Journal of Economic Behavior & Organization 101 (2014) 1–23

Contents lists available at ScienceDirect

Journal of Economic Behavior & Organization

journal homepage: www.elsevier.com/locate/jebo

Positional preferences in time and space: Optimal income taxation with dynamic social comparisons[☆,☆☆]

Thomas Aronsson[a], Olof Johansson-Stenman[b,*]

[a] Department of Economics, Umeå School of Business and Economics, Umeå University, S-901 87 Umeå, Sweden
[b] Department of Economics, School of Business, Economics and Law, University of Gothenburg, S-405 30 Gothenburg, Sweden

ARTICLE INFO

Article history:
Received 15 January 2013
Received in revised form
19 November 2013
Accepted 14 January 2014
Available online 7 February 2014

JEL classification:
D62
H21
H23

Keywords:
Optimal income taxation
Asymmetric information
Relative consumption
Status
Habit formation
Positional goods

ABSTRACT

This paper concerns optimal redistributive non-linear income taxation in an OLG model, where people care about their own consumption relative to (i) other people's current consumption, (ii) own past consumption, and (iii) other people's past consumption. We show that both (i) and (iii) affect the marginal income tax structure whereas (ii) does not. We also derive conditions under which atemporal and intertemporal consumption comparisons give rise to exactly the same tax policy responses. On the basis of the available empirical estimates, comparisons with other people's current and past consumption tend to substantially increase the optimal marginal labor income tax rates. Yet, such comparisons may either increase or decrease the optimal marginal capital income tax rates.

1. Introduction

A rapidly growing body of evidence suggests that people have positional preferences in the sense of deriving utility from their own consumption relative to that of others.[1] Alongside this development, a corresponding literature dealing with

☆☆ We would like to thank three anonymous referees, Thomas Gaube, Katarina Nordblom, Ola Olsson, Måns Söderbom, and Joakim Westerlund, as well as participants at the IIPF conference in Uppsala and the National Conference of Swedish Economists in Lund, and seminar participants at the Center of Economic Research at ETH, Zurich, for helpful comments and suggestions. Research grants from the Swedish Research Council (ref. 421-2010-1420) are gratefully acknowledged.

* Corresponding author. Tel.: +46 31 786 25 38.
E-mail addresses: Thomas.Aronsson@econ.umu.se (T. Aronsson), Olof.Johansson@economics.gu.se (O. Johansson-Stenman).

[1] This includes evidence from happiness research (e.g., Easterlin, 1995, 2001; Blanchflower and Oswald, 2004; Ferrer-i-Carbonell, 2005; Luttmer, 2005; Clark and Senik, 2010) and questionnaire-based studies (e.g., Johansson-Stenman et al., 2002; Solnick and Hemenway, 2005; Carlsson et al., 2007). See also, e.g., Marmot (2004) and Daly and Wilson (2009) for evidence based on physiological studies, and Fliessbach et al. (2007) and Dohmen et al. (2011) for evidence based on brain science. Bowles and Park (2005) and Oh et al. (2012) analyze variations in work hours between countries and over time, and find that social comparisons are important driving forces behind changes in work hours.

0167-2681/$ – see front matter
http://dx.doi.org/10.1016/j.jebo.2014.01.004

optimal policy responses to positional concerns has evolved, showing that such concerns may have a substantial effect on the incentive structure underlying public policy. Within the literature on optimal income taxation,[2] it has for example been shown that social comparisons may motivate substantially higher marginal income tax rates than without such comparisons; see, e.g., Boskin and Sheshinski (1978), Layard (1980), Oswald (1983), Tuomala (1990), Blomquist (1993), Ireland (2001), Aronsson and Johansson-Stenman (2008, 2010, 2013), Wendner and Goulder (2008) and Wendner (2010a).

Yet, almost all earlier studies on optimal policy responses to positional concerns that we are aware of assume that people only make "atemporal" consumption comparisons, by valuing their own current consumption relative to other people's current consumption. A much more general approach has recently been presented by Rayo and Becker (2007). According to their evolutionary model, selfish genes would prefer that the humans they belong to are motivated by their own current consumption relative to (i) their own past consumption, (ii) other people's current consumption, and (iii) other people's past consumption. In the macroeconomic literature of dynamic consumption behavior, (i) corresponds to what is typically denoted *habit formation* (sometimes denoted *internal habit formation*), (ii) corresponds to *keeping up with the Joneses*, while (iii) corresponds to *catching up with the Joneses* (sometimes denoted *external habit formation*).[3] The present paper takes these three types of consumption comparisons as a point of departure in a study of optimal income taxation in a dynamic economy.[4]

We develop and analyze an overlapping generations (OLG) model with endogenous labor supply and savings, where the consumers are concerned with their relative consumption and where nonlinear taxes of labor income and capital income are used for purposes of externality correction and redistribution. A dynamic model allows us to explore intertemporal aspects of consumption comparisons, and provides a natural framework for studying capital income taxation. The latter is important not least due to the difficulties of explaining the widespread use of capital taxes with conventional public economics models. Earlier research shows that relative consumption concerns may motivate such taxes (Aronsson and Johansson-Stenman, 2010), and one might perhaps conjecture such concerns to be particularly important when the concept of relative consumption has more than one dimension, as we assume here.

The literature on optimal redistributive taxation under relative consumption concerns is scarce, and almost all earlier studies are based on static models. The only exception that we are aware of is Aronsson and Johansson-Stenman (2010), who analyze optimal nonlinear income taxation in a dynamic economy where each consumer compares his/her own current consumption with other people's current consumption. Hence, their study neglects internal habit formation as well as the catching up with the Joneses type of comparison mentioned above, and focuses solely on consumption comparisons based on keeping up with the Joneses preferences. The present paper, in contrast, addresses the implications of such atemporal comparisons for optimal income taxation simultaneously with the implications of relative consumption comparisons over time. Another study related to ours is Ljungqvist and Uhlig (2000), who consider optimal labor income taxation in a dynamic representative agent model where the consumer preference for relative consumption is driven by a catching up with the Joneses motive. We generalize their approach in several different ways by (1) considering a broader set of tax instruments, (2) analyzing redistribution policy alongside externality correction, and (3) allowing keeping up and catching up with the Joneses mechanisms to be operative simultaneously.

These extensions are important. In addition to the empirical evidence for between-people comparisons mentioned above, there is evidence suggesting that people also make comparisons with their own past consumption (e.g., Loewenstein and Sicherman, 1991; Frank and Hutchens, 1993); indeed, such comparisons were discussed already by Veblen (1899). It also makes intuitive sense that old people compare their own consumption with several different reference levels, including what they recall about their own and others' consumption when they were young. Moreover, when growing up, most people are likely to receive information from parents and grandparents about the consumption (and other living conditions) characterizing earlier generations. The results from happiness studies have also documented that people's happiness adapts to income changes, consistent with the idea that the reference income increases over time when actual income increases; see, e.g., Stutzer (2004) and Di Tella et al. (2010). Specifically, Senik (2009) presents recent estimates regarding the importance of different kinds of comparisons over time, showing that subjective well-being is dependent on one's own standard of living

[2] Other issues dealt with include public good provision (Ng, 1987; Aronsson and Johansson-Stenman, 2008; Wendner and Goulder, 2008; Wendner, 2014), social insurance (Abel, 2005), growth (Corneo and Jeanne, 1997, 2001; Wendner, 2010b, 2011), environmental externalities (Howarth, 1996, 2006; Brekke and Howarth, 2002; Wendner, 2005), stabilization policy (Ljungqvist and Uhlig, 2000), distributional consequences (Alvarez-Cuadrado and Long, 2012) and tax evasion (Goerke, 2013). See also Frank (1999, 2005, 2007, 2008) for extensive and illuminating informal discussions of relative consumption concerns and how the society should deal with them.

[3] The notion "keeping up with the Joneses" is unfortunately used with different meanings in the literature. It is either used to indicate social comparisons in the sense that my utility depends in part on my current consumption relative to your current consumption, as in our case, or it is used with more specific meanings, e.g., if you consume more now I will also consume more now. Similarly, the notion "catching up with the Joneses" may either, as here, simply mean that my utility today depends on my current consumption relative to your previous consumption, or it may reflect something more specific such that my consumption today increases with your previous consumption. No results in the present paper depend directly on the direction of people's consumption and leisure adjustment in response to a change in the reference consumption.

[4] The macroeconomics literature referred to above rarely analyzes the optimal policy responses to the externalities induced by relative consumption concerns. Ljungqvist and Uhlig (2000) and Gomez (2006) are two noteworthy exceptions.

relative to both internal and external reference points. Such comparisons are also consistent with the empirical pattern of some financial puzzles,[5] and (as mentioned) they are in line with recent research based on evolutionary models.

Section 2 presents the model and the outcome of private optimization. In Section 3, we consider optimal income taxation in a first-best setting with full information about individual productivity, where the government can deal with its distributional objectives through lump-sum taxes. This provides a simple benchmark by which to compare our results with some earlier literature on optimal labor income taxation under positional concerns (which is based on similar models). Our results show that relative comparisons with one's own past consumption (internal habit formation) do not directly affect the policy *rules* for marginal income taxation (although they may, of course, influence the *levels* of marginal income tax rates). The intuition is that such comparisons are fully internalized at the individual level and do not generate any externalities. However, positional concerns governed by comparisons with other people's current *and* past consumption give rise to externalities and will, therefore, also directly affect the incentive structure underlying marginal income taxation. We show that positional concerns with respect to other people's current and past consumption tend to increase the optimal marginal labor income tax rates. We also show how the marginal capital income tax rates are governed by differences in positional concerns over the individual life cycle, where the relevant measure of reference consumption is again based on both the current and past consumption of others.

In general, positional concerns governed by other people's past consumption give rise to much more complex policy responses than do comparisons based on other people's current consumption. This is so because consumption comparisons over time give rise to an intertemporal chain reaction with welfare effects in the entire future, whereas comparisons with other people's current consumption only lead to "atemporal externalities." We can nevertheless derive strong results for a natural benchmark case in which the concerns for relative consumption are constant over time, implying that relative consumption comparisons over time (based on the catching up with the Joneses preferences) give rise to exactly the same marginal tax rate responses as comparisons with other people's current consumption (based on the keeping up with the Joneses preferences).

In practice, informational limitations are likely to prevent governments from using differentiated lump-sum taxation as a basis for redistribution. Therefore, in Section 4, we introduce asymmetric information between the government and the private sector with respect to individual ability (worker productivity), where the public decision problem is described by a variant of the two-type optimal income tax model originally developed by Stern (1982) and Stiglitz (1982). Although simple, the two-type model provides a powerful framework for analyzing externality correction and redistribution simultaneously. In such a second-best framework, tax distortions are the outcome of an optimal choice made by the government, subject to informational limitations, and not of any arbitrary restrictions on the tax instruments (such as linearity) or the necessity to raise revenue per se. Therefore, our approach enables us to capture that the optimal income tax responses to positional concerns may involve purely corrective as well as redistributive elements.

In a second-best setting where ability is private information, there is also another policy incentive involved beyond positional externalities: the government may relax the incentive constraint by exploiting differences in positional concerns across ability types. Our results show that while the externality-correcting mechanism unambiguously works to increase the marginal labor income tax rates, independently of whether individuals compare their own current consumption with other people's current or past consumption (or use a combination of these two reference measures), the direction of the mechanism through the incentive constraint is ambiguous. We both present general optimal taxation results and derive sufficient conditions for when the overall net effect of positional concerns works to increase the marginal labor income tax rates. Section 5 illustrates with a particular Cobb-Douglas functional form and shows, based on parameter estimates from the literature, that positional preferences of both the keeping up with the Joneses and the catching up with the Joneses types substantially increase the optimal marginal labor income tax rates. Section 6 summarizes and concludes the paper; proofs are presented in the Appendix.

2. Consumers, firms, and market equilibrium

We start this section by describing the OLG framework and people's preferences, followed by the definition of some useful measures of the extent to which people care about relative consumption. We then present the individual optimality conditions for labor supply and savings, followed by the corresponding profit maximization conditions for the firms and the conditions for market equilibrium.

2.1. The OLG framework and positional preferences

Consider an OLG model where each individual lives for two periods and works during the first but not during the second. Since each individual only works during the first period of life, there is no evolution of productivity over time for a single individual, as in Kocherlakota (2005), although we allow for technical progress (discussed subsequently) that makes labor productivity increase over time. Individuals differ in ability, as measured by the before-tax wage rate. The number of

[5] This includes various kinds of asset pricing puzzles, such as the equity premium puzzle; see, e.g., Abel (1990), Constantinides (1990), Campbell and Cochrane (1999), Chan and Kogan (2002), and Díaz et al. (2003).

individuals of ability-type i in generation t, i.e., who were born at the beginning of period t, is denoted n_t^i. Each such individual derives utility from his/her consumption when young, c_t^i, consumption when old, x_{t+1}^i, and use of leisure when young, z_t^i, given by a time endowment normalized to one less the hours of work, l_t^i (when old, all available time is leisure). For further use, we define the average consumption in the economy as a whole in period t as $\bar{c}_t = \left[\sum_i n_t^i c_t^i + \sum_i n_{t-1}^i x_t^i\right] / \sum_i \left[n_t^i + n_{t-1}^i\right]$.

People also care about their own consumption relative to that of others.[6] In accordance with the bulk of earlier comparable literature, we focus on *difference* comparisons, where relative consumption is defined by the difference between the individual's own consumption and a measure of reference consumption.[7] The appropriate measure of reference consumption at the individual level is, of course, an empirical question; yet, as indicated above, there is very little information available. Our approach is to follow the recent contribution by Rayo and Becker (2007), who argue in the context of an evolutionary model of happiness that the reference point of an individual might be determined by three components: (i) other people's current consumption, (ii) his/her own past consumption, and (iii) other people's past consumption. In terms of our model, we interpret these three components such that people care about four different kinds of relative consumption: the first three are their own current consumption relative to (i) the current average consumption such that $c_t^i - \bar{c}_t$ and $x_{t+1}^i - \bar{c}_{t+1}$ are the corresponding measures of relative consumption when young and when old, respectively; (ii) their own consumption one period earlier, i.e., $x_{t+1}^i - c_t^i$; and (iii) the average consumption one period earlier such that $c_t^i - \bar{c}_{t-1}$ and $x_{t+1}^i - \bar{c}_t$ are the corresponding measures of relative consumption when young and when old, respectively. In addition, following Rayo and Becker (2007), we may also allow for the change in relative consumption to matter, such that individual utility depends on $(x_{t+1}^i - \bar{c}_{t+1}) - (c_t^i - \bar{c}_t)$. This fourth comparison type is interpretable as habit formation in relative consumption (compared with others' current consumption), i.e., that the individual prefers increased relative consumption over time, ceteris paribus. All in all, we can then write the utility function of ability-type i in generation t in terms of absolute consumption, leisure, and six different relative consumption comparisons as follows:

$$U_t^i = V_t^i(c_t^i, z_t^i, x_{t+1}^i, x_{t+1}^i - c_t^i, c_t^i - \bar{c}_t, x_{t+1}^i - \bar{c}_{t+1}, c_t^i - \bar{c}_{t-1}, x_{t+1}^i - \bar{c}_t, (x_{t+1}^i - \bar{c}_{t+1}) - (c_t^i - \bar{c}_t)). \tag{1a}$$

The function $V_t^i(\cdot)$ is assumed to be increasing in each argument, i.e., in leisure, absolute consumption when young and when old, and the six measures of relative consumption.

