CAPITAL TAXATION
IN A SIMPLE FINITE-HORIZON OLG MODEL

Charles Blackorby

And

Craig Brett

No 709

WARWICK ECONOMIC RESEARCH PAPERS

DEPARTMENT OF ECONOMICS

THE UNIVERSITY OF WARWICK

Capital Taxation in a Simple Finite-Horizon OLG Model

by

Charles Blackorby and Craig Brett

April 2004

Charles Blackorby: Department of Economics, University of Warwick, Coventry, CV4 7AL, UK and GREQAM, Centre de la Vieille Charité, 2 rue de la Charité, 13002 Marseille, France. Email: c.blackorby@warwick.ac.uk

Craig Brett: Canada Research Chair in Canadian Public Policy, Department of Economics, Mount Allison University, Sackville, New Brunswick, E4L 1A7, Canada and CEPR. Email: cbrett@mta.ca

Acknowledgments: This research has been generously supported by the Social Science Research Council of Canada.

April 29, 2004

Abstract

In a simple finite-horizon overlapping-generations model where the government has the power to levy commodity taxes and to implement uniform lump-sum transfers, and individuals as well as the government can purchase units of a storable good in order to transfer resources from the present to the future, we derive the equations that implicitly define the taxes and subsidies that are part of the second-best Pareto optima. In this context we first show that there is production efficiency. We then show that taxes on capital income/savings are required at almost all Pareto optima. Finally we show that there are no restriction on preferences or technologies that are consistent with a general exemption of capital income/savings from the tax base.

Journal of Economic Literature Classification Numbers: D5, D6, D91, H2.

Keywords: overlapping generations, capital taxes, tax-reform.

Capital Income/Savings Taxation in a Finite-Horizon OLG Model

by

Charles Blackorby and Craig Brett

1. Introduction

The study of optimal taxation has been used to shed light on many problems in policy design, among them, capital income taxation. That is, when should capital income form part of the direct income tax base? A main theme of this literature is that distortionary taxation of capital income is unwarranted because the production efficiency result of Diamond and Mirrlees [1971] implies that there should be no taxes within the production sector.[1] However, capital income is also a form of savings for consumers and this aspect of the issue is addressed in the arguments in favor of a consumption tax as opposed to an income tax. The putative advantage of a consumption tax is that before-tax and after-tax returns on savings are equal. This absence of distortion in the capital market is seen as a prerequisite for overall efficiency and the enhancement of economic growth.[2]

Our view is that a proper examination of the capital taxation hypothesis requires several tests. First, the model must be intertemporal and have some second-best feature in order that commodity taxes are part of almost all Pareto optima. Otherwise, an exemption for capital from taxation is an immediate consequence of the suboptimality of any form of commodity taxes. Second, individuals should have life-times that are shorter than the horizon of the economy in order to avoid intertemporal neutrality results and there should be more than one consumer in each time period so that intertemporal transfers are not tantamount to individual lump-sum transfers. Lastly, the horizon should be finite in order prevent the government from engaging in some sort of Ponzi scheme.

The model we adopt satisfies these conditions and is conceptually simple. There are three periods: a start-up generation, and two generations that live for two periods each.[3] We allow the government to transfer abstract purchasing power to the old individual alive at time t, but we restrict these transfers to be of the same magnitude for all generations. This rules out the presence of lump–sum taxation, and makes our model second–best; that is,

[1] See Auerbach [1989] and Feldstein [1978, 1990]. In the overlapping generations context, Ordover and Phelps [1979] show that a zero marginal tax rate on savings income is a feature of an optimal nonlinear income tax when future consumption is separable from (current) labour supply in the direct utility function.

[2] See Auerbach [1994], Aurerbach and Kotlikoff [1987], Bradford [1987], Browning and Burbidge [1990], Kesselman [1994], and Meade [1978], for example. See also Chamley [1986], and Judd [1987,1999]. These models range from steady-state analysis of OLG models to infinitely-lived representative agent paradigms. An in-depth survey of these results is available in Bernheim [2002].

[3] There is nothing special about three; any finite horizon would do as long as the generations have lifetimes less than the horizon of the model.

some commodity taxes are optimal at almost all Pareto optima.[4] Both individuals — there is but one per generation — and the government can buy a storable good, capital, in order to transfer resources to the future.[5] There is a competitive firm that uses capital, labour, and potentially other inputs, to produce a vector of outputs and the storable commodity. The government levies a one-hundred per cent profit tax on these firms.[6] Thus, just as in the standard Ramsey optimal taxation problem, tax changes may have distributional consequences which cannot be neutralized by compensating adjustment in individual lump–sums. We shall see that capital taxes are no exception to this general rule.

In this context we show that there is production efficiency — the shadow prices in the economy are proportional to producer prices — but that there is non-zero taxation of capital income/savings at almost all Pareto optima. In addition, we show that there are no restrictions on preferences or technologies that would entail zero capital taxation along any segment of the Pareto frontier.

The intuition for these results is reasonably straightforward. Consider period one: except for labour, generation zero and generation one are consuming the same goods and providing the same services and hence facing the same prices. Because the lump-sum transfer is uniform the government cannot distinguish between the generations using either commodity taxes or the lump-sum tax. However, the storable good is purchased only by generation one (Generation zero has no reason to save.) and generation zero does not work. Hence the consumer price of the storable good and the after-tax wage rate are personalized prices to generation one. Thus, in principle, the tax on either could be used to distinguish generation one from generation zero. However, because labour enters the utility function, changing its after-tax price generates substitution effects within period one as well as intertemporal substitution effect by changing the first period budget constraint. Because the storable good is not consumed it does not enter the preferences of generation one and so a tax on it simply moves its period one budget constraint in or out. Thus the tax on capital acts as an imperfect substitute for an individual lump-sum tax on generation one in period one. It is, in general, an imperfect substitute for an individual lump-sum transfer because generation one can adjust its demand for the storable good in response to price changes thus generating intertemporal substitution effects. Nevertheless, the tax on capital permits the planner to move closer to the first-best frontier than would be possible without the tax on capital. A similar argument holds in period two. Only generation two purchases the storable good and again, its price is a personalized price to generation two and can be used in lieu of a lump-sum transfer.

The remainder of the paper is organized as follows. Section 2 outlines the model. This is followed by a tax reform analysis.[7] We start from an economy that is at an arbitrary

---

4 A similar situation would hold if we allowed many individuals per generation, unrestricted transfers across generations, but a common transfer within generations.

5 In fact, this common transfer plus the purchases of the storable could be replaced by a bond market with an equal endowment of bonds given to each generation, but at some cost in complexity.

6 Alternatively, we could have assumed a constant-returns-to-scale technology which would have eliminated this instrument.

7 See Guesnerie [1977,1995], Diewert [1978], Weymark [1979].

equilibrium and ask if there exist changes in taxes, transfers, and producer prices that are strictly Pareto-improving and equilibrium preserving. By describing the feasible, Pareto–improving directions of policy reform, we also furnish a characterization of the (local) second–best Pareto optima. Our main result is presented and interpreted in Section 4. We offer some concluding remarks in Section 5, and collect many of the mathematical details in an appendix.

2. The Model

We consider the simplest possible overlapping generations model.[8] The economy lasts for three periods: a start-up phase, a single period of the type usually examined in overlapping generations models, and a shut-down period.

2.1. Goods and Consumers

There is a single consumer in each generation, so consumer and generation are used interchangeably. Consumers have preferences over a vector $\alpha \in \mathcal{R}^n$ in the first period and $\bar{\alpha} \in \mathcal{R}^{n-1}$ in the second period of non–storable goods and services; good n is labour which is only supplied when young.[9] A single storable good, $\kappa \geq 0$, is the sole means of transferring resources forward in time and can be purchased when young to be resold when old.

An initial generation, denoted 0, is born old. It enters at date 1. It consumes goods and services, $\bar{\alpha}_1^0 \in \mathcal{R}_+^{n-1}$, and receives common lump-sum transfer, m. Also alive at time 1 is a generation born young. This generation lives for two periods. During period 1, it consumes (or supplies) $\alpha_1^1 \in \mathcal{R}^n$ and may also purchases an amount of the storable good, κ_1^1, to carry forward with it into the second period. In period 2 it spends its accumulated wealth and its lump-sum transfer, m, on the consumption of $\bar{\alpha}_2^1 \in \mathcal{R}_+^{n-1}$. A second young generation is born in period 2; it works, consumes, and saves. In period 3, the final period of our model, this generation sells its capital stock, κ_2^2, receives its lump-sum transfer,m, and consumes $\bar{\alpha}_3^3 \in \mathcal{R}_+^{n-1}$.

