Present-Biased Preferences and Publicly Provided Private Goods*

Thomas Aronsson and David Granlund

Submitted April 17, 2013; re-sumbitted September 29, 2013; final submission, November 7, 2013

Post-refereeing-version. The final version is available at FinanzArchiv/Public Finance Analysis 70(2), 169-199, 2014 via http://dx.doi.org/10.1628/001522114X681333.

Abstract

This paper analyzes the welfare effects of a publicly provided private good with long-term consequences for individual well-being, in an economy where consumers have "present-biased" preferences due to quasi-hyperbolic discounting. The analysis is based on a two-type model with asymmetric information between the government and the private sector, and each consumer lives for three periods. We present formal conditions under which public provision to the young and middle-aged generation,

*Aronsson: Department of Economics, Umeå School of Business and Economics, Umeå University, SE-901 87 Umeå, Sweden (thomas.aronsson@econ.umu.se); Granlund: Department of Economics, Umeå School of Business and Economics, Umeå University, SE-901 87 Umeå, Sweden, and HUI Research, SE-103 29 Stockholm, Sweden (david.granlund@econ.umu.se). The authors would like to thank the editor Ronnie Schöb, two anonymous referees, and Tomas Sjögren for helpful comments and suggestions. Research grants from the Bank of Sweden Tercentenary Foundation, the Swedish Council for Working Life and Social Research, and the Swedish Tax Agency are also gratefully acknowledged.

respectively, leads to higher welfare. Our results show that quasi-hyperbolic discounting provides a strong incentive for public provision to the young generation; especially if the consumers are naive (instead of sophisticated).
Keywords: Public provision of private goods, health care, hyperbolic discounting, intertemporal model, asymmetric information.
JEL classification: D03, D61, H42, I18.

1 Introduction

There is now a considerable amount of research based on experiments suggesting that consumers make dynamically inconsistent choices. The underlying behavioral failure is a self-control problem caused by "present-biased" preferences, i.e. a tendency for the individual to give less weight to the future welfare consequences of today's actions than would be optimal for the individual himself/herself in a longer time-perspective. A mechanism that generates this behavior is quasi-hyperbolic discounting, where the individual, at any time t, attaches a higher utility discount rate to tradeoffs between periods t and $t+1$ than to similar tradeoffs in the more distant future.[1] The resulting self-control problem might be exemplified by a tendency to undersave or underinvest in health capital or human capital; all of which may have serious welfare consequences.

The present paper develops a dynamic general equilibrium model, where each consumer lives for three periods and suffers from a self-control problem generated by quasi-hyperbolic discounting. The purpose is to analyze the welfare consequences of a publicly provided private investment good with long-term consequences for individual well-being, exemplified by health care services. However, health care services are also interesting in their own right, as governments devote much resources to the provision of health care in many countries. Since some of the benefits to the individual of such investments are likely to arise in the future (in the form of increased health capital), whereas the costs arise at the time

[1] Experimental evidence pointing in this direction can be found in, e.g., Thaler (1981), Kirby and Marakovic (1997), Kirby (1997), Viscusi, Huber and Bell (2008) and Brown, Chua and Camerer (2009). In the latter two studies, estimates of the "hyperbolic parameter" (referred to as "β" below) are in the interval $0.5 - 0.8$ (instead of 1 as under exponential discounting). See also Fredrick, Loewenstein and O'Donoghue (2002) for a review of empirical research on intertemporal choice, and Rubinstein (2003) for a critical view of the evidence for hyperbolic discounting.

the investment is made, quasi-hyperbolic discounting is likely to imply that the investment made by the individual becomes too small from his/her own long-term perspective.[2] Therefore, one would normally expect public provision to be welfare improving; yet, the results presented below show that this is not necessarily the case. One important reason is that the consumers may decrease both their current and future private purchases in response to an anticipated future increase in the public provision. The intertemporal adjustments further imply that the welfare effects of public provisions to different generations interact.

As emphasized in earlier research, publicly provided private goods are important tools for redistribution. Indeed, and based on model-economies where the consumers are fully rational, Blomquist and Christiansen (1995, 1998) and Boadway and Marchand (1995) show how the welfare effects of public provision depend on whether such policies facilitate or hinder redistribution. Therefore, to examine how self-control problems affect the usefulness of publicly provided private goods, it is vital that corrective and redistributive aspects of public policy are addressed simultaneously. We will follow this earlier literature in assuming that individual ability (productivity) is private information, whereas income is observable and can be used as a basis for revenue collection and redistribution subject to an incentive (self-selection) constraint. As such, our model is an extension of

[2] Although in contexts different from ours, empirical evidence shows that time-inconsistent preferences may have an influence on health decisions at the individual level; for instance, that smokers are more prone to hyperbolic discounting than non-smokers (Bickel et al., 1999; Odum et al., 2002); that food purchase (measured in terms of calories) among food stamp recepients is consistent with hyperbolic discounting (Shapiro, 2005); and that women's preferences for medical treatment during child-birth may undergo reverals during the birth-process in a way consistent with hyperbolic discounting (Christensen-Szalanski, 1984). Also, Pol and Cairns (2002) find evidence for hyperbolic discounting with respect to health outcomes based on a stated preference approach.

the two-type model originally developed by Stern (1982) and Stiglitz (1982), which is here extended to allow for a time-inconsistent preference for immediate gratification through quasi-hyperbolic discounting. Assuming that the tax revenue is raised through optimal nonlinear taxes on labor income and capital income, our results show that the corrective and redistributive motives for public provision are aligned if low-income consumers (either the low-ability type or high-ability agents mimicking the low-ability type) purchase less health care than high-income consumers.

Hyperbolic discounters are often described in terms of naivety or sophistication.[3] A naive consumer does not realize that future selves will be subject to the same self-control problem, i.e., expects the time-inconsistent preference to vanish in the future. As such, the naive consumer is characterized by time-inconsistent behavior in the sense that he/she may revise the savings-investment plan in each subsequent period. A sophisticated consumer, on the other hand, realizes that the same self-control problem also arises in future periods and will, therefore, act strategically vis-a-vis his/her future selves. Since the distinction between naivety and sophistication matters for public policy, i.e., the policy rule for public provision under naivety differs from the corresponding policy rule under sophistication, we consider both cases below.[4]

The paper also relates to the literature on tax policy (or subsidy) responses, as well

[3]For a more thorough discussion, see, e.g., O'Donoghue and Rabin (1999).

[4]To the best of our knowledge, the empirical evidence gives no clear guidance here. Based on experimental data, Hey and Lotito (2009) find that the majority of agents were either naive or resolute (where resolute means that agents stick to the plan preferred ex-ante), whereas sophistication was a less common strategy. Evidence in favor of naivety is also presented in, e.g., DellaVigna and Malmendier (2006) and Skiba and Tobacman (2008). See also Gine, Karlan and Zinman (2010) for behavioral patterns consistent with sophistication. The review by DellaVigna (2009) shows behavioral patterns consistent with both naivety and sophistication.

as responses in terms of public goods, to quasi-hyperbolic discounting (e.g., Gruber and Köszegi, 2004; O'Donoghue and Rabin, 2003, 2006; Aronsson and Thunström, 2008; Aronsson and Sjögren, 2009; Aronsson and Granlund, 2011). To our knowledge, however, there are no earlier studies on public provision of private goods in this particular context. Our study serves to bridge this gap by considering the supplemental role of publicly provided private goods when the income tax is optimal. Pirttilä and Tenhunen (2008) also address public provision of private goods under optimal income taxation in an economy where agents suffer from bounded rationality. However, their study is based on a static model combined with a "non-welfarist" approach, where the objective function of the government differs from that faced by the consumers. As such, their analysis is neither able to address the intertemporal adjustment effect described above nor the distinction between naivety and sophistication. In Section 3 below, we compare our results with those derived by Pirttilä and Tenhunen.

Our study is closely related to a paper by Aronsson and Sjögren (2009), which deals with optimal mixed taxation under asymmetric information, in an economy where the consumers suffer from the same kind of self-control problem as in the present study. Therefore, as the implications of quasi-hyperbolic discounting for optimal taxation are analyzed at some length in their study, we focus on public provision here. The outline of the study is as follows. Section 2 presents the model and characterizes the outcome of private optimization. In Section 3, we present the cost benefit rules for public provision of private goods to the young and middle-aged generation, respectively, as well as relate these policy rules to whether the consumers are characterized by naivety or sophistication. The results are summarized and discussed in Section 4.

2 The Model

To simplify the analysis as much as possible, the production side of the model is represented by a linear technology. This means that the producer prices and factor prices (before-tax hourly wage rates and interest rate) are fixed in each period,[5] although not necessarily constant over time.

Turning to the consumption side, we assume that each consumer lives for three periods: works in the first and second, and is retired in the third. At least three periods are required to model the time-inconsistent preference that quasi-hyperbolic discounting gives rise to.[6] The consumers differ with respect to productivity and can be divided in two ability-types: a low-ability type (denoted by superindex 1) earning wage rate w^1 and a high-ability type (denoted by superindex 2) earning wage rate $w^2 > w^1$. For simplicity, we abstract from population growth and normalize the number of consumers of each ability-type and generation to one. The instantaneous utility functions facing ability-type i ($i = 1, 2$) of generation t - who is young in period t, middle-aged in period $t+1$ and old in period $t+2$ - can be written as

[5] Similar assumptions have been used in some of the previous studies referred to above (e.g., Blomquist and Christiansen, 1995, 1998; Pirttilä and Tenhunen, 2008). For exceptions, see Pirttilä and Tuomala (2002) dealing with productivity and relative wage effects of publicly provided private goods, and Aronsson et al. (2005) analyzing how the appearance of equilibrium unemployment affects the incentives for public provision.

[6] To be more specific, to capture the preference reversals implied by quasi-hyperbolic discounting, the model must allow the consumer to make intertemporal (saving and investment) choices at least twice during the life-cycle.

$$
\begin{aligned}
u^i_{0,t} &= v(c^i_{0,t}, z^i_{0,t}) + f(h^i_{0,t}) && (1)\\
u^i_{1,t+1} &= v(c^i_{1,t+1}, z^i_{1,t+1}) + f(h^i_{1,t+1}) && (2)\\
u^i_{2,t+2} &= v(c^i_{2,t+2}, \bar{l}\,) + f(h^i_{2,t+2}), && (3)
\end{aligned}
$$

where c denotes the consumption of a numeraire good, z leisure and h the stock of health capital. Subindices 0, 1 and 2 indicate that the consumer is young, middle-aged and old, respectively. The functions $v(\cdot)$ and $f(\cdot)$ are increasing in their arguments, strictly concave, and all goods are assumed to be normal. Leisure is defined as a time endowment, $\bar{l}$, less the time spent in market work, l. For analytical convenience, the health capital stock is assumed to enter the utility function in a separable way.

An alternative formulation would have been to assume - as did Grossman (1972) - that health capital also (in addition to the direct effects displayed above) enters utility via the time constraint by reducing the time lost (from market work and non-market activities) due to illness. Such an extension would divide the marginal benefit of health capital into two components, i.e. a direct marginal benefit (as above) and an indirect marginal benefit via the preference for leisure, which is most likely realistic. Yet, except for this modification, such an extension would not change the cost benefit rules for publicly provided health care, which leads us to rely on the simpler utility formulation given by equation (1)-(3).

Following the approach developed by Phelps and Pollak (1968), Laibson (1997) and O'Donoghue and Rabin (2003), the intertemporal objective of any generation t is given by

$$U^i_{0,t} = u^i_{0,t} + \beta^i \sum_{j=1}^{2} \Theta^j u^i_{j,t+j}, \tag{4}$$

where $\Theta^j = 1/(1+\theta)^j$ is a conventional (exponential) utility discount factor with utility discount rate θ, whereas $\beta^i \in (0,1)$ is a type-specific time-inconsistent preference for immediate gratification.[7,8]

Our concern is to analyze whether the disincentive to invest in health capital due to quasi-hyperbolic discounting may justify public provision of health care services. As a consequence, we focus attention on the intertemporal aspects of such investments, by assuming that the investment in health capital (i.e. the use of health care services) in period t affects the stock of health capital in period $t+1$, while suppressing any atemporal (within-period) relationship between the use of health care services and the stock of health capital (which is not directly distorted by quasi-hyperbolic discounting).[9] The health capital stock facing the young ability-type i is fixed at $h^i_{0,t}$. For the middle-aged and old, respectively, the health capital stock depends on past investments according to

$$h^i_{1,t+1} = h^i_{0,t}\delta + m^i_{0,t} \tag{5}$$

$$h^i_{2,t+2} = h^i_{1,t+1}\delta + m^i_{1,t+1} \tag{6}$$

[7]With hyperbolic discounting, the discount factor is represented by a hyperbolic function, where the discount rate decreases as the tradeoff occurs further in the future. Quasi-hyperbolic discounting is a discrete approximation, such that the tradeoff between the present period and future periods is discounted at a higher rate than the rate currently applied to tradeoffs in the more distant future.

[8]It is not important for the qualitative results derived below that β differs between the ability-types. Yet, we still allow for such heterogeneity since it gives a more general model, while having negligible costs in terms of additional notational complexity.

