The effects of population aging on optimal redistributive taxes in an overlapping generations model

Brett, Craig
Mount Allison University

25. March 2008

The Effects of Population Aging on Optimal Redistributive Taxes in an Overlapping Generations Model

Craig Brett

Department of Economics,
Mount Allison University,
144 Main Street, Sackville NB, E4L 1A7, Canada
(e-mail: cbrett@mta.ca)

March 25, 2008

*This research was generously supported by the Canada Research Chairs Program. I am grateful to Frank Strain and John Weymark for helpful comments. Errors, omissions and views are entirely my own.

Abstract:

"The Effects of Population Aging on Optimal Redistributive Taxes in an Overlapping Generations Model"

by

Craig Brett

The impact of population aging on the steady state solution to a Ordover and Phelps (1979) overlapping generations optimal nonlinear income tax problem with two types of workers and quasilinear-in-leisure preferences is investigated. A decrease in the rate of population growth, which leads to an aging population, increases the relative price of consumption per person in retirement, which tends to decrease optimal consumption for retirees of both skill types. It is also shown that the optimal steady state rate of interest equals the rate of population growth. As a result, the steady state interest rate unambiguously declines when the rate of population growth declines. The resulting adjustments in production plans has an ambiguous effect on the aggregate wage rate. This article identifies factors contributing to an increase in the aggregate wage when the population ages, namely normality of consumption in retirement, complementarity between capital and labor in production, and a large capital deepening effect relative to the increase in dependency owing to demographic change. Depending on the sign of this wage effect, ambiguities may arise in the direction of change in the optimal steady state consumption and production plans. It is also shown that the optimal marginal income tax rates are independent of the rate of population growth.

JEL classification: D82, H21

Keywords: optimal income taxation, overlapping generations model, population aging

1 Introduction

The effects of population aging on public finances are potentially profound. As Visco (2001) argues, population aging is expected to exert upward pressure on government expenditure. These pressures on expenditure may be cause for policy concern in that they call for redistribution of resources among generations.[1] Moreover, governments must call upon distortionary taxation to fund expenditures. Thus, it is also important to understand how population aging affects the revenue side of the public budget.

This article addresses the effects of population aging on optimal distortionary income taxes using a model that embeds the Mirrlees (1971) personal income tax framework into an overlapping generations model. The distortionary effects of personal income taxation are modeled as arising out of information asymmetry between the taxation authority and individuals. The population of workers is divided into two classes, differing in exogenous labor productivity. The two types of workers are assumed to be perfect substitutes in production. Following the standard set of assumptions, the taxation authority is assumed to observe only market earnings, which are a mixture of innate ability and hours of work. Because the goal of this paper is to examine the effects of parameter changes on the optimal tax schedule, and not to elaborate on further properties of the tax schedule itself, the government is assumed to use only the nonlinear income tax to further its redistributive goals.

The dynamic structure is equally simple, deriving in a straightforward way from a commonly used deterministic overlapping generations model. Population aging is modeled by allowing the number of workers in each generation to grow at a constant rate per generation, and allowing the rate of growth of new workers to decline. The model is very similar to the one used by Ordover and Phelps (1979) to describe optimal income taxes with a continuum of workers.[2] A two–class version of the Ordover-Phelps model,

[1] See, however, McDaniel (2003) for a critical assessment of "apocalyptic demography" in the Canadian context.

[2] See Myles (1995, pp. 509–514) for a textbook treatment of this analysis.

allowing for the possibility of endogenous relative wages, was introduced by Pirttilä and Tuomala (2001) in order to analyze capital taxation and public good provision.

The effects of population aging are demonstrated by deriving how the steady state optimal income tax changes in response to a change in the rate of population growth. The steady state envisioned is one in which the capital per worker of each type is constant over time. Given the assumptions on technology, capital per unit of labor in efficiency units is constant at a steady state. Confining attention to changes in steady states renders the analysis similar in form and in spirit to the literature on the comparative statics of nonlinear taxation, pioneered by Weymark (1987) and recently extended by Hamilton and Pestieau (2005), Boadway and Pestieau (2007), Simula (2007), Brett and Weymark (2008a), and Brett and Weymark (2008c). The overlapping generations framework endogenizes many of the variables that are assumed to be exogenous in a static nonlinear income tax framework, such as the amount of labor in efficiency units required to produce one unit of the consumption good. However, the relative wage rates and the rate of population growth are exogenous. Thus, it is reasonable to carry out comparative steady state analysis with respect to these variables. I choose not to consider the effects of changes in the relative wage rate, because the model is similar enough to that of Weymark (1987) and Brett and Weymark (2008a) that their insights should carry over with only minor modifications. Changes in the age structure of the population, however, have yet to receive attention in the nonlinear taxation literature.

Population aging has multiple effects on the model economy presented in this article. There is the usual capital deepening effect, whereby the capital stock is used in conjunction with relatively fewer workers. There is also the standard dependency effect, as the relative number of retirees increases. This effect also acts to increase the price of consumption in retirement relative to consumption when working. Because relative wages are fixed and preferences are assumed to be separable between labor and consumption, there are no capital market distortions in the steady state optimum. In particular, the optimal steady state rate of interest equals the rate of population growth. Thus, population aging leads to a decline in the optimal steady state interest rate. There is an ambiguous effect on the aggregate wage rate. I show that the aggregate wage tends to increase when the population ages when some combination of the following factors is sufficiently strong in the neighborhood of the initial steady state optimum: normality of consumption in retirement, complementarity between capital and labor in production, or a large capital deepening effect relative to the increase in dependency.

When the wage rate increases in response to population aging, few unambiguous comparative steady state results are available. The implicit marginal income tax rates remain unchanged. In addition, consumption when working increases under the additional assumption of time-separable utility. However, when the wage rate decreases in response to population aging (due, for example, to a large dependency effect) and utility is time-separable, it is possible to deduce the direction of change in most of the variables of interest: consumption falls for individuals of both skill types when working and when retired; the per-capita capital stock and aggregate effective labor rise; optimal implicit marginal income tax rates remain unchanged.

