Present-Biased Preferences and Publicly Provided Health Care*

Thomas Aronsson[1] and David Granlund[1,2]

[1]Department of Economics, Umeå University,
SE-901 87 Umeå, Sweden

[2]The Swedish Retail Institute (HUI),
SE-103 29 Stockholm, Sweden.

September 2010

Abstract

In this paper, we analyze the welfare effects of publicly provided health care in an economy where the 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, 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, hyperbolic discounting, intertemporal model, asymmetric information.

JEL classification: D03, D61, H42.

*The authors would like to thank Tomas Sjögren for helpful comments and suggestions. Generous research grants from The Bank of Sweden Tercentenary Foundation, The Swedish Council for Working Life and Social Research, and The Swedish National Tax Board are also gratefully acknowledged.

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; both of which may have serious welfare consequences.

The present paper develops a dynamic general equilibrium model, where the consumers suffer from a self-control problem generated by quasi-hyperbolic discounting. The purpose is to analyze the welfare effects of publicly provided health care services, which exemplify private goods provided by the public sector. We present two reasons as to why this is interesting. First, as 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 the investment is made, the appearance of quasi-hyperbolic discounting implies that the investment made by the individual might become too small from the perspective of his/her future preferences. Therefore, a paternalistic government may want to make sure that agents reach an optimal level - or at least a minimum level - of health capital in a longer time-perspective. Second, there is already a literature - albeit small - dealing with optimal tax (or subsidy) responses to quasi-hyperbolic discounting (Gruber and Köszegi, 2004; O'Donoghue and Rabin, 2003, 2006; Aronsson and Thunström, 2008; Aronsson and Sjögren, 2009), while there are no earlier studies on public provision of private goods in this

[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 Rubenstein (2003) for a critical view of the evidence for hyperbolic discounting.

particular context. Our study serves to bridge this gap by considering the supplemental role of publicly provided health care services when the income tax is optimal. The only related study that we are aware of - dealing with public provision of private goods under optimal income taxation in an economy where agents suffer from bounded rationality - is Pirttilä and Tenhunen (2008), which 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 (for whatever reason). In Section 3 below, we compare our results with those derived by Pirttilä and Tenhunen.

To be more specific, we develop an overlapping generations (OLG) model, in which the consumers differ in ability, where ability is private information, and where each consumer lives for three periods (the minimum number of periods required to analyze the consequences of quasi-hyperbolic discounting). Following Stern (1982) and Stiglitz (1982), we simplify the analysis by considering a framework with two ability-types. The instantaneous utility facing each consumer depends on the current consumption of a numeraire good, the use of leisure, and the stock of health capital[2], respectively, where the latter accumulates via the consumption of a specific private good referred to as "health care" in what follows. Furthermore, we allow the two ability-types to differ with respect to the preference for "immediate gratification". The policy instruments faced by the government are nonlinear taxes on labor income and capital income as well as publicly provided health care, which the consumers may "top up" via private purchases. Therefore, the present study also relates to earlier literature on public provision of private goods under asymmetric information between the government and the private sector, where publicly provided private goods are tools - in addition to the income tax - for relaxation of the self-selection constraint (that places restraint on redistribution policy).[3] While this earlier literature is typically based on static models, we extend the analysis to a dynamic model to be able to capture the policy incentives following from a dynamically inconsistent preference structure. The two-type model constitutes a simple - yet powerful - framework for studying corrective and redistributive aspects of public policy simultaneously; an approach which is arguably realistic in the sense that a government attempting to correct

[2]By describing health as a capital concept, our model bears some resemblance to the classical health economics model developed by Grossman (1972).

[3]See e.g. Blomquist and Christiansen (1995, 1998) and Boadway and Marchand (1995).

for a behavioral failure may also want to redistribute in the most efficient way.

Earlier studies dealing with quasi-hyperbolic discounting often distinguish between naive and sophisticated consumers.[4] At any time, a naive consumer erroneously expects to be time-consistent in the future, meaning that he/she may have an incentive to revise the optimal plan in each subsequent period. A sophisticated consumer, on the other hand, recognizes that the self-control problem also arises in future periods, and implements a plan that his/her future selves will follow (see, e.g., Laibson 1997). We consider both naivety and sophistication in what follows, which is important for at least two reasons. First, it is not a priori clear whether agents in real world economies are better described by naivety than sophistication or vice versa.[5] Second, the distinction between naivety and sophistication matters for the optimal public provision of health care; to be more specific, the policy rule derived under naivety is a technical special case of the corresponding policy rule associated with sophistication.

Our study is closely related to a paper by Aronsson and Sjögren (2009), which deals with optimal mixed taxation (i.e. the optimal combination of income and commodity 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, where a distinction is made between naive and sophisticated consumers. In Section 3, we present the cost benefit rules for public provision of health care 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.

[4] See O'Donoghue and Rabin (1999) for an excellent article about the distinction between naivety and sophistication.

[5] To our knowledge, the empirical evidence here is scarce. Although in a different context than ours, Hey and Lotito (2009) analyze dynamically inconsistent decision-making, and distinguish between naive, sophisticated and resolute agents. The latter category, which is not represented in our study, means that agents do not let their inconsistency affect their behavior, i.e. the agents stick to a plan that is best from an ex-ante perspective. Based on data from an experiment, the authors find that the majority of agents were either naive or resolute (with slightly more agents being naive), whereas sophistication was a less common strategy.