[9]We realize, of course, that most aspects of health care may have both atemporal (within-period) and intertermporal effects of the stock of health capital. However, since the appearance of quasi-hyperbolic discounting only affects the incentives underlying investment behavior, and not those underlying atemporal within-period decisions, adding within-period relationships between the use of health care services and the stock of health capital would not affect our understanding of how quasi-hyperbolic discounting may motivate public provision of health care services.

where $\delta \in (0, 1)$ is a depreciation factor - defined as "one minus the depreciation rate" - while $m^i_{0,t}$ and $m^i_{1,t+1}$ are the flow-services of health care used by the young and middle-aged selves.[10] We have simplified by assuming a linear relationship between the use of flow-services of health care and the health capital stock in the next period. The same qualitative results as those derived below will also apply in a more general model where the marginal effect of m is decreasing.

Flow-services of health care may be privately purchased on the market or publicly provided free of charge; each consumer may, therefore, top up the level that the government provides via his/her own private purchases. This means that the flow-services of health care used by the young consumer can be characterized as $m^i_{0,t} = g_{0,t} + x^i_{0,t}$, where $g_{0,t}$ is the amount publicly provided and $x^i_{0,t}$ the private purchase. An analogous definition applies for the middle-aged. Notice that the government is not allowed to provide different levels of health care to the two ability-types, although it may target different age-groups differently. Throughout the paper, we assume that health care services cannot be resold.

Let s denote savings and r denote the interest rate; in addition, let $y^i_{0,t} = w^i_{0,t} l^i_{0,t}$ and $y^i_{1,t+1} = w^i_{1,t+1} l^i_{1,t+1}$ denote the labor income of the young and middle-aged, respectively, and $I^i_{1,t+1} = s^i_{0,t} r_{t+1}$ and $I^i_{2,t+2} = s^i_{1,t+1} r_{t+2}$ denote the capital income facing the middle-aged and old, respectively. There are no bequests here; the initial endowment of wealth by each young consumer is zero. Using this notation, the income tax payment for each of the three phases of the life-cycle can be written as $T^i_{0,t} = T_{0,t}\left(y^i_{1,t}, 0\right)$, $T^i_{1,t+1} = T_{1,t+1}\left(y^i_{1,t+1}, I^i_{1,t+1}\right)$ and $T^i_{2,t+2} = T_{2,t+2}\left(0, I^i_{2,t+2}\right)$.

[10] Although it is likely that the depreciation rate increases with age, we refrain from modeling aged-dependent depreciation here. The reason is that using different depreciation factors in equations (5) and (6) would not alter our qualitative results.

The individual budget constraint is then given by

$$y^i_{0,t} - T^i_{0,t} - s^i_{0,t} = c^i_{0,t} + x^i_{0,t} \tag{7}$$

$$s^i_{0,t} + I^i_{1,t+1} + y^i_{1,t+1} - T^i_{1,t+1} - s^i_{1,t+1} = c^i_{1,t+1} + x^i_{1,t+1} \tag{8}$$

$$s^i_{1,t+1} + I^i_{2,t+2} - T^i_{2,t+2} = c^i_{2,t+2} \tag{9}$$

where the prices of c and x have been normalized to one. Notice that the old consumer does not invest in health capital in our model, since there would be no future benefit associated with such investments.

2.1 Consumer choices

As mentioned above, it is not a priori clear how consumers deal with their self-control problems, and we shall, therefore, make a distinction between naivety and sophistication. In technical terms, naivety is a special case of a model with sophisticated consumers, since the first order conditions for consumption and savings faced by a naive consumer are interpretable as special cases of those obeyed by their sophisticated counterparts. Therefore, to shorten the presentation as much as possible, we use sophistication as a reference case and then discuss how the first order conditions simplify in the special case of naivety.

Also, as the sophisticated consumer implements a time-consistent consumption/savings plan, we begin by analyzing the behavior of the middle-aged generation and then continue with the young generation. For the middle-aged, there is no technical distinction between naivety and sophistication. The reason is, of course, that the old self does not make any forward-looking decisions, implying that the middle-aged self has no direct incentive to modify the behavior of the old self. In fact, in the model described above, the old generation makes no active decision; each old consumer just uses his/her remaining assets for consumption. We have used this particular set up for simplicity, as the possible (atemporal) trade-offs faced by the elderly are not affected by discounting.

2.1.1 Decisions Made by the Middle-Aged Generation

Following the literature on optimal mixed taxation under asymmetric information (e.g., Edwards, Keen and Tuomala, 1994), the consumer-choices are analyzed in two stages; in the first, we derive commodity demand functions (for c and x) conditional on the hours of work and savings; in the second, we derive the labor supply and savings functions. The reason for using this particular approach is that the conditional demand functions will be useful in the policy-problem presented below.

For the middle-aged, the first stage problem means choosing $c^i_{1,t+1}$ and $x^i_{1,t+1}$ to maximize $u^i_{1,t+1} + \beta^i \Theta u^i_{2,t+2}$, i.e. the remaining life-time utility when middle-aged, subject to the nonnegativity constraint $x^i_{1,t+1} \geq 0$; the health capital function (6); and the following budget constraint:

$$b^i_{1,t+1} = c^i_{1,t+1} + x^i_{1,t+1} \tag{10}$$

$$b^i_{2,t+2} = c^i_{2,t+2}, \tag{11}$$

where b is fixed net income adjusted for savings (see below). By substituting equations (10)-(11) into the utility function and using that $\partial h^i_{2,t+2}/\partial x^i_{1,t+1} = 1$, we can then write the Kuhn-Tucker conditions for $x^i_{1,t+1}$ as

$$-\frac{\partial u_{1,t+1}^i}{\partial c_{1,t+1}^i} + \beta^i \Theta \frac{\partial u_{2,t+2}^i}{\partial h_{2,t+2}^i} = A_{1,t+1}^i \leq 0 \tag{12}$$

$$x_{1,t+1}^i A_{1,t+1}^i = 0. \tag{13}$$

In the Kuhn-Tucker conditions (12) and (13), the self-control problem shows up as an adjustment of the weight attached to the future marginal utility of health capital (through the parameter β^i).[11]

If the nonnegativity constraint does not bind, equations (10)-(13) imply the following conditional demand functions:

$$n_{1,t+1}^i = n_1^i \left(b_{1,t+1}^i, z_{1,t+1}^i, (x_{0,t}^i + g_{0,t})\delta + g_{1,t+1} \right) \text{ for } n = c, x. \tag{14}$$

Equation (14) relates the demand (for the numeraire good and private health care services) by the middle-aged consumer to his/her current levels of disposable income (adjusted for savings) and leisure, as well as to the level of publicly provided health care to the middle-aged and the total past consumption of health care services (publicly provided as well as privately purchased). Equation (14) is also a reaction function, as it describes how the young consumer can influence the consumption choices made by his/her middle-aged self.

[11] As indicated above, if we like Grossman (1972) were to assume that health capital also affects the time lost due to illness or injury, ψ, the time constraint of the old consumer would become $z_{2,t+2}^i = \bar{l} - \psi_{2,t+2}^i(h_{2,t+2}^i)$, such that first order condition (12) changes to read

$$-\frac{\partial u_{1,t+1}^i}{\partial c_{1,t+1}^i} + \beta^i \Theta \left[\frac{\partial u_{2,t+2}^i}{\partial h_{2,t+2}^i} + \frac{\partial u_{2,t+2}^i}{\partial z_{2,t+2}^i} \frac{\partial z_{2,t+2}^i}{\partial \psi_{2,t+2}^i} \frac{\partial \psi_{2,t+2}^i}{\partial h_{2,t+2}^i} \right] \leq 0,$$

where $\partial z_{2,t+2}^i / \partial \psi_{2,t+2}^i = -1$ and $\partial \psi_{2,t+2}^i / \partial h_{2,t+2}^i < 0$. The expression within the square-bracket represents the total marginal utility of health capital, which in this case is decomposable into two positive terms; the first representing a direct effect of health capital and the second an indirect effect through the time constraint. The latter effect is not included in the simpler expression (12) above.

The labor supply and savings behavior of the middle-aged consumer is analyzed by choosing $l_{1,t+1}^i$ and $s_{1,t+1}^i$ to maximize $u_{1,t+1}^i + \beta^i \Theta u_{2,t+2}^i$ subject to the health capital function (6), the conditional demand functions (14), and the following budget constraint:

$$b_{1,t+1}^i = s_{0,t}^i (1 + r_{t+1}) + w_{1,t+1}^i l_{1,t+1}^i - T_{1,t+1}^i - s_{1,t+1}^i \quad (15)$$

$$b_{2,t+2}^i = s_{1,t+1}^i (1 + r_{t+2}) - T_{2,t+2}^i. \quad (16)$$

If we define the marginal net wage rate $\omega_{1,t+1}^i = w_{1,t+1}^i \left(1 - \partial T_{1,t+1}^i / \partial y_{1,t+1}^i\right)$ and marginal net interest rate $\rho_{2,t+2}^i = r_{2,t+2}^i \left(1 - \partial T_{2,t+2}^i / \partial I_{2,t+2}^i\right)$, the first order conditions for hours of work and savings can be written as

$$\frac{\partial u_{1,t+1}^i}{\partial c_{1,t+1}^i} \omega_{1,t+1}^i - \frac{\partial u_{1,t+1}^i}{\partial z_{1,t+1}^i} = 0 \quad (17)$$

$$-\frac{\partial u_{1,t+1}^i}{\partial c_{1,t+1}^i} + \beta^i \Theta \left(1 + \rho_{2,t+2}^i\right) \frac{\partial u_{2,t+2}^i}{\partial c_{2,t+2}^i} = 0. \quad (18)$$

Since quasi-hyperbolic discounting does not distort the atemporal tradeoff between consumption and leisure, equation (17) is a standard labor supply condition. Equation (18) shows that the middle-aged consumer saves less than he/she would have done without quasi-hyperbolic discounting, i.e., if $\beta^i = 1$. Equations (17) and (18) imply the following labor supply and saving functions (in which variables other than those decided upon by the consumer's young self have been suppressed)

$$l_{1,t+1}^i = l_1^i \left(s_{0,t}^i, m_{0,t}^i\right) \quad (19)$$

$$s_{1,t+1}^i = s_1^i \left(s_{0,t}^i, m_{0,t}^i\right). \quad (20)$$

By analogy to equation (14) above, equations (19) and (20) are also interpretable as reaction functions, relating the middle-aged consumer's labor supply and savings behavior to his/her savings and health investments as young. The partial derivatives of equations (19) and (20) with respect to $s^i_{0,t}$ and $m^i_{0,t}$ cannot be signed unambiguously.

2.1.2 Decisions Made by the Young Generation

Turning to the young generation, the distinction between naivety and sophistication becomes important. As we mentioned above, a sophisticated consumer recognized that the self-control problem will also appear in future periods, and the young sophisticated consumer will act strategically to influence the incentives faced by his/her middle-aged self. This motive for strategic behavior is absent under naivety (as the young naive consumer erroneously expects the self-control problem to vanish in the future). In the following, we derive the optimality conditions obeyed by sophisticated consumer, and then explain how these conditions simplify under naivety.

The objective function faced by the young ability-type i is given by

$$U^i_{0,t} = u^i_{0,t} + \beta^i \Theta V^i_{1,t+1}, \tag{21}$$

where

$$V^i_{1,t+1} = u^i_{1,t+1} + \Theta u^i_{2,t+2} \tag{22}$$

is the intertemporal objective that the young consumer would like his/her middle-aged self to maximize (which the middle-aged self does not, as his/her objective is given by $u^i_{1,t+1} + \beta^i \Theta u^i_{2,t+2}$). In particular, note that equation (22) does not contain the parameter β^i.

To derive conditional demand functions for the numeraire good and health care services, we maximize equation (21) with respect to $c^i_{0,t}$ and $x^i_{0,t}$ subject to equations (5)-(6), (10)-(11), (14), (15)-(16), and (19)-(20) as well as subject to the following budget constraint

$$b^i_{0,t} = c^i_{0,t} + x^i_{0,t}. \tag{23}$$

By substituting the budget constraint into the objective function and using $\partial h^i_{2,t+2}/\partial x^i_{1,t+1} = \partial h^i_{1,t+1}/\partial x^i_{0,t} = 1$, the Kuhn-Tucker condition for $x^i_{0,t}$ becomes

$$\begin{aligned}
& -\frac{\partial u^i_{0,t}}{\partial c^i_{0,t}} + \beta^i \left[\Theta \frac{\partial u^i_{1,t+1}}{\partial h^i_{1,t+1}} + \Theta^2 \frac{\partial u^i_{2,t+2}}{\partial h^i_{2,t+2}} \delta \right] \\
& + \left(1 - \beta^i\right) \beta^i \Theta^2 \frac{\partial u^i_{2,t+2}}{\partial h^i_{2,t+2}} \frac{\partial x^i_{1,t+1}}{\partial x^i_{0,t}} \\
& + \left(1 - \beta^i\right) \beta^i \Theta^2 \left[\left(1 + \rho^i_{2,t+2}\right) \frac{\partial u^i_{2,t+2}}{\partial c^i_{2,t+2}} - \frac{\partial u^i_{2,t+2}}{\partial h^i_{2,t+2}} \frac{\partial x^i_{1,t+1}}{\partial b^i_{1,t+1}} \right] \frac{\partial s^i_{1,t+1}}{\partial x^i_{0,t}} \\
& - \left(1 - \beta^i\right) \beta^i \Theta^2 \frac{\partial u^i_{2,t+2}}{\partial h^i_{2,t+2}} \frac{\partial \tilde{x}^i_{1,t+1}}{\partial z^i_{1,t+1}} \frac{\partial l^i_{1,t+1}}{\partial x^i_{0,t}} = A^i_{0,t} \leq 0
\end{aligned} \tag{24}$$

$$x^i_{0,t} A^i_{0,t} = 0 \tag{25}$$

where $\partial \tilde{x}^i_{1,t+1}/\partial z^i_{1,t+1} = \partial x^i_{1,t+1}/\partial z^i_{1,t+1} - MRS^i_{z,c,t+1}[\partial x^i_{1,t+1}/\partial b^i_{1,t+1}]$ measures the change in the conditional compensated demand for health care services following increased use of leisure, while $MRS^i_{z,c,t+1} = (\partial u^i_{1,t+1}/\partial z^i_{1,t+1})/(\partial u^i_{1,t+1}/\partial c^i_{1,t+1})$ is the marginal rate of substitution between leisure and the numeraire good faced by the consumer's middle-aged self.[12]

