Economics Working Papers

2012-12

# Optimal Tax-Transfer Policies, Life-Cycle Labour Supply and Present-Biased Preferences

Lasse Frisgaard Gunnersen and Bo Sandemann Rasmussen

As of January 2012 ECON-ASB working papers are included in Economics Working Papers

AARHUS UNIVERSITY
BUSINESS AND SOCIAL SCIENCES
DEPARTMENT OF ECONOMICS AND BUSINESS

# Optimal Tax-Transfer Policies, Life-Cycle Labour Supply and Present-Biased Preferences*

Lasse Frisgaard Gunnersen† Bo Sandemann Rasmussen‡

May 16, 2012

**Abstract**

Using a two-period model with two types of agents that are characterized by present-biased preferences second-best optimal tax-transfer policies are considered. The paternalistic optimal tax-transfer policy has two main concerns: Income redistribution from high to low ability households and correction of undersaving due to present-biasedness. Policies must comply with incentive-compatibility constraints that restricts both how much income redistribution that can take place and how much savings should be subsidized. A main result is that the degree of present-biasedness has important consequences not only for optimal subsidies to savings but also for optimal marginal income taxes.

*JEL classification*: H21, H23, H24.
*Keywords*: Optimal tax-transfer policy, paternalistic government, age-dependent taxes, labour supply, present-biasedness, redistribution.

# 1 Introduction

The rationales for fiscal intervention through tax-transfer policies may be many. In economies where agents are heterogeneous due to differences in abilities there is the standard call for income redistribution from high- to low-productivity agents as described in the seminal work of Mirrlees (1971). A more recent line of analysis

*Comments from Torben M. Andersen, David Dreyer Lassen, Thomas Sjögren and Allan Sørensen are gratefully acknowledged.

†Department of Economics and Business, University of Aarhus, Building 1322, DK-8000 Aarhus C, Denmark. e-mail: lgunnersen@econ.au.dk.

‡Department of Economics and Business, University of Aarhus, Building 1322, DK-8000 Aarhus C, Denmark. e-mail: brasmussen@econ.au.dk.

considers how agents with self-control problems tend to undersave because their preferences give rise to time-inconsistent choices with a preference for immediate gratification (this literature in mainly based on the insights of Laibson (1997)). If the government is aware of the self-control problems of private households the tax-transfer policies can be used to correct the insufficient savings of households by subsidizing savings. Of course, the interesting case is when there is both a desire to redistribute income due to heterogeneity in abilities and a desire to correct for undersaving due to present-biased preferences of private agents. In such a setting it becomes important to study how the attempts to cater both concerns simultaneously give rise to interdependencies among the various tax policy instruments.

This is the point of departure of the present paper. We take an economy with two types of households both facing self-control problems. Households differ only in productivity and following the standard assumptions in optimal tax theory these productivities are private information. Hence, with an imperfectly informed government that only observes market transactions of agents, a second-best optimal tax-transfer policy must satisfy incentive-compatibility constraints. Stiglitz (1982) considers a simple two-type model where the government would like to discriminate between high and low ability types but observes only the earned incomes of individuals. Following the revelation principle (see e.g. Salanié (2005)) a second-best optimal policy must induce individuals to reveal their true type so that ex post the government is aware of the productivities of individuals. However, it is not costless to induce self-selection, and in the second-best world marginal rates of substitution are typically distorted. In particular, low ability households face a strictly positive marginal tax on earned income, while there is no distortion at the top. Brett and Weymark (2005) extend the Stiglitz-model to a two period economy where agents work in both periods. They assume that agents and government use the same discount factor, and consequently they find that the optimal tax on savings is zero.

When agents have present-biased preferences this is no longer true. It may be optimal for the government to induce households to save more than they want to and although it is a distortion of the intertemporal allocation of consumption, agents will ex post be grateful to the government for the induced additional savings. Cremer et al. (2009) consider a two-period model where agents differ in both degree of present-biasedness and in earnings ability, where consumption, labour supply and savings decisions are made in the first period while savings are consumed in the second period. They show that a paternalistic solution does not necessarily imply forced savings of the myopics as paternalistic considerations could be outweighed by incentive effects.

Tenhunen and Tuomala (2010) also use a two period set up and assume that productivities and subjective discount factors are correlated (and both are unknown ex ante to the government). They extend the Stiglitz-set up by introducing more than two types of households and focus on how optimal policies affect the distribution of consumption and labour supply across the different types of individuals.

In the present paper we also consider a two-period model but allow households to supply labour in both periods. Contrary to Cremer et al. (2009), we assume that all households have identical present-biased preferences whereas we stick to the usual assumption of having both high- and low-productivity households. By assuming a declining age-profile of productivities the majority of lifetime income becomes concentrated in the first period implying a desire for positive savings for consumption smoothing purposes. Still, with positive labour supply at old age we can study the interaction of the possible subsidies granted to savings and the disincentive effects from labour income redistribution on old age labour supply. The government designs a non-linear tax-transfer policy based on observables: Labour income and savings. Moreover, since the age of households is perfectly observable it is possible to specify age-dependent tax-transfer schedules. The desirability of using tags that are correlated with the unobserved productivity in optimal taxation has been recognized since Akerlof (1978). More recently Blomquist and Micheletto (2008) have shown that age-dependent tax schedules Pareto-dominates age-independent taxation, and we build on these insights by explicitly using age-dependent tax schedules.

In this model set-up the optimal tax-transfer policy has to deal with two main issues: Induce more household savings and redistribute from high- to low-productivity agents. With myopic households we find that it is always optimal to distort the intertemporal marginal rate of substitution to induce agents to favour future consumption more. This holds both in a first-best and a second-best environment. But when self-selection must be satisfied, we find that it is optimal to force savings of the low-productivity households relatively more than that of the high-productivity household (a similar result is found by Tenhunen and Tuomala (2010) in a setting where subjective (present-biased) discount rates are correlated with productivities). As a consequence, in a second-best environment, the high-ability households save too little and low-ability households save too much. The explanation of this result is that is follows from the incentive-constraint: When low ability households are forced to save more the allocation of the low-productivity household becomes less attractive to the high ability household (because of his present-biased preferences), so the optimal intertemporal wedge - and hence the optimal marginal subsidy to savings - is larger for the low-ability household.

It is not only the intertemporal decision that is distorted (in an asymmetric way), the temporal decisions are as well. Given that savings are observed, a high ability household that wishes to mimick a low ability household must do so in the first period. With only one incentive constraint we obtain the Stiglitz-result that the marginal income tax of the high ability households is zero while it is positive for low ability households. Since the degree of present-biasedness affects the optimal allocation, it implicitly also affects the optimal marginal taxes on low incomes.

As usual in this literature there is limited information to be derived from the analytical solution to the second-best optimal tax-transfer problem. Hence, doing numerical simulations of the model is useful for illustrating how the design of the optimal tax-transfer scheme depends on keys parameters of the model. One of the numerical results that is surprising at a first glance is that the optimal marginal tax on income of low ability young agents seems to increase while that of the low ability old agents decreases when the degree of present-biasedness become more pronounced. Again, this effect can be explained by the incentive constraint: For a higher degree of present-biasedness, ceteris paribus, the allocation of the low ability young becomes *more* attractive to a mimicking high ability household, while at the same time the allocation of the low ability old household becomes less attractive, necessitating a larger distortion of low ability young, but allowing for a smaller distortion of the old households. Hence, there seems to be important implications for the optimal marginal income taxes at different points in time of how strong the present-biasedness of household preferences is.

The rest of the paper is organized as follows. Section 2 presents the basic model while section 3 describes the optimal tax problem. The analytical solution to the second-best optimal tax problem is derived in section 4 where after numerical solutions are considered in section 5. Section 6 provides a discussion of the results derived for both the analytical and numerical parts while section 7 discusses extensions and generalizations of the main model of the paper. Finally, section 8 offers some concluding remarks.

# 2 The Model

Consider a two-period model with two types of households differing in productivity only and where productivities may change from young to old age, so $w_j^i$ is labour productivity of a type $i \in \{H, L\}$ (high, low) of a household of age $j \in \{y, o\}$ (young, old) with $w_j^H > w_j^L$, $j \in \{y, o\}$. Young households supply labour, enjoy consumption of goods and save and to allow for present-biased preferences the discount factor governing these choice is less than the "true" or long-run discount factor. Old households also supply labour and consume, and the attitudes towards the leisure-consumption trade-off generally differ between young and old households. This will allow for different labour supply responses to taxes and transfers of young and old households. The technology is linear in labour inputs with constant marginal productivities.

The government chooses a tax-transfer policy that focuses on two issues: It aims at redistributing income from high- to low-productivity households and at correcting the distortions in intertemporal consumption allocations that follows from the present-biased nature of household preferences. To what extent it succeeds in its endeavours depends on the information available to the government.

Under perfect information about all household attributes and choices - like productivity and individual labour supply - the government would be able to implement the first-best allocation where the tax-transfer policy avoids introducing distortionary effects on market choices.

In the more likely scenario the government is imperfectly informed about productivity and individual labour supply in which case the tax-transfer policy must be conditioned on other variables that can be observed. This leads to the traditional second-best policies where tax-transfer schemes are based on labour income and possibly other observable household variables. Quite non-controversially the age of each household is assumed to be observable[1] while it is less obvious whether or not individual savings should be regarded as perfectly observable to the government. In some parts of this literature, e.g. Blomquist and Micheletto (2008), it is argued that individual households should be able to conceal the true level of their savings if they desired to do so, and under that assumption the possible strategies of high-productivity household attempting to avoid the redistributing efforts of the government becomes quite complicated, especially in a set-up where households supply labour at both ages. To avoid these complications we follow Cremer et al. (2009) by assuming that individual savings can be observed. Hence, under the second-best tax-transfer policy the net tax liability of a household depends on its labour income, savings and age.

## 2.1 Households

To make the model manageable an additively separable preference relation for households is specified. Household decisions are governed by

$$\mathcal{U}\left(c_y^i, c_o^i, l_y^i, l_o^i\right) = u\left(c_y^i\right) - v\left(l_y^i\right) + \beta\left(U\left(c_o^i\right) - V\left(l_o^i\right)\right), \quad i \in \{H, L\}, \tag{1}$$

where $c_j^i$ and $l_j^i$ are consumption and labour supply of type $i \in \{H, L\}$ for a household of age $j \in \{y, o\}$, while $\beta \in \left]0, 1\right[$ is the discount factor reflecting the present-biasedness of households, while the "true" discount factor is set equal to unity for simplicity.[2] Subutility functions are denoted by lower-case letters for

[1] The possible desirability of using tags - like age - in optimal tax schemes has been know since Akerlof (1978) and has been used e.g. by Blomquist and Micheletto (2008) in an OLG-setting that shares some similarities with our model.

[2] Since we only have two periods we do not present household discounting as quasi-hyperbolic. For a proper quasi-hyperbolic discounting model we would need at least three periods (see section

choices of the young and capital letters for the old with $u' > 0$, $u'' < 0$, $U' > 0$, $U'' < 0$, $v' > 0$, $v'' > 0$, $V' > 0$ and $V'' > 0$.