The effects of population aging on steady state consumption has received attention in models with fixed per-person labor supply, time-separable utility and no within-cohort heterogeneity. Cutler et al. (1990) provide a detailed analysis of anticipated changes in steady state consumption owing to demographic changes in the United States. Meijdam and Verbon (1997) compute the effects of an aging population on steady state consumption in the presence of public pension schemes. In their model, only capital deepening and dependency effects arise, with the latter dominating. Consequently, population aging reduces steady state consumption.

The remainder of this article is organized as follows. Section 2 provides a description of the model, paying careful attention to the information assumptions contained therein. Section 3 derives some qualitative features of optimal taxation in this environment. Section 4 provides a mathematical formulation of the comparison among steady states and offers verbal statements of the effects of population aging on the optimal tax schedule.

Some concluding remarks are then offered. Proofs are gathered in an Appendix.

2 The Model

There are two types of workers born each period. During the first period of their lives, they supply labor elastically and they consume. In the second period of life, each individual retires. Within a generation, individuals differ in productivity. Denote the productivity of a person of type i by a_i, $i = 1, 2$, $a_1 < a_2$.[3] Thus, if a person supplies l_i units of labor, her effective labor is $y_i := a_i l_i$. At any date (apart from the start–up period), t, the following types of individuals are alive: young individuals, some of type a_1,the others with productivity a_2; retired individuals, born at time $t-1$, living off the proceeds of their savings. I assume that the number of workers varies from period-to-period, but that the within-period composition of workers is fixed. For simplicity, I assume that exactly half the workers in each time period are of type 1, and denote the number of such workers by N^t. The number of workers of each type evolves according to the equation

$$N^t = (1+n)N^{t-1}, \tag{1}$$

which states that the population grows at a constant rate n. The focus of this paper is to investigate how changes in n affect the optimal tax system in the steady state.

Total output at any date t is a function of the capital stock, K^t, and total effective labor,

$$Y^t := N^t(y_1^t + y_2^t). \tag{2}$$

Let $F(K^t, Y^t)$ be the production function, assumed to exhibit constant returns to scale and to be strictly concave, with isoquants that do not intersect the coordinated axes, for all positive levels of output. The prices of inputs are determined by the profit-

[3]Throughout this analysis, subscripts are used to denote the type of an individual and superscripts denote the date of birth of an individual. Quantities denoted without subscripts are within-period aggregates.

maximization conditions

$$r^t = F_k(k^t, y^t); \quad w^t = F_y(k^t, y^t), \tag{3}$$

where w^t is the price of effective labor and r^t is the rental price of capital. The before-tax income of an individual is given by

$$z_i^t := w^t a_i l_i^t = w^t y_i^t. \tag{4}$$

Total consumption at time t is made up of consumption by the young born at that date, denoted by the symbol c, and the spending in retirement of those born at date $t-1$, denoted by x. Depreciation is assumed away, so that the capital stock evolves according to the equation

$$K^{t+1} = F(K^t, Y^t) + K^t - N^t(c_1^t + c_2^t) - N^{t-1}(x_1^{t-1} + x_2^{t-1}). \tag{5}$$

That is, capital next period equals current output plus current capital less total consumption of those currently alive. Because production exhibits constant returns to scale, the evolution of the capital stock per young worker of each type can be tracked with the equation

$$(1+n)k^{t+1} = f(k^t, y^t) + k^t - c_1^t - c_2^t - \frac{1}{1+n}(x_1^{t-1} + x_2^{t-1}), \tag{6}$$

where lowercase quantities are their respective uppercase analogs divided by N^t and

$$f(k^t, y^t) = F\big(\frac{K^t}{N^t}, \frac{Y^t}{N^t}\big) = \frac{1}{N^t} F(K^t, Y^t). \tag{7}$$

The government can observe both z and w, but cannot observe l or a. This accords with the standard assumptions of nonlinear tax theory. It is equivalent to say that the planner can observe y. Implicitly, then, the planner can also observe k. Because l is unobserved, the planner must resort to distortionary taxation. Exactly which tax instruments are available to the planner depend on the further assumptions one makes about the use of non-income information. It is assumed that the planner knows the age of each individual, so that the young cannot pretend to be old, nor can the old pretend to be young. The old do not work, so there is no direct interaction between them and the income tax schedule. Thus, the only concern is that the young may have incentive to misrepresent their ability. Given that information about type is revealed when young, the planner can distinguish between retirees of the same generation. Thus, without loss of generality, it is assumed that the tax on consumption of the old is pre-paid at the end of the first period of life. Because retirees simply consume their after-tax savings, one need not worry about the potential ratchet effect arising from disclosure of information in the first period.[4]

Individuals derive utility from consumption when young and consumption during retirement. Moreover, they are assumed to have a disutility of labor. All individuals have a common utility function, assumed to be quasi-linear in labor supply, so that preferences are represented by

$$V(c,x,l) = v(c,x) - l. \tag{8}$$

The function v is assumed to be twice continuously differentiable at all $(c,x) \neq (0,0)$, continuous and nondecreasing on $\mathbb{R}^2_+$, strictly increasing on $\mathbb{R}^2_{++}$, and strictly concave on $\mathbb{R}^2_{++}$ with $v(0,0) = 0$, $v_c(0,x) = \infty$ for all $x > 0$, $v_x(c,0) = \infty$ for all $c > 0$, $v_c(c,x) \to 0$ as $c \to \infty$ for all $x \geq 0$, and $v_x(c,x) \to 0$ as $x \to \infty$ for all $c \geq 0$. The limiting assumptions on v ensure that the optimal tax problem has a solution and that individuals of both types have positive consumption of both goods at this solution.

