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 7, 2011

*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. Nevertheless, it is also shown that the optimal marginal income tax rates are independent of the rate of population growth. In addition, the steady state interest rate unambiguously declines when the rate of population growth declines. Resulting adjustments in production plans have 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. However, when the dependency effect is sufficiently strong, it is possible to sign the direction of change in all production and consumption plans. Moreover, regardless of the direction of change in optimal consumption plans, the absolute value of the changes in consumption plans are smaller for low-skilled workers than for high-skilled when utility is time-separable and preferences exhibit decreasing absolute risk aversion. Adopting, instead, a quasilinear-in-consumption specification of preferences sharpens the comparative statics of consumption allocations, but introduces ambiguity into the effect of the rate of population growth on the optimal marginal income tax rate.

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. One important ingredient of this understanding is the question of how population aging might influence optimal income tax schemes.

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) show that an aging population leads to a reduction in steady state consumption in the presence of public pension schemes funded by non-distorting taxes.

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

---

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

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, 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 (2010), and Brett and Weymark (2008a,c). 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.

---

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

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 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. Moreover, the individuals of the the lower skill-type face smaller decreases in consumption than do their higher-skill type counterparts when preferences exhibit decreasing absolute risk aversion. However, 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 and consumption when working increases under the additional assumption of time-separable utility. In this case, individuals of the lower type have a smaller increase in consumption than do the more highly skilled workers when preferences exhibit decreasing absolute risk aversion.

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 details the effects of population aging on the optimal tax schedule. Section 5 contains two elaborations. First, the results are reconsidered under the the additional assumption of time-separable utility. Second, a quasilinear-in-consumption version of the model is sketched in order to illustrate how the choice of functional form affects the some of the results. Some concluding remarks are then offered. Some of the more technical results and the proofs of all the results 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.

---

[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.

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 coordinate axes, for all positive levels of output. The prices of inputs are determined by the profit-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

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

By assuming an aggregate intertemporal resource constraint, I am allowing the government to engage in unrestricted debt and/or pension policies to move resources across time. Because the focus of this paper is on optimal income taxes, the exact form of these policies is not specified.

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

---

[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.

$\mathbb{R}^2_{++}$ 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.

In static economies, the assumption of preferences that are quasi-linear in leisure is well-known to place restrictions on the optimal nonlinear tax schedule. Weymark (1987) and Brett and Weymark (2008a) find 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 quasilinear 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. Nevertheless, given the current state of the literature, it is not possible to carry out comparative static exercises in optimal nonlinear taxation problems without recourse to assumptions of quasilinearity.

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)

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

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 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

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

---

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

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).

The quasilinear 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 1. *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]

---

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

Lemma 1 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]

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.

$$\beta_i v_{c_i} - \mu = 0, \quad i = 1, 2; \tag{19}$$

$$\beta_i v_{x_i} - \frac{\mu}{1+n} = 0, \quad i = 1, 2; \tag{20}$$

$$-1 + \mu f_y = 0; \tag{21}$$

$$f_k - n = 0. \tag{22}$$

In fact, the solution is completely described by the necessary conditions (19)–(22) and the resource constraint (11). It follows directly from (21) that $\tilde{\mu} > 0$. Moreover, the qualitative properties of the optimal tax system, including its implied behavioral distortions, can be derived from equations (19)–(22). 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,*

$$\frac{v_c(\tilde{c}_1, \tilde{x}_1)}{v_x(\tilde{c}_1, \tilde{x}_1)} = \frac{v_c(\tilde{c}_2, \tilde{x}_2)}{v_x(\tilde{c}_2, \tilde{x}_2)} = (1+n). \tag{23}$$

---

[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.