The first row of (24) - which is analogous to (12) faced by the middle-aged consumer - comprises the marginal efficiency condition for $x^i_{0,t}$ that would characterize a young naive

[12] To derive expression (24), note that by substituting the consumer's budget constraint into equation

consumer. It shows that a naive consumer will choose $c^i_{0,t}$ and $x^i_{0,t}$ so that the marginal utility of numeraire consumption equals the sum of discounted marginal utilities of health capital when middle-aged and old, respectively. The intuition is that the health investment made when young will directly affect the health capital stock both when middle-aged and old according to equations (5) and (6). The second, third and fourth rows are due to sophistication and show how the young consumer will adjust his/her consumption of health care services to influence the consumption, savings and labor supply decisions made by his/her middle-aged self. In particular, notice that all these terms are proportional to $1 - \beta^i$: the intuition is that the middle-aged individual discounts his/her future utility by the discount factor $\beta^i\Theta$, whereas the young self wants the middle-aged self to use the discount factor Θ. As such, $1 - \beta^i$ is the "weight" that the young consumer attaches to this discrepancy. The reason as to why the second, third and fourth rows vanish under naivety is that a naive consumer has no incentive to affect the choices made by his/her middle-aged self, as the naive consumer erroneously expects not to be subject to this self-control problem in the future. Another - yet related - difference between naivety and sophistication, therefore, is that the naive consumer underestimates the future marginal utility of health (as he/she overestimates the future stock of health capital).

To provide some further intuition, notice that the second row of (24) is negative, since

(21) and then differentiating with respect to $x^i_{0,t}$, we obtain the first order condition

$$\begin{aligned} 0 \quad \geq \quad & -\frac{\partial u^i_{0,t}}{\partial c^i_{0,t}} + \beta^i \left[\Theta \frac{\partial u^i_{1,t+1}}{\partial h^i_{1,t+1}} + \Theta^2 \frac{\partial u^i_{2,t+2}}{\partial h^i_{2,t+2}} \delta \right] \\ & + \beta^i \Theta \left[\frac{\partial V^i_{1,t+1}}{\partial x^i_{1,t+1}} \frac{\partial x^i_{1,t+1}}{\partial x^i_{0,t}} + \frac{\partial V^i_{1,t+1}}{\partial s^i_{1,t+1}} \frac{\partial s^i_{1,t+1}}{\partial x^i_{0,t}} + \frac{\partial V^i_{1,t+1}}{\partial l^i_{1,t+1}} \frac{\partial l^i_{1,t+1}}{\partial x^i_{0,t}} \right]. \end{aligned}$$

Expression (24) can then be derived by writing the derivatives $\partial V^i_{1,t+1}/\partial x^i_{1,t+1}$, $\partial V^i_{1,t+1}/\partial s^i_{1,t+1}$ and $\partial V^i_{1,t+1}/\partial l^i_{1,t+1}$ in terms of the private first order condition for $x^i_{1,t+1}$, $s^i_{1,t+1}$ and $l^i_{1,t+1}$, respectively.

$\partial x^i_{1,t+1}/\partial x^i_{0,t} \in (-\delta, 0)$.[13] As such, this component contributes to reduce the sophisticated young consumer's health investment, ceteris paribus; a choice made to induce his/her middle-aged self to invest more in health capital.[14] Furthermore, this effect is reinforced if we simplify by adding the assumption that $l^i_{1,t+1}$ is fixed, in which case we can derive $\partial s^i_{1,t+1}/\partial x^i_{0,t} \in (0, -\partial x^i_{1,t+1}/\partial x^i_{0,t})$, such that the sum of the second and third rows of (24) can be written as

$$\left(1-\beta^i\right)\Theta\frac{\partial u^i_{1,t+1}}{\partial c^i_{1,t+1}}\left[\frac{\partial x^i_{1,t+1}}{\partial x^i_{0,t}}+\left(1-\frac{\partial x^i_{1,t+1}}{\partial b^i_{1,t+1}}\right)\frac{\partial s^i_{1,t+1}}{\partial x^i_{0,t}}\right]<0. \quad (26)$$

The sign of the fourth row of (24) depends on whether the use of health care services by the middle-aged self is complementary with, or substitutable for, leisure, and how an increase in $x^i_{0,t}$ affects the hours of work supplied by the middle-aged ability-type i. Therefore, in the general case, the strategic incentive faced by sophisticated young consumers may either contribute to larger or smaller investments in health capital.

If the nonnegativity constraint does not bind, equations (23) and (24) imply the following conditional demand functions (if defined conditional on the use of leisure both when young and when middle-aged);

$$n^i_{0,t}=n^i_0\left(b^i_{0,t},z^i_{0,t},b^i_{1,t+1},z^i_{1,t+1},g_{0,t},g_{1,t+1}\right)\ \text{ for } n=c,x. \quad (27)$$

As mentioned in the introduction, although the present study presupposes that the income taxes are optimally chosen, we do not discuss income tax policy in what follows.

[13] Recall from equation (14) that

$$x^i_{1,t+1}=x^i_1\left(b^i_{1,t+1},z^i_{1,t+1},(x^i_{0,t}+g_{0,t})\delta+g_{1,t+1}\right),$$

where $\partial x^i_{1,t+1}/\partial(x^i_{0,t}\delta)\in(-1,0)$. Therefore, $\partial x^i_{1,t+1}/\partial x^i_{0,t}=\delta[\ \partial x^i_{1,t+1}/\partial(x^i_{0,t}\delta)]\in(-\delta,0)$.

[14] A related result of strategic undersaving was found by Diamond and Köszegi (2004), where the agent reduces his/her savings in order to induce his/her future self to work.

Therefore, to shorten the presentation, we present the first order conditions for labor supply and savings faced by the young consumer in the Appendix, as these conditions will not be used in the study of costs and benefits of publicly provided private goods.

3 Public Provision of Private Goods

The government wants to redistribute as well as correct for the self-control problem described above. Therefore, since the (paternalist) government would like the consumers to behave as if the self-control problem were absent, the government does not discount the future hyperbolically (meaning that $\beta^1 = \beta^2 = 1$ from the point of view of the government).[15] Accordingly, and if we assume a utilitarian type of policy objective, the contribution to social welfare by ability-type i of generation t can be written as

$$V_{0,t}^i = u_{0,t}^i + \sum_{j=1}^{2} \Theta^j u_{j,t+j}^i. \tag{28}$$

Since the consumers are assumed to discount the future hyperbolically, equation (28) differs from the corresponding utility function faced by the young ability-type i of generation t, $U_{0,t}^i$, given by equation (4). The social welfare function becomes

$$W = \sum_t \sum_i \Theta^t V_{0,t}^i. \tag{29}$$

For all t, the resource constraint can be written as

[15] This way of modeling the objective of a paternalist government is in line with earlier comparable literature; see, e.g., O'Donoghue and Rabin (2003, 2006), Aronsson and Thunström (2008) and Aronsson and Granlund (2011).

$$\sum_i \left[w^i_{0,t} l^i_{0,t} + w^i_{1,t} l^i_{1,t}\right] + K_t(1+r_t) - K_{t+1}$$
$$-\sum_i \left[c^i_{0,t} + c^i_{1,t} + c^i_{2,t} + m^i_{0,t} + m^i_{1,t}\right] = 0, \tag{30}$$

where K_t is the capital stock at the beginning of period t, which depends on savings in period $t-1$. Since the government can make lump-sum payments between periods as well as control the capital stock via the nonlinear income taxes, it is not necessary to include the government's budget constraint in the public decision-problem, given that the resource constraint is included (Atkinson and Sandmo 1980, Pirttilä and Tuomala 2001).

We make the conventional assumptions about information: the government can observe income, whereas ability is private information. We also assume that the government wants to redistribute from the high-ability to the low-ability type. As a consequence, it must prevent the high-ability type from pretending to be a low-ability type, i.e., from becoming a mimicker. This is accomplished by imposing a self-selection constraint, implying that the high-ability type (at least weakly) prefers the combination of disposable income and hours of work intended for him/her over the combination intended for the low-ability type.

Note that the hours of work that the high-ability type needs to supply in order to reach the same labor income as the low-ability type is given by $\widehat{l}^2_{0,t} = \left(w^1_{0,t}/w^2_{0,t}\right) l^1_{0,t}$ when young and by $\widehat{l}^2_{1,t+1} = \left(w^1_{1,t+1}/w^2_{1,t+1}\right) l^1_{1,t+1}$ when middle-aged. In the following, all variables with a hat, $\widehat{\ }$, refer to the mimicker. In the same way as for the true low and high-ability types, we can, if the non-negative constraints for x do not bind, define the conditional demand functions for the mimicker as

$$\widehat{n}^2_{0,t} = n^2_0\left(b^1_{0,t}, \widehat{z}^2_{0,t}, b^1_{1,t+1}, \widehat{z}^2_{1,t+1}, g_{0,t}, g_{1,t+1}\right) \text{ for } n = c, x \tag{31}$$

$$\widehat{n}^2_{1,t+1} = n^2_1\left(b^1_{1,t+1}, \widehat{z}^2_{1,t+1}, (\widehat{x}^2_{0,t} + g_{0,t})\delta + g_{1,t+1}\right) \text{ for } n = c, x \tag{32}$$

where, $\widehat{z}^2_{0,t} = \bar{l} - \widehat{l}^2_{0,t}$ and $\widehat{z}^2_{1,t+1} = \bar{l} - \widehat{l}^2_{1,t+1}$. The mimicker receives the same labor and capital income as the low-ability type. However, as the mimicker is more productive than the low-ability type, the mimicker spends more time on leisure, meaning that equations (31) and (32) generally differ from the corresponding conditional demand functions faced by the low-ability type. This means, in turn, that the mimicker and the low-ability type have different health capital stocks when middle-aged and old, respectively, even if their initial stocks were to coincide.

The self-selection constraint can be written as

$$U^2_{0,t} = u^2_{0,t} + \beta^2 \sum_{j=1}^{2} \Theta^j u^2_{j,t+j} \geq \widehat{U}^2_{0,t} = \widehat{u}^2_{0,t} + \beta^2 \sum_{j=1}^{2} \Theta^j \widehat{u}^2_{j,t+j} \tag{33}$$

where the definitions of $\widehat{u}^2_{j,t+j}$ for $j = 0, 1, 2$, are analogous to those for true low- and high-ability types given by equations (1)-(3).

If defined conditional on the publicly provided private good (the cost benefit rule for which will be addressed later), the second best problem will be to choose $l^i_{0,t}$, $b^i_{0,t}$, $l^i_{1,t}$, $b^i_{1,t}$, $b^i_{2,t}$ (for $i = 1, 2$) and K_t for all t to maximize the social welfare function given by equation (29), subject to the accumulation equations for health capital (5)-(6), the self-selection constraint (33), the resource constraint (30), and the conditional demand functions (14),(27) and (31)-(32).[16]

[16] As the government is equipped with nonlinear taxes on labor and capital income by assumption, it is able to implement any desired combination of work hours and disposable income for each ability-type and

The Lagrangean corresponding to this policy problem is presented in the Appendix together with the associated first order conditions for work hours, disposable income and the capital stock, which reflect an optimal income tax policy implemented for generation t. Our concern is then to analyze the welfare effects of publicly provided private goods given that the income taxes are optimal. Also, we follow some of the earlier literature on optimal nonlinear income taxation in dynamic models in assuming that the government can credibly commit to the announced tax and expenditure policies.[17]

We start by analyzing public provision to the young generation and then continue with public provision to the middle-aged.

3.1 Public Provision to the Young

To facilitate comparison with earlier research, we begin by briefly discussing public provision under the assumption that the consumers do not discount the future hyperbolically, i.e. behave as if $\beta^i = 1$ for $i = 1, 2$. We will then return to the assumption that the consumers discount hyperbolically and examine the welfare effect of publicly provided private goods to the young generation under naivety as well as sophistication.

generation, as well as an optimal path for the capital stock, subject to the self-selection, health capital and resource constraints. It is, therefore, convenient to write the second best problem as a direct decision-problem where the government (or social planner) directly decides upon work hours and disposable income for each ability-type and generation as well as the capital stock.

[17] See, e.g., Brett (1997), Pirttilä and Tuomala (2001), Aronsson et al. (2009) and Aronsson and Johansson-Stenman (2010). Situations where the government implements a time-consistent policy without commitment are analyzed by, e.g., Brett and Weymark (2008) and Aronsson and Sjögren (2009).