Differences in ability generate differences in preferences over consumption and effective labor, which, following Weymark (1987) are conveniently represented by the type-specific monotonic transformation of (8)

$$U^i(c,x,y) = a_i v(c,x) - y. \tag{9}$$

[4]See Dillén and Lundholm (1996) for an exposition of a two-period model in which the taxation authority sets an optimal linear tax schedule for workers who supply labor in both time periods. Apps and Rees (2006), Berliant and Ledyard (2005), and Brett and Weymark (2008b) study nonlinear income taxes with labor supply in two periods and the potential for a ratchet effect.

Equation (9) describes preferences over variables that the planner can observe. This representation is linear in both y and in the unobserved characteristic a. This linearity is heavily exploited in the analysis of Section 3.

The taxation authority is assumed to select a tax system that specifies an amount of tax to be paid on labor income, along with a levy on the amount of savings. Equivalently, it can be modeled as choosing the consumption levels and effective labor time for each type of worker at each date in time, subject to incentive compatibility constraints. I analyze only the case in which a person of high ability may wish to misrepresent its type. That is, at each date, only one form of self-selection constraints is considered, namely

$$a_2 v(c_2^t, x_2^t) - y_2^t \geq a_2 v(c_1^t, x_1^t) - y_1^t \quad t = 1, 2, \ldots \tag{10}$$

This is the case most commonly analyzed in the literature. Moreover, this is the form of the self-selection constraint that can easily be shown to bind under the assumptions used in Section 3 below.

3 Optimal Taxation In a Steady State

I consider only taxation in the steady state, defined as a state in which all variables per worker of each type remain constant over time. At a steady state, the aggregate resource constraint (6) reduces to

$$f(k, y) - nk = c_1 + c_2 + \frac{1}{1+n}(x_1 + x_2), \tag{11}$$

where variables without time superscripts denote steady state values.

The government is assumed to maximize a weighted sum of steady-state utilities[5]

$$\mathcal{W} = \alpha_1[v(c_1, x_1) - l_1] + \alpha_2[v(c_2, x_2) - l_2]. \tag{12}$$

The welfare function (12) is equivalent to the weighted average utilitarian criterion, where the weights are over the two types. In order for the sum in (12) to be meaningful, the

[5] The total population size, N, can be incorporated into the welfare weights.

utility function v must have cardinal significance. Welfare can be re-expressed in terms of observable variables as

$$\mathcal{W} = \lambda_1[a_1 v(c_1, x_1) - y_1] + \lambda_2[a_2 v(c_2, x_2) - y_2], \tag{13}$$

where $\lambda_i = \alpha_i / a_i$ is the skill-normalized welfare weight assigned to individuals of type i. I assume that $\lambda_1 > \lambda_2$, which implies that a redistribution of before-tax income (labor supply) from individuals of type 1 to individuals of type 2 is always welfare improving. Thus, the self-selection constraint (10) must bind at a solution to the planner's problem. In the steady state, this binding constraint is given by the equation

$$a_2 v(c_2, x_2) - y_2 = a_2 v(c_1, x_1) - y_1. \tag{14}$$

Following Weymark (1986), I also assume that the skill-normalized welfare weights sum to the number of types of individuals in the economy; that is,

$$\lambda_1 + \lambda_2 = 2. \tag{15}$$

The Steady State Optimal Nonlinear Income Tax Problem. The government chooses an allocation $(c_1, c_2, x_1, x_2, k, y_1, y_2) \in \mathbb{R}^7_+$ to maximize the social welfare function (13) subject to the resource constraint (11) and the binding self-selection constraint (14).

In order to carry out any comparative static exercise, it is first necessary to show that the problem at hand has a unique solution. Lemma 1 establishes that this is so for the Steady State Optimal Nonlinear Income Tax Problem.

Lemma 1. *The Steady State Optimal Nonlinear Income Tax Problem has a unique solution.*

The quasi-linear form of the utility function allows for a straightforward substitution of the self-selection constraint (14) into the social welfare function. The result of this substitution is summarized in the following Lemma.

Lemma 2. *Let* $(\tilde{c}_1, \tilde{c}_2, \tilde{x}_1, \tilde{x}_2, \tilde{k}, \tilde{y}_1, \tilde{y}_2)$ *solve the Steady State Optimal Nonlinear Income Tax Problem. Then* $(\tilde{c}_1, \tilde{c}_2, \tilde{x}_1, \tilde{x}_2, \tilde{k}, \tilde{y})$ *solves:*

$$\max_{(c_1,c_2,x_1,x_2,k,y)} \beta_1 v(c_1, x_1) + \beta_2 v(c_2, x_2) - y \quad \text{subject to} \quad (11), \tag{16}$$

where

$$\beta_1 = \lambda_1 a_1 + (1 - \lambda_1)(a_2 - a_1) \tag{17}$$

and

$$\beta_2 = a_2. \tag{18}$$

Brett and Weymark (2008c) call β_1 and β_2 the reduced form welfare weights. These weights measure the marginal social value of an increase in the utility from consumption (in either or both periods) of the individuals of the two types. The normalization $\lambda_1+\lambda_2 = 2$ and the assumptions that $\lambda_1 > \lambda_2$ and $a_2 > a_1$ imply that $\beta_2 > a_1 > \beta_1$. The social value of the utility of individuals of type 1 is less than the raw welfare weight a_1 because this utility brings with it added temptation for persons of type 2 to mimic those of type 1.[6]

Lemma 2 establishes that all of the components of the solution to the Steady State Optimal Nonlinear Tax Problem, except the effective labor supplies, can be found by solving the simpler maximization problem (16). The solution to (16) can be substituted into the definition of aggregate effective labor and into (14) in order to compute the effective labor supplies.[7] Performing these substitutions yields Lemma 3.

