Future-biased government*

Francisco M. Gonzalez, Itziar Lazkano and Sjak A. Smulders

October 2015

Abstract

We argue that governments are future biased when they aggregate the preferences of overlapping generations. Future bias, which involves preference reversals favoring future over current consumption, explains why governments legislate old-age transfers at the expense of capital accumulation and growth, even if generations are altruistic.

JEL classification: D71; D72; H55.
Keywords: collective preferences, time inconsistency; future bias; altruism; intergenerational redistribution; growth.

*Gonzalez: University of Waterloo (francisco.gonzalez@uwaterloo.ca). Lazkano: University of Wisconsin-Milwaukee (lazkano@uwm.edu). Smulders: Tilburg University (J.A.Smulders@uvt.nl). We have benefited from comments by Stefan Ambec, Hippolyte d'Albis, Matt Doyle, John Hassler, Per Krusell, David Levine, Fabien Postel-Vinay, Victor Rios-Rull and Randy Wright. Gonzalez gratefully acknowledges financial support from the Social Sciences and Humanities Research Council of Canada.

1 Introduction

For a government, the relevant calculus of optimal intergenerational redistribution is such that the values of the present-day society are to be used to evaluate the distribution of consumption between present and future generations. Intergenerational disagreements create potentially serious problems. One is that governments have to strike a balance between the preferences of presently young and old people. Another is that the effect of current policies on future social outcomes can be undone by future policies.

Phelps and Pollak (1968), and more recently Krusell et al. (2002), analyze second-best capital accumulation under the assumption that governments' preferences are *present biased*, favoring current over future consumption. In their non-overlapping generations setting, which is the standard setting in analyses of second-best growth with intergenerational disagreement, present-biased social preferences are a natural consequence of parents' imperfect altruism towards children.[1] This approach is consistent with the common view that present bias is inherent to all utilitarian aggregation of heterogeneous time preferences (Jackson and Yariv, 2014).[2] Weighted utilitarian functions are relevant because they are the most prominent class of collective utility functions and they can be derived as the outcome of political competition (Lindbeck and Weibull 1987).

By contrast, in this paper we show that, when they aggregate the preferences of overlapping generations, utilitarian governments suffer from a systematic *future bias*. Intuitively, whereas present bias is characterized by the possibility of preference reversals favoring current over future consumption, future bias is characterized by the possibility of preference reversals favoring future over current consumption.

Future bias in governments' preferences matters for policy making. In contrast with present bias, which causes systematic undersaving and favors pro-growth policy, future bias provides an incentive to legislate intergenerational redistribution to increase future old-age transfers at the expense of growth.

We trace the source of future bias to intergenerational disagreement stemming from the

[1] Kohlberg (1976), Bernheim and Ray (1987), Ray (1987) and Barro (1999) are well known examples.
[2] Also see Caplin and Leahy (2004) and Gollier and Zeckhauser (2005).

combination of children's imperfect altruism towards parents and the overlap of generations that takes place in actual economies. The combination of these two elements translates into disagreement between coexisting generations about the distribution of current aggregate consumption. It is well known that this kind of intergenerational disagreement renders plausible social welfare functions time-inconsistent, even if parents are perfectly altruistic towards children.[3] However, today's young generations are tomorrow's old, and so it is unclear how intergenerational disagreement at each point in time translates into intertemporal allocations. In particular, neither the specific form of time inconsistency that one may expect to afflict social preferences nor its implications for government intervention are well understood.

For simplicity, we develop our argument in the case where children are not altruistic towards parents (e.g., Barro 1974). In order to focus on the equilibrium interaction between present and future policy makers we assume that policy makers have sufficient instruments to implement their desired allocation. This allows us to abstract away from the role of private economic decisions. The simplicity of our stylized example allows us to characterize the specific form of time inconsistency that arises and its consequences.

To understand the source of future bias, and its implications, it is helpful to consider it first in the context of non-altruistic generations. Consider a sequence of governments that live for one period and seek to maximize a weighted sum of the utilities of living generations. Suppose a generation is born every period and lives for two periods, so only two generations coexist every period. A future bias arises because the date-t government does not value the well-being of the generation born at date $t+1$, whereas the date-$t+1$ government does. Accordingly, there will be both too little redistribution towards the old and too much saving at date $t+1$ from the viewpoint of the date-t government. Hence, the date-t government has an incentive to increase redistribution towards the old at date $t+1$. But if the current government cannot control future old-age transfers directly, it will have an incentive to influence future income instead, which can be achieved by distorting current investment.

[3] Burbidge (1983) and Calvo and Obstfeld (1988) recognize the potential time-inconsistency, but they do not address its consequences. Bernheim (1989) and Hori (1997) are two interesting analyses of this problem.

Now suppose that the weight of the young in social preferences (i.e., their political weight) is relatively large. Then, not only do current governments not care about future generations, but their main concern is to make sure that the current young get enough consumption when they are old. While increasing current investment and growth is costly, governments will be willing to do so if they have a strong enough preference for consumption smoothing.

When we analyze the properties of equilibrium growth, we conclude that the future bias in social preferences is conducive to growth, rather than inimical to it. We show that this effect of future bias implies that the present-day government has an incentive to legislate old-age transfers at the expense of growth. Moreover, future governments have an incentive to sustain the legislation, because its negative impact on their wealth is due to lower investments that will be sunk from the viewpoint of every future government.

The above intuition helps identifying the role of future bias but it is misleading in that it suggests, incorrectly, that the source of future bias is that the old have relatively little political power and do not care about the young. By contrast, our analysis shows that both future bias and the resulting incentive to legislate and sustain old-age transfers at the expense of growth arise even when the old are altruistic towards the young. Moreover, this is so regardless of the relative political weight of the old and the young, as long as both of them have positive weight.

Even though perfectly altruistic individuals have time-consistent preferences, governments are future biased, because the current young do not value the consumption of the current old, and social welfare puts positive weight on the current young. Since the current generations value future consumption equally, the current government systematically favors future over current aggregate consumption. It will become clear that the source of future bias is that children do not care *enough* about their parents, not that they do not care at all.

Our analysis offers a novel perspective on the widespread legislation of pay-as-you-go social security despite recognition of its negative effects on capital accumulation (Auerbach and Kotlikoff 1987). Grossman and Helpman (1998) also consider a sequence of governments that cannot precommit future redistributive policy. Social security arises in equilibrium because the preferences of politicians are biased towards the old. By contrast, our theory does not depend on the relative political weigh of young and old agents, as long as both of them have some weight. Veall (1986) and Hansson and Stuart (1989) presume that children are sufficiently altruistic towards parents and, in the absence of social security, the current young would have an incentive to undersave in order to elicit old-age transfers from the future young.[4] By contrast, we argue that, in the absence of social security, policy makers would have an incentive to oversave in order to increase future old-age transfers, precisely because children are insufficiently altruistic towards parents.

The next section presents the basic model. Section 3 illustrates the logic of our argument in the simplest case, where individuals are non-altruistic. Section 4 considers the case of altruistic generations. We first illustrate the future bias in social preferences in the context of the first-best allocation for an arbitrary government. Then, we analyze the symmetric Markov perfect equilibrium in linear strategies and discuss the main implications of our analysis. Section 5 concludes. Technical proofs are in the Appendix.

2 The model

Consider an economy with overlapping altruistic generations. A unit mass of individuals are born every period $t \geq 0$, each individual lives for two periods and individuals born at date t have preferences given by

$$\begin{aligned} u_t &= u\left(c_t^y\right) + u\left(c_{t+1}^o\right) + \delta u_{t+1} \\ &= u\left(c_t^y\right) + u\left(c_{t+1}^o\right) + \sum_{s=1}^{\infty} \delta^s \left(u\left(c_{t+s}^y\right) + u\left(c_{t+1+s}^o\right)\right), \end{aligned} \tag{1}$$

with $\delta \in (0,1)$, where c_t^y is the consumption of young agents at date t, and c_{t+1}^o is their consumption when old. For simplicity, we assume that individuals do not discount their

[4] Tabellini (2000) also presumes that children are sufficiently altruistic towards parents, but he stresses the fact that social security redistributes both across and within generations.

second-period felicity, and also that felicity functions each period are isoelastic, with

$$u(c) = \begin{cases} \frac{c^{1-\sigma}}{1-\sigma} & \text{if } \sigma \neq 1, \quad \sigma > 0 \\ \ln(c) & \text{if } \sigma = 1. \end{cases} \tag{2}$$

Output is linear in the capital stock at the aggregate level, where k_t units of capital produce Ak_t units of output that become available at date $t+1$, with $A > 0$. The aggregate resources constraint in the economy is given by

$$Ak_t \geq c_t^y + c_t^o + k_{t+1} - k_t, \tag{3}$$

where we have ignored depreciation of the capital stock. We assume that $\delta(A+1) > 1$ in order to ensure positive equilibrium growth rates. We also assume that $\delta(A+1)^{1-\sigma} < 1$ in order to ensure that growth is not so fast that it leads to unbounded utility.

We consider a sequence of governments. The government at date t seeks to maximize a weighted sum of the present generations' utilities:

$$v_t = u_{t-1} + au_t, \tag{4}$$

where $a > 0$. The assumption of a utilitarian welfare objective is meant to capture in a simple manner the fact that democratic governments are unlikely to be immune to disagreement between coexisting generations. It can be interpreted as the outcome of political competition in a probabilistic voting model (Lindbeck and Weibull 1987, Grossman and Helpman 1998).