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

$$MRS_{2,lc} := \frac{1}{v_c(\tilde{c}_2, \tilde{x_2})} = wa_2. \tag{24}$$

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

$$IMTR_1 := 1 - \frac{1}{wa_1 v_c(\tilde{c}_1, \tilde{x}_1)} = (\lambda_1 - 1)\left(\frac{a_2 - a_1}{a_1}\right). \tag{25}$$

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 (24) and (25) 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. Lemma A.1 and Proposition A.1, both in the Appendix, establish that it is possible to carry out this comparative static

---

[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.7) used in the proof of Proposition 1.

analysis. Moreover, Proposition A.1 makes it possible to derive explicit formulae for the responses of the optimal allocation to a change in rate of population growth. These expressions are contained in Results 1–4 below.

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

$$\frac{d\tilde{\mu}}{dn} = \theta \left[\sum_{i=1}^{2} \left\{ \frac{\tilde{\mu}}{\Delta_i (1+n)^2} \left(\frac{v_{c_i c_i}}{1+n} - v_{c_i x_i} \right) \right\} - \frac{\tilde{\mu}}{\Delta_f} f_y^2 f_{ky} \right] - \theta \left[\tilde{k} - \frac{\tilde{x}}{(1+n)^2} \right]. \tag{26}$$

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* $\dfrac{d\tilde{\mu}}{dn} > 0$.

Result 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 (26). 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 countervailing effects on the resource constraint. First, there is a capital spreading effect, as more workers arrive to work with the existing capital stock. This is captured by $(1+n)^2\tilde{k}$. Second, there is a reduced dependency effect, as the relative number of retirees falls. This is captured by $\tilde{x}$. Capital spreading reduces productivity per person, increasing the social marginal cost of consumption. Reduced dependency decreases the social marginal cost of consumption. 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 Result posits that the capital spreading effect is stronger than the dependency effect.[9]

---

[9] A few back-of-the-envelope calculations reveal that it really is an empirical question as to which of these terms would dominate. The capital stock is roughly three times GDP in Canada. A 2% annual growth rate, compounded over 20 years, results in $n \approx 0.5$. This results in a capital spreading effect of roughly 4.5 times GDP. Retirement-in-consumption is meant to capture the entire post-work period. If, say, 20% of the population is retired and the retirement period is 20 years, then $\tilde{x}$ might be four times the value of yearly consumption, which is somewhat less than four times GDP. Given the closeness of

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 (26).

It follows from (3) and (21) that $\tilde{\mu} = 1/\tilde{w}$. Hence, the shadow value of the resource constraint varies inversely with the aggregate wage rate at the optimum.

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 Result 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. Result 2 provides a formal summary of this discussion.

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

$$\frac{d\tilde{c}_i}{dn} = \frac{1}{\Delta_i}\left[\left((v_{x_i x_i} - \frac{v_{c_i x_i}}{1+n}\right)\frac{d\tilde{\mu}}{dn} + \frac{\tilde{\mu} v_{c_i x_i}}{(1+n)^2}\right], \quad i = 1, 2. \tag{27}$$

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

---

these “ballpark” figures, it is difficult to say which effects dominate for a particular economy without careful empricial investigation.

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.[10] Thus, it is impossible to sign the effect of an increase in n on x. Result 3 gives an algebraic rendering of the ambiguous effect of n on the optimal consumption in retirement.

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

$$\frac{d\tilde{x}_i}{dn} = \frac{1}{\Delta_i}\left[\left((-v_{c_i x_i} + \frac{v_{c_i c_i}}{1+n}\right)\frac{d\tilde{\mu}}{dn} - \frac{\tilde{\mu} v_{c_i c_i}}{(1+n)^2}\right], \quad i = 1, 2. \tag{28}$$

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. Result 4 displays the potentially offsetting terms.

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

$$\frac{d\tilde{y}}{dn} = \frac{1}{\Delta_f}\left[-f_{kk} f_y \frac{d\tilde{\mu}}{dn} - \tilde{\mu} f_{ky}\right], \tag{29}$$

and

$$\frac{d\tilde{k}}{dn} = \frac{1}{\Delta_f}\left[f_{ky} f_y \frac{d\tilde{\mu}}{dn} + \tilde{\mu} f_{yy}\right]. \tag{30}$$

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.[11] Without further restrictions on technology, it is impossible to tell which of these levers the taxation

---

[10] Formally, both the numerator and denominator in the final term on the left-hand side of (20) increase.