Lemma 3. *Let* $(\tilde{c}_1, \tilde{c}_2, \tilde{x}_1, \tilde{x}_2, \tilde{k}, \tilde{y})$ *solve (16). Then the solution to the Steady State Optimal Nonlinear Income Tax Problem is* $(\tilde{c}_1, \tilde{c}_2, \tilde{x}_1, \tilde{x}_2, \tilde{k}, \tilde{y}_1, \tilde{y}_2)$ *where*

$$\tilde{y}_1 = \frac{1}{2}\left(\tilde{y} - a_2[v(\tilde{c}_2, \tilde{x}_2) - v(\tilde{c}_1, \tilde{x}_1)]\right), \tag{19}$$

[6] Weymark (1987, p. 1171) provides a detailed discussion justifying the exact form of the reduced form welfare weight.

[7] One technical complication remains. There is no guarantee that the solution procedure outlined here guarantees that $\tilde{y}_1 > 0$. I assume this to be the case throughout the remainder of the analysis. With this assumption, all elements of the optimal program can be shown to be positive.

and

$$\tilde{y}_2 = \frac{1}{2}\left(\tilde{y} + a_2[v(\tilde{c}_2, \tilde{x}_2) - v(\tilde{c}_1, \tilde{x}_1)]\right). \tag{20}$$

After introducing the variable μ to describe the shadow value of the constraint (11), the solution to (16) can be easily described in terms of the following first-order conditions.

$$\beta_i v_{c_i} - \mu = 0, \quad i = 1, 2; \tag{21}$$

$$\beta_i v_{x_i} - \frac{\mu}{1+n} = 0, \quad i = 1, 2; \tag{22}$$

$$-1 + \mu f_y = 0; \tag{23}$$

$$f_k - n = 0. \tag{24}$$

In fact, the solution is completely described by the necessary conditions (21)–(24) and the resource constraint (11). It follows directly from (23) that $\tilde{\mu} > 0$. Moreover, the qualitative properties of the optimal tax system, including its implied behavioral distortions, can be derived from equations (21)–(24). These properties are summarized in Proposition 1.

Proposition 1. *The following statements hold at the solution to the Steady State Optimal Nonlinear Tax Problem.*

(i) *The rate of interest is equal to the biological rate of interest; that is,* $\tilde{r} = f_r(\tilde{k}, \tilde{y}) = n$.

(ii) *There are no distortions in saving behavior; that is,*

$$\frac{v_c(\tilde{c}_1, \tilde{x}_1)}{v_x(\tilde{c}_1, \tilde{x}_1)} = \frac{v_c(\tilde{c}_2, \tilde{x}_2)}{v_x(\tilde{c}_2, \tilde{x}_2)} = (1+n). \tag{25}$$

(iii) *The labor supply of individuals of type* 2 *is not distorted; that is,*

$$MRS_{2,lc} := \frac{1}{v_c(\tilde{c}_2, \tilde{x}_2)} = wa_2. \tag{26}$$

(iv) *The implicit marginal tax rate (IMTR) on the labor income of individuals of type* 1 *is positive; specifically,*

$$IMTR_1 := 1 - \frac{1}{wa_1 v_c(\tilde{c}_1, \tilde{x}_1)} = (\lambda_1 - 1)\left(\frac{a_2 - a_1}{a_1}\right). \tag{27}$$

Parts (i) and (ii) arise because the planner has no reason to distort savings decisions at the margin. Because preferences are separable between consumption and labor supply, low-productivity workers and high-productivity workers considering the possibility of mimicking low-productivity workers are each willing to trade consumption across time at the same implicit prices. Thus, the taxation authority can gain no informational advantage by distorting this margin. Part (iii) is the traditional no distortion result for workers of the higher type. Part (iv) implies that low-skilled individuals face a positive implicit marginal tax rate. Naturally, the specific form of the marginal tax rate is similar to the form found by Weymark (1987).[8] An immediate consequence of (26) and (27) is that changes in n have no effect on the optimal implicit marginal income tax rates faced by both types of individuals.

4 The Effects of Aging on the Optimal Tax Schedule

In order to assess the effects of an aging population on the solution to the Steady State Optimal Nonlinear Income Tax Problem, it is necessary to describe how its solution varies with the population growth parameter n. The lower the value of n, the larger is the cohort of retirees relative to the cohort of workers. Proposition 2 establishes that it is possible to carry out this comparative static analysis.

Proposition 2. *The optimality conditions (11) and (21)–(24) define a continuously differentiable solution function* $F\colon \mathbb{R}_+ \to \mathbb{R}^7_{++}$ *of the problem (16) with* $n \mapsto (\tilde{c}_1, \tilde{x}_1, \tilde{c}_2, \tilde{x}_2, \tilde{y}, \tilde{k}, \tilde{\mu})$.

[8] Weymark (1987) does not give an explicit statement of the analogous result. However, combining his equations (37) and (A.1) for any unbunched individuals yields a generalization of equation (A.8) used in the proof of Proposition 1.

For all $n \in \mathbb{R}_+$, the derivative DF of F at n is given by

$$DF(n) = (A^{-1}b)(n), \tag{28}$$

where

$$A(n) = \begin{bmatrix} \beta_1 v_{c_1c_1} & \beta_1 v_{c_1x_1} & 0 & 0 & 0 & 0 & -1 \\ \beta_1 v_{c_1x_1} & \beta_1 v_{x_1x_1} & 0 & 0 & 0 & 0 & -(1+n)^{-1} \\ 0 & 0 & \beta_2 v_{c_2c_2} & \beta_2 v_{c_2x_2} & 0 & 0 & -1 \\ 0 & 0 & \beta_2 v_{c_2x_2} & \beta_2 v_{x_2x_2} & 0 & 0 & -(1+n)^{-1} \\ 0 & 0 & 0 & 0 & \mu f_{yy} & \mu f_{ky} & f_y \\ 0 & 0 & 0 & 0 & f_{ky} & f_{kk} & 0 \\ -1 & -(1+n)^{-1} & -1 & -(1+n)^{-1} & f_y & 0 & 0 \end{bmatrix} \tag{29}$$

and

$$b(n) = \begin{bmatrix} 0 \\ -(1+n)^{-2}\tilde{\mu} \\ 0 \\ -(1+n)^{-2}\tilde{\mu} \\ 0 \\ 1 \\ \tilde{k} - (1+n)^{-2}\tilde{x} \end{bmatrix}, \tag{30}$$

and where all expressions on the right-hand sides of (29) and (30) are evaluated at the solution to (16).

