revenue gain of the tax increase.

On the other hand, if $\beta > 1$, then $\mu\beta > \rho$ and workers will want anticipated redistributive taxation in the distant future. This can be seen from (18) by noting that for distant tax increases, $H(\mu)H(\rho)^{-1}$ will be large and dominate (18), and utility will increase for distant tax increases if τ is initially zero but fall if we are in a steady state associated with a high tax rate. Note that $\beta > 1$ is also the case where capitalists save in response to future tax increases, with the short-run immediate capital accumulation raising wages immediately. Hence, if we are in the untaxed steady state, this short-run wage effect is an additional benefit of the distant tax increase.

In summary, we have proved theorem 1.

Theorem 1. In the steady state corresponding to no taxation, workers will want a perfectly anticipated increase in the capital income tax in the distant future if and only if $\beta > 1$. Also, they will always want both temporary and permanent tax increases which begin immediately. In steady states associated with sufficiently high tax rates, workers will desire immediately enacted temporary and permanent tax decreases.

Theorem 1 tells us exactly when the revenue gains from these kinds of tax increases will be exactly offset by the wage losses due to the induced capital decumulation. It is not surprising that β be an important factor since it influences the rate at which the capitalists respond to tax increase and how they respond immediately to a tax change. The main point of theorem 1 is that unanticipated immediate tax increases may be useful for redistributive purposes if the current tax is not too high, whereas perfectly anticipated tax increases are possibly more likely to harm workers.

To gain some perspective on the quantitative nature of our analysis, we next ask how high τ must be for the workers to want a reduction in τ. Let τ^c be the critical value of the tax rate such that workers will want neither a permanent unanticipated decrease nor increase in τ when the economy is in the steady state corresponding to τ^c. τ^c is computed by solving for τ when (20) is equated to zero. If $\tau = \tau^c$ then the workers will want increases, whereas if $\tau > \tau^c$, workers will want a decrease. Table 1 gives values of τ^c for various values of σ and β, assuming $\theta_K = 0.25$. The values of σ and β represent the range of estimates from Lucas (1969), Berndt and Christensen (1973), Hansen and Singleton (1982), and Weber (1970, 1975). Note that τ^c is neither trivial nor unrealistically large, ranging from 0.40 to 0.80. τ^c is greater as β is larger and σ is smaller, an intuitive result since capital decumulation is slower in both cases.

Table 1

β \ σ	0.3	0.5	0.8	1.0	1.3
0.5	0.45	0.43	0.41	0.40	0.40
1.0	0.50	0.50	0.50	0.50	0.50
2.0	0.56	0.58	0.59	0.60	0.61
5.0	0.58	0.64	0.70	0.73	0.75
10.0	0.60	0.68	0.77	0.81	0.84

The entry corresponding to each σ–β pair is τ^c if workers cannot save. The calculations assume that $\theta_K = 0.25$.

The neoclassical model most similar to our model is the two-class model where workers save nothing and capitalists save a fixed proportion of after-tax income. This neoclassical model looks like our with regard to workers' savings and also implies 100 percent shifting of the capital income tax, that is, the long-run net return on capital is unaffected by the tax. However, in our dynamic optimization model workers may gain from a redistributive tax. Since capital is fixed in the short run, we expect a short-run tax increase to benefit workers. What is surprising is that under some conditions workers will benefit from a tax increase which comes only in the very distant future where this 100 percent shifting occurs. This effect is absent in neoclassical growth models since they have no anticipation effects. It is intuitive that a sufficiently concave capitalist utility function will yield this result since it implies that the capital stock adjusts slowly to tax changes and the tax burden is slowly shifted to labor.

3.2. The optimal redistributive taxation

These impact analyses of long-range tax changes lead us to inquire as to the long-run nature of an optimal tax on capital from the viewpoint of a social planner. Let $\bar{r}(t)$ be the rate of return net of both taxes and depreciation realized by capitalists at t. For any such tax law, the laws of motion for the capitalist class are constraints from the point of view of a policymaker and are given by