2 The Model

The production side of the model follows the bulk of earlier literature on optimal taxation and public provision of private goods under asymmetric information in assuming that the output is produced 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, 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. 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

$$u^i_{0,t} = v(c^i_{0,t}, z^i_{0,t}) + f(h^i_{0,t}) \tag{1}$$

$$u^i_{1,t+1} = v(c^i_{1,t+1}, z^i_{1,t+1}) + f(h^i_{1,t+1}) \tag{2}$$

$$u^i_{2,t+2} = v(c^i_{2,t+2}, \bar{l}\,) + f(h^i_{2,t+2}), \tag{3}$$

where c denotes the consumption of a numeraire good, z leisure and h the stock of health capital. Leisure is defined as a time endowment, $\bar{l}$, less the time spent in market work, l. 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. For simplicity, we assume that the instantaneous utility is additively separable in the health capital stock.

The concept of present-biased preferences is operationalized by using the approach developed by Phelps and Pollak (1968) and later used by, for example, Laibson (1997) and O'Donoghue and Rabin (2003). The intertemporal objective of any generation t is given by

$$U_{0,t}^{i} = u_{0,t}^{i} + \beta^{i} \sum_{j=1}^{2} \Theta^{j} u_{j,t+j}^{i}, \tag{4}$$

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.[6]

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 disregarding any atemporal (within-period) relationship between the use of health care services and the stock of health capital. The health capital stock facing the young ability-type i is fixed at $h_{0,t}^i$. For the middle-aged and old, respectively, the health capital stock depends on past investments according to

$$h_{1,t+1}^{i} = h_{0,t}^{i}\delta + m_{0,t}^{i} \tag{5}$$

$$h_{2,t+2}^{i} = h_{1,t+1}^{i}\delta + m_{1,t+1}^{i} \tag{6}$$

where $\delta \in (0,1)$ is a depreciation factor - defined as "one minus the depreciation rate" - while $m_{0,t}^i$ and $m_{1,t+1}^i$ are the flow-services of health care used by the young and middle-aged selves. 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 $m_{0,t}^i = g_{0,t} + x_{0,t}^i$, where $g_{0,t}$ is the amount publicly provided and $x_{0,t}^i$ the private purchase. An analogous definition apply for the middle-aged. Notice that the government is not allowed to provide different levels

[6] It would add no important insight into the consequences of quasi-hyperbolic discounting if we were to assume that the conventional utility discount factor differs between ability-types.

of health care to the two ability-types; an assumption which is in accordance with earlier comparable literature on publicly provided private goods, 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 capital 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 $T^i_{2,t+2} = T_{2,t+2}\left(0, I^i_{2,t+2}\right)$.

The individual budget constraint is then given by

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

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

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

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 the consumers deal with their self-control problems, and we shall, therefore, make a distinction between naivety and sophistication. In the former case, the consumer does not recognize that his/her future selves are faced by the same self-control problem as the current self, whereas sophistication means that each consumer implements a time-consistent plan that his/her future selves will follow. As we show below, this means that the young sophisticated consumer will act as strategic leader vis-a-vis his/her middle-aged self. In technical terms, naivety is a special case of sophistication in the sense that the first order conditions for consumption and savings that a naive consumer obeys are special cases of those obeyed by a sophisticated consumer. 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 earlier literature on optimal nonlinear taxation under asymmetric information, 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 and x to maximize the objective represented by equation (4) subject to the nonnegativity constraint $x^i_{1,t+1} \geq 0$; the health capital function (6); and the following budget constraint:

$$b^i_{1,t+1} = c^i_{1,t+1} + x^i_{1,t+1} \quad (10)$$

$$b^i_{2,t+2} = c^i_{2,t+2}, \quad (11)$$

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

$$-\frac{\partial u^i_{1,t+1}}{\partial c^i_{1,t+1}} + \beta^i \Theta \frac{\partial u^i_{2,t+2}}{\partial h^i_{2,t+2}} = A^i_{1,t+1} \leq 0; \tag{12}$$

$$x^i_{1,t+1} A^i_{1,t+1} = 0. \tag{13}$$

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).

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

$$n^i_{1,t+1} = n^i_1 \left(b^i_{1,t+1}, z^i_{1,t+1}, (x^i_{0,t} + g_{0,t})\delta + g_{1,t+1} \right) \text{ for } n = c, x. \tag{14}$$

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.

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

$$b^i_{1,t+1} = s^i_{0,t}(1 + r_{t+1}) + w^i_{1,t+1} l^i_{1,t+1} - T^i_{1,t+1} - s^i_{1,t+1} \tag{15}$$

$$b^i_{2,t+2} = s^i_{1,t+1}(1 + r_{t+2}) - T^i_{2,t+2}. \tag{16}$$

If we define the marginal net wage rate $\omega^i_{1,t+1} = w^i_{1,t+1}\left(1 - \partial T^i_{1,t+1}/\partial y^i_{1,t+1}\right)$ and marginal net interest rate $\rho^i_{2,t+2} = r^i_{2,t+2}\left(1 - \partial T^i_{2,t+2}/\partial I^i_{2,t+2}\right)$, the first order conditions for hours of work and savings can be written as

$$\frac{\partial u^i_{1,t+1}}{\partial c^i_{1,t+1}}\omega^i_{1,t+1} - \frac{\partial u^i_{1,t+1}}{\partial z^i_{1,t+1}} = 0 \quad (17)$$

$$-\frac{\partial u^i_{1,t+1}}{\partial c^i_{1,t+1}} + \beta^i \Theta \left(1+\rho^i_{2,t+2}\right) \frac{\partial u^i_{2,t+2}}{\partial c^i_{2,t+2}} = 0. \quad (18)$$

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. where $\beta^i = 1$. 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)

$$l^i_{1,t+1} = l^i_1\left(s^i_{0,t}, m^i_{0,t}\right) \quad (19)$$

$$s^i_{1,t+1} = s^i_1\left(s^i_{0,t}, m^i_{0,t}\right). \quad (20)$$

By analogy to equation (14) above, equations (19) and (20) are also interpretable as reaction functions, showing how the young consumer may influence the labor supply and savings behavior of his/her middle-aged self.

