Present-Biased Preferences and Optimal Taxation of Parental Transfers*

Nicola Pavoni† and Hakki Yazici‡

September 16, 2013

Abstract

We study optimal taxation of bequests and inter vivos transfers in a model with altruistic parents facing children with present-biased preferences and provide an efficiency reason to tax parental wealth transfers. This tax result is independent of whether the government has access to saving taxes, which enables it to discipline the life-cycle saving behavior of the offspring. Interestingly, the optimality of a positive tax in the case where the government has limited instruments is motivated by the presence of a pure price externality. Cautioned by the technical complications present in this class of models, our normative prescriptions do not rely on the assumption of differentiability of the agents' policy functions. The mechanism that we outline here is potentially able to generate taxes of the same magnitude that we observe in the data regarding existing bequest and inter vivos taxation.

JEL classification: E21, H21, D03, D91.

Keywords: Transfer taxation, price externalities, quasi-hyperbolic discounting, altruism.

*We thank Per Krusell, Marco Ottaviani, seminar participants at Bogazici University, Goethe University in Frankfurt, IIES, the SED meetings in Ghent, University of Alicante, University of Bologna, University of Oxford, and the Eighth CSEF-IGIER Symposium on Economics and Institutions for their comments and suggestions. Pavoni gratefully acknowledges financial support from the European Research Council, Starting Grant #210908. Yazici gratefully acknowledges financial support from the European Community Framework Programme through the Marie Curie International Reintegration Grant #268457.

†Department of Economics, Bocconi University, via Roentgen 1, 20136 Milan; IFS, IGIER, and CEPR.

‡Sabanci University, Faculty of Arts and Social Sciences, Istanbul.

1 Introduction

Transfer taxation has long been a highly controversial issue. Except for a brief period of time, the U.S. government has maintained a positive bequest tax ever since it was first introduced in 1916. The tax was eliminated in 2010 and reintroduced in 2011. The nature of the policy debate, however, seems to go in only one direction: in all developed countries, we indeed observe either *zero or positive taxation* of parental transfers.[1]

This regularity on observed tax systems around the world contrasts with the lack of a clean theoretical justification for a positive tax on bequests from an efficiency point of view.[2] In virtually all traditional models of bequests, efficiency calls for either *zero or a negative tax* on parental transfers.[3] Perhaps the most widely analyzed model of bequest taxation is the altruistic model, where the motive for bequests comes from the assumption that parents care about their children's welfare. A maintained assumption in the altruistic model – much like most other models of intertemporal choice – is that people agree about intertemporal trade-offs, with their future selves or their descendants. The implication of the standard altruistic model for intergenerational wealth transfers - such as bequests and inter vivos - is simple. An altruistic parent knows that the offspring will save according to his optimal plan, which is also optimal from the parent's point of view. As a result, parents do not have any paternalistic concerns regarding the offspring's savings choice and the whole dynasty acts as if it is a single individual. If the society cares only about the parent directly, this would imply that parental transfers are socially optimal and should remain undistorted, given that there are no other reasons for taxation such as financing government expenditures. Following the same arguments, whenever society attaches some direct weight to future generations, parental transfers should actually be *subsidized* according to the classical altruistic model.[4]

Recently, however, laboratory and field studies on intertemporal choice have suggested that people display present-biased preferences, which means they may be tempted to divert from the plans that are optimal from a long-term perspective.[5] When parents face children

[1] For an overview on parental transfer taxation in OECD countries, see Cremer and Pestieau (2010) and OECD (2008).

[2] It is, of course, possible to justify taxation of parental transfers based on equality grounds. See, for instance, Piketty and Saez (2013).

[3] We discuss these models and their implications for bequest taxation in detail in the related literature section.

[4] See Kaplow (1995) and Farhi and Werning (2010) for a discussion of optimality of bequest subsidies under social preferences that weigh future generations directly.