The production sector is composed of an aggregate profit–maximizing firm whose technology may change over time. During periods 1 and 2, this firm can produce non–durables in amounts a and the storable good in amount b, using a and k as inputs. In period 3 it does not produce any b.[10]

There is a set of consumer prices and a set of producer prices for each good and service at each date in time. Let p_t be the producer price vector for a_t.[11] π_t denotes the corresponding consumer price vector. r_t is the producer price of the storable good at time t, while ρ_t is

[8] This is a simple version of the model introduced by Allais [1947], Samuelson [1958], and analyzed by Diamond [1965]. The restriction to three periods simplifies the notation considerably with no loss in generality.

[9] If individuals were allowed to work in both periods nothing of substance would change.

[10] We use roman letters to indicate quantities produced and greek letters to indicate quantities consumed. That is, the symbols α and a refer to goods of identical characteristics. The same correspondence applies to κ and k. An inconsistency in notation arises in that the supply of κ is denoted b.

[11] We express all prices in present value form.

its consumer price. In addition the firm buys, at time t, the capital stock from generation $t-1$ at a price s_{t-1} while generation $t-1$ receives σ_{t-1}. Taxes and/or subsidies are defined implicitly by

$$\pi_t = p_t + \tau_t^a, \quad \rho_t = r_t + \tau_t^b, \quad \text{and} \quad \sigma_t = s_t + \tau_t^k. \tag{2.1}$$

We assume that the government may bestow a common lump–sum income transfer upon each generation. Specifically, at date t it transfers m to the old generation. These transfers are financed by a combination of commodity taxes and a one–hundred per cent pure profit tax.[12] The government may also purchase the storable good in period t for resale in the next period. We denote these purchases by κ_1^g in period 1 and κ_2^g in period 2. Note that the government's redistribution of income across time is mediated by the capital market.

Given the prices prevailing in period 1, and its lump-sum transfer, the budget constraint of generation 0 is[13]

$$\bar{\pi}_1^T \bar{\alpha}_1^0 \leq m. \tag{2.2}$$

Its preferences are represented by an indirect utility function given by

$$u_0 = V^0(\bar{\pi}_1, m). \tag{2.3}$$

Generation 1 faces an inter–temporal decision problem. It allocates its period 1 service income among goods and services consumed in that period and purchases of the storable good. At the beginning of the second period, it sells its capital to the firm, receiving σ_1 per unit. It combines this with its lump–sum payment from the government to finance (net) transactions in period 2. Thus, its behaviour is consistent with the joint budget constraints:

$$\pi_1^T \alpha_1^1 + \rho_1 \kappa_1^1 \leq 0 \quad \text{and} \quad \bar{\pi}_2^T \bar{\alpha}_2^1 \leq \sigma_1 \kappa_1^1 + m. \tag{2.4}$$

If generation 1 has positive savings,[14] its behaviour is also consistent with the single budget constraint

$$\frac{\sigma_1}{\rho_1} \pi_1^T \alpha_1^1 + \bar{\pi}_2^T \bar{\alpha}_2^1 \leq m. \tag{2.5}$$

When this is the case, its indirect utility function is given by

$$u_1 = V^1(\tilde{\pi}_1, \bar{\pi}_2, m) \quad \text{where} \quad \tilde{\pi}_1 := \frac{\sigma_1}{\rho_1} \pi_1. \tag{2.6}$$

Similarly, the value function of generation 2—conditional on positive savings—is given by

$$u_2 = V^2(\tilde{\pi}_2, \bar{\pi}_3, m). \tag{2.7}$$

We assume that the preferences are such that the indirect utility functions are differentially strongly quasi-convex.[15]

[12] Alternatively, one could assume constant returns-to-scale, implying zero profits.

[13] We use subscripts to denote the date at which a commodity is produced or consumed. When ambiguity is possible, we use superscripts to denote the birth date of the consuming agent. $\bar{\pi}_t$ is the consumer price vector minus the nth price, the price of labour.

[14] This is the only case that we analyze because we are interested in capital/savings taxation.

[15] See Blackorby and Diewert [1979].

In each period, the firm uses the capital it purchases from the old generation in combination with the services supplied by the young generation to produce a vector of (net) outputs. We assume that the within–period profit functions of the firm are twice continuously differentiable and strongly convex in each period.[16] The supply functions are the derivatives of these profit functions and are denoted

$$a_1 = A^1(p_1, r_1), \quad \text{and} \quad b_1 = B^1(p_1, r_1), \tag{2.8}$$

$$a_2 = A^2(s_1, p_2, r_2), \quad b_2 = B^2(s_1, p_2, r_2), \quad \text{and} \quad k_2 = K^2(s_1, p_2, r_2), \tag{2.9}$$

and

$$a_3 = A^3(s_2, p_3), \quad \text{and} \quad k_3 = K^3(s_2, p_3). \tag{2.10}$$

We have suppressed the level of the accumulated capital stock in the supply functions as a matter of notation. Because the technology is not assumed to be the same in each period, this formulation is consistent with any rate of capital depreciation or technological progress. To ease notation we let $\alpha_1^0 = (\bar{\alpha}_1^0, 0)$, $\alpha_2^1 = (\bar{\alpha}_2^1, 0)$ and $\alpha_3^2 = (\bar{\alpha}_3^2, 0)$.

A collection of prices give rise to an equilibrium if

$$\begin{aligned}
-\alpha_1^0 - \alpha_1^1 + a_1 &\geq 0, \\
-\kappa_1^1 - \kappa_1^g + b_1 &\geq 0, \\
\kappa_1^1 + \kappa_1^g - k_2 &\geq 0, \\
-\alpha_2^1 - \alpha_2^2 + a_2 &\geq 0, \\
-\kappa_2^2 - \kappa_2^g + b_2 &\geq 0, \\
\kappa_2^2 + \kappa_2^g - k_3 &\geq 0, \\
-\alpha_3^2 + a_3 &\geq 0, \\
\kappa_1^g &\geq 0, \quad \text{and} \\
\kappa_2^g &\geq 0.
\end{aligned} \tag{2.11}$$

That is, all markets — for both storable and non–storable commodities — clear. The capital market clearing conditions include the demand for capital purchases by the government. Walras' law guarantees that the government budget is balanced.[17]

[16] See Diewert, Avriel, and Zang [1981].

[17] See Chapter 2 in Guesnerie [1995] for a general discussion of this issue.

3. The Second-Best Pareto Optima

We assume that the taxation authority has control over the uniform lump–sum transfer, commodity taxes and its capital purchases. Producer prices adjust to changes in taxes, but these adjustments are captured by the equilibrium conditions. We wish to investigate if, starting at an initial tight equilibrium[18] with positive saving by consumers, there exist directions of policy reform that are strictly Pareto improving and equilibrium preserving. If no such directions exist, then the economy is at a local second-best optimum. We use this to characterize the set of all Pareto optima.

We denote a direction of policy reform by γ, where

$$\gamma^T := [\gamma_p^T, \gamma_\tau^T, \gamma_m^T, \gamma_\kappa^T] \tag{3.1}$$

and

$$\begin{aligned}\gamma_p^T &:= [dp_1^T, dr_1, ds_1, dp_2^T, dr_2, ds_2, dp_3^T];\\ \gamma_\tau^T &:= [d\tau_1^{aT}, d\tau_1^b, d\tau_1^k, d\tau_2^{aT}, d\tau_2^b, d\tau_2^k, d\tau_3^{aT}];\\ \gamma_m^T &:= [dm] \quad \text{and} \quad \gamma_\kappa^T := [d\kappa_1^g, d\kappa_2^g].\end{aligned} \tag{3.2}$$

The partition of γ corresponds to changes in producer prices, taxes, the common lump-sum transfer, and government capital purchases.

Consumer utility depends on income and consumer prices alone. The latter are the sum of producer prices and taxes. The set of strictly Pareto-improving changes in producer prices, taxes, and demogrant is given by

$$P_\pi\gamma_p + P_\pi\gamma_\tau + P_m\gamma_m + 0_{3\times 2}\gamma_\kappa \gg 0 \tag{3.3}$$

where P_π and P_m are the matrices that define the directions of strict Pareto improvements from any arbitrary equilibrium with respect to consumer prices and the demogrant.[19] A direction is equilibrium-preserving if and only if

$$[E_\pi + E_p]\gamma_p + E_\pi\gamma_\tau + E_m\gamma_m + E_\kappa\gamma_\kappa \geq 0 \tag{3.4}$$

where E_π, E_p, E_m,and E_κ are the directions of change that preserve equilibrium with respect to consumer prices, producer prices, the demogrant, and government capital purchases. In addition, the capital constraints on government must be satisfied before and after the changes so that

$$d\kappa_1^g + \kappa_1^g \geq 0 \quad \text{and} \quad d\kappa_2^g + \kappa_2^g \geq 0. \tag{3.5}$$

[18] An equilibrium is said to be tight if all relations in (2.11) (except possibly the last two) hold with equality.
[19] These matrices as well as those that follow are explicitly defined in the Appendix.