Equations (28)–(30) characterize, albeit opaquely, how the optimal allocation changes in response to a change in the rate of population growth. As is shown in the proof of Proposition 2, the structure of the matrix $A(n)$ makes it possible to derive an explicit formula for its inverse. This formula, in turn, can be used to derive expressions for the terms in (28). These expressions are contained in Corollaries 1–4 below.

Corollary 1. *There exists numbers $\Delta_1, \Delta_2, \Delta_f > 0$ and $\theta < 0$ such that*

$$\frac{d\tilde{\mu}}{dn} = \theta \left[\sum_{i=1}^{2} \left\{ \frac{\tilde{\mu}}{\Delta_i (1+n)^2} \left(\frac{v_{c_ic_i}}{1+n} - v_{c_ix_i} \right) \right\} - \frac{\tilde{\mu}}{\Delta_f} f_y f_{ky} \right] - \theta \left[\tilde{k} - \frac{\tilde{x}}{(1+n)^2} \right]. \tag{31}$$

Moreover, if x is (in the neighborhood of the initial optimum) a normal good for each individual and $(1+n)^2\tilde{k} > \tilde{x}$ then $\dfrac{d\tilde{\mu}}{dn} > 0$.

Corollary 1 shows that the sign of the effect of n on shadow value of the resource constraint is, in general, ambiguous. Normality of x is sufficient to sign the first term in (31). The greater source of ambiguity is the final term, which captures the direct effect of a change in n on the steady state resource constraint. As in all overlapping generations models, an increase in n has both a capital spreading effect, as more workers arrive to work with the existing capital stock, and a reduced dependency effect, as the relative number of retirees falls. While Meijdam and Verbon (1997) are able to sign the relative magnitudes of the capital spreading and dependency effects in their model of public pensions supported by lump-sum taxation, it does not appear possible to do so in the current second-best framework. The condition expressed in the Corollary posits that the capital spreading effect is stronger than the dependency effect. It is, however, possible for the optimal μ to increase with n when the dependency effect dominates capital spreading, provided the dependency effect does not also outweigh the first term in (31). It follows from (3) and (23) that $\tilde{\mu} = 1/\tilde{w}$. Hence, the shadow value of the resource constraint varies inversely with the aggregate wage rate at the optimum. It seems plausible to expect that an increase in n, which raises the supply of workers, decreases the aggregate wage rate. If this is so, then it is plausible that $\tilde{\mu}$ increases with n.

At the solution to the Steady State Optimal Nonlinear Income Tax Problem, the marginal net social value of consumption when young is equal to the marginal social cost of acquiring the resources to finance that consumption, $\tilde{\mu}$. As Corollary 1 shows, an increase in n typical changes the marginal social cost of consumption. For concreteness, suppose that $\tilde{\mu}$ increases. Then, there exists an incentive to economize on the now socially more expensive consumption, and one might expect the optimal consumption when young to fall for all individuals. This intuition must be modified, however, if preferences over consumption are not additive across time periods. The taxation authority can restore the balance between the marginal benefits and marginal costs of consumption by any combination of changes in c and x that produce an appropriate increase in the marginal social value of consumption when young. Corollary 2 provides a formal summary of this discussion.

Corollary 2. *For the same $\Delta_1, \Delta_2 > 0$ as in Corollary 1,*

$$\frac{d\tilde{c}_i}{dn} = \frac{1}{\Delta_i}\left[\left((v_{x_i x_i} - \frac{v_{c_i x_i}}{1+n}\right)\frac{d\tilde{\mu}}{dn} + \frac{\tilde{\mu} v_{c_i x_i}}{(1+n)^2}\right], \quad i = 1, 2. \tag{32}$$

Moreover, if (in the neighborhood of the initial optimum) $v_{c_i x_i} = 0$ then $\frac{d\tilde{c}_i}{dn}$ has the opposite sign of $\frac{d\tilde{\mu}}{dn}$, for $i = 1, 2$.

An increase in n has two, potentially offsetting, effects on the marginal social cost of consumption in retirement. Because the socially optimal interest rate equals the rate of population growth, an increase in n lowers the opportunity cost of x. In other words, the reduced dependency effect makes the consumption of retirees relatively less expensive. On the other hand, if the social value of resources $\tilde{\mu}$ increases with n, then all consumption, including consumption in retirement, becomes more socially expensive. The net effect on the opportunity cost of x is ambiguous.[9] Thus, it is impossible to sign the effect of an increase in n on x. Corollary 3 gives an algebraic rendering of the ambiguous effect of n on the optimal consumption in retirement.

Corollary 3. *For the same $\Delta_1, \Delta_2 > 0$ as in Corollary 1,*

$$\frac{d\tilde{x}_i}{dn} = \frac{1}{\Delta_i}\left[\left((-v_{c_i x_i} + \frac{v_{c_i c_i}}{1+n}\right)\frac{d\tilde{\mu}}{dn} - \frac{\tilde{\mu} v_{c_i c_i}}{(1+n)^2}\right], \quad i = 1, 2. \tag{33}$$

It is possible to sign the effect of n on x_i when $v_{c_i x_i}$ is sufficiently small and the dependency effect is sufficiently large (or x sufficiently inferior for the other type of individual) that $\tilde{\mu}$ decreases with n. In this case, the marginal social cost of x_i falls and the only effective way to reduce the marginal social benefit commensurately is to increase x_i.

It is impossible to sign the effects of an increase in n on the production side of the economy. Corollary 4 displays the potentially offsetting terms.

[9] Formally, both the numerator and denominator in the final term on the left-hand side of (22) increase.