However, note that the internal (habit-based) consumption comparison can be expressed in terms of c_t^i and x_{t+1}^i, which are decision variables of the individual. As such, we can without loss of generality rewrite Eq. (1a) as the following "reduced form" function:

$$U_t^i = h_t^i(c_t^i, z_t^i, x_{t+1}^i, c_t^i - \bar{c}_t, x_{t+1}^i - \bar{c}_{t+1}, c_t^i - \bar{c}_{t-1}, x_{t+1}^i - \bar{c}_t, (x_{t+1}^i - \bar{c}_{t+1}) - (c_t^i - \bar{c}_t)), \tag{1b}$$

where the internal habit formation component, $x_{t+1}^i - c_t^i$, is now embedded in the effects of c_t^i and x_{t+1}^i. This implies that the partial derivative of $h_t^i(\cdot)$ with respect to c_t^i reflects both the direct utility gain of increased absolute consumption when young and the utility loss due to lower relative consumption when old compared to when young; correspondingly, the partial derivative with respect to x_{t+1}^i reflects the direct utility gain of increased absolute consumption when old plus the gain of increased consumption when old compared to when young. Therefore, all analytical results derived in a model where individuals do not compare their own current and past consumption will continue to hold also in the case where people make such comparisons. Intuitively, rational utility-maximizing individuals will internalize such comparisons perfectly, and there is no externality involved in internal habit formation.[8]

Note also that the last component on the right-hand side of Eq. (1b), i.e., the change in the relative consumption over the life cycle, is given by the difference between the fifth $(x_{t+1}^i - \bar{c}_{t+1})$ and fourth $(c_t^i - \bar{c}_t)$ arguments, such that we can obtain a further reduced form as follows:

$$U_t^i = v_t^i(c_t^i, z_t^i, x_{t+1}^i, c_t^i - \bar{c}_t, x_{t+1}^i - \bar{c}_{t+1}, c_t^i - \bar{c}_{t-1}, x_{t+1}^i - \bar{c}_t). \tag{1c}$$

The intuition is that the effect of changed relative consumption over the life cycle, i.e., $(x_{t+1}^i - \bar{c}_{t+1}) - (c_t^i - \bar{c}_t)$, is now embodied in the effects through $c_t^i - \bar{c}_t$ and $x_{t+1}^i - \bar{c}_{t+1}$. Thus, if an individual derives utility from increased relative consumption

[6] We follow earlier comparable literature in assuming that people do not care about their relative leisure; see Arrow and Dasgupta (2009) and Aronsson and Johansson-Stenman (2013) for analysis of the case where also relative leisure matters.

[7] See, e.g., Akerlof (1997), Corneo and Jeanne (1997), Ljungqvist and Uhlig (2000), Bowles and Park (2005), Carlsson et al. (2007), and Alvarez-Cuadrado and Long (2011, 2012). Alternative approaches include ratio comparisons (Boskin and Sheshinski, 1978; Layard, 1980; Abel, 2005; Wendner and Goulder, 2008) and comparisons of ordinal rank (Frank, 1985; Hopkins and Kornienko, 2004, 2009). Dupor and Liu (2003) consider a flexible functional form that includes the difference and ratio comparisons as special cases. Mujcic and Frijters (2013) compare models with difference comparisons, ratio comparisons, and rank-based comparisons without being able to discriminate between them.

[8] This approach is in principle innocuous as long as the government commits to its future tax policy, as we will assume below. Without such commitment, internal habit formation may influence the optimal tax structure in a second-best setting; see Guo and Krause (2011). However, whether people in reality do internalize such adaptation perfectly is another matter. Loewenstein et al. (2003) analyze how systematic errors in anticipating adaptation processes affect the consumption-savings choices people make, and show that anticipation bias may lead people to make decisions that are not in their own best interest. Tax policy implications of anticipation bias are addressed by Aronsson and Schöb (2013), where the corrective policies that a paternalist government may undertake are embedded into a framework with redistribution under asymmetric information. Further exploration of the case with such bounded rationality is a recommended task for future research.

over the life cycle, this will in Eq. (1c) be reflected in a larger partial derivative with respect to $x^i_{t+1} - \bar{c}_{t+1}$ and a smaller partial derivative with respect to $c^i_t - \bar{c}_t$, ceteris paribus. This has an important implication: analytical results derived based on Eq. (1c) hold irrespective of whether people derive utility from increased relative consumption over the life cycle. In the following, and in order not to make the analysis overly complex, we will not deal with the effect through $(x^i_{t+1} - \bar{c}_{t+1}) - (c^i_t - \bar{c}_t)$ directly but rely on the formulation in Eq. (1c). Instead, we return to the effect of increased or decreased relative consumption over the life cycle when discussing interpretations and orders of magnitudes in Section 5.

For further use, note finally that the utility function can be rewritten as a "fully reduced form" often found in classic externality problems as follows:

$$U^i_t = u^i_t(c^i_t, z^i_t, x^i_{t+1}, \bar{c}_{t-1}, \bar{c}_t, \bar{c}_{t+1}). \tag{1d}$$

Eq. (1d) can also be thought of as the most general formulation of the utility function, since it contains no information about the structure of the social comparisons, beyond that others' consumption levels cause negative externalities. We assume that the function $v^i_t(\cdot)$ is increasing in each argument, meaning that $u^i_t(\cdot)$ is decreasing in $\bar{c}_t$, $\bar{c}_{t-1}$, and $\bar{c}_{t+1}$, and that Eqs. (1c) and (1d) are twice continuously differentiable in their respective arguments and strictly concave.

As will be demonstrated, for some results we do not need any stronger assumptions regarding the preference structure than in Eq. (1d). Yet, we need the more restrictive utility formulation based on the function $v^i_t(\cdot)$ in Eq. (1c), where we specify that people care about additive comparisons, to establish a relationship between the optimal tax policy and the degree to which the utility gain from higher consumption is associated with increased relative consumption.[9] The definition of such measures is the issue to which we turn next.

2.2. The degree of current versus intertemporal consumption positionality

Since much of the subsequent analysis is focused on relative consumption concerns, it is useful to introduce measures of the degree to which such concerns matter for each individual. By using the function $v^i_t(\cdot)$ in Eq. (1c), we can define the *degree of current consumption positionality* when young and when old, respectively, as

$$\alpha^{i,c}_t = \frac{v^i_{t,c^i_t-\bar{c}_t}}{v^i_{t,c^i_t-\bar{c}_t} + v^i_{t,c^i_t-\bar{c}_{t-1}} + v^i_{t,c}}, \tag{2a}$$

$$\alpha^{i,x}_{t+1} = \frac{v^i_{t,x^i_{t+1}-\bar{c}_{t+1}}}{v^i_{t,x^i_{t+1}-\bar{c}_{t+1}} + v^i_{t,x^i_{t+1}-\bar{c}_t} + v^i_{t,x}}, \tag{2b}$$

where subscripts attached to the utility function (except for the time indicator) denote partial derivatives, i.e., $v^i_{t,c} = \partial v^i_t(\cdot)/\partial c^i_t$ and similarly for the other terms. The variables $\alpha^{i,c}_t$ and $\alpha^{i,x}_{t+1}$ are interpretable as the fraction of the overall utility increase from an additional dollar spent when young in period t and old in period $t+1$, respectively, that is due to the increased consumption relative to other people's current consumption (measured by the average consumption in the same period). As indicated above, these measures are also interpretable as if they accommodate the utility gain to the individual of increased relative consumption over the life cycle.

By analogy, we can define the *degree of intertemporal consumption positionality* when young and when old, respectively, as

$$\beta^{i,c}_t = \frac{v^i_{t,c^i_t-\bar{c}_{t-1}}}{v^i_{t,c^i_t-\bar{c}_t} + v^i_{t,c^i_t-\bar{c}_{t-1}} + v^i_{t,c}}, \tag{3a}$$

$$\beta^{i,x}_{t+1} = \frac{v^i_{t,x^i_{t+1}-\bar{c}_t}}{v^i_{t,x^i_{t+1}-\bar{c}_{t+1}} + v^i_{t,x^i_{t+1}-\bar{c}_t} + v^i_{t,x}} \tag{3b}$$

Eqs. (3a) and (3b) have interpretations similar to Eqs. (2a) and (2b), yet with the obvious modification of reflecting consumption comparisons over time: the variables $\beta^{i,c}_t$ and $\beta^{i,x}_{t+1}$ measure the fraction of the overall utility increase from an additional dollar spent in period t and $t+1$ (i.e., when young and when old), respectively, that is due to the increased consumption relative to other people's past consumption. By our earlier assumptions, $0 < \alpha^{i,c}_t, \alpha^{i,x}_{t+1}, \beta^{i,c}_t, \beta^{i,x}_{t+1} < 1$ for all t.

[9] We do not attempt to explain why people care about relative consumption. We do believe that signaling of some attractive characteristics constitutes a likely reason why people care about their own consumption relative to that of others (cf. e.g. Ireland, 2001), but we still follow the considerably simpler modeling strategy of simply assuming that people's preferences directly depend on relative consumption.

Let us next define the notions of *the average degree of current consumption positionality* and *the average degree of intertemporal consumption positionality*, which are given by

$$\bar{\alpha}_t = \sum_i \alpha_t^{i,x} \frac{n_{t-1}^i}{N_t} + \sum_i \alpha_t^{i,c} \frac{n_t^i}{N_t} \in (0,1), \tag{4a}$$

$$\bar{\beta}_t = \sum_i \beta_t^{i,x} \frac{n_{t-1}^i}{N_t} + \sum_i \beta_t^{i,c} \frac{n_t^i}{N_t} \in (0,1), \tag{4b}$$

respectively, where $N_t = \sum_i [n_{t-1}^i + n_t^i]$. Note that both $\bar{\alpha}_t$ and $\bar{\beta}_t$ are measured among those alive in period t.

2.3. The optimality conditions for individuals and firms

The consumer's intertemporal budget constraint is summarized by the following two equations (for $i = 1, 2$):

$$w_t^i l_t^i - T_t(w_t^i l_t^i) - s_t^i = c_t^i, \tag{5a}$$

$$s_t^i(1 + r_{t+1}) - \Phi_{t+1}(s_t^i r_{t+1}) = x_{t+1}^i, \tag{5b}$$

where w_t^i is the before-tax wage rate, implying that $w_t^i l_t^i$ is the before-tax labor income, s_t^i is savings, r_{t+1} is the market interest rate, and $T_t(\cdot)$ and $\Phi_{t+1}(\cdot)$ denote the payments of labor income and capital income taxes, respectively.[10] Thus, consumption levels when young are given by gross labor income net of labor income taxes and savings, whereas consumption levels when old are given by the sum of savings and capital income net of capital income taxes.

Although the measures of reference consumption are endogenous in our model, we assume that each individual treats them as exogenous, which is the conventional equilibrium assumption in models with externalities. To be more specific, and with reference to Eqs. (1) above, this means that ability-type i of generation t treats $\bar{c}_{t-1}$, $\bar{c}_t$, and $\bar{c}_{t+1}$ as exogenous. The first-order conditions for the hours of work and savings can then be written as

$$u_{t,c}^i w_t^i [1 - T_t'(w_t^i l_t^i)] - u_{t,z}^i = 0, \tag{6}$$

$$-u_{t,c}^i + u_{t,x}^i [1 + r_{t+1}[1 - \Phi_{t+1}'(s_t^i r_{t+1})]] = 0, \tag{7}$$

where $u_{t,c}^i = \partial u_t^i / \partial c_t^i$, $u_{t,z}^i = \partial u_t^i / \partial z_t^i$, and $u_{t,x}^i = \partial u_t^i / \partial x_{t+1}^i$, while $T_t'(w_t^i l_t^i)$ and $\Phi_{t+1}'(s_t^i r_{t+1})$ are the marginal labor income tax rate and marginal capital income tax rate, respectively.

The production sector consists of identical competitive firms producing a homogenous good with constant returns to scale; the number of firms is normalized to one for notational convenience. Following Aronsson and Johansson-Stenman (2010), the production function is given by

$$F(L_t, K_t; t) = g\left(\sum_i \theta^i L_t^i, K_t; t\right), \tag{8}$$

where $L_t^i = n_t^i l_t^i$ is the total number of hours of work supplied by ability-type i in period t, $L_t = [L_t^1, L_t^2, \ldots]$ is a vector whose elements reflect the total number of work hours by each ability-type, K_t is the capital stock in period t, and θ^i (for all i) is a positive constant. The direct time dependency implies that we allow for exogenous technological change. Note that the functional form assumption implicit in $g(\cdot)$ means that the relative wage rates, i.e., $w_t^j / w_t^k = \theta^j / \theta^k$ for all j and k, are fixed. This assumption simplifies the calculations in Section 4, where the public decision problem is affected by asymmetric information between the government and the private sector, but is not important for the policy incentives created by relative consumption comparisons (which are the major concerns here).

The firm obeys the necessary optimality conditions

$$F_{L^i}(L_t, K_t; t) = \frac{\partial g}{\partial \left(\sum_i \theta^i L_t^i\right)} \theta^i = w_t^i \quad \text{for all} \quad i, \tag{9}$$

$$F_K(L_t, K_t; t) = \frac{\partial g}{\partial K_t} = r_t. \tag{10}$$

3. First-best taxation

In this section, we begin by specifying the social objective function, which is maximized subject to the overall resource constraint. Then we present the results in terms of optimal taxation, starting with the case where the average positionality degrees, population size, and the interest rate in the economy are all constant over time. This framework is technically

[10] Note that both the labor and capital income taxation functions are nonlinear and may include both intercept and slope parameters. For example, it seems reasonable that the optimal labor income tax payment is negative at zero labor income.

convenient and allows us to characterize the basic policy incentives in an intuitive way. It also facilitates comparison with earlier studies. Finally, we present the optimal taxation results for the more general model where these assumptions are relaxed.

3.1. The social decision problem

For convenience, as well as comparability with Section 4, we assume that there are two types in each generation, type 1 and type 2.[11] The government faces a general intertemporal social welfare function as follows:

$$W = W(n_0^1 U_0^1, n_0^2 U_0^2, n_1^1 U_1^1, n_1^2 U_1^2, \ldots), \tag{11}$$

which is increasing in each argument. Since the optimum conditions are expressed for any such social welfare function, they are necessary optimum conditions for a Pareto-efficient allocation.

Note that nonlinear taxation of labor and capital income allows the government to implement any desired combination of consumption, savings, and work hours for each individual. Therefore, we follow earlier literature on optimal nonlinear income taxation in dynamic economies by formulating the public decision problem as a direct decision problem in terms of private consumption, work hours, and the capital stock – an approach that will be particularly convenient in Section 4, where we introduce asymmetric information between the government and the private sector.[12] The marginal income tax rates implicit in the optimal resource allocation can then be derived by combining the social first-order conditions with the first-order conditions characterizing the private sector.

The resource constraint for the economy as a whole can be written as

$$F(L_t, K_t; t) + K_t - \sum\nolimits_i \left[n_t^i c_t^i + n_{t-1}^i x_t^i\right] - K_{t+1} = 0, \tag{12}$$

such that production equals consumption plus investment in each period. We abstract from capital depreciation, since such depreciation is of no significance for the qualitative results derived below. The Lagrangian associated with the public decision problem then becomes

$$\mathscr{L} = W + \sum_t \gamma_t \left[F(L_t, K_t; t) + K_t - \sum\nolimits_i [n_t^i c_t^i + n_{t-1}^i x_t^i] - K_{t+1}\right]. \tag{13}$$

Note finally that the government is assumed to treat the measures of reference consumption as endogenous, i.e., the government recognizes and incorporates into its decision problem how the measures of reference consumption change in response to public policy.