The welfare of each generation is influenced by the actions of different governments. Consequently, each government's optimal behavior depends on its expectation of future governments' behavior. Since every government can affect future aggregate economic conditions, equilibrium allocations depend on the interaction between current and future governments. We consider the following problem. Every period t the government chooses investment $(k_{t+1} - k_t)$, and consumption (c_t^y and c_t^o) in order to maximize (9) subject to (2) and (3), taking as given the strategies of all other governments. A Markov strategy of the date-t government consists of an investment policy $i^t(k_t)$ and consumption policies $c_y^t(k_t)$ and $c_o^t(k_t)$ that are only functions of the payoff-relevant state variable k_t. A sequence of Markov strategies $\left\{f_t\left(i^t(k_t), c_y^t(k_t), c_o^t(k_t)\right)\right\}_{t=0}^{\infty}$ is a symmetric Markov perfect equilibrium if it is a subgame perfect equilibrium for every realization of the state variable k_t, and all governments follow the same strategy, that is, if $f_t\left(i^t(k_t), c_y^t(k_t), c_o^t(k_t)\right) = f\left(i(k_t), c_y(k_t), c_o(k_t)\right)$, for all t. We will restrict attention to symmetric Markov perfect equilibria in linear strategies.

3 The case of non-altruistic generations

It will be useful to consider first the special case where individuals do not care about the well-being of future generations. Accordingly, suppose that $\delta = 0$, with $A < \infty$. Letting $\tau_t = c_t^y / c_t$, with $c_t = c_t^y + c_t^o$, the date-t government's preferences are given by

$$v_t = u\left((1-\tau_t)\,c_t\right) + a\left[u\left(\tau_t c_t\right) + u\left((1-\tau_{t+1})\,c_{t+1}\right)\right],$$

with $c_t = Ak_t - (k_{t+1} - k_t)$, for all $t \geq 0$. The optimal intergenerational allocation of consumption at date t solves the static problem

$$\max_{\tau_t} \left\{au\left(\tau_t c_t\right) + u\left((1-\tau_t)\,c_t\right)\right\},$$

and so the young's share of aggregate consumption is given by

$$\tau^* = \frac{1}{1 + a^{-1/\sigma}}, \tag{5}$$

for all t. The symmetric Markov equilibrium in linear strategies is easy to construct. If the date-t government takes as given that $i_{t+1} = \widehat{g} k_{t+1}$, the optimal investment decision at date t solves the problem

$$\max_{i_t} \alpha\left(\tau^*\right) u\left[Ak_t - i_t\right] + u\left[\left(A - \widehat{g}\right)\left(k_t + i_t\right)\right],$$

where

$$\alpha(\tau) = \frac{1}{a} + \left(\frac{\tau}{1-\tau}\right)^{1-\sigma}.$$

One can verify that the best investment response to the anticipation of $i_{t+1} = \widehat{g} k_{t+1}$ is given by $i_t = g k_t$, where

$$1 + g = \frac{A + 1}{1 + \left(\alpha\left(\tau^*\right) / \left(A - \widehat{g}\right)^{1-\sigma}\right)^{1/\sigma}}. \tag{6}$$

The unique pair (τ^*, g^*) such that τ^* satisfies (5) and g^* is a fixed point of the mapping in (6) characterizes a symmetric equilibrium, where g^* solves

$$\alpha\left(\tau^*\right)\left(1 + g^*\right)^{\sigma} = \left(A - g^*\right). \tag{7}$$

Clearly, there is an equilibrium with $g^* \in (-1, A)$. Moreover, $g^* > 0$ if and only if $a^{-1}\left(1 + a^{1/\sigma}\right) < A$.

Now, suppose that the date-t government can precommit future transfers, provided that it treats current and future generations symmetrically, by making (proportional) transfers identical at all points in time. For instance, this may be the case if the date-t government can legislate a transfer τ at date t that is sufficiently costly to change in the future (Boadway and Wildasin 1989), or if the stationary transfer τ is supported by the threat of collapse of the system if any government repeals the legislation (Cooley and Soares 1999).[5]

Furthermore, suppose that the date-t government can control *current,* but not future, investment. Thus, once transfers are legislated, consecutive governments will choose investment unilaterally, taking into account the investment strategies of future governments. Our previous analysis then implies that equilibrium investment is given by

$$\alpha(\tau)(1 + g)^{\sigma} = (A - g),$$

and, letting $g(\tau)$ be a solution to this equation, that the date-t government's optimal choice of τ solves the following problem:

$$\max_{\tau \in [0,1]} \left\{ \begin{array}{c} u\left((1-\tau)\left(A - g(\tau)\right) k_t\right) \\ +a\left[u\left(\tau\left(A - g(\tau)\right) k_t\right) + u\left((1-\tau)\left(A - g(\tau)\right)\left(1 + g(\tau)\right) k_t\right)\right] \end{array} \right\}. \tag{8}$$

[5] Azariadis and Galasso (2002) argue that giving current voters or policy makers some veto power over changes in future policies acts like a commitment device.

Let $\overline{\tau}$ denote a solution to this problem. We have the following result.

Proposition 1 *Suppose that $\delta = 0$ and $\sigma > 1$. The date-t government legislates old-age transfers ($\overline{\tau} < \tau^*$), even though the legislation will depress equilibrium growth ($g(\overline{\tau}) < g^*$).*

Proposition 1 rests on the fact that consecutive governments have very different preferences. In particular, the date-t government does not care about the generation born in period $t+1$, while the date-$t+1$ government does. This introduces a *future bias.* The date-t government would prefer that the date-$t+1$ government consume immediately, saving nothing, whereas when the time comes, the date-$t+1$ government will rather postpone consumption.

Intuitively, next-period old-age transfers are systematically too low from the perspective of current policy makers. If they cannot influence those future transfers, then they will have an incentive to influence future income instead. If $\sigma > 1$, current governments have an incentive to increase current growth in order to raise the young's future consumption, as income effects dominate substitution effects. Since current and future investments are strategic complements when $\sigma > 1$ (see (6)), the incentive for every policy maker to increase growth is self-enforcing. As a result, the economy's growth rate is too high from the viewpoint of the current government. Hence, if a pay-as-you-go system of intergenerational transfers could be legislated in order to redistribute from the young to the old every period, the current government would choose to do so. In turn the strategic value of raising growth would be reduced and the new growth rate would be lower than before the legislation.

Although the equilibrium before the new legislation is not Pareto efficient, it is not the case that equilibrium investment is dynamically inefficient to begin with. As usual, we say that an investment allocation is dynamically efficient if there is no alternative allocation that provides more aggregate consumption in one period and at least the same consumption in every other period. It is not difficult to verify that investment in the Markov equilibrium is dynamically efficient. To see why, note that this is the case if the growth rate is lower than the *social* return to investment, that is, if $g^* < A$.[6] That this condition must hold follows immediately from the aggregate resources constraint: $Ak_t \geq c_t + k_{t+1} - k_t$.

[6] The following proof replicates the argument in Saint Paul (1992). Consider an allocation $\left\{\widetilde{k}_t\right\}$ with $\widetilde{k}_s < k_s$, for some s, with $\widetilde{c}_t \geq c_t$ for $t \geq s$. Since $\widetilde{k}_{t+1} = (A+1)\widetilde{k}_t - \widetilde{c}_t$ and $k_{t+1} = (A+1)k_t - c_t$, for

The legislation that is optimal from the viewpoint of the government that introduces it also hurts an infinite number of future generations. Yet, if a future government unexpectedly had the chance to legislate stationary intergenerational redistribution, it would choose to sustain the current legislation. This is because the reduction in wealth caused by the original legislation is due to lower investments that will be sunk from the viewpoint of the future government.

At first pass, Proposition 1 may seem to depend on the fact that individuals are not altruistic. If only they cared about the well-being of future generations, one may expect that they would not ever have an incentive to sacrifice growth. However, we show below that this is not the case.

4 The case of altruistic generations

Now suppose that $\delta \in (0,1)$. In this case, the utility of individuals born in period t, as given by (1), can be re-written as

$$u_t = u\left(c_t^y\right) + \sum_{s=1}^{\infty} \delta^s \left(u\left(c_{t+s}^y\right) + \delta^{-1} u\left(c_{t+s}^o\right)\right).$$

The difference between individuals born at date $t-1$ and those born at date t is that the latter do not care about the former. Otherwise, their preferences are time consistent: the trade-off between dates t and $t+1$ is perceived the same way by all individuals at date $t-1$ and at date t. Accordingly, the date-t government's preferences are given by

$$v_t = u\left(c_{t-1}^y\right) + u\left(c_t^o\right) + (\delta + a) u\left(c_t^y\right) + (\delta + a) \sum_{s=1}^{\infty} \delta^s \left(u\left(c_{t+s}^y\right) + \delta^{-1} u\left(c_{t+s}^o\right)\right). \tag{9}$$

Consider the right side of equation (9). The government at date t must treat the old individuals' felicity at date $t-1$ (first term) as sunk. The felicity from consumption of old

$t \geq s$, it must be that $k_{t+1} - \widetilde{k}_{t+1} \geq (A+1)\left(k_t - \widetilde{k}_t\right)$, for $t \geq s$. In turn this implies that $k_{s+T} - \widetilde{k}_{s+T} \geq (A+1)^T \left(k_s - \widetilde{k}_s\right)$, and thus $\widetilde{k}_{s+T} \leq (1+g^*)^T k_s - (A+1)^T \left(k_s - \widetilde{k}_s\right)$, for any $T \geq 1$. Clearly, if $g^* < A$, the right side of the inequality becomes negative for T sufficiently large, contradicting the hypothesis that there is a feasible deviation $\widetilde{k}_s < k_s$, for some s, with $\widetilde{c}_t \geq c_t$ for $t \geq s$. This concludes the proof.

individuals at date t (second term) enters with weight 1, the weight at which the government values the old individuals' utility. However, the felicity of young individuals at date t (third term) enters not only with the direct weight on their young individuals' utility, given by a, but also indirectly because the old care about the young (with weight δ) and the government cares about the old (with weight 1). The last term reflects that both young and old individuals care about future consumption through altruism, and the young care about their own future old-age consumption.