There are strict Pareto-improving changes that are simultaneously equilibrium-preserving if and only if (3.3), (3.4), and (3.5) have a solution. Together these constitute a non–homogeneous system of linear equalities and inequalities. It is shown in the appendix that this system is equivalent to the following homogeneous system

$$\begin{bmatrix} P_\pi & P_\pi & P_m & 0_{3\times 2} & 0_3 \\ 0_{3n+4}^T & 0_{3n+4}^T & 0 & 0_2^T & 1 \end{bmatrix} \begin{bmatrix} \gamma \\ \gamma_\eta \end{bmatrix} \gg 0 \tag{3.6}$$

and

$$\begin{bmatrix} E_\pi + E_p & E_\pi & E_m & E_\kappa & 0_{3n+4} \\ 0_{2\times(3n+4)} & 0_{2\times(3n+4)} & 0_2 & I_{2\times 2} & \kappa^g \end{bmatrix} \begin{bmatrix} \gamma \\ \gamma_\eta \end{bmatrix} \geq 0 \tag{3.7}$$

where γ_η is the dummy variable that permits the conversion of the non–homogeneous system into a homogeneous one.

If there is no solution to (3.6) and (3.7) then there are no Pareto-improving feasible changes and the economy is at a second-best optimum. Using Motzkin's Theorem[20] the economy is at a second-best optimum if and only if there exists some collection of multipliers such that $0 \neq [\xi^T, \theta] \geq 0^T$ and $[v^T, \eta^T] \geq 0^T$ satisfying

$$\xi^T P_\pi + v^T [E_\pi + E_p] = 0, \tag{3.8}$$

$$\xi^T P_\pi + v^T E_\pi = 0, \tag{3.9}$$

$$\xi^T P_m + v^T E_m = 0, \tag{3.10}$$

$$v^T E_\kappa + \eta^T = 0, \tag{3.11}$$

and

$$\theta + \eta^T \kappa^g = 0. \tag{3.12}$$

The relations (3.8)–(3.12) contain the first–order conditions for a Pareto optimum, and implicitly define the efficient taxes and demogrant. The multipliers correspond to shadow values. Specifically, v is the vector of the social marginal values of commodities (the shadow prices), and ξ is the vector of social marginal values of the incomes of the various generations (consumers).[21] That is, if the planner were maximizing a social welfare function $W(u_0, u_1, u_2)$ subject to the equilibrium conditions, then,

$$\xi_t = \frac{\partial W}{\partial u_t} \frac{\partial V^t}{\partial m}. \tag{3.13}$$

We first show that production efficiency obtains in the model; that is, that the shadow prices in this economy are proportional to producer prices. Subtracting (3.9) from (3.8) yields

$$v^T E_p = 0. \tag{3.14}$$

20 See Mangasarian [1969, pp. 28-29] for a statement and proof of this result.

21 See Guesnerie [1995] or Myles [1995].

From (3.14) and the strong convexity of the firm's profit functions, it follows that[22]

Theorem 1: *Given our regularity conditions and positive savings by generations one and two, there is production efficiency at all Pareto optima. Per period shadow prices are proportional to producer prices,*

$$v^T = \left[\mu_1(p_1^T, r_1), \mu_2(s_1, p_2^T, r_2), \mu_3(s_2, p_3^T)\right] \tag{3.15}$$

where μ_t is a function of period t producer prices only.

Although shadow prices in the economy are proportional to producer prices in each period, no such claim can be made directly for intertemporal shadow prices which depend upon the values of μ_t. This, however, is not necessary for production efficiency; the production sector does not engage directly in intertemporal production but only indirectly by producing the storable good that is resold in the next period to augment the capital stock.

The relation (3.12) has no analogue in the static Ramsey problem. Expanding it yields

$$\theta + \eta_1\kappa_1^g + \eta_2\kappa_2^g = 0. \tag{3.16}$$

Because each term in (3.16) is nonnegative, they are all zero yielding

$$\eta_1\kappa_1^g = 0 \quad \text{and} \quad \eta_2\kappa_2^g = 0, \tag{3.17}$$

the standard complementary slackness conditions associated with the constraints on government capital purchases. These conditions partition the set of Pareto optima into four regions, depending upon the timing of government saving. The region of $\kappa_1^g = 0$ and $\eta_1 > 0$ corresponds to a situation where the government would like to transfer more resources into period one, but is prevented from doing so by the nonnegativity constraint on government capital purchases. The region in which $\eta_1 = \eta_2 = 0$ corresponds to a case when the planner is not capital constrained and is saving in both periods. In this case, (3.11) implies that

$$\mu_{t+1}s_t = \mu_t r_t \quad \text{for} \quad t = 1, 2. \tag{3.18}$$

That is, the ratio of intertemporal shadow prices is equal to the ratio of the price that firms must pay for the stored good at the beginning of period t to what it can sell the storable good for at the end of the period.

(3.9) and (3.10) contain the information commonly contained in Ramsey formulae. More specifically, expanding and manipulating them under the assumption that (3.18) holds yields (See the Appendix.)

$$\xi_1 + \xi_2 + \xi_3 = \mu_1 p_1^T \nabla_m \alpha_1^0 + \mu_1 p_1^T \nabla_m \alpha_1 + \mu_2 p_2^T \nabla_m \alpha_2 + \mu_3 p_3^T \nabla_m \alpha_3^2, \tag{3.19}$$

[22] The proof is in the Appendix.

$$\left[\, \xi_1 \alpha_1^{0T} + \xi_2 \frac{\sigma_1}{\rho_1} \alpha_1^{1T} \quad \xi_2 \alpha_2^{1T} + \xi_3 \frac{\sigma_2}{\rho_2} \alpha_2^{2T} \quad \xi_3 \alpha_3^{2T} \,\right]$$

$$= - \left[\, \mu_1 p_1^T \quad \mu_2 p_2^T \quad \mu_3 p_3^T \,\right] \begin{bmatrix} \nabla_{\pi_1} \alpha_1 & \nabla_{\pi_2} \alpha_1^1 & \mathbf{0} \\ \nabla_{\pi_1} \alpha_2^1 & \nabla_{\pi_2} \alpha_2 & \nabla_{\pi_3} \alpha_2^2 \\ \mathbf{0} & \nabla_{\pi_2} \alpha_3^2 & \nabla_{\pi_3} \alpha_3^2 \end{bmatrix}, \tag{3.20}$$

$$-\xi_2 \rho_1 \kappa_1^1 = -\mu_1 p_1^T \nabla_{\tilde{\pi}_1} \alpha_1^1 \pi_1 - \mu_2 p_2^T \nabla_{\tilde{\pi}_1} \alpha_2^1 \pi_1 \tag{3.21}$$

and

$$-\xi_3 \rho_2 \kappa_2^2 = -\mu_2 p_2^T \nabla_{\tilde{\pi}_2} \alpha_2^2 \pi_2 - \mu_3 p_3^T \nabla_{\tilde{\pi}_2} \alpha_3^2 \pi_2. \tag{3.22}$$

Expression (3.20) is close to the standard Mirrlees' result if one ignores (3.21) and (3.22).[23] That is, the social value of the demands and supplies is equal to the cost of the change in commodity and service taxes valued at producer prices. Using the Slutsky equation for each generation allows one to rewrite this to obtain the many person Ramsey rule.

From (3.21) and (3.22) we obtain a result relating the value of savings of generation one to changes in the consumer cost of changes in consumption, again valued at producer prices. This means that the social value of the savings of generation one must be equal to the value, at consumer prices, of the induced change in the demand for goods and services times the shadow prices of these commodities and services. To interpret this result it is useful to consider the effect of giving generation one a voucher that it can spend on capital purchases only. The social value of this would be minus the left–hand side of (3.21). Generation one would spend this windfall on goods and services, perhaps some in each period of its life. The first two terms of (3.21) account for the costs, at the social shadow prices prevailing in the period of consumption, of these changes. Note that (3.21) contains only information about generation 1. Because generation zero does not purchase any of the storable good, this equation allows the planner to distinguish between the two generations even though it does not have access to lump sum transfers. From (3.22) we obtain a similar result for the value of the savings of generation two.

In the other three regions — those in which (3.18) does not hold — we obtain results that are quite different. The index of discouragement in consumption must be adjusted for the induced changes in savings which are valued at the differences in the social value of capital in two adjacent time periods, $\mu_{t+1} s_1 - \mu_t r_t$. This is the result of the fact that in these regions the government is moving into regions of the Pareto-set where it wishes that it could move capital from the future to the present but is unable to do so. This in turn entails that the shadow value of capital is different in the two periods.