Corollary 4. *For the same $\Delta_f > 0$ as in Corollary 1,*

$$\frac{d\tilde{y}}{dn} = \frac{1}{\Delta_f}\left[-f_{kk}f_y\frac{d\tilde{\mu}}{dn} - \tilde{\mu}f_{ky}\right], \tag{34}$$

and

$$\frac{d\tilde{k}}{dn} = \frac{1}{\Delta_f}\left[f_{ky}f_y\frac{d\tilde{\mu}}{dn} + \tilde{\mu}f_{yy}\right]. \tag{35}$$

An increase in the rate of population growth induces an increase in the rate of interest at the optimum. This increase in the rate of interest can be brought about by either a decrease in the capital stock or an increase in aggregate effective labor.[10] Without further restrictions on technology, it is impossible to tell which of these levers the taxation authority would pull. Ambiguity on the production side of the economy disappears when the dependency effect is sufficiently strong (or x sufficiently inferior) so the $\tilde{\mu}$ decreases with n. In this event, both capital and aggregate effective labor are optimally reduced in the steady state.

Table 1 collects the results of this analysis for the special case of utility that is additive across time. The results are re-phrased in terms of a decrease in n in order that they may cast direct light on the effects of population aging on the solution to the Steady State Optimal Nonlinear Income Tax Problem. The results are most clear-cut when the dependency effect is very strong near the initial optimum. In that case, population aging leads to a decrease in optimal consumption in each period for all individuals, to an increase in the steady state capital stock, and to an increase in aggregate effective labor. Moreover, the steady state aggregate wage falls, so that labor supply must increase for at least one type of individual. Nevertheless, the optimal implicit marginal tax rate remains unchanged. When the dependency effect is more muted, consumption when young increases for workers of both types, as does the aggregate wage, while the optimal implicit marginal tax rate remains unchanged. It is not possible to sign the directions of change in any other variables.

[10]Linear homogeneity and strict concavity of f imply $f_{ky} > 0$.

Table 1: The effects of a decrease in n when utility is additive across time

	Direction of the effect of a decrease in n	
Variable	Weak dependency effect ($\frac{d\tilde{\mu}}{dn} > 0$)	Strong dependence effect ($\frac{d\tilde{\mu}}{dn} < 0$)
$\tilde{c}_i$	increase	decrease
$\tilde{x}_i$	ambiguous	decrease
$\tilde{k}$	ambiguous	increase
$\tilde{y}$	ambiguous	increase
$\tilde{w}$	increase	decrease
$IMTR$	no change	no change

5 Concluding Remarks

The rate of population growth affects both the consumption and production sides of the model economy presented in this article. The real price of consumption in retirement is directly affected by the demographic make-up of the population, and the optimal aggregate wage rate is indirectly affected by the relative numbers of workers and retirees. For some goods and some initial configurations, the price and wage effects reinforce one another; for others, they offset. The results presented in this article identify the competing forces and give them a precise formulation. Moreover, the structure of the matrix A presented in Proposition 2 makes it apparent how to generalize the results to an arbitrary, finite number of skill-types.

It is striking that the implicit marginal income tax rate faced by low-skill workers is invariant to the rate of population growth. However, this result is easily reconciled with the finding by Weymark (1987) and Brett and Weymark (2008a) that for an arbitrary finite number of skill-types, optimal marginal income tax rates depend only on the distribution of skills and the relative welfare weights when preferences are quasi-linear in leisure. Using the same class of preferences, a continuum of skill types and a utilitarian objective, Boadway et al. (2000) find that optimal marginal income tax rates depend only on the distribution of skills. Thus, the invariance result presented here is, in part, an artifact of the form of preferences, but also a consequence of exogenous relative wages.

A natural extension to this work would be an analysis with endogenous relative wages. The model of Pirttilä and Tuomala (2001) could serve as a natural starting point. There are several challenges posed by such an extension. There is the obvious task of describing how changes in demographics might change relative wages. There is also the technical challenge of analyzing the Weymark model without recourse to skill-normalized welfare weights, because it is inappropriate to impose a normalization rule containing endogenous variables. Moreover, even when preferences are separable between consumption and leisure, there may exist a motivation for capital market distortions in the Pirttilä–Tuomala model. It is not immediately obvious, but potentially worthwhile to find out, how these distortions respond to demographic change.

Appendix

Proof of Lemma 1. Solving (14) for $a_2 v(c_2, x_2)$ and substituting into (13) yields

$$\mathcal{W} = \lambda_1 a_1 v(c_1, x_1) - \lambda_1 y_1 + \lambda_2 [a_2 v(c_1, x_1) - y_1 + y_2] - \lambda_2 y_2. \tag{A.1}$$

Employing the normalization $\lambda_1 + \lambda_2 = 2$ along with (A.1) yields

$$\mathcal{W} = \lambda_1 a_1 v(c_1, x_1) + (1 - \lambda_1) a_2 v(c_1, x_1) + a_2 v(c_1, x_1) - 2y_1. \tag{A.2}$$

Solving (14) for $a_2 v(c_1, x_1)$ and substituting into the penultimate term in (A.2) yields,

$$\mathcal{W} = [\lambda_1 a_1 + (1 - \lambda_1) a_2] v(c_1, x_1) + a_2 v(c_2, x_2) - y_1 - y_2. \tag{A.3}$$

Thus, the Steady State Optimal Nonlinear Income Tax Problem is equivalent to maximizing the objective (A.3) subject to the constraint (11). The curvature and boundary conditions on v and f guarantee a unique solution for the vector $(c_1, c_2, x_1, x_2, k, y_1 + y_2)$. I show in Lemma 3 how to compute unique solution values of y_1 and y_2 from the uniquely determined $(c_1, c_2, x_1, x_2, k, y)$. □

Proof of Lemma 2. Rearranging (A.3) yields

$$\mathcal{W} = [a_1 + (1-\lambda_1)(a_2 - a_1)]v(c_1, x_1) + a_2 v(c_2, x_2) - y. \tag{A.4}$$