3.1.1 Without the Self-Control Problem

The total consumption of health care services by the young ability-type i is given by $m^i_{0,t} = g_{0,t} + x^i_{0,t}$. Now, let

$$\frac{dm^i_{0,t}}{dg_{0,t}} = 1 + \frac{\partial x^i_{0,t}}{\partial g_{0,t}} - \frac{\partial x^i_{0,t}}{\partial b^i_{0,t}} \tag{34}$$

$$\frac{dm^i_{1,t+1}}{dg_{0,t}} = \frac{\partial x^i_{1,t+1}}{\partial g_{0,t}} + \frac{\partial x^i_{1,t+1}}{\partial x^i_{0,t}} \left[\frac{\partial x^i_{0,t}}{\partial g_{0,t}} - \frac{\partial x^i_{0,t}}{\partial b^i_{0,t}} \right] \tag{35}$$

denote how $m^i_{0,t}$ and $m^i_{1,t+1}$, respectively, responds to a <u>tax-financed</u> increase in $g_{0,t}$. The responses by the mimicker are analogous. Then, if the consumers behave as if $\beta^1 = \beta^2 = 1$, we show in the Appendix that the welfare effect of an increase in $g_{0,t}$ can be written as

$$\begin{aligned}
\frac{\partial W}{\partial g_{0,t}} = {} & \Theta^t \sum_i \left\{ \left(\frac{\partial V^i_{0,t}}{\partial m^i_{0,t}} - \frac{\partial V^i_{0,t}}{\partial c^i_{0,t}} \right) \frac{dm^i_{0,t}}{dg_{0,t}} + \left(\frac{\partial V^i_{0,t}}{\partial m^i_{1,t+1}} - \frac{\partial V^i_{0,t}}{\partial c^i_{1,t+1}} \right) \frac{dm^i_{1,t+1}}{dg_{0,t}} \right\} \\
& + \lambda_t \left[\left(\frac{\partial V^2_{0,t}}{\partial m^2_{0,t}} - \frac{\partial V^2_{0,t}}{\partial c^2_{0,t}} \right) \frac{dm^2_{0,t}}{dg_{0,t}} - \left(\frac{\partial \widehat{V}^2_{0,t}}{\partial \widehat{m}^2_{0,t}} - \frac{\partial \widehat{V}^2_{0,t}}{\partial \widehat{c}^2_{0,t}} \right) \frac{d\widehat{m}^2_{0,t}}{dg_{0,t}} \right] \\
& + \lambda_t \left[\left(\frac{\partial V^2_{0,t}}{\partial m^2_{1,t+1}} - \frac{\partial V^2_{0,t}}{\partial c^2_{1,t+1}} \right) \frac{dm^2_{1,t+1}}{dg_{0,t}} - \left(\frac{\partial \widehat{V}^2_{0,t}}{\partial \widehat{m}^2_{1,t+1}} - \frac{\partial \widehat{V}^2_{0,t}}{\partial \widehat{c}^2_{1,t+1}} \right) \frac{d\widehat{m}^i_{1,t+1}}{dg_{0,t}} \right].
\end{aligned} \tag{36}$$

As we assume away quasi-hyperbolic discounting here, the consumer objective, $U^i_{0,t}$, becomes equal to the individual contribution to the social welfare function, $V^i_{0,t}$, for each ability-type. Note first that an increase in $g_{0,t}$ affects the instantaneous utility via the consumption of health care services both when young and when middle-aged, i.e. via $m^i_{0,t}$ and $m^i_{1,t+1}$, respectively, which explains the first row of equation (36). The second and third rows appear because a change in $g_{0,t}$ affects the self-selection constraint via the consumption of health care services by the young high-ability type and young mimicker (the second row), and via the consumption of health care services by the middle-aged high-ability type and middle-aged mimicker (the third row).[18]

Equation (36) is just an intertemporal analogue to formulas derived in static models. If $g_{0,t}$ is small enough to imply that the nonnegativity constraint attached to $x^i_{0,t}$ does not bind, then the first term within brackets on the right hand side of equation (36) vanishes because the consumer has made an optimal choice, i.e.

$$\frac{\partial V^i_{0,t}}{\partial m^i_{0,t}} - \frac{\partial V^i_{0,t}}{\partial c^i_{0,t}} = \Theta \frac{\partial u^i_{1,t+1}}{\partial h^i_{1,t+1}} + \Theta^2 \frac{\partial u^i_{2,t+2}}{\partial h^i_{2,t+2}}\delta - \frac{\partial u^i_{0,t}}{\partial c^i_{0,t}} = 0. \tag{37}$$

Analogous results apply for the young mimicker if $\widehat{x}^2_{0,t} > 0$, as well as for the middle-aged true ability-types (if $x^i_{1,t+1} > 0$, for $i = 1, 2$) and the middle-aged mimicker (if $\hat{x}^2_{1,t+1} > 0$), respectively. Furthermore, with $x^i_{0,t} > 0$, it also follows that $dm^i_{0,t}/dg_{0,t} = 1 + \partial x^i_{0,t}/\partial g_{0,t} - \partial x^i_{0,t}/\partial b^i_{0,t} = 0$, simply because each consumer adjusts his/her own private consumption of health care services such that the total consumption remains unchanged.

As $g_{0,t}$ continues to increase, one of the nonnegativity constraints will eventually become binding. For instance, at the point where the young ability-type i becomes crowded out, we have $\partial V^i_{0,t}/\partial m^i_{0,t} - \partial V^i_{0,t}/\partial c^i_{0,t} < 0$ and $dm^i_{0,t}/dg_{0,t} = 1$, meaning that the first term on the right hand side of equation (36) contributes to lower welfare (as ability-type i is forced to consume more health care services than he/she prefers). Similarly, if the young mimicker becomes crowded out, then the second term in the second row contributes to higher welfare, i.e. $-\lambda_t[\partial \hat{V}^2_{0,t}/\partial \hat{m}^2_{0,t} - \partial \hat{V}^2_{0,t}/\partial \hat{c}^2_{0,t}] > 0$. The intuition is that decreased utility for the mimicker leads to a relaxation of the self-selection constraint. The components

[18] With reference to foonote 11, note that equation (36) would take the same general form if we were to assume that health capital also affects the time available for market work and nonmarket activities. The only effect would be to add an extra component to the marginal utility of health capital in equations such as (37).

referring to the middle-aged in equation (36) have analogous interpretations. In other words, public provision is welfare improving if the mimicker becomes crowded out first, which is analogous to results derived in earlier literature on public provision of private goods under optimal income taxation.

As we mentioned above, Pirttilä and Tuomala (2002) also address public provision of private goods under nonlinear income taxation in an OLG model. In their study, the consumers live for two periods, and the government provides education (the consumption of which the individuals can top up through private purchases). They derive a formula that resembles equation (36); yet with two important modifications. First, since education leads to human capital accumulation that (partly) carries over to future generations in their study, public provision at time t also affects the welfare of future generations (through endogenous relative wages and human capital externalities). Second, due to their focus on human capital and assumption that agents only live for two periods, they do not distinguish between behavioral responses of different age groups, as we do. In our framework, the consumers may adjust their consumption also in the intertemporal dimension: if the young consumer becomes crowded out, this effect is partly offset via adjustments made by the middle-aged self, given that the nonnegativity constraint faced by the middle-aged self does not bind. As will be explained in greater detail below, this reduces the size of the welfare effect, although it does not change the qualitative result.

3.1.2 Naive Consumers with Present-Biased Preferences

If the consumers have present-biased preferences, we show in the Appendix that the analogue to equation (36) can be written as

$$\frac{\partial W}{\partial g_{0,t}} = \Theta^t \sum_i \left\{ \left(\frac{\partial V^i_{0,t}}{\partial m^i_{0,t}} - \frac{\partial V^i_{0,t}}{\partial c^i_{0,t}} \right) \frac{dm^i_{0,t}}{dg_{0,t}} + \left(\frac{\partial V^i_{0,t}}{\partial m^i_{1,t+1}} - \frac{\partial V^i_{0,t}}{\partial c^i_{1,t+1}} \right) \frac{dm^i_{1,t+1}}{dg_{0,t}} \right\}$$
$$+\lambda_t \left[\left(\frac{\partial U^2_{0,t}}{\partial m^2_{0,t}} - \frac{\partial U^2_{0,t}}{\partial c^2_{0,t}} \right) \frac{dm^2_{0,t}}{dg_{0,t}} - \left(\frac{\partial \widehat{U}^2_{0,t}}{\partial \widehat{m}^2_{0,t}} - \frac{\partial \widehat{U}^2_{0,t}}{\partial \widehat{c}^2_{0,t}} \right) \frac{d\widehat{m}^2_{0,t}}{dg_{0,t}} \right] \tag{38}$$
$$+\lambda_t \left[\left(\frac{\partial U^2_{0,t}}{\partial m^2_{1,t+1}} - \frac{\partial U^2_{0,t}}{\partial c^2_{1,t+1}} \right) \frac{dm^2_{1,t+1}}{dg_{0,t}} - \left(\frac{\partial \widehat{U}^2_{0,t}}{\partial \widehat{m}^2_{1,t+1}} - \frac{\partial \widehat{U}^2_{0,t}}{\partial \widehat{c}^2_{1,t+1}} \right) \frac{d\widehat{m}^i_{1,t+1}}{dg_{0,t}} \right].$$

The difference compared to equation (36) is that the first and second rows of equation (38) contain derivatives of $U^2_{0,t}$ and $\widehat{U}^2_{0,t}$ instead of $V^2_{0,t}$ and $\widehat{V}^2_{0,t}$. These components serve to prevent mimicking and are associated with the self-selection constraint; as such, they reflect the actual consumer-objective (not the social welfare function). This is important because the objective function facing ability-type i, $U^i_{0,t}$, will differ from his/her contribution to the social welfare function, $V^i_{0,t}$, if the consumers have present-biased preferences.

Earlier studies based on model-economies where the consumers are fully rational show that public provision of private goods can be welfare improving if the mimicker is crowded out first (e.g., Boadway and Marchand, 1995; Blomquist and Christiansen, 1995). We have derived the following result from equation (38):

Proposition 1 *Suppose that either the mimicker or the low-ability type is crowded out first, i.e., $x^2_{0,t} > \min\left\{x^1_{0,t}, \hat{x}^2_{0,t}\right\}$ when $g_{0,t} = 0$. Then, if the consumers have present-biased preferences and are naive, there exists a level of $g_{0,t} > 0$ for which the welfare is strictly higher than without public provision.*

The proof of Proposition 1 is straight forward. Suppose first that $g_{0,t}$ is small enough to imply $x^i_{0,t} > 0$ and $x^i_{1,t+1} > 0$. This means that the first row of equation (38) is zero, because $dm^i_{0,t}/dg_{0,t} = 0$ and $dm^i_{1,t+1}/dg_{0,t} = 0$ (as in the absence of the self-control problem), because the consumer adjusts his/her private consumption of health care services to maintain the total consumption of health care services at the desired level. However, the expressions within parenthesis in the first row are no longer equal to zero, since the self-control problem discussed here implies that each consumer uses less health care services than preferred by the paternalistic government, i.e.[19]

$$\frac{\partial V_{0,t}^i}{\partial m_{0,t}^i} - \frac{\partial V_{0,t}^i}{\partial c_{0,t}^i} = \left(1-\beta^i\right)\left(\Theta \frac{\partial u_{1,t+1}^i}{\partial h_{1,t+1}^i} + \Theta^2 \frac{\partial u_{2,t+2}^i}{\partial h_{2,t+2}^i}\delta\right) > 0 \tag{39}$$

$$\frac{\partial V_{0,t}^i}{\partial m_{1,t+1}^i} - \frac{\partial V_{0,t}^i}{\partial c_{1,t+1}^i} = \left(1-\beta^i\right)\Theta^2 \frac{\partial u_{2,t+2}^i}{\partial h_{2,t+2}^i} > 0. \tag{40}$$

Therefore, at the point where the nonnegativity constraint becomes binding we have $dm_{0,t}^i/dg_{0,t} = 1$, which in combination with equation (39) means that welfare increases via the first term on the right hand side of equation (38). This welfare increase is, in turn, partly (yet not fully) offset by the intertemporal adjustment made by the middle-aged self when $x_{0,t}^i = 0$. To see this, note that $dm_{1,t+1}^i/dg_{0,t} = \partial x_{1,t+1}^i/\partial g_{0,t}$ if $x_{0,t}^i = 0$. We can then write the first row of equation (38) as follows by combining equations (39) and (40);

$$\begin{aligned} &\left(\frac{\partial V_{0,t}^i}{\partial m_{0,t}^i} - \frac{\partial V_{0,t}^i}{\partial c_{0,t}^i}\right)\frac{dm_{0,t}^i}{dg_{0,t}} + \left(\frac{\partial V_{0,t}^i}{\partial m_{1,t+1}^i} - \frac{\partial V_{0,t}^i}{\partial c_{1,t+1}^i}\right)\frac{dm_{1,t+1}^i}{dg_{0,t}} \\ &= \left(1-\beta^i\right)\left[\Theta \frac{\partial u_{1,t+1}^i}{\partial h_{1,t+1}^i} + \Theta^2 \frac{\partial u_{2,t+2}^i}{\partial h_{2,t+2}^i}\left(\delta + \frac{\partial x_{1,t+1}^i}{\partial g_{0,t}}\right)\right], \end{aligned} \tag{41}$$

[19] To derive equation (39), take the the derivative of equation (28) with respect to $m_{0,t}^i$, while using equations (5) and (6). Then, take the derivative of equation (28) with respect to $c_{0,t}^i$ and substitute for $\partial u_{0,t}^i/\partial c_{0,t}^i$ using the first order condition of the consumer, (24), keeping in mind that the second, third and fourth rows vanish under naivety. Equation (40) can be derived in the same general way.

which is positive even if $\partial x^i_{1,t+1}/\partial g_{0,t}$ approaches $-\delta$.[20] That the intertemporal adjustment effect does not eliminate the gain of public provision is explained by the assumption that the consumers' instantaneous utilities are time-separable, meaning that increased health when middle-aged does not reduce the utility of health when old. Therefore, a consumer's behavioral response to a larger health capital stock when middle-aged will never lead to a smaller health capital stock when old. Note finally that if the middle-aged self is crowded out first, so $x^i_{1,t+1} = 0$, then $dm^i_{1,t+1}/dg_{0,t} = 0$ and the (negative) intertemporal adjustment effect vanishes.