23 See Mirrlees [1986] and Guesnerie [1995].

4. Capital Income/Savings Taxation

We are now in a position to pose the principal question of the paper: what conditions on preferences and technologies are necessary and sufficient to imply that the taxation of capital income is unnecessary at efficient equilibria? It is well-known that in certain classes of models with sufficient structure capital taxation is redundant.[24] The model set out above is ideal to test this intuition. Each generation has a lifetime that is shorter than that of the economy. Commodity and service taxes are an essential part of almost all Pareto-efficient outcomes because the government does not have access to individual lump-sum transfers, but only a uniform lump-sum transfer—a demogrant. Given that one of the components of α_t^t must be negative, this is equivalent to an anonymous affine income-tax schedule with linear taxes for all other commodities.

Before we can pose the question that motivates this paper we need first to establish the number of prices and or taxes that can be normalized without changing the set of Pareto optima. The following result, which is proved in the appendix, establishes this.

Theorem 2: *Given our regularity conditions and positive savings by generations one and two, at every Pareto optimum one can normalize one producer price in each period, either the tax on capital or the tax on savings in each period, but not both, and one consumer price (not per period).*

To effectuate Theorem 2 we employ the following normalizations in the rest of the paper. Let

$$p_{11} = 1, \quad s_1 = 1, \quad s_2 = 1, \quad \tau_{11}^a = 0, \quad \tau_1^k = 0, \quad \text{and} \quad \tau_2^k = 0. \tag{4.1}$$

That is, we have chosen arbitrarily to set the producer price of good one in period one, of capital in period two and three equal to one, the tax on good one in period one and the tax on savings in periods two and three equal to zero.

We ask if there are any restrictions on preferences or technologies that entail zero capital taxes along the Pareto frontier. To pose the question, we ask under what circumstances can we—without loss of generality—set the tax on the purchases of the storable good equal to zero in addition to the normalizations given above. That is, under what conditions are efficient equilibria consistent with (4.1) and

$$\tau_1^b = 0 \quad \textit{and} \quad \tau_2^b = 0. \tag{4.2}$$

The following result is established in the Appendix.

Theorem 3: *Given the regularity conditions, there are no restrictions on preferences or technologies such that zero taxes on capital and savings hold along a segment of the second best Pareto frontier in a finite-horizon OLG model with positive individual savings.*

---

24 These are either models with a single infinitely-lived consumer or a study of steady-state behaviour. See Auerbach and Hines [2002] for a discussion and for references.

Theorem 3 is a strong result but is relatively easily explained. In periods one and two there are two consumers for all good and services except labour; they face the same consumer prices except for the after-tax wage rate, and have the same lump-sum income. Because the planner does not have access to generation-specific lump-sum transfers, the optimal commodity taxes can only depend upon the aggregate elasticities of demand. However, only generation one purchases the storable good in period one and only generation two buys the storable good in period two. Thus, for generation one, the price of the storable good in period one is a personalized price. Because the storable good does not appear in the preferences of generation one, the only effect of a change in the personalized price to generation one is to move the period one budget constraint of generation one in or out.[25] Similarly, the price of the storable good in period two is a personalized price for generation two. The planner can, in part, substitute the tax on the storable good for the unavailable lump-sum tax. Of course, this substitution is less than perfect because the individuals can adjust their purchases of the storable good in response to price changes. If generations one and two simply possessed a stock of the storable good and could not buy or sell it, then, we would be able to retrieve a first-best optimum. Because we are only interested in the regions where generations one and two are saving, and because this is the only intertemporal link for these generations, the planner uses the tax on savings to move in the direction of lump-sum taxation.

In order to see this argument more formally consider, for simplicity, the region where the government purchases of the storable good are positive so that (3.18) holds. Considering this case also makes it clear that Theorem 3 depends in no way on the relationship between inter-temporal shadow prices and rates of return on the storable commodity. Expanding the first equation of (3.20) and and using the first period budget constraint of generation one, (3.20) and (3.21) can be rewritten as

$$\xi_1 \alpha_1^{0T} + \xi_2 \frac{\sigma_1}{\rho_1} \alpha_1^{1T} = -\mu_1 p_1^T \nabla_{\pi_1} \alpha_1^0 - \mu_1 p_1^T \nabla_{\tilde{\pi}_1} \alpha_1^1 \frac{\sigma_1}{\rho_1} - \mu_2 p_2^T \nabla_{\tilde{\pi}_1} \alpha_2^1 \frac{\sigma_1}{\rho_1} \tag{4.3}$$

and

$$\xi_2 \pi_1^T \alpha_1^1 = -\mu_1 p_1^T \nabla_{\tilde{\pi}_1} \alpha_1^1 \pi_1 - \mu_2 p_2^T \nabla_{\tilde{\pi}_1} \alpha_2^1 \pi_1. \tag{4.4}$$

Postmultiplying (4.3) by π_1 and collecting terms yields

$$\xi_1 \alpha_1^{0T} \pi_1 + \mu_1 p_1^T \nabla_{\pi_1} \alpha_1^0 \pi_1 = -\frac{\sigma_1}{\rho_1} \left[\xi_2 \alpha_1^{1T} \pi_1 + \mu_1 p_1^T \nabla_{\tilde{\pi}_1} \alpha_1^1 \pi_1 + \mu_2 p_2^T \nabla_{\tilde{\pi}_1} \alpha_2^1 \pi_1 \right] \tag{4.5}$$

Using (4.4), the term in square brackets on the right side of (4.5) is zero yielding

$$\xi \alpha_1^{0T} \pi_1 + \mu_1 p_1^T \nabla_{\pi_1} \alpha_1^0 \pi_1 = 0 \tag{4.6}$$

which, using the budget constraint of generation zero and the homogeneity of degree zero in (π_1, m) of generations zero's demand functions becomes

$$\xi \alpha_1^{0T} = \mu_1 p_1^T \nabla_m \alpha_1^0. \tag{4.7}$$

[25] See the left side of (2.4).

However, (4.7) is exactly the first-order condition that would obtain if there were a generation-specific lump-sum transfer to generation zero that would permit the planner to distinguish it from generations one and two. That is, taxing the storable good acts indirectly just like a generation-specific lump-sum transfer and thus allows the second-best optima to be moved closer to the first-best than would be possible if there were no capital taxation. A similar, but less complete, argument explains the taxation of the capital good in period two. It allows the planner to distinguish between generations one and two because only generation two buys the storable good. It is less complete because consumption of generation one in period one cannot be controlled by manipulating the period-two budget constraint of generation two but only the consumption of period two commodities by generation one. This explains Theorem 3, the tax on saving should be nonzero almost everywhere.

Our argument for the existence of a tax on saving is similar to the standard argument for differential commodity taxes, in that the savings tax allows the planner a measure of targeted redistribution. It is possible to find restrictions on preferences, like those outlined in Atkinson and Stiglitz [1976], that render commodity demands similar across types of consumers, so that differential commodity taxes have no redistributive effect. However, at each date, savings taxes are necessarily targeted at the (in our model, unique) generation doing the saving. Under the assumption of non-zero saving a change in the price of saving has an income effect on exactly one generation, regardless of its preferences.

One might imagine that these results derive from some sort of dynamic inconsistency. This question is more difficult in the current context because we are trying to describe properties that hold at all Pareto optima and not just the result of a particular social welfare maximization. Of course, each point on the Pareto frontier would be chosen by some social welfare function so that in principle one could analyze this problem point by point. This would involve not only examining the changes in the planner's behaviour but also the individual generation's optimal responses to it. There is no obvious way to do this while looking at the entire set of Pareto optima, however, a partial response is possible.

In the original set of Pareto optima all three generations received the same lump-sum transfer. At the beginning of period two, given the work/consumption and saving decision taken by generations zero and one, we can compute a (potentially) new set of Pareto optima constraining generations one and two to receive the same lump-sum transfer but possibly different from that received by generation zero. This preserves the second-best nature of the problem but permits us to ask if the tax on the storable good is still a necessary instrument at the potentially new Pareto optima. The solutions that we found to the original problem are still consistent with the new problem but there may of course be solutions to the new problem that were not solutions to the old. We have shown in the original problem that the capital tax was non-zero almost everywhere and that it was an integral part of the tax system. At the beginning of period two, a tax has already been levied on the storable good in period one. Beginning in period two and constraining the lump-sum transfers to generations one and two to be equal but not necessarily equal to that of generation zero, we can show that the capital tax is still non-zero and that it is an essential element of the indirect tax system for the same reasons provided above, namely that the price of the storable good in period to is a personalized price to generation two.

Theorem 4: *Constraining the lump-sum payments to generations one and two to be equal at the beginning of period two, the result of Theorem 3 is preserved at the resulting Pareto-optima.*

The proof of Theorem 4 is exactly like the proof of Theorem 3.