Inspection of equation (9) indicates that the date-t government's preferences are time inconsistent as long as $a > 0$, that is, as long as the date-t government's preferences put any weight on the utility of currently young individuals. The weight of old-age consumption at t relative to old-age consumption at $t+1$ equals $1/(\delta + a)$, and this is smaller than the weight on the felicities of old-age consumption at $t+s$ relative to old-age consumption at $t+1+s$, which equals $1/\delta$. This stems from the fact that the currently old care about the future consumption of the currently young, but the young do not care about the currently old. However, both the currently young and currently old care about the consumption of all future generations. As a result, consumption of the currently old gets a relatively small weight.

It should be noted that our assumption that the current young do not place any weight at all on the current old is made for simplicity. The essential feature is that the current young do not place *sufficient* weight on the current old.[7]

4.1 Benchmark commitment solution

In order to understand the impact of the time inconsistency of social preferences on equilibrium behavior, it is useful to consider first a benchmark problem for an arbitrary government

[7] The time inconsistency of the date-t government's preferences would still arise in the presence of two-sided altruism. Suppose that $u_t = u(c_t^y) + u(c_{t+1}^o) + \delta_F u_{t+1} + \delta_B u_{t-1} = \tilde{u}_t + \delta_F u_{t+1} + \delta_B u_{t-1}$. Kimball (1987) shows that this generates, under certain conditions, the following time-consistent individual preferences: $u_t = \sum_{b=1}^{\infty} (\lambda_B)^{-b} \tilde{u}_{t-b} + \sum_{f=0}^{\infty} (\lambda_F)^f \tilde{u}_{t+f}$, where λ_B and λ_F are functions of δ_B and δ_F. Following the same procedure as above, one can verify that time inconsistency of the date-t government's preferences arises if and only if $\lambda_B \neq \lambda_F$ and that the currently old get too small a weight whenever $\lambda_B > \lambda_F$, which is the natural case. This formulation would generate the same conclusions as our simpler case of one-sided altruism.

under the assumption that it can control future allocations. The goal of this section is to analyze the solution to this benchmark problem. The nature of this first-best solution is clarified by formulating the date-t government problem recursively. We will simplify notation by avoiding time subscripts and using primes to denote next-period values whenever possible.

First, consider the static intergenerational allocation of consumption every period from the viewpoint of the date-t government. At date t, only the second and third term in the social preferences given by equation (9) are relevant; therefore, the optimal intergenerational allocation of consumption solves the static problem

$$\max_{c^y, c^o} \left\{ u\left(c^y\right) + (\delta + a)^{-1} u\left(c^o\right) \right\} \text{ subject to } c^y + c^o \leq c, \text{ with } c^y, c^o \geq 0, \tag{10}$$

and so the young's share of aggregate consumption is given by

$$\tau_c \equiv \frac{c^y}{c} = 1 - \frac{c^o}{c} = \frac{1}{1 + (\delta + a)^{-1/\sigma}}, \tag{11}$$

where $c = c^y + c^o$.

By contrast, from date $t+1$ onwards, the date-t government would choose intergenerational consumption allocations differently than future governments would actually do. Instead, the date-t government's optimal allocation would solve the static problem

$$\max_{c^y, c^o} \left\{ u\left(c^y\right) + \delta^{-1} u\left(c^o\right) \right\} \text{ subject to } c^y + c^o \leq c, \text{ with } c^y, c^o \geq 0. \tag{12}$$

Accordingly, the young's share of aggregate consumption at every future date would be

$$\overline{\tau}_c \equiv \frac{c^y}{c} = 1 - \frac{c^o}{c} = \frac{1}{1 + \delta^{-1/\sigma}}, \tag{13}$$

where $c = c^y + c^o$. It is easy to see that $\tau_c > \overline{\tau}_c$. The date-$t$ government prefers to allocate a larger share of aggregate consumption to the current young than the share he would like to allocate to the young in every future period.

Now consider the date-t government's investment problem. Taking into account the optimal intergenerational allocations of aggregate consumption every period, by substituting the consumption shares given by equations (11) and (13) into the social preferences given by equation (9), we express the relevant preferences for the date-t government in terms of aggregate consumption levels as

$$\widetilde{v}_t = q(\tau_c, a)\, u(c_t) + q(\overline{\tau}_c, 0) \sum_{s=1}^{\infty} \delta^s u(c_{t+s}), \tag{14}$$

where

$$q(\tau, a) = \tau^{1-\sigma} + (\delta + a)^{-1} (1-\tau)^{1-\sigma}. \tag{15}$$

Note that $\widetilde{v}_t$ is simply a positive linear transformation of the preferences for the date-t government given in equation (9). The representation of social welfare in equation (14) reveals the key to understanding the governments' problem. It shows that social preferences are time inconsistent whenever $q(\tau_c, a) \neq q(\overline{\tau}_c, 0)$, which in turn is the case if and only if $a > 0$. Even though all individuals have standard time-consistent, geometric preferences, governments have time-inconsistent preferences over aggregate consumption streams. The interesting question concerns the nature of the biases this introduces in policy making.

Definition 1 *Let* $w_t = u(c_t) + \sum_{s=1}^{\infty} d(t+s)\, u(c_{t+s})$. *(i)* w_t *is present biased it for any* $s \geq 1$ *and* $k \geq 1$*, there exist* c *and* c' *such that* $u(c) > d(t+k)\, u(c')$ *and* $d(t+s)\, u(c) < d(t+s+k)\, u(c')$. *(ii)* w_t *is future biased it for any* $s \geq 1$ *and* $k \geq 1$*, there exist* c *and* c' *such that* $u(c) < d(t+k)\, u(c')$ *and* $d(t+s)\, u(c) > d(t+s+k)\, u(c')$.

This definition emphasizes the possibility of preference reversals. When preferences are present biased, it is possible that as of date t, date-$t+2$ consumption is preferred to date-$t+1$ consumption, but the reverse is true when date $t+1$ arrives. When preferences are future biased, it is possible that as of date t, date-$t+1$ consumption is preferred to date-$t+2$ consumption, but the reverse is true when date $t+1$ arrives.

In our context, discounting is given by

$$d(t+s) = \beta\delta^s, \text{ with } \beta \equiv \frac{q(\overline{\tau}_c, 0)}{q(\tau_c, a)}$$

for all $s \geq 1$, in which case each government discounts consumption streams starting then according to the sequence 1, $\beta\delta$, $\beta\delta^2$, $\beta\delta^3$, ... and so the discount factor between the current period and the next is $\beta\delta$ whereas the discount factor between any two future periods is δ. Accordingly, every government is subject to preference reversals whenever $\beta \neq 1$, because then the marginal rate of substitution between consecutive consumption levels changes with the planning date. In particular, note that

$$\frac{\beta\delta^2 u'(c_{t+2})}{\beta\delta u'(c_{t+1})} > \frac{\beta\delta u'(c_{t+2})}{u'(c_{t+1})} \text{ if and only if } \beta < 1,$$

where the left side of the first inequality is the marginal rate of substitution between date-$t+1$ consumption and date-$t+2$ consumption from the viewpoint of date t and the right side is the same marginal rate of substitution from the viewpoint of date $t+1$. Present bias corresponds to falling marginal rates of substitution as the planning date evolves, whereas future bias corresponds to rising marginal rates of substitution. Intuitively, future-biased governments are overly eager to postpone consumption whereas present biased governments are overly eager to postpone saving.

Thus, the social preferences given in (14) are present biased if and only if $\beta < 1$, in which case governments would have quasi-hyperbolic preferences over aggregate consumption streams of the form used by Phelps and Pollak (1968), Laibson (1997) and Krusell, Kuruşçu and Smith (2002). In our setting, however, it is always the case that $q(\tau_c, a) < q(\overline{\tau}_c, 0)$ and thus $\beta > 1$.

Proposition 2 *The preferences of the date-t government are future biased, for all t.*

Thus, social preferences are necessarily future biased and so governments are overly eager to postpone current consumption. This result is in sharp contrast with the common view that a present bias is inherent to all utilitarian aggregation of heterogeneous time preferences (Jackson and Yariv, 2014). In our model, it is the overlapping generations demographic structure and the fact that current generations care insufficiently about previous generations that imply that governments' preferences, which aggregate the preferences of the generations currently alive, will be future-biased. This is the source of our results below.