$$\dot{k}=\bar{r}k-C, \tag{21a}$$

$$\dot{C}=-C(\rho-\bar{r})/\beta, \tag{21b}$$

$$0<\lim_{t\to\infty} k(t), C(t)<\infty. \tag{21c}$$

Suppose that a government concerned only with the welfare of the workers determines a tax policy for all future time and that such revenues must cover a constant stream of government consumption, G, as well as lump-sum transfers to the workers. Furthermore, suppose that it has only two instruments: capital income taxation and lump-sum transfers to or taxation of workers. If the capital income tax does not raise enough to finance G, then lump-sum taxes are imposed on the workers. (Since an equal amount of labor is supplied inelastically by all workers, this analysis would be unchanged if we allowed labor taxes.) If γ is the relative weight put on worker welfare by the social planner, his control problem becomes:

$$\max_{\bar{r}(t)} \gamma \int_0^\infty e^{-\rho t}u(f(k)-\bar{r}k-G)\,dt+(1-\gamma)\int_0^\infty e^{-\rho t}u(\bar{r}k)\,dt$$

$$\begin{aligned}\text{s.t.}\quad &\dot{k}=\bar{r}k-C\\ &\dot{C}=-C(\rho-\bar{r})/\beta\\ &\bar{r}\geqq 0\\ &0<\lim_{t\to\infty} C(t), k(t)<\infty.\end{aligned} \tag{R}$$

The current-value Hamiltonian for this problem is

$$H=\gamma u(f-\bar{r}k-G)+(1-\gamma)u(\bar{r}k)+q_1(\bar{r}k-C)-q_2C(\rho-\bar{r})/\beta+\eta\bar{r}, \tag{22}$$

where q_1, q_2, and η are the current-value multipliers of the state variables k and C, and the $\bar{r}\geqq 0$ constraint, respectively. The laws of motion for solutions to this problem are

$$\dot{q}_1 = \rho q_1 - u'(f' - \bar{r}) - q_1 \bar{r}, \tag{23a}$$

$$\dot{q}_2 = \rho q_2 + q_1 + q_2(\rho - \bar{r})/\beta, \tag{23b}$$

$$0 = -\gamma u'(f - \bar{r}k - G) + (1-\gamma)u'(\bar{r}k) + q_1 k + q_2 C/\beta + \eta, \tag{23c}$$

$$0 = \eta \bar{r}. \tag{23d}$$

The steady-state conditions for this problem are therefore

$$\dot{k} = 0, \tag{24a}$$

$$\dot{C} = 0, \tag{24b}$$

$$\dot{q}_1 = \dot{q}_2 = 0. \tag{24c}$$

Conditions (21b) and (24b) imply that $\bar{r} = \rho$. This, together with (23a) and (24c), implies that $f' = \bar{r}$. Hence, $f' = \rho$ and we have proved theorem 2.

Theorem 2. If the redistributive capital taxation program maximizing a Paretian social welfare function converges, if both classes have the same pure rate of time preference, and it only capitalists save, then the optimal capital income tax vanishes asymptotically. Specifically, there should be no redistribution in the limit and any government consumption should be financed by lump-sum taxation of workers.

It should be kept in mind that this does not just follow from efficiency considerations alone, i.e. from the domination of lump-sum taxation over distortionary taxation. This is because theorem 1 shows that some redistributive taxation is desirable for workers, and hence is present in the solution of (R). Theorem 2 says that in any convergent program, redistributive capital taxation disappears asymptotically, even if the social planner cares only about the workers.

A related fact is that any gain in worker utility from any change in tax policy must come from an unanticipated jump in the capitalists' shadow price of capital, $u'(c)$. If $u'(c)$ does not change in response to a policy change then $C_\varepsilon(0) = 0$, implying that $H(\mu)(-\mu\beta/\rho + 1)$ must be zero, and that y_ε^w equals $-H(\rho)kf'\sigma\tau/(\theta_L(1-\tau))$, which is negative if τ is positive and $h(t)$ is non-negative, and zero in the untaxed steady state. This demonstrates lemma 1.

Lemma 1. If the workers were limited to choosing a policy which left the capitalists' shadow value of capital, $u'(c)$, initially unchanged, then there is no first-order gain to workers' utility if they are in the untaxed steady state and there is a first-order loss if they are in a taxed steady state.

Lemma 1 shows that the gain from imposing a tax in the distant future is solely due to the unanticipated nature of that change. Had that change been anticipated, there would be no jump in the marginal utility of consumption of capitalists since it is the private shadow price for capital and shadow prices must be continuous along any anticipated path.