5. Conclusion

We have shown that a tax on capital is required at almost all Pareto optima and that there are no restrictions on preferences or technology that render these taxes unnecessary. These theorems depend explicitly on the OLG structure of the model. Theorems 3 and 4 do not claim that there are no Pareto efficient tax structures with zero capital taxation. The Walrasian equal-income equilibrium is Pareto efficient. However, it is a *single point* on the Pareto frontier. Even points on the Pareto frontier near the laissez faire outcome require some form of capital market intervention.

Capital taxation can have far–reaching effects on the economy. Like all forms of taxation, it affects the relative prices consumers face, and the real incomes at their disposal. By their very nature, the incidence of capital income taxes differs among agents. In the absence of optimal inter–generational transfer schemes, this differential incidence can be exploited to implement some parts of the Pareto frontier that would otherwise be unattainable. This paper has uncovered, perhaps surprisingly, that this intuition is more robust than it might first appear. No set of restrictions on preferences is sufficient to render the possible redistributive effects of capital taxation inoperable.

6. Appendix

This sections contains several technical arguments used in the text.

6.1. Pareto-Improving and Equilibrium-Preserving Directions

The matrices that define Pareto-improving directions with respect to consumer prices and the demogrant are given by

$$P_\pi := \begin{bmatrix} -\alpha_1^{0T} & 0 & 0 & 0_n^T & 0 & 0 & 0_n^T \\ -\frac{\sigma_1}{\rho_1}\alpha_1^{1T} & \frac{\sigma_1}{\rho_1^2}\pi_1^T\alpha_1^1 & -\frac{1}{\rho_1}\pi_1^T\alpha_1^1 & -\alpha_2^{1T} & 0 & 0 & 0_n^T \\ 0_n^T & 0 & 0 & -\frac{\sigma_2}{\rho_2}\alpha_2^{2T} & \frac{\sigma_2}{\rho_2^2}\pi_2^T\alpha_2^2 & -\frac{1}{\rho_2}\pi_2^T\alpha_2^2 & -\alpha_3^{2T} \end{bmatrix} \tag{6.1}$$

and

$$P_m := \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}. \tag{6.2}$$

In order to describe the feasible directions of policy reform, we define a collection of matrices.[26] First,

$$E_\pi :=$$

$$\begin{bmatrix} -\nabla_{\pi_1}\alpha_1 & -\nabla_{\rho_1}\alpha_1^1 & -\nabla_{\sigma_1}\alpha_1^1 & -\nabla_{\pi_2}\alpha_1^1 & 0_n & 0_n & 0_{n\times n} \\ -\nabla_{\pi_1}\kappa_1^1 & -\nabla_{\rho_1}\kappa_1^1 & -\nabla_{\sigma_1}\kappa_1^1 & -\nabla_{\pi_2}\kappa_1^1 & 0 & 0 & 0_n^T \\ +\nabla_{\pi_1}\kappa_1^1 & +\nabla_{\rho_1}\kappa_1^1 & +\nabla_{\sigma_1}\kappa_1^1 & +\nabla_{\pi_2}\kappa_1^1 & 0 & 0 & 0_n^T \\ -\nabla_{\pi_1}\alpha_2^1 & -\nabla_{\rho_1}\alpha_2^1 & -\nabla_{\sigma_1}\alpha_2^1 & -\nabla_{\pi_2}\alpha_2 & -\nabla_{\rho_2}\alpha_2^2 & -\nabla_{\sigma_2}\alpha_2^2 & -\nabla_{\pi_3}\alpha_2^2 \\ 0_n^T & 0 & 0 & -\nabla_{\pi_2}\kappa_2^2 & -\nabla_{\rho_2}\kappa_2^2 & -\nabla_{\sigma_2}\kappa_2^2 & -\nabla_{\pi_3}\kappa_2^2 \\ 0_n^T & 0 & 0 & +\nabla_{\pi_2}\kappa_2^2 & +\nabla_{\rho_2}\kappa_2^2 & +\nabla_{\sigma_2}\kappa_2^2 & +\nabla_{\pi_3}\kappa_2^2 \\ 0_{n\times n} & 0_n & 0_n & -\nabla_{\pi_2}\alpha_3^2 & -\nabla_{\rho_2}\alpha_3^2 & -\nabla_{\sigma_2}\alpha_3^2 & -\nabla_{\pi_3}\alpha_3^2 \end{bmatrix}. \tag{6.3}$$

Each row of (6.3) corresponds to a relation in (2.11), with elements corresponding to the change in the left–hand side in (2.11) to an infinitesimal change in a consumer price. To deal with the demogrant, producer price changes and government capital purchases, we introduce:

$$E_m := \begin{bmatrix} -\nabla_m\alpha_1^0 - \nabla_m\alpha_1^1 \\ -\nabla_m\kappa_1^1 \\ \nabla_m\kappa_1^1 \\ -\nabla_m\alpha_2^1 - \nabla_{m_3}\alpha_2^2 \\ -\nabla_m\kappa_2^2 \\ \nabla_m\kappa_2^2 \\ -\nabla_m\alpha_3^2 \end{bmatrix}; \tag{6.4}$$

$$E_p :=$$

$$\begin{bmatrix} \nabla_{p_1}a_1 & \nabla_{r_1}a_1 & 0_n & 0_{n\times n} & 0_n & 0_n & 0_{n\times n} \\ \nabla_{p_1}b_1 & \nabla_{r_1}b_1 & 0 & 0_n^T & 0 & 0 & 0_n^T \\ 0_n^T & 0 & -\nabla_{s_1}k_2 & -\nabla_{p_2}k_2 & -\nabla_{r_2}k_2 & 0 & 0_n^T \\ 0_{n\times n} & 0_n & \nabla_{s_1}a_2 & \nabla_{p_2}a_2 & \nabla_{r_2}a_2 & 0 & 0_n^T \\ 0_n^T & 0 & \nabla_{s_1}b_2 & \nabla_{p_2}b_2 & \nabla_{r_2}b_2 & 0 & 0_n^T \\ 0_n^T & 0 & 0 & 0_n^T & 0 & -\nabla_{s_2}k_3 & -\nabla_{p_3}k_3 \\ 0_{n\times n} & 0_n & 0_n & 0_{n\times n} & 0_n & \nabla_{s_2}a_3 & \nabla_{p_3}a_3 \end{bmatrix}; \tag{6.5}$$

and

$$E_\kappa := \begin{bmatrix} 0 & 0 \\ -1 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & -1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}. \tag{6.6}$$

[26] Variables without superscripts denote within–period aggregate demands.

6.2. Non Homogeneous Motzkin Theorem

We want to show that the non–homogeneous system of relations,

$$Ax \gg 0, \quad Bx \geq 0, \quad Dx \geq \xi, \quad \text{and} \quad Cx = 0, \tag{6.7}$$

has a solution exactly when there exists a solution to the following homogeneous system:

$$\begin{bmatrix} A & 0_a \\ 0_n^T & \theta \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} \gg 0, \quad \begin{bmatrix} B & 0_b \\ D & -\xi \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} \geq 0, \quad \text{and} \quad \begin{bmatrix} C & 0_c \\ 0_n^T & 0 \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} = 0 \tag{6.8}$$

where $x \in \mathcal{R}^n$ and 0_a has the same number of rows as A.

If (6.7) has a solution, then (6.8) has a solution with $z = 1$. If (6.8) has a solution, then, dividing through by z shows that (6.7) has a solution.

6.3. Proof of Theorem 1

From (3.14) we have

$$v^T E_p = 0. \tag{6.9}$$

From (6.5) this is

$$v^T \begin{bmatrix} \nabla_{p_1} a_1 & \nabla_{r_1} a_1 & 0_n & 0_{n\times n} & 0_n & 0_n & 0_{n\times n} \\ \nabla_{p_1} b_1 & \nabla_{r_1} b_1 & 0 & 0_n^T & 0 & 0 & 0_n^T \\ 0_n^T & 0 & -\nabla_{s_1} k_2 & -\nabla_{p_2} k_2 & -\nabla_{r_2} k_2 & 0 & 0_n^T \\ 0_{n\times n} & 0_n & \nabla_{s_1} a_2 & \nabla_{p_2} a_2 & \nabla_{r_2} a_2 & 0 & 0_n^T \\ 0_n^T & 0 & \nabla_{s_1} b_2 & \nabla_{p_2} b_2 & \nabla_{r_2} b_2 & 0 & 0_n^T \\ 0_n^T & 0 & 0 & 0_n^T & 0 & -\nabla_{s_2} k_3 & -\nabla_{p_3} k_3 \\ 0_{n\times n} & 0_n & 0_n & 0_{n\times n} & 0_n & \nabla_{s_2} a_3 & \nabla_{p_3} a_3 \end{bmatrix} = 0. \tag{6.10}$$

This matrix is block diagonal. The first block is the $N+1$ by $N+1$ Hessian of the first period profit function which is assumed to be differentially strongly convex and hence has rank N. This means that the zero eigen-vector is unique up to positive scalar multiplication and is proportional to the first period producer price vector, (p_1^T, r_1). It follows that the factor of proportionality can only be a function of period one producer prices. The second block is the Hessian of the second period profit function and the third block that of the third period profit function. Similar arguments establish the theorem. ■

6.4. Proof of Theorem 2:

This model admits six independent price normalizations: one producer price per period can be fixed; one consumer price — but not one per period — may be fixed; and either the capital input taxes or the taxes on savings (but not both) may be set to zero. We show that here only six are possible:

$$p_{11} = 1, \quad s_1 = 1, \quad s_2 = 1, \quad \tau_{11}^a = 0 \quad \tau_1^k = 0, \quad \text{and} \quad \tau_2^k = 0. \tag{6.11}$$

Given our tax reform perspective, it is necessary to translate the normalizations and (possibly) binding restrictions into statements about the possible directions of policy reform.