It is now straight forward to see that Proposition 1 applies. Suppose that we were to increase the public provision up to a point where either the low-ability type or the mimicker (or both of them) is crowded out. If the low-ability type is crowded out first, the welfare gain is given by equation (41) above; if the mimicker is crowded out first, there is a welfare gain due to the relaxation of the self-selection constraint (as discussed in the previous subsection). However, if we were to assume that the high-ability type is crowded out first, we have two counteracting effects; a welfare gain described by equation (41) and a welfare loss due to a tighter self-selection constraint.

3.1.3 Sophisticated Consumers with Present-Biased Preferences

Note that equation (38) provides a general characterization of the welfare effect of increased public provision, and is written on a format that applies irrespective of whether the consumers are naive or sophisticated. However, the signs of the expressions in parentheses (i.e. the difference between the marginal utility of health care and the marginal utility of numeraire consumption) may depend on the distinction between naivety and

[20] To see that $\partial x^i_{1,t+1}/\partial g_{0,t} > -\delta$, recall from equation (14) that $\partial x^i_{1,t+1}/\partial(g_{0,t}\delta) \in (-1, 0)$ and, as a consequence, $\partial x^i_{1,t+1}/\partial g_{0,t} = \delta[\partial x^i_{1,t+1}/\partial(g_{0,t}\delta)] > -\delta$.

sophistication.

To see this, we may rewrite the young consumer's first order condition for health care services as follows (given that $x_{0,t}^i > 0$):

$$-\frac{\partial u_{0,t}^i}{\partial c_{0,t}^i} + \beta^i \left[\Theta \frac{\partial u_{1,t+1}^i}{\partial h_{1,t+1}^i} + \Theta^2 \frac{\partial u_{2,t+2}^i}{\partial h_{2,t+2}^i} \delta \right] + \Gamma_{0,t}^i = 0 \tag{42}$$

where $\Gamma_{0,t}^i = 0$ under naivety (as the young naive consumer does not act strategically vis-a-vis his/her middle-aged self), whereas

$$\begin{aligned}
\Gamma_{0,t}^i &= \left(1 - \beta^i\right) \beta^i \Theta^2 \frac{\partial u_{2,t+2}^i}{\partial h_{2,t+2}^i} \frac{\partial x_{1,t+1}^i}{\partial x_{0,t}^i} \\
&+ \left(1 - \beta^i\right) \beta^i \Theta^2 \left[\left(1 + \rho_{2,t+2}^i\right) \frac{\partial u_{2,t+2}^i}{\partial c_{2,t+2}^i} - \frac{\partial u_{2,t+2}^i}{\partial h_{2,t+2}^i} \frac{\partial x_{1,t+1}^i}{\partial b_{1,t+1}^i} \right] \frac{\partial s_{1,t+1}^i}{\partial x_{0,t}^i} \\
&- \left(1 - \beta^i\right) \beta^i \Theta^2 \frac{\partial u_{2,t+2}^i}{\partial h_{2,t+2}^i} \frac{\partial \tilde{x}_{1,t+1}^i}{\partial z_{1,t+1}^i} \frac{\partial l_{1,t+1}^i}{\partial x_{0,t}^i}
\end{aligned} \tag{43}$$

is generally nonzero under sophistication and reflects an incentive faced by the young consumer to affect choices made by his/her middle-aged self. We can then derive the following result:

Proposition 2 *Suppose that either the mimicker or the low-ability type is crowded out first, i.e.,* $x_{0,t}^2 > \min\left\{x_{0,t}^1, \hat{x}_{0,t}^2\right\}$ *when* $g_{0,t} = 0$*. Then, if the consumers have present-biased preferences and are sophisticated, and if* $\Gamma_{0,t}^1 \leq 0$*, there exists a level of* $g_{0,t} > 0$ *for which the welfare is strictly higher than without public provision.*

Proposition 2 follows by analogy to Proposition 1 by observing that

$$\frac{\partial V_{0,t}^i}{\partial m_{0,t}^i} - \frac{\partial V_{0,t}^i}{\partial c_{0,t}^i} = \left(1 - \beta^i\right) \left(\Theta \frac{\partial u_{1,t+1}^i}{\partial h_{1,t+1}^i} + \Theta^2 \frac{\partial u_{2,t+2}^i}{\partial h_{2,t+2}^i} \delta \right) - \Gamma_{0,t}^i > 0, \text{ if } \Gamma_{0,t}^i \leq 0. \tag{44}$$

Note that $\Gamma^1_{0,t} \leq 0$ is a sufficient - not necessary - condition for the right hand side of equation (44) to be positive. As a consequence, the qualitative result indicated by Proposition 2 also applies if $\Gamma^1_{0,t} > 0$ and small enough in absolute value.

By comparison with the cost benefit rule for public provision derived in the previous subsection, it follows that the strategic incentive faced by the young sophisticated consumer may either strengthen or counteract the result presented in Proposition 1. As we indicated in Section 2, the first row of equation (43) is negative. Therefore, if the sum of the second and third row of equation (43) is either negative or small in absolute value, then $\Gamma^i_{0,t} < 0$.

An interesting example as to when the right hand side of equation (44) is positive is where the nonnegativity constraint faced by each middle-aged self binds at a lower level of $g_{0,t}$ than the corresponding nonnegativity constraint faced by the young self, meaning that equation (38) should be evaluated for $x^1_{1,t+1} = x^2_{1,t+1} = \hat{x}^1_{1,t+1} = 0$. In this case, the right hand side of equation (43) is equal to zero. The intuition is that if $x^i_{1,t+1} = 0$ - and with $s^i_{0,t}$ (which the government controls via the income tax system) held constant - there is no channel via which $x^i_{0,t}$ may affect the first order conditions for $l^i_{1,t+1}$ and $s^i_{1,t+1}$ presented in equations (17) and (18). This means that the policy rule for public provision takes the same form as under naivety.

Pirttilä and Tenhunen (2008) have also examined paternalistic motives for publicly provided private goods under optimal income taxation. Their study is based on a static model where the objective of the social planner differs from the objective faced by the consumers. They find that it is welfare improving to publicly provide a private good that is "undervalued" by the consumers (in the sense that the social marginal willingness to pay exceeds the private marginal willingness to pay), if this good is either substitutable for leisure (in which case the mimicker is crowded out before the low-ability type) or if leisure is weakly separable from the other goods in the utility function. Although this result has important similarities to our Proposition 1 above, an important difference is that we also find that public provision might be welfare improving if the low-ability type is crowded out first. In a way similar to our study, Pirttilä and Tenhunen also give an example where the consumers attach less value to their future health than preferred by the government (which they interpret as hyperbolic discounting); yet, as they use a static model, they are unable to distinguish between naivety and sophistication and, therefore, identify how the strategic incentives faced by the consumers affect the policy incentives underlying publicly provided private goods. Furthermore, a static model does not capture intertemporal consumption-adjustments over the individual life-cycle. As we indicated above, it matters for the welfare effect of public provision to the young generation whether or not the individual's young self becomes crowded out before his/her middle-aged self. These intertemporal adjustments will be even more important in the context of public provision to the middle-aged, to which we turn next.

3.2 Public Provision to the Middle-Aged

Without hyperbolic discounting, the conditions under which public provision of health care to the middle-aged leads to higher welfare are analogous to those described for public provision towards the young generation above. Therefore, we only examine the policy rule for public provision under quasi-hyperbolic discounting here.

In a way similar to the notation used above, let

$$\frac{dm^i_{0,t}}{dg_{1,t+1}} = \frac{\partial x^i_{0,t}}{\partial g_{1,t+1}} - \frac{\partial x^i_{0,t}}{\partial b^1_{1,t+1}} \tag{45}$$

$$\frac{dm^i_{1,t+1}}{dg_{1,t+1}} = 1 + \left(\frac{\partial x^i_{1,t+1}}{\partial g_{1,t+1}} - \frac{\partial x^i_{1,t+1}}{\partial b^i_{1,t+1}} \right) + \frac{\partial x^i_{1,t+1}}{\partial m^i_{0,t}} \frac{dm^i_{0,t}}{dg_{1,t+1}} \tag{46}$$

denote how the total consumption of health care services by ability-type i, when young and when middle-aged, is affected by a tax-financed increase in the provision of health care services to the middle-aged in period $t+1$, $g_{1,t+1}$. We show in the Appendix that the cost benefit rule for $g_{1,t+1}$ can be written as

$$\begin{aligned} \frac{\partial W}{\partial g_{1,t+1}} = {} & \Theta^t \sum_i \left\{ \left(\frac{\partial V^i_{0,t}}{\partial m^i_{0,t}} - \frac{\partial V^i_{0,t}}{\partial c^i_{0,t}} \right) \frac{dm^i_{0,t}}{dg_{1,t+1}} + \left(\frac{\partial V^i_{0,t}}{\partial m^i_{1,t+1}} - \frac{\partial V^i_{0,t}}{\partial c^i_{1,t+1}} \right) \frac{dm^i_{1,t+1}}{dg_{1,t+1}} \right\} \\ & + \lambda_t \left[\left(\frac{\partial U^2_{0,t}}{\partial m^2_{0,t}} - \frac{\partial U^2_{0,t}}{\partial c^2_{0,t}} \right) \frac{dm^2_{0,t}}{dg_{1,t+1}} - \left(\frac{\partial \widehat{U}^2_{0,t}}{\partial \widehat{m}^2_{0,t}} - \frac{\partial \widehat{U}^2_{0,t}}{\partial \widehat{c}^2_{0,t}} \right) \frac{d\widehat{m}^2_{0,t}}{dg_{1,t+1}} \right] \\ & + \lambda_t \left[\left(\frac{\partial U^2_{0,t}}{\partial m^2_{1,t+1}} - \frac{\partial U^2_{0,t}}{\partial c^2_{1,t+1}} \right) \frac{dm^2_{1,t+1}}{dg_{1,t+1}} - \left(\frac{\partial \widehat{U}^2_{0,t}}{\partial \widehat{m}^2_{1,t+1}} - \frac{\partial \widehat{U}^2_{0,t}}{\partial \widehat{c}^2_{1,t+1}} \right) \frac{d\widehat{m}^2_{1,t+1}}{dg_{1,t+1}} \right] \end{aligned} \tag{47}$$

We can then use equation (47) to derive the following result:

Proposition 3 *Suppose that* $x^2_{1,t+1} > \min \left\{ x^1_{1,t+1}, \hat{x}^2_{1,t+1} \right\}$ *without any public provision. Then, if the consumers have present-biased preferences, and irrespective of whether they are characterized by naivety or sophistication, there exists a level of* $g_{1,t+1} > 0$ *for which the welfare is strictly higher than without public provision, if the young generation is crowded out before the middle-aged generation.*

To see this result more clearly, suppose that all young agents have become crowded out at $g^*_{1,t+1}$, meaning that $dm^1_{0,t}/dg_{1,t+1} = dm^2_{0,t}/dg_{1,t+1} = d\hat{m}^2_{0,t}/dg_{1,t+1} = 0$ for $g_{1,t+1} \geq g^*_{1,t+1}$.

Then, if the middle-aged low-ability type becomes crowded out at, say, $g^{**}_{1,t+1} > g^{*}_{1,t+1}$, and if the middle-aged mimicker is not yet crowded out at this point, meaning that $\hat{x}^2_{1,t+1} > 0$ at $g_{1,t+1} = g^{**}_{1,t+1}$, equation (47) reduces to read

$$\frac{\partial W}{\partial g_{1,t+1}} = \Theta^t \left(\frac{\partial V^1_{0,t}}{\partial m^1_{1,t+1}} - \frac{\partial V^1_{0,t}}{\partial c^1_{1,t+1}} \right) \frac{dm^1_{1,t+1}}{dg_{1,t+1}} = \left(1 - \beta^1\right) \Theta^{t+2} \frac{\partial u^1_{2,t+2}}{\partial h^1_{2,t+2}} > 0. \tag{48}$$

By analogy, if the middle-aged mimicker is crowded out, while the middle-aged low-ability type is not, equation (47) simplifies to

$$\frac{\partial W}{\partial g_{1,t+1}} = -\lambda_t \left(\frac{\partial \widehat{U}^2_{0,t}}{\partial \widehat{m}^2_{1,t+1}} - \frac{\partial \widehat{U}^2_{0,t}}{\partial \widehat{c}^2_{1,t+1}} \right) \frac{d\widehat{m}^2_{1,t+1}}{dg_{1,t+1}} > 0 \tag{49}$$

as crowding out here means $\partial \widehat{U}^2_{0,t}/\partial \widehat{m}^2_{1,t+1} - \partial \widehat{U}^2_{0,t}/\partial \widehat{c}^2_{1,t+1} < 0$ and $d\widehat{m}^2_{1,t+1}/dg_{1,t+1} = 1$. The intuition as to why these results apply both under naivety and sophistication is, of course, that the welfare effect of public provision is governed solely by the instantaneous utility change and behavioral response associated with the middle-aged low-ability type or mimicker. Sophistication only gives rise to a strategic motive faced by the young consumers (not the middle-aged), which are already crowded out by assumption.