We are concerned with the optimal tax policy implemented for any ability-type i of any generation t, which is based on the social first-order conditions for l_t^i, c_t^i, x_{t+1}^i, and K_{t+1}. These social first-order conditions are given by (for all i and t)

$$-\frac{\partial W}{\partial(n_t^i U_t^i)} n_t^i u_{t,z}^i + \gamma_t n_t^i w_t^i = 0, \tag{14a}$$

$$\frac{\partial W}{\partial(n_t^i U_t^i)} n_t^i u_{t,c}^i - \gamma_t n_t^i + \frac{n_t^i}{N_t}\frac{\partial \mathscr{L}}{\partial \bar{c}_t} = 0, \tag{14b}$$

$$\frac{\partial W}{\partial(n_t^i U_t^i)} n_t^i u_{t,x}^i - \gamma_{t+1} n_t^i + \frac{n_t^i}{N_{t+1}}\frac{\partial \mathscr{L}}{\partial \bar{c}_{t+1}} = 0, \tag{14c}$$

$$\gamma_{t+1}[1 + r_{t+1}] - \gamma_t = 0, \tag{14d}$$

where we have used that $w_t^i = F_{L^i}(L_t, K_t; t)$ and $r_t = F_K(L_t, K_t; t)$ from the first-order conditions of the firm. For notational convenience, we have written Eqs. (14b) and (14c) such that the left-hand side contains the partial derivative of the Lagrangian with respect to the appropriate measure of reference consumption, i.e., the measure of reference consumption that is affected by a change in c_t^i and x_{t+1}^i, respectively. The derivative $\partial\mathscr{L}/\partial\bar{c}_t$ will be referred to as the *positionality effect* in period t and plays a crucial role in the subsequent analysis of optimal taxation.

3.2. Optimal first-best taxation with constant average positionality degrees

Let us start with some less general assumptions than those outlined above, and more specifically that:

[11] We show in an earlier version of the paper that all results in this section are straightforward to generalize to an arbitrary number of types.

[12] See, e.g., Brett (1997) and Aronsson and Johansson-Stenman (2010).

i. The population size is constant over time, such that $N_t = N$ for all t.
ii. The average (current and intertemporal) degrees of positionality are constant over time, such that $\bar{\alpha}_t = \bar{\alpha}$ and $\bar{\beta}_t = \bar{\beta}$ for all t.
iii. The interest rate is constant over time, such that $r_t = r$ for all t.

Although these assumptions of course reflect limitations, similar (or stronger) assumptions are typically made in the "catching up with the Joneses" literature, where one also often assumes specific functional forms; see, e.g., Campbell and Cochrane (1999) and Díaz et al. (2003). It should also be noted that the model is still general enough to reflect different preferences between types, including different positionality degrees.

Assumption (iii) above implies from Eq. (14d) that $\gamma_{t+k} = \gamma_t/[1+r]^k$. This special case, which makes it easy to relate our results to findings in earlier literature, is either interpretable in terms of a steady state[13] – provided that a steady state exists – or may follow as a consequence of adding additional assumptions about the preferences and technology (see Section 5).

Now, let us define the *average degree of time-inclusive consumption positionality*, in present value terms, as follows:

$$\bar{\rho} = \bar{\alpha} + \frac{\bar{\beta}}{1+r}.$$

Intuitively, $\bar{\rho}$ reflects the overall social loss of consuming an additional dollar today, due to the associated increase in current and future reference consumption levels. The first term, $\bar{\alpha}$, reflects the part of this loss that will occur through current consumption positionality, whereas the second term, $\bar{\beta}/[1+r]$, reflects the loss due to intertemporal consumption positionality. The reason why the latter loss is discounted is, of course, that it will occur in the next period. We can then derive

$$\frac{\partial \mathscr{L}}{\partial \bar{c}_t} = -N\gamma_t \frac{\bar{\alpha} + \bar{\beta}/[1+r]}{1 - \bar{\alpha} - \bar{\beta}/[1+r]} = -N\gamma_t \frac{\bar{\rho}}{1-\bar{\rho}} \tag{15}$$

Although intertemporal consumption comparisons give rise to welfare costs in the future (as will be shown explicitly in Section 3.3), the forward-looking component reduces to a single variable under assumptions i-iii, which explains the simple form of Eq. (15). Using Eqs. (6), (7), (14a)–(14d) and (15), we are ready to present the following benchmark results:

Proposition 1. *Under assumptions (i)–(iii), the first-best marginal labor and capital income tax rates, respectively, can be written as (for i = 1, 2)*

$$T_t'(w_t^i l_t^i) = \bar{\rho} > 0, \tag{16}$$

$$\Phi_{t+1}'(s_t^i r_{t+1}) = 0. \tag{17}$$

Proof. See the Appendix. □

To interpret Proposition 1, note first that since the government may reach its distributional objectives by using lump-sum taxes/subsidies, there is no distributional reason for using distortionary taxation. As such, the non-zero marginal labor income tax rates are solely due to the externalities that positional concerns give rise to. By analogy to earlier comparable literature (see Section 1), we find that positional concerns motivate positive marginal labor income tax rates. The novelty here is that the marginal labor income tax rate is given by the average degree of *time-inclusive* positionality; there are no additional effects associated with the two separate components $\bar{\alpha}$ and $\bar{\beta}/(1+r)$. This has a strong implication: current positionality (reflecting the keeping up with the Joneses motive) and intertemporal positionality (reflecting the catching up with the Joneses motive) affect the marginal labor income tax rates in *exactly* the same way. Note also that Eq. (16) nests corresponding results derived by, e.g., Dupor and Liu (2003) (where $\bar{\rho} = \bar{\alpha}$ and $T_t'(w_t^i l_t^i) = \bar{\alpha}$) and Ljungqvist and Uhlig (2000) (where $\bar{\rho} = \bar{\beta}/(1+r)$ and $T_t'(w_t^i l_t^i) = \bar{\beta}/(1+r)$).[14]

Second, capital income taxation plays no corrective role here. The intuition is that the positionality effects in periods t and $t+1$ cancel out. Therefore, as long as assumptions (i)–(iii) hold, the marginal capital income tax rates are zero in the first best, irrespective of whether the consumers have keeping up or catching up with the Joneses preferences (or a mix of them).

3.3. Optimal first-best taxation with time-varying degrees of positionality

As indicated above, the optimal tax policy depends on the consumption externalities, which in turn depend on the positionality effect. As long as we made the simplifying assumptions (i)–(iii), both the positionality effect and the optimal

[13] This requires that the preferences and technology do not change over time, and that the economy approaches a stationary equilibrium in which l_t^i, c_t^i, x_t^i (for all i), and K_t all remain constant over time.

[14] Ljungqvist and Uhlig use a model with infinite time horizons to describe the consumer behavior, and assume that the reference consumption relevant today is a geometric average of the consumption in earlier periods. The result discussed above refers to their special case where the reference consumption is given by the average consumption in the previous period. The study by Dupor and Liu is based on a static model.

marginal income tax formulas became simple and easily interpretable. Not surprisingly, when relaxing these assumptions, things become more complex. The positionality effect is now given by:

$$\frac{\partial \mathcal{L}}{\partial \bar{c}_t} = -\frac{N_t\gamma_t\bar{\alpha}_t}{1-\bar{\alpha}_t} - \frac{N_{t+1}\gamma_{t+1}\bar{\beta}_{t+1}}{1-\bar{\alpha}_t} - \sum_{k=1}^{\infty}\left[\frac{N_{t+k}\gamma_{t+k}\bar{\alpha}_{t+k}}{1-\bar{\alpha}_{t+k}} + \frac{N_{t+k+1}\gamma_{t+k+1}\bar{\beta}_{t+k+1}}{1-\bar{\alpha}_{t+k}}\right]D_{t,k} < 0, \tag{18}$$

where

$$D_{t,k} = \prod_{j=1}^{k}\bar{\beta}_{t+j}/(1-\bar{\alpha}_{t+j-1}) > 0.$$

Eq. (18) comprises three distinct negative effects. The first term on the right-hand side, $-N_t\gamma_t\bar{\alpha}_t/(1-\bar{\alpha}_t) < 0$, measures the direct welfare loss in period t of an increase in $\bar{c}_t$; the intuition is that an increase in $\bar{c}_t$, ceteris paribus, leads to lower utility for all consumers via the argument $c_t^i - \bar{c}_t$ in the function $v_t^i(\cdot)$ in Eq. (1c). This effect depends on the average degree of *current* positionality. The analogous second term, $-N_{t+1}\gamma_{t+1}\bar{\beta}_{t+1}/(1-\bar{\alpha}_t) < 0$, is interpretable as the direct welfare loss in period $t+1$ of an increase in $\bar{c}_t$, and the underlying mechanism here is that $\bar{c}_t$ affects individual utility negatively via the argument $x_{t+1}^i - \bar{c}_t$ in the function $v_t^i(\cdot)$. This effect captures the intertemporal consumption externality and depends on the average degree of *intertemporal* positionality.

Finally, the third term on the right-hand side of Eq. (18) reflects an intertemporal chain reaction. The intuition is that the intertemporal aspect of the consumption comparisons, i.e., that other people's past consumption affects utility, means that the welfare effects of changes in the reference consumption are not time-separable (as they would be without intertemporal comparisons). This is so because a change in the reference consumption today means behavioral adjustments in the future, which in turn influence the reference consumption levels relevant for future generations. As such, when we relax assumptions (i)–(iii) in the previous subsection, the welfare costs of intertemporal consumption comparisons can no longer be summarized by a single variable. In the absence of relative comparisons over time, i.e., if $\bar{\beta}_t = 0$ for all t, the right-hand side of Eq. (18) collapses to $-N_t\gamma_t\bar{\alpha}_t/(1-\bar{\alpha}_t)$, which takes the same form as Eq. (15).

Before presenting the optimal marginal income tax rates, define the marginal rate of substitution between leisure and private consumption in period t, $\mathrm{MRS}_{z,c}^{i,t} = u_{t,z}^i/u_{t,c}^i$, and the marginal rate of substitution between consumption in periods t and $t+1$, $\mathrm{MRS}_{c,x}^{i,t} = u_{t,c}^i/u_{t,x}^i$, for ability-type i of generation t, and let $1+\eta_{t+1} = N_{t+1}/N_t$ denote the population growth factor. The optimal tax structure can then be characterized as follows:

Proposition 2. *The first-best marginal income tax rates faced by ability-type i can be written as (for i = 1, 2)*

$$T_t'(w_t^i l_t^i) = -\frac{\mathrm{MRS}_{z,c}^{i,t}}{\gamma_t w_t^i N_t}\frac{\partial \mathcal{L}}{\partial \bar{c}_t} > 0, \tag{19}$$

$$\Phi_{t+1}'(s_t^i r_{t+1}) = \frac{1}{\gamma_{t+1} r_{t+1} N_t}\left[\frac{\partial \mathcal{L}}{\partial \bar{c}_t} - \frac{\mathrm{MRS}_{c,x}^{i,t}}{1+\eta_{t+1}}\frac{\partial \mathcal{L}}{\partial \bar{c}_{t+1}}\right] \tag{20}$$

Proof. See the Appendix. □

Again, since the government may reach its distributional objectives by using lump-sum taxes/subsidies, there is no distributional reason for using distortionary taxation, and the non-zero marginal labor income tax rates are therefore solely due to the negative externalities that positional concerns give rise to. Therefore, the marginal labor income tax rates are positive for all types. Yet, since the positionality effects become very complex in the case where the average positionality degrees vary over time, it is not possible to express the optimal marginal labor income tax rates in a simple way in terms of the degrees of positionality.

To interpret Eq. (20), note that the marginal capital income tax rates reflect a desired tradeoff for society between present and future consumption. As a consequence, the right-hand side of Eq. (20) is decomposable into two parts (in Section 3.2, these two parts canceled out due to assumptions (i)–(iii). The basic intuition is that each individual generates positional externalities both when young and when old. Therefore, whether positional concerns lead to a positive or negative marginal capital income tax rate in period $t+1$ depends on the difference between the positionality effect in period t and the discounted positionality effect in period $t+1$. Again, this result holds regardless of whether the preferences for relative consumption are governed by a keeping up or catching up with the Joneses motive, or by a mix of them.

3.4. Briefly on age-specific consumption comparisons

The analysis carried out so far assumes that each individual compares his/her consumption with economy-wide averages, i.e., the relevant measures of reference consumption are given by the average consumption in the present period (for the keeping up with the Joneses comparison) and average consumption in the previous period (for the catching up with the Joneses comparison), respectively. Although mean-value comparisons are very common in earlier comparable studies on tax and other policy implications of relative consumption concerns,[15] empirical evidence suggests that people may use more narrow social reference groups based on similarities (e.g., Runciman, 1966; McBride, 2001). Following McBride (2001), we briefly examine how the results may change if each consumer compares his/her own consumption with that of other persons of the same age (instead of with economy-wide averages).[16]

Define the average consumption among the young and old alive in period t as follows:

$$\bar{c}_t^{\text{young}} = \frac{\sum_i n_t^i c_t^i}{\sum_i n_t^i} \quad \text{and} \quad \bar{x}_t^{\text{old}} = \frac{\sum_i n_{t-1}^i x_t^i}{\sum_i n_{t-1}^i}.$$

Eq. (1c) is then replaced with

$$U_t^i = v_t^i(c_t^i, z_t^i, x_{t+1}^i, c_t^i - \bar{c}_t^{\text{young}}, x_{t+1}^i - \bar{x}_{t+1}^{\text{old}}, c_t^i - \bar{c}_{t-1}^{\text{young}}, x_{t+1}^i - \bar{x}_t^{\text{old}}) \quad \text{for} \quad i = 1, 2. \tag{1c$'$}$$

Based on Eq. (1c′), the degrees of current positionality characterizing ability-type i of generation t (who is young in period t and old in period $t+1$) are given by

$$\alpha_t^{i,\text{young}} = \frac{v^i_{t,c_t^i-\bar{c}_t^{\text{young}}}}{v^i_{t,c_t^i-\bar{c}_t^{\text{young}}} + v^i_{t,c_t^i-\bar{c}_{t-1}^{\text{young}}} + v^i_{t,c}},$$

$$\alpha_{t+1}^{i,\text{old}} = \frac{v^i_{t,x_{t+1}^i-\bar{x}_{t+1}^{\text{old}}}}{v^i_{t,x_{t+1}^i-\bar{x}_{t+1}^{\text{old}}} + v^i_{t,x_{t+1}^i-\bar{x}_t^{\text{old}}} + v^i_{t,x}},$$

when young and old, respectively, while the corresponding degrees of intertemporal positionality become

$$\beta_t^{i,\text{young}} = \frac{v^i_{t,c_t^i-\bar{c}_{t-1}^{\text{young}}}}{v^i_{t,c_t^i-\bar{c}_t^{\text{young}}} + v^i_{t,c_t^i-\bar{c}_{t-1}^{\text{young}}} + v^i_{t,c}},$$

$$\beta_{t+1}^{i,\text{old}} = \frac{v^i_{t,x_{t+1}^i-\bar{x}_t^{\text{old}}}}{v^i_{t,x_{t+1}^i-\bar{x}_{t+1}^{\text{old}}} + v^i_{t,x_{t+1}^i-\bar{x}_t^{\text{old}}} + v^i_{t,x}}.$$