Clearly, the components of γ corresponding to a change in a normalized quantity must be zero. We can introduce these restrictions with the help of the following matrices:

$$\mathcal{I} = \begin{bmatrix} 1, 0_{n-1}^T & 0 & 0 & 0_n^T & 0 & 0 & 0_n^T \\ 0_n^T & 0 & 1 & 0_n^T & 0 & 0 & 0_n^T \\ 0_n^T & 0 & 0 & 0_n^T & 0 & 1 & 0_n^T \end{bmatrix} \tag{6.12}$$

and

$$\tilde{\mathcal{I}} = \begin{bmatrix} 1, 0_{n-1}^T & 0 & 0 & 0_n^T & 0 & 0 & 0_n^T \\ 0_n^T & 0 & 1 & 0_n^T & 0 & 0 & 0_n^T \\ 0_n^T & 0 & 0 & 0_n^T & 0 & 1 & 0_n^T \end{bmatrix}. \tag{6.13}$$

The rows of the matrices correspond to the order in which the constraints are imposed by (6.11); for example, the first row of $\tilde{\mathcal{I}}$ imposes the constraint $d\tau_{11}^a = 0$.

These normalizations are imposed by

$$\begin{bmatrix} \mathcal{I} \\ 0_{3\times(3n+4)} \end{bmatrix} \gamma_p + \begin{bmatrix} 0_{3\times(3n+4)} \\ \tilde{\mathcal{I}} \end{bmatrix} \gamma_\tau + \begin{bmatrix} 0_3 \\ 0_3 \end{bmatrix} \gamma_m + \begin{bmatrix} 0_{3\times 2} \\ 0_{3\times 2} \end{bmatrix} \gamma_\kappa + \begin{bmatrix} 0_3 \\ 0_3 \end{bmatrix} \gamma_\eta = 0. \tag{6.14}$$

There are strict Pareto-improving changes that are simultaneously equilibrium-preserving with the six normalizations if and only (3.6), (3.7), and (6.14) have a solution. If there is no such solution we are at a second-best optimum. Using Motzkin's Theorem the economy is at a second-best optimum if and only if

$$\begin{bmatrix} \xi^T & \theta \end{bmatrix} \begin{bmatrix} [P_\pi & P_\pi & P_m & \mathbf{0} & 0 \\ 0_n^T & 0_n^T & 0 & 0_2^T & 1 \end{bmatrix} + \begin{bmatrix} v^T & \eta^T \end{bmatrix} \begin{bmatrix} E_\pi + E_p & E_\pi & E_m & E_\kappa & 0 \\ 0_{2\times n} & 0_{2\times n} & 0_2 & I_{2\times 2} & \kappa^g \end{bmatrix}$$
$$+ (w^T, z^T) \begin{bmatrix} \mathcal{I} & 0_{3\times(3n+4)} & 0_3 & 0_{3\times 2} & 0_3 \\ 0_{3\times(3n+4)} & \tilde{\mathcal{I}} & 0_3 & 0_{3\times 2} & 0_3 \end{bmatrix} = 0 \tag{6.15}$$

where $0 \neq [\xi^T, \theta] \geq 0^T$ and $[v^T, \eta^T] \geq 0^T$.

Expanding (6.15) yields

$$\xi^T P_\pi + v^T (E_\pi + E_p) + w^T \mathcal{I} = 0, \tag{6.16}$$

$$\xi^T P_\pi + v^T E_\pi + z^T \tilde{\mathcal{I}} = 0, \tag{6.17}$$

$$\xi^T P_m + v^T E_m = 0, \tag{6.18}$$

$$v^T E_\kappa + \eta^T = 0, \tag{6.19}$$

$$\theta + \eta^T \kappa^g = 0. \tag{6.20}$$

Let $v^T = (v_1^T, v_2, v_3, v_4^T, v_5, v_6, v_7^T)$ where the first, fourth, and last element are n-tuples. Expanding yields (6.18) and then (6.17) yields

$$\sum_t \xi_t - v_1^T \nabla_m \alpha_1 + (v_3 - v_2) \nabla_m \kappa_1^1 - v_4^T \nabla_m \alpha_2 + (v_6 - v_5) \nabla_m \kappa_2^2 - v_7^T \nabla_m \alpha_3^2 = 0, \tag{6.21}$$

$$-\xi_1\alpha_1^{0T} - \xi_2\frac{\sigma_1}{\rho_1}\alpha_1^{1T} - v^T\nabla_{\pi_1}\alpha_1 + (v_3 - v_2)\nabla_{\pi_1}\kappa_1^1 - v_4^T\nabla_{\pi_1}\alpha_2^1 + z_1 = 0, \tag{6.22}$$

$$\xi_2\frac{\sigma_1}{\rho_1^2}\pi_1^T\alpha_1^1 - v_1^T\nabla_{\rho_1}\alpha_1^1 + (v_3 - v_2)\nabla_{\rho_1}\kappa_1^1 - v_4^T\nabla_{\rho_1}\alpha_2^1 = 0, \tag{6.23}$$

$$-\xi_2\frac{1}{\rho_1}\pi_1^T\alpha_1^1 - v_1^T\nabla_{\sigma_1}\alpha_1^1 + (v_3 - v_2)\nabla_{\sigma_1}\kappa_1^1 - v_4^T\nabla_{\sigma_1}\alpha_2^1 + z_2 = 0, \tag{6.24}$$

$$-\xi_2\alpha_2^{1T} - \xi_3\frac{\sigma_2}{\rho_2}\alpha_2^{2T} - v_1^T\nabla_{\pi_2}\alpha_1^1 + (v_3 - v_2)\nabla_{\pi_2}\kappa_1^1 - v_4^T\nabla_{\pi_2}\alpha_2 + (v_6 - v_5)\nabla_{\pi_2}\kappa_2^2 - v_7^T\nabla_{\pi_2}\alpha_3^2 = 0, \tag{6.25}$$

$$\xi_3\frac{\sigma_2}{\rho_2^2}\pi_2^T\alpha_2^2 - v_4^T\nabla_{\rho_2}\alpha_2^2 + (v_6 - v_5)\nabla_{\rho_2}\kappa_2^2 - v_7^T\nabla_{\rho_2}\alpha_3^2 = 0, \tag{6.26}$$

$$-\xi_3\frac{1}{\rho_2}\pi_2^T\alpha_2^2 - v_4^T\nabla_{\sigma_2}\alpha_2^2 + (v_6 - v_5)\nabla_{\sigma_2}\kappa_2^2 - v_7^T\nabla_{\sigma_2}\alpha_3^2 + z_3 = 0, \tag{6.27}$$

and

$$-\xi_3\alpha_3^{2T} - v_4^T\nabla_{\pi_3}\alpha_2^2 + (v_6 - v_5)\nabla_{\pi_3}\kappa_2^2 - v_7^T\nabla_{\pi_3}\alpha_3^2 = 0. \tag{6.28}$$

Multiplying (6.23) by ρ_1, (6.24) by σ_1, adding and using the homogeneity of κ_1^1 yields

$$\sigma_1 z_2 = 0 \tag{6.29}$$

so that z_2 is identically zero and the constraint that $t_1^k = 0$ is not binding; hence a free normalization. Similarly multiply (6.26) by ρ_2, (6.25) by σ_2, adding and using the homogeneity of κ_2^2 yields

$$\sigma_2 z_3 = 0 \tag{6.30}$$

so that z_3 is identically zero and the constraint that $t_2^k = 0$ is not binding. Next, multiply (6.21) by m, (6.22) by π_1, (6.25) by π_2, (6.28) by π_3, add, use the homogeneity of the demand equations and the budget constraints to obtain

$$z_1\pi_{11} = 0. \tag{6.31}$$

Thus the tax on one consumer price can be set equal to zero. Next, using the proof of Theorem 1, shows that one producer price can be normalized in each period. ■