[5] Present biased preferences are intended to represent situations where agents have short term discount

with present-bias problems, they foresee that the children will succumb to temptation and end up saving less than what is optimal for them. Now, intergenerational wealth transfers are motivated not only by equating marginal utilities of wealth across generations but also by paternalistic concerns. In this paper, we analyze optimal wealth transfer taxation in such a world and find a genuine *efficiency reason to impose a positive tax on parental transfers.*

We study an otherwise standard overlapping generations model where altruistic parents face offspring with present-biased preferences. We focus on Markov equilibria. Every period agents make consumption saving decisions. In the last period of their lives, they choose how much to consume and transfer to their offspring, knowing that the latter have present-bias problems. To simplify the exposition, in the main body of the paper, we assume transfers are made only in terms of bequests.[6] To isolate the implications of present-bias problems for bequest taxation, we assume away other reasons for intervention: namely, there is a representative dynasty (no horizontal inequality), there are no exogenous government expenditures to be financed, and government cares only about the parent's welfare directly. Thus, there is no reason for a government to distort the economy except to correct for the present-bias problems of the offspring.

In this environment, we study optimal taxation considering two separate notions of efficiency. First, we analyze what we denote as 'Ramsey' (or sometimes unconstrained efficient) allocation, defined as the allocation that would arise if people had no present-bias problems regarding their life-cycle saving decisions. This is probably the most widely adopted benchmark in the literature. We characterize the equilibrium transfer behavior of parents when they face offspring with present-bias problems and compare it to the Ramsey transfer behavior. It is customary to make such a comparison by defining and measuring a parental transfer wedge which represents the discrepancy between individual and social marginal returns to transfers. However, such an analysis requires differentiability of policy functions that describe people's life-cycle saving behavior, and in our environment with sequences of agents disagreeing over intertemporal trade-offs, it is not possible to establish differentiability properties in general. Therefore, first, without making any differentiability assumptions, we compare equilibrium

rates that are larger than long term ones. See DellaVigna (2009) for a survey of field studies and Frederick, Loewenstein, and O'Donoghue (2002) for a survey of experimental studies. For evidence of hyperbolic discounting in consumption-asset panel data, see Laibson, Repetto, and Tobacman (2007).

[6] In Section 7, we consider an extended version of the model where there is a period in which parents and offspring coexist, and in addition to bequests, parents can make inter vivos transfers in the period coexistence. There we show that it is efficient to tax both types of transfers.

and Ramsey *levels* of parental transfers and show that parents want to transfer too much relative to the Ramsey level in equilibrium. Second, under the assumption of differentiability, we show that this result about the levels of transfers translates into a positive parental transfer wedge, indicating that private returns to transfers are larger than social returns. Finally, we show that, when we focus our attention to linear Markov equilibria, the parental transfer wedge implies optimality of transfer taxation: the Ramsey allocation can be implemented using positive linear wealth transfer taxes as long as the government has access to (linear) life-cycle saving taxation to offset offspring's present-bias problems. We also show that the linear Markov equilibria assumption is innocuous by proving that, under the constant elasticity of intertemporal substitution utility function (CEIS), such equilibria exist.

The brief intuition for why parents transfer too much and hence should be taxed is as follows. Due to present-bias problem, in the laissez-faire equilibrium, the offspring save less than the Ramsey level for their old age, which induces parents to transfer more than the Ramsey level in order to compensate for the offspring's undersaving behavior. In order to make children save the correct (Ramsey) amount, the government uses linear subsidies on savings and uses lump-sum taxes to finance these subsidies. However, from the perspective of the parents, who take the saving subsidy as given, their offspring are still undersaving under the new - subsidized - interest rate. In other words, since parents take the lump-sum tax as given, they do not internalize the fact that the subsidy is there to discipline the saving behavior of the next generation and does not actually change the gross return to their offspring's savings. As a result, the parents still transfer too much to their offspring and hence should be taxed.

Next, we investigate whether the optimality of positive taxes on transfers depends on the government's ability to correct offspring's present-bias problems through saving taxes. We solve a planning problem where a planner chooses the level of bequests to maximize parental (time-consistent) welfare, but cannot control offsprings' saving decisions. The constraint that the planner cannot control offsprings' savings reflects the limitation imposed on government that it cannot tax life-cycle savings. We call the solution to this planning problem 'constrained efficient' allocation. Again, we first compare equilibrium and constrained efficient levels of parental transfers. We show that in equilibrium parents want to transfer too much relative to the constrained efficient level of transfers as well. Second, we show that there is a positive transfer wedge, meaning private marginal returns to transfers exceed social marginal returns. Then, we prove that, under the assumption of linear Markov equilibria, the government can implement the constrained efficient allocation through positive, linear taxes on bequests.