Assuming for simplicity a zero interest rate the budget constraints of household $i$ become

$$c_y^i = w_y^i l_y^i - s^i - \mathcal{T}_y(\cdot), \quad i \in \{H, L\} \tag{2}$$
$$c_o^i = w_o^i l_o^i + s^i - \mathcal{T}_o(\cdot), \quad i \in \{H, L\}\,, \tag{3}$$

where $\mathcal{T}_y(\cdot)$ and $\mathcal{T}_o(\cdot)$ are the net taxes paid by a young and an old household, respectively.

## 2.2 Government

The government chooses its tax-transfer policy to maximize a measure of social welfare. Following much of the literature the social welfare function is taken to be utilitarian and it is furthermore assumed that the government aims at correcting the "mistakes" made by households due to their present-biased preferences. Hence, as in Cremer et al. (2009) we take a paternalistic approach by which the government evaluates outcomes using the "true" household preference relation which is the household utility function in (1) evaluated at $\beta = 1$,

$$\mathcal{U}^*\left(c_y^i, c_o^i, l_y^i, l_o^i\right) = u\left(c_y^i\right) - v\left(l_y^i\right) + \left(U\left(c_o^i\right) - V\left(l_o^i\right)\right), \quad i \in \{H, L\}\,. \tag{4}$$

Of course, in the design of the optimal policy the policy-maker acknowledges that market behaviour reflects present-biased choices so part of the role of the optimal tax-transfer policy is to correct for the present-biased choices of the households regarding savings and the timing of labour supply. Another role for the tax-transfer policy is to redistribute in the classical way from high-productivity to low-productivity households.

Denoting population shares by $\pi^i \in \,]0, 1[$, $i \in \{H, L\}$ so $\pi^H + \pi^L = 1$, the paternalistic, utilitarian government objective function reads

$$W = \sum_{i \in \{H,L\}} \pi^i \left(u\left(c_y^i\right) - v\left(l_y^i\right) + U\left(c_o^i\right) - V\left(l_o^i\right)\right), \tag{5}$$

while the government budget constraint reads

$$\sum_{i \in \{y,o\}} \sum_{j \in \{H,L\}} \pi^i \mathcal{T}_j(\cdot) = 0. \tag{6}$$

7 for more details).

Given the linear technology the aggregate resource constraint reads

$$\sum_{i\in\{H,L\}}\sum_{j\in\{y,o\}}\pi^i c_j^i = \sum_{i\in\{H,L\}}\sum_{j\in\{y,o\}}\pi^i w_j^i l_j^i, \tag{7}$$

and by Walras' Law the government budget constraint will be redundant once the aggregate resource constraint is imposed. Hence, in the policy problems to follow we will omit the government budget constraint from the government optimal tax-transfer problem.

### 2.2.1 First-Best Optimum

As a reference point consider the case of perfect information where optimal policy intervention allows for the first-best allocation to be implemented.

**Proposition 1** *At the first-best solution* $c_y^L = c_y^H, c_o^L = c_o^H, l_y^L < l_y^H$ *and* $l_o^L < l_o^H$.

**Proof.** See Appendix A. ■

Given the assumption of additively separable preferences, the paternalistic utilitarian solution to the first-best problem implies that consumption levels at all ages are equalized across ability types (but generally not over the life-cycle) and that at all ages the more able households supply more labour than less able households of the same age.[3]

Designing a tax-transfer scheme that decentralizes this allocation has two distinct features. First, a lump-sum element is needed in order to transfer income from high- to low-productivity households (utilitarianism). Second, incentives must be provided to induce myopic households to save the amount desired by the social planner (paternalism). This second feature can be obtained either by subsidizing savings or by a public pension benefit scheme (forced savings).

# 3 Second-Best Tax-Transfer Policy Problem

With imperfect information about household endowments of labour productivity and household labour supply, the policy-maker cannot implement the first-best

[3] This leads to a "curse of being productive": The high productivity households work more than the low productivity households while both types have the same consumption levels. Therefore, the low productivity households are strictly better off than the high productivity households. This result follows from the combination of additively linear individual utility functions and a utilitarian social welfare function. If e.g. the social welfare function were strictly concave in individual utilities (like with a CES social welfare function) the first-best allocation would involve less redistribution and smaller differences in labour supplies between high- and low-productivity households.

outcome (since the lump-sum tax required to obtain that is conditional on labour productivities). In stead, the second-best tax-transfer policy must be based on observables, in our case being labour income, savings and age. Hence, the second-best tax functions are specified as

$$\mathcal{T}_j^i = T_j(I_j^i, s^i), \quad i \in \{H, L\}\,,\ j \in \{y, o\}\,, \tag{8}$$

where $I_j^i \equiv w_j^i l_j^i$ is pre-tax labour income of a type $i$ household of age $j$, and $T_j(I_j^i, s^i)$ is a general non-linear tax-transfer function. Marginal tax rates on labour income and savings are defined as

$$T_{j,I}^i \equiv \frac{\partial T_j(I_j^i, s^i)}{\partial I_j^i}, \quad j \in \{y, o\}\,, i \in \{H, L\} \tag{9}$$

$$T_{j,s}^i \equiv \frac{\partial T_j(I_j^i, s^i)}{\partial s^i}, \quad j \in \{y, o\}\,, i \in \{H, L\}\,. \tag{10}$$

## 3.1 Household Behaviour

Household behaviour is determined by the present-biased preferences where consumption and labour supply when old are valued using the effective discount factor $\beta < 1$. The household optimization problem for a type $i$ household is

$$\max_{c_y^i, l_y^i, c_o^i, l_o^i} u\left(c_y^i\right) - v\left(l_y^i\right) + \beta\left(U\left(c_o^i\right) - V\left(l_o^i\right)\right)$$

$$s.t.$$

$$\sum_{j \in \{y,o\}} c_j^i = \sum_{j \in \{y,o\}} \left(w_j^i l_j^i - T_y(w_j^i l_j^i, s^i)\right).$$

The first-order conditions from this optimization problems can be expressed as

$$\frac{v'(l_y^i)}{u'(c_y^i)} = w_y^i(1 - T_{y,I}^i) \tag{11}$$

$$\frac{V'(l_o^i)}{U'(c_o^i)} = w_o^i(1 - T_{o,I}^i) \tag{12}$$

$$\frac{u'(c_y^i)}{U'(c_o^i)} = \beta\frac{1 - T_{o,s}^i}{1 + T_{y,s}^i}, \tag{13}$$

where the first two conditions are the static optimality conditions determining consumption-labour supply combinations when young and old, while the last condition is the intertemporal optimality condition.[4] Obviously, the present-biasedness

[4] The intertemporal optimality condition is here stated in terms of the optimal sequence of consumption, but it could equally well be stated in terms of the optimal sequence of labour supplies.

($\beta < 1$) only has a direct effect on the intertemporal decisions and does so by making households saving too little and supplying too much labour when they are old (relative to their labour supply when young).

## 3.2 Incentive-Compatibility Constraints

When governments engage in second-best redistribution the policy choices must respect incentive-compatibility constraints. In principle there are four different types of households (young or old, high or low productivity), but since age is observable and the tax-transfer system is explicitly age-dependent a young high-productivity household can only mimic the young low-productivity household and not any of the old types of households. Hence, the incentive constraints should be stated in present discounted values using the present-biased preferences, and with redistribution from the high- to the low-productivity households, it will be the incentive constraint related to the high-productivity household that will be binding.

The relevant incentive constraint is then that the lifetime utility of a high-productivity household when choosing the consumption-income bundles that the policy-maker intends for him is no less than the lifetime utility obtained when choosing the consumption-income bundles intended for the low-productivity household. Rewriting the utility function in terms of observables, $c_j^i$ and $I_j^i$, $i \in \{H, L\}$ and $j \in \{y, o\}$, the incentive compatibility constraint reads

$$u\left(c_y^H\right) - v\left(\frac{I_y^H}{w_y^H}\right) + \beta\left[U\left(c_o^H\right) - V\left(\frac{I_o^H}{w_o^H}\right)\right] \geq u\left(c_y^L\right) - v\left(\frac{I_y^L}{w_o^H}\right) + \beta\left[U\left(c_o^L\right) - V\left(\frac{I_o^L}{w_o^H}\right)\right] \quad (14)$$

By assuming perfect foresight regarding age-dependent productivities there is no individual uncertainty and agents' true types are revealed in the first period (given incentive compatibility).[5]

[5] The '*New Dynamic Public Finance*' literature pioneered by Narayana Kocherlakota (e.g. Kocherlakota (2010)) allows skills to change over time with probabilities that are common knowledge. Thus, in such a framework it is possible to have a household that has low productivity in the first period but has high productivity in the second. Obviously, this allows a vast number of mimicker strategies, but to focus the analysis on the joint taxation of savings and labour income with potential undersavings, we abstract from such sofisticated specifications of the model here. In other words, the optimal tax design problem remains an adverse selection problem à la Mirrlees (1971) rather than repeated moral hazard problem.

# 4 Second-Best Tax-Transfer Policy

To obtain the second-best optimal tax-transfer policy we first derive the second-best optimal allocation and then consider how the optimal allocation can be implemented by the tax-transfer system consisting of the age-dependent net tax functions $T_j(I_j^i, s^i)$ for $i \in \{H, L\}$ and $j \in \{y, o\}$.

## 4.1 The Second-Best Allocation

The Lagrangian $\mathcal{L}$ associated with the finding the second-best allocation in terms of observables $I_i^j$ and $c_i^j$, ($i \in \{H, L\}$ and $j \in \{y, o\}$) is then

$$\begin{aligned}\mathcal{L} = \sum_{i\in\{L,H\}} \pi^i \left\{u\left(c_y^i\right) - v\left(\frac{I_y^i}{w_y^i}\right) + \left(U\left(c_o^i\right) - V\left(\frac{I_o^i}{w_o^i}\right)\right)\right\} \\ + \lambda\left[u\left(c_y^H\right) - v\left(\frac{I_y^H}{w_y^H}\right) + \beta\left[U\left(c_o^H\right) - V\left(\frac{I_o^H}{w_o^H}\right)\right] - u\left(c_y^L\right)\right. \\ \left. + v\left(\frac{I_y^L}{w_y^H}\right) - \beta\left[U\left(c_o^L\right) - V\left(\frac{I_o^L}{w_o^H}\right)\right]\right] \\ + \mu \sum_{j\in\{y,o\}} \sum_{i\in\{L,H\}} \pi^i \left(I_j^i - c_j^i\right) \end{aligned} \quad (15)$$

where $\lambda$ and $\mu$ are the multipliers associated with the incentive compatibility and the resource constraints, respectively. The first order conditions to this problem are given in the appendix.

### 4.1.1 Labour Supply for High-Productivity Households

Looking into the first order conditions for optimal consumption and labour supply of high-productivity young individuals (equations (58b) and (58f) in the appendix) we find that the marginal rate of substitution of income for consumption is 1

$$\frac{v'\left(\frac{I_y^H}{w_y^H}\right)}{u'\left(c_y^H\right)} \frac{1}{w_y^H} = 1. \quad (16)$$

In other words, at the margin there is no distortion of the consumption-leisure choice on high-productivity young households. The same applies for high-productivity old households using equations (58d) and (58h) in the appendix,

$$\frac{V'\left(\frac{I_o^H}{w_o^H}\right)}{U'\left(c_o^H\right)} \frac{1}{w_o^H} = 1. \quad (17)$$

The no-distortion result of the labour supply of the high-productivity households follow the standard results in the literature: There should be no distortion of labour supply at the highest possible income level (cf. Sadka (1976) and Stiglitz (1982)), and since the types of households are fully revealed by incentive compatibility at young age the no-distortion result holds for the high-productivity households at both ages. Thus, in our model any distortionary income taxes must be related to the low-productivity households only.