We can then define average degrees of current and intertemporal positionality such that

$$\bar{\alpha}_t^{\text{young}} = \frac{\sum_i n_t^i \alpha_t^{i,\text{young}}}{\sum_i n_t^i},\ \bar{\alpha}_t^{\text{old}} = \frac{\sum_i n_{t-1}^i \alpha_t^{i,\text{old}}}{\sum_i n_{t-1}^i},\ \bar{\beta}_t^{\text{young}} = \frac{\sum_i n_t^i \beta_t^{i,\text{young}}}{\sum_i n_t^i},\ \text{and}\ \bar{\beta}_t^{\text{old}} = \frac{\sum_i n_{t-1}^i \beta_t^{i,\text{old}}}{\sum_i n_{t-1}^i}.$$

The general case with time-varying average degrees of positionality provides little insight beyond those presented in Section 3.3. Let us, therefore, return to the special case examined in Section 3.2, which was based on (the additional) assumptions (i)–(iii), and also modify assumption (ii) such that the average positionality degrees when comparing with others of the same age are constant over time, i.e., $\bar{\alpha}_t^{\text{young}} = \bar{\alpha}^{\text{young}}$, $\bar{\alpha}_t^{\text{old}} = \bar{\alpha}^{\text{old}}$, $\bar{\beta}_t^{\text{young}} = \bar{\beta}^{\text{young}}$ and $\bar{\beta}_t^{\text{old}} = \bar{\beta}^{\text{old}}$ for all t. Finally, and by analogy to the analysis carried out above, define average degrees of time-inclusive positionality based on the positionality concept examined here

$$\bar{\rho}^{\text{young}} = \bar{\alpha}^{\text{young}} + \frac{\bar{\beta}^{\text{young}}}{1+r} \quad \text{and} \quad \bar{\rho}^{\text{old}} = \bar{\alpha}^{\text{old}} + \frac{\bar{\beta}^{\text{old}}}{1+r}$$

It is now straightforward to show that the optimal tax policy summarized by Eqs. (16) and (17) in Proposition 1 is modified as follows:

$$T_t'(w_t^i l_t^i) = \bar{\rho}^{\text{young}} > 0, \tag{16$'$}$$

$$\Phi_{t+1}'(s_t^i r_{t+1}) = \frac{1+r}{r}(1-\bar{\rho}^{\text{old}}) \left(\frac{\bar{\rho}^{\text{old}}}{1-\bar{\rho}^{\text{old}}} - \frac{\bar{\rho}^{\text{young}}}{1-\bar{\rho}^{\text{young}}} \right). \tag{17$'$}$$

In the Appendix, we briefly describe how Eqs. (16′) and (17′) are derived. The intuition behind Eqs. (16′) and (17′) follows by observing that age-specific comparisons allow for a distinction between average degrees of time-inclusive positionality among the young and old. Therefore, since only young individuals supply labor in this model, the marginal labor income

[15] Aronsson and Johansson-Stenman (2010) constitutes a recent exception, where the optimal tax policy implications of upward and within-generation comparisons (in addition to mean value comparison) are addressed. Yet, that study only allows the consumers to compare their current consumption with other people's current consumption (i.e., keeping up with the Joneses), irrespective of whether the measures of reference consumption are defined to accommodate mean-value, upward-, or within-generation comparisons.

[16] We are grateful to one of the referees for suggesting this extension. Alvarez-Cuadrado and Long (2012) use an OLG-model and heterogeneous agents with relative comparisons of this kind to derive several interesting results, in particular regarding distributional effects due to social comparisons.

tax rate equals the average degree of time-inclusive positionality among the young, which reflects the externality that each young consumer imposes on his/her referent others.

The marginal capital income tax rate typically differs from zero, and its sign depends on whether the old consumers are, on average, more or less positional than their young counterparts. This result accords well with the general characterization in Section 3.3, and is here expressed in terms of the difference in the average degree of time-inclusive positionality between the old and the young age groups (while there was no such difference in Section 3.2). Clearly, when these measures of time-inclusive positionality are the same for the young and the old, such that $\bar{\rho}^{\text{old}} = \bar{\rho}^{\text{young}}$, it follows that the optimal marginal capital income tax is zero.

Finally, note that the optimal marginal income tax rates are also in this case fully characterized by average degrees of *time-inclusive* positionality. Once again, therefore, relative consumption concerns based on comparisons with other people's current and past consumption affect the marginal income tax structure in exactly the same way.

4. Optimal second-best taxation

In reality, governments are not likely to be able to redistribute on a lump-sum basis, since individual ability is private information. This may, in turn, have important implications for how the tax system should be used in response to positional concerns. In this section, we will generalize the findings obtained in Sections 3.2 and 3.3 to the case where asymmetric information prevents redistribution through lump-sum taxes.

4.1. The social decision problem

As before, there are two types of individuals, but we now explicitly state that individuals of type 1 are less productive than those of type 2 (measured in terms of the before-tax wage rate), and refer to type 1 as the "low-ability type" and type 2 as the "high-ability type." Following the convention in earlier literature on optimal nonlinear taxation, we assume that the government is able to observe income, that ability (and, consequently, work hours) is private information, and that the government wants to redistribute from the high-ability to the low-ability type. To prevent the high-ability type from becoming a mimicker, we impose the following self-selection constraint:

$$U_t^2 = u_t^2(c_t^2, z_t^2, x_{t+1}^2, \bar{c}_{t-1}, \bar{c}_t, \bar{c}_{t+1}) \geq u_t^2(c_t^1, 1 - \phi l_t^1, x_{t+1}^1, \bar{c}_{t-1}, \bar{c}_t, \bar{c}_{t+1}) = \hat{U}_t^2, \tag{21}$$

which means that any high-ability type individual weakly prefers the allocation intended for his/her type over the allocation intended for the low-ability type. The variable $\phi = w_t^1/w_t^2 = \theta^1/\theta^2 < 1$ denotes the wage ratio, which is a constant by the assumptions about the technology made earlier. The left-hand side of the weak inequality in (21) measures the utility faced by the high-ability type if revealing his/her true ability, while the right-hand side represents the utility of the high-ability mimicker, i.e., a high-ability type who chooses the same income-consumption points as the low-ability type. Although the mimicker enjoys the same consumption as the low-ability type in each period, he/she reaches this consumption with less work-effort implying more leisure ($1 - \phi l_t^1 > z_t^1$ denotes the mimicker's leisure).[17]

The social welfare function, production technology, and resource constraint are the same as in Section 3. Therefore, the Lagrangian can be written as

$$\mathscr{L} = W + \sum_t \lambda_t [U_t^2 - \hat{U}_t^2] + \sum_t \gamma_t \left[F(L_t, K_t; t) + K_t - \sum_{i=1}^{2} [n_t^i c_t^i + n_{t-1}^i x_t^i] - K_{t+1} \right], \tag{22}$$

where λ_t is the Lagrange multiplier associated with the self-selection constraint faced by generation t. As above, the government treats the measures of reference consumption as endogenous.

Following the bulk of earlier literature on optimal nonlinear income taxation in dynamic economies, we assume that the government commits to its future tax policy.[18] Let $\hat{u}_t^2 = u_t^2(c_t^1, 1 - \phi l_t^1, x_{t+1}^1, \bar{c}_{t-1}, \bar{c}_t, \bar{c}_{t+1})$ denote the utility of the mimicker based on the utility formulation in Eq. (1d). The direct decision variables relevant for generation t are $l_t^1, c_t^1, x_{t+1}^1, l_t^2, c_t^2, x_{t+1}^2$, and K_{t+1}, and the social first-order conditions are given by

$$-\frac{\partial W}{\partial (n_t^1 U_t^1)} n_t^1 u_{t,z}^1 + \phi \lambda_t \hat{u}_{t,z}^2 + \gamma_t n_t^1 w_t^1 = 0, \tag{23a}$$

$$\frac{\partial W}{\partial (n_t^1 U_t^1)} n_t^1 u_{t,c}^1 - \lambda_t \hat{u}_{t,c}^2 - \gamma_t n_t^1 + \frac{n_t^1}{N_t} \frac{\partial \mathscr{L}}{\partial \bar{c}_t} = 0, \tag{23b}$$

[17] Given the set of available policy instruments assumed here, it is possible for the government to control the present and future consumption as well as the hours of work of each ability type. As a consequence, in order to be a mimicker, the high-ability type must mimic the labor income and capital income of the low-ability type, and thus consume the same amount as the low-ability type in both periods.

[18] See Brett (1997), Pirttilä and Tuomala (2001), Aronsson et al. (2009), and Aronsson and Johansson-Stenman (2010). For recent studies of time-consistent optimal nonlinear income taxation without commitment, see Brett and Weymark (2008) and Guo and Krause (2011). See also Acemoglu et al. (2011) for a political economy model with taxation of capital and labor when self-interested politicians cannot commit to future policies.

$$\frac{\partial W}{\partial (n_t^1 U_t^1)} n_t^1 u_{t,x}^1 - \lambda_t \hat{u}_{t,x}^2 - \gamma_{t+1} n_t^1 + \frac{n_t^1}{N_{t+1}} \frac{\partial \mathcal{L}}{\partial \bar{c}_{t+1}} = 0, \tag{23c}$$

$$-\left[\frac{\partial W}{\partial (n_t^2 U_t^2)} n_t^2 + \lambda_t\right] u_{t,z}^2 + \gamma_t n_t^2 w_t^2 = 0, \tag{23d}$$

$$\left[\frac{\partial W}{\partial (n_t^2 U_t^2)} n_t^2 + \lambda_t\right] u_{t,c}^2 - \gamma_t n_t^2 + \frac{n_t^2}{N_t} \frac{\partial \mathcal{L}}{\partial \bar{c}_t} = 0, \tag{23e}$$

$$\left[\frac{\partial W}{\partial (n_t^2 U_t^2)} n_t^2 + \lambda_t\right] u_{t,x}^2 - \gamma_{t+1} n_t^2 + \frac{n_t^2}{N_{t+1}} \frac{\partial \mathcal{L}}{\partial \bar{c}_{t+1}} = 0, \tag{23f}$$

$$\gamma_{t+1}[1 + r_{t+1}] - \gamma_t = 0, \tag{23g}$$

where (as before) $u_{t,c}^i = \partial u_t^i / \partial c_t^i$, $u_{t,z}^i = \partial u_t^i / \partial z_t^i$, and $u_{t,x}^i = \partial u_t^i / \partial x_{t+1}^i$, and where $\partial \mathcal{L} / \partial \bar{c}_t$ denotes the positionality effect in period t, i.e., the partial welfare effect of an increase in $\bar{c}_t$. For further use, we define the following measures of differences in the degree of current and intertermporal positionality, respectively, between the mimicker and the low-ability type in period t:

$$\alpha_t^d = \frac{\lambda_{t-1} \hat{u}_{t-1,x}^2}{\gamma_t N_t} [\hat{\alpha}_t^{2,x} - \alpha_t^{1,x}] + \frac{\lambda_t \hat{u}_{t,c}^2}{\gamma_t N_t} [\hat{\alpha}_t^{2,c} - \alpha_t^{1,c}], \tag{24a}$$

$$\beta_t^d = \frac{\lambda_{t-1} \hat{u}_{t-t,x}^2}{\gamma_t N_t} [\hat{\beta}_t^{2,x} - \beta_t^{1,x}] + \frac{\lambda_t \hat{u}_{t,c}^2}{\gamma_t N_t} [\hat{\beta}_t^{2,c} - \beta_t^{1,c}], \tag{24b}$$

where the symbol "$\hat{}$" denotes "mimicker" (as before), while the super-script "d" stands for "difference." Note that α_t^d and β_t^d reflect positionality differences between the young mimicker and the young low-ability type, and between the old mimicker and the old low-ability type, respectively. Note also that the variables α_t^d and β_t^d are related to the self-selection constraint, since each component in Eqs. (24a) and (24b), respectively, is proportional to the Lagrange multiplier of the self-selection constraint (either the constraint facing generation $t-1$ or the constraint facing generation t). As such, α_t^d and β_t^d are fundamentally related to the second-best framework with asymmetric information set out here (they would vanish in a first-best economy, where $\lambda_t = 0$ for all t).

4.2. *Optimal second-best taxation and time-invariant positionality degrees*

Consider first the second-best analog to the stationary regime addressed in Section 3.2, where the degrees of positionality are constant over time. This special case facilitates comparison with earlier literature on optimal second-best taxation under relative consumption as well as provides straightforward (second-best) analogs to Eqs. (16) and (17). Therefore, and by analogy to Section 3.2, let us again make assumptions (i), (ii), and (iii), such that, for all t, $N_t = N$, $\bar{\alpha}_t = \bar{\alpha}$, $\bar{\beta}_t = \bar{\beta}$, and $r_t = r$, and in addition make the following assumption:

iv. The indicators of positionality differences between the mimicker and the low-ability type are constant over time in the sense that $\alpha_t^d = \alpha^d$ and $\beta_t^d = \beta^d$ for all t.

As before, this case is either interpretable in terms of a steady state or may follow as a consequence of adding additional assumptions about the preferences and technology. By analogy to the average degree of time-inclusive consumption positionality defined in Section 3.2, i.e., $\bar{\rho} = \bar{\alpha} + \bar{\beta}/(1 + r)$, we define *the difference in the time-inclusive degree of consumption positionality between the mimicker and the low-ability type* (also in present value terms) as

$$\rho^d = \alpha^d + \frac{\beta^d}{1 + r}.$$

We can then express the positionality effect, which characterizes the partial welfare effect of an increase in $\bar{c}_t$, as

$$\frac{\partial \mathcal{L}}{\partial \bar{c}_t} = -N \gamma_t \frac{\bar{\rho} - \rho^d}{1 - \bar{\rho}}. \tag{25}$$

With Eq. (25) at our disposal, we can relate the marginal income tax rates to the average degree of time-inclusive consumption positionality and to the difference in this measure of positionality between the mimicker and the low-ability type. Starting with the marginal labor income tax rates, we combine Eqs. (6), (23a), (23b) and (25) to derive the marginal labor income tax rate faced by the low-ability type, and Eqs. (6), (23d), (23e) and (25) to derive the marginal labor income tax rate faced by the high-ability type. To simplify the notations, we use τ_t^1 and τ_t^2 to denote the optimal marginal labor income tax rates in the original Stiglitz (1982) model, i.e., the expressions that would follow in the absence of relative consumption concerns:

$$\tau_t^1 = \frac{\lambda_t^*}{w_t^1 n_t^1}[\mathrm{MRS}_{z,c}^{1,t} - \phi \hat{\mathrm{MRS}}_{z,c}^{2,t}] \quad \text{and} \quad \tau_t^2 = 0,$$

$$\text{where} \quad \lambda_t^* = \lambda_t \hat{u}_{t,c}^2/\gamma_t \quad \text{and} \quad \hat{\mathrm{MRS}}_{z,c}^{2,t} = \hat{u}_{t,z}^2/\hat{u}_{t,c}^2.$$

We can then characterize the marginal labor income tax rates as follows:

Proposition 3. *Under assumptions (i)–(iv), the second-best marginal labor income tax rate can, for each ability type, be written in the following additive form (for i = 1, 2):*

$$T_t'(w_t^i l_t^i) = \tau_t^i + [1-\tau_t^i]\bar{\rho} - [1-\tau_t^i][1-\bar{\rho}]\frac{\rho^d}{1-\rho^d}. \tag{26}$$

Proof. See the Appendix. □

The first term on the right-hand side is the expression for marginal labor income taxation that would follow in the standard optimal income tax model and is well understood from earlier research (Stiglitz, 1982).[19] The second term measures the marginal external cost of consumption as reflected by the average degree of time-inclusive consumption positionality, although its contribution to the marginal labor income tax rates is modified compared with the first-best formula given by Eq. (16). The intuition is that the fraction of marginal income that is already taxed away does not give rise to positional externalities. Therefore, if $\tau_t^1 > 0$, this "second-best modification" tends to reduce the externality-correcting component in the formula for the low-ability type.