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

$$U_{0,t}^i = u_{0,t}^i + \beta^i \Theta V_{1,t+1}^i, \tag{21}$$

where

$$V_{1,t+1}^i = u_{1,t+1}^i + \Theta u_{2,t+2}^i \tag{22}$$

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 $u_{1,t+1}^i + \beta^i \Theta u_{2,t+2}^i$). 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 $c_{0,t}^i$ and $x_{0,t}^i$ subject to equations (5)-(6), (10)-(11), (14), (15)-(16), and (19)-(20) as well as subject to the following budget constraint

$$b_{0,t}^i = c_{0,t}^i + q_t x_{0,t}^i. \tag{23}$$

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

$$\begin{aligned}
& -\frac{\partial u_{0,t}^i}{\partial c_{0,t}^i} + \beta^i \left[\Theta \frac{\partial u_{1,t+1}^i}{\partial h_{1,t+1}^i} + \Theta^2 \frac{\partial u_{2,t+2}^i}{\partial h_{2,t+2}^i} \delta \right] \\
& + \left(1 - \beta^i\right) \beta^i \Theta^2 \frac{\partial u_{2,t+2}^i}{\partial h_{2,t+2}^i} \frac{\partial x_{1,t+1}^i}{\partial x_{0,t}^i} \\
& + \left(1 - \beta^i\right) \beta^i \Theta^2 \left[\left(1 + \rho_{2,t+2}^i\right) \frac{\partial u_{2,t+2}^i}{\partial c_{2,t+2}^i} - \frac{\partial u_{2,t+2}^i}{\partial h_{2,t+2}^i} \frac{\partial x_{1,t+1}^i}{\partial b_{1,t+1}^i} \right] \frac{\partial s_{1,t+1}^i}{\partial x_{0,t}^i} \\
& - \left(1 - \beta^i\right) \beta^i \Theta^2 \frac{\partial u_{2,t+2}^i}{\partial h_{2,t+2}^i} \frac{\partial \tilde{x}_{1,t+1}^i}{\partial z_{1,t+1}^i} \frac{\partial l_{1,t+1}^i}{\partial x_{0,t}^i} = A_{0,t}^i \leq 0;
\end{aligned} \tag{24}$$

$$x_{0,t}^i A_{0,t}^i = 0 \tag{25}$$

where $\partial \tilde{x}_{1,t+1}^i / \partial z_{1,t+1}^i = \partial x_{1,t+1}^i / \partial z_{1,t+1}^i - MRS_{z,c,t+1}^i (\partial x_{1,t+1}^i / \partial b_{1,t+1}^i)$ measures the change in the conditional compensated demand for health care services following increased use of leisure, while $MRS_{z,c,t+1}^i = (\partial u_{1,t+1}^i / \partial z_{1,t+1}^i)/(\partial u_{1,t+1}^i / \partial c_{1,t+1}^i)$ is the marginal rate of substitution between leisure and the numeraire good faced by the consumer's middle-aged self.

The first row of (24) - which is analogous to (12) faced by the middle-aged agent - comprises the marginal efficiency condition for $x^i_{0,t}$ that would characterize a young naive consumer, whereas the additional 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 $\left(1-\beta^i\right)$: 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. 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).

Note that the variables $x^i_{1,t+1}$, $s^i_{1,t+1}$, and $l^i_{1,t+1}$ are decided upon simultaneously by the middle-aged consumer, and the effect that an increase $x^i_{0,t}$ will have on each of these variables is, in general, ambiguous. For purposes of interpretation, we will, nevertheless, discuss a possible scenario. First, note that the partial derivative $\partial x^i_{1,t+1}/\partial x^i_{0,t} \in (-\delta, 0)$ is likely to be relatively large in absolute value, as increased consumption of health care services when young leads to a lower marginal utility of health capital when middle-aged and old, ceteris paribus.[7] Second, if lower expenditures on health care services when middle-aged (due to an increase in $x^i_{0,t}$) means increased saving <u>and</u> increased expenditure on the numeraire good - and with $l^i_{1,t+1}$ held constant - we have $\partial s^i_{1,t+1}/\partial x^i_{0,t} \in (0, -\ \partial x^i_{1,t+1}/\partial x^i_{0,t})$. The intuition is that the stock of health capital does not directly affect the first order condition for $s^i_{1,t+1}$. In this hypothetical - let be plausible - scenario, the second and third row of (24) will sum to a negative number, i.e.