#### 4.1.2 Labour Supply for Low-Productivity Households

Turning to the optimal consumption-labour bundle for young low-productivity households, we have from the first order conditions (58a) and (58e) in the appendix that

$$u'\left(c_y^L\right)\left(1-\frac{\lambda}{\pi^L}\right)=v'\left(\frac{I_y^L}{w_y^L}\right)\frac{1}{w_y^L}-\frac{\lambda}{\pi^L}v'\left(\frac{I_y^L}{w_y^H}\right)\frac{1}{w_y^H} \tag{18}$$

and since $w_y^H > w_y^L$ and $v'' > 0$, we have that (see the appendix)

$$\frac{v'\left(\frac{I_y^L}{w_y^L}\right)}{u'\left(c_y^L\right)}\frac{1}{w_y^L}<1. \tag{19}$$

The fact that the marginal rate of substitution of income for consumption is strictly less than 1 at the optimal allocation, implies that the marginal tax rate on labour income of the low-productivity young must be strictly positive in the optimal design of the tax scheme.

A similar argument using equations (58c) and (58g) applies to the allocation of low-productivity old households

$$\frac{V'\left(\frac{I_o^L}{w_o^L}\right)}{U'\left(c_o^L\right)}\frac{1}{w_o^L}<1, \tag{20}$$

so again, at the margin, labour income of the low-productivity old is taxed at a positive rate. Of course, as long as no further restricting assumptions are imposed on the subutility functions it is impossible to deduce from the second-best optimality conditions whether the young or the old low-productivity households should be taxed at the highest rate (in general, however, the tax rates at the two ages should differ).

#### 4.1.3 Optimal Savings

The existence of present-biasedness has not affected the formulas for optimal labour wedges as these wedges do not depend directly on the value of $\beta$. However, since

$\beta$ affects the consumption and labour supply choices of individuals there will be indirect effects of the degree of present-biasedness on the second-best optimal labour income tax rates (as indeed the numerical results will reveal, see below).

The second-best optimal allocation does imply a wedge between the social intertemporal marginal rate of substitution and the individual marginal rate of substitution between current and future consumption when $\beta < 1$. Thus it appears to be socially optimal to distort the savings choice in order to force households to consume (relatively) more and work less when they are old.

For low-productivity households the marginal rate of intertemporal substitution of consumption at the second-best optimal allocation follows by combining equations (58a) and (58c) from the appendix which yields

$$\frac{u'\left(c_y^L\right)}{U'\left(c_o^L\right)} = \frac{\pi^L - \lambda\beta}{\pi^L - \lambda}. \tag{21}$$

Since $u'$ and $U'$ are both strictly positive we have $\pi^L > \lambda$,[6] implying that at the second-best optimal allocation the intertemporal marginal rate of substitution is *larger* than 1,

$$\frac{u'\left(c_y^L\right)}{U'\left(c_o^L\right)} = \frac{\pi^L - \lambda\beta}{\pi^L - \lambda} > 1, \quad \text{for } \beta < 1,\ \pi^L > \lambda. \tag{22}$$

Now, at the first-best allocation this intertemporal marginal rate of substitution equals one, implying that at the second-best allocation the low-productivity households consume relatively more when they are old compared to the first-best situation. Another interesting point of reference is the laissez-faire equilibrium (no taxes) where the individual's intertemporal marginal rate of substitution equals $\beta < 1$. Hence, since the intertemporal marginal rate of substitution in the second-best environment exceeds 1 the low-productivity agents consume relatively more when they are old at the second-best allocation than they would both at the laissez-faire equilibrium, *and* in the first-best equilibrium.

The interesting thing here is that the difference arises only through the incentive constraint as an unbinding incentive constraint (set $\lambda = 0$ in equation (21)) generates an intertemporal marginal rate of substitution equal to 1 (so we return to the first-best). Thus, in order to ensure that the high-productivity households do not have an incentive to mimic low-productivity households, the low-productivity

[6]It follows from the first order condition (58a) that

$$u'\left(c_y^L\right)\left(1 - \frac{\lambda}{\pi^L}\right) = \mu,$$

so since $u' > 0$ and the resource constraint binds at the optimum ($\mu > 0$) it follows that $\pi^L > \lambda$.

households must be forced to save even more at the second-best allocation than at the first-best.

Turning next to the intertemporal marginal rate of substitution of high-productivity households, combining equations (58b) and (58d) we find that the second-best optimal allocation must satisfy

$$\frac{u'\left(c_y^H\right)}{U'\left(c_o^H\right)} = \frac{\pi^H + \lambda\beta}{\pi^H + \lambda}. \tag{23}$$

Again, it immediately follows that the first-best solution is to have the intertemporal marginal rate of substitution equal to 1, and that absence of present-biasedness in the individual's savings decision (i.e. when $\beta = 1$) would imply achievement of the first-best allocation. Unlike the case for low-productivity households, we find that for the high type[7]

$$\beta < \frac{u'\left(c_y^H\right)}{U'\left(c_o^H\right)} < 1. \tag{24}$$

Thus, at the second-best optimal allocation high-productivity households are forced to save *more* than they would do at the laissez-faire equilibrium, but unlike the low-productivity type they save *less* than they would in a first-best environment. A similar result is found by Tenhunen and Tuomala (2010) although in a slightly different setting where the degree of present-biasedness varies among household types and is correlated with productivity (such that low-productivity households suffer from the largest degree of present-biasedness). Hence, our result reveals that the oversaving of low-productivity households and undersaving of the high type (relative to the first-best) are not dependent on low-ability households being more myopic than the high-ability households. Rather, it is a more general property of second-best optimal tax policies when households lack self-control. Intuitively, the result follows from the incentive-constraint facing the policy-maker: To avoid the high-productivity households from mimicking the low-productivity households the second-best optimal tax-transfer policy must induce the low-productivity households to save so much more that the high-productivity households regard mimicking behaviour as sub-optimal.

[7] The right-hand side of the inequality folws immediately from $\lambda > 0$ and $\beta < 1$. The left-hand side of the inequality follows directly from equation (23), as

$$\begin{aligned} \beta < \frac{\pi^H + \lambda\beta}{\pi^H + \lambda} \quad &\Longleftrightarrow \\ \left(\pi^H + \lambda\right)\beta < \pi^H + \lambda\beta \quad &\Longleftrightarrow \\ \pi^H(\beta - 1) < 0 \quad &\Longleftrightarrow \\ \beta < 1. & \end{aligned}$$

## 4.2 Implementation of the Optimal Tax-Transfer Policy

Having derived the second-best optimal allocation, we turn to the design of a tax-transfer scheme that implements this allocation. The tax functions specified in section 3 are general, non-linear functions of labour income and savings and initially we will be able to characterize analytically what the marginal income and savings taxes should be under a second-best optimal tax-transfer policy.[8] As usual in optimal tax policy analysis, however, the conclusions to be derived from the analytical results are quite limited and therefore we subsequently have to rely on numerical examples to gain further insights into what the second-best optimal tax-transfer policy looks like.

### 4.2.1 Optimal Taxes for High-Productivity Households

For high-productivity households, the intertemporal wedge in the second-best environment was described above in equation (23), and the individually rational choice in equation (13) for $i = H$. Combining the two gives

$$\frac{u'\left(c_y^H\right)}{U'\left(c_o^H\right)} = \beta \frac{1 - T_{o,s}^H}{1 + T_{y,s}^H} = \frac{\pi^H + \lambda\beta}{\pi^H + \lambda}. \tag{25}$$

A first thing to note is that for the tax-transfer system to implement the second-best optimal allocation we must have that $T_{y,s}^H \neq -T_{o,s}^H$, that is the marginal tax on savings must differ at young and old age to induce individuals to save a sufficient amount. Hence, the tax treatment of savings must be asymmetric in the two periods. Normalizing the marginal tax on savings of the old to zero, $T_{o,s}^H = 0$, the optimal marginal tax on savings of the young is

$$T_{y,s}^H = \frac{(\beta - 1)\pi^H}{\pi^H + \lambda\beta} < 0. \tag{26}$$

Thus, as expected savings should be subsidized to overcome the present-biasedness in intertemporal consumption choices of households.

[8] It is implicitly assumed that the policy-maker can commit to the tax-transfer scheme selected in the initial period. Without commitment there generally exists a time-inconsistency problem as the policy-maker will have an incentive to deviate from the original tax-transfer scheme at the beginning of the second period. Guo and Krause (2011) analyze optimal time-consistent policies without commitment in a model similar to ours and show that two types of equilibria exist: One where productivities are fully revealed in the first period such that the first-best outcome can be implemented in the second period, and another where nothing is revealed in the first period leading to a standard one-period incentive compatibility problem in the second period. Hence, in both cases the equilibrium outcomes are qualitatively significantly different in the two periods which would seem at odds with empirical evidence. For that reason we stick to the assumption of commitment being possible for the policy-maker.

The optimal taxes on labour income in the two periods are clearly zero at the margin for high-productivity households, since it follows from equations (16) and (17) that there is no temporal wedge, which combined with equations (11) and (12) from section 3 imply that $T_{y,I}^{H} = T_{o,I}^{H} = 0$. As noted above, this is just the standard result in optimal tax theory that the marginal income tax rate at the highest possible income level is zero (the only special result here is that it applies to the marginal income tax rates at both young and old age of the high-productivity households due to types being revealed when agents are young given incentive compatibility).

#### 4.2.2 Optimal Taxes for Low-Productivity Households

For low productivity households, the intertemporal wedge is given by equation (21), and to find the implementing taxes it is equated with the individually rational intertemporal choice in equation (13) for $i = L$, yielding

$$\beta \frac{1 - T_{o,s}^{L}}{1 + T_{y,s}^{L}} = \frac{\pi^{L} - \lambda\beta}{\pi^{L} - \lambda}. \tag{27}$$

Again we can normalize one of the marginal tax rates, so setting $T_{o,s}^{L} = 0$ the resulting optimal tax formula becomes

$$T_{y,s}^{L} = \frac{(\beta - 1)\pi^{L}}{\pi^{L} - \lambda\beta} < 0. \tag{28}$$

As for the high-productivity households we need savings to be subsidized.