The neoclassical analysis closest in spirit to this exercise was carried out by Hamada (1967). He examined the optimal transfer from capitalists to workers when workers cannot save and capitalists have a fixed savings rate, s. He showed that if the initial capital stock was small then the workers would accept a small transfer until a critical level of capital stock, k^*, was reached at which point the transfer is increased to $\rho k^*/s>0$ and capital stock becomes stationary. While confirmation awaits numerical analysis, our analysis indicates that the optimal program here has a quite different character, with large transfers initially and no transfer asymptotically. Theorem 2 shows that there would be no transfer asymptotically and theorem 1 indicates that in the short run some transfer is desirable for the workers. Therefore, the intertemporal pattern of transfers differs between the intertemporal maximization framework and the neoclassical savings framework.

The obvious weaknesses of our optimal redistributive tax analysis are that we have no idea as to how long it will take to reach the zero tax on capital income and how much redistribution is accomplished in terms of lifetime utility. It is doubtful that there are any tractable robust examples where one can explicitly calculate the optimal tax schedule. Resolution of these questions await numerical analysis which could give some insight.

4. Case II: All agents own capital

We next examine the case where all agents participate in the capital market and differ only in their endowments of capital. In particular, all agents supply one unit of labor. Minor adjustments in the above equilibrium analysis show that the equilibrium equations will aggregate to

$$\begin{aligned} \dot{C} &= -C(\rho-(1-\tau)f'(k))/\beta, \\ \dot{k} &= f(k)-C-G, \end{aligned} \tag{25}$$

where C is now mean consumption of all agents and k is still the aggregate capital–labor ratio. Again we suppose that we are in the steady state associated with some constant rate of taxation and that agents will have to evaluate the desirability of various tax changes. The perturbation analysis conducted above when applied here yields:

$$K_\varepsilon(\rho)=\frac{C}{\beta}f'\frac{H(\rho)-H(\mu)}{(\rho-\lambda)(\rho-\mu)}, \tag{26}$$

where $\varepsilon h(t)$ again is the increase in the tax rate at t, and where now λ and μ are the negative and positive eigenvalues of the linearization of (28) around its steady state. One feature of this system which differs from the case of no savings by workers is that $\rho<\mu$ always;[4] hence $H(\rho)$ exceeds $H(\mu)$ in magnitude whenever $h(t)$ is of one sign. More intuitively, this fact says that the rate of divergence away from the steady state along any divergent path exceeds the rate of discount. If taxes are increased, then $H(\rho)>H(\mu)$. Since $\lambda<0<\rho<\mu$, (26) then shows that $K_\varepsilon(\rho)$ would be negative, that is, a tax increase causes a decline in the discounted value of the capital stock.

We assume that the revenues are lump-sum rebated uniformly to all agents since all agents are both capitalists and workers. For each, the discounted value of the change in wages plus the change in rebate is equal to

$$kf'H(\rho)+K_\varepsilon(\rho)f'(\tau+(1-\tau)\theta_L/\sigma).$$

However, an agent possessing k^w units of capital losses $H(\rho)k^w f'$ on the capital he holds. His net gain is therefore equal to

$$y_\varepsilon^w=(k-k^w)f'\left(H(\rho)+(1-\tau)\frac{K_\varepsilon(\rho)}{k}\frac{\theta_L}{\sigma}\right)+\tau K_\varepsilon(\rho)f'. \tag{27}$$

Theorem 3 follows from direct computation as did theorem 1.

Theorem 3. Suppose that the economy is in the steady state associated with the current tax rate. An immediate, permanent tax increase is desired by an agent if and only if his capital holding is less than the average holding and the current tax rate is sufficiently small. An immediate temporary tax increase is desired in any steady state if his capital holding is below average and the duration is sufficiently short. No agent would find a perfectly anticipated tax increase in the distant future desirable.

The major difference between theorems 1 and 3 is the unanimous agreement that a perfectly foreseen tax on capital income is undesirable. This follows from (26), (27), and the fact that μ exceeds ρ always. Intuitively, no agent will want a tax increase in the distant future if the economy is currently in the untaxed steady state since the tax increase will cause so much capital decumulation in the short run, since $\mu>\rho$, that the wage decrease offsets the future subsidy for all, even for those owning no capital.

[4]The eigenvalues in this model are

$$\mu,\lambda=\frac{\rho}{2(1-\tau_K)}\left(1\pm\sqrt{1+4(1-\tau_K)\theta_L/(\sigma\theta_K\beta)}\right)$$

In the interest of keeping the notation clean, we have defined μ and λ twice. This is excusable here because it will always be clear from context which μ (positive eigenvalue) and which λ (negative eigenvalue) is meant.