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

$$\left(1-\beta^{i}\right)\Theta\frac{\partial u_{1,t+1}^{i}}{\partial c_{1,t+1}^{i}}\left[\frac{\partial x_{1,t+1}^{i}}{\partial x_{0,t}^{i}}+\left(1-\frac{\partial x_{1,t+1}^{i}}{\partial b_{1,t+1}^{i}}\right)\frac{\partial s_{1,t+1}^{i}}{\partial x_{0,t}^{i}}\right]<0. \tag{26}$$

This discussion suggests that the second and third rows of (24) contribute to reduce $x_{0,t}^{i}$, ceteris paribus; a choice made by the young consumer to induce his/her middle-aged self to increase the consumption of health care services. This incentive may, in turn, either be reinforced or counteracted by the fourth row of (24); the sign of which depends on whether the use of health care services by the middle-aged self is complementary with, or substitutable for, leisure, i.e. whether $\partial\tilde{x}_{1,t+1}^{i}/\partial z_{1,t+1}^{i}$ is positive or negative, and how an increase in $x_{0,t}^{i}$ affects the hours of work supplied by the middle-aged ability type i. If the nonnegativity constraint does not bind, equations (23) and (25) imply the following conditional demand functions (if defined conditional on the use of leisure both when young and when middle-aged);

$$n_{0,t}^{i}=n_{0}^{i}\left(b_{0,t}^{i},z_{0,t}^{i},b_{1,t+1}^{i},z_{1,t+1}^{i},g_{0,t},g_{1,t+1}\right)\text{ for }n=c,x. \tag{27}$$

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. 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 health care.

3 Public Provision of Health Care

The government aims at redistributing as well as correcting for the self-control problem discussed above. We assume that $\beta^{1}=\beta^{2}=1$ from the perspective of the (paternalistic) government,[8] and that the government faces a utilitarian social welfare function. Therefore, the contribution by ability-type i of generation t to the social welfare function can be written

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

[8] This assumption is in line with earlier comparable literature; see, e.g., O'Donoghue and Rabin (2003, 2006), Aronsson and Thunström (2008) and Aronsson and Sjögren (2009).

Therefore, as 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

$$W = \sum_{s}\sum_{i}\Theta^{s}V_{0,s}^{i}. \tag{29}$$

For all t, the resource constraint can be written as

$$\sum_{i}\left[w_{0,t}^{i}l_{0,t}^{i} + w_{1,t}^{i}l_{1,t}^{i}\right] + K_{t}(1+r_{t}) - K_{t+1}$$
$$-\sum_{i}\left[c_{0,t}^{i} + c_{1,t}^{i} + c_{2,t}^{i} + m_{0,t}^{i} + m_{1,t}^{i} + m_{2,t}^{i}\right] = 0, \tag{30}$$

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 follow much earlier literature in concentrating a normal case, where the government wants to redistribute from the high-ability to the low-ability type. As a consequence, one would like to prevent the high-ability type from pretending to be a low-ability type, i.e. 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 $\widehat{l}_{0,t}^{2} = \left(w_{0,t}^{1}/w_{0,t}^{2}\right)l_{0,t}^{1}$ when young and by $\widehat{l}_{1,t+1}^{2} = \left(w_{1,t+1}^{1}/w_{1,t+1}^{2}\right)l_{1,t+1}^{1}$ when middle-aged. 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

$$\widehat{n}_{0,t}^2 = n_0^2\left(b_{0,t}^1, \widehat{z}_{0,t}^2, b_{1,t+1}^1, \widehat{z}_{1,t+1}^2, g_{0,t}, g_{1,t+1}\right) \text{ for } n = c, x \tag{31}$$

$$\widehat{n}_{1,t+1}^2 = n_1^2\left(b_{1,t+1}^1, \widehat{z}_{1,t+1}^2, (\widehat{x}_{0,t}^2 + g_{0,t})\delta + g_{1,t+1}\right) \text{ for } n = c, x \tag{32}$$

where, $\widehat{z}_{0,t}^2 = \bar{l} - \widehat{l}_{0,t}^2$ and $\widehat{z}_{1,t+1}^2 = \bar{l} - \widehat{l}_{1,t+1}^2$. 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

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

where the definitions of $\widehat{u}_{j,t+j}^2$ 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 $l_{0,t}^i$, $b_{0,t}^i$, $l_{1,t}^i$, $b_{1,t}^i$, $b_{2,t}^i$ (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).[9]

[9] 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 generation, as well as an optimal path for the capital stock, subject to the self-selection, health capital and resource constraints. Following earlier literature on optimal nonlinear taxation in dynamic economies, 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. The marginal income tax structure that implements the second best resource allocation can then be derived by combining the first order conditions characterizing the second best problem with those faced by the consumers.

The Lagrangean corresponding to this policy problem is presented in the Appendix together with the associated first order conditions reflecting an optimal income tax policy implemented for generation t. Our concern is then to analyze the welfare effects of publicly provided health care given that the income taxes are optimal. Note that there is a potential time-inconsistency problem here, since the government may have an incentive to change its announced policies after that the consumers have revealed their abilities (which they do at the end of the first period of life given the appropriate incentives). Although we recognize this potential problem, we follow the bulk of 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.[10]

We start by analyzing public provision of health care 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 $\beta^i = 1$. We will then return to the assumption that the consumers behave as if $\beta^i < 1$ and examine the welfare effect of publicly provided health care to the young generation under naivety as well as sophistication.