The third term on the right-hand side of Eq. (26) reflects the self-selection constraint. Suppose first that $\rho^d > 0$, meaning that the mimicker has a higher degree of time-inclusive positionality than the low-ability type. In this case, increased reference consumption gives rise to a larger utility loss for the mimicker than for the low-ability type. The government may then relax the self-selection constraint through a tax policy that leads to increased reference consumption, i.e., through lower marginal labor income tax rates.[20] Correspondingly, if $\rho^d < 0$, increased reference consumption tightens the self-selection constraint such that the third term on the right-hand side contributes to increase the marginal labor income tax rate.

Eq. (26) also reveals that the average degree of current positionality, $\bar{\alpha}$, and the present value of the average degree of intertemporal positionality, $\bar{\beta}/(1+r)$, affect each marginal labor income tax rate in exactly the same way. The same applies to the measures of positionality differences, i.e., α^d and $\beta^d/(1+r)$. Special cases of Eq. (26) are derived by Aronsson and Johansson-Stenman (2010), where the preferences for relative consumption are based solely on the keeping up with the Joneses motive ($\bar{\rho} = \bar{\alpha}$ and $\rho^d = \alpha^d$), and by Ljungqvist and Uhlig (2000), where the consumption comparisons are based on the catching up with the Joneses motive and there is no asymmetric information ($\bar{\rho} = \bar{\beta}/(1+r)$ and $\rho^d = 0$).

Turning to the capital income tax structure, we can derive the marginal capital income tax rate implemented for the low-ability type by combining Eqs. (7), (23b), (23c), (23g) and (25), and the marginal capital income tax rate implemented for the high-ability type by combining Eqs. (7), (23e), (23f), (23g) and (25). To simplify the presentation of the results, let δ_t^i denote the expression for marginal capital income taxation of ability type i that would follow in the absence of relative consumption concerns, where

$$\delta_t^1 = \frac{\lambda_t \hat{u}_{t,x}^2}{\gamma_{t+1} n_t^1 r_{t+1}}[\mathrm{MRS}_{c,x}^{1,t} - \hat{\mathrm{MRS}}_{c,x}^{2,t}] \quad \text{and} \quad \delta_t^2 = 0.$$

We can then derive the following result:

Proposition 4. *Under assumptions (i)–(iv), the second-best marginal capital income tax rate can, for each ability type, be written as (for i = 1, 2):*

$$\Phi_{t+1}'(s_t^i r_{t+1}) = \frac{1-\bar{\rho}}{1-\rho^d}\delta_t^i. \tag{27}$$

Proof. See the Appendix. □

The variable δ_t^i on the right-hand side of Eq. (27), which would also be present in a standard two-type model without consumption externalities, is due to the self-selection constraint and is well understood and explained in earlier research (e.g.,

[19] For the low-ability type, this component is typically positive (at least if the consumers share a common utility function), while it is zero for the high-ability type due to that the relative wage rate is constant. The intuition is that the government may relax the self-selection constraint by taxing low-ability labor, since the low-ability type attaches a higher marginal value to leisure than the mimicker, whereas no such option exists for the high-ability type.

[20] Thus, if ρ^d is positive and sufficiently large, relative consumption concerns may actually contribute to reduce the marginal labor income tax rates (although this scenario seems unlikely).

Brett, 1997).[21] As in Section 3.2, there is no direct effect of relative consumption concerns in Eq. (27) since the positionality effects in periods t and $t+1$ largely cancel out when the degrees of positionality are constant over time. Yet, there is a second-best adjustment of the marginal capital income tax rate faced by the low-ability type due to positional concerns, which was not present in Section 3.2. Consider the case where $MRS_{c,x}^{1,t} > M\hat{R}S_{c,x}^{2,t}$, in which the mimicker values an additional dollar today in terms of consumption tomorrow less than does the low-ability type, implying that $\delta_t^1 > 0$ and $\Phi'_{t+1}(s_t^1 r_{t+1}) > 0$. The term $1-\bar{\rho}$ then serves to reduce the effect that δ_t^1 would otherwise have on the marginal capital income tax rate, since capital income taxation leads to an increase in $\bar{c}_t$ and, therefore, in the externality that $\bar{c}_t$ gives rise to. The effect through $1-\rho^d$ in the denominator is because increased reference consumption may either relax ($\rho^d > 0$) or tighten ($\rho^d < 0$) the self-selection constraint, depending on whether a mimicking high-ability type is more or less positional than the low-ability type in terms of time-inclusive positionality. Analogous results and interpretations hold for the case where $MRS_{c,x}^{1,t} < M\hat{R}S_{c,x}^{2,t}$.

It is easy to see that all qualitative results in this subsection hold in the special case without consumption comparisons over time, i.e., where $\bar{\rho} = \bar{\alpha}$ and $\rho^d = \alpha^d$, which is the case addressed by Aronsson and Johansson-Stenman (2010). Similarly, all qualitative results hold in the other extreme situation where $\bar{\rho} = \bar{\beta}/(1+r)$ and $\rho^d = \beta^d/(1+r)$, in which there are no comparisons with other people's current consumption.

4.2.1. When the optimal capital income taxes vanish

An important issue in the literature on capital income taxation is to determine conditions for when it is optimal not to use such taxes. In this brief subsection, we add some further assumptions that together are sufficient for an optimal tax structure without capital income taxes. In doing so, we will relate to a classical result by Ordover and Phelps (1979), regarding when it is optimal not to use capital taxation at all on the margin. From Proposition 4, it is straightforward to derive such conditions also in our model. The following result is an immediate consequence of Proposition 4:

Corollary 1. *Under assumptions (i)–(iv), and if leisure is weakly separable from private consumption in the sense that* $U_t^i = q_t^i(f_t(c_t^i, x_{t+1}^i, c_t^i - \bar{c}_t, x_{t+1}^i - \bar{c}_{t+1}, c_t^i - \bar{c}_{t-1}, x_{t+1}^i - \bar{c}_t), z_t^i)$ *describes the utility function, both optimal marginal capital income tax rates are zero.*

This result follows from acknowledging that the mimicker and the low-ability type differ only with respect to preferences and use of leisure. Given the assumption that leisure is weakly separable from the other goods in the utility function, and that the consumers share a common sub-utility function $f_t(\cdot)$, it follows that $MRS_{c,x}^{1,t} = M\hat{R}S_{c,x}^{2,t}$, implying that both marginal capital income tax rates are equal to zero according to Eq. (27). The quite remarkable consequence of Corollary 1 is that the separability result by Ordover and Phelps (1979) continues to apply under a fairly general formulation of the relative consumption concerns, which allow for both the keeping up and catching up with the Joneses mechanisms.[22]

To see the intuition behind this result, note first that a time-dependent lump-sum tax element would in a first-best world handle intertemporal allocation issues efficiently, such that capital income taxes are not needed (as we saw in Section 3.2). In a second-best world, the additional question is whether (and if so how) the government may relax the self-selection constraint through capital income taxation, i.e., by exploiting that the low-ability type might value the tradeoff between present and future consumption in a different way than the mimicker. With a common utility function in the sense of Corollary 1, the only reason why this tradeoff may differ between the mimicker and the low-ability type is that the mimicker is more productive and, therefore, enjoys more leisure than the low-ability type. However, since leisure separability means that the labor supply does not directly affect the marginal rate of substitution between present and future consumption, it follows that the mimicker and low-ability type will face the same intertemporal consumption tradeoff, in which case it is no longer possible to relax the self-selection constraint by distorting the savings behavior.[23] This is so whether the model incorporates positional externalities or not.

4.3. Optimal second-best taxation with time-varying degrees of positionality

Let us finally – as we also did in Section 3.3 – relax the stationarity assumptions (assumptions (i)–(iv)). While the policy rules will be more complex, we are still able to show some important findings for how the relative consumption concerns

[21] The government may relax the self-selection constraint by exploiting that the mimicker and the low-ability type differ with respect to the marginal rate of substitution between present and future consumption (in which case $\delta_t^1 \neq 0$), whereas no such option exists with regard to the marginal capital income tax rate of the high-ability type ($\delta_t^2 = 0$).

[22] The Ordover and Phelps (1979) result can also be seen as an intertemporal analog to the Atkinson and Stiglitz (1976) result, saying that under optimal nonlinear income taxation, differentiated consumption taxes are not needed under leisure separability.

[23] The result that optimal capital income tax rates are zero under leisure separability (while typically non-zero otherwise) in an OLG model has also been derived under optimal age-dependent (yet linear) labor income taxation; see Erosa and Gervais (2002) and Mathieu-Bolh (2006). Yet, none of these studies address relative consumption concerns.

affect the optimal marginal income tax rates. We show in the Appendix that a general characterization of the optimal marginal income tax rates in the second best is given as (for $i = 1, 2$)

$$T_t'(w_t^i l_t^i) = \tau_t^i - \frac{\mathrm{MRS}_{z,c}^{i,t}}{\gamma_t w_t^i N_t} \frac{\partial \mathcal{L}}{\partial \bar{c}_t}, \tag{28}$$

$$\Phi_{t+1}'(s_t^i r_{t+1}) = \delta_t^i + \frac{1}{\gamma_{t+1} r_{t+1} N_t} \left[\frac{\partial \mathcal{L}}{\partial \bar{c}_t} - \frac{\mathrm{MRS}_{c,x}^{i,t}}{1+\eta_{t+1}} \frac{\partial \mathcal{L}}{\partial \bar{c}_{t+1}} \right], \tag{29}$$

where we have used the short notations τ_t^i and δ_t^i defined above. Eqs. (28) and (29) are straightforward generalizations of the corresponding first-best formulas given by Eqs. (19) and (20).

To be able to say more about the relationship between the relative consumption concerns and the marginal income tax rates, we must explore the positionality effect in more detail. By using the expressions for the differences in the degree of current (α_t^d) and intertemporal (β_t^d) positionality between the mimicker and the low-ability type in period t, as given by Eqs. (24a) and (24b), together with the short notations

$$A_{t+k} = \frac{N_{t+k}\gamma_{t+k}[\alpha_{t+k}^d - \bar{\alpha}_{t+k}]}{1-\bar{\alpha}_{t+k}},$$

$$B_{t+k} = \frac{N_{t+k+1}\gamma_{t+k+1}[\beta_{t+k+1}^d - \bar{\beta}_{t+k+1}]}{1-\bar{\alpha}_{t+k}},$$

we obtain the following second-best analog to Eq. (18):

$$\frac{\partial \mathcal{L}}{\partial \bar{c}_t} = A_t + B_t + \sum_{k=1}^{\infty} [A_{t+k} + B_{t+k}] \prod_{j=1}^{k} \frac{\bar{\beta}_{t+j}}{1-\bar{\alpha}_{t+j-1}}. \tag{30}$$

Eq. (30) extends Eq. (25) to the case where neither the degrees of positionality, nor the population and interest rate, are (necessarily) constant over time. We can see that Eq. (30) has the same structure as Eq. (18) above, with the exception that the terms $-\bar{\alpha}_{t+k}$ and $-\bar{\beta}_{t+k+1}$ in Eq. (18) are here replaced with $\alpha_{t+k}^d - \bar{\alpha}_{t+k}$ and $\beta_{t+k+1}^d - \bar{\beta}_{t+k+1}$, respectively. The only important difference between Eqs. (25) and (30) is that we can no longer make use of the (analytically convenient) time-inclusive degrees of positionality in the same way as before. As a consequence, although comparisons with other people's past consumption (the catching up motive) give rise to the same qualitative policy implications as those associated with comparisons with other people's current consumption (the keeping up motive), the welfare consequences of intertemporal consumption comparisons are, in general, much more complex than those associated with atemporal consumption comparisons. This is seen by recognizing that in the special case without intertemporal consumption comparisons, i.e., when $\bar{\beta}_t = \beta_t^d = 0$ for all t, Eq. (30) reduces to

$$\frac{\partial \mathcal{L}}{\partial \bar{c}_t} = -N_t\gamma_t \frac{\bar{\alpha}_t - \alpha_t^d}{1-\bar{\alpha}_t},$$

for all t, which takes exactly the same form as Eq. (25) above.

From Eq. (30), we can derive the following result regarding the conditions for when the sign of the positionality effect is unambiguously negative:

Lemma 1. *If, from period t and onwards, the low-ability type is at least as positional as the mimicker on average, or if the positionality differences are sufficiently small, in any of the following senses:*

(i) $\frac{N_t\gamma_t\alpha_t^d + N_{t+1}\gamma_{t+1}\beta_{t+1}^d}{1-\bar{\alpha}_t} + \sum_{k=1}^{\infty} \frac{N_{t+k}\gamma_{t+k}\alpha_{t+k}^d + N_{t+k+1}\gamma_{t+k+1}\beta_{t+k+1}^d}{1-\bar{\alpha}_{t+k}} \prod_{j=1}^{k} \frac{\bar{\beta}_{t+j}}{1-\bar{\alpha}_{t+j-1}} \leq 0,$

(ii) $\alpha_{t+k}^d < \bar{\alpha}_{t+k}$ and $\beta_{t+k+1}^d < \bar{\beta}_{t+k+1}$ $\forall k \geq 0$,

(iii) $\alpha_{t+k}^d \leq 0$ and $\beta_{t+k+1}^d \leq 0$ $\forall k \geq 0$,

then increased reference consumption in period t reduces the welfare.

Given that the individual degrees of positionality (both in the current and intertemporal dimensions) are always between zero and one, (i) gives a sufficient condition for when increased reference consumption in period t leads to lower welfare. Yet, condition (i) is not necessary, since the terms in Eq. (30) that solely reflect the average degrees of positionality (i.e., the pure externality terms) contribute to lower welfare as well. Condition (*ii*) is not necessary either, since $\partial \mathcal{L}/\partial \bar{c}_t$ can clearly be negative even if (*ii*) does not hold for some k. Note finally that condition (iii), which we refer to because of its straightforward interpretation, is actually redundant since it implies condition (ii).

By combining Lemma 1 with Eq. (28), we obtain the following result:

Proposition 5. *If any of the conditions in Lemma 1 hold, so that increased reference consumption leads to lower welfare, ceteris paribus, then the positionality effect in period t contributes to increase the marginal labor income tax rates for both ability types in period t*

The interpretation of Proposition 5 is straightforward. If the low-ability type is at least as positional as the mimicker on average, or if loosely speaking the positionality differences are sufficiently small, and given that the individual degrees of positionality are always between zero and one, then we obtain from Eq. (30) that $\partial\mathscr{L}/\partial\bar{c}_t < 0$. As such, Proposition 5 also extends the result of Oswald (1983) – that "marginal tax rates are higher than in the conventional model when the population is predominantly jealous" (page 82) – to a model with consumption comparisons over time.[24]

Similarly, by combining Lemma 1 with Eq. (29), we can derive the following result for how positional concerns contribute to the marginal capital income tax rates:

Proposition 6. *Suppose that any of the conditions in Lemma 1 hold, so that increased reference consumption leads to lower welfare, ceteris paribus Then, if the preferences become less (more) positional over time in the sense that*

$$\left|\frac{\partial\mathscr{L}}{\partial\bar{c}_t}\right| > (<)\frac{\mathrm{MRS}^{i,t}_{c,x}}{(1+\eta_{t+1})}\left|\frac{\partial\mathscr{L}}{\partial\bar{c}_{t+1}}\right|,$$

i.e., the positionality effect in period t dominates (is dominated by) the positionality effect in period t + 1, then the joint contribution of the positionality effects in periods t and t + 1 is to decrease (increase) the marginal capital income tax rate for ability-type i in period t + 1.