To what extent is it possible to compare the relative subsidies to high- and low-productivity households? Using $\pi^{H} = 1 - \pi^{L}$ it follows that numerically[9]

$$\left|T_{y,s}^{H}\right| < \left|T_{y,s}^{L}\right|$$

so low-productivity households should receive a higher marginal subsidy to savings than high-productivity households. The difference occurs through the term $\beta\lambda$, so both the degree of present-biasedness and the shadow price of imposing incentive compatibility matter for the difference between the subsidy to savings to the low-productivity and that to the high-productivity households. In particular, the more

[9] We have

$$T_{y,s}^{L} = \frac{(\beta - 1)\pi^{L}}{\pi^{L} - \lambda\beta} < \frac{(\beta - 1)\left(1 - \pi^{L}\right)}{(1 - \pi^{L}) + \lambda\beta} = T_{y,s}^{H}$$

which reduces to $\beta\lambda > 0$.

costly it is induce self-selection the larger is the wedge between the two (marginal) subsidies.[10]

The wedge on labour income for the young low-productivity households is given by

$$\frac{v'\left(l_y^L\right)}{u'\left(c_y^L\right)} = w_y^L \left[\left(1 - \frac{\lambda}{\pi^L}\right) + \frac{\lambda}{\pi^L}\frac{v'\left(I_y^L/w_y^H\right)}{w_y^H u'\left(c_y^L\right)}\right], \tag{29}$$

which is then equated to $w_y^L\left(1 - T_{y,I}^L\right)$, yielding a marginal income tax rate of

$$T_{y,I}^L = \frac{\lambda}{\pi^L}\left(1 - \frac{v'\left(I_y^L/w_y^H\right)}{w_y^H u'\left(c_y^L\right)}\right). \tag{30}$$

Since $I_y^L/w_y^H < l_y^L$ and $v(\cdot)$ is convex, we have that

$$\frac{v'\left(I_y^L/w_y^H\right)}{w_y^H u'\left(c_y^L\right)} < \frac{v'\left(l_y^L\right)}{w_y^H u'\left(c_y^L\right)} < \frac{v'\left(l_y^L\right)}{w_y^L u'\left(c_y^L\right)} < 1, \tag{31}$$

so the optimal marginal tax on labour income of the young low-productivity households is positive[11].

A similar expression holds for low-productivity old households, namely

$$T_{o,I}^L = \frac{\beta\lambda}{\pi^L}\left(1 - \frac{V'\left(I_o^L/w_o^H\right)}{w_o^H U'\left(c_o^L\right)}\right), \tag{32}$$

which basically differs in three respects from the optimal marginal tax on the young low-productivity households: It is evaluated at a different productivity level (the productivity levels of the mimickers being $w_y^H$ or $w_o^H$, respectively); different marginal utility function applies ($v'(\cdot)$ and $u'(\cdot)$ for the young households, and $V'(\cdot)$ and $U'(\cdot)$ for the old households), and there is a direct effect of the degree of present-biasedness, $\beta$, on the marginal income tax of the old low-productivity households.

The analytical solutions to the optimal marginal tax on low-productivity households may not appear to be very informative with respect to what determines the size of the marginal income tax rates, essentially because many of the variables

[10] The numerical analysis reveals that the size of $\lambda$ is (almost) unaffected by changes in $\beta$. This suggests that the changes in subsidies are driven primarily by the changes in the degree of present-biasedness, and not by changes in the incentive relation between the two types. When we vary the share of low-productivity workers, our numerical calculations show that $\lambda$ increases with $\pi^L$, so the increase in the difference between the two marginal subsidies is due to the increased cost of inducing self-selection.

[11] Since $\lambda < \pi^L$, the tax is bounded above by 1.

entering the optimal tax formulas are endogenous (like the Lagrange multiplier on the incentive constraint, $\lambda$, and the marginal rates of substitution between consumption and leisure). Therefore, as usual in this literature, we have to rely on numerical solutions for gaining further insights into what the optimal tax-transfer policy may look like.

# 5 Numerical Results

To illustrate the analytical results above, we now turn to numerical simulations of the model in different scenarios. Analytically, we were able to obtain results only for the second-best optimal wedges, whereas numerically we can obtain insights into the entire allocation across generations and household types and look for how the second-best optimal allocations - and the implied optimal tax-transfer schemes - depend on key parameters of the model.[12] Of course, given the highly stylized structure of the model we will not attempt to calibrate the model to match empirical moments of typical real world economies. Instead, we will look for more qualitative properties of the second-best optimal tax-transfer policies, in particular by considering which types of policies will not belong to the set of optimal ones.

For the choice of parametrization we select quite standard functional forms with instantaneous utility from consumption of goods being logaritmic and instantaneous disutility of work being iso-elastic. Hence, for consumption of goods the individual utilities from private consumption are given by

$$u\left(c_y^j\right) = \ln c_y^j,\ U\left(c_o^j\right) = \ln c_o^j, \tag{33}$$

while the disutilities from supplying labour are iso-elastic

$$v\left(l_y^j\right) = \frac{\left(l_y^j\right)^{1+1/\varepsilon_y}}{1+1/\varepsilon_y}, V\left(l_o^j\right) = \frac{\left(l_o^j\right)}{1+1/\varepsilon_o}^{1+1/\varepsilon_o}, \tag{34}$$

where we explicitly allow for different labour supply responses to tax changes for young and old households, respectively. For these isoelastic preferences, $\varepsilon_i$ is the Frisch elasticity of labour supply for $i \in \{y, o\}$. Very similar parameterizations have been employed by e.g. Cremer et al. (2009) and Bastani et al. (2010).

The second-best allocation is found by solving a system of ten non-linear equations in ten unknowns (four consumption levels, four income levels and two multipliers). From these levels taxes can be deduced, and we can calculate the optimal

[12] We can determine the optimal second-best total lifetime tax-transfers of the various households, but we cannot determine the timing of these tax-transfers. This is because the model exhibits Ricardian Equivalence so any changes in the timing of tax-transfers will be met by changes in the opposite direction of private savings. However, the full allocation of consumption levels and labour supplies will be determined in the numerical solutions.

wedges[13]. We also compute the first-best and laissez-faire allocations for comparison.

## 5.1 Scenario 1: Benchmark

As mentioned above we do no attempt to calibrate the model to match any empirical moments of an actual economy. In stead, parameter values will be chosen within a set of "reasonable" values to reveal if there exists interesting relations between these values and the optimal second-best tax-transfer policy. The following parameter values must be chosen:

1. Shares of high- and low-productivity households ($\pi^i$, $i \in \{H, L\}$).
2. Labour productivities of the two types of households when young and old, respectively ($w_j^i$, $i \in \{H, L\}$, $j \in \{y, o\}$).
3. Labour supply elasticities of young and old households ($\varepsilon_j$, $j \in \{y, o\}$).
4. The degree of present-biasedness ($\beta$).

In the benchmark scenario the parameter values are set to

| Model parameters | | | |
|---|---|---|---|
| $\pi^L$ | 0.5 | $w_y^L$ | 5 |
| $\pi^H$ | 0.5 | $w_o^L$ | 2 |
| $\varepsilon_y$ | 0.5 | $w_y^H$ | 8 |
| $\varepsilon_o$ | 0.5 | $w_o^H$ | 5 |
| $\beta$ | 0.5 | | |

Table 1: Parameter values in the baseline scenario.

Hence, we start out assuming the two types of households are present in equally sized groups. Both types of households experience a declining labour productivity time profile so young age could be interpreted as prime age in terms of earnings potential, and old age as the age where retirement becomes an option.[14] The labour supply elasticities are formally Frisch elasticities, but due to the two period set up where a single period could reflect 20-30 calender years, the labour supply

[13]The numerical exercise is carried out in MATLAB using the fsolve.m routine to find the numerical solution. Files are available upon request.

[14]Since labour productivities decline with age labour supply of the old will be considerably less than the labour supply of the young. This could be interpreted - as discussed later - as reflecting retirement within the period of old age.

| Regime | *Laissez-faire* | *First-best* | *Second-best* |
|---|---|---|---|
| **Allocation** | | | |
| $c_y^L$ | 4.6787 | 5.2281 | 3.9659 |
| $c_y^H$ | 8.6920 | 5.2281 | 6.1455 |
| $c_o^L$ | 2.3394 | 5.2281 | 4.5108 |
| $c_o^H$ | 4.3460 | 5.2281 | 5.6006 |
| $I_y^L$ | 5.1688 | 4.8897 | 5.1086 |
| $I_y^H$ | 7.6749 | 9.8961 | 9.1276 |
| $I_o^L$ | 1.8493 | 1.2370 | 1.2623 |
| $I_o^H$ | 5.3630 | 4.8897 | 4.7243 |
| $\mu$ | 0.0000 | 0.1913 | 0.1078 |
| $\lambda$ | 0.0000 | 0.0000 | 0.1978 |
| **Taxes** | | | |
| $\bar{T}^L$ | 0 | $-0.707$ | $-0.331$ |
| $\bar{T}^H$ | 0 | 0.293 | 0.152 |
| $T_{y,I}^L$ | 0 | 0 | 0.172 |
| $T_{o,I}^L$ | 0 | 0 | 0.102 |
| $T_{y,s}^L$ | 0 | $-0.500$ | $-0.560$ |
| $T_{y,s}^H$ | 0 | $-0.500$ | $-0.451$ |
| **Welfare** | | | |
| $W$ | 2.3436 | 2.6414 | 2.6036 |

Table 2: Allocation in the laissez-faire, the first-best and the second-best regimes in the baseline case.

elasticities are more like macro elasticities (steady state elasticities). As shown by Rogerson and Wallenius (2009) macro elasticities are basically unrelated to the underlying micro (Frisch) elasticities and a value of one half may not be inconsistent with macro evidence. Finally, the discount factor showing the degree of present-biasedness is set to one half which is consistent with and annual present-biased discount factor of around 0.98.

Using these parameter values we can calculate the equilibrium allocations of consumption and labour income and the marginal rates of substitution (measured as wedges) in the laissez-faire, the first-best and the second-best settings, and we can derive the second-best optimal tax-transfers and marginal tax rates. The allocations in the benchmark case are presented in Table 2.[15]

[15] In the tables $\overline{T}^L$ and $\overline{T}^H$ are the average lifetime tax *rates* of low- and high-productivity households, respectively, calculated as total taxes paid by the household type divided by household lifetime income. Of course, as the government budget balances the sum of total *taxes* paid equals zero.

By comparing the laissez-faire and the first-best allocations it becomes obvious what the tax-transfer policy is intended to correct. In relative terms households consume too much and work too little when young due to their present-biased choices under laissez-faire.[16] Thus, the second-best optimal tax-transfer policy should distort consumption and labour supply choices such that both types work relative more when young and have a more smooth consumption time profile. Secondly, the optimal policy should also transfer resources from the high- to the low-productivity households. Looking carefully at the table it is revealed that the second-best policy actually succeeds on both accounts. Consumption smoothing becomes much more pronounced than under laissez-faire and households are induced to work relative more hours when they are young compared to their old age labour supply, and the low-productivity households receive positive transfers (measured in present value terms) from the high-productivity households.

As shown in the analytical section only the low-productivity households should be distorted in their labour supply decisions. Hence, we only have non-zero marginal income taxes on the low-productivity households and in the benchmark case it is optimal to tax the young low-productivity households at a higher rate than the old low-productivity households ($T^L_{y,I} = 0.17 > T^L_{o,I} = 0,10$). Although it could seem counter-intuitive to impose a higher marginal tax on the young households - who are more productive than the old households - it is required for the incentive-compatibility constraint to be fulfilled. If the young low-productivity households' income was taxed at a lower rate - which would be efficiency enhancing - the young high-productivity households would find it desirable to mimick the behaviour if the low-productivity households and this would be in conflict with a second-best optimal policy.

Regarding the marginal taxes on savings that are needed to compensate for the present-biased choices made by the households the numerical results confirm the analytical results: Savings of low-productivity households are subsidized at a higher rate than the savings of their high-productivity peers.