3.1.1 Without Hyperbolic Discounting

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

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

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

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

denote how $m^i_{0,t}$ and $m^i_{1,t+1}$, respectively, responds to a tax-financed increase in $g_{0,t}$. The responses by the mimicker are analogous. Then, 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

$$\begin{aligned}\frac{\partial W}{\partial g_{0,t}} &= \Theta^t\sum_i\left\{\left(\frac{\partial V^i_{0,t}}{\partial m^i_{0,t}}-\frac{\partial V^i_{0,t}}{\partial c^i_{0,t}}\right)\frac{dm^i_{0,t}}{dg_{0,t}}+\left(\frac{\partial V^i_{0,t}}{\partial m^i_{1,t+1}}-\frac{\partial V^i_{0,t}}{\partial c^i_{1,t+1}}\right)\frac{dm^i_{1,t+1}}{dg_{0,t}}\right\}\\ &+\lambda_t\left[\left(\frac{\partial V^2_{0,t}}{\partial m^2_{0,t}}-\frac{\partial V^2_{0,t}}{\partial c^2_{0,t}}\right)\frac{dm^2_{0,t}}{dg_{0,t}}-\left(\frac{\partial \widehat{V}^2_{0,t}}{\partial \widehat{m}^2_{0,t}}-\frac{\partial \widehat{V}^2_{0,t}}{\partial \widehat{c}^2_{0,t}}\right)\frac{d\widehat{m}^2_{0,t}}{dg_{0,t}}\right]\\ &+\lambda_t\left[\left(\frac{\partial V^2_{0,t}}{\partial m^2_{1,t+1}}-\frac{\partial V^2_{0,t}}{\partial c^2_{1,t+1}}\right)\frac{dm^2_{1,t+1}}{dg_{0,t}}-\left(\frac{\partial \widehat{V}^2_{0,t}}{\partial \widehat{m}^2_{1,t+1}}-\frac{\partial \widehat{V}^2_{0,t}}{\partial \widehat{c}^2_{1,t+1}}\right)\frac{d\widehat{m}^i_{1,t+1}}{dg_{0,t}}\right]\end{aligned} \quad (34)$$

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 (34). 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).

Equation (34) is just an intertemporal analogue to, and has the same interpretation as, formulas derived in earlier literature. 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 (34) vanishes because the consumer has made an optimal choice, i.e.

$$\frac{\partial V^i_{0,t}}{\partial m^i_{0,t}}-\frac{\partial V^i_{0,t}}{\partial c^i_{0,t}}=\Theta\frac{\partial u^i_{1,t+1}}{\partial h^i_{1,t+1}}+\Theta^2\frac{\partial u^i_{2,t+2}}{\partial h^i_{2,t+2}}\delta-\frac{\partial u^i_{0,t}}{\partial c^i_{0,t}}=0. \quad (35)$$

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 (34) 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 referring to the middle-aged in equation (34) 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.

Note finally that 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 (34) can be written as

$$\frac{\partial W}{\partial g_{0,t}} = \Theta^t \sum_i \left\{ \left(\frac{\partial V_{0,t}^i}{\partial m_{0,t}^i} - \frac{\partial V_{0,t}^i}{\partial c_{0,t}^i} \right) \frac{dm_{0,t}^i}{dg_{0,t}} + \left(\frac{\partial V_{0,t}^i}{\partial m_{1,t+1}^i} - \frac{\partial V_{0,t}^i}{\partial c_{1,t+1}^i} \right) \frac{dm_{1,t+1}^i}{dg_{0,t}} \right\}$$
$$+\lambda_t \left[\left(\frac{\partial U_{0,t}^2}{\partial m_{0,t}^2} - \frac{\partial U_{0,t}^2}{\partial c_{0,t}^2} \right) \frac{dm_{0,t}^2}{dg_{0,t}} - \left(\frac{\partial \widehat{U}_{0,t}^2}{\partial \widehat{m}_{0,t}^2} - \frac{\partial \widehat{U}_{0,t}^2}{\partial \widehat{c}_{0,t}^2} \right) \frac{d\widehat{m}_{0,t}^2}{dg_{0,t}} \right]$$
$$+\lambda_t \left[\left(\frac{\partial U_{0,t}^2}{\partial m_{1,t+1}^2} - \frac{\partial U_{0,t}^2}{\partial c_{1,t+1}^2} \right) \frac{dm_{1,t+1}^2}{dg_{0,t}} - \left(\frac{\partial \widehat{U}_{0,t}^2}{\partial \widehat{m}_{1,t+1}^2} - \frac{\partial \widehat{U}_{0,t}^2}{\partial \widehat{c}_{1,t+1}^2} \right) \frac{d\widehat{m}_{1,t+1}^i}{dg_{0,t}} \right] \quad (36)$$

Notice that, if the consumers have present-biased preferences, the objective function facing ability-type i, $U_{0,t}^i$, will differ from his/her contribution to the social welfare function, $V_{0,t}^i$. Note also that the components of the cost benefit rule that are associated with the self-selection constraint - the second and third rows - reflect the actual consumer-objective (not the social welfare function). Throughout the paper, we assume that $x_{0,t}^2 > x_{0,t}^1$, $\hat{x}_{0,t}^2$ when $g_{0,t} = 0$. We can then derive the following result from equation (36);

Proposition 1 *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 $x_{0,t}^i > 0$ and $x_{1,t+1}^i > 0$. This means that the first row of equation (36) is zero, because $dm_{0,t}^i/dg_{0,t} = 0$ and $dm_{1,t+1}^i/dg_{0,t} = 0$ (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 optimal level. However, the terms within parenthesis 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.