The constrained planner is identical to a parent in the sense that neither can control the offspring's saving decision directly. Therefore, much like a parent, the planner also finds it optimal to choose a bequest level that is higher than the Ramsey level. In fact, in a partial equilibrium version of our economy where the interest rates are exogenous, parental bequeathing behavior is constrained efficient, implying that bequests should go untaxed. The story is different under general equilibrium, however. The planner still agrees with the parent about bequeathing more than the Ramsey level, but they do not agree on the exact amount. Due to a *price externality*, parents want to bequeath too much relative to the constrained efficient level. Intuitively, the government realizes that increasing parental bequests increases the savings of the offspring, which then depresses the interest rate they face assuming there is decreasing marginal returns to capital. This in turn makes offspring decrease their savings. An individual parent does not realize this second effect since the parent – unlike the planner – is small and takes prices as given. Consequently, parents bequest more than the constrained efficient level. To prevent this, the government should tax bequests.

To comprehend whether our mechanism has a quantitative potential of providing a rationale for observed tax rates on bequests, we compute taxes for a parameterized version of our economy. The current estate tax system in the United States has a progressive nature, where the top tax rate is 55%. In the United Kingdom, there is a flat 40% tax on inheritance.[7] These tax rates are in the range of the optimal bequest tax rates generated by the model: varying the level of the severity of present-bias problems in the range estimated by Laibson, Repetto, and Tobacman (2007), optimal taxes vary from 10% to 48%.

Finally, this paper makes a methodological contribution by deriving normative implications of models with present-biased preferences in the absence of differentiability assumptions. This is in contrast to the approach taken both in general in the public finance literature, and more specifically within the literature on optimal fiscal policy under present-bias problems. As described above, our results are presented over three levels of analysis. In addition to analyzing optimal wedges - which requires the assumption of differentiability - and optimal taxes - which requires further restrictive assumptions implying concavity of the agents' problems - we provide a non-differentiable analysis of the discrepancies between the equilibrium allocation and the efficient allocation. Deriving normative prescriptions at this level of generality is particularly important for models with present-biased preferences for at least two reasons. First, it is well known that, in general, these models may not have equilibria with differentiable policies (e.g.,

[7]Both countries have exemption levels below which wealth transfers go untaxed. In the United States, currently the exemption level is 1 million US dollars, whereas in the United Kingdom it is 325,000 pounds.

see Harris and Laibson (2002)). Second, even when a differentiable equilibrium exists, models with multiple selves often admit multiple (Markov) equilibria. It is important in such cases to understand whether a policy implication emerges from a general principle or, instead, it is linked to a specific equilibrium (or equilibrium property such as differentiability or linearity of the policy).

Related Literature. This paper is related to two strands of literature. First, as we discussed above, it is related to the literature on optimal taxation of bequests and inter vivos transfers.[8] Our contribution here is to provide a novel, pure efficiency argument for taxing parental transfers. In addition to the model with altruism that we discussed above, a widely used model of bequest is the warm-glow (or "joy of giving") model. In this model as well, either the optimal tax of bequests is zero or parental transfers should be subsidized to internalize the positive externality that wealth transfers induce on future generations (e.g., Kopczuk (2010)). Another framework considered in the literature is the model with exchange motives for bequests. In this class of game-theoretical models, the normative predictions crucially depend on the details of the game played between the parents and the offspring (e.g., Laitner (1997)). Finally, we have the accidental bequests model, where taxing (accidental) bequests is non-distortionary.[9] According to this model, bequest taxes are simply a good way to finance positive government expenditures when the government has no lump-sum taxes available. This model does not imply an optimal positive tax, at least not in the way we define optimality in this paper. Specifically, there is no equilibrium inefficiency to be corrected by taxes on bequests or gifts.