## 5.2 Scenario 2: Less Present-Biasedness

To investigate the role of the degree of present-biasedness in household preferences for the optimal tax-transfer policy we calculate the equilibrium allocation and the required taxes needed to implement the second-best optimal policy for higher

[16] In absolute terms the low-productivity households actually work more at both ages under laissez-faire than at the first-best which is a consequence of the lack of transfers under laissez-faire, implying that the low-productivity households have to work quite a lot to obtain the desired balance between consumption of goods and consumption of leisure. Under the first-best allocation there is no explicit link between individual labour supply and individual consumption as the social planner costlessly transfers resources from the high- to the low-productivity households.

values of $\beta$ (reflecting less present-biasedness in household preferences). All the other parameter values are as in the baseline case, and alternative values for the individual's discount factor, $\beta$, are chosen to be 0.75 and 0.9, respectively. The full allocation is given in the appendix while the tax-transfers and marginal tax rates are presented in Table 3.

| **Tax** | **Baseline** | $\beta = \mathbf{0.75}$ | $\beta = \mathbf{0.9}$ |
|---|---|---|---|
| $\overline{T}^L$ | $-0.331$ | $-0.317$ | $-0.318$ |
| $\overline{T}^H$ | 0.152 | 0.145 | 0.146 |
| $T^L_{y,I}$ | 0.172 | 0.155 | 0.143 |
| $T^L_{o,I}$ | 0.102 | 0.139 | 0.154 |
| $T^L_{y,s}$ | $-0.560$ | $-0.293$ | $-0.119$ |
| $T^H_{y,s}$ | $-0.451$ | $-0.218$ | $-0.086$ |

Table 3: Average (lifetime) tax rates for low and high ability workers, marginal income tax rates for low ability workers and marginal subsidies on savings of low and high ability (young) households for values of the present-bias parameter $\beta$.

As expected a smaller tendency to bias decisions towards the present implies that households work relatively more when young and plan more intertemporal consumption smoothing. Also quite naturally, this implies that the savings subsidies should be reduced under the second-best optimal tax-transfer policy.

Increasing the value of $\beta$ also affects the marginal income tax rates on the low-productivity households, although it is not entirely clear how the relation should be (see footnote 10). At least for the present parameterization it seems as if the marginal income tax rates of the old low-productivity households increase with $\beta$ (so the direct effect of $\beta$ on $T^L_{o,I}$ as evident in equation (32) dominates any indirect effects) while the marginal income tax rate of the young low-productivity households falls, thereby increasing the efficiency of the tax-transfer policy.[17] An explanation of this phenomenon could be that with less present-biasedness it becomes less attractive for the high-productivity young household to mimic the young low-productivity household (because he now favours present consumption relatively less), which then allows the paternalistic planner to reduce the distortion on the behaviour of the low-productivity young households.

It should also be noted that in this scenario it is possible for the marginal income tax on the old low-productivity household to exceed the marginal tax rate on the young household of the same type, so there can be no general ranking on who should face the highest marginal tax rate, young or old households.

[17] It is efficiency enhancing to make young households work more and old households work less (since the young households have higher productivity). By reducing the marginal income tax on the young and increasing it on the old households we get higher efficiency.

## 5.3 Scenario 3: Smaller Share of High-Productivity Households

Now, suppose that a larger share of the population is endowed with low productivity. All the remaining parameters are set as in the baseline case expect for the share of low productivity households which is set to $\pi^L = 0.75$ and $\pi^L = 0.9$, respectively. The full allocation is given in the appendix while the taxes are presented in table 4.

| **Tax** | **Baseline** | $\pi^L = \mathbf{0.75}$ | $\pi^L = \mathbf{0.9}$ |
|---|---|---|---|
| $\bar{T}^L$ | $-0.331$ | $-0.152$ | $-0.057$ |
| $\bar{T}^H$ | 0.152 | 0.219 | 0.254 |
| $T^L_{y,I}$ | 0.172 | 0.097 | 0.042 |
| $T^L_{o,I}$ | 0.102 | 0.059 | 0.026 |
| $T^L_{y,s}$ | $-0.560$ | $-0.533$ | $-0.514$ |
| $T^H_{y,s}$ | $-0.451$ | $-0.421$ | $-0.401$ |

Table 4: Average (lifetime) tax rates for low and high ability workers, marginal income tax rates for low ability workers and marginal subsidies on savings of low and high ability (young) households for values of the population share parameter $\pi^L$.

There seems to be two main consequences of increasing the share of low-productivity households in the economy. First, the per capita net transfers received by the low-productivity households are decreased (by a factor of 6 when the share of low-productivity households is increased from 50 % to 90 %), and secondly the marginal income tax rates of the low-productivity households are reduced (this time by a factor of 4).

That it becomes optimal to reduce the extent of redistribution is simply a matter of the mass of high-productivity households being so small that there are few resources available per low-productivity household. This implies, on the other hand, that - for given marginal income tax rates - the welfare level of high-productivity households greatly exceeds that of the low-productivity households, compared to the baseline scenario. Hence, even for very low marginal income tax rates on the low-productivity households the incentives for high-productivity households to pursue a mimicking strategy are rather weak, leading to low optimal marginal tax rates on the low-productivity households. It is worth noting that the pre-tax income of high-productivity households increases with $\pi^L$ while consumption levels decrease clearly indicating that the net burden on the high-productivity households is increased.

The effect on the optimal intertemporal wedges is similar. When the low income strategy is less attractive to the high-productivity household less distortion of the allocation of the low-productivity households is needed. Since the second-best allocation implies 'oversaving' for low-productivity households in the baseline scenario, the scenario with a larger fraction of low-productivity households implies less 'oversaving' of these households.

## 5.4 Scenario 4: Different Labour Supply Elasticity for Young and Old Households

In the general model we have allowed for different attitudes towards leisure between young and old households and this is modelled in the parameterized model as possibly different Frisch labour supply elasticities, $\varepsilon_y$ and $\varepsilon_o$, respectively. Usually, in optimal taxation behavioural elasticities are important for the optimal size of tax rates, so it is of interest to investigate whether this is also the case in the present model. In the baseline case both elasticities are set at 0.50 and in the alternative scenarios the labour supply elasticity of the young households is kept at this level while the elasticity of the old households is set at $\varepsilon_o = 0.25$ or $\varepsilon_o = 1.0$.[18] All the remaining parameters are set as in the baseline case. The full allocations are given in the appendix while the optimal taxes are presented in table 5.

| **Tax** | $\varepsilon_o = \mathbf{0.25}$ | **Baseline** | $\varepsilon_o = \mathbf{1}$ |
|---|---|---|---|
| $\bar{T}^L$ | −0.325 | −0.331 | −0.344 |
| $\bar{T}^H$ | 0.154 | 0.152 | 0.151 |
| $T^L_{y,I}$ | 0.156 | 0.172 | 0.192 |
| $T^L_{o,I}$ | 0.097 | 0.102 | 0.102 |
| $T^L_{y,s}$ | −0.554 | −0.506 | −0.568 |
| $T^H_{y,s}$ | −0.455 | −0.451 | −0.447 |

Table 5: Average (lifetime) tax rates for low and high ability workers, marginal income tax rates for low ability workers and marginal subsidies on savings of low and high ability (young) households for values of the elasticity parameter $\varepsilon_o$.

Perhaps somewhat surprisingly neither optimal savings subsidies nor optimal labour income taxes seem to be very dependent on these changes in relative labour supply elasticities. The biggest quantitative effect is actually on the marginal labour income tax of the young low-productivity households whose marginal tax is increasing in the labour supply elasticity of the *old* households. Since the optimal marginal labour income tax rate of the old low-productivity households is slightly increasing in the elasticity of labour supply of the old, we do get the expected

[18] At noted earlier the labour supply elasticities in this model are basically macro elasticities that easily can take values in the range [0.25 − 1.0], see also Kocherlakota (2010).

result in relative terms: When old households react more elastically to changes in after-tax labour income the optimal tax-transfer policy taxes the labour income of the young more heavily. Hence, if empirically labour supply elasticities tend to increase with age we would expect this separate effect to increase marginal taxes of the young relative to the marginal taxes of the old households.

# 6 Discussion of the Results and Policy Recommendations

Given the relatively simple structure of the model no claims will be made that very strong policy recommendations can be supported by the results of the model. Instead, we will first discuss what kind of insights can be derived form the analytical results and subsequently what can be inferred from the numerical results.

## 6.1 Discussion of Analytical Results

It follows straightforwardly from the assumption of having two types of households in the economy that only the low type will be subject to distortionary labour income taxation. Hence, it comes as no surprise that the optimal marginal tax rates on the incomes of high-productivity households are zero (cf. Stiglitz (1982)). Moreover, it will generally be the case that young and old low-productivity households should not face the same marginal income tax rates, so the assumed age-dependency of the tax system is useful for creating the right incentives for life cycle labour supply decisions. However, it is not possible to determine which age group should face the highest marginal income tax rate.

Regarding the tax treatment of savings it is generally the case that as long as households possess present-biased preferences, it is optimal to subsidize household savings irrespective of the productivity level of the household. It also follows that the low-productivity households should be induced to increase their savings relative to the first-best level of savings whereas the high-productivity households should only be induced to increase savings relative to laissez-faire. Hence, there is a tendency for the optimal policy to induce "oversaving" of the low-productivity households so savings subsidies should be regressive (i.e. largest in absolute value for the low-productivity households).

Another feature of the optimal second-best tax-transfer policy is that the two main issues the tax-transfer policy should attend to - income redistribution from high- to low-productivity households and correction of present-biasedness in household preferences - cannot be treated independently of each other. As seen from equation (32) the optimal marginal income tax rate of the old low-productivity households depends explicitly on the degree of present-biasedness. So even when the policy-maker has a sufficient number of tax instruments, optimal income redistribution of income across productivity types of households cannot be treated separately from the issue of correcting for present-biasedness. Intuitively, this follows because the degree of present-biasedness affects the incentive-compatibility constraint and thereby by how much incentives can be distorted through marginal income taxes.

As an example of what generally cannot constitute a second-best optimal tax-transfer system consider the following age-dependent, non-linear tax-transfer functions that only depends on a single measure of income (so income and savings do not enter the tax functions as separate variables). Hence, let the tax functions be $T_y(\cdot)$ and $T_o(\cdot)$ and let the tax base of young households be labour income net of savings while the tax base of the old is labour income plus savings (including the return to savings if the interest rate were positive). Having savings being tax deductible in the first period and taxed in the second period is similar to having pre-tax contributions in the US 401(k) pension system. This would imply household budget constraints like

$$c_y^i = w^i l_y^i - s^i - T_y(w^i l_y^i - s^i) \tag{35}$$

$$c_o^i = w^i l_o^i + s^i - T_o(w^i l_o^i + s^i). \tag{36}$$

In this case the optimality conditions from the household optimization problem become (for $i \in \{H, L\}$)

$$\frac{v'(l_y^i)}{u'(c_y^i)} = w_y^i(1 - T_{y,I}^i) \tag{37}$$

$$\frac{V'(l_o^i)}{U'(c_o^i)} = w_o^i(1 - T_{o,I}^i) \tag{38}$$

$$\frac{u'(c_y^i)}{U'(c_o^i)} = \beta \frac{1 - T_{o,I}^i}{1 - T_{y,I}^i}, \tag{39}$$

where $T_{j,I}^i$ is the derivative of the tax function $T_j(\cdot)$ evaluated at the taxable income level of household type $i \in \{H, L\}$. Now, using the optimality conditions for the second-best tax-transfer policy problem for, say, the low-productivity households, we obtain

$$T_{y,I}^L = \frac{\lambda}{\pi^L}\left(1 - \frac{v'\left(I_y^L / w_y^H\right)}{w_y^H u'\left(c_y^L\right)}\right) \tag{40}$$

$$T_{o,I}^L = \frac{\beta\lambda}{\pi^L}\left(1 - \frac{V'\left(I_o^L / w_o^H\right)}{w_o^H U'\left(c_o^L\right)}\right) \tag{41}$$

$$\beta \frac{1 - T_{o,I}^L}{1 - T_{y,I}^L} = \frac{\pi^L - \lambda\beta}{\pi^L - \lambda}. \tag{42}$$

It follows straightforwardly that it is generally not possible to determine the two marginal tax rates $T_{y,I}^{L}$ and $T_{o,I}^{L}$ such that all three optimality conditions are satisfied simultaneously. There is simply one tax instrument lacking (the marginal tax of savings). Thus, it is essential for implementation of the second-best optimal tax-transfer policy that labour income and savings enter separately in the tax functions.