$$\frac{\partial V_{0,t}^i}{\partial m_{0,t}^i} - \frac{\partial V_{0,t}^i}{\partial c_{0,t}^i} = \left(1 - \beta^i\right) \left(\Theta \frac{\partial u_{1,t+1}^i}{\partial h_{1,t+1}^i} + \Theta^2 \frac{\partial u_{2,t+2}^i}{\partial h_{2,t+2}^i} \delta \right) > 0 \quad (37)$$
$$\frac{\partial V_{0,t}^i}{\partial m_{1,t+1}^i} - \frac{\partial V_{0,t}^i}{\partial c_{1,t+1}^i} = \left(1 - \beta^i\right) \Theta^2 \frac{\partial u_{2,t+2}^i}{\partial h_{2,t+2}^i} > 0. \quad (38)$$

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 (37) means that welfare increases via the first term on the right hand side of equation (36). This welfare increase is, in turn, partly (yet not fully) offset by the intertemporal adjustment made by the middle-aged self when $x^i_{0,t} = 0$. To see this, note that $dm^i_{1,t+1}/dg_{0,t} = \partial x^i_{1,t+1}/\partial g_{0,t}$ if $x^i_{0,t} = 0$. We can then write the first row of equation (36) as follows by combining equations (37) and (38);

$$\left(\frac{\partial V^i_{0,t}}{\partial m^i_{0,t}} - \frac{\partial V^i_{0,t}}{\partial c^i_{0,t}}\right)\frac{dm^i_{0,t}}{dg_{0,t}} + \left(\frac{\partial V^i_{0,t}}{\partial m^i_{1,t+1}} - \frac{\partial V^i_{0,t}}{\partial c^i_{1,t+1}}\right)\frac{dm^i_{1,t+1}}{dg_{0,t}}$$

$$= \left(1-\beta^i\right)\left[\Theta\frac{\partial u^i_{1,t+1}}{\partial h^i_{1,t+1}} + \Theta^2\frac{\partial u^i_{2,t+2}}{\partial h^i_{2,t+2}}\left(\delta + \frac{\partial x^i_{1,t+1}}{\partial g_{0,t}}\right)\right], \tag{39}$$

which is positive even if $\partial x^i_{1,t+1}/\partial g_{0,t}$ approaches $-\delta$.[11] Therefore, although the intertemporal adjustment effect reduces the gain of public provision, it does not eliminate it. 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. If the low-ability type is crowded out first, the welfare gain is given by equation (39) above. If, on the other hand, the mimicker is crowded out first, there is a welfare gain due to 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 (39) and a welfare loss due to a tighter self-selection constraint.

3.1.3 Sophisticated Consumers with Present-Biased Preferences

Note that equation (36) 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 clearly depend on the distinction between naivety and sophistication.

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

[11] 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, $\partial x^i_{1,t+1}/\partial g_{0,t} = \delta[\partial x^i_{1,t+1}/\partial(g_{0,t}\delta)] > -\delta$.

$$-\frac{\partial u_{0,t}^{i}}{\partial c_{0,t}^{i}}+\beta^{i}\left[\Theta\frac{\partial u_{1,t+1}^{i}}{\partial h_{1,t+1}^{i}}+\Theta^{2}\frac{\partial u_{2,t+2}^{i}}{\partial h_{2,t+2}^{i}}\delta\right]+\Gamma_{0,t}^{i}=0 \tag{40}$$

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

$$\begin{aligned}\Gamma_{0,t}^{i}&=\left(1-\beta^{i}\right)\beta^{i}\Theta^{2}\frac{\partial u_{2,t+2}^{i}}{\partial h_{2,t+2}^{i}}\frac{\partial x_{1,t+1}^{i}}{\partial x_{0,t}^{i}}\\&+\left(1-\beta^{i}\right)\beta^{i}\Theta^{2}\left[\left(1+\rho_{2,t+2}^{i}\right)\frac{\partial u_{2,t+2}^{i}}{\partial c_{2,t+2}^{i}}-\frac{\partial u_{2,t+2}^{i}}{\partial h_{2,t+2}^{i}}\frac{\partial x_{1,t+1}^{i}}{\partial b_{1,t+1}^{i}}\right]\frac{\partial s_{1,t+1}^{i}}{\partial x_{0,t}^{i}}\\&-\left(1-\beta^{i}\right)\beta^{i}\Theta^{2}\frac{\partial u_{2,t+2}^{i}}{\partial h_{2,t+2}^{i}}\frac{\partial \tilde{x}_{1,t+1}^{i}}{\partial z_{1,t+1}^{i}}\frac{\partial l_{1,t+1}^{i}}{\partial x_{0,t}^{i}}\end{aligned} \tag{41}$$

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 analogy to Proposition 1 (again under the assumption that $x_{0,t}^{2}>x_{0,t}^{1},\hat{x}_{0,t}^{2}$);

Proposition 2 *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

$$\frac{\partial V_{0,t}^{i}}{\partial m_{0,t}^{i}}-\frac{\partial V_{0,t}^{i}}{\partial c_{0,t}^{i}}=\left(1-\beta^{i}\right)\left(\Theta\frac{\partial u_{1,t+1}^{i}}{\partial h_{1,t+1}^{i}}+\Theta^{2}\frac{\partial u_{2,t+2}^{i}}{\partial h_{2,t+2}^{i}}\delta\right)-\Gamma_{0,t}^{i}>0,\text{ if }\Gamma_{0,t}^{i}\leq 0. \tag{42}$$

Note that $\Gamma_{0,t}^{i}\leq 0$ is a sufficient - not necessary - condition for the right hand side of equation (42) to be positive. As a consequence, the qualitative result indicated by Proposition 2 also applies if $\Gamma_{0,t}^{1}>0$ 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 and second row of equation (41) may under reasonable assumptions sum to a negative number. In that case, and if the third row of equation (41) is either negative or small in absolute value, then $\Gamma^i_{0,t} < 0$. The latter applies if $(\partial \tilde{x}^i_{1,t+1}/\partial z^i_{1,t+1})(\partial l^i_{1,t+1}/\partial x^i_{0,t}) \geq 0$; e.g., if leisure is substitutable for health care and $l^i_{1,t+1}$ depends negatively on $x^i_{0,t}$, or $\partial l^i_{1,t+1}/\partial x^i_{0,t}$ is close enough to zero.