Substituting (17) and (18) into (A.4) yields (16). In so doing, one constraint appearing in the Steady State Optimal Nonlinear Income Tax Problem has been substituted into its objective, and the variables y_1 and y_2 have been eliminated. However, the variable y is inserted and the constraint (11) remains. The Lemma follows. □

Proof of Lemma 3. In light of (14), the effective labor supplies can be found by solving the following linear system in the variables y_1 and y_2.

$$\begin{aligned} y_1 + y_2 &= \tilde{y} \\ -y_1 + y_2 &= a_2\left[v(\tilde{c}_2, \tilde{x}_2) - v(\tilde{c}_1, \tilde{x}_1)\right]. \end{aligned} \tag{A.5}$$

It is easy to check that (19) and (20) give the solution to the system (A.5). □

Proof of Proposition 1. Part (i) follows directly from equations (3) and (24). Part (ii) follows from dividing (21) by (22) for individuals of each type. By (23),

$$\mu = \frac{1}{f_y}. \tag{A.6}$$

Part (iii) follows from substituting (A.6) and (18) into (21) for individuals of type 2 and rearranging.

Using (3) in conjunction with the definition of $IMTR_1$ given in (27) yields

$$IMTR_1 = 1 - \frac{1}{a_1 f_y v_c(\tilde{c}_1, \tilde{x}_1)}. \tag{A.7}$$

Substituting (A.6) and (21) into (A.7) yields

$$IMTR_1 = 1 - \frac{1}{\frac{a_1}{\beta_1}} = \frac{a_1 - \beta_1}{a_1}. \tag{A.8}$$

Recalling the definition of β_1 from (17) and rearranging yields (27). □

Proof of Proposition 2. By Lemma 1, the first order necessary conditions define a solution function. Differentiating the first order conditions and the resource constraint yields

$$A \begin{bmatrix} dc_1 & dx_1 & dc_2 & dx_2 & dy & dk & d\mu \end{bmatrix}^\top = b\, dn, \tag{A.9}$$

where dependence on the parameter n is now expunged from the notation. The first zero in the final line of (29) follows from (24). In order to establish the Proposition, it suffices to show that the matrix A is invertible. To that end, introduce the partition

$$A = \begin{bmatrix} H & p \\ p^\top & 0 \end{bmatrix}, \tag{A.10}$$

where H is the upper 6×6 block of A, p is a column of length 6 containing all but the last element of the seventh column of A, and the zero in (A.10) is a scalar.

The matrix H is block-diagonal. I now show that each of its blocks is invertible, so that H^{-1} exists.[11] Specifically,

$$H = \begin{bmatrix} H_1 & 0 & 0 \\ 0 & H_2 & 0 \\ 0 & 0 & H_f \end{bmatrix} \quad \longrightarrow \quad H^{-1} = \begin{bmatrix} H_1^{-1} & 0 & 0 \\ 0 & H_2^{-1} & 0 \\ 0 & 0 & H_f^{-1} \end{bmatrix}, \tag{A.11}$$

where each block in the partition of H is 2×2 and

$$H_i^{-1} = \frac{1}{\beta_i \left(v_{c_i c_i} v_{x_i x_i} - v_{c_i x_i}^2 \right)} \begin{bmatrix} v_{x_i x_i} & -v_{c_i x_i} \\ -v_{c_i x_i} & v_{c_i c_i} \end{bmatrix} := \frac{1}{\Delta_i} \begin{bmatrix} v_{x_i x_i} & -v_{c_i x_i} \\ -v_{c_i x_i} & v_{c_i c_i} \end{bmatrix}, \quad i = 1, 2, \tag{A.12}$$

and

$$H_f^{-1} = \frac{1}{\mu \left(f_{yy} f_{kk} - f_{ky}^2 \right)} \begin{bmatrix} f_{kk} & -\mu f_{ky} \\ -f_{ky} & \mu f_{yy} \end{bmatrix} := \frac{1}{\Delta_f} \begin{bmatrix} f_{kk} & -\mu f_{ky} \\ -f_{ky} & \mu f_{yy} \end{bmatrix}. \tag{A.13}$$

Strict concavity of v and f imply that $\Delta_i > 0$, $i = 1, 2, f$. Indeed, the curvature properties imply that H is negative-definite.

[11]The calculations presented here are more than the minimum required to prove the Proposition. However, they are needed later on.

It is straightforward to check that[12]

$$A^{-1} = \begin{bmatrix} H^{-1} - \theta H^{-1} p p^\top H^{-1} & \theta H^{-1} p \\ \theta p^\top H^{-1} & -\theta \end{bmatrix}, \tag{A.14}$$

where

$$\theta = \frac{1}{p^\top H^{-1} p}. \tag{A.15}$$

Incidentally, because H is negative-definite, so is H^{-1}; therefore, $\theta < 0$. $\square$

Proof of Corollary 1. Using the bottom line of (A.14), (28) and (30) yields

$$\frac{d\mu}{dn} = \theta \begin{bmatrix} -1 & -(1+n)^{-1} & -1 & -(1+n)^{-1} & f_y & 0 \end{bmatrix} H^{-1} \begin{bmatrix} 0 \\ -\frac{\mu}{(1+n)^2} \\ 0 \\ -\frac{\mu}{(1+n)^2} \\ 0 \\ 1 \end{bmatrix} - \theta \left[k - (1+n)^{-2} x \right]. \tag{A.16}$$

Substituting (A.11)–(A.13) into (A.16) and performing the matrix multiplication gives (31).