## 6.2 Discussion of Numerical Results

The numerical results revealed some tendencies that may be of some interest. First, the relative distortions on life cycle labour supply of the low-productivity households seem to depend on the degree of present-biasedness in the manner suggested by the direct effect in equation (32): Less severe present-biasedness increases the relative distortion of labour supply of the old low-productivity households. Even in absolute terms it becomes possible to generate higher marginal income taxes on the old than on the young low-productivity households. This may suggest that if insufficient savings for the old age is considered to be a prominent policy concern leading to large savings subsidies one would expect that marginal income taxes should decline with age. On the other hand, if present-biasedness seems to be less of a problem there is no support from the current model for the optimal policy to have reduced labour supply distortions at old age. This exemplifies how the distortions coming from present-biased savings decisions influence the efficiency losses of labour income taxation. Since the young households are more productive than the old households it would be preferable on efficiency grounds to tax the young households at a lower rate than the old households, but as present-biasedness becomes more severe (a lower value of $\beta$) the higher is the optimal relative marginal income tax on the young households.

A second result of interest is that although the size of labour supply elasticities matter for the optimal tax rates, the quantitative effects seem quite small. In our numerical analysis a doubling of the labour supply elasticity of the old households has hardly any effect on the optimal marginal income tax on the old low-productivity household and increases the marginal income tax rate on the young low-productivity household by 2 percentage points (from 17% to 19%). Of course, this may just reflect a particularity of the parameterization of our model so future work may be needed to verify the robustness of this result.

Finally, regarding the optimal subsidy to savings the marginal savings subsidies are fairly constant across different parameterizations except for different values of $\beta$. Hence, it is predominantly the degree of present-biasedness in household preferences that govern how much savings should be subsidized. Since the degree of present-biasedness also has strong implications for the optimal labour supply distortions, $\beta$ may be the single most important parameter for the design of second-best optimal tax-transfer policy.

# 7 Extension and Generalization

In the model households work at both ages but since productivity is assumed to be declining with age household supply of labour is also declining with age. This could be interpreted as if households are partly retiring at old age. Moreover, a perhaps little unusual feature of our model is that - due to quasi-hyperbolic discounting and only two periods - households plan to work too little when young and too much when old, so within a retirement interpretation this amounts to later rather than earlier retirement. Within a three period model with mandatory retirement in the final period and voluntary retirement in the middle period Diamond and Köszegi (2003) have studied the implications of quasi-hyperbolic preferences for the retirement decision and they show that early retirement (rather than later retirement) may be the equilibrium outcome. Hence, it would be of interest to analyze what a tendency to early retirement due to present-biasedness implies for the optimal tax-transfer policy.

There is a growing literature trying to incorporate the basic insights of Laibson's (1997) quasi-hyperbolic discounting savings model into a variety of public finance settings.[19] A crucial feature of these models is whether agents are "naive" or "sophisticated". However, for the difference between naive and sophisticated behaviour to occur we need at least three periods in the model when agents use quasi-hyperbolic discounting. In our set up this could be accomplished by splitting up the second period into two subperiods with mandatory retirement in the last part of the second period and an endogenous labour supply decision in the first part. With effectively three periods and sophisticated households that take into account that future "selves" have different preferences than the current self the models gets extremely complicated as different selves interact in a quite complicated game (see e.g. Diamond and Köszegi (2003) where it especially follows that extending the model from three to four periods complicates the analysis considerably). Add to such a model households of different types that are unknown to the policy-maker and have the policy-maker engaging in an income redistribution exercise as in our main model, and we end up with a rather unwieldy model. Hence, to keep matters a bit more simple we assume that households are naive in the sense that current selves expect future selves not to reoptimize (since no new information is revealed through time).

[19] Some examples include: Diamond and Köszegi (2003) focus on the effects on endogenous retirement; O'Donoghue and Rabin (2006) look at optimal sin taxes; Aronsson and Sjögren (2011) investigate the consequences for optimal taxation more broadly and look at both closed and open economies.

## 7.1 Present-Biased Preferences: Three Period Model with Naive Households

To model this more formally we now introduce three periods by dividing the second period in two: One where the household consume and (possibly) works and one where work is not an option (mandatory retirement). Moreover, preferences are now explicit modelled as quasi-hyperbolic so that the plan made by a young household for consumption and labour supply later in life can be subject to reoptimization in the middle period.

The three periods are now denoted $i \in \{y, o, r\}$ (young, old, retirement) and the household preferences at young age are governed by

$$\mathcal{U}_y \left(c_y^i, c_o^i, l_y^i, l_o^i, c_r^i\right) = u\left(c_y^i\right) - v\left(l_y^i\right) \\ + \beta\delta\left(U\left(c_o^i\right) - V\left(l_o^i\right)\right) + \beta\delta^2 U(c_r^i), \quad i \in \{H, L\}, \quad (43)$$

where we now introduce the subjective discount factor $\delta \in ]0, 1[$ while $\beta$ as before measures the degree of present-biasedness. The budget constraints can now be written as

$$c_y^i = w_y^i l_y^i - s_y^i - T_y(w_y^i l_y^i, s_y^i), \qquad i \in \{H, L\} \quad (44)$$

$$c_o^i = w_o^i l_o^i + s_y^i - s_o^i - T_o(w_o^i l_o^i, s_y^i, s_o^i), \qquad i \in \{H, L\} \quad (45)$$

$$c_r^i = s_o^i - T_r(s_o^i), \qquad i \in \{H, L\}, \quad (46)$$

where we now have to distinguish between savings made by the young, $s_y^i$, and by the old, $s_o^i$.[20] The solution to the optimization problem is stated in the appendix.

At old age preferences are now governed by

$$\mathcal{U}_o\left(\widetilde{c}_o^i, \widetilde{l}_o^i, \widetilde{c}_r^i\right) = U\left(\widetilde{c}_o^i\right) - V(\widetilde{l}_o^i) + \beta\delta U(\widetilde{c}_r^i), \quad i \in \{H, L\}, \quad (47)$$

where $\left(\widetilde{c}_o^i, \widetilde{l}_o^i, \widetilde{c}_r^i\right)$ denote the actually chosen values of consumption and labour supply when old and consumption when retired. The budget constraint are

$$\widetilde{c}_o^i = w_o^i \widetilde{l}_o^i + s_y^i - \widetilde{s}_o^i - T_o(w_o^i \widetilde{l}_o^i, s_y^i, \widetilde{s}_o^i), \quad i \in \{H, L\} \quad (48)$$

$$\widetilde{c}_r^i = \widetilde{s}_o^i - T_r(\widetilde{s}_o^i), \quad i \in \{H, L\}. \quad (49)$$

We can now compare the actual consumption and labour supply choices with their planned values at young age. Comparing equations (81) and (83) it follows

[20] Of course, $s_o^i$ is planned savings of the old while the actual savings generally will differ from the planned level due to the present-biased preferences.

immediately that - assuming no differences in marginal savings tax rates between planned and actual incomes (so $T^i_{j,s_o} = T^i_{j,\widetilde{s}_o}$ for $i \in \{H, L\}$ and $j \in \{o, r\}$)

$$\frac{U'(\widetilde{c}^i_o)}{U'(\widetilde{c}^i_r)} < \frac{U'(c^i_o)}{U'(c^i_r)}, \tag{50}$$

so as expected the actual consumption choices will be biased toward old age and away from the retirement age. Regarding the actual labour supply choice of the old it may appear - assuming no differences in marginal income tax rates between planned and actual incomes (so $T^i_{o,I} = T^i_{o,\widetilde{I}}$) - that actual labour supply is undistorted at old age (compare equations (79) and (82)). However, as $\widetilde{c}^i_o > c^i_o$, $U'' < 0$ and $V'' > 0$ it follows that $\widetilde{l}^i_o < l^i_o$. That is, labour supply at old age is downward biased relative to the "true" preferences of the households.

Hence, by extending the model to a three period setting with mandatory retirement in the final period we actually get downward bias of labour supply at old age. For the second-best optimal tax-transfer policy this must have implications, most likely by reducing the marginal income tax rate at old age. Of course, a formal treatment of the second-best optimal tax-transfer policy problem that include the relevant incentive-compatibility constraints is needed before more firm conclusions can be drawn but this is certainly an interesting avenue for future research.

# 8 Concluding Remarks

This paper has dealt with how classical tax-based income redistribution from high- to low-ability individuals interacts with intertemporal redistribution tailored at inducing myopic individuals to increase savings for old age. Doing so in a second-best framework where the abilities of individuals are revealed through self-selection generates some interesting insights. Previous studies - like Tenhunen and Tuomala (2010) - have found that when low-ability households are more myopic than high-ability households the second-best optimal policy quite naturally stimulates savings of the low-ability households more strongly than for the high-ability type. However, we show that this result also holds when all households are equally myopic, so it is not the differences in the degree of myopia that is important for this result. Instead, it reflects some fundamental properties of the incentive-compatibility constraint facing the policy-maker: In order to avoid the high-ability type from mimicking the low-ability type, the latter must be induced to increase savings substantially so mimicking becomes an inferior strategy. This is accomplished by offering higher marginal subsidies to low-ability households than to high-ability households.

The need for subsidizing savings through the tax-transfer system implies that savings must be subject to asymmetric tax treatment across time. Tax deductibility for savings made when young (e.g. for retirement purposes) and subsequently savings being subject to income taxation when it becomes available for consumption at old age will not allow for the second-best optimal allocation to be implemented, or at least this type of tax treatment of savings should be supplemented by explicit subsidies to savings. Moreover, savings subsidies should not be uniform across households types but rather be biased towards being larger for low-ability types.

More generally, there appears to be important interdependencies between the two main concerns in designing the optimal tax-transfer policy. The degree of present-biasedness seems to impact on both the optimal level of marginal income tax rates and on the relative marginal income taxes imposed on young and old households, in a way such that stronger degrees of myopia increases the relative marginal income tax rate of the young. Hence, with strong myopia among households generous savings subsidies to the low-ability type should not be counter-acted by having higher marginal income taxes at old age for this type.