On the other hand, if each middle-aged consumer is crowded out before his/her young self, Proposition 3 no longer applies. In that case, $dm^i_{1,t+1}/dg_{1,t+1} = 1$ and $dm^i_{0,t}/dg_{1,t+1}$ is (most likely) negative at the point where the middle-aged ability-type i is crowded out, suggesting that the first row on the right hand side of equation (47) can be either positive or negative. Then, if $g_{1,t+1}$ continues to increase, and we eventually reach the point where the young consumer becomes crowded out, we may already have passed the level of $g_{1,t+1}$ at which $\partial V^i_{0,t}/\partial m^i_{1,t+1} - \partial V^i_{0,t}/\partial c^i_{1,t+1}$ switches sign from positive to negative. As a consequence, the welfare effect of public provision remains ambiguous here.

It is worth noticing that, although Proposition 3 applies both for naive and sophisticated consumers, the distinction between naivety and sophistication is still important for the outcome. Whether or not the consumers are first crowded out when young instead of when middle-aged (meaning that the condition on which Proposition 3 is based will apply) might depend on whether they are naive or sophisticated. In Section 2, we gave some intuition as to why the young sophisticated consumer may reduce his/her own investment in health care to provide incentives for his/her middle-aged self to spend more resources on health care services. Alternatively, a young naive consumer may spend less resources on health care services than a young sophisticated consumer, simply because the naive consumer underestimates his/her future marginal utility of health capital. These two mechanisms can work in opposite directions. It is, therefore, inconclusive whether the condition in Proposition 3 is more likely to apply for naive than sophisticated consumers or vice versa.

Finally, Propositions 1, 2 and 3 together give a strong argument for public provision of health care services both to the young and middle-aged. To see this, note that an increase in $g_{0,t}$ up to the point where the young low-ability type or mimicker is crowded out makes it more likely that the condition for welfare improving public provision to the middle-aged in Proposition 3 is fulfilled, if the middle-aged generation is not yet crowded out.

4 Summary and Discussion

This paper develops an OLG model with two ability-types, where the consumers suffer from a self-control problem generated by quasi-hyperbolic discounting, to analyze the welfare effects of publicly provided private goods with long-term consequences for individual well-being. Health care services are used to exemplify private goods with an explicit intertemporal dimension: the benefits (or at least some of them) following the use of such services are likely to arise in the future in the form of increased health capital, while the cost arises at the time the investment is made. Therefore, the appearance of quasi-hyperbolic discounting means that the health capital stock might become too small for the individual himself/herself in a longer time-perspective, which provides a paternalistic motive for public provision. The policy instruments faced by the government are non-linear taxes on labor income and capital income as well as the expenditures associated with publicly provided health care. To our knowledge, this is the first study dealing with publicly provided private goods under quasi-hyperbolic discounting.

In our model, each consumer lives for three periods, which allows us to distinguish between public provision to the young and the middle-aged as well as between naivety and sophistication in terms of consumer behavior. We find that publicly provided health care to the young generation is welfare improving under optimal income taxation, if the consumers have present-biased preferences and are naive; a result which applies independently of whether the mimicker is crowded out before the low-ability type or vice versa. The intuition is that quasi-hyperbolic discounting leads the consumer to spend too little resources on health care, while naivety means that the policy incentives are not distorted by strategic consumer behavior. With sophistication, on the other hand, the young consumer acts strategically vis-a-vis his/her middle-aged self which may, in turn, either increases or decreases the demand for health care as young. If the strategic incentives contribute to reduce the demand for private health care services among the young, then the policy incentives underlying public provision are analogous to those under naivety. However, if the strategic consumer behavior increases the demand for health care services, public provision to the young generation is not necessarily welfare improving.

The policy incentives for public provision of health care services to the middle-aged generation differ from those described above. We find that public provision to the middle-aged is welfare improving if the young generation is crowded out before the middle-aged generation. Furthermore, this result holds independently of whether the consumers are naive or sophisticated, as this distinction only affects the incentives facing the young generation (which is already crowded out by assumption). If the middle-aged are crowded out first, there will be a counteracting effect following as the young may reduce their own private consumption of health care in response to the anticipated policy-induced increase when middle-aged.

We interpret our results to give a strong argument for publicly provided health care to the young and the middle-aged. Public provision to the young leads by itself (most likely) to higher welfare as well as increases the likelihood that the conditions for welfare improving public provision to the middle-aged are fulfilled (by crowding out the private demand for health care among the young).

Future research may take several directions, and we briefly discuss two of them here. First, it would be interesting to analyze the consequences of heterogeneity with respect to the self-control problem in greater detail, such that this problem is only present for part of the population, and where the self-control problem is not perfectly correlated with ability (as it is in our paper). Our conjecture is that this scenario might change the relevant tradeoffs for public policy in fundamental ways, since the government must, in this case, balance the incentive to correct for the self-control problem against the welfare cost for those that do not suffer from this problem (see O'Donoghue and Rabin, 2006, for a study of optimal sin taxes in a similar context). Yet, the appearance of several sources of unobserved heterogeneity is also likely to require a much more complex model than in the present paper. Second, real world tax instruments may differ from those assumed here; for instance, a linear capital income tax makes the government unable to perfectly control the capital stock. In that case, public provision might also serve as an indirect instrument to affect the savings behavior. We leave these extensions for future study.

5 Appendix

Labor Supply and Savings Behavior by the Young Consumer

The first order conditions for work hours and saving, respectively, can be written as

$$\frac{\partial u_{0,t}^{i}}{\partial c_{0,t}^{i}}\omega_{0,t}^{i}-\frac{\partial u_{0,t}^{i}}{\partial z_{0,t}^{i}}=0 \tag{A1}$$

$$\begin{aligned}
&-\frac{\partial u_{0,t}^{i}}{\partial c_{0,t}^{i}}+\beta^{i}\Theta\left(1+\rho_{1,t+1}^{i}\right)\frac{\partial u_{1,t+1}^{i}}{\partial c_{1,t+1}^{i}}\\
&+\beta^{i}\left(1-\beta^{i}\right)\Theta^{2}\frac{\partial u_{2,t+2}^{i}}{\partial h_{2,t+2}^{i}}\frac{\partial x_{1,t+1}^{i}}{\partial b_{1,t+1}^{i}}\left(1+\rho_{2,t+2}^{i}\right)\\
&+\beta^{i}\left(1-\beta^{i}\right)\Theta^{2}\left(\left(1+\rho_{2,t+2}^{i}\right)\frac{\partial u_{2,t+2}^{i}}{\partial c_{2,t+2}^{i}}-\frac{\partial u_{2,t+2}^{i}}{\partial h_{2,t+2}^{i}}\frac{\partial x_{1,t+1}^{i}}{\partial b_{1,t+1}^{i}}\right)\frac{\partial s_{1,t+1}^{i}}{\partial s_{0,t}^{i}}\\
&-\beta^{i}\left(1-\beta^{i}\right)\Theta^{2}\frac{\partial u_{2,t+2}^{i}}{\partial h_{2,t+2}^{i}}\left(\frac{\partial x_{1,t+1}^{i}}{\partial z_{1,t+1}^{i}}-\frac{\partial u_{1,t+1}^{i}/\partial z_{1,t+1}^{i}}{\partial u_{1,t+1}^{i}/\partial c_{1,t+1}^{i}}\frac{\partial x_{1,t+1}^{i}}{\partial b_{1,t+1}^{i}}\right)\frac{\partial l_{1,t+1}^{i}}{\partial s_{0,t}^{i}}=0,
\end{aligned} \tag{A2}$$

where $\omega_{0,t}^{i}=w_{0,t}^{i}\left(1-\partial T_{0,t}^{i}/\partial y_{0,t}^{i}\right)$ and $\rho_{1,t+1}^{i}=r_{1,t+1}^{i}\left(1-\partial T_{1,t+1}^{i}/\partial I_{1,t+1}^{i}\right)$.

First Order Conditions for the Government

The Lagrangean corresponding to the optimization problem facing the government can be written as

$$
\begin{aligned}
L &= \sum_s \sum_i \Theta^s \left(u_{0,s}^i + \sum_{j=1}^{2} \Theta^j u_{j,s+j}^i \right) \\
&+ \sum_s \lambda_s \left\{ u_{0,s}^2 + \beta^2 \sum_{j=1}^{2} \Theta^j u_{j,s+j}^2 - \widehat{u}_{0,s}^2 - \beta^2 \sum_{j=1}^{2} \Theta^j \widehat{u}_{j,s+j}^2 \right\} \\
&+ \sum_s \gamma_s \{ \sum_i \left[w_{0,s}^i l_{0,s}^i + w_{1,s}^i l_{1,s}^i \right] - g_{0,s} - g_{1,s} \\
&+ K_s(1+r_s) - K_{s+1} - \sum_i \left[b_{0,s}^i + b_{1,s}^i + b_{2,s}^i \right] \} \\
&+ \sum_s \sum_{i} \mu_{0,s}^i \left\{ x_{0,s}^i - x_{0,s}^i \left(b_{0,s}^i, z_{0,s}^i, b_{1,s+1}^i, z_{1,s+1}^i, g_{0,s}, g_{1,s+1} \right) \right\} \\
&+ \sum_s \widehat{\mu}_{0,s}^2 \left\{ \widehat{x}_{0,s}^2 - \widehat{x}_{0,s}^2 \left(b_{0,s}^1, \widehat{z}_{0,s}^2, b_{1,s+1}^i, \widehat{z}_{1,s+1}^2, g_{0,s}, g_{1,s+1} \right) \right\} \\
&+ \sum_s \sum_i \mu_{1,s+1}^i \left\{ x_{1,s+1}^i - x_{1,s+1}^i \left(b_{1,s+1}^i, z_{1,s+1}^i, (x_{0,s}^i + g_{0,s})\delta + g_{1,s+1} \right) \right\} \\
&+ \sum_t \widehat{\mu}_{1,s+1}^2 \left\{ \widehat{x}_{1,s+1}^2 - \widehat{x}_{1,s+1}^2 \left(b_{1,s+1}^1, \widehat{z}_{1,s+1}^2, (\widehat{x}_{0,s}^2 + g_{0,s})\delta + g_{1,s+1} \right) \right\}. \qquad \text{(A3)}
\end{aligned}
$$

Instead of substituting the conditional commodity demand functions into the objective function, we have followed the equivalent approach of introducing the conditional commodity demand function for one of the two goods, x, as separate restrictions. Then, by using $c = b - x$, the first order conditions faced by generation t can be written as

$$
\frac{\partial L}{\partial b_{0,t}^1} = \Theta^t \frac{\partial u_{0,t}^1}{\partial c_{0,t}^1} - \lambda_t \frac{\partial \widehat{u}_{0,t}^2}{\partial \widehat{c}_{0,t}^2} - \gamma_t - \mu_{0,t}^1 \frac{\partial x_{0,t}^1}{\partial b_{0,t}^1} - \widehat{\mu}_{0,t}^2 \frac{\partial \widehat{x}_{0,t}^2}{\partial b_{0,t}^1} = 0 \qquad \text{(A4)}
$$

$$
\frac{\partial L}{\partial b_{0,t}^2} = \frac{\partial u_{0,t}^2}{\partial c_{0,t}^2} \left(\Theta^t + \lambda_t \right) - \gamma_t - \mu_{0,t}^2 \frac{\partial x_{0,t}^2}{\partial b_{0,t}^2} = 0 \qquad \text{(A5)}
$$

$$\frac{\partial L}{\partial b^1_{1,t+1}} = \Theta^{t+1}\frac{\partial u^1_{1,t+1}}{\partial c^1_{1,t+1}} - \lambda_t\beta^1\frac{\partial \widehat{u}^2_{1,t+1}}{\partial \widehat{c}^2_{1,t+1}} - \gamma_{t+1} - \mu^1_{0,t}\frac{\partial x^1_{0,t}}{\partial b^1_{1,t+1}}$$

$$-\mu^1_{1,t+1}\frac{\partial x^1_{1,t+1}}{\partial b^1_{1,t+1}} - \widehat{\mu}^2_{0,t}\frac{\partial \widehat{x}^2_{0,t}}{\partial b^1_{1,t+1}} - \widehat{\mu}^2_{1,t+1}\frac{\partial \widehat{x}^2_{1,t+1}}{\partial b^1_{1,t+1}} = 0 \tag{A6}$$

$$\frac{\partial L}{\partial b^2_{1,t+1}} = \frac{\partial u^2_{1,t+1}}{\partial c^2_{1,t+1}}\left(\Theta^{t+1} + \lambda_t\beta^2\right) - \gamma_{t+1} - \mu^2_{0,t}\frac{\partial x^2_{0,t}}{\partial b^2_{1,t+1}} - \mu^2_{1,t+1}\frac{\partial x^2_{1,t+1}}{\partial b^2_{1,t+1}} = 0 \tag{A7}$$

$$\frac{\partial L}{\partial b^1_{2,t+2}} = \Theta^{t+2}\frac{\partial u^1_{2,t+2}}{\partial c^1_{2,t+2}} - \lambda_t\beta^1\frac{\partial \widehat{u}^2_{2,t+2}}{\partial \widehat{c}^2_{2,t+2}} - \gamma_{t+2} = 0 \tag{A8}$$

$$\frac{\partial L}{\partial b^2_{2,t+2}} = \frac{\partial u^2_{2,t+2}}{\partial c^2_{2,t+2}}\left(\Theta^{t+2} + \beta^2\lambda_t\right) - \gamma_{t+2} = 0 \tag{A9}$$