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

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.

5 Refinements and Extensions

5.1 Additive Utility

As Results 2 and 3 demonstrate, consumption adjustments are somewhat nuanced. However, they come into sharper focus when utility is additive across time, that is, when

$$v(c, x) = g(c) + h(x). \tag{31}$$

Table 1 collects the results of this analysis when utility has the form given by (31). 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.

When the dependency effect is strong, both aggregate inputs are used in greater quantity, so output increases. Yet, everyone's consumption decreases. This is a caricature of "apocalyptic demography": younger generations work harder to support more retirees, still everyone consumes less. In models with exogenous labor supply, only the consumption response comes into focus. Capital deepening exerts countervailing forces on both

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

consumption and input usage. The added capital per worker increases the marginal productivity of labor, thereby exerting an upward pressure on output and on wage. There is a resulting downward pressure on optimal labor usage. Capital deepening may be sufficiently strong to undo the effects of dependency on the optimal use of of labor and capital.

The respective consumptions of the two different types of individuals may change by different amounts in response to population aging. Corollary 1 shows that when utility is additive across time, it is possible to sign the relative magnitudes of the consumption responses of the two types of individuals.

Corollary 1. *When utility is additive across time and the functions g and h display decreasing (increasing) absolute risk aversion, c_1 and x_1 are less (more) responsive to changes in n than are c_2 and x_2. Specifically,*

(i) $\left|\dfrac{d\tilde{c}_1}{dn}\right| > (<) \left|\dfrac{d\tilde{c}_2}{dn}\right|$ *when g displays increasing (decreasing) absolute risk aversion in the neighborhood of the initial steady state.*

(ii) $\left|\dfrac{d\tilde{x}_1}{dn}\right| > (<) \left|\dfrac{d\tilde{x}_2}{dn}\right|$ *when h displays increasing (decreasing) absolute risk aversion i the neighborhood of the initial steady state.*

For concreteness, consider the situation in which the dependency effect dominates and the plausible case that the degree of absolute risk aversion decreases with consumption. According to Table 1, everyone's consumption falls in both periods of life. Corollary 1 states that the fall in consumption is of a higher magnitude for high-skilled individuals than is the corresponding fall in consumption for low-skilled individuals. In this case, population aging results in an increase in the shadow value of resources $\tilde{\mu}$, thereby increasing the marginal cost of consumption in the same proportion for each type of individual. The government optimally reduces consumption levels, thereby increasing the marginal values of the various type- and age-specific consumption plans. The lower is the degree of absolute risk aversion, the larger is the required change in consumption needed to bring about a fixed change in the marginal utility of consumption. High-skilled individuals consume more and, therefore, when absolute risk aversion is decreasing, have a lower degree of absolute risk aversion. Thus, the government has to reduce the consumption of high-skilled individuals more than the consumption of low-skilled individuals in order to re-establish the balance between the marginal benefit of consumption for each type of individual and the marginal cost of consumption.

5.2 Quasilinear-in-Consumption Utility

Adopting the quasilinear in consumption approach of Simula (2010) is problematic in dynamic settings because the optimal intertemporal allocation of consumption for an individual is indeterminate when utility is linear in consumption for each period. It is, however, possible to arrive at a tractable model when utility is quasilinear in consumption in one of the two periods.[12] The more *a priori* plausible of the two possibilities is a utility function of the form

$$U_i(c, x, y) = g\left(c, \frac{y}{a_i}\right) + x, \tag{32}$$

---

[12] This sub-section is meant to illustrate how some of the results depend on the form of the utility function. A full analysis of the alternative model is beyond the scope of this article.

where the function g assumed to be twice continuously differentiable at all $(c, y) \neq (0, 0)$, continuous on $\mathbb{R}^2_+$, nondecreasing in c and nonincreasing in y (or l) and these monotonicity properties are strict on $\mathbb{R}^2_{++}$. In addition, g is assumed to be strictly concave on $\mathbb{R}^2_{++}$ with $g(0, 0) = 0$, $g_c(0, x) = \infty$ for all $x > 0$, and $g_c(c, x) \to 0$ as $c \to \infty$ for all $l \geq 0$. These limiting assumptions ensure that c is always positive at an optimum.