One drawback of our analysis - which is common to much of this literature - is that many of the results can only be derived from the numerical solutions to the optimal tax problem, so the generality of the results is limited. What does hold more generally is that it is not possible to separate optimal income redistribution from the need to the induce myopic households to save more, even in a model where the degree of myopia is uniform across household types and perfectly known to the policy-maker.

# References

Akerlof, George A. (1978), 'The economics of "tagging" as applied to the optimal income tax, welfare programs, and manpower planning', *The American Economic Review* **68**(1), 8–19.

Aronsson, Thomas and Tomas Sjögren (2011), Tax policy and present-biased preferences: Paternalism under international capital mobility, Umeå Economic Studies 833, Umeå University, Department of Economics.

Bastani, Spencer, Sören Blomquist and Luca Micheletto (2010), The welfare gains of age related optimal income taxation, CESifo Working Paper Series 3225, CESifo Group Munich.

Blomquist, Sören and Luca Micheletto (2008), 'Age-related optimal income taxation', *Scandinavian Journal of Economics* **110**(1), 45–71.

Brett, Craig and John A. Weymark (2005), Optimal nonlinear taxation of income and savings in a two class economy, Working Papers 0525, Department of Economics, Vanderbilt University.

Cremer, Helmuth, Philippe De Donder, Dario Maldonado and Pierre Pestieau

(2009), 'Forced saving, redistribution, and nonlinear social security schemes', *Southern Economic Journal* **76**(1), 86–98.
Diamond, Peter and Botond Köszegi (2003), 'Quasi-hyperbolic discounting and retirement', *Journal of Public Economics* **87**(9-10), 1839–1872.
Kocherlakota, Narayana R. (2010), *The New Dynamic Public Finance*, Toulouse Lectures in Economics, Princeton University Press, Princeton, New Jersey.
Laibson, David (1997), 'Golden eggs and hyperbolic discounting', *The Quarterly Journal of Economics* **112**(2), 443–77.
Mirrlees, James A. (1971), 'An exploration in the theory of optimum income taxation', *Review of Economic Studies* **38**, 175–208.
O'Donoghue, Ted and Matthew Rabin (2006), 'Optimal sin taxes', *Journal of Public Economics* **90**(10-11), 1825–1849.
Rogerson, Richard and Johanna Wallenius (2009), 'Micro and macro elasticities in a life cycle model with taxes', *Journal of Economic Theory* **144**(6), 2277–2292.
Sadka, Efraim (1976), 'On income distribution, incentive effects and optimal income taxation', *Review of Economic Studies* **43**(2), 261–67.
Salanié, Bernard (2005), *The Economics of Contracts*, 2 edn, The MIT Press, Cambridge, MA.
Stiglitz, Joseph E. (1982), 'Self-selection and pareto efficient taxation', *Journal of Public Economics* **17**(2), 213–240.
Tenhunen, Sanna and Matti Tuomala (2010), 'On optimal lifetime redistribution policy', *Journal of Public Economic Theory* **12**(1), 171–198.

# A Appendix to Section 2

## A.1 Proof of Proposition 1

The first-best allocation is derived by maximizing the government objective function (equation (5)) subject to the aggregate resource constraint (equation (7)). The Lagrangian of this constrained maximization problem is

$$\begin{aligned}\mathcal{L} = {} & \pi \left\{u\left(c_y^L\right) - v\left(l_y^L\right) + \beta\left[U\left(c_o^L\right) - V\left(l_o^L\right)\right]\right\} \\ & + (1-\pi)\left\{u\left(c_y^H\right) - v\left(l_y^H\right) + \beta\left[U\left(c_o^H\right) - V\left(l_o^H\right)\right]\right\} && (51) \\ & + \lambda\Big[\pi\left\{w^L l_y^L + w^L l_o^L - c_y^L - c_o^L\right\} && (52) \\ & + (1-\pi)\left\{w^H l_o^H + w^H l_y^H - c_y^H - c_o^H\right\}\Big], && (53)\end{aligned}$$

and the first order conditions for a maximum with respect to the four consumption levels are

$$\pi u'\left(c_y^L\right) - \lambda\pi = 0 \tag{54a}$$

$$\pi\beta U'\left(c_o^L\right) - \lambda\pi = 0 \tag{54b}$$

$$(1-\pi)\,u'\left(c_y^H\right) - \lambda\,(1-\pi) = 0 \tag{54c}$$

$$(1-\pi)\,\beta U'\left(c_o^H\right) - \lambda\,(1-\pi) = 0, \tag{54d}$$

from which it follows directly that $u'\left(c_y^L\right) = u'\left(c_y^H\right) = \lambda$ and $\beta U'\left(c_o^L\right) = \beta U'\left(c_o^H\right) = \lambda$ so at the first-best solution we have $c_y^L = c_y^H$ and $c_o^L = c_o^H$. The first order conditions with respect to the four labour supply levels are

$$-\pi v'\left(l_y^L\right) + \lambda\pi w^L = 0 \tag{55a}$$

$$-\pi\beta V'\left(l_o^L\right) + \lambda\pi w^L = 0 \tag{55b}$$

$$-\left(1-\pi\right) v'\left(l_y^H\right) + \lambda\,(1-\pi)\,w^H = 0 \tag{55c}$$

$$-\left(1-\pi\right)\beta V'\left(l_o^H\right) + \lambda\,(1-\pi)\,w^H = 0. \tag{55d}$$

It follows that $v'\left(l_y^L\right) = \lambda w^L$ and $v'\left(l_y^H\right) = \lambda w^H$, and since $w_y^L < w_y^H$ we have

$$v'\left(l_y^L\right) < v'\left(l_y^H\right) \Rightarrow l_y^L < l_y^H, \tag{56}$$

where the second inequality follows from the fact that $v' > 0$ and $v'' > 0$. Similarly it follows that $l_o^L < l_o^H$. ■

# B Appendix to Section 4

The Lagrangian $\mathcal{L}$ associated with the second-best tax design problem in terms of observables $I_i^j$ and $c_i^j$, $(i = y, o,\ j = L, H)$ is then

$$\begin{aligned}
\mathcal{L} = {} & \pi^L \left\{ u\left(c_y^L\right) - v\left(\frac{I_y^L}{w_y^L}\right) + \left(U\left(c_o^L\right) - V\left(\frac{I_o^L}{w_o^L}\right)\right)\right\} \\
& + \pi^H \left\{ u\left(c_y^H\right) - v\left(\frac{I_y^H}{w_y^H}\right) + \left(U\left(c_o^H\right) - V\left(\frac{I_o^H}{w_o^H}\right)\right)\right\} \\
& + \lambda \left[ u\left(c_y^H\right) - v\left(\frac{I_y^H}{w_y^H}\right) + \beta\left[U\left(c_o^H\right) - V\left(\frac{I_o^H}{w_o^H}\right)\right] - u\left(c_y^L\right) \right. \\
& \left. + v\left(\frac{I_y^L}{w_y^H}\right) - \beta\left[U\left(c_o^L\right) - V\left(\frac{I_o^L}{w_o^H}\right)\right]\right] \\
& + \mu\left\{\pi^L\left(I_y^L + I_o^L - c_y^L - c_o^L\right) + \pi^H\left(I_y^H + I_o^H - c_y^H - c_o^H\right)\right\},
\end{aligned} \tag{57}$$

where $\mu$ and $\lambda$ are the multipliers associated with the resource and incentive compatibility constraints, respectively.

The first order conditions to this problem are

$$\frac{\partial \mathcal{L}}{\partial c_y^L} = \pi^L u'\left(c_y^L\right) - \lambda u'\left(c_y^L\right) - \pi^L \mu = 0 \tag{58a}$$

$$\frac{\partial \mathcal{L}}{\partial c_y^H} = \pi^H u'\left(c_y^H\right) + \lambda u'\left(c_y^H\right) - \pi^H \mu = 0 \tag{58b}$$

$$\frac{\partial \mathcal{L}}{\partial c_o^L} = \pi^L U'\left(c_o^L\right) - \lambda \beta U'\left(c_o^L\right) - \pi^L \mu = 0 \tag{58c}$$

$$\frac{\partial \mathcal{L}}{\partial c_o^H} = \pi^H U'\left(c_o^H\right) + \lambda \beta U'\left(c_o^H\right) - \pi^H \mu = 0 \tag{58d}$$

$$\frac{\partial \mathcal{L}}{\partial I_y^L} = -\pi^L v'\left(\frac{I_y^L}{w_y^L}\right)\frac{1}{w_y^L} + \lambda v'\left(\frac{I_y^L}{w_y^H}\right)\frac{1}{w_y^H} + \pi^L \mu = 0 \tag{58e}$$

$$\frac{\partial \mathcal{L}}{\partial I_y^H} = -\pi^H v'\left(\frac{I_y^H}{w_y^H}\right)\frac{1}{w_y^H} - \lambda v'\left(\frac{I_y^H}{w_y^H}\right)\frac{1}{w_y^H} + \pi^H \mu = 0 \tag{58f}$$

$$\frac{\partial \mathcal{L}}{\partial I_o^L} = -\pi^L V'\left(\frac{I_o^L}{w_o^L}\right)\frac{1}{w_o^L} + \lambda \beta V'\left(\frac{I_o^L}{w_o^H}\right)\frac{1}{w_o^H} + \pi^L \mu = 0 \tag{58g}$$

$$\frac{\partial \mathcal{L}}{\partial I_o^H} = -\pi^H V'\left(\frac{I_o^H}{w_o^H}\right)\frac{1}{w_o^H} - \lambda \beta V'\left(\frac{I_o^H}{w_o^H}\right)\frac{1}{w_o^H} + \pi^H \mu = 0 \tag{58h}$$

## B.1 MRS for High-Productivity Households

Looking into the first order conditions for optimal consumption and labour supply of high-productivity young individuals we find that the marginal rate of substitution of income for consumption is 1

$$u'\left(c_y^H\right) + \frac{\lambda}{\pi^H} u'\left(c_y^H\right) = v'\left(\frac{y_y^H}{w_y^H}\right)\frac{1}{w_y^H} + \frac{\lambda}{\pi^H} v'\left(\frac{y_y^H}{w_y^H}\right)\frac{1}{w_y^H}$$

$$\frac{v'\left(\frac{y_y^H}{w_y^H}\right)}{u'\left(c_y^H\right)}\frac{1}{w_y^H} = 1.$$

In other words, at the margin there is no distortion on high-productivity young households. The same applies for high-productivity old households

$$U'\left(c_o^H\right)+\frac{\lambda}{\pi^H}\beta U'\left(c_o^H\right)=V'\left(\frac{I_o^H}{w_o^H}\right)\frac{1}{w_o^H}+\frac{\lambda}{\pi^H}\beta V'\left(\frac{I_o^H}{w_o^H}\right)\frac{1}{w_o^H}$$
$$\frac{V'\left(\frac{I_o^H}{w_o^H}\right)}{U'\left(c_o^H\right)}\frac{1}{w_o^H}=1.$$