6.5. Proof of Theorem 3:

In order to impose the six normalizations plus (4.2) let

$$\mathcal{I} = \begin{bmatrix} 1, 0_{n-1}^T & 0 & 0 & 0_n^T & 0 & 0 & 0_n^T \\ 0_n^T & 0 & 1 & 0_n^T & 0 & 0 & 0_n^T \\ 0_n^T & 0 & 0 & 0_n^T & 0 & 1 & 0_n^T \end{bmatrix}. \tag{6.32}$$

$$\hat{\mathcal{I}} = \begin{bmatrix} 1, 0_{n-1}^T & 0 & 0 & 0_n^T & 0 & 0 & 0_n^T \\ 0_n^T & 1 & 0 & 0_n^T & 0 & 0 & 0_n^T \\ 0_n^T & 0 & 1 & 0_n^T & 0 & 0 & 0_n^T \\ 0_n^T & 0 & 0 & 0_n^T & 1 & 0 & 0_n^T \\ 0_n^T & 0 & 0 & 0_n^T & 0 & 1 & 0_n^T \end{bmatrix}. \tag{6.33}$$

Then, the six normalizations plus (4.2) are imposed by

$$\begin{bmatrix} \mathcal{I} \\ 0_{5\times(3n+4)} \end{bmatrix} \gamma_p + \begin{bmatrix} 0_{3\times(3n+4)} \\ \hat{\mathcal{I}} \end{bmatrix} \gamma_\tau + \begin{bmatrix} 0_3 \\ 0_5 \end{bmatrix} \gamma_m + \begin{bmatrix} 0_{3\times 2} \\ 0_{5\times 2} \end{bmatrix} \gamma_\kappa + \begin{bmatrix} 0_3 \\ 0_5 \end{bmatrix} \gamma_\eta = 0. \tag{6.34}$$

Using Motzkin's Theorem the economy is at a second-best optimum if and only if

$$\begin{bmatrix} \xi^T & \theta \end{bmatrix} \begin{bmatrix} [P_\pi & P_\pi & P_m & \mathbf{0} & 0 \\ 0_n^T & 0_n^T & 0 & 0_2^T & 1 \end{bmatrix} + \begin{bmatrix} v^T & \eta^T \end{bmatrix} \begin{bmatrix} E_\pi + E_p & E_\pi & E_m & E_\kappa & 0 \\ 0_{2\times n} & 0_{2\times n} & 0_2 & I_{2\times 2} & \kappa^g \end{bmatrix}$$

$$+ (w^T, z^T) \begin{bmatrix} \mathcal{I} & 0_{3\times(3n+4)} & 0_3 & 0_{3\times 2} & 0_3 \\ 0_{5\times(3n+4)} & \hat{\mathcal{I}} & 0_5 & 0_{5\times 2} & 0_5 \end{bmatrix} = 0 \tag{6.35}$$

where $0 \neq [\xi^T, \theta] \geq 0^T$ and $[v^T, \eta^T] \geq 0^T$.

Expanding (6.35) yields

$$\xi^T P_\pi + v^T(E_\pi + E_p) + w^T \mathcal{I} = 0, \tag{6.36}$$

$$\xi^T P_\pi + v^T E_\pi + z^T \hat{\mathcal{I}} = 0, \tag{6.37}$$

$$\xi^T P_m + v^T E_m = 0, \tag{6.38}$$

$$v^T E_\kappa + \eta^T = 0, \tag{6.39}$$

$$\theta + \eta^T \kappa^g = 0. \tag{6.40}$$

In order for capital taxation to be redundant, it must be that w and z are both identically zero. However, if z is not zero, then w cannot be. Therefore a necessary condition for the redundancy of capital taxation is that z be equal to zero. Expanding (6.38) and (6.37) yields

$$\sum_t \xi_t - v_1^T \nabla_m \alpha_1 + (v_3 - v_2)\nabla_m \kappa_1^1 - v_4^T \nabla_m \alpha_2 + (v_6 - v_5)\nabla_m \kappa_2^2 - v_7^T \nabla_m \alpha_3^2 = 0, \tag{6.41}$$

$$-\xi_1 \alpha_1^{0T} - \xi_2 \alpha_1^{1T} - v^T \nabla_{\pi_1} \alpha_1 + (v_3 - v_2)\nabla_{\pi_1} \kappa_1^1 - v_4^T \nabla_{\pi_1} \alpha_2^1 + (z_1, 0_{n-1}^T) = 0, \tag{6.42}$$

$$\xi_2 \frac{\sigma_1}{\rho_1^2} \pi_1^T \alpha_1^1 - v_1^T \nabla_{\rho_1} \alpha_1^1 + (v_3 - v_2)\nabla_{\rho_1} \kappa_1^1 - v_4^T \nabla_{\rho_1} \alpha_2^1 + z_2 = 0, \tag{6.43}$$

$$-\xi_2 \frac{1}{\rho_1} \pi_1^T \alpha_1^1 - v_1^T \nabla_{\sigma_1} \alpha_1^1 + (v_3 - v_2)\nabla_{\sigma_1} \kappa_1^1 - v_4^T \nabla_{\sigma_1} \alpha_2^1 + z_3 = 0, \tag{6.44}$$

$$-\xi_2 \alpha_2^{1T} - \xi_3 \frac{\sigma_2}{\rho_2} \alpha_2^{2T} - v_1^T \nabla_{\pi_2} \alpha_1^1 + (v_3 - v_2)\nabla_{\pi_2} \kappa_1^1 - v_4^T \nabla_{\pi_2} \alpha_2 + (v_6 - v_5)\nabla_{\pi_2} \kappa_2^2 - v_7^T \nabla_{\pi_2} \alpha_3^2 = 0, \tag{6.45}$$

$$\xi_3 \frac{\sigma_2}{\rho_2^2} \pi_2^T \alpha_2^2 - v_4^T \nabla_{\rho_2} \alpha_2^2 + (v_6 - v_5)\nabla_{\rho_2} \kappa_2^2 - v_7^T \nabla_{\rho_2} \alpha_3^2 + z_4 = 0, \tag{6.46}$$

$$-\xi_3 \frac{1}{\rho_2} \pi_2^T \alpha_2^2 - v_4^T \nabla_{\sigma_2} \alpha_2^2 + (v_6 - v_5)\nabla_{\sigma_2} \kappa_2^2 - v_7^T \nabla_{\sigma_2} \alpha_3^2 + z_5 = 0, \tag{6.47}$$

and

$$-\xi_3\alpha_3^{2T} - v_4^T\nabla_{\pi_3}\alpha_2^2 + (v_6 - v_5)\nabla_{\pi_3}\kappa_2^2 - v_7^T\nabla_{\pi_3}\alpha_3^2 = 0. \tag{6.48}$$

From the argument in the proof of Theorem 2, z_1 is equal to zero.[27] To find the conditions under which z_2 and z_3 are zero, multiply (6.43) by ρ_1, (6.44) by σ_1, and add. Using the homogeneity of the demand functions yields

$$z_2\rho_1 + z_3\sigma_1 = 0. \tag{6.49}$$

Hence we cannot in general normalize both the capital tax and savings tax to be zero. In fact, by (6.43), z_2 and z_3 are identically zero if and only if

$$\xi_2\frac{1}{\rho_1}\pi_1^T\alpha_1^1 - v_1^T\nabla_{\sigma_1}\alpha_1^1 - (v_3 - v_2)\nabla_{\sigma_1}\kappa_1^1 - v_4^T\nabla_{\sigma_1}\alpha_2^1 = 0. \tag{6.50}$$

To see the implications of (6.50) first differentiate the period-one budget constraint of generation one with respect to σ_1 to obtain

$$\pi_1^T\nabla_{\sigma_1}\alpha_1^1 + \rho_1\nabla_{\sigma_1}\kappa_1^1 = 0. \tag{6.51}$$

Substituting this into (6.50) and rearranging terms yields

$$\xi_2\frac{1}{\rho_1}\pi_1^T\alpha_1^1 = \left[v_1^T + (v_3 - v_2)\frac{\pi_1^T}{\rho_1}\right]\nabla_{\sigma_1}\alpha_1^1 + v_4^T\nabla_{\sigma_1}\alpha_2^1. \tag{6.52}$$