$$\frac{\partial L}{\partial l^1_{0,t}} = -\Theta^t\frac{\partial u^1_{0,t}}{\partial z^1_{0,t}} + \lambda_t\frac{w^1_{0,t}}{w^2_{0,t}}\frac{\partial \widehat{u}^2_{0,t}}{\partial \widehat{z}^2_{0,t}} + \gamma_t w^1_{0,t} + \mu^1_{0,t}\frac{\partial x^1_{0,t}}{\partial z^1_{0,t}} + \widehat{\mu}^2_{0,t}\frac{w^1_{0,t}}{w^2_{0,t}}\frac{\partial \widehat{x}^2_{0,t}}{\partial \widehat{z}^2_{0,t}} = 0 \tag{A10}$$

$$\frac{\partial L}{\partial l^2_{0,t}} = -\left(\Theta^t + \lambda_t\right)\frac{\partial u^2_{0,t}}{\partial z^2_{0,t}} + \gamma_t w^2_{0,t} + \mu^2_{0,t}\frac{\partial x^2_{0,t}}{\partial z^2_{0,t}} = 0 \tag{A11}$$

$$\frac{\partial L}{\partial l^1_{1,t+1}} = -\Theta^{t+1}\frac{\partial u^1_{1,t+1}}{\partial z^1_{1,t+1}} + \lambda_t\beta^1\frac{w^1_{1,t+1}}{w^2_{1,t+1}}\frac{\partial \widehat{u}^2_{1,t+1}}{\partial z^1_{1,t+1}} + \gamma_{t+1}w^1_{1,t+1} + \mu^1_{0,t}\frac{\partial x^1_{0,t}}{\partial z^1_{1,t+1}}$$

$$+\mu^1_{1,t+1}\frac{\partial x^1_{1,t+1}}{\partial z^1_{1,t+1}} + \widehat{\mu}^2_{0,t}\frac{w^1_{1,t+1}}{w^2_{1,t+1}}\frac{\partial \widehat{x}^2_{0,t}}{\partial \widehat{z}^2_{1,t+1}} + \widehat{\mu}^2_{1,t+1}\frac{w^1_{1,t+1}}{w^2_{1,t+1}}\frac{\partial \widehat{x}^2_{1,t+1}}{\partial \widehat{z}^2_{1,t+1}} = 0 \qquad \text{(A12)}$$

$$\frac{\partial L}{\partial l^2_{1,t+1}} = -\left(\Theta^{t+1} + \lambda_t\beta^2\right)\frac{\partial u^2_{1,t+1}}{\partial z^2_{1,t+1}} + \gamma_{t+1}w^2_{1,t+1} + \mu^2_{0,t}\frac{\partial x^2_{0,t}}{\partial z^2_{1,t+1}} + \mu^2_{1,t+1}\frac{\partial x^2_{1,t+1}}{\partial z^2_{1,t+1}} = 0 \qquad \text{(A13)}$$

$$\begin{aligned}\frac{\partial L}{\partial x^1_{0,t}} &= \Theta^t\left[-\frac{\partial u^1_{0,t}}{\partial c^1_{0,t}} + \left(\Theta\frac{\partial u^1_{1,t+1}}{\partial h^1_{1,t+1}} + \Theta^2\frac{\partial u^1_{2,t+2}}{\partial h^1_{2,t+2}}\delta\right)\right] \\ &\quad +\mu^1_{0,t} - \mu^1_{1,t+1}\frac{\partial x^1_{1,t+1}}{\partial x^1_{0,t}} = 0 \qquad \text{(A14)}\end{aligned}$$

$$\begin{aligned}\frac{\partial L}{\partial x^2_{0,t}} &= -\left(\Theta^t + \lambda_t\right)\frac{\partial u^2_{0,t}}{\partial c^2_{0,t}} + \left(\Theta^t + \lambda_t\beta^2\right)\left(\Theta\frac{\partial u^2_{1,t+1}}{\partial h^2_{1,t+1}} + \Theta^2\frac{\partial u^2_{2,t+2}}{\partial h^2_{2,t+2}}\delta\right) \\ &\quad +\mu^2_{0,t} - \mu^2_{1,t+1}\frac{\partial x^2_{1,t+1}}{\partial x^2_{0,t}} = 0 \qquad \text{(A15)}\end{aligned}$$

$$\begin{aligned}\frac{\partial L}{\partial \widehat{x}^2_{0,t}} &= \lambda_t\frac{\partial \widehat{u}^2_{0,t}}{\partial \widehat{c}^2_{0,t}} - \lambda_t\beta^2\left(\Theta\frac{\partial \widehat{u}^2_{1,t+1}}{\partial \widehat{h}^2_{1,t+1}} + \Theta^2\frac{\partial \widehat{u}^2_{2,t+2}}{\partial \widehat{h}^2_{2,t+2}}\delta\right) \\ &\quad +\widehat{\mu}^2_{0,t} - \hat{\mu}^2_{1,t+1}\frac{\partial \widehat{x}^2_{1,t+1}}{\partial \widehat{x}^2_{0,t}} = 0 \qquad \text{(A16)}\end{aligned}$$

$$\frac{\partial L}{\partial x_{1,t+1}^{1}}=\Theta^{t}\left[-\Theta\frac{\partial u_{1,t+1}^{1}}{\partial c_{1,t+1}^{1}}+\Theta\frac{\partial u_{2,t+2}^{1}}{\partial h_{2,t+2}^{1}}\right]+\mu_{1,t+1}^{1}=0 \tag{A17}$$

$$\frac{\partial L}{\partial x_{1,t+1}^{2}}=\left(\Theta^{t}+\lambda_{t}\beta\right)^{2}\left(-\Theta\frac{\partial u_{1,t+1}^{2}}{\partial c_{1,t+1}^{2}}+\Theta^{2}\frac{\partial u_{2,t+2}^{2}}{\partial h_{2,t+2}^{2}}\right)+\mu_{1,t+1}^{2}=0 \tag{A18}$$

$$\frac{\partial L}{\partial \widehat{x}_{1,t+1}^{2}}=\lambda_{t}\beta^{2}\left(\Theta\frac{\partial \widehat{u}_{1,t+1}^{2}}{\partial \widehat{c}_{1,t+1}^{2}}-\Theta^{2}\frac{\partial \widehat{u}_{2,t+2}^{2}}{\partial \widehat{h}_{2,t+2}^{2}}\right)+\hat{\mu}_{1,t+1}^{2}=0 \tag{A19}$$

$$\frac{\partial L}{\partial K_{t+j}}=-\gamma_{t+j-1}+\gamma_{t+j}(1+r_{t+j})=0 \text{ for } j=1,2. \tag{A20}$$

Welfare Effects of Public Provision

If the income taxes are optimal, i.e. the first order conditions given by equations (A4)-(A20) are fulfilled, the welfare effect of increased public provision of health care services to the young generation is given by

$$\begin{aligned}\frac{\partial L}{\partial g_{0,t}} \quad = \quad & \Theta^{t}\sum_{i}\left(\Theta\frac{\partial u_{1,t+1}^{i}}{\partial h_{1,t+1}^{i}}+\Theta^{2}\frac{\partial u_{2,t+2}^{i}}{\partial h_{2,t+2}^{i}}\right)\\ & +\lambda_{t}\beta^{2}\left[\left(\Theta\frac{\partial u_{1,t+1}^{2}}{\partial h_{1,t+1}^{2}}+\Theta^{2}\frac{\partial u_{2,t+2}^{2}}{\partial h_{2,t+2}^{2}}\right)-\left(\Theta\frac{\partial \widehat{u}_{1,t+1}^{2}}{\partial \widehat{h}_{1,t+1}^{2}}+\Theta^{2}\frac{\partial \widehat{u}_{2,t+2}^{2}}{\partial \widehat{h}_{2,t+2}^{2}}\right)\right]-2\gamma_{t}\\ & -\sum_{i}\left(\mu_{0,t}^{i}\frac{\partial x_{0,t}^{i}}{\partial g_{0,t}}+\mu_{1,t+1}^{i}\frac{\partial x_{1,t+1}^{i}}{\partial g_{0,t}}\right)-\widehat{\mu}_{0,t}^{2}\frac{\partial \widehat{x}_{0,t}^{2}}{\partial g_{0,t}}-\hat{\mu}_{1,t+1}^{2}\frac{\partial \widehat{x}_{1,t+1}^{2}}{\partial g_{0,t}}.\end{aligned} \tag{A21}$$

Use equations (A4) and (A5) to solve for γ_t and substitute into equation (A21)

$$
\begin{aligned}
\frac{\partial L}{\partial g_{0,t}} \quad = \quad & \Theta^t \sum_i \left(\Theta \frac{\partial u^i_{1,t+1}}{\partial h^i_{1,t+1}} + \Theta^2 \frac{\partial u^i_{2,t+2}}{\partial h^i_{2,t+2}} - \frac{\partial u^i_{0,t}}{\partial c^i_{0,t}} \right) \\
& + \lambda_t [\left(\beta^2 \Theta \frac{\partial u^2_{1,t+1}}{\partial h^2_{1,t+1}} + \beta^2 \Theta^2 \frac{\partial u^2_{2,t+2}}{\partial h^2_{2,t+2}} - \frac{\partial u^2_{0,t}}{\partial c^2_{0,t}} \right) \\
& - \left(\beta^2 \Theta \frac{\partial \widehat{u}^2_{1,t+1}}{\partial \widehat{h}^2_{1,t+1}} + \beta^2 \Theta^2 \frac{\partial \widehat{u}^2_{2,t+2}}{\partial \widehat{h}^2_{2,t+2}} - \frac{\partial \widehat{u}^2_{0,t}}{\partial \widehat{c}^2_{0,t}} \right)] \\
& - \sum_i \left[\mu^i_{0,t} \left(\frac{\partial x^i_{0,t}}{\partial g_{0,t}} - \frac{\partial x^i_{0,t}}{\partial b^i_{0,t}} \right) + \mu^i_{1,t+1} \frac{\partial x^i_{1,t+1}}{\partial g_{0,t}} \right] \\
& - \widehat{\mu}^2_{0,t} \left(\frac{\partial \widehat{x}^2_{0,t}}{\partial g_{0,t}} - \frac{\partial \widehat{x}^2_{0,t}}{\partial b^1_{0,t}} \right) - \hat{\mu}^2_{1,t+1} \frac{\partial \widehat{x}^2_{1,t+1}}{\partial g_{0,t}}. \qquad \text{(A22)}
\end{aligned}
$$

Then, use equations (A14), (A15) and (A16) to solve for $\mu^1_{0,t}$, $\mu^2_{0,t}$ and $\hat{\mu}^1_{0,t}$, respectively, and substitute into equation (A22)

$$
\begin{aligned}
\frac{\partial L}{\partial g_{0,t}} &= \Theta^t \sum_i \left(\Theta \frac{\partial u^i_{1,t+1}}{\partial h^i_{1,t+1}} + \Theta^2 \frac{\partial u^i_{2,t+2}}{\partial h^i_{2,t+2}} - \frac{\partial u^i_{0,t}}{\partial c^i_{0,t}} \right) \\
&+ \lambda_t [\left(\beta^2 \Theta \frac{\partial u^2_{1,t+1}}{\partial h^2_{1,t+1}} + \beta^2 \Theta^2 \frac{\partial u^2_{2,t+2}}{\partial h^2_{2,t+2}} - \frac{\partial u^2_{0,t}}{\partial c^2_{0,t}} \right) \\
&- \left(\beta^2 \Theta \frac{\partial \widehat{u}^2_{1,t+1}}{\partial \widehat{h}^2_{1,t+1}} + \beta^2 \Theta^2 \frac{\partial \widehat{u}^2_{2,t+2}}{\partial \widehat{h}^2_{2,t+2}} - \frac{\partial \widehat{u}^2_{0,t}}{\partial \widehat{c}^2_{0,t}} \right)] \\
&+ \Theta^t \left[- \frac{\partial u^1_{0,t}}{\partial c^1_{0,t}} + \left(\Theta \frac{\partial u^1_{1,t+1}}{\partial h^1_{1,t+1}} + \Theta^2 \frac{\partial u^1_{2,t+2}}{\partial h^1_{2,t+2}} \right) \right] \left(\frac{\partial x^1_{0,t}}{\partial g_{0,t}} - \frac{\partial x^1_{0,t}}{\partial b^1_{0,t}} \right) \\
&+ \left[- \left(\Theta^t + \lambda_t \right) \frac{\partial u^2_{0,t}}{\partial c^2_{0,t}} + \left(\Theta^t + \lambda_t \beta^2 \right) \left(\Theta \frac{\partial u^2_{1,t+1}}{\partial h^2_{1,t+1}} + \Theta^2 \frac{\partial u^2_{2,t+2}}{\partial h^2_{2,t+2}} \right) \right] \left(\frac{\partial x^2_{0,t}}{\partial g_{0,t}} - \frac{\partial x^2_{0,t}}{\partial b^2_{0,t}} \right) \\
&+ \left[\lambda_t \frac{\partial \widehat{u}^2_{0,t}}{\partial \widehat{c}^2_{0,t}} - \lambda_t \beta^2 \left(\Theta \frac{\partial \widehat{u}^2_{1,t+1}}{\partial \widehat{h}^2_{1,t+1}} + \Theta^2 \frac{\partial \widehat{u}^2_{2,t+2}}{\partial \widehat{h}^2_{2,t+2}} \right) \right] \left(\frac{\partial \widehat{x}^2_{0,t}}{\partial g_{0,t}} - \frac{\partial \widehat{x}^2_{0,t}}{\partial b^1_{0,t}} \right) \\
&- \mu^1_{1,t+1} \left[\frac{\partial x^1_{1,t+1}}{\partial g_{0,t}} + \frac{\partial x^1_{1,t+1}}{\partial x^1_{0,t}} \left(\frac{\partial x^1_{0,t}}{\partial g_{0,t}} - \frac{\partial x^1_{0,t}}{\partial b^1_{0,t}} \right) \right] \\
&- \mu^2_{1,t+1} \left[\frac{\partial x^2_{1,t+1}}{\partial g_{0,t}} + \frac{\partial x^2_{1,t+1}}{\partial x^2_{0,t}} \left(\frac{\partial x^2_{0,t}}{\partial g_{0,t}} - \frac{\partial x^2_{0,t}}{\partial b^2_{0,t}} \right) \right] \\
&- \hat{\mu}^2_{1,t+1} \left[\frac{\partial \widehat{x}^2_{1,t+1}}{\partial g_{0,t}} + \frac{\partial \widehat{x}^2_{1,t+1}}{\partial \widehat{x}^2_{0,t}} \left(\frac{\partial \widehat{x}^2_{0,t}}{\partial g_{0,t}} - \frac{\partial \widehat{x}^2_{0,t}}{\partial b^1_{0,t}} \right) \right] \qquad \text{(A23)}
\end{aligned}
$$