Our paper is also related to a number of recent papers that have explored the implications of self-control problems for optimal taxation. O'Donoghue and Rabin (2003) analyze a model of paternalistic taxation for unhealthy goods. More closely related is Krusell, Kuruscu, and Smith (2010), which analyzes properties of linear taxes on life-cycle savings that implement the Ramsey allocation. Pavoni and Yazici (2012) also focus on optimal Ramsey taxation of life-cycle savings within a quasi-hyperbolic discounting model and allow for age changing present-bias problems. The current paper is different from these papers in at least three ways. First, none of these papers analyze the implications of self-control problems for bequest and inter

[8]See Cremer and Pestieau (2010), Kaplow (2001), and Kopczuk (2010) for excellent surveys on the literature on optimal transfer taxation.

[9]There is an obvious theoretical assumption - not yet carefully tested empirically - that would justify positive taxation of all sorts of wealth (not only of parental wealth transfers). It is the assumption that wealth concentration generates negative externalities (Kopczuk (2010)).

vivos transfers taxation. The efficiency of positive transfer taxation is crucially based on the presence of a positive wedge in equilibrium transfer decisions and a price externality, which, to the best of our knowledge, has not been cleanly identified before in this class of models. Second, in all of these papers, being focused on Ramsey taxation, the rationale behind government intervention is paternalism: the government uses taxation to correct the behavior of agents who have self-control problems. In the current paper, when considering constrained efficiency, we analyze optimal taxation of parental transfers which do not stem from parents' present-bias problems but rather from their inability to internalize prices. Third, we make a methodological contribution by abstracting from differentiability of policy functions, a particularly delicate issue in models with agents having present-biased preferences.

The paper is organized as follows. Section 2 introduces the main model, and section 3 characterizes the equilibrium bequest behavior of parents in the absence of government intervention. In Section 4, we compare equilibrium and Ramsey bequest behavior and provide a tax implementation of the Ramsey allocation. In Section 5, we characterize the constrained efficient allocation, compare it to the equilibrium allocation, and provide implementation. In Section 6, we briefly scrutinize the potential quantitative significance of our mechanism. In Section 7, we verify the robustness of our results by extending the model in various directions. Section 8 concludes.

2 Model

The economy is populated by a continuum of a unit measure dynasties who live for a countable infinity of periods, $t = 0, 1, ...$, where each agent within a dynasty is active for two periods. In the first period of their lives, agents are young adults, and they make consumption saving decisions facing present-bias problems. In the second period, they become (old) parents, decide how much to consume and bequeath, and die. In the period after they die, the child becomes a young adult and receives the bequests. This is a model of non-overlapping generations.[10] Parents are altruistic and sophisticated in the sense that they anticipate their

[10]In Section 7.2, we allow for a longer life cycle and show that our main results regarding bequest taxation are robust to such extension. In this extension, we also model periods in which parents and their offspring are alive together and analyze inter vivos behavior and taxation. We show that - much like bequests - parents transfer too much through inter vivos relative to what is socially optimal and, hence, it is optimal to tax inter vivos transfers.

children's present-bias problems.[11] Agents have one unit of time that they supply inelastically to the market. Consider a parent in some calendar year t. Her preference over dynastic allocation is given by

$$V_t = u(c_t) + \gamma \left[u(c_{t+1}) + \delta V_{t+2} \right],$$

where V_t represents the dynastic welfare of the parent in period t, c_t is parental consumption, and c_{t+1} is the first period consumption of the offspring. The instantaneous utility function, u, has the usual properties: strictly increasing, strictly concave, and twice differentiable, with $\lim_{c \to 0} u'(c) = +\infty$. The parameters $\delta, \gamma \in [0, 1)$ are the long-run discount factor and the altruism factor, respectively.