We next use (27) to get a quantitative handle on just how high τ can be before a permanent immediate decrease is desirable. Agents' preferences will depend on their relative wealth since the sign of y_ε^w depends on $\theta^w \equiv k^w/k$. Table 2 shows the tax rate, τ^c, such that if the current capital stock is the steady-state level for τ^c then an agent holding θ^w times as much capital as the average will have his utility unaffected by a permanent and immediate decrease in τ, assuming various values for β and σ. Again, we assume $\theta_K = 0.25$. We again see the same biases towards a high τ^c, but note that these critical tax rates are lower here than in the model where workers do not participate in the capital markets. The appropriate comparison is between table 1 and the $\theta^w = 0$ columns of table 2 since the agent being examined holds no capital in both cases. The lower value of τ^c when the worker can invest is intuitive since a tax cut in that case will induce him to invest and accumulate, whereas if he cannot invest, the gain of a tax cut is reduced.

Table 2

	$\sigma =$ 0.4			0.65			1.00		
β	$\theta^w =$ 0.0	0.5	0.9	0.0	0.5	0.9	0.0	0.5	0.9
0.5	0.35	0.20	0.05	0.28	0.17	0.04	0.25	0.15	0.04
1.0	0.40	0.27	0.07	0.36	0.23	0.06	0.33	0.21	0.05
3.0	0.55	0.41	0.13	0.53	0.38	0.12	0.51	0.35	0.11

The entry corresponding to each σ-β-θ^w triple is τ^c when workers save. Again, $\theta_K = 0.25$ is assumed.

At this point, we should compare our results with earlier neoclassical growth analyses, as in Feldstein, Grieson, Boadway, and Bernheim. Both approaches assume a concave production function with no adjustment costs; the only difference is in their ad hoc savings specifications versus our dynamic optimization approach to saving behavior. While the Feldstein and Grieson analyses indicated substantial steady-state losses for all from capital taxation, Boadway shows that the adjustment process is so slow that the short-run gains may dominate at plausible discount rates. In neoclassical growth analysis, the discount rate plays no role in determining investment, so when it comes to present value calculations one has to be introduced independent of saving behavior. Therefore, the two approaches are somewhat difficult to compare. However, the capital stock dynamics are qualitatively similar since in both cases the capital stock converges linearly to a new steady state. The differences are that the supply of capital stock in the long run is more sensitive in our analysis, where the long-run elasticity with respect to net return is infinite, than in neoclassical models, and also the convergence to the steady state is more rapid in our examples. The last point holds since in our model the half-life of the adjustment process varies from 5 to 15 years, whereas Boadway's examples have half-lives of roughly 20–30 years, where we use the same production function in this comparison. These differences explain why the redistributive potential of capital taxation differs in these dynamic analysis.

Next we examine the asymptotic character of the optimal redistributive tax from the perspective of any agent. Theorem 4 is shown exactly as theorem 2 was proved.

Theorem 4. If workers and capitalists have the same constant rate of time preference, and both have access to perfect capital markets, then the optimal redistributive tax on capital for any worker is asymptotically zero if it converges.

Theorem 4 implies that, in a Pareto-efficient program, there is no effort to equalize income asymptotically and should be compared to the analysis of Pestieau and Possen (1978). They analyzed a model where a social planner has an instantaneous utility function over average consumption and the distribution of income, which is discounted at a constant rate. They assume, however, that private investment is described by a constant savings rate. They find that with labor taxation, capital taxation and bonds, there will be no income inequality asymptotically if the marginal value of income equality is positive in the steady state. Since it is straightforward to show that theorem 4 remains true even if bonds are allowed, we see that the Pestieau–Possen results depend critically on the constant savings rate formulation. If bonds are not available to the planner, then Pestieau and Possen show that some redistribution will generally be desired asymptotically, whereas theorem 4 shows the opposite for out model. Again we see that the character of optimal redistribution changes substantially when the analysis is conducted in an intertemporal maximizing framework.

In discussing neoclassical models, we have concentrated on fixed savings rate models. If the savings rate were variable, as in Grieson (1975), when one solves the optimal tax program for the worker, one finds that the long-run redistributive capital income tax is less. However, it disappears only when workers do not discount or when the savings function is infinitely elastic (see the appendix). Therefore, the only way for the neoclassical models to behave in the long run like ours is to assume no discounting by workers, an absurdity, or to assume an absurdly high short-run savings elasticity.