Finally, use equations (A17), (A18) and (A19) to solve for $\mu^1_{1,t+1}$, $\mu^2_{1,t+1}$ and $\hat{\mu}^1_{1,t+1}$, respectively, substitute into equation (A23) and rearrange

$$
\begin{aligned}
\frac{\partial L}{\partial g_{0,t}} \;=\;& \Theta^t \sum_i \left(\Theta \frac{\partial u^i_{1,t+1}}{\partial h^i_{1,t+1}} + \Theta^2 \frac{\partial u^i_{2,t+2}}{\partial h^i_{2,t+2}} - \frac{\partial u^i_{0,t}}{\partial c^i_{0,t}} \right) \left(1 + \frac{\partial x^i_{0,t}}{\partial g_{0,t}} - \frac{\partial x^i_{0,t}}{\partial b^i_{0,t}} \right) \\
& + \lambda_t \left(\beta^2 \Theta \frac{\partial u^2_{1,t+1}}{\partial h^2_{1,t+1}} + \beta^2 \Theta^2 \frac{\partial u^2_{2,t+2}}{\partial h^2_{2,t+2}} - \frac{\partial u^2_{0,t}}{\partial c^2_{0,t}} \right) \left(1 + \frac{\partial x^2_{0,t}}{\partial g_{0,t}} - \frac{\partial x^2_{0,t}}{\partial b^2_{0,t}} \right) \\
& - \lambda_t \left(\beta^2 \Theta \frac{\partial \widehat{u}^2_{1,t+1}}{\partial \widehat{h}^2_{1,t+1}} + \beta^2 \Theta^2 \frac{\partial \widehat{u}^2_{2,t+2}}{\partial \widehat{h}^2_{2,t+2}} - \frac{\partial \widehat{u}^2_{0,t}}{\partial \widehat{c}^2_{0,t}} \right) \left(1 + \frac{\partial \widehat{x}^2_{0,t}}{\partial g_{0,t}} - \frac{\partial \widehat{x}^2_{0,t}}{\partial b^1_{0,t}} \right) \\
& + \Theta^t \sum_i \left(\Theta^2 \frac{\partial u^i_{2,t+2}}{\partial h^i_{2,t+2}} - \Theta \frac{\partial u^i_{1,t+1}}{\partial c^i_{1,t+1}} \right) \left[\frac{\partial x^i_{1,t+1}}{\partial g_{0,t}} + \frac{\partial x^i_{1,t+1}}{\partial x^i_{0,t}} \left(\frac{\partial x^i_{0,t}}{\partial g_{0,t}} - \frac{\partial x^i_{0,t}}{\partial b^i_{0,t}} \right) \right] \\
& + \lambda_t \beta^2 \left(\Theta^2 \frac{\partial u^2_{2,t+2}}{\partial h^2_{2,t+2}} - \Theta \frac{\partial u^2_{1,t+1}}{\partial c^2_{1,t+1}} \right) \left[\frac{\partial x^2_{1,t+1}}{\partial g_{0,t}} + \frac{\partial x^2_{1,t+1}}{\partial x^2_{0,t}} \left(\frac{\partial x^2_{0,t}}{\partial g_{0,t}} - \frac{\partial x^2_{0,t}}{\partial b^2_{0,t}} \right) \right] \\
& - \lambda_t \beta^2 \left(\Theta^2 \frac{\partial \widehat{u}^2_{2,t+2}}{\partial \widehat{h}^2_{2,t+2}} - \Theta \frac{\partial \widehat{u}^2_{1,t+1}}{\partial \widehat{c}^2_{1,t+1}} \right) \left[\frac{\partial \widehat{x}^2_{1,t+1}}{\partial g_{0,t}} + \frac{\partial \widehat{x}^2_{1,t+1}}{\partial \widehat{x}^2_{0,t}} \left(\frac{\partial \widehat{x}^2_{0,t}}{\partial g_{0,t}} - \frac{\partial \widehat{x}^2_{0,t}}{\partial b^1_{0,t}} \right) \right]
\end{aligned} \tag{A24}
$$

which is equation (38) in the main text. Equation (36) appears in the special care where $\beta^1 = \beta^2 = 1$.

By analogy, the welfare effect of public provision of health care services to the middle-aged can be written as

$$
\begin{aligned}
\frac{\partial L}{\partial g_{1,t+1}} \;=\;& \Theta^{t+2} \sum_i \frac{\partial u^i_{2,t+2}}{\partial h^i_{2,t+2}} + \lambda_t \beta^2 \Theta^2 \left[\frac{\partial u^2_{2,t+2}}{\partial h^2_{2,t+2}} - \frac{\partial \widehat{u}^2_{2,t+2}}{\partial \widehat{h}^2_{2,t+2}} \right] - 2\gamma_{t+1} \\
& - \sum_i \left(\mu^i_{0,t} \frac{\partial x^i_{0,t}}{\partial g_{1,t+1}} + \mu^i_{1,t+1} \frac{\partial x^i_{1,t+1}}{\partial g_{1,t+1}} \right) - \widehat{\mu}^2_{0,t} \frac{\partial \widehat{x}^2_{0,t}}{\partial g_{1,t+1}} - \hat{\mu}^2_{1,t+1} \frac{\partial \widehat{x}^2_{1,t+1}}{\partial g_{1,t+1}}.
\end{aligned} \tag{A25}
$$

Equation (47) can then be derived in the same general way as equation (38). To derive equation (47), solve equation (A7) for γ_{t+1} and substitute into equation (A25). Then, use equations (A14), (A15) and (A16) to solve for $\mu^1_{0,t}$, $\mu^2_{0,t}$ and $\hat{\mu}^1_{0,t}$, respectively, and substitute into the equation derived in the first step. Finally, in the new equation, substitute for $\mu^1_{1,t+1}$, $\mu^2_{1,t+1}$ and $\hat{\mu}^1_{1,t+1}$ by using equations (A17), (A18) and (A19), and rearrange to obtain equation (47).

References

Aronsson, T. and Granlund, D. (2011) Public Goods and Optimal Paternalism under Present-Biased Preferences. Economics Letters 113, 54-57.

Aronsson, T. and Johansson-Stenman, O. (2010) Positional Concerns in an OLG Model: Optimal Labor and Capital Income Taxation. International Economic Review 51, 1071-1095.

Aronsson, T., Markström, M. and Sjögren, T. (2005) Public Provision of Private Goods and Equilibrium Unemployment. FinanzArchiv 61, 353-367, 2005.

Aronsson, T. and Thunström, L. (2008) A Note on Optimal Paternalism and Health Capital Subsidies. Economic Letters 101, 241-242.

Aronsson, T. and Sjögren, T. (2009) Quasi-Hyperbolic Discounting and Mixed Taxation. Umeå Economic Studies 778, Umeå University.

Aronsson, T., Sjögren, T. and Dalin, T. (2009) Optimal Taxation and Redistribution in an OLG Model with Unemployment. International Tax and Public Finance 16, 198-218.

Atkinson, A. and Sandmo, A. (1980) Welfare Implications of the Taxation of Savings.Economic Journal 90, 529-548.

Bickel, W. K., Odum, A. L., and Madden, G. J. (1999) Impulsivity and cigarette smoking: Delay discounting in current, never, and ex-smokers. Psychopharmacology 146, 447–454.

Blomquist, S. and Christiansen, V. (1995) Public Provision of Private Goods as a Redistributive Device in an Optimum Income Tax Model. Scandinavian Journal of Economics 97, 547-567.

Blomquist, S. and Christiansen, V. (1998), Topping up or Opting out? The Optimal Design of Public Provision Schemes. International Economic Review 39, 399-411.

Boadway, R. and Marchand, M. (1995) The use of Public Expenditures for Redistributive Purposes. Oxford Economic Papers 47, 45-59.

Brett, C. (1997) A Note on Nonlinear Taxation in an Overlapping Generations Model. Mimeo.

Brett, C. and Weymark, J. (2008) Optimal Nonlinear Taxation of Income and Savings Without Commitment. Working paper 09-W05, Department of Economics, Vanderbilt University.

Brown, A., Chua, Z., and Camerer, C. (2009) Learning and Visceral Temptation in Dynamic Saving Experiments. Quarterly Journal of Economics 124, 197-231.

Christensen-Szalanski, J.J. (1984) Discount functions and the measurement of patients' values. Women's decisions during childbirth. Medical Decision Making 4, 47-58.

DellaVigna, S. and Malmendier, U. (2006) Paying Not to Go to the Gym. American Economic Review 96, 694–719.

DellaVigna, S. (2009) Psychology and Economics: Evidence from the Field. Journal of Economic Literature 47, 315–372.

Diamond, P. and Köszegi, B. (2004) Quasi-Hyperbolic Discounting and Retirement. Journal of Public Economics 87, 1839-1872.

Edwards, J., Keen, M. and Tuomala, M. (1994) Income Tax, Commodity Taxes and Public Good Provision: A Brief Guide. FinanzArchiv 51, 472-487.

Fredreick, S., Loewenstein, G. and O'Donoghue, T. (2002) Time Discounting and Time

Preference: A Critical Review. Journal of Economic Literature 11, 351-401.

Gine, X., Karlan, D. and Zinman, J. (2010) Put Your Money Where Your Butt Is: A Commitment Contract for Smoking Cessation. American Economic Journal: Applied Economics 2, 213-235.

Grossman, M. (1972). On the concept of health capital and the demand for health. Journal of Political Economy 80, 223-255.

Hey, J. and Lotito, G. (2009) Naive, resolute or Sophisticated? A Study of Dynamic Decision Making. Journal of Risk and Uncertainty 38, 1-25.

Laibson, D. (1997) Golden Eggs and Hyperbolic Discounting. Quarterly Journal of Economics 112, 443-477.

O'Donoghue, T. and Rabin, M. (1999) Doing it Now or Later. American Economic Review, March, 103-124.

O'Donoghue, T. and Rabin, M. (2003) Studying Optimal Paternalism, Illustrated by a Model of Sin Taxes. American Economic Review, May, 186-191.

O'Donoghue, T. and Rabin, M. (2006) Optimal Sin Taxes. Journal of Public Economics 90, 1825-1849.

Odum, A. L., Madden, G. J., and Bickel, W. K. (2002) Discounting of delayed health gains and losses by current, never-and ex-smokers of cigarettes. Nicotine & Tobacco Research 4, 295-303.

Phelps, E. and Pollak, R. (1968) On Second Best National Saving and Game Equilibrium Growth. Review of Economic Studies 35, 185-199.

Pirttilä, J. and Tenhunen, S. (2008) Pawns and Queens Revisited: Public Provision of Private Goods when Individuals Make Mistakes. International Tax and Public Finance 15, 599-616.

Pirttilä, J. and Tuomala, M. (2001) On Optimal Non-Linear Taxation and Public

Good Provision in an Overlapping Generations Economy. Journal of Public Economics 79, 485-501.

Pirttilä, J. and Tuomala, M. (2002) Publicly Provided Private Goods and Redistribution: A General Equilibrium Analysis. Scandinavian Journal of Economics 104, 173-188.

Rubinstein, A. (2003) "Economics and Psychology"? The Case of Hyperbolic Discounting. International Economic Review 44, 1207-1216.

Shapiro, J. M. (2005) Is there a daily discount rate? Evidence from the food stamp nutrition cycle. Journal of Public Economics 89, 303–325.

Skiba, P. and Tobacman, J. (2008) Payday Loans, Uncertainty, and Discounting: Explaining Patterns of Borrowing, Repayment, and Default. Vanderbilt Law and Economics Research Paper 08-33, Vanderbilt University.

Stern, N.H. (1982) Optimum Taxation with Errors in Administration. Journal of Public Economics 17, 181-211.

Stiglitz, J.E. (1982). Self-selection and Pareto efficient taxation. Journal of Public Economics 17, 213–240.

Thaler, R. (1981) Some Empirical Evidence on Dynamic Inconsistency. Economic Letters 8, 201-207.

Viscusi, K., Huber, J. and Bell, J. (2008) Estimating Discount Rates for Environmental Quality from Utility-Based Choice Experiments. Journal of Risk and Uncertainty 37, 199-220.