Normality of x implies that the terms inside the summation sign on the right-hand side of (31) are negative. Linear homogeneity of f implies that f_y is homogeneous of degree zero. Hence, by Euler's Theorem

$$y f_{yy} + k f_{ky} = 0. \tag{A.17}$$

But $f_{yy} < 0$, so $f_{ky} > 0$. Hence, the entire expression inside the square bracket is negative when x is normal. Because $\theta < 0$, the first term is positive. Clearly, $(1+n)^2 k > x$ is sufficient for the final term to be positive as well. $\square$

[12] See Intriligator (1971, p. 158) for an analogous calculation in the context of consumer theory.

Proof of Corollaries 2–4. It is possible to use equations (A.11)–(A.14) to directly compute the results presented in Corollaries 2–4. However, it is instructive to use a more heuristic solution method. The top six lines of (A.9) can be written

$$H \begin{bmatrix} dc_1 \\ dx_1 \\ dc_2 \\ dx_2 \\ dy \\ dk \end{bmatrix} = \begin{bmatrix} d\mu \\ (1+n)^{-1}d\mu - (1+n)^{-2}\mu dn \\ d\mu \\ (1+n)^{-1}d\mu - (1+n)^{-2}\mu dn \\ -f_y d\mu \\ dn \end{bmatrix}. \tag{A.18}$$

Given the block-diagonal structure of H, (A.18) can be decomposed into the following three matrix equations,

$$\begin{aligned} H_i \begin{bmatrix} dc_i \\ dx_i \end{bmatrix} &= \begin{bmatrix} d\mu \\ (1+n)^{-1}d\mu - (1+n)^{-2}\mu dn \end{bmatrix}, \quad i = 1, 2; \\ H_f \begin{bmatrix} dy \\ dk \end{bmatrix} &= \begin{bmatrix} -f_y d\mu \\ dn \end{bmatrix}. \end{aligned} \tag{A.19}$$

Using (A.12) and (A.13) to compute the solutions to the equations (A.19) gives

$$\begin{bmatrix} dc_i \\ dx_i \end{bmatrix} = \frac{1}{\Delta_i} \begin{bmatrix} v_{x_i x_i} d\mu - (1+n)^{-1} v_{c_i x_i} d\mu + (1+n)^{-2} v_{c_i x_i} \mu dn \\ -v_{c_i x_i} d\mu + (1+n)^{-1} v_{c_i c_i} d\mu - (1+n)^{-2} v_{c_i c_i} \mu dn \end{bmatrix} \tag{A.20}$$

and

$$\begin{bmatrix} dy \\ dk \end{bmatrix} = \frac{1}{\Delta_f} \begin{bmatrix} -f_{kk} f_y d\mu - \mu f_{ky} dn \\ f_{ky} f_y d\mu + \mu f_{yy} dn \end{bmatrix}. \tag{A.21}$$

Equations (32)–(35) follow from "dividing" the appropriate entries in (A.20) and (A.21) through by dn.

The final sentence of Corollary 2 is immediate. □

References

Apps, P., Rees, R., 2006. Repeated optimal nonlinear income taxation, unpublished manuscript, University of Sydney and University of Munich.

Berliant, M., Ledyard, J. O., 2005. Optimal dynamic nonlinear income taxes with no commitment, downloadable from the Working Paper Archive at Washington University at St. Louis, `http://econpapers.repec.org/paper/wpawuwppe/0403004.htm`.

Boadway, R., Cuff, K., Marchand, M., 2000. Optimal income taxation with quasi-linear preferences revisited. Journal of Public Economic Theory 2, 435–460.

Boadway, R., Pestieau, P., 2007. Tagging and redistributive taxation. Annales d'Économie et de Statistique 83–84, 123–147.

Brett, C., Weymark, J. A., 2008a. The impact of changing skill levels on optimal nonlinear income taxes. Journal of Public Economics, forthcoming.

Brett, C., Weymark, J. A., 2008b. Optimal nonlinear taxation of income and savings without commitment. Discussion paper 08-W05, Vanderbilt University.

Brett, C., Weymark, J. A., 2008c. Public good provision and the comparative statics of optimal nonlinear income taxation. International Economic Review 49, 255–290.

Cutler, D. M., Poterba, J. M., Sheiner, L. M., Summers, L. H., 1990. An aging society: Opportunity or challenge? Brookings Papers on Economic Activity 1, 1–56.

Dillén, M., Lundholm, M., 1996. Dynamic income taxation, redistribution, and the ratchet effect. Journal of Public Economics 59, 69–93.

Hamilton, J., Pestieau, P., 2005. Optimal income taxation and the ability distribution: Implications for migration equilibria. International Tax and Public Finance, 29–45.

Intriligator, M. D., 1971. Mathematical Optimization and Economic Theory. Prentice-Hall, Englewood Cliffs.

McDaniel, S. A., 2003. Toward disentangling policy implications of economic and demographic changes in Canada's aging population. Canadian Public Policy 29, 491–509.

Meijdam, L., Verbon, H. A. A., 1997. Aging and public pensions in an overlapping-generations model. Oxford Economic Papers 49, 29–42.

Mirrlees, J. A., 1971. An exploration in the theory of optimum income taxation. Review of Economic Studies 38, 175–208.

Myles, G. D., 1995. Public Economics. Cambridge University Press, Cambridge.

Ordover, J., Phelps, E., 1979. The concept of optimal taxation in the overlapping–generations model of capital and wealth. Journal of Public Economics 12, 1–26.

Pirttilä, J., Tuomala, M., 2001. On optimal non-linear taxation and public good provision in an overlapping generations economy. Journal of Public Economics 79, 485–501.

Simula, L., 2007. Optimality conditions and comparative static properties of non-linear income taxes revisited, unpublished manuscript, Paris School of Economics.

Visco, I., 2001. The fiscal implications of ageing populations in oecd countries, Organization for Economic Cooperation and Development, presented at the Oxford Centre on Population Ageing Pensions Symposium, June.

Weymark, J. A., 1986. A reduced-form optimal nonlinear income tax problem. Journal of Public Economics 30, 199–217.

Weymark, J. A., 1987. Comparative static properties of optimal nonlinear income taxes. Econometrica 55, 1165–1185.