5. Pareto-efficient taxation

Heretofore we have assumed constant rates of time preference and an inelastic labor supply. This was an appropriate simplifying assumption for our comparative dynamics analysis. However, the strong results of the optimal taxation analyses may look special and sensitive to these special assumptions. Therefore we next turn to the problem of Pareto-efficient taxation with two classes of infinitely-lived agents with elastic labor supplies and heterogeneous and flexible time preferences. We demonstrate that the results of theorems 2 and 4 are due to neither the inelastic labor supply, the infinitely elastic long-run supply of capital, nor the possibility of lump-sum taxation of workers to finance government consumption. We find that in any convergent Pareto-efficient tax program the tax rate on capital income is asymptotically zero, showing that the capital income tax has no role for either redistributive or efficiency purposes in the long run.

We assume that an individual in class i, $i=1,2$, has a utility functional of the form:

$$U=\int_0^\infty e^{-R^i}u^i(c^i,l^i,X^i)\,dt,$$

where $\dot{R}^i=\phi^i(c^i,l^i,X^i)$ is the instantaneous rate of time preference as a function of representative class i consumption, c^i, and class i labor, l^i. This is a generalization of Uzawa (1968). R represents a cumulative discount rate, and X represents a state variable for cumulative consumption and possibly affects the contemporaneous marginal rate of substitution between consumption and leisure. This form is intended to represent a rich class of intertemporal preferences. Everything below holds if X were a vector, but we assume X to be a scalar to economize on notation.

Suppose that the representative class i agent holds $K^i(t)$ units of capital at t. Then he solves the following problem (for convenience, we drop the i superscripts at this point):

$$\max_{c,l}\int_0^\infty e^{-R}u(c,l,X)\,dt$$

$$\text{s.t.}\quad \dot{K}=\bar{r}K-c+\bar{w}l,$$

$$\dot{R}=\phi(c,l,X),$$

$$\dot{X}=\psi(c,l,X).$$

The present-value Hamiltonian for his problem is

$$H=e^{-R}u(c,l,X)+q_1(\bar{r}K-c+\bar{w}l)+q_2\phi(c,l,X)+q_3\psi(c,l,X), \tag{28}$$

where q_1, q_2, and q_3 are the costates for the state variables, R, K, and X, respectively. The equations of motion for the optimal path are

$$\dot{Q}_1 = Q_1(\phi(c,l,X) - \bar{r}), \tag{29a}$$

$$\dot{Q}_2 = u(c,l,X) + Q_2\phi(c,l,X), \tag{29b}$$

$$\dot{Q}_3 = Q_3(\phi(c,l,X) - \psi_X(c,l,X)) - u_X(c,l,X) - Q_2\phi_X(c,l,X), \tag{29c}$$

$$0 = u_c - Q_1 + Q_2\phi_c + Q_3\psi_c, \tag{29d}$$

$$0 = u_l + Q_1\bar{w} + Q_2\phi_l + Q_3\psi_l, \tag{29e}$$

where $Q_i = q_i \mathrm{e}^R$, $i = 1,2,3$, are the current value costates. $\phi(c,l,X) = \bar{r}$ and $0 = \psi(c,l,X)$ in the steady state of this system, showing that the steady-state net return to capital may vary with steady-state consumption and labor. Hence, the long-run factor prices are not fixed, and long-run supply curves of both factors may have finite and nonzero elasticities. [See Uzawa (1968) for a more complete analysis of such utility functionals.]

We assume that the government has a social welfare function which is a positively weighted average of individual utilities. We also assume that wage taxes may be imposed, and that class-specific lump-sum rebates are allowed, but no lump-sum taxes. Let $\bar{w}$ be the after-tax wage and S_i be the lump-sum rebate for class i. The government's problem is then to choose $\bar{r}(t)$, $S_i(t)$ and $\bar{w}(t)$ so as to maximize social welfare, subject to the constraint that the economy is in equilibrium and that current revenues cover current rebates and current government consumption, G. (The addition of a bond market changes no asymptotic result and is assumed away to eliminate the possibility that our results hold because revenue is zero asymptotically.) That problem is:

$$\max_{S_i, \bar{r}, \bar{w}, c^i, l^i} \gamma \int_0^\infty \mathrm{e}^{-R^1} u^1(c^1, l^1, X^1)\,\mathrm{d}t + (1-\gamma) \int_0^\infty \mathrm{e}^{-R^2} u^2(c^2, l^2, X^2)\,\mathrm{d}t \tag{S}$$

$$\begin{aligned}
\text{s.t.}\quad & \dot{K}^i = \bar{r}K^i + \bar{w}l^i - c^i + S_i, \\
& \dot{R}^i = \phi^i(c^i, l^i), \\
& \dot{X}^i = \psi^i(c^i, l^i, X^i), \\
& \dot{Q}^i_1 = Q^i_1(\phi^i(c^i, l^i) - \bar{r}), \\
& \dot{Q}^i_2 = u^i(c^i, l^i, X^i) + Q^i_2\phi^i(c^i, l^i, X^i), \\
& \dot{Q}^i_3 = Q^i_3(\phi^i - \psi^i_{X^i}) - u^i_{X^i} - Q_2\psi^i_{X^i}, \\
& 0 = u^i_{c^i} - Q^i_1 + Q^i_2\phi^i_{c^i} + Q^i_3\psi^i_{l^i},
\end{aligned}$$

$$0 = u_l^i - Q_1^i \bar{w} + Q_2^i \phi_l^i,$$

$$0 = (f(k) - \bar{r}k - \bar{w})(l^1 + l^2) - S_1 - S_2 - G,$$

$$S_i \geqq 0,$$

where $i = 1, 2$. Also γ, the social weight on class 1, is between 0 and 1, and $k = (K^1 + K^2)/(l^1 + l^2)$ is the aggregate capital–labor ratio.

The equations of motion for the optimal problem include

$$\dot{\pi}_i = -\bar{r}\pi_i - (f' - \bar{r})m, \tag{29}$$

where π_i is the social costate of K^i, $i = 1, 2$, and m is the social shadow price of the balanced budget constraint, hence nonzero if G is positive. The value of (S) presumably declines as G increases since increases in G reduce the feasible set. Therefore we assume $m > 0$.

Define the current values of π_1, π_2, m:

$$\Pi_i \equiv \pi_i \mathrm{e}^R, \quad i = 1, 2; \qquad M \equiv m\,\mathrm{e}^R.$$

Then the equations of motion in (29) can be expressed as

$$\dot{\Pi}_i = (\phi^i - \bar{r})\Pi_i - (f' - \bar{r})M. \tag{30}$$

In a steady state of the optimal problem,

$$\dot{Q}_1^i = 0 \Rightarrow \bar{r} = \phi^i(c^i, l^i), \quad i = 1, 2, \tag{31}$$

$$\dot{\Pi}_i = 0 \Rightarrow (f' - \bar{r})M = 0, \tag{32}$$

which implies that $\bar{r} = f'$ if M is positive. This demonstrates theorem 5.

Theorem 5. If the solution to (S) converges to a steady state where the shadow price of G is positive, then the tax rate on capital is zero in the long run.

Note that convergence to a steady state implies that the rate of time preference for the two classes must be equated. We are therefore implicitly assuming that $\phi^1(c^1, l^1, X^1)$ and $\phi^2(c^2, l^2, X^2)$ can be equal for some values of c^1, c^2, l^1, l^2, X^1, and X^2 consistent with the steady state. To allow different steady-state discount rates would necessitate the imposition of borrowing constraints on the class with the higher discount rate, violating the perfect capital market spirit of this exercise.

We next give an intuitive explanation for theorem 5 along the lines suggested by Bradford (1974). Using differentials, (30) can be rewritten:

$$\Pi_i(t) = (1 - \phi^i \,\mathrm{d}t)((1 + \bar{r}\,\mathrm{d}t)\Pi_i(t + \mathrm{d}t) + (f' - \bar{r})M(t + \mathrm{d}t)), \tag{33}$$

where negligible terms of order $(\mathrm{d}t)^2$ have been included to help with the intuition. The term $\Pi_i(t)$ is the social value of class i's capital stock at the end of period t. If there were an extra unit of capital in class i's hands at the beginning of period $t+\mathrm{d}t$ social welfare would rise by $\Pi_i(t+\mathrm{d}t)(1+\bar{r}\,\mathrm{d}t)$ since the end-of-period stock rises by the initial increment plus class i's net income in period $t+\mathrm{d}t$, $\bar{r}\,\mathrm{d}t$. At $t+\mathrm{d}t$, capital tax revenue increases by $(f'-\bar{r})\,\mathrm{d}t$, improving welfare by $(f'-\bar{r})M(t+\mathrm{d}t)\,\mathrm{d}t$. The sum of these $t+\mathrm{d}t$ values discounted by *class* i's discount factor, $1-\phi^i\,\mathrm{d}t$, equals the social value of K^i at t, $\Pi(t)$. In particular, this says that society should use class i's discount rate when evaluating changes in K^i.