When preferences are of the form (32), it is possible to solve the budget constraint (11) and the binding self-selection constraint (14) for x_1 and x_2 in terms of the other variables in the government's optimization problem. Substituting these expressions into the objective function yields the following result.

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 with quasilinear-in-consumption preferences (32). Then $(\tilde{c}_1, \tilde{c}_2, \tilde{y}_1, \tilde{y}_2, \tilde{k})$ solves:*

$$\begin{aligned} \max_{(c_1,c_2,y_1,y_2,k)} \lambda_1 g\left(c_1, \frac{y_1}{a_1}\right) + (1 - \lambda_1) g\left(c_1, \frac{y_1}{a_2}\right) + g\left(c_2, \frac{y_2}{a_2}\right) \\ + (1 + n)\left[f(k, y - 1 + y_2) - nk - c_1 - c_2\right]. \end{aligned} \tag{33}$$

The reduced-form optimization problem (33) is, on the surface, much simpler in structure than the one that obtains with quasilinear in leisure preferences — or even the reduced forms obtained in Weymark (1986) or Simula (2010). This simplicity arises from the ability to completely eliminate consumption when old from the reduced form. Unlike previous reduced form problems, however, it is not formulated solely in terms of the utilities of the agents in the economy. The first-period utility of "mimickers" appears in the the second term of (33). Because $\lambda_1 > 1$, increases a mimicker's utility reduces social welfare. This is the vehicle through which (33) encodes the social cost of tightening the self-selection constraint. The factor $(1 + n)$ in the final term is the opportunity cost of resources consumed in the first period of life in terms consumption in retirement, which is also the units of utility measurement for the quasilinear form (32).

The qualitative features of the solution can be anticipated from the literature, especially the work of Ordover and Phelps (1979). As in the model with quasilinear-in-leisure preferences, there is no reason to distort production decisions in this model, because relative wages are fixed. There are, of course, no distortions at the top, either. Savings taxes are justified by differences in the marginal rate of intertemporal substitution between individuals of different types. Individuals of type 1 and their potential mimickers differ in their labor supply only, with the mimickers needing to supply less labor to achieve y_1 than do individuals of type 1. If, for example, g_{cl} is positive, then mimickers have a lower marginal utility of consumption when young than do agents of type 1. As a result, the government can deter mimicking by subsidizing consumption when young or, equivalently, by taxing savings. Finally, there is a positive marginal income tax rate at the bottom of the distribution.

This version of the model does not give rise to an explicit formula for the optimal marginal tax rate in terms of model parameters alone. For this reason, it is not possible to provide unambiguous statements concerning how the marginal rate of income taxation responds to changes in n. On the other hand, the comparative static results for consumption are somewhat sharper in this formulation of the model. This is especially so when preferences are further specialized to an additive form

$$U_i(c, x, y) = u(c) - h\left(\frac{y}{a_i}\right) + x. \tag{34}$$

With this additional assumption, the following result holds.

Proposition 2. *When preferences are quasilinear in consumption when old and additively separable as in (34) the consumption when young of each individual falls as* n *increases.*

Proposition 2 holds because, according to (33), the opportunity cost of consumption when young is increasing in n. Put differently, the more young people there are, the more units of the numeraire have to be sacrificed to increase the consumption of each of them. This opportunity cost effect leads to the natural substitution away from the more costly form of consumption. Despite its obvious economic interpretation, this result might cut against the grain of folk wisdom. It says: an aging population — at least in the sense defined in this paper — induces an increase in the per capita consumption of the young.