Another interesting example as to when the right hand side of equation (42) 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 (36) 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 (41) 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.

As we mentioned in the introduction, 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

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

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

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, $g_{1,t+1}$. We show in the Appendix that the cost benefit rule for $g_{1,t+1}$ can be written as

$$\begin{aligned}\frac{\partial W}{\partial g_{1,t+1}} = {} & \Theta^t \sum_i \left\{ \left(\frac{\partial V^i_{0,t}}{\partial m^i_{0,t}} - \frac{\partial V^i_{0,t}}{\partial c^i_{0,t}}\right)\frac{dm^i_{0,t}}{dg_{1,t+1}} + \left(\frac{\partial V^i_{0,t}}{\partial m^i_{1,t+1}} - \frac{\partial V^i_{0,t}}{\partial c^i_{1,t+1}}\right)\frac{dm^i_{1,t+1}}{dg_{1,t+1}} \right\} \\ & + \lambda_t \left[\left(\frac{\partial U^2_{0,t}}{\partial m^2_{0,t}} - \frac{\partial U^2_{0,t}}{\partial c^2_{0,t}}\right)\frac{dm^2_{0,t}}{dg_{1,t+1}} - \left(\frac{\partial \widehat{U}^2_{0,t}}{\partial \widehat{m}^2_{0,t}} - \frac{\partial \widehat{U}^2_{0,t}}{\partial \widehat{c}^2_{0,t}}\right)\frac{d\widehat{m}^2_{0,t}}{dg_{1,t+1}} \right] \\ & + \lambda_t \left[\left(\frac{\partial U^2_{0,t}}{\partial m^2_{1,t+1}} - \frac{\partial U^2_{0,t}}{\partial c^2_{1,t+1}}\right)\frac{dm^2_{1,t+1}}{dg_{1,t+1}} - \left(\frac{\partial \widehat{U}^2_{0,t}}{\partial \widehat{m}^2_{1,t+1}} - \frac{\partial \widehat{U}^2_{0,t}}{\partial \widehat{c}^2_{1,t+1}}\right)\frac{d\widehat{m}^2_{1,t+1}}{dg_{1,t+1}} \right] \end{aligned} \quad (43)$$

By analogy to the analysis of public provision to the young generation carried out above, we assume here that $x^2_{1,t+1} > x^1_{1,t+1}, \hat{x}^2_{1,t+1}$ without any public provision of health care. We can then use equation (43) to derive the following result;

Proposition 3 *When 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 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 (43) reduces to read

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

By analogy, if the middle-aged mimicker becomes crowded out before the middle-aged low-ability type, equation (43) simplifies to read

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

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 does no longer apply. 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 (43) 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 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.[12]

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 health care services. Health care exemplifies a private good with an

[12] Note also that when a sophisticated consumer becomes crowded out as middle-aged, his/her strategic motives to hold down the investment in health care when young will vanish. As a consequence, it is possible that he/she may actually increase the investment when young in response to being crowded out as middle-aged. Therefore, even thought it seems likely that the effect that $m^i_{1,t+1}$ has on the future marginal utility of health is more important for the choice of $x^i_{0,t}$ than the strategic incentive, this mechanism will, nevertheless, prevent us from concluding that $dx^i_{0,t}/dg_{1,t+1} = dm^i_{0,t}/dg_{1,t+1} < 0$. In addition, it is possible that naive consumers will not alter their consumption choices when young in response to being crowded out as middle-aged. The reason is that a young naive consumer erroneously expects his/her middle-aged self to demand more health care services than the middle-aged self actually does.

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 investment made by the individual might be too small from the perspective of his/her future preferences, which provides a paternalistic motive for public provision. The policy instruments faced by the government are nonlinear 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. A naive consumer erroneously expects to be time-consistent in the future, meaning that he/she may have an incentive to revise the consumption plan in each subsequent period, whereas a sophisticated consumer recognizes that the self-control problem also arises in future periods and implements a plan that his/her future selves will follow. 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 increase or decrease the demand for health care as young. If the strategic incentives contribute to reduce the demand for private health care 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, 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 provide a strong case 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, the degree to which the preferences are "present-biased" may not only vary over productivity types (as we assume here); it may also vary in other dimensions. Such a change of assumption is likely to affect the policy incentives underlying publicly provided private goods (as well as the use of other policy instruments). 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

$$\frac{\partial u^i_{0,t}}{\partial c^i_{0,t}}\omega^i_{0,t} - \frac{\partial u^i_{0,t}}{\partial z^i_{0,t}} = 0 \tag{A1}$$