In steady state $\phi^i = \bar{r}$, implying that $(1-\phi^i\,\mathrm{d}t)(1-\bar{r}\,\mathrm{d}t) = 1$ to $O(\mathrm{d}t)$. Therefore, (33) reduces to $0 = f' - \bar{r}$ if $M > 0$. We see that this simple present value calculation yields theorem 5. The crucial detail to note is that there is no one social discount rate that is appropriate to use in evaluating marginal changes in capital stock but rather that class i's rate should be used when evaluating K^i.

Theorem 5 shows that no Paretian social welfare function desires redistributive capital taxation in the long run, independent of long-run factor supply responses, as long as the optimal program converges to a steady state of consumption, leisure and assets. The assumption of equal long-run discount rates is reasonable in this context since it is a necessary condition for the existence of a steady-state distribution of capital.

Chamley (1980) came to the same conclusion for the special case of ϕ being constant, implying an infinitely elastic long-run supply curve of capital, and of a single class, thereby examining efficient taxation only. We see that the zero long-run interest tax result is quite robust, even when we allow the possibility of redistribution, heterogeneity in tastes, and arbitrary long-run elasticities of supply for both factors. Given the arbitrary nature of the u^i and ϕ^i functions, it is clear that the asymptotic distribution of wealth and income could be highly unequal. Yet the steady state of the optimal tax program involves no capital income tax. This forcefully shows that capital income taxation is useless as a redistributive tool in the long run for this broad class of utility functions, under the assumption of stability.

Another paper which reaches a similar conclusion is Brito (1981). He shows in an intergenerational model that the optimal tax program will eventually not tax life-cycle capital but will eliminate bequests, independent of how the planner values different generations. Brito adopts a utilitarian social welfare function when valuing the utility of individuals within a generation. Here we get the no capital tax result when we have an arbitrary *intragenerational* social welfare function. Together, these papers indicate the generality of the asymptotic inefficiency of capital income taxation.

This result for infinitely-lived agents model should also be compared to comparable analysis for the two-period life overlapping generation model. Atkinson and Sandmo (1980), among others, have shown (assuming global asymptotic stability) that in the steady state of the optimal policy, the tax on interest depends crucially on cross-elasticities among consumption goods and leisure. Here we have shown that the interest tax should be zero without imposing the usual separability assumptions on preferences. These models differ in the extent of intergenerational bequest motives and the amount of intertemporal aggregation. Examination of continuous-time overlapping generation models is needed to indicate which approach better approximates the real world of finite lives and frequent transactions.

5. Conclusions

In this paper we have studied the redistributive potential of capital income taxation in a model where investment behavior is based on the maximization of some intertemporal utility function. The major finding is that if the economy were to converge to a steady state where all agents have a common rate of time preference, no agent will asymptotically choose redistributive capital income taxation, independent of his initial and asymptotic level of wealth. This holds even when agents have a non-additive utility functional where the long-run supply curve of capital may not be perfectly elastic and is also independent of the ability of workers to participate in the capital markets. In summary, we have seen that redistribution through capital income taxation may be ineffective in the long run in a utility-maximizing model of capital accumulation.

Appendix

Lemma 2. If capitalists save out of gross income at rate $s(\bar{r})$, where $\bar{r}$ is the allowed gross return on capital, then the optimal long run tax on capital income from the workers' point of view is zero only if ρ is zero or s' is infinite.