6 Concluding Remarks

The model introduced here as two features that are required for any discussion of nonlinear income taxation, but not universal in the literature on policy choices with overlapping generations. First, labor is supplied elastically. Thus, if output adjustment is required in response to demographic change, then both labor and capital adjustments are possible. Indeed, much of the ambiguity in the signs of the responses of input usage to population aging is due to the ability of the planner to adjust both labor and capital rather than just capital alone. Second, there is heterogeneity of workers within each generation. Thus, it is possible to answer questions about the within-generation incidence of adjustments to population aging, at least under the additional assumptions presented in Section 5.1.

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. When utility is additive across time, it is also possible to tell which type of worker faces the larger adjustments in consumption when demographics change. Moreover, the structure of the matrix A presented in Proposition A.1 makes it apparent how to generalize the results for quasilinear-in-leisure preferences to an arbitrary, finite number of skill-types.

It is striking that, for some specifications of preferences, the implicit marginal income tax rate faced by low-skill workers is invariant to the rate of population growth. One may interpret this result as stating that it is not necessarily the case that population aging forces the government to rely on more (or less) distortionary forms of taxation. Also striking is the frailty of the notion that a decrease in the population growth rate, and the resulting increasing concentration of retirees in the economy, causes consumption of the working to decrease. This outcome is possible in the model economy presented in this article, but by no means assured. Moreover, it is impossible in the (admittedly special) time-separable and quasilinear-in-consumption variant of the model.

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. This technical challenge may be sidestepped by adopting the quasilinear-in-consumption model, but at the cost of more cumbersome expressions for optimal marginal tax rates. 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

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 A.1 establishes that this is so for the Steady State Optimal Nonlinear Income Tax Problem.

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

Proof of Lemma A.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 $(c_1, c_2, x_1, x_2, k, y_1 + y_2)$. Using the binding self-selection constraint (14) and $y = y_1 + y_2$, it is straightforward 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 1. 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 Proposition 1. Part (i) follows directly from equations (3) and (22). Part (ii) follows from dividing (19) by (20) for individuals of each type. By (21),

$$\mu = \frac{1}{f_y}. \tag{A.5}$$

Part (iii) follows from substituting (A.5) and (18) into (19) for individuals of type 2 and rearranging.

Using (3) in conjunction with the definition of $IMTR_1$ given in (25) yields

$$IMTR_1 = 1 - \frac{1}{a_1 f_y v_c(\tilde{c}_1, \tilde{x}_1)}. \tag{A.6}$$

Substituting (A.5) and (19) into (A.6) yields

$$IMTR_1 = 1 - \frac{1}{\frac{a_1}{\beta_1}} = \frac{a_1 - \beta_1}{a_1}. \tag{A.7}$$

Recalling the definition of β_1 from (17) and rearranging yields (25). □

Proposition A.1. *The optimality conditions (11) and (19)–(22) define a continuously differentiable solution function* $F\colon \mathbb{R}_+ \to \mathbb{R}^7_{++}$ *of the problem (16) with* $n \mapsto (\tilde{c}_1, \tilde{x}_1, \tilde{c}_2, \tilde{x}_2, \tilde{y}, \tilde{k}, \tilde{\mu})$*. For all* $n \in \mathbb{R}_+$*, the derivative* DF *of* F *at* n *is given by*