From date $t+1$ onwards, the date-t government has time-consistent preferences, and so it would solve the following problem:

$$W(k) = \max_{0 \leq k' \leq Ak} \left\{ q(\overline{\tau}_c, 0)\, u(Ak - k' + k) + \delta W(k') \right\}, \tag{16}$$

where $q(\tau, a)$ is given by equation (15) and $\overline{\tau}_c$ is given by equation (13). It is easy to verify that the corresponding first-order condition equates the marginal disutility from additional investment and the marginal value of additional capital next period:

$$-q(\overline{\tau}_c, 0) \frac{\partial u(c)}{\partial c} \frac{\partial c}{\partial k'} = \delta \frac{\partial W(k')}{\partial k'}. \tag{17}$$

Furthermore, since preferences are time consistent from date $t+1$ onwards, the solution to the above problem satisfies the familiar envelope condition

$$\frac{\partial W(k)}{\partial k} = q(\overline{\tau}_c, 0) \frac{\partial u(c)}{\partial c} \frac{\partial c}{\partial k} \tag{18}$$

every period. Combining equation (18), evaluated one period ahead, and equation (17), it is easy to see that the intertemporal allocation of aggregate consumption satisfies the familiar Euler equation

$$\frac{\partial u(c) / \partial c}{\delta \left(\partial u(c') / \partial c' \right)} = \frac{-\partial c' / \partial k'}{\partial c / \partial k'},$$

which equates the marginal rate of substitution between current and next-period consumption and the corresponding marginal rate of transformation. Taking derivatives and noting that consumption and capital grow at the common rate $\overline{g}_c$, it can be verified that the solution to the above standard dynamic programming problem implies that investment from date $t+1$ onwards is given by $k' - k = \overline{g}_c k$, where

$$1 + \overline{g}_c = \left[\delta (A + 1) \right]^{1/\sigma}. \tag{19}$$

Finally, the investment problem at date t can be formulated as:

$$W_0(k) = \max_{0 \le k' \le Ak} \left\{ q(\tau_c, a)\, u(Ak - k' + k) + \delta W(k') \right\}, \tag{20}$$

where $q(\tau, a)$ is given by equation (15) and τ_c is given by equation (11). Once again, the corresponding first-order condition equates the marginal disutility incurred from additional investment and the marginal value of additional capital next period:

$$-q(\tau_c, a) \frac{\partial u(c)}{\partial c} \frac{\partial c}{\partial k'} = \delta \frac{\partial W(k')}{\partial k'}. \tag{21}$$

Noting that the value of future capital is derived from the appropriately weighted sum of utilities from the future consumption stream it generates, that is,

$$\delta W(k_{t+1}) = \sum_{s=1}^{\infty} \delta^s q(\overline{\tau}_c, 0)\, u(c_{t+s}),$$

and noting that consumption from date $t+1$ onwards grows at the constant growth rate $\overline{g}_c$, as given by equation (19), one can verify that the value of future capital is such that

$$W(k') = \left(\frac{q(\overline{\tau}_c, 0)}{1 - \delta (1 + \overline{g}_c)^{1-\sigma}} \right) u(c'),$$

therefore, the marginal value of additional capital next period is given by

$$\frac{\partial W(k')}{\partial k'} = \left(\frac{q(\overline{\tau}_c, 0)}{1 - \delta (1 + \overline{g}_c)^{1-\sigma}} \right) \frac{\partial u(c')}{\partial c'} \frac{\partial c'}{\partial k'}. \tag{22}$$

Combining equations (21) and (22), it is easy to see that time inconsistency influences the date-t government's allocation of aggregate consumption at date t by introducing a wedge between the marginal rate of substitution between current and next-period consumption and the corresponding marginal rate of transformation:

$$\frac{\partial u(c) / \partial c}{\delta (\partial u(c') / \partial c')} = \left(\frac{q(\overline{\tau}_c, 0) / q(\tau_c, a)}{1 - \delta (1 + \overline{g}_c)^{1-\sigma}} \right) \frac{-\partial c' / \partial k'}{\partial c / \partial k'}. \tag{23}$$

The magnitude of the wedge takes into account that $c_t^y = \tau_c c$ and $c_{t+1}^y = \overline{\tau}_c c'$, through the term $q(\overline{\tau}_c, 0) / q(\tau_c, a) = (\overline{\tau}_c / \tau_c)^{-\sigma}$, and applies an effective discount rate equal to

$\delta\left(1+\overline{g}_c\right)^{1-\sigma}$, because the young's share of aggregate consumption in all future periods is equal to $\overline{\tau}_c$ rather than τ_c. To interpret the wedge, note that $\delta q\left(\overline{\tau}_c,0\right)/q\left(\tau_c,a\right)$ specifies the relative weight placed on $u\left(c_{t+1}\right)$ rather than $u\left(c_t\right)$ by the social welfare function in equation (14).

Using the facts that felicity functions (given in equation (2)) are isoelastic, the aggregate resources constraint (given in equation (3)) holds with equality, and investment from date $t+1$ onwards is given by $i\left(k\right)=\overline{g}_c k$, to evaluate equation (23), one can verify that the solution to problem (20) implies that investment at date t is given by $i\left(k\right)=g_c k$, with

$$1+g_c=\frac{A+1}{1+\left(\frac{q(\tau_c,a)}{q(\overline{\tau}_c,0)}\frac{\delta^{-1}-(1+\overline{g}_c)^{1-\sigma}}{(A-\overline{g}_c)^{1-\sigma}}\right)^{1/\sigma}}, \tag{24}$$

where $q\left(\tau,a\right)$ is given by equation (15) and τ_c, $\overline{\tau}_c$ and $\overline{g}_c$ are given by equations (11), (13), and (19), respectively.

In the Appendix we show that $g_c>\overline{g}_c$, for all $\sigma>0$. If the date-t government could commit future allocations, it would choose a current growth rate that is larger than the growth rate it would dictate to future generations. This is because the current government cares more about the future old than it does about the current old generation, and so $\frac{q(\tau_c,a)}{q(\overline{\tau}_c,0)}<1$, for $a>0$. In turn, this occurs because the current government puts positive weight on the current young, but the current young does not care about the current old. Indeed, it can be verified that the right side of equation (24) is equal to $1+\overline{g}_c$ if and only if $a=0$ (see Appendix). It is also easy to verify that the solution to the above problem is continuous at $\delta=0$.

The following proposition summarizes our previous discussion.

Proposition 3 *If the date-t government could precommit future allocations, optimal allocations would be given by* $i(k)=gk$, $c_y(k)=\tau\left(A-g\right)k$, *and* $c_o(k)=\left(1-\tau\right)\left(A-g\right)k$, *with*

$$(\tau,g)=\begin{cases}(\tau_c,g_c) & \textit{in the first period}\\ (\overline{\tau}_c,\overline{g}_c) & \textit{in every future period,}\end{cases}$$

where τ_c *and* $\overline{\tau}_c$ *are given by equations (11) and (13), respectively, with* $\tau_c > \overline{\tau}_c$*; while* g_c *and* $\overline{g}_c$ *are given by equations (24) and (19), respectively, with* $g_c > \overline{g}_c$*.*

Of course, the problem with the above solution is that it is time inconsistent. Accordingly, each government needs to take into account that future governments will deviate from the allocation that the current government would dictate if it could control future allocations. In the following section we consider equilibrium behavior when current governments recognize that future allocations will be chosen optimally by future governments.

4.2 Markov perfect equilibrium

In this section we characterize the unique symmetric Markov perfect equilibrium in linear strategies. Now each government recognizes that every future government will choose the same optimal intergenerational allocation of consumption each period as the one chosen in the current period by the current government. This is the allocation that solves the static problem given by equation (10) and so the young's share of aggregate consumption is now given by

$$\tau^* \equiv \frac{c^y}{c} = 1 - \frac{c^o}{c} = \frac{1}{1 + (\delta + a)^{-1/\sigma}}. \tag{25}$$

every period. Of course, τ^* is equal to the young's share of aggregate consumption that the date-t government would have chosen at date t even if it could control future allocations. That is, we have $\tau^* = \tau_c$ every period.

Suppose that the current government anticipates that every future government follows the linear investment policy $i' = \widehat{g}k'$, with $\delta\,(1 + \widehat{g})^{1-\sigma} < 1$. Then, the current investment decision solves the following problem:

$$V_0\,(k) = \max_{0 \le k' \le Ak} \{q\,(\tau^*, a)\,u\,(Ak - k' + k) + \delta V\,(k')\}\,, \tag{26}$$

with

$$V\,(k) = q\,(\tau^*, 0)\,u\,(Ak - (1 + \widehat{g})\,k + k) + \delta V\,((1 + \widehat{g})\,k)\,, \tag{27}$$

where $q(\tau, a)$ is given by equation (15), and τ^* is given by equation (25). An investment policy $i(k) = gk$ that is part of a symmetric Markov perfect equilibrium must have $g = \widehat{g}$.