Proof. The maximization problem is

$$\max_{\bar{r}} \int_0^\infty e^{-\rho t} u(f(k) - \bar{r}k)\,dt,$$

$$\dot{k} = s(\bar{r})\bar{r}k - \delta k.$$

The Hamiltonian is

$$H = u(F(k) - \bar{r}k) + \xi(s(\bar{r})\bar{r}k - \delta k).$$

Then

$$\dot{\xi}=\rho\xi-u'(F'-\bar{r})-\xi(s(\bar{r})\bar{r}-\delta),$$

$$0=-u'+\xi(s'(\bar{r})\bar{r}+s(\bar{r})).$$

In the steady state, $\dot{k}=0$ and $\dot{\xi}=0$, implying

$$s(\bar{r})\bar{r}=\delta,$$

$$\rho\xi=u'(F'-\bar{r}),$$

$$u'=\xi(s'(\bar{r})\bar{r}/s(\bar{r})+1)\,s(\bar{r}).$$

Therefore, $F'=\bar{r}=0$, i.e. the optimal tax is asymptotically zero, only if $\rho=0$ or $s'\bar{r}/s$ is infinite.

References

Atkinson, Anthony B. and Agnar Sandmo, 1980, Welfare implications of the taxation of savings, Economic Journal 90.

Benveniste, Laurence and José Scheinkman, 1982, Duality theory for dynamic optimization models: The continuous time case, Journal of Economic Theory.

Berndt, Ernst R. and Lau R. Christensen, 1973, The translog function and the substitution of equipment, structures and labor in U.S. manufacturing, 1929–1968, Journal of Econometrics.

Bernheim, B. Douglas, 1982, A note on dynamic tax incidence, Quarterly Journal of Economics 96, 705–723.

Boadway, Robin, 1979, Long-run tax incidence: A comparative dynamic approach, Review of Economic Studies 46, 505–511.

Bradford, David F., 1974, Constraints on government investment opportunities and the choice of discount rates, American Economic Review 65, 887–899.

Brito, Dagobert L., 1981, Optimal taxation, the distribution of wealth and the ownership of capital, Murphy Institute Discussion Paper 81–2.

Brock, William and Stephen Turnovsky, 1981, The analysis of macroeconomic policies in perfect foreshight equilibrium, International Economic Review.

Chamley, Christophe, 1980, Optimal intertemporal taxation and the public debt, Cowles Foundation Discussion Paper 554.

Feldstein, Martin S., 1974, Incidence of a capital income tax in a growing economy with variable savings rate, Review of Economic Studies, 505–513.

Grieson, Ronald E., 1975, The incidence of profit taxes in a neo-classical growth model, Journal of Public Economics 4, 75–85.

Hamada, Koichi, 1967, On the optimal transfer and income distribution in a growing economy, Review of Economic Studies, 295–299.

Hansen, Lars and Kenneth Singleton, 1982, Generalized instrumental variables estimation of nonlinear rational expectations models, Econometrica 50, 1269–1286.

Homma, Masaaki, 1981, A dynamic analysis of the differential incidence of capital and labor taxes in a two-class economy, Journal of Public Economics 15, 363–378.

Judd, Kenneth L., 1982, An alternative to steady state comparisons in perfect foresight models, Economic Letters.

Judd, Kenneth L., 1984, The welfare cost of factor taxation in a perfect foresight model, CMSEMS Discussion Paper 642.

Judd, Kenneth L., 1985, Short-run analysis of fiscal policy in a simple perfect foresight model, Journal of Political Economy 93, 298–319.

Lucas, Robert E., 1969, Labor-capital substitution in U.S. manufacturing, in: A.C. Harberger and M.J. Bailey, eds., The taxation of income from capital, Brookings Institution.

Oniki, Hajime, 1971, Comparative dynamics (sensitivity analysis) in optimal control theory, Journal of Economic Theory.

Pestieau, P. and U.M. Possen, 1978, Optimal growth and distribution policies, Journal of Public Economics 9, 357–372.

Solow, Robert M., 1956, A contribution to the theory of economic growth, Quarterly Journal of Economics 70, 65–94.

Stiglitz, Joseph E., 1978, Notes on estate taxes, redistribution and the concept of balanced growth path incidence, Journal of Political Economy 86, S137–S150.

Uzawa, Hirofumi, 1968, Time preference, the consumption function and optimum asset holdings, in: J.N. Wolfe, ed., Value, capital and growth: Papers in honour of Sir John Hicks (Edinburgh University Press, Edinburgh), 485–504.

Weber, Warren E., 1970, The effect of interest rates on aggregate consumption, American Economic Review 60, 591–600.

Weber, Warren E., 1975, Interest rates, inflation and consumer expenditures, American Economic Review 65, 843–858.