$$DF(n) = (A^{-1}b)(n), \tag{A.8}$$

where

$$A(n) = \begin{bmatrix} \beta_1 v_{c_1c_1} & \beta_1 v_{c_1x_1} & 0 & 0 & 0 & 0 & -1 \\ \beta_1 v_{c_1x_1} & \beta_1 v_{x_1x_1} & 0 & 0 & 0 & 0 & -(1+n)^{-1} \\ 0 & 0 & \beta_2 v_{c_2c_2} & \beta_2 v_{c_2x_2} & 0 & 0 & -1 \\ 0 & 0 & \beta_2 v_{c_2x_2} & \beta_2 v_{x_2x_2} & 0 & 0 & -(1+n)^{-1} \\ 0 & 0 & 0 & 0 & \mu f_{yy} & \mu f_{ky} & f_y \\ 0 & 0 & 0 & 0 & f_{ky} & f_{kk} & 0 \\ -1 & -(1+n)^{-1} & -1 & -(1+n)^{-1} & f_y & 0 & 0 \end{bmatrix} \tag{A.9}$$

and

$$b(n) = \begin{bmatrix} 0 \\ -(1+n)^{-2}\tilde{\mu} \\ 0 \\ -(1+n)^{-2}\tilde{\mu} \\ 0 \\ 1 \\ \tilde{k} - (1+n)^{-2}\tilde{x} \end{bmatrix}, \tag{A.10}$$

and where all expressions on the right-hand sides of (A.9) and (A.10) are evaluated at the solution to (16).

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

$$A \begin{bmatrix} dc_1 & dx_1 & dc_2 & dx_2 & dy & dk & d\mu \end{bmatrix}^\top = b\, dn, \tag{A.11}$$

where dependence on the parameter n is now expunged from the notation. The first zero in the final line of (A.9) follows from (22). In order to establish the Proposition, it suffices to show that the matrix A is invertible. To that end, introduce the partition

$$A = \begin{bmatrix} H & p \\ p^\top & 0 \end{bmatrix}, \tag{A.12}$$

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.12) is a scalar.

The matrix H is block-diagonal. I now show that each of its blocks is invertible, so that H^{-1} exists.[13] Specifically,

$$H = \begin{bmatrix} H_1 & 0 & 0 \\ 0 & H_2 & 0 \\ 0 & 0 & H_f \end{bmatrix} \quad \longrightarrow \quad H^{-1} = \begin{bmatrix} H_1^{-1} & 0 & 0 \\ 0 & H_2^{-1} & 0 \\ 0 & 0 & H_f^{-1} \end{bmatrix}, \tag{A.13}$$

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

$$H_i^{-1} = \frac{1}{\beta_i \left(v_{c_i c_i} v_{x_i x_i} - v_{c_i x_i}^2 \right)} \begin{bmatrix} v_{x_i x_i} & -v_{c_i x_i} \\ -v_{c_i x_i} & v_{c_i c_i} \end{bmatrix} := \frac{1}{\Delta_i} \begin{bmatrix} v_{x_i x_i} & -v_{c_i x_i} \\ -v_{c_i x_i} & v_{c_i c_i} \end{bmatrix}, \quad i = 1, 2, \tag{A.14}$$

and

$$H_f^{-1} = \frac{1}{\mu \left(f_{yy} f_{kk} - f_{ky}^2 \right)} \begin{bmatrix} f_{kk} & -\mu f_{ky} \\ -f_{ky} & \mu f_{yy} \end{bmatrix} := \frac{1}{\Delta_f} \begin{bmatrix} f_{kk} & -\mu f_{ky} \\ -f_{ky} & \mu f_{yy} \end{bmatrix}. \tag{A.15}$$

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.

---

[13] 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[14]

$$A^{-1} = \begin{bmatrix} H^{-1} - \theta H^{-1} p p^\top H^{-1} & \theta H^{-1} p \\ \theta p^\top H^{-1} & -\theta \end{bmatrix}, \tag{A.16}$$

where

$$\theta = \frac{1}{p^\top H^{-1} p}. \tag{A.17}$$

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

Proof of Result 1. Using the bottom line of (A.16), (A.8) and (A.10) yields

$$\begin{aligned} \frac{d\mu}{dn} = \theta \begin{bmatrix} -1 & -(1+n)^{-1} & -1 & -(1+n)^{-1} & f_y & 0 \end{bmatrix} H^{-1} \begin{bmatrix} 0 \\ -\frac{\mu}{(1+n)^2} \\ 0 \\ -\frac{\mu}{(1+n)^2} \\ 0 \\ 1 \end{bmatrix} \\ - \theta \left[k - (1+n)^{-2} x \right]. \end{aligned} \tag{A.18}$$

Substituting (A.13)–(A.15) into (A.18) and performing the matrix multiplication gives (26).