If Lemma 1 applies, such that the positionality effect is negative in all periods, the interpretation of the first-best policy rule for marginal capital income taxation in Eq. (20) carries over with some modifications to the second-best framework addressed here. Therefore, if the positionality effect in period *t* dominates the corresponding effect in period *t* + 1, there is an incentive for the government to discourage the consumption in period *t* relative to the consumption in period *t* + 1. The difference by comparison with the first-best policy analyzed in Section 3.3 is, of course, that the policy incentives analyzed here are due to both externality correction and redistribution effects of positional concerns through the self-selection constraint. This has been discussed at some length above. The analogous policy incentive to encourage the consumption in period *t* relative to the consumption in period *t* + 1 arises if the positionality effect in period *t* + 1 dominates. Again, these insights follow irrespective of whether the positional concerns are driven by the keeping up or catching up mechanism or a mix between them.

An interesting implication of the proposition is that it would be optimal with increasing marginal capital income taxation over time in an economy where the preferences become more positional over time (i.e., if we tend to attach a higher value to increased relative consumption than to increased absolute consumption as time passes). Such a pattern is actually broadly consistent with some empirical evidence: Clark et al. (2008) analyze the impact of relative income on happiness and conclude that the concern for relative income tends to increase as the average income in a country increases. Note also that we can interpret the component $\mathrm{MRS}^{i,t}_{c,x}/(1+\eta_{t+1})$ as the effective discount factor for ability-type *i*, which is used to discount the positionality effect in period *t* + 1 to period *t*.

5. Results based on a Cobb–Douglas utility function

In order to more clearly illustrate some implications of the relative consumption comparisons for optimal income taxation, let us assume that the population and interest rate are constant (as we did several times above), but in addition consider the following Cobb–Douglas utility function:

$$U^i_t = k^i (z^i_t)^{k^i_z} (c^i_{\mathrm{net},t})^{k_c} (x^i_{\mathrm{net},t+1})^{k_x}, \tag{31}$$

where $k^i, k^i_z, k_c, k_x > 0$ are constants and $k^i_z + k_c + k_x < 1$; $c^i_{\mathrm{net},t}$ and $x^i_{\mathrm{net},t+1}$ reflect what we may think of as consumption net of relative consumption concerns, when young and when old, for an individual of ability-type *i* born in period *t*, as defined below:

$$c^i_{net,t} = [1-a-b]c^i_t + a[c^i_t - \bar{c}_t] + b[c^i_t - \bar{c}_{t-1}] = c^i_t - a\bar{c}_t - b\bar{c}_{t-1}, \tag{32a}$$

$$x^i_{net,t+1} = [1-a'-b']x^i_{t+1} + a'[x^i_{t+1} - \bar{c}_{t+1}] + b'[x^i_{t+1} - \bar{c}_t] = x^i_{t+1} - a'\bar{c}_{t+1} - b'\bar{c}_t. \tag{32b}$$

By substituting Eqs. (32a) and (32b) into Eq. (31), we obtain:

$$U^i_t = k^i [z^i_t]^{k^i_z} [c^i_t - a\bar{c}_t - b\bar{c}_{t-1}]^{k_c} [x^i_{t+1} - a'\bar{c}_{t+1} - b'\bar{c}_t]^{k_x}. \tag{33}$$

[24] By "jealousy," Oswald meant that an increase in the reference consumption leads to decreased utility for the individual consumer (as compared to "altruism," which has the opposite effect). He analyzed nonlinear taxation of commodities instead of income, and the result referred to above is based on a utility function that is separable in the reference measure; an assumption that limits the influence of the incentive constraint on the marginal tax rates.

Although the utility functions are allowed to differ between the ability types, through the parameters k^i and k_z^i, the individual degrees of current and intertemporal consumption positionality are clearly the same between types, and are also constant over time. The degrees of current consumption positionality when young and when old for each type are equal to a and a', respectively, whereas the corresponding degrees of intertemporal consumption positionality are given by b and b'.

We can then easily consider the special case where individuals only care about the keeping up with the Joneses type of comparison, i.e., that only comparisons with others' current consumption matter. This is the case where $b = b' = 0$ above, implying that Eq. (33) reduces to

$$U_t^i = k^i [z_t^i]^{k_z^i} [c_t^i - a\bar{c}_t]^{k_c} [x_{t+1}^i - a'\bar{c}_{t+1}]^{k_x}$$

Returning to the general case with both keeping and catching up with the Joneses preferences, we have from Eq. (26) that the optimal marginal labor income tax rate for ability-type i is given by

$$T_t^{'}(w_t^i l_t^i) = \tau_t^i + [1 - \tau_t^i] \left[\frac{a + a'}{2} + \frac{b + b'}{2(1 + r)} \right], \tag{34}$$

where the first expression in brackets thus represents the average degree of current consumption positionality, and the second the (one period discounted) average degree of intertemporal consumption positionality. As before, τ_t^i is used as a short notation of the formula for marginal labor income taxation that would follow without any relative consumption concerns.

Regarding optimal capital income taxation, it is easy to see that the separability assumption in Corollary 1 above is fulfilled by the utility function in Eq. (33). Therefore, we know that the optimal marginal capital income tax rate is zero for each ability type and in all time periods, irrespective of the parameter values of the utility function.

5.1. Orders of magnitude

Let us now briefly discuss possible orders of magnitude of the optimal marginal income taxes. A couple of studies have attempted to measure the average degree of current consumption positionality, corresponding to $(a + a')/2$ in Eq. (34). According to the survey-experimental evidence of Solnick and Hemenway (1998), Johansson-Stenman et al. (2002), Alpízar et al. (2005), and Carlsson et al. (2007), the average degree of current consumption positionality appears to be in the order of magnitude of 0.5. Wendner and Goulder (2008) argue, based on the existing empirical evidence, for a value between 0.2 and 0.4, whereas evidence from happiness studies such as Luttmer (2005) suggests a much larger value in the order of magnitude of 0.8.

There is less direct evidence regarding the average degree of intertemporal consumption positionality, corresponding to $(b + b')/2$ in Eq. (34). Alvarez-Cuadrado et al. (2004) refer to a benchmark value used by Carroll et al. (1997), with a value of a parameter that can be interpreted as an intertemporal degree of consumption positionality equal to 0.5.[25] As a sensitivity analysis, they use a value of 0.8, based on Fuhrer (2000).

As an illustrative example, consider the case where $\tau_t^i = 0.3$ and where both the average degree of current consumption positionality and the average degree of intertemporal consumption positionality are also 0.3, i.e., $\tau_t^i = (a + a')/2 = (b + b')/2 = 0.3$. Then, if the real interest rate between the periods is given by $r = 1$,[26] it follows that the optimal marginal labor income tax rate is equal to $T_t^{'}(w_t^i l_t^i) = 0.3 + 0.7[0.3 + 0.3/2] = 0.615$. In other words, the optimal marginal labor income tax rate would be above 60% rather than 30% as suggested by the standard tax formula.

While the underlying estimates of the current and intertermporal degrees of positionality presented above are highly uncertain and can hardly be interpreted as completely independent of each other,[27] it nevertheless seems as if their joint effect on the marginal labor income tax rates may be substantial.

5.2. Explicit treatment of habit formation in own consumption and relative consumption

Let us finally return to the most general utility function we started from, i.e., Eq. (1a), where people also compare their consumption when old with their own consumption when young, as well as their relative consumption (compared with others' current consumption) when old with their relative consumption when young. Consider the following Cobb-Douglas utility function:

$$U_t^i = k^i (z_t^i)^{k_z^i} (C_{\text{net},t}^i)^{k_C}, \tag{35}$$

[25] In Carroll et al. (1997), the reference consumption is not others' average consumption one period earlier (since their study is not based on an OLG model), but instead a weighted average of others' average consumption where the weight is larger the closer to the present the consumption takes place.

[26] This corresponds to an annual real interest rate of slightly less than 2 percent if we assume 40 years between the periods.

[27] We are not aware of any study that simultaneously attempts to estimate the average degree of current and intertemporal consumption positionality.

where $k^i, k_z^i, k_C > 0$ are constants, $k_z^i + k_C < 1$, and

$$\begin{aligned} C^i_{\mathrm{net},t} &= [1-a-b+b''+b''']c^i_t + a[c^i_t - \bar{c}_t] + b[c^i_t - \bar{c}_{t-1}] + [1-a'-b'-b''-b''']x^i_{t+1} + a'[x^i_{t+1} - \bar{c}_{t+1}] \\ &\quad + b'[x^i_{t+1} - \bar{c}_t] + b''[(x^i_{t+1} - c^i_t)] + b'''[(x^i_{t+1} - \bar{c}_{t+1}) - (c^i_t - \bar{c}_t)] \\ &= c^i_t + x^i_{t+1} - b\bar{c}_{t-1} - [a+b'-b''']\bar{c}_t - [a'+b''']\bar{c}_{t+1} \end{aligned} \tag{36}$$

denotes the overall consumption net of relative consumption concerns. Clearly a, b, a', b', b'', and b''' reflect, respectively, the importance of comparisons with others' present consumption when young, comparisons with others' previous consumption when young, comparisons with others' present consumption when old, comparisons with others' previous consumption when old, comparisons with one's own previous consumption when old, and finally comparisons with one's own previous relative consumption when old.

Yet, as suggested in Eqs. (1b) and (1c) above, Eq. (36) can be rewritten as a reduced form in terms of current and intertemporal consumption comparisons as follows:

$$\begin{aligned} C^i_{\mathrm{net},t} &= [1-a-b+b''']c^i_t + [a-b'''][c^i_t - \bar{c}_t] + b[c^i_t - \bar{c}_{t-1}] + [1-a'-b'-b''']x^i_{t+1} \\ &\quad + [a'+b'''][x^i_{t+1} - \bar{c}_{t+1}] + b'[x^i_{t+1} - \bar{c}_t], \end{aligned} \tag{37}$$

where it is clear that the current and intertemporal degrees of positionality when young are given by $a - b'''$ and b, respectively. Correspondingly, the current and intertemporal degrees of positionality when old are given by $a' + b'''$ and b'. This means that we can again use Eq. (26), implying also that the optimal marginal labor income taxes are given by Eq. (34). Thus, not surprisingly, there is no effect on marginal taxation of comparisons with own previous consumption. Moreover, although the interpretations of the concepts of current and intertemporal degrees of positionality have to be modified due to concerns for changes in relative consumption over time, these effects cancel out such that there is no effect on optimal taxation.

6. Conclusion

The present paper simultaneously recognizes three mechanisms behind relative consumption concerns: comparisons with (i) other people's current consumption (keeping up with the Joneses), (ii) own past consumption (habit formation), and (iii) other people's past consumption (catching up with the Joneses). We are not aware of any previous normative economic analysis in such a setting.

We start by deriving a first-best tax policy to correct for the positional externalities in the case where the government is able to redistribute through lump-sum instruments. We show that comparisons with one's own past consumption do not affect the optimal policy rules, since such comparisons are internalized by each individual (although internal habit formation may, of course, affect the levels of marginal income tax rates), whereas comparisons with other people's current and past consumption generate positional externalities. In a stationary regime where the degrees of positionality are time-invariant, the optimal tax policy is derived in terms of the average degree of time-inclusive consumption positionality, which is essentially the sum of the average degree of current consumption positionality and the average degree of intertemporal consumption positionality. Results derived in earlier literature such as Ljungqvist and Uhlig (2000) and Dupor and Liu (2003) follow as special cases of our first-best model. We also show that the optimal marginal labor income tax rates become larger the more positional people are on average, in terms of the average degree of time-inclusive consumption positionality.

The second-best analysis carried out in Section 4 is based on the two-type optimal income tax model with asymmetric information between the government and the private sector. In this case, the net effects of relative consumption concerns also depend on whether the low-ability type is more or less positional (broadly speaking) than the mimicker. The reason is that this determines whether an increase in the reference consumption works to relax or tighten the self-selection constraint. If the degrees of positionality are constant over time, there are no direct effects of relative consumption concerns on the marginal capital income tax rates; in a second-best setting, such concerns will, nevertheless, affect the marginal capital income tax structure through the self-selection constraint. We are also able to reproduce the well-known result of Ordover and Phelps (1979) for when there should be no capital income taxes on the margin, in a model where people compare their own current consumption with several different measures of reference consumption.

When we generalize the model to allow for time variation also with regard to the positionality degrees, the population size, and the interest rate, the optimal policy responses become considerably more complex and the optimal policy rules are no longer possible to express in a simple way in terms of time-inclusive positionality degrees. This applies both to the first-best and the second-best model. Yet, we were able to obtain important findings regarding the qualitative effects of positional concerns on the optimal marginal income tax rates, and in particular, when such concerns unambiguously work to increase or decrease these tax rates.

Finally we illustrate with a Cobb–Douglas functional form and show, based on parameter estimates from the literature, that positional preferences of both the keeping up with the Joneses and catching up with the Joneses types substantially increase the optimal marginal labor income tax rates for both types. Since the leisure separability conditions are fulfilled for this form, the optimal marginal capital income tax rates are consequently zero for both types.

We believe that the research area consisting of normative economic analysis when relative consumption matters is still underexplored. Examples of important issues that remain to be analyzed include multi-country settings, public provision of private (non-positional) goods, public good provision in a dynamic economy, and long-term social discounting.

Appendix.

To save space, we have chosen to derive the more general results first, such that the results following from more restrictive formulations of the model appear as special cases.

Labor Income Taxation: derivation of Eqs. (19) and (28)

Consider the tax formula for the low-ability type. By combining Eqs. (23a) and (23b), we obtain

$$\frac{u^1_{t,z}}{u^1_{t,c}}\left[\lambda_t\hat{u}^2_{t,c}+\gamma_t n^1_t-\frac{n^1_t}{N_t}\frac{\partial\mathcal{L}}{\partial\bar{c}_t}\right]=\lambda_t\phi\hat{u}^2_{t,z}+\gamma_t n^1_t w^1_t. \tag{A1}$$

By substituting $T_t'(w^1_t l^1_t)w^1_t=w^1_t-u^1_{t,z}/u^1_{t,c}$ from Eq. (6) into Eq. (A1) and rearranging, we obtain Eq. (28) for the low-ability type. The corresponding formula for the high-ability type can be derived in the same general way by combining Eqs. (6), (23d) and (23e). Eq. (19) follows as the special case where $\lambda_t=0$.