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

$$+\beta^i \left(1 - \beta^i\right) \Theta^2 \frac{\partial u_{2,t+2}^i}{\partial h_{2,t+2}^i} \frac{\partial x_{1,t+1}^i}{\partial b_{1,t+1}^i} \left(1 + \rho_{2,t+2}^i\right)$$

$$+\beta^i \left(1 - \beta^i\right) \Theta^2 \left(\left(1 + \rho_{2,t+2}^i\right) \frac{\partial u_{2,t+2}^i}{\partial c_{2,t+2}^i} - \frac{\partial u_{2,t+2}^i}{\partial h_{2,t+2}^i} \frac{\partial x_{1,t+1}^i}{\partial b_{1,t+1}^i} \right) \frac{\partial s_{1,t+1}^i}{\partial s_{0,t}^i}$$

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

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

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

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

$$\frac{\partial L}{\partial b^1_{0,t}} = \Theta^t \frac{\partial u^1_{0,t}}{\partial c^1_{0,t}} - \lambda_t \frac{\partial \widehat{u}^2_{0,t}}{\partial \widehat{c}^2_{0,t}} - \gamma_t - \mu^1_{0,t} \frac{\partial x^1_{0,t}}{\partial b^1_{0,t}} - \widehat{\mu}^2_{0,t} \frac{\partial \widehat{x}^2_{0,t}}{\partial b^1_{0,t}} = 0 \tag{A4}$$

$$\frac{\partial L}{\partial b^2_{0,t}} = \frac{\partial u^2_{0,t}}{\partial c^2_{0,t}} \left(\Theta^t + \lambda_t\right) - \gamma_t - \mu^2_{0,t} \frac{\partial x^2_{0,t}}{\partial b^2_{0,t}} = 0 \tag{A5}$$

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

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

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

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

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

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

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

$$\begin{aligned}\frac{\partial L}{\partial l^1_{1,t+1}} &= -\Theta^{t+1}\frac{\partial u^1_{1,t+1}}{\partial z^1_{1,t+1}} + \lambda_t\beta^1\frac{w^1_{1,t+1}}{w^2_{1,t+1}}\frac{\partial \widehat{u}^2_{1,t+1}}{\partial z^1_{1,t+1}} + \gamma_{t+1}w^1_{1,t+1} + \mu^1_{0,t}\frac{\partial x^1_{0,t}}{\partial z^1_{1,t+1}} \\ &+\mu^1_{1,t+1}\frac{\partial x^1_{1,t+1}}{\partial z^1_{1,t+1}} + \widehat{\mu}^2_{0,t}\frac{w^1_{1,t+1}}{w^2_{1,t+1}}\frac{\partial \widehat{x}^2_{0,t}}{\partial \widehat{z}^2_{1,t+1}} + \widehat{\mu}^2_{1,t+1}\frac{w^1_{1,t+1}}{w^2_{1,t+1}}\frac{\partial \widehat{x}^2_{1,t+1}}{\partial \widehat{z}^2_{1,t+1}} = 0\end{aligned} \tag{A12}$$

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

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

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

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

$$+\widehat{\mu}_{0,t}^{2} - \hat{\mu}_{1,t+1}^{2} \frac{\partial \widehat{x}_{1,t+1}^{2}}{\partial \widehat{x}_{0,t}^{2}} = 0 \qquad \text{(A16)}$$

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

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

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

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

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

$$
\begin{aligned}
\frac{\partial L}{\partial g_{0,t}} &= \Theta^t \sum_i \left(\Theta \frac{\partial u^i_{1,t+1}}{\partial h^i_{1,t+1}} + \Theta^2 \frac{\partial u^i_{2,t+2}}{\partial h^i_{2,t+2}} \right) \\
&\quad + \lambda_t \beta^2 \left[\left(\Theta \frac{\partial u^2_{1,t+1}}{\partial h^2_{1,t+1}} + \Theta^2 \frac{\partial u^2_{2,t+2}}{\partial h^2_{2,t+2}} \right) - \left(\Theta \frac{\partial \widehat{u}^2_{1,t+1}}{\partial \widehat{h}^2_{1,t+1}} + \Theta^2 \frac{\partial \widehat{u}^2_{2,t+2}}{\partial \widehat{h}^2_{2,t+2}} \right) \right] - 2\gamma_t \\
&\quad - \sum_i \left(\mu^i_{0,t} \frac{\partial x^i_{0,t}}{\partial g_{0,t}} + \mu^i_{1,t+1} \frac{\partial x^i_{1,t+1}}{\partial g_{0,t}} \right) - \widehat{\mu}^2_{0,t} \frac{\partial \widehat{x}^2_{0,t}}{\partial g_{0,t}} - \hat{\mu}^2_{1,t+1} \frac{\partial \widehat{x}^2_{1,t+1}}{\partial g_{0,t}}. \qquad \text{(A21)}
\end{aligned}
$$

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

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

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[\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) \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] \quad \text{(A24)}
\end{aligned}
$$

which is equation (36) in the main text. Equation (34) 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

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

Equation (43) can then be derived in the same general way as equation (36). To derive equation (43), 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 (43).

References

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. 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.

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.

Fredreick, S., Loewenstein, G. and O'Donoghue, T. (2002) Time Discounting and Time Preference: A Critical Review. *Journal of Economic Literature* **11**, 351-401.

Gruber, J. and Köszegi, B. (2004) Tax incidence when Individuals are Time-Inconsistent: the Case of Cigarette Excise Taxes. *Journal of Public Economics* **88**, 1959-1987

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.

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.

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

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.