Normality of x implies that the terms inside the summation sign on the right-hand side of (26) are negative. Linear homogeneity of f implies that f_y is homogeneous of degree zero. Hence, by Euler's Theorem

$$y f_{yy} + k f_{ky} = 0. \tag{A.19}$$

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. □

---

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

Proof of Results 2–4. It is possible to use equations (A.13)–(A.16) 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.11) can be written

$$H \begin{bmatrix} dc_1 \\ dx_1 \\ dc_2 \\ dx_2 \\ dy \\ dk \end{bmatrix} = \begin{bmatrix} d\mu \\ (1+n)^{-1}d\mu - (1+n)^{-2}\mu dn \\ d\mu \\ (1+n)^{-1}d\mu - (1+n)^{-2}\mu dn \\ -f_y d\mu \\ dn \end{bmatrix}. \tag{A.20}$$

Given the block-diagonal structure of H, (A.20) can be decomposed into the following three matrix equations,

$$\begin{aligned} H_i \begin{bmatrix} dc_i \\ dx_i \end{bmatrix} &= \begin{bmatrix} d\mu \\ (1+n)^{-1}d\mu - (1+n)^{-2}\mu dn \end{bmatrix}, \quad i = 1, 2; \\ H_f \begin{bmatrix} dy \\ dk \end{bmatrix} &= \begin{bmatrix} -f_y d\mu \\ dn \end{bmatrix}. \end{aligned} \tag{A.21}$$

Using (A.14) and (A.15) to compute the solutions to the equations (A.21) gives

$$\begin{bmatrix} dc_i \\ dx_i \end{bmatrix} = \frac{1}{\Delta_i} \begin{bmatrix} v_{x_i x_i} d\mu - (1+n)^{-1} v_{c_i x_i} d\mu + (1+n)^{-2} v_{c_i x_i} \mu dn \\ -v_{c_i x_i} d\mu + (1+n)^{-1} v_{c_i c_i} d\mu - (1+n)^{-2} v_{c_i c_i} \mu dn \end{bmatrix} \tag{A.22}$$

and

$$\begin{bmatrix} dy \\ dk \end{bmatrix} = \frac{1}{\Delta_f} \begin{bmatrix} -f_{kk} f_y d\mu - \mu f_{ky} dn \\ f_{ky} f_y d\mu + \mu f_{yy} dn \end{bmatrix}. \tag{A.23}$$

Equations (27)–(30) follow from “dividing” the appropriate entries in (A.22) and (A.23) through by dn.

The final sentence of Result 2 is immediate. □

Proof of Corollary 1. When utility is of the form (31), (27) can be written

$$\frac{d\tilde{c}_i}{dn} = \frac{h''(\tilde{x}_i)}{\Delta_i} \frac{d\tilde{\mu}}{dn}, \quad = 1, 2. \tag{A.24}$$

Also, (31) and (A.14) imply

$$\Delta_i = \beta_i g''(\tilde{c}_i) h''(\tilde{x}_i), \quad i = 1, 2. \tag{A.25}$$

Moreover, when (31) holds, the first order condition (19) becomes

$$\beta_i g'(\tilde{c}_i) = \tilde{\mu}, \quad i = 1, 2. \tag{A.26}$$

Substituting (A.25) and (A.26) into (A.24) and simplifying yields

$$\frac{d\tilde{c}_i}{dn} = \frac{d\tilde{\mu}}{dn} \frac{1}{\tilde{\mu}} \frac{g'(\tilde{c}_i)}{g''(\tilde{c}_i)}, \quad i = 1, 2. \tag{A.27}$$