In order to appreciate the role of commitment problems, first note that the first-order condition with respect to k' at date t is given by

$$-q(\tau^*, a) \frac{\partial u(c)}{\partial c} \frac{\partial c}{\partial k'} = \delta \frac{\partial V(k')}{\partial k'}, \tag{28}$$

which equates the marginal disutility incurred from additional investment and the marginal value of additional capital next period. Solving the recursion in equation (27) it can be verified that

$$V(k') = \left(\frac{q(\tau^*, 0)}{1 - \delta (1 + \widehat{g})^{1-\sigma}} \right) u(c')$$

and so we have

$$\frac{\partial V(k')}{\partial k'} = \left(\frac{q(\tau^*, 0)}{1 - \delta (1 + \widehat{g})^{1-\sigma}} \right) \frac{\partial u(c')}{\partial c'} \frac{\partial c'}{\partial k'}. \tag{29}$$

The first-order condition at date t if the date-t government could control future allocations would be the same as equation (28), except that V is replaced by W. The main difference lies in the marginal effect of current investment on the value function next period. The difference can be understood as follows. Each government recognizes that a marginal increase in current investment results in extra income next period that will in turn influence investment next period. Since current and future governments disagree about future investment decisions, current governments have an incentive to manipulate future investment decisions via current investment.

By contrast, if the current government could control future allocations, time consistency of the date-t government's preferences from date $t+1$ onwards would ensure that the familiar envelope condition holds, which ensures that the above effect of current on future investment can be ignored when making current investment decisions. In turn, this guarantees that investment in every future period will be given by $\overline{g}_c k$. The difference between equations (29) and (22) lies in that, in the Markov perfect equilibrium, the current government anticipates intergenerational disagreement in every future period.

Let us return to the characterization of the Markov perfect equilibrium. Combining equations (28) and (29) we have

$$\frac{\partial u(c)/\partial c}{\delta(\partial u(c')/\partial c')} = \left(\frac{q(\tau^*,0)/q(\tau^*,a)}{1-\delta(1+\widehat{g})^{1-\sigma}}\right)\frac{-\partial c'/\partial k'}{\partial c/\partial k'}. \tag{30}$$

Although the intertemporal allocation of aggregate consumption in the Markov perfect equilibrium is time consistent by construction, there is a wedge between the marginal rate of substitution between current and next-period consumption and the corresponding marginal rate of transformation. The magnitude of the wedge takes into account that the young's share of aggregate consumption every period is equal to $\tau^* = \tau_c > \overline{\tau}_c$ rather than $\overline{\tau}_c$, and also anticipates that investment in all future periods is given by $i(k') = \widehat{g}k'$.

Recognizing that

$$\frac{\partial c'}{\partial k'} = A - \frac{\partial i(k')}{\partial k'} = A - \widehat{g},$$

since $c' = Ak' - i(k')$ and $i(k') = \widehat{g}k'$, and using the facts that felicity functions (given in equation (2)) are isoelastic and the aggregate resources constraint (given in equation (3)) holds with equality, it is now straightforward to write the above Euler equation as

$$\left(\frac{k'}{Ak-k'+k}\right)^{\sigma} = \frac{q(\tau^*,0)(A-\widehat{g})^{1-\sigma}}{q(\tau^*,a)\left(\delta^{-1}-(1+\widehat{g})^{1-\sigma}\right)},$$

which describes the best response k' to the anticipation of $\widehat{g}$, for given k. Clearly, the best response to any given $\widehat{g}$ is linear in k. Consequently, we obtain the best-response mapping

$$1+g = \frac{A+1}{1+\left(\frac{q(\tau^*,a)}{q(\tau^*,0)}\frac{\delta^{-1}-(1+\widehat{g})^{1-\sigma}}{(A-\widehat{g})^{1-\sigma}}\right)^{1/\sigma}} \equiv 1+B(\tau^*,\tau^*,\widehat{g}). \tag{31}$$

The best response function $g = B(\tau,\tau',\widehat{g})$ characterizes the best investment response by a government that allocates a share τ of current consumption to the current young and anticipates that future governments will allocate a share τ' of consumption to the young and invest according to $i(k') = \widehat{g}k'$. Note that the structure of the best response function in equation (31) is identical to that in equation (24). The above commitment solution has $g_c = B(\tau_c, \overline{\tau}_c, \overline{g}_c)$, whereas the Markov equilibrium has $g^* = B(\tau^*, \tau^*, g^*)$.

Proposition 4 *(i) There is a unique symmetric, interior, Markov perfect equilibrium in linear strategies. The equilibrium is characterized by $i(k) = g^* k$, $c_y(k) = \tau^*(A - g^*)k$, and $c_o(k) = (1 - \tau^*)(A - g^*)k$, where τ^* is given by equation (25); $g^* = B(\tau^*, \tau^*, g^*) \in (\overline{g}_c, A)$, where B is given by equation (31) and $\overline{g}_c$ is given by equation (19). (ii) For all $\widehat{g} \in (\overline{g}_c, A)$, $\partial B(\tau^*, \tau^*, \widehat{g}) / \partial \widehat{g} \geq 0$ if and only if $\sigma \geq 1$, with equality if and only if $\sigma = 1$.*

Part (i) characterizes the unique symmetric, interior, Markov perfect equilibrium in linear strategies. Part (ii) provides additional insight into the role of commitment problems. Note that the disagreement between governments about investment decisions takes the particular form that the date-$(t+1)$ government invests too much from the viewpoint of the date-t government. The best response mapping (31) indicates how each government will attempt to manipulate investment next period. Part (ii) of the proposition implies that locally around the equilibrium current and next-period investments are "strategic" complements if $\sigma > 1$ and "strategic" substitutes if $\sigma < 1$. The panels in Figure 1 plot the different types of best responses.

[FIGURE 1]

Panel (1) shows that $B(\tau^*, \tau^*, \widehat{g})$ decreases at first, reaching a minimum at $\overline{g}_c$, and then increases, when $\sigma > 1$. Panel (2) shows that the best response is flat when $\sigma = 1$. In this case, $g^* = B(\tau^*, \tau^*, g^*)$ has a closed-form solution and the equilibrium growth rate is given by

$$1 + g^* = \frac{A + 1}{1 + \left(\frac{\delta + a + 1}{\delta + a}\right)\left(\frac{1 - \delta}{1 + \delta}\right)}. \tag{32}$$

Panel (3) in the above figure illustrates that $B(\tau^*, \tau^*, \widehat{g})$ increases at first, peaking at $\overline{g}_c$, and then decreases, when $\sigma < 1$.

The role of the elasticity of intertemporal substitution, given by $1/\sigma$, is worth noting. With respect to a generation's lifetime, higher values of σ indicate greater aversion to differences in consumption over the life cycle. However, since individuals are altruistic, higher values of σ also indicate greater aversion to unequal consumption across generations. With balanced growth, the higher the value of σ, the less individuals are willing to tolerate larger positive, or smaller negative, growth rates. Indeed, it is not difficult to show that the equilibrium growth rate decreases as individuals are less willing to substitute consumption intertemporally.

4.3 A positive theory of intergenerational redistribution

In this section we consider the analogue of Proposition 1 in the case where individuals are altruistic. As before, suppose that the date-t government can legislate (symmetric/stationary) intergenerational redistribution, as given by a transfer τ and suppose that changing the legislation is sufficiently costly, so the introducing government will in effect precommit future transfers, before equilibrium investments are determined. Furthermore, suppose that the date-t government can control *current,* but not future, investment. Thus, once transfers are legislated, consecutive governments will choose investment unilaterally, taking into account the investment strategies of future governments.

Conditional on the transfer τ, our previous analysis implies that there is a symmetric Markov perfect equilibrium in linear strategies such that the equilibrium growth rate is given by

$$\frac{q(\tau,a)}{q(\tau,0)} = \frac{\delta(1+g)^{-\sigma}}{1-\delta(1+g)^{1-\sigma}}(A-g).$$

Letting $g(\tau)$ be a solution to this equation, one can easily verify that our previous analysis also implies that the date-t government's optimal choice of τ solves the following problem:

$$\max_{\tau\in[0,1]} \left\{ q(\tau,a)\, u((A-g(\tau))k) + \frac{\delta q(\tau,0)}{1-\delta(1+g(\tau))^{1-\sigma}} u((A-g(\tau))(1+g(\tau))k) \right\}. \quad (33)$$

Let $\overline{\tau}$ denote a solution to this problem. We have the following analogue of Proposition 1.

Proposition 5 *Suppose that $\delta \in (0,1)$ and $\sigma > 1$. The date-t government legislates old-age transfers ($\overline{\tau} < \tau^*$), even though the legislation will depress equilibrium growth ($g(\overline{\tau}) < g^*$) and so it will hurt an infinite number of future generations.*

Proposition 5 provides a positive theory of intergenerational redistribution. It may seem at first counterintuitive that individuals who are perfectly altruistic towards future generations would support institutions that will necessarily harm future generations. The main insight here is that they do so because future old-age transfers are too low and future growth is too high *from the perspective of currently living generations*, not because equilibrium investment is dynamically inefficient. The same arguments we used in the case of non-altruistic generations continue to apply here, implying that equilibrium investment is dynamically efficient when generations are altruistic.