Capital income taxation: derivation of Eqs. (20) and (29)

Consider first the formula for the low-ability type. By combining Eqs. (23b) and (23c), we obtain

$$\mathrm{MRS}^{1,t}_{c,x}\left[\lambda_t\hat{u}^2_{t,x}+\gamma_{t+1}n^1_t-\frac{n^1_t}{N_{t+1}}\frac{\partial\mathcal{L}}{\partial\bar{c}_{t+1}}\right]=\lambda_t\hat{u}^2_{t,c}+\gamma_t n^1_t-\frac{n^1_t}{N_t}\frac{\partial\mathcal{L}}{\partial\bar{c}_t}. \tag{A2}$$

We can then use Eqs. (7) and (23g) to derive $\mathrm{MRS}^{1,t}_{c,x}=1+r_{t+1}-r_{t+1}\Phi'_{t+1}(s^1_t r_{t+1})$ and $\gamma_t=\gamma_{t+1}[1+r_{t+1}]$, respectively. Substituting into Eq. (A2) and rearranging, we obtain Eq. (29) for the low-ability type. We can derive the corresponding expression for the high-ability type in the same general way by combining Eqs. (7), (23e), (23f) and (23g). Eq. (20) follows as the special case where $\lambda_{t-1}=\lambda_t=0$.

Derivation of Eqs. (15), (18), (25) and (30)

Consider first Eq. (30). By using Eq. (22), we can derive

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial\bar{c}_t}&=\sum_{i=1}^{2}\frac{\partial W}{\partial(n^i_{t-1}U^i_{t-1})}n^i_{t-1}u^i_{t-1,\bar{c}_t}+\sum_{i=1}^{2}\frac{\partial W}{\partial(n^i_t U^i_t)}n^i_t u^i_{t,\bar{c}_t}\\&\quad+\sum_{i=1}^{2}\frac{\partial W}{\partial(n^i_{t+1}U^i_{t+1})}n^i_{t+1}u^i_{t+1,\bar{c}_t}+\lambda_{t-1}[u^2_{t-1,\bar{c}_t}-\hat{u}^2_{t-1,\bar{c}_t}]+\lambda_t[u^2_{t,\bar{c}_t}-\hat{u}^2_{t,\bar{c}_t}]+\lambda_{t+1}[u^2_{t+1,\bar{c}_t}-\hat{u}^2_{t+1,\bar{c}_t}].\end{aligned} \tag{A3}$$

From Eqs. (1c) and (1d), we have

$$u^i_{t,c}=v^i_{t,c}+v^i_{t,c^i_t-\bar{c}_t}+v^i_{t,c^i_t-\bar{c}_{t-1}},$$

$$u^i_{t,x}=v^i_{t,x}+v^i_{t,x^i_{t+1}-\bar{c}_{t+1}}+v^i_{t,x^i_{t+1}-\bar{c}_t},$$

$$u^i_{t,\bar{c}_t}=-v^i_{t,c^i_t-\bar{c}_t}-v^i_{t,x^i_{t+1}-\bar{c}_t},$$

$$u^i_{t,\bar{c}_{t-1}}=-v^i_{t,c^i_t-\bar{c}_{t-1}},$$

$$u^i_{t,\bar{c}_{t+1}}=-v^i_{t,x^i_{t+1}-\bar{c}_{t+1}},$$

so

$$u^i_{t,\bar{c}_t}=-\alpha^{i,c}_t u^i_{t,c}-\beta^{i,x}_{t+1}u^i_{t,x}, \tag{A4}$$

$$u^i_{t,\bar{c}_{t-1}}=-\beta^{i,c}_t u^i_{t,c}, \tag{A5}$$

$$u^i_{t,\bar{c}_{t+1}} = -\alpha^{i,x}_{t+1} u^i_{t,x}, \tag{A6}$$

which substituted into Eq. (A3) imply

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \bar{c}_t} &= -\sum_{i=1}^{2} \frac{\partial W}{\partial (n^i_{t-1} U^i_{t-1})} n^i_{t-1} \alpha^{i,x}_t u^i_{t-1,x} - \sum_{i=1}^{2} \frac{\partial W}{\partial (n^i_t U^i_t)} n^i_t [\alpha^{i,c}_t u^i_{t,c} + \beta^{i,x}_{t+1} u^i_{t,x}] \\ &\quad - \sum_{i=1}^{2} \frac{\partial W}{\partial (n^i_{t+1} U^i_{t+1})} n^i_{t+1} \beta^{i,c}_{t+1} u^i_{t+1,c} + \lambda_{t-1}[-\alpha^{2,x}_t u^2_{t-1,x} + \hat{\alpha}^{2,x}_t \hat{u}^2_{t-1,x}] \\ &\quad + \lambda_t[-\alpha^{2,c}_t u^2_{t,c} - \beta^{2,x}_t u^2_{t+1,x} + \hat{\alpha}^{2,c}_t \hat{u}^2_{t,c} + \hat{\beta}^{2,x}_{t+1} \hat{u}^2_{t,x}] + \lambda_{t+1}[-\beta^{2,c}_{t+1} u^2_{t+1,c} + \hat{\beta}^{2,c}_{t+1} \hat{u}^2_{t+1,c}].\end{aligned} \tag{A7}$$

From Eqs. (23b), (23c), (23e) and (23f), we have

$$\frac{\partial W}{\partial (n^1_t U^1_t)} n^1_t u^1_{t,c} = \lambda_t \hat{u}^2_{t,c} + \gamma_t n^1_t - \frac{n^1_t}{N_t} \frac{\partial \mathcal{L}}{\partial \bar{c}_t}, \tag{A8}$$

$$\frac{\partial W}{\partial (n^2_t U^2_t)} n^2_t u^2_{t,c} = -\lambda_t u^2_{t,c} + \gamma_t n^2_t - \frac{n^2_t}{N_t} \frac{\partial \mathcal{L}}{\partial \bar{c}_t}, \tag{A9}$$

$$\frac{\partial W}{\partial (n^1_{t-1} U^1_{t-1})} n^1_{t-1} u^1_{t-1,x} = \lambda_{t-1} \hat{u}^2_{t-1,x} + \gamma_t n^1_{t-1} - \frac{n^1_{t-1}}{N_t} \frac{\partial \mathcal{L}}{\partial \bar{c}_t}, \tag{A10}$$

$$\frac{\partial W}{\partial (n^2_{t-1} U^2_{t-1})} n^2_{t-1} u^2_{t-1,x} = -\lambda_{t-1} u^2_{t-1,x} + \gamma_t n^2_{t-1} - \frac{n^2_{t-1}}{N_t} \frac{\partial \mathcal{L}}{\partial \bar{c}_t}. \tag{A11}$$

By substituting Eqs. (A8)–(A11) into Eq. (A7), and collecting terms, we obtain

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \bar{c}_t} &= \frac{\partial \mathcal{L}}{\partial \bar{c}_{t+1}} \frac{\bar{\beta}_{t+1}}{1-\bar{\alpha}_t} - N_t \gamma_t \frac{\bar{\alpha}_t}{1-\bar{\alpha}_t} - N_{t+1}\gamma_{t+1} \frac{\bar{\beta}_{t+1}}{1-\bar{\alpha}_t} + \frac{\lambda_{t-1} \hat{u}^2_{t-1,x}}{1-\bar{\alpha}_t}[\hat{\alpha}^{2,x}_t - \alpha^{1,x}_t] + \frac{\lambda_t \hat{u}^2_{t,c}}{1-\bar{\alpha}_t}[\hat{\alpha}^{2,c}_t - \alpha^{1,c}_t] + \frac{\lambda_t \hat{u}^2_{t,x}}{1-\bar{\alpha}_t}[\hat{\beta}^{2,x}_{t+1} - \beta^{1,x}_{t+1}] \\ &\quad + \frac{\lambda_{t+1} \hat{u}^2_{t+1,c}}{1-\bar{\alpha}_t}[\hat{\beta}^{2,c}_{t+1} - \beta^{1,c}_{t+1}] = \frac{1}{1-\bar{\alpha}_t}\left[\bar{\beta}_{t+1} \frac{\partial \mathcal{L}}{\partial \bar{c}_{t+1}} + N_t \gamma_t [\alpha^d_t - \bar{\alpha}_t] + N_{t+1}\gamma_{t+1}[\beta^d_{t+1} - \bar{\beta}_{t+1}]\right],\end{aligned} \tag{A12}$$

where we have used the short notations a^d_t and β^d_t as defined earlier. Using the short notations

$$A_t = \frac{N_t \gamma_t [\alpha^d_t - \bar{\alpha}_t]}{1-\bar{\alpha}_t},$$

$$B_t = \frac{N_{t+1}\gamma_{t+1}[\beta^d_{t+1} - \bar{\beta}_{t+1}]}{1-\bar{\alpha}_t},$$

$$\varphi_t = \frac{\bar{\beta}_{t+1}}{1-\bar{\alpha}_t},$$

the recursive Eq. (A12) can more conveniently be rewritten and expanded as

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \bar{c}_t} &= A_t + B_t + \varphi_t \frac{\partial \mathcal{L}}{\partial \bar{c}_{t+1}} = A_t + B_t + \varphi_t \left[A_{t+1} + B_{t+1} + \varphi_{t+1} \frac{\partial \mathcal{L}}{\partial \bar{c}_{t+2}}\right] \\ &= A_t + B_t + \varphi_t \left[A_{t+1} + B_{t+1} + \varphi_{t+1}\left[A_{t+2} + B_{t+2} + \varphi_{t+2} \frac{\partial \mathcal{L}}{\partial \bar{c}_{t+3}}\right]\right] \\ &= A_t + B_t + [A_{t+1} + B_{t+1}]\varphi_t + [A_{t+2} + B_{t+2}]\varphi_t\varphi_{t+1} + [A_{t+3} + B_{t+3}]\varphi_t\varphi_{t+1}\varphi_{t+2}\ldots \\ &= A_t + B_t + \sum_{i=1}^{\infty}[A_{t+i} + B_{t+i}]\prod_{j=1}^{i}\varphi_{t+j-1}\end{aligned} \tag{A13}$$

Substituting back $\varphi_t = \bar{\beta}_{t+1}/[1-\bar{\alpha}_t]$ into Eq. (A13) implies Eq. (30). Eq. (18) follows as the special case where $\lambda_t = 0$ for all t.

Finally, by adding assumptions (i)–(iv) (i.e., *i.* $N_t = N$; (ii) $\bar{\alpha}_t = \bar{\alpha}$ and $\bar{\beta}_t = \bar{\beta}$; (iii) $r_t = r$; and iv. $\bar{\alpha}_t^d = \bar{\alpha}^d$ and $\bar{\beta}_t^d = \bar{\beta}^d$, for all t), Eq. (30) reduces to Eq. (25) and Eq. (18) reduces to Eq. (15). To see this, use assumptions i–iv in Eq. (30) to derive

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \bar{c}_t} &= \frac{N\gamma_t}{1-\bar{\alpha}}\left[\alpha^d - \bar{\alpha} + \frac{\beta^d - \bar{\beta}}{1+r}\right]\sum_{i=0}^{\infty}\left[\frac{\bar{\beta}}{(1-\bar{\alpha})(1+r)}\right]^i \\ &= N\gamma_t \frac{\alpha^d - \bar{\alpha} + (\beta^d - \bar{\beta})/(1+r)}{1-\bar{\alpha}-\bar{\beta}/(1+r)}\end{aligned}.$$

The second line presupposes that $0 < \bar{\beta} < (1-\bar{\alpha})(1+r)$, such that the series converges. Using the definitions of $\bar{\rho}$ and ρ^d gives Eq. (25).

Derivation of Eqs. (16) and (26)

Consider first Eq. (26) for the low-ability type. By combining Eqs. (25) and (28), we obtain

$$T_t'(w_t^1 l_t^1) = \frac{\lambda_t^*}{w_t^1 n_t^1}[\mathrm{MRS}_{z,c}^{1,t} - \phi \mathrm{M\hat{R}S}_{z,c}^{2,t}] - \frac{\mathrm{MRS}_{z,c}^{1,t}}{w_t^1}\frac{\rho^d - \bar{\rho}}{1-\bar{\rho}}. \tag{A14}$$

Then, by using $\mathrm{MRS}_{z,c}^{1,t}/w_t^1 = 1 - T_t'(w_t^1 l_t^1)$ and rearranging, we obtain Eq. (26) for the low-ability type. The corresponding marginal income tax rate for the high-ability type is derived in a similar way. Eq. (16) follows as the special case where $\lambda_t = 0$.

Derivation of Eqs. (17) and (27)

Consider first Eq. (27). Substituting Eq. (25), for period t and period $t+1$, into Eq. (29), we obtain

$$\Phi_{t+1}'(s_t^i r_{t+1}) = \delta_t^i + \frac{1}{r}\frac{\rho^d - \bar{\rho}}{1-\bar{\rho}}\left[\frac{\gamma_t}{\gamma_{t+1}} - \mathrm{MRS}_{c,x}^{i,t}\right], \tag{A15}$$

for $i = 1, 2$, where we have used the short notations δ_t^1 and δ_t^2 as defined earlier. Using $\mathrm{MRS}_{c,x}^{i,t} = 1 + r - r\Phi_{t+1}'(s_t^i r_{t+1})$ together with $\gamma_t/\gamma_{t+1} = 1 + r$ in Eq. (A15) and rearranging, we obtain Eq. (27). Eq. (17) follows as the special case where $\lambda_{t-1} = \lambda_t = 0$.

Brief derivations of Eqs. (16′) and (17′)

The utility function facing ability-type i of generation t is given by

$$\begin{aligned}U_t^i &= v_t^i(c_t^i, z_t^i, x_{t+1}^i, c_t^i - \bar{c}_t^{\text{young}}, x_{t+1}^i - \bar{x}_{t+1}^{\text{old}}, c_t^i - \bar{c}_{t-1}^{\text{young}}, x_{t+1}^i - \bar{x}_t^{\text{old}}) \\ &= u_t^i(c_t^i, z_t^i, x_{t+1}^i, \bar{c}_t^{\text{young}}, \bar{c}_{t-1}^{\text{young}}, \bar{x}_{t+1}^{\text{old}}, \bar{x}_t^{\text{old}})\end{aligned}. \tag{A16}$$

The resource constraint is given by Eq. (12) and the Lagrangian by Eq. (13), in which the social welfare function is based on Eq. (A16) instead of on Eqs. (1). The social first-order conditions for the hours of work and the capital stock remain as in Eqs. (14a) and (14d), respectively, while the social first-order conditions for private consumption change to (where $n_t = n_t^1 + n_t^2$)

$$\frac{\partial W}{\partial (n_t^i U_t^i)} n_t^i u_{t,c}^i - \gamma_t n_t^i + \frac{n_t^i}{n_t}\frac{\partial \mathcal{L}}{\partial \bar{c}_t^{\text{young}}} = 0, \tag{A17}$$

$$\frac{\partial W}{\partial (n_t^i U_t^i)} n_t^i u_{t,x}^i - \gamma_{t+1} n_t^i + \frac{n_t^i}{n_{t+1}}\frac{\partial \mathcal{L}}{\partial \bar{x}_{t+1}^{\text{old}}} = 0. \tag{A18}$$

By using Eqs. (6), (7), (14a), (14d), (A17) and (A18), we can derive the following analogs to Eqs. (19) and (20):

$$T_t'(w_t^i l_t^i) = -\frac{\mathrm{MRS}_{z,c}^{i,t}}{\gamma_t w_t^i n_t}\frac{\partial \mathcal{L}}{\partial \bar{c}_t^{\text{young}}}, \tag{A19}$$

$$\Phi_{t+1}'(s_t^i r_{t+1}) = \frac{1}{\gamma_{t+1} r_{t+1} n_t}\left[\frac{\partial \mathcal{L}}{\partial \bar{c}_t^{\text{young}}} - \mathrm{MRS}_{c,x}^{i,t}\frac{\partial \mathcal{L}}{\partial \bar{x}_{t+1}^{\text{old}}}\right]. \tag{A20}$$