## B.2 MRS for Low-Productivity Households

Turning to the optimal consumption-labour bundle for young low-productivity households, we have from the first order conditions that

$$u'\left(c_y^L\right)\left(1-\frac{\lambda}{\pi^L}\right)=v'\left(\frac{I_y^L}{w_y^L}\right)\frac{1}{w_y^L}-\frac{\lambda}{\pi^L}v'\left(\frac{I_y^L}{w_y^H}\right)\frac{1}{w_y^H} \tag{59}$$

and since $w_y^H > w_y^L$ and $v'' > 0$, we have

$$u'\left(c_y^L\right)\left(1-\frac{\lambda}{\pi^L}\right)=v'\left(\frac{I_y^L}{w_y^L}\right)\frac{1}{w_y^L}-\frac{\lambda}{\pi^L}v'\left(\frac{I_y^L}{w_y^H}\right)\frac{1}{w_y^H} \tag{60}$$

$$>v'\left(\frac{I_y^L}{w_y^L}\right)\frac{1}{w_y^L}-\frac{\lambda}{\pi^L}v'\left(\frac{I_y^L}{w_y^H}\right)\frac{1}{w_y^L} \tag{61}$$

$$>v'\left(\frac{I_y^L}{w_y^L}\right)\frac{1}{w_y^L}-\frac{\lambda}{\pi^L}v'\left(\frac{I_y^L}{w_y^L}\right)\frac{1}{w_y^L}, \tag{62}$$

which implies

$$u'\left(c_y^L\right)\left(1-\frac{\lambda}{\pi^L}\right)>v'\left(\frac{I_y^L}{w_y^L}\right)\frac{1-\frac{\lambda}{\pi^L}}{w_y^L} \tag{63}$$

$$\frac{v'\left(\frac{I_y^L}{w_y^L}\right)}{u'\left(c_y^L\right)}\frac{1}{w_y^L}<1 \tag{64}$$

Note that $u'\left(c_y^L\right)\left(1-\frac{\lambda}{\pi^L}\right)=\mu>0$, so the sign of the inequality is preserved. With the marginal rate of substitution income for consumption less than 1 implies that the marginal tax rate on labour income of the low-productivity young is strictly positive.

A similar argument applies to the allocation of low-productivity old households:

$$U'\left(c_o^L\right) - \frac{\lambda}{\pi^L}\beta U'\left(c_o^L\right) = V'\left(\frac{I_o^L}{w_o^L}\right)\frac{1}{w_o^L} - \frac{\lambda}{\pi^L}\beta V'\left(\frac{I_o^L}{w_o^H}\right)\frac{1}{w_o^H} \tag{65}$$

$$U'\left(c_o^L\right)\left(1 - \frac{\lambda}{\pi^L}\beta\right) = V'\left(\frac{I_o^L}{w_o^L}\right)\frac{1}{w_o^L} - \frac{\lambda}{\pi^L}\beta V'\left(\frac{I_o^L}{w_o^H}\right)\frac{1}{w_o^H} \tag{66}$$

$$> V'\left(\frac{I_o^L}{w_o^L}\right)\frac{1}{w_o^L} - \frac{\lambda}{\pi^L}\beta V'\left(\frac{I_o^L}{w_o^H}\right)\frac{1}{w_o^L} \tag{67}$$

$$> V'\left(\frac{I_o^L}{w_o^L}\right)\frac{1}{w_o^L} - \frac{\lambda}{\pi^L}\beta V'\left(\frac{I_o^L}{w_o^L}\right)\frac{1}{w_o^H}, \tag{68}$$

since $w_o^L < w_o^H$ and $V'' > 0$, and we have again

$$\frac{V'\left(\frac{I_o^L}{w_o^L}\right)}{U'\left(c_o^L\right)}\frac{1}{w_o^L} < 1, \tag{69}$$

so at the margin, labour income of the low-productivity young is taxed at a positive rate.

## B.3 Intertemporal MRS

Solving for the optimal rate of substitution of consumption when young for consumption when old for low-productivity households yields

$$\pi^L u'\left(c_y^L\right) - \lambda u'\left(c_y^L\right) = \pi^L U'\left(c_o^L\right) - \lambda\beta U'\left(c_o^L\right) \tag{70}$$

$$u'\left(c_y^L\right)\left(\pi^L - \lambda\right) = U'\left(c_o^L\right)\left(\pi^L - \lambda\beta\right) \tag{71}$$

$$\frac{u'\left(c_y^L\right)}{U'\left(c_o^L\right)} = \frac{\pi^L - \lambda\beta}{\pi^L - \lambda}. \tag{72}$$

Turning to the intertemporal marginal rate of substitution of high-productivity households, we find

$$\pi^H u'\left(c_y^H\right) + \lambda u'\left(c_y^H\right) = \pi^H U'\left(c_o^H\right) + \lambda\beta U'\left(c_o^H\right) \tag{73}$$

$$\left(\pi^H + \lambda\right) u'\left(c_y^H\right) = \left(\pi^H + \lambda\beta\right) U'\left(c_o^H\right) \tag{74}$$

$$\frac{u'\left(c_y^H\right)}{U'\left(c_o^H\right)} = \frac{\pi^H + \lambda\beta}{\pi^H + \lambda}. \tag{75}$$

The intertemporal optimality condition could easily be expressed in terms of the optimal sequence of labour supply which yields exactly the same condition that

$$\frac{v'\left(\frac{I_y^H}{w_y^H}\right)\frac{1}{w_y^H}}{V'\left(\frac{I_o^H}{w_o^H}\right)\frac{1}{w_o^H}} = \frac{\pi^H + \lambda\beta}{\pi^H + \lambda} \tag{76}$$

for high productivity households, and

$$\frac{v'\left(\frac{I_y^L}{w_y^L}\right)\frac{1}{w_y^L}}{V'\left(\frac{I_o^L}{w_o^L}\right)\frac{1}{w_o^L}} = \frac{\pi^L - \lambda\beta}{\pi^L - \lambda} \tag{77}$$

for low productivity households.

# C Appendix to Section 5

| Benchmark | | Present bias | | Population share | | Old age labour | |
|---|---|---|---|---|---|---|---|
| $\beta = \pi^L = \varepsilon_o = 0.5$ | | $\beta = 0.75$ | $\beta = 0.9$ | $\pi^L = .75$ | $\pi^L = 0.9$ | $\varepsilon_o = 0.25$ | $\varepsilon_o = 1$ |
| Allocation | | | | | | | |
| $c_y^L$ | 3.966 | 4.064 | 4.141 | 3.791 | 3.707 | 4.140 | 3.749 |
| $c_y^H$ | 6.146 | 6.041 | 5.970 | 5.955 | 5.855 | 6.158 | 6.105 |
| $c_o^L$ | 4.511 | 4.312 | 4.232 | 4.061 | 3.814 | 4.645 | 4.338 |
| $c_o^H$ | 5.601 | 5.794 | 5.879 | 5.143 | 4.888 | 5.653 | 5.516 |
| $I_y^L$ | 5.109 | 5.097 | 5.086 | 5.456 | 5.684 | 5.049 | 5.191 |
| $I_y^H$ | 9.128 | 9.206 | 9.260 | 9.273 | 9.351 | 9.118 | 9.158 |
| $I_o^L$ | 1.262 | 1.264 | 1.265 | 1.362 | 1.430 | 1.579 | 0.828 |
| $I_o^H$ | 4.724 | 4.645 | 4.611 | 4.930 | 5.057 | 4.849 | 4.532 |
| $\mu$ | 0.108 | 0.098 | 0.090 | 0.094 | 0.049 | 0.194 | 0.120 |
| $\lambda$ | 0.198 | 0.198 | 0.198 | 0.231 | 0.255 | 0.098 | 0.203 |
| Lifetime Average Tax Rates | | | | | | | |
| $Tax^L$ | −0.331 | −0.317 | −0.318 | −0.152 | −0.057 | −0.325 | −0.344 |
| $Tax^H$ | 0.152 | 0.145 | 0.146 | 0.219 | 0.254 | 0.154 | 0.151 |
| Wedges | | | | | | | |
| $MRS_y^L$ | 0.828 | 0.845 | 0.857 | 0.903 | 0.958 | 0.844 | 0.808 |
| $MRS_o^L$ | 0.898 | 0.861 | 0.846 | 0.941 | 0.974 | 0.903 | 0.898 |
| $MRS_y^H$ | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| $MRS_o^H$ | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| $MRS_{int}^L$ | 2.275 | 1.414 | 1.136 | 2.143 | 2.058 | 2.244 | 2.314 |
| $MRS_{int}^H$ | 1.823 | 1.279 | 1.094 | 1.727 | 1.670 | 1.936 | 1.807 |

Table 6: Second-best allocation, average taxes and wedges in the scenarios considered in section 5.

# D Appendix to Section 7

Solution to the optimality problem facing the young household:

$$\frac{v'(l_y^i)}{u'(c_y^i)} = w_y^i(1 - T_{y,I}^i) \tag{78}$$

$$\frac{V'(l_o^i)}{U'(c_o^i)} = w_o^i(1 - T_{o,I}^i) \tag{79}$$

$$\frac{u'(c_y^i)}{U'(c_o^i)} = \beta\delta\frac{1 - T_{o,s_y}^i}{1 + T_{y,s_y}^i} \tag{80}$$

$$\frac{U'(c_o^i)}{U'(c_r^i)} = \delta\frac{1 - T_{r,s_o}^i}{1 + T_{o,s_o}^i}. \tag{81}$$

The solution to the optimality problem of the old household:

$$\frac{V'(\widetilde{l}_o^i)}{U'(\widetilde{c}_o^i)} = w_o^i(1 - T_{o,\widetilde{I}}^i) \tag{82}$$

$$\frac{U'(\widetilde{c}_o^i)}{U'(\widetilde{c}_r^i)} = \beta\delta\frac{1 - T_{r,\widetilde{s}_o}^i}{1 + T_{o,\widetilde{s}_o}^i}. \tag{83}$$

**Economics Working Papers**

2011-18: Philipp J.H. Schröder and Allan Sørensen: A welfare ranking of multilateral reductions in real and tariff trade barriers when firms are heterogenous

2012-01: Maria K. Humlum, Jannie H.G. Kristoffersen and Rune Vejlin: Timing of College Enrollment and Family Formation Decisions

2012-02: Kenneth L. Sørensen and Rune Vejlin: Return to Experience and Initial Wage Level: Do Low Wage Workers Catch Up?

2012-03: Peter Arendorf Bache: A Dynamic Model of Trade with Heterogeneous Firms

2012-04: Marianna Marino, Pierpaolo Parrotta and Dario Pozzoli: Does Labor Diversity Promote Entrepreneurship?

2012-05: Peter Arendorf Bache and Anders Laugesen: Trade Liberalization, Mergers and Acquisitions, and Intra-Industry Reallocations

2012-06: Gustaf Bruze, Michael Svarer and Yoram Weiss: The Dynamics of Marriage and Divorce

2012-07: Allan Sørensen: Additive versus multiplicative trade costs and the gains from trade

2012-08: Niclas Berggren and Christian Bjørnskov: Does Religiosity Promote Property Rights and the Rule of Law?

2012-09: Daniel Monte and Norovsambuu Tumennasan: Centralized Allocation in Multiple Markets

2012-10: Daniel Monte and Norovsambuu Tumennasan: Matching with Quorums

2012-11: Ioana Neamtu and Niels Westergård-Nielsen: How are firms affected by the crisis and how do they react?

2012-12: Lasse Frisgaard Gunnersen and Bo Sandemann Rasmussen: Optimal Tax-Transfer Policies, Life-Cycle Labour Supply and Present-Biased Preferences