Moreover, it should be noted that the Markov perfect equilibrium, both before and after the legislation, is not Pareto efficient, because the private and the social return to investment are different. A Pareto improvement would result from investing optimally from the viewpoint of the currently young generation at the socially optimal rate of return, without changing the allocation for any other generation. This is in contrast with the common perception that perfect altruism towards the following generation must lead to Pareto efficiency (Streufert, 1993). This is the case in the non-overlapping generations models studied in the literature, because perfect altruism then amounts to time-consistent preferences. However, with time-inconsistent social preferences, as is the case here, the private return to investment is necessarily lower than the social return, because the incentive to manipulate future investment does not disappear.

Proposition 5 rests on the fact that the aggregation of the preferences of the young and the old implies that every government has future biased preferences. Consequently, future old-age transfers are too low from the perspective of the current government (and that of both living generations). Accordingly, the current government has an incentive to legislate intergenerational redistribution to increase future old-age transfers at the expense of growth. When $\sigma > 1$, current and future investments are strategic complements and so the legislation ends up lowering investment and growth. This will obviously hurt future generations. Yet, if a future government unexpectedly had the chance to legislate stationary intergenerational redistribution, it would choose to sustain the current legislation, because the reduction in wealth caused by the original legislation is due to lower investments that will be sunk from the viewpoint of the future government.

Now consider the reason why future old-age transfers are too low from the perspective of the current government. First note that the relationship between transfers and growth is such that $g'(\tau) > 0$ if and only if $\sigma > 1$. Thus, when income effects dominate substitution effects, old-age transfers, given by $1 - \tau$, and growth are negatively related, as one may expect.

Next, compare the growth rate in the Markov equilibrium before the legislation of old-age transfers (see Proposition 4) and the growth rate that governments would choose if they could commit all future allocations (see Proposition 3). In particular, Proposition 6 shows that even if the current government could control all future allocations, instead of just being able to control old-age transfers, it would choose to lower growth below the equilibrium rate. This also shows that Proposition 5 does not rely on the fact that the government has limited instruments of intergenerational redistribution.

Proposition 6 *$g^* > g_c$ if and only if $\sigma > 1$, that is, whenever the elasticity of intertemporal substitution is less than one ($1/\sigma < 1$), the equilibrium growth rate is higher than the growth rate every generation would set if they were able to control future allocations.*

Proposition 6 is a striking result for two reasons. First, note that $g_c > \overline{g}_c$, where $\overline{g}_c$ is the first-best growth rate from the viewpoint of the old, and it is also the growth rate that every young generation, and every government, would dictate on every future generation, if they could do so. In this sense, commitment problems lead to equilibrium growth that is too high, relative to the preferences of all generations. Second, the private return to investment is lower than the social return to investment. The latter is given by the constant marginal product of capital A, whereas the former is given by $A - \partial(g^* k)/\partial k = A - g^*$.

In order to understand the source of the result stated in Proposition 6, it is useful to consider the relationship between the current investment decisions of a government with and without commitment. Note that the investment problem of the date-t government at date t, given by equation (26), can be written as

$$V_0(k) = \max_{0 \le k' \le Ak} \left\{ \widetilde{W}_0(k, k') - \delta \left(W(k') - V(k') \right) \right\} \tag{34}$$

where W and V are given by equation (16) and equation (27), respectively, and where $\widetilde{W}_0(k, k')$ is precisely the objective to be maximized at date t under the assumption that the date-t government can commit future allocations (see equation (20)), that is,

$$W_0(k) = \max_{0 \le k' \le Ak} \widetilde{W}_0(k, k') \,. \tag{35}$$

Clearly, it must be that $V_0(k) < W_0(k)$, since the commitment solution from date $t+1$ onwards is the date-t government's first-best solution (i.e., since $W(k') - V(k') > 0$). Thus, whenever the current government anticipates $(\tau^*, \widehat{g})$ to deviate from $(\overline{\tau}_c, \overline{g}_c)$ in the future, it anticipates a welfare loss. Accordingly it has an incentive to invest strategically to compensate for this loss. Formally, the marginal effect of additional current investment on the welfare loss associated with the difference between $\widehat{g}$ and $\overline{g}_c$ in the future is given by $\partial W(k')/\partial k' - \partial V(k')/\partial k'$, where $\partial W(k')/\partial k'$ is given by equation (22) and $\partial V(k')/\partial k'$ is given by equation (29). It is easy to verify that $\partial W(k')/\partial k' - \partial V(k')/\partial k' \le 0$ if and only if $B(\tau^*, \tau^*, \widehat{g}) \ge B(\tau_c, \overline{\tau}_c, \overline{g}_c)$, where recall that the best-response mapping $B(\tau^*, \tau^*, \widehat{g})$ is given by equation (31), with $B(\tau^*, \tau^*, g^*) = g^*$ and $B(\tau_c, \overline{\tau}_c, \overline{g}_c) = g_c$.

To understand the above "strategic-compensation effect", note that the anticipated discrepancy between $(\tau^*, \widehat{g})$ and $(\overline{\tau}_c, \overline{g}_c)$ gives rise to two opposing effects. The problem arises because future governments weigh future consumption too little relative to the current government, that is, $q(\tau^*, a) < q(\tau^*, 0)$. On the one hand, for given next-period consumption, next-period utility is anticipated to be lower because next-period's government will misallocate consumption over the two generations. The current government can compensate for this loss by strategically raising investment in order to increase next-period aggregate consumption. On the other hand, transferring wealth to the future has a lower return, because the increase in production is misallocated over the two coexisting generations: by strategically lowering investment the current government can substitute intertemporally away from misallocated future investment.

With log utility, the current government is in effect unable to use current investment strategically to its advantage, as the welfare loss from misallocation of future investment exactly offsets the welfare gain from additional future consumption. Consequently, the current government's best response to any future growth rate $\widehat{g} \geq \overline{g}_c$ is given by $i(k) = g_c k$ when $\sigma = 1$. Since every future government faces the same problem, each government will choose the growth rate that it would be chosen in the first period if future allocations could be controlled. Thus, the resulting equilibrium growth rate g^* must be equal to g_c. This outcome reflects the pro-growth bias inherent to the social preferences that aggregate the preferences of coexisting generations that disagree about current investment: the young would like faster current growth than the old. This is the "imperfect-altruism effect" that underlaid the relatively high *short-run* growth rate associated with the benchmark commitment solution. Now, however, this effect translates into higher *long-run* equilibrium growth, relative to the commitment solution.

If inequality aversion is large enough (i.e., if $\sigma > 1$), the welfare loss from misallocation of future investment cannot offset the welfare gain from additional future consumption and thus, each government has a strategic incentive to overinvest, relative to $i(k) = g_c k$. This explains the strategic complementarity between current and next-period investments (see Proposition 4), which leads to a long-run equilibrium growth rate g^* that is not only higher than $\overline{g}_c$, but also higher than g_c.

It is also worth noting the following limiting results.

Proposition 7 *Suppose that $\sigma > 1$. Then,*

$$\lim_{\delta \to 1} g^* > \overline{g}_c, \text{ for } a > 0\text{; and } \lim_{a \to 0} g^* = \overline{g}_c < A = \lim_{a \to \infty} g^*, \text{ for } 0 < \delta \leq 1.$$

Thus, even in the limit as the discount rate on future generations approaches zero, the equilibrium growth rate is strictly higher than $\overline{g}_c$. This is because as long as governments puts weight on the current young the time inconsistency of their preferences creates a nontrivial problem, which does not disappear as δ approaches 1. Furthermore, as the weight governments put on the current old becomes negligible, the equilibrium growth rate becomes arbitrarily close to A (for $\sigma > 1$), and so the savings rate approaches 1. In this sense, the equilibrium growth rate when future old-age transfers cannot be precommitted can be arbitrarily higher than the growth rate that is preferred by all generations, which explains why the current government legislates old-age transfers at the expense of growth even when parents value children's well-being exactly the same as their own.

5 Conclusion

We have argued that governments are future biased when they aggregate the preferences of overlapping generations. Future bias, which involves preference reversals favoring future over current consumption, explains why governments legislate and sustain old-age transfers at the expense of capital accumulation and growth, even if generations are altruistic.

Strotz's (1956) seminal work, and more recently Laibson's (1997) demonstrate the general relevance of economic agents' time inconsistency for the design of institutions that can cope with intertemporal disagreement by facilitating commitments. With respect to this, an implication of our analysis is that the availability of commitment mechanisms to cope with intergenerational disagreement can harm future generations.

In terms of concrete policies, we have offered a novel perspective on the widespread legislation of pay-as-you-go social security despite recognition of its negative effects on capital accumulation. Our analysis implies that social security legislation systematically favors current generations at the expense of future generations. Yet, future governments do not have an incentive to repeal the legislation, because its negative impact on their own wealth is due to lower investments that will be sunk from the viewpoint of every future government.