The way we model present-biased preferences is inspired by the quasi-hyperbolic discounting framework of Strotz (1955), Phelps and Pollak (1968), and Laibson (1997). Even though the offspring's preference from a long-term perspective agrees with that of the parent, at the time of savings, the offspring behaves according to the following preference:

$$u(c_{t+1}) + \beta\delta V_{t+2}.$$

When $\beta = 1$, true preference and behavioral preference agree; there is no present-bias problem. Whenever $\beta < 1$, people in the short term discount future at a higher rate (according to $\beta\delta$) than their long-term perspective δ. In this benchmark version of the model, with agents deciding over only two periods, present-bias does not imply time inconsistency within one's life time.[12] In Section 7, we analyze the general multi-period version of the model. When agents leave for at least three periods, our model naturally converges to the quasi-hyperbolic discounting framework as formalized by Laibson (1997). There, agents have time-inconsistent preferences. Our benchmark model has the important virtue of simplicity, and, at the same time, it shows how the discrepancy in intertemporal preferences between parents and offsprings (which is fully captured here) is the key force responsible for our normative prescriptions.

Production takes place at the aggregate level according to the function $F(k_t, l_t)$, where k_t and l_t are aggregate levels of capital stock and labor in period t, and F is a neoclassical concave production function with the usual properties: $F_1, F_2 > 0$ and $F_{11}, F_{22} \leq 0$. Since

[11] In Section 7.1, we analyze what happens if parents are unaware of their offspring's present-bias problems and show that our main results carry over to this case as well.

[12] Note that, even tough there is no time-inconsistency within an agent's life time, agents have an intertemporal aspect of 'inconsistency': a young adult at the time of deciding his savings discounts at the rate $\beta\delta$; in contrast, when he thinks about his future offspring at the same age as him, the young adult believes the future young adult should save according to the long-run preferences with $\delta-$discounting.

each agent supplies one unit of labor inelastically, for all t, we have

$$l_t = 1.$$

Letting θ be the depreciation rate, this allows us to write the production function as

$$f(k) = F(k, 1) + (1 - \theta)k.$$

Letting $f(k_0)$ be the amount period-0 parent is endowed with, the feasibility for any $t \geq 0$ is

$$c_t + k_{t+1} = f(k_t).$$

As evident from the feasibility condition above, we assume there is one representative dynasty, which implies that in any calendar year there is only one age group alive. We could, instead, allow for members of different dynasties to be at different points in their life cycles. Moreover, we could also allow for income heterogeneity by assuming, for instance, that people have different skill levels and that effective labor is given by labor times the skill level, similar to Mirrlees (1971). We abstract from such distributional issues in order to isolate our mechanism. As we explain in Section 7.4, the mechanism behind our results is robust to both of these extensions.

3 Equilibrium

In this section, we characterize the equilibrium parental transfer behavior. Let b_{t+1} and b_{t+2} denote the bequest made by the parent in period t and the offspring's saving level in $t+1$, respectively. Let R_t, w_t be the interest rate and the wage rate in period t. Let $Q := \{R_t, w_t\}_{t=0}^{\infty}$ be the sequence of prices that decision makers take as given. Finally, let $Q_t := \{R_s, w_s\}_{s=t}^{\infty}$ be continuation of prices from period t onwards.

Define $V(a_t, Q_t)$ as the value of the problem of an agent who is a parent in calendar year t with $a_t := R_t b_t + w_t$ units of wealth and who faces the price sequence Q_t. The parent's problem is given by

$$V(a_t, Q_t) = \max_{b_{t+1} \geq -B(Q_{t+1})} u(c_t) + \gamma \left[u\left(c_{t+1}(b_{t+1}, Q_{t+1})\right) + \delta V\left(a_{t+2}(b_{t+1}, Q_{t+1}), Q_{t+2}\right) \right], \qquad (1)$$

subject to the budget constraints and the definition of wealth

$$\begin{aligned} c_t &= a_t - b_{t+1}, \\ c_{t+1}(b_{t+1}, Q_{t+1}) &= R_{t+1} b_{t+1} + w_{t+1} - b_{t+2}(b_{t+1}, Q_{t+1}), \\ a_{t+2}(b_{t+1}, Q_{t+1}) &:= R_{t+2} b_{t+2}(b_{t+1}, Q_{t+1}) + w_{t+2}, \end{aligned}$$

together with the condition defining the policy of the offspring:[13]

$$b_{t+2}(b_{t+1}, Q_{t+1}) = \arg \max_{\tilde{b}_{t+2} \geq -B(Q_{t+2})} u\left(R_{t+1}b_{t+1} + w_{t+1} - \tilde{b}_{t+2}\right) + \beta\delta V(R_{t+2}\tilde{b}_{t+2} + w_{t+2}, Q_{t+2}). \quad (2)$$