Note also that

$$\frac{\partial \mathcal{L}}{\partial \bar{c}_t^{\text{young}}} = \sum_{i=1}^{2} \frac{\partial W}{\partial (n_t^i U_t^i)} n_t^i u^i_{t,\bar{c}_t^{\text{young}}} + \sum_{i=1}^{2} \frac{\partial W}{\partial (n_{t+1}^i U_{t+1}^i)} n_{t+1}^i u^i_{t+1,\bar{c}_t^{\text{young}}}$$
$$= -\sum_{i=1}^{2} \frac{\partial W}{\partial (n_t^i U_t^i)} n_t^i u^i_{t,c} \alpha_t^{i,\text{young}} - \sum_{i=1}^{2} \frac{\partial W}{\partial (n_{t+1}^i U_{t+1}^i)} n_{t+1}^i u^i_{t+1,c} \beta_{t+1}^{i,\text{young}}, \tag{A21}$$

$$\frac{\partial \mathcal{L}}{\partial \bar{x}_{t+1}^{\text{old}}} = \sum_{i=1}^{2} \frac{\partial W}{\partial (n_t^i U_t^i)} n_t^i u^i_{t,\bar{x}_{t+1}^{\text{old}}} + \sum_{i=1}^{2} \frac{\partial W}{\partial (n_{t+1}^i U_{t+1}^i)} n_{t+1}^i u^i_{t+1,\bar{x}_{t+1}^{\text{old}}}$$
$$= -\sum_{i=1}^{2} \frac{\partial W}{\partial (n_t^i U_t^i)} n_t^i u^i_{t,x} \alpha_{t+1}^{i,\text{old}} - \sum_{i=1}^{2} \frac{\partial W}{\partial (n_{t+1}^i U_{t+1}^i)} n_{t+1}^i u^i_{t+1,x} \beta_{t+2}^{i,\text{old}}. \tag{A22}$$

Substituting Eqs. (A17) and (A18) and their counterparts for generation $t+1$ into Eqs. (A21) and (A22), rearranging, and using the variants of assumptions (i)–(iii) discussed in Section 3.4 give

$$\frac{\partial \mathcal{L}}{\partial \bar{c}_t^{\text{young}}} = -\gamma_t n \frac{\bar{\rho}^{\text{young}}}{1-\bar{\rho}^{\text{young}}}, \tag{A23}$$

$$\frac{\partial \mathcal{L}}{\partial \bar{x}_{t+1}^{\text{old}}} = -\gamma_{t+1} n \frac{\bar{\rho}^{\text{old}}}{1-\bar{\rho}^{\text{old}}}. \tag{A24}$$

Substituting Eqs. (A23) and (A24) into Eqs. (A19) and (A20) gives Eqs. (16′) and (17′).

References

Abel, A.B., 1990. Asset prices under habit formation and catching up with the Joneses. American Economic Review, Papers and Proceedings 80, 38–42.
Abel, A.B., 2005. Optimal taxation when consumers have endogenous benchmark levels of consumption. Review of Economic Studies 72, 1–19.
Acemoglu, D., Golosov, M., Tsyvinski, A., 2011. The political economy of Ramsey taxation. Journal of Public Economics 95, 467–475.
Akerlof, G.A., 1997. Social distance and social decisions. Econometrica 65, 1005–1027.
Alvarez-Cuadrado, F., Monteiro, G., Turnovsky, S., 2004. Habit formation, catching up with the joneses, and economic growth. Journal of Economic Growth 9, 47–80.
Alvarez-Cuadrado, F., Long, N.V., 2011. The relative income hypothesis. Journal of Economic Dynamics and Control 35, 1489–1501.
Alvarez-Cuadrado, F., Long, N.V., 2012. Envy and inequality. Scandinavian Journal of Economics 114, 949–973.
Alpízar, F., Carlsson, F., Johansson-Stenman, O., 2005. How much do we care about absolute versus relative income and consumption? Journal of Economic Behavior and Organization 56, 405–421.
Aronsson, T., Johansson-Stenman, O., 2008. When the Joneses' consumption hurts: optimal public good provision and nonlinear income taxation. Journal of Public Economics 92, 986–997.
Aronsson, T., Johansson-Stenman, O., 2010. Positional concerns in an OLG model: optimal labor and capital income taxation. International Economic Review 51, 1071–1095.
Aronsson, T., Johansson-Stenman, O., 2013. Conspicuous leisure: optimal income taxation when both relative consumption and relative leisure matter. Scandinavian Journal of Economics 115, 155–175.
Aronsson, T., Schöb, R., 2013. Consumption Adaptation, Anticipation-Bias and Optimal Income Taxation. Umeå University, Mimeo.
Aronsson, T., Sjögren, T., Dalin, T., 2009. Optimal taxation and redistribution in an OLG model with unemployment. International Tax and Public Finance 16, 198–218.
Arrow, J.K., Dasgupta, P., 2009. Conspicuous consumption, inconspicuous leisure. Economic Journal 119, F497–F516.
Atkinson, A.B., Stiglitz, J.E., 1976. The design of tax structure: direct versus indirect taxation. Journal of Public Economics 6, 55–75.
Blanchflower, D.G., Oswald, A.J., 2004. Well-being over time in Britain and the USA. Journal of Public Economics 88, 1359–1386.
Blomquist, S., 1993. Interdependent behavior and the effect of taxes. Journal of Public Economics 51, 211–218.
Boskin, M.J., Sheshinski, E., 1978. Individual welfare depends upon relative income. Quarterly Journal of Economics 92, 589–601.
Bowles, S., Park, Y.-J., 2005. Inequality, emulation, and work hours: was Thorsten Veblen right? Economic Journal 15, F397–F413.
Brekke, K.-A., Howarth, R.B., 2002. Affluence, Well-Being and Environmental Quality. Edward Elgar, Cheltenham.
Brett, C., 1997. A Note on Nonlinear Taxation in an Overlapping Generations Model Mimeo. Department of Economics, University of Essex.
Brett, C., Weymark, J.A., 2008. Optimal Nonlinear Taxation of Income and Savings Without Commitment. Working paper 08-W05. Department of Economics, Vanderbilt University.
Campbell, J.Y., Cochrane, J.H., 1999. By force of habit: a consumption-based explanation of aggregate stock market behavior. Journal of Political Economy 107, 205–251.
Carlsson, F., Johansson-Stenman, O., Martinsson, P., 2007. Do you enjoy having more than others? Survey evidence of positional goods. Economica 74, 586–598.
Carroll, C.D., Overland, J.R., Weil, D.N., 1997. Comparison utility in a growth model. Journal of Economic Growth 2, 339–367.
Chan, Y.L., Kogan, L., 2002. Catching up with the joneses: heterogeneous preferences and the dynamics of asset prices. Journal of Political Economy 110, 1255–1285.
Clark, A.E., Frijters, P., Shields, M.A., 2008. Relative income, happiness and utility: an explanation for the easterlin paradox and other puzzles. Journal of Economic Literature 46, 95–144.
Clark, A., Senik, C., 2010. Who compares to whom? The anatomy of income comparisons in Europe. Economic Journal 120, 573–594.
Constantinides, G.M., 1990. Habit formation: a resolution of the equity premium puzzle. Journal of Political Economy 98, 519–543.
Corneo, G., Jeanne, O., 1997. Conspicuous consumption, snobbism and conformism. Journal of Public Economics 66, 55–71.
Corneo, G., Jeanne, O., 2001. Status, the distribution of wealth, and growth. Scandinavian Journal of Economics 103, 283–293.
Daly, M., Wilson, D.J., 2009. Happiness, unhappiness, and suicide: an empirical assessment. Journal of the European Economic Association 7, 539–549.

Di Tella, R., Haisken-De New, J., MacCulloch, R., 2010. Happiness adaptation to income and to status in an individual panel. Journal of Economic Behavior and Organization 76, 834–852.
Díaz, A., Pijoan-Mas, J., Ríos-Rull, J.V., 2003. Precautionary savings and wealth distribution under habit formation preferences. Journal of Monetary Economics 50, 1257–1291.
Dohmen, T., Falk, A., Fliessbach, K., Sunde, U., Weber, B., 2011. Relative versus absolute income, joy of winning, and gender: brain imaging evidence. Journal of Public Economics 70, 279–285.
Dupor, B., Liu, W.F., 2003. Jealousy and overconsumption. American Economic Review 93, 423–428.
Easterlin, R.A., 1995. Will raising the incomes of all increase the happiness of all? Journal of Economic Behavior and Organization 27, 35–47.
Easterlin, R.A., 2001. Income and happiness: towards a unified theory. Economic Journal 111, 465–484.
Erosa, A., Gervais, M., 2002. Optimal taxation in life-cycle economies. Journal of Economic Theory 105, 338–369.
Ferrer-i-Carbonell, A., 2005. Income and well-being: an empirical analysis of the comparison income effect. Journal of Public Economics 89, 997–1019.
Fliessbach, K., Weber, B., Trautner, P., Dohmen, T., Sunde, U., Elger, C.E., Falk, A., 2007. Social comparison affects reward-related brain activity in the human ventral striatum. Science 318, 1305–1308.
Frank, R.H., 1985. The demand for unobservable and other nonpositional goods. American Economic Review 75, 101–116.
Frank, R.H., 1999. Luxury Fever: Money and Happiness in an Era of Excess. Princeton University Press, Princeton.
Frank, R.H., 2005. Positional externalities cause large and preventable welfare losses. American Economic Review 95, 137–141.
Frank, R.H., 2007. Falling Behind: How Rising Inequality Harms the Middle Class. University of California Press, Berkeley.
Frank, R.H., 2008. Should public policy respond to positional externalities? Journal of Public Economics 92, 1777–1786.
Frank, R.H., Hutchens, R.M., 1993. Wages, seniority, and the demand for rising consumption profiles. Journal of Economic Behavior and Organization 21, 251–276.
Fuhrer, J.C., 2000. Habit formation in consumption and its implications for monetary-policy models. American Economic Review 90, 367–390.
Goerke, L., 2013. Relative consumption and tax evasion. Journal of Economic Behavior and Organization 87, 52–65.
Gomez, M.A., 2006. Optimal consumption taxation in a model of endogenous growth with external habit formation. Economics Letters 93, 427–435.
Guo, J-T., Krause, A., 2011. Optimal nonlinear income taxation with habit formation. Journal of Public Economic Theory 13, 463–480.
Hopkins, E., Kornienko, T., 2004. Running to keep in the same place: consumer choice as a game of status. American Economic Review 94, 1085–1107.
Hopkins, E., Kornienko, T., 2009. Status, affluence, and inequality: rank-based comparisons in games of status. Games and Economic Behavior 67, 552–568.
Howarth, R.B., 1996. Status effects and environmental externalities. Ecological Economics 16, 25–34.
Howarth, R.B., 2006. Optimal environmental taxation under relative consumption effects. Ecological Economics 58, 209–219.
Ireland, N.J., 2001. Optimal income tax in the presence of status effects. Journal of Public Economics 81, 193–212.
Johansson-Stenman, O., Carlsson, F., Daruvala, D., 2002. Measuring future grandparents' preferences for equality and relative standing. Economic Journal 112, 362–383.
Kocherlakota, N., 2005. Zero expected wealth taxes a Mirrlees approach to dynamic optimal taxation. Econometrica 73, 1587–1621.
Layard, R., 1980. Human satisfaction and public policy. Economic Journal 90, 737–750.
Ljungqvist, L., Uhlig, H., 2000. Tax policy and aggregate demand management under catching up with the Joneses. American Economic Review 90, 356–366.
Loewenstein, G., O'Donoghue, T., Rabin, M., 2003. Projection bias in predicting future utility. Quarterly Journal of Economics 118, 1209–1248.
Loewenstein, G.F., Sicherman, N., 1991. Do workers prefer increasing wage profiles? Journal of Labor Economics 9, 67–84.
Luttmer, E.F.P., 2005. Neighbors as negatives: relative earnings and well-being. Quarterly Journal of Economics 120, 963–1002.
Marmot, M., 2004. The Status Syndrome: How Your Social Standing Directly Affects Your Health. Bloomsbury, London.
Mathieu-Bolh, N., 2006. Optimal taxation and finite horizon. Economics Letters 91, 215–221.
McBride, M., 2001. Relative-income effects on subjective well-being in the cross-section. Journal of Economic Behavior and Organization 45, 251–278.
Mujcic, R., Frijters, P., 2013. Economic choices and status: measuring preferences for income rank. Oxford Economic Papers 65, 47–73.
Ng, Y.-K., 1987. Relative income effects and the appropriate level of public expenditure. Oxford Economic Papers 39, 293–300.
Oh, S.-Y., Park, Y.-J., Bowles, S., 2012. Veblen effects, political representation, and the reduction in working time over the 20th century. Journal of Economic Behavior and Organization 83, 218–242.
Ordover, J.E., Phelps, E.S., 1979. The concept of optimal taxation in the overlapping generations model of capital and wealth. Journal of Public Economics 12, 1–26.
Oswald, A., 1983. Altruism, jealousy and the theory of optimal non-linear taxation. Journal of Public Economics 20, 77–87.
Pirttilä, J., Tuomala, M., 2001. On optimal non-linear taxation and public good provision in an overlapping generations economy. Journal of Public Economics 79, 485–501.
Rayo, L., Becker, G.S., 2007. Evolutionary efficiency and happiness. Journal of Political Economy 115, 302–337.
Runciman, W.G., 1966. Relative Deprivation and Social Justice. Routledge and Kegan Paul, London.
Senik, C., 2009. Direct evidence on income comparisons and their welfare effects. Journal of Economic Behavior and Organization 72, 408–424.
Solnick, S., Hemenway, D., 1998. Is more always better?: A survey on positional concerns. Journal of Economic Behavior and Organization 37, 373–383.
Solnick, S., Hemenway, D., 2005. Are positional concerns stronger in some domains than in others? American Economic Review, Papers and Proceedings 45, 147–151.
Stern, N.H., 1982. Optimum taxation with errors in administration. Journal of Public Economics 17, 181–211.
Stiglitz, J.E., 1982. Self-selection and pareto efficient taxation. Journal of Public Economics 17, 213–240.
Stutzer, A., 2004. The role of income aspirations in individual happiness. Journal of Economic Behavior and Organization 54, 89–110.
Tuomala, M., 1990. Optimal Income Tax and Redistribution. Clarendon Press, Oxford.
Veblen, T., 1899. The Theory of the Leisure Class. Macmillan, New York.
Wendner, R., 2005. Frames of reference, the environment, and efficient taxation. Economics of Governance 6, 13–31.
Wendner, R., 2010a. Conspicuous consumption and generation replacement in a model with perpetual youth. Journal of Public Economics 94, 1093–1107.
Wendner, R., 2010b. Growth and keeping up with the Joneses. Macroeconomic Dynamics 14, 176–199.
Wendner, R., 2011. Will the consumption externalities' effects in the ramsey model please stand up? Economics Letters 111, 210–212.
Wendner, R., 2014. Ramsey, Pigou, heterogeneous agents, and non-atmospheric consumption externalities. Journal of Public Economic Theory (forthcoming).
Wendner, R., Goulder, L.H., 2008. Status effects, public goods provision, and the excess burden. Journal of Public Economics 92, 1968–1985.