Appendix

Proof of Proposition 1

Assume that $\delta = 0$ and $\sigma > 1$. Let $U(\tau, g(\tau))$ denote the objective function in (8). Since $U(\tau, g(\tau))$ is a continuous function of τ on $[0,1]$, it must have a maximum. Moreover, U is differentiable on $(0,1)$ with

$$\frac{dU}{d\tau} = \frac{\partial U}{\partial \tau} + \frac{\partial U}{\partial g} g'(\tau).$$

Now we prove that $dU/d\tau < 0$ for all $\tau \geq \tau^*$. First, note that $\frac{\partial U}{\partial \tau} < 0$ for all $\tau \geq \tau^*$, because $\tau^{-\sigma} - \frac{1}{a}(1-\tau)^{-\sigma} \leq 0$ for all $\tau \geq \tau^*$ and

$$\frac{\partial U}{\partial \tau} = ((A-g)\, k_t)^{1-\sigma} a \left(\tau^{-\sigma} - \frac{1}{a}(1-\tau)^{-\sigma} - (1+g)^{1-\sigma}(1-\tau)^{-\sigma} \right).$$

Then, note that $\frac{\partial U(\tau,g)}{\partial g} < 0$, for $g = g(\tau)$ and for all $\tau \in [0,1]$, because

$$\begin{aligned} \frac{\partial U(\tau,g)}{\partial g} &= ((A-g)\, k_t)^{-\sigma} k_t a (1-\tau)^{1-\sigma} (1+g)^{-\sigma} \left[-\alpha(\tau)(1+g)^{\sigma} + A - g - (1+g) \right] \\ &= -((A-g)\, k_t)^{-\sigma} k_t a (1-\tau)^{1-\sigma} (1+g)^{1-\sigma}, \end{aligned}$$

where the second equality follows from the fact that $\alpha(\tau)(1+g(\tau))^{\sigma} = A - g(\tau)$. Differentiating this last equation with respect to τ and g, we have that $g'(\tau) > 0$ if and only if $\sigma > 1$. Hence, $dU/d\tau < 0$ for all $\tau \geq \tau^*$, which implies that $\overline{\tau} < \tau^*$.

From (8), it is easy to verify that $\lim_{\tau \to 0} U(\tau, g(\tau)) = -\infty$. Therefore, we have $\overline{\tau} \in (0, \tau^*)$, which implies $g(\overline{\tau}) \in (-1, g^*)$. This concludes the proof. **QED**

Proof of Proposition 2

It follows from the fact that $q(\overline{\tau}_c, 0) > q(\tau_c, a)$, which implies that one can always find c and c' such that

$$u(c) < \left(\frac{q(\overline{\tau}_c, 0)}{q(\tau_c, a)} \right) \delta^k u(c') \text{ and } u(c) > \delta^k u(c'),$$

for any $s \geq 1$ and $k \geq 1$. **QED**

Proof of Proposition 3

All statements in the proposition are proven in the main text, except for the inequality $g_c > \bar{g}_c$. To prove this, let

$$1 + \tilde{B}(\hat{g}, Q) \equiv \frac{A+1}{1 + Q\left(\frac{\delta^{-1} - (1+\widehat{g})^{1-\sigma}}{(A-\widehat{g})^{1-\sigma}}\right)^{1/\sigma}},$$

and inspect equation (24) to note that $\tilde{B}(\bar{g}_c, Q_c) = g_c$, where

$$Q_c \equiv \left(\frac{q(\tau_c, a)}{q(\overline{\tau}_c, 0)}\right)^{1/\sigma}.$$

Using equations (11) and (13), one can verify that $Q_c = \overline{\tau}_c / \tau_c < 1$. Moreover, one can easily verify that $\tilde{B}(\bar{g}_c, 1) = \bar{g}_c$. Since $\partial \tilde{B}(\hat{g}, Q)/\partial Q < 0$, it follows that $\tilde{B}(\bar{g}_c, Q_c) = g_c > \tilde{B}(\bar{g}_c, 1) = \bar{g}_c$, as required. **QED**

Proof of Proposition 4

Consider Part (ii) first. From equation (31), $B(\tau^*, \tau^*, \hat{g})$ can be rewritten as $\tilde{B}(\hat{g}, Q^*)$, with

$$Q^* \equiv \left(\frac{q(\tau^*, a)}{q(\tau^*, 0)}\right)^{1/\sigma},$$

where $\tilde{B}$ is defined in the proof of Proposition 3. Using equation (25), one can verify that

$$Q^* = \left(\frac{1}{1 + (1-\tau^*)a/\delta}\right)^{1/\sigma} < 1.$$

Next, note that the sign of $\partial \tilde{B}(\hat{g}, Q)/\partial \hat{g}$ is given by the sign of $(\sigma - 1)\left[(1+\widehat{g})^{\sigma} - \delta(A+1)\right]$. It is easy to verify that, for given Q^*, $\tilde{B}(\hat{g}, Q^*)$ has a global minimum at $\widehat{g} = \bar{g}_c$ if $\sigma > 1$; it has a global maximum at $\widehat{g} = \bar{g}_c$ if $\sigma < 1$; and it is flat at $\widehat{g} = g_c$ if $\sigma = 1$. This proves Part (ii) of the proposition.

Now consider Part (i). To prove existence of a unique fixed point $g^* \in (\overline{g}_c, A)$, evaluate $g = \tilde{B}(\hat{g}, Q)$ at $\hat{g} = g$ and rewrite it as

$$\delta^{-1} Q^{\sigma}(1+g)^{\sigma} + (1 - Q^{\sigma})(1+g) - (A+1) = 0. \tag{36}$$

As long as $Q \leq 1$, the left side of the equation is increasing in g. Moreover, it is negative when $g = -1$ and positive when $g = A$. Hence, there is exactly one fixed point, $\tilde{g}(Q) < A$, where $g^* = \tilde{g}(Q^*)$. Differentiating equation (36), one can verify that $\partial \tilde{g}(Q)/\partial Q < 0$ if and only if $1 > \delta\left(1 + \tilde{g}(Q)\right)^{1-\sigma}$. For $\sigma \leq 1$, the latter inequality holds since $g \leq A$ and we have assumed $1 > (A+1)^{1-\sigma}\delta$. For $\sigma > 1$, first note that $\tilde{g}(1) = \bar{g}_c$ and $1 > \delta\left(1 + \tilde{g}(1)\right)^{1-\sigma}$, so $\partial \tilde{g}(1)/\partial Q < 0$; hence for all $Q \leq 1$ the inequality $1 > \delta\left(1 + \tilde{g}(Q)\right)^{1-\sigma}$ holds. Since $Q^* < 1$, we have $\tilde{g}(Q^*) = g^* > \tilde{g}(1) = \bar{g}_c$. Therefore, $A > g^* > \bar{g}_c$. All other statements in the proposition are proven in the main text. **QED**

Proof of Proposition 5

Assume that $\delta \in (0,1)$ and $\sigma > 1$. Let $U(\tau, g(\tau))$ denote the objective function in (33). Since $U(\tau, g(\tau))$ is a continuous function of τ on $[0,1]$, it must have a maximum. Moreover, U is differentiable on $(0,1)$ with

$$\frac{dU}{d\tau} = \frac{\partial U}{\partial \tau} + \frac{\partial U}{\partial g} g'(\tau).$$

Now we prove that $dU/d\tau < 0$ for all $\tau \geq \tau^*$. First, note that $\frac{\partial U}{\partial \tau} < 0$ for all $\tau \geq \tau^*$, because $\tau^{-\sigma} - \frac{1}{\delta + a}(1-\tau)^{-\sigma} \leq 0$ and $\tau^{-\sigma} - \frac{1}{\delta}(1-\tau)^{-\sigma} < 0$ for all $\tau \geq \tau^*$ and

$$\frac{\partial U}{\partial \tau} = \left((A-g)k\right)^{1-\sigma}\left(\tau^{-\sigma} - \frac{1}{\delta + a}(1-\tau)^{-\sigma} + \left(\tau^{-\sigma} - \frac{1}{\delta}(1-\tau)^{-\sigma}\right)\left(\frac{\delta(1+g)^{1-\sigma}}{1-\delta(1+g)^{1-\sigma}}\right)\right).$$

Then, note that $\frac{\partial U(\tau,g)}{\partial g} < 0$, for $g = g(\tau)$ and for all $\tau \in [0,1]$, because $q(\tau, a) < q(\tau, 0)$ and

$$\begin{aligned}
\frac{\partial U(\tau, g)}{\partial g} &= k^{1-\sigma}(A-g)^{-\sigma}\left[\begin{array}{c} -q(\tau,a) \\ + \left(\frac{\delta q(\tau,0)(1+g)^{-\sigma}}{1-\delta(1+g)^{1-\sigma}}\right)\left(A - g - (1+g) + \frac{(A-g)\delta(1+g)^{1-\sigma}}{1-\delta(1+g)^{1-\sigma}}\right) \end{array}\right] \\
&= k^{1-\sigma}(A-g)^{-\sigma}\left(\frac{\delta q(\tau,0)(1+g)^{1-\sigma}}{1-\delta(1+g)^{1-\sigma}}\right)\left(\frac{q(\tau,a)}{q(\tau,0)} - 1\right),
\end{aligned}$$

where the second equality follows from the fact that $g(\tau)$ satisfies

$$\frac{q(\tau,a)}{q(\tau,0)} = \frac{\delta(1+g)^{-\sigma}}{1-\delta(1+g)^{1-\sigma}}(A-g).$$

Differentiating this last equation with respect to τ and g, we have that $g'(\tau) > 0$ if and only if $\sigma > 1$. Hence, $dU/d\tau < 0$ for all $\tau \geq \tau^*$, which implies that $\overline{\tau} < \tau^*$.