Thus,

$$\left| \frac{d\tilde{c}_1}{dn} \right| \geq \left| \frac{d\tilde{c}_2}{dn} \right| \quad \longleftrightarrow \quad \left| \frac{g'(\tilde{c}_1)}{g''(\tilde{c}_1)} \right| \geq \left| \frac{g'(\tilde{c}_2)}{g''(\tilde{c}_2)} \right| \quad \longleftrightarrow \quad -\frac{g''(\tilde{c}_1)}{g'(\tilde{c}_1)} \leq \frac{g''(\tilde{c}_2)}{g'(\tilde{c}_2)}. \tag{A.28}$$

Because $\beta_2 > \beta_1$, it follows from (A.26) on strict concavity of g that $\tilde{c}_1 < \tilde{c}_2$. Thus, part (i) of the Corollary follows from (A.28).

The proof of part (ii) of the Corollary is analogous. □

Proof of Lemma 2. The steady state resource constraint (11) and the self-selection constraint (14) imply

$$\begin{aligned} (1+n)[f(k,y) - nk - c_1 - c_2] &= x_1 + x_2; \\ g\left(c_2, \frac{y_2}{a_2}\right) - g\left(c_1, \frac{y_1}{a_2}\right) &= x_1 - x_2. \end{aligned} \tag{A.29}$$

Solving the system (A.29) for x_1 and x_2 yields:

$$\begin{aligned} x_1 &= \frac{1}{2}\left[g\left(c_2, \frac{y_2}{a_2}\right) - g\left(c_1, \frac{y_1}{a_2}\right)\right] + \frac{(1+n)}{2}[f(k,y) - nk - c_1 - c_2]; \\ x_2 &= -\frac{1}{2}\left[g\left(c_2, \frac{y_2}{a_2}\right) - g\left(c_1, \frac{y_1}{a_2}\right)\right] + \frac{(1+n)}{2}[f(k,y) - nk - c_1 - c_2]. \end{aligned} \tag{A.30}$$

Social welfare is given by

$$\mathcal{W} = \lambda_1 \left[g\left(c_1, \frac{y_1}{a_1}\right) + x_1\right] + \lambda_2 \left[g\left(c_2, \frac{y_2}{a_2}\right) + x_2\right]. \tag{A.31}$$

Substituting (A.30) into (A.31) and simplifying yields

$$\begin{aligned}\mathcal{W} =&\lambda_1 g\left(c_1, \frac{y_1}{a_1}\right) + \left(\lambda_2 + \frac{\lambda_1 - \lambda_2}{2}\right) g\left(c_2, \frac{y_2}{a_2}\right) + \left(\frac{\lambda_2 - \lambda_1}{2}\right) g\left(c_1, \frac{y_1}{a_2}\right) \\ &+ (1+n)[f(k,y) - nk - c_1 - c_2].\end{aligned} \tag{A.32}$$

Applying the normalization $\lambda_1 + \lambda_2 = 2$ to (A.32) yields the objective function in (33). □

Proof of Proposition 2. When preferences have the form (34), the objective function in (33) becomes

$$\mathcal{W} = u(c_1) + u(c_2) - \lambda_1 h\left(\frac{y_1}{a_1}\right) - (1-\lambda_1) h\left(\frac{y_1}{a_2}\right) - h\left(\frac{y_2}{a_2}\right) + (1+n)[f(k,y) - nk - c_1 - c_2]. \tag{A.33}$$

The first order condition for the choice of c_i is, therefore,

$$u'(c_i) = 1 + n. \tag{A.34}$$

The Proposition follows immediately from applying the Implicit Function Theorem to (A.34). □

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 92, 1765–1771.

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 12, 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., 2010. Optimal nonlinear income tax and nonlinear pricing: Optimality conditions and comparative static properties. Social Choice and Welfare 35, 199–220.

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.