From Theorem 1, a necessary condition for all eight normalizations to be harmless yields in addition to the above that

$$v^T = (v_1^T, v_2, v_3, v_4^T, v_5, v_6, v_7^T) = \left[\mu_1(p_1^T, r_1), \mu_2(s_1, p_2^T, r_2), \mu_3(s_2, p_3^T)\right] \tag{6.53}$$

where μ_t is a function only of period t producer prices. Thus, (6.52) can be rewritten as

$$\xi_2\frac{1}{\rho_1}\pi_1^T\alpha_1^1 = \left[\mu_1 p_1^T + (\mu_2 s_1 - \mu_1 r_1)\frac{\pi_1^T}{\rho_1}\right]\nabla_{\sigma_1}\alpha_1^1 + \mu_2 p_2^T\nabla_{\sigma_1}\alpha_2^1. \tag{6.54}$$

Carrying out the indicated differentiation and canceling ρ_1 yields

$$\xi_2\pi_1^T\alpha_1^1 = \left[\mu_1 p_1^T + (\mu_2 s_1 - \mu_1 r_1)\frac{\pi_1^T}{\rho_1}\right]\nabla_{\tilde{\pi}_1}\alpha_1^1\pi_1 + \mu_2 p_2^T\nabla_{\tilde{\pi}_1}\alpha_2^1\pi_1. \tag{6.55}$$

Using (6.53) it follows from (6.48) that ξ_3 is homogeneous of degree minus one in (π_1, π_2, π_3, m) and from (6.45) it follows that ξ_2 has the same property. The left side of (6.55) is homogeneous of degree zero in (π_1, π_2, m). This is true of the right side if and only if $\mu_2 s_1 - \mu_1 r_1 = 0$; that is, we are in the region where the government has positive savings. This leaves us with

$$\xi_2\pi_1^T\alpha_1^1 = \mu_1 p_1^T\nabla_{\sigma\tilde{\pi}_1}\alpha_1^1\pi_1 + \mu_2 p_2^T\nabla_{\tilde{\pi}_1}\alpha_2^1\pi_1. \tag{6.56}$$

[27] This follows from the homogeneity of demand functions and the budget constraints.

If $\xi_2 > 0$, then dividing (6.54) by ξ_2 leaves the left side independent of producer prices while the right side depends upon them. Hence both sides must be zero which yields a contradiction because we have assumed positive savings by generations one and two and hence that $-\pi_1^T \alpha_1^1 = \rho_1 \kappa_1^1 > 0$. If $\xi_2 = 0$, then the right side of (6.54) must equal zero. Rearranging the right side of (6.54) so that we have but period one producer prices on one side of the equation and period two producer prices on the other side shows that

$$\nabla_{\tilde{\pi}_1} \alpha_1^1 \tilde{\pi} = 0 \quad \text{and} \quad \nabla_{\tilde{\pi}_1} \alpha_2^1 \tilde{\pi} = 0. \tag{6.57}$$

That is, the demand functions of generation one are homogeneous of degree zero in $\tilde{\pi}_1$ as well as being homogeneous of degree zero in all of its arguments. Thus, we have

$$\begin{aligned} \alpha_1^1(\tilde{\pi}_1, \pi_2, m) &= \alpha_1^1\left(\frac{\tilde{\pi}_1}{m}, \pi_2, m\right) \\ &= \alpha_1^1\left(\frac{\tilde{\pi}_1}{m^2}, \frac{\pi_2}{m}, 1\right). \end{aligned} \tag{6.58}$$

From (6.58) it is clear that the demand functions are no longer homogeneous of degree zero in $(\tilde{\pi}_1, \pi_2, m)$ yielding a contradiction. ■

6.6. Proof of Theorem 4

Computing anew the set of feasible Pareto improvements, the set of new Pareto-optima are described by

$$\xi_2 + \xi_3 = \mu_2 p_2^T \nabla_m \alpha_2 + (\mu_3 s_2 - \mu_2 r_2) \nabla_m \kappa_2^2 + \mu_3 p_3^T \nabla_m \alpha_3^2, \tag{6.59}$$

$$\xi_2 \alpha_2^{1T} + \xi_3 \frac{\sigma_2}{\rho_2} \alpha_2^{2T} = -\mu_2 p_2^T \nabla_{\pi_2} \alpha_2 + (\mu_3 s_2 - \mu_2 r_2) \nabla_{\pi_2} \kappa_2^2 + \mu_3 p_3^T \nabla_{\pi_2} \alpha_3^2, \tag{6.60}$$

plus (6.46)—(6.48). Multiply (6.46) by ρ_2, (6.47) by σ_2 and add to obtain

$$z_4 \rho_2 + z_5 \sigma_2 = 0 \tag{6.61}$$

which shows that in general one cannot normalize both the tax on the storable good and tax on it sale to be zero. Repeating the argument of Theorem 3 using the budget constraint of generation two establishes the result. ■

REFERENCES

Allais, M., *Economie et Intérêt*, Paris: Imprimerie Nationale, 1947.

Atkinson, A. and J. Stiglitz (1976) "The Design of Tax Structure: Direct Versus Indirect Taxation", *Journal of Public Economics* **6**, 1976, 55-75.

Auerbach, A., "The Deadweight Loss from Non-Neutral Capital Income Taxation", *Journal of Public Economics* **40**, 1989, 1-36.

Auerbach, A., "Public Sector Dynamics", in *Modern Public Finance* edited by J. Quigley and E. Smolensky, Cambridge, Mass: Harvard, 1994.

Auerbach, A. and J. Hines, "Taxation and Economic Efficiency" in *Handbook of Public Economics* Vol. 3, edited by A. Auerbach and M. Feldstein, Amsterdam: North-Holland, 2002.

Auerbach, A. and L. Kotlikoff, *Dynamic Fiscal Policy*, Cambridge, Mass: Harvard, 1987.

Bernheim, B. D., "Taxation and Saving", in *Handbook of Public Economics*, Vol. 3, edited by A. Auerbach and M. Feldstein, Amsterdam: North-Holland, 2002.

Blackorby, C. and W. Diewert, "Expenditure Functions, Local Duality, and Second Order Approximations", *Econometrica* **47**, 1979, 579-601.

Bradford, D., "On the Incidence of Consumption Taxes", in *The Consumption Tax: A Better Alternative?* edited by C. Walker and M. Bloomfield, Cambridge, Mass: Ballinger, 1987.

Browning, M. and J. Burbidge, "Consumption and Income Taxation", *Oxford Economic Papers* **42**, 1990, 281-292.

Chamley, C., "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives", *Econometrica* **54**, 1986, 607–622.

Diamond, P. and J. Mirrlees, (1971) "Optimal Taxation and Public Production 1: Production Efficiency" *American Economic Review* **61**, 1971, 8-27.

Diamond, P., "National Debt in a Neoclassical Growth Model", *American Economic Review* **55**, 1965, 1125-1150.

Diewert, W., "Optimal Tax Perturbations", *Journal of Public Economics* **10**, 1978, 139-177.

Diewert, W., M. Avriel, and I. Zang, "Nine Kinds of Quasi-Concavity and Concavity", *Journal of Economic Theory* **25**, 1981, 397-420.

Feldstein, M., "The Welfare Cost of Capital Income Taxation", *The Journal of Political Economy* **86**, 1978, S29-S51.

Feldstein, M., "The Second-Best Theory of Differential Capital Taxation", *Oxford Economic Papers* **42**, 1990, 256-267.

Guesnerie, R., "On The Direction of Tax Reform", *Journal of Public Economics* **7**, 1977, 179–202.

Guesnerie, R., *A Contribution to the Pure Theory of Taxation*, Cambridge, UK: Cambridge University Press, 1995.

Judd, K., "The Welfare Cost of Factor Taxation in a Perfect Foresight Model", *Journal of Political Economy* **95**, 1987, 675–709.

Judd, K., "Optimal Taxation and Spending in General Competitive Growth Models", *Journal of Public Economics* **71**, 1999, 1–26.

Kesselman, J. "Assessing a Direct Consumption Tax to Replace the GST", *Canadian Tax Journal* **42**, 1994, 709-803.

Mangasarian, O., *Nonlinear Programming*, New York: McGraw-Hill, 1969.

Meade, J., *The Structure and Reform of Direct Taxation*, London: Allen and Unwin, 1978.

Mirrlees, J., "The Theory of Optimal Taxation" in the *Handbook of Mathematical Economics*, Vol. 3, edited by K. Arrow and M. Intrilligator, Amsterdam: North Holland, 1985.

Myles, G., *Public Economics*, Cambridge, UK: Cambridge University Press, 1995.

Ordover, J. and E. Phelps, "The Concept of of Optimal Taxation in the Overlapping-Generations Model of Capital and Wealth", *Journal of Public Economics* **12**, 1979, 1-26.

Samuelson, P., "An Exact Consumption-Loan Model of Interest With or Without the Social Contrivance of Money", *Journal of Political Economy* **46**, 1958, 467-482.

Weymark, J., "A Reconciliation of Recent Results in Optimal Taxation", *Journal of Public Economics* **12**, 1979,171-189.