It is easy to verify that the proof of Proposition 4 implies that $g(\overline{\tau}) > \bar{g}_c$, with $\overline{g}_c =$

$[\delta\,(A+1)]^{1/\sigma}-1$. Therefore, we have $\overline{\tau}\in(0,\tau^{*})$ and $g\,(\overline{\tau})\in(\bar{g}_c,g^{*})$, as required. **QED**

Proof of Proposition 6

First note that

$$\frac{Q_c}{Q^*}=\left(\frac{q\,(\tau_c,0)}{q\,(\overline{\tau}_c,0)}\right)^{1/\sigma}=\left(\frac{\tau_c^{1-\sigma}+\delta^{-1}(1-\tau_c)^{1-\sigma}}{\overline{\tau}_c^{1-\sigma}+\delta^{-1}(1-\overline{\tau}_c)^{1-\sigma}}\right)^{1/\sigma},$$

where the first equality follows from the definitions of Q_c and Q^* given in the previous two proofs, and the fact that $\tau^*=\tau_c$, and the second equality follows from the definition of $q\,(\tau,a)$ given in equation (15). Next, note that the right side of the second equality above is increasing in τ_c for $\sigma>1$ and is decreasing in τ_c for $\sigma<1$. Since $\tau_c>(1+\delta^{-1/\sigma})^{-1}=\overline{\tau}_c$, it follows that $Q_c>Q^*$ if $\sigma>1$ and $Q_c<Q^*$ if $\sigma<1$. Hence, we have the following: (1) If $\sigma>1$, then $\tilde{B}(g^*,Q^*)=g^*>\tilde{B}(g^*,Q_c)>\tilde{B}(\bar{g}_c,Q_c)=g_c$, where the first inequality follows from the fact that $\partial\tilde{B}(\hat{g},Q)/\partial Q<0$, and the second one from the fact that $g^*>\bar{g}_c$ and $\partial\tilde{B}(g,Q)/\partial g>0$ for $g>\bar{g}_c$. (2) If $\sigma<1$, then $\tilde{B}(g^*,Q^*)=g^*<\tilde{B}(g^*,Q_c)<\tilde{B}(\bar{g}_c,Q_c)=g_c$, where the first inequality follows from the fact that $\partial\tilde{B}(\hat{g},Q)/\partial Q<0$, and the second one from the fact that $g^*>\bar{g}_c$ and $\partial\tilde{B}(g,Q)/\partial g<0$ for $g>\bar{g}_c$. It follows that $g^*>g_c$ if and only if $\sigma>1$, as required. **QED**

Proof of Proposition 7

Suppose that $\sigma>1$. First, note that the limit of g^* as δ approaches 1 is well defined for all $\sigma>1$. That $g^*>\bar{g}_c$, for $a>0$ and $\delta\leq 1$ follows from the proof of Proposition 4. Next, note that $\lim_{a\to 0}Q^*=1$, so $\lim_{a\to 0}g^*=\lim_{a\to 0}\tilde{g}(Q^*)=\tilde{g}(1)=\bar{g}_c$, as required. Finally, one can apply l'Hopital's rule to find $\lim_{a\to\infty}Q^*$. If $\sigma>1$, $\lim_{a\to\infty}Q^*=0$, and $\lim_{a\to\infty}g^*=\tilde{g}(0)=A$, where the last equality follows from substituting $Q=0$ into equation (36). This concludes the proof. **QED**

References

Auerbach, A.J. and L.J.Kotlikoff (1987): *Dynamic Fiscal Policy*, Cambridge: Cambridge University Press.

Azariadis C. and V. Galasso (2002): "Fiscal Constitutions", *Journal of Economic Theory* 103, 255-281.

Barro, R.J. (1974): "Are government bonds net wealth?", *Journal of Political Economy* 82, 1095-1117.

Barro, R.J. (1999): "Ramsey meets Laibson in the Neoclassical growth model", *Quarterly Journal of Economics* 114, 1125-1152.

Bernheim, B.D. (1989): "Intergenerational altruism, dynastic equilibria and social welfare", *Review of Economic Studies* 56, 119-128.

Bernheim, B.D. and D. Ray (1987): "Economic growth with intergenerational altruism", *Review of Economic Studies* 54, 227-241.

Boadway, R.W. and D.E. Wildasin (1989): "A median voter model of social security", *International Economic Review* 30, 307-328.

Burbidge, J.B. (1983): "Government debt in an overlapping-generations model with bequests and gifts", *American Economic Review* 73, 222-227.

Calvo, G.A. and M. Obstfeld (1988): "Optimal time-consistent fiscal policy with finite lifetimes", *Econometrica*, 56, 411-432.

Caplin, A. and J. Leahy (2004): "The social discount rate", *Journal of Political Economy* 112, 1257-1268.

Cooley, T.F. and J. Soares (1999): "A positive theory of social security based on reputation", *Journal of Political Economy* 107, 135-160.

Gollier, C. and R. Zeckhauser (2005): "Aggregation of heterogeneous time preferences", *Journal of Political Economy* 113, 878-896.

Grossman, G.M., Helpman, E. (1998): "Intergenerational redistribution with short-lived governments", *Economic Journal* 108, 1299-1329.

Hansson, I. and C. Stuart (1989): "Social security as trade among living generations", *American Economic Review* 79, 1182-1195.

Hori, H. (1997): 'Dynamic allocation in an altruistic overlapping generations economy", *Journal of Economic Theory* 73, 292-315.

Jackson, M.O. and L. Yariv (2014): "Present bias and collective dynamic choice in the lab", *American Economic Review* 104, 4184-4204.

Kimball, M.S. (1987): "Making sense of two-sided altruism", *Journal of Monetary Economics* 20, 301-326.

Kohlberg, E. (1976): "A model of economic growth with altruism between generations", *Journal of Economic Theory* 13, 1-13.

Krusell, P., B. Kuruşçu and A.A. Smith (2002): "Equilibrium welfare and government policy with quasi-geometric discounting", *Journal of Economic Theory* 105, 42-72.

Laibson, D. (1997): "Golden eggs and hyperbolic discounting", *Quarterly Journal of Economics* 112, 443-447.

Lindbeck, A. and J.W. Weibull (1987): "Balanced-budget redistribution as the outcome of political competition", *Public Choice* 52, 273-297.

Phelps, E. and R. Pollak (1968): "On second-best national saving and game-equilibrium growth", *Review of Economic Studies* 35, 185-199.

Ray, D. (1987): "Nonpaternalistic intergenerational altruism", *Journal of Economic Theory* 41, 112-132.

Saint Paul, G. (1992): "Fiscal policy in an endogenous growth model", *Quarterly Journal of Economics* 107, 1243-1259.

Streufert, P.A. (1993): "Consistent preferences and intergenerational equilibria in Markov strategies," in B. Dutta, S. Gangopadhyay, D. Mookerjee, and D. Ray (eds.), *Theoretical Issues in Economic Development*, 263-287, Bombay: Oxford University Press.

Strotz, R.H. (1956): "Myopia and inconsistency in dynamic utility maximization", *Review of Economic Studies* 23, 165-180.

Tabellini, G. (2000): "A positive theory of social security", *Scandinavian Journal of Economics* 102, 523-545.

Veall, M.R. (1986): "Public pensions as optimal social contracts", *Journal of Public Economics* 31, 237-251.

(1) $\sigma > 1$ (2) $\sigma = 1$ (3) $\sigma < 1$

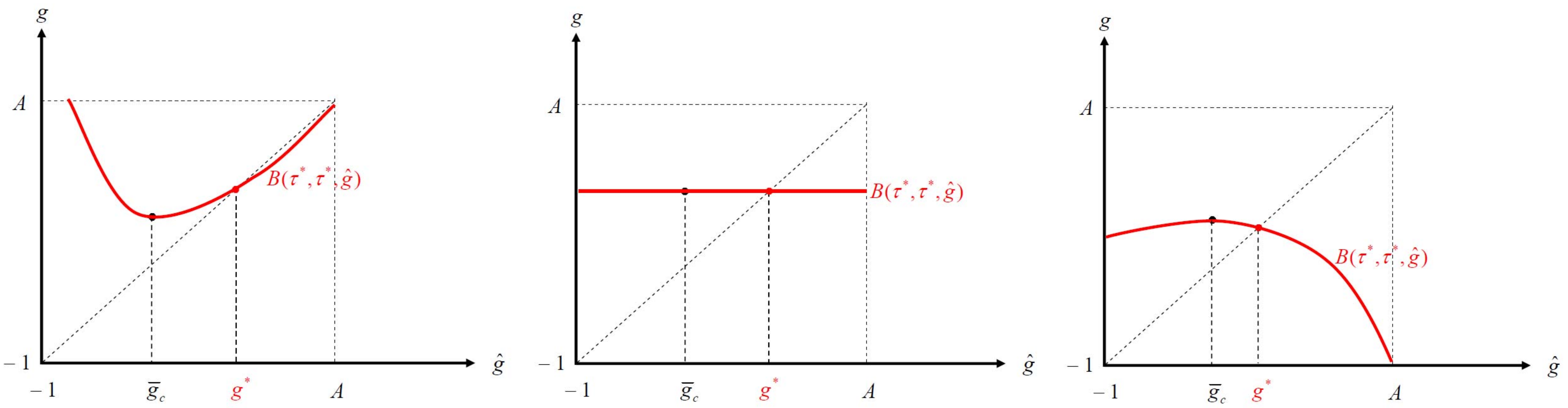

Figure 1