$B(Q_t)$ is the 'natural' (and never binding) borrowing limit defined by requiring consumption to be non-negative at all periods:

$$B(Q_t) := \sum_{s=t}^{\infty} \frac{w_s}{\Pi_{p=t}^{s} R_s}.$$

In equilibrium, prices are given by

$$\begin{aligned} R_t &= f'(k_t), \\ w_t &= f(k_t) - f'(k_t)k_t, \end{aligned} \quad (3)$$

and aggregate capital and saving levels satisfy the market clearing condition

$$k_t = b_t.$$

The parent chooses his bequest level b_{t+1} taking into account the choice rule of his offspring, $b_{t+2}(\cdot, Q_{t+1})$. The parent is sophisticated in the sense that he correctly guesses his children's choice, and that is why he takes into account condition (2), we see in his problem. Define $b_{t+1}(b_t, Q_t)$ as the policy function describing parental optimal bequeathing behavior as a function of his period $t-1$ savings and the price sequence.

A Markov equilibrium *consists of a sequence of capital levels* $\{k_t\}_{t=0}^{\infty}$*, a sequence of prices* Q*, value functions* $V(\cdot; Q_t)$ *and policy functions* $\{b_{t+1}(\cdot, Q_t), b_{t+2}(\cdot, Q_{t+1})\}_{t=0,2,4,...}$ *such that: (i) the prices satisfy (3); (ii) the value function and the policies are consistent with the parent's problem; (iii) markets clear:* $b_t = k_t$ *for all* t*.*

Proposition 1 below characterizes equilibrium parental transfer behavior. Proving Proposition 1 would be relatively easier if we could assume differentiability of the value function, $V(\cdot)$, and the policy functions, $b_{t+2}(\cdot)$, which describe how offspring's saving choice changes as a function of parental bequests. However, it is well known that, in environments with present-bias problems, we cannot guarantee even the continuity of the policy functions in general.[14] Therefore, to show that the result in Proposition 1 is quite general, we prove it without making any differentiability or continuity assumptions about the value or policy functions.

[13]To save notation, we indicate the policy as a function. In case there were multiple solutions to the offspring's problem, $b_{t+2}(\cdot)$ should be intended as a selection from the policy correspondence.

[14]See Morris and Postlewaite (1997), Krusell and Smith (2003), and Harris and Laibson (2002) for examples of economies with quasi-hyperbolic discounters where policy functions are discontinuous.

Proposition 1 *Suppose $\beta < 1$. Then, in equilibrium,*

$$u'(c_t) \geq R_t \gamma u'(c_{t+1}), \tag{4}$$

with strict inequality whenever the offspring's optimal saving policy, $b_{t+2}(\cdot, Q_{t+1})$, is strictly monotone in the amount of the bequests received, b_{t+1}.
If $\beta = 1$, then

$$u'(c_t) = R_t \gamma u'(c_{t+1}). \tag{5}$$

Proof. Relegated to the Appendix. ■

To get a better grasp of what the proposition says, first focus on the second part. If the child does not have a present-bias problem, meaning $\beta = 1$, then the parent chooses the level of transfers to equate the marginal cost of his forgone consumption (left-hand side of (5)) to the marginal benefit of his child's increased consumption in period $t+1$ (right-hand side of (5)). However, when $\beta < 1$, meaning the child has present-bias issues, then, as seen from (4), the parent keeps increasing transfers even after the marginal cost is equated to the marginal benefit from increased child consumption in period $t+1$. Why does the parent do that? This is because, from the parent's perspective, increasing bequests carries an additional benefit when the offspring has present-bias problems. We will give an intuitive explanation of this extra benefit in the next subsectio where we assume differentiability of the value and policy functions. The differentiability assumption will also allow us to provide a sharper marginal characterization of equilibrium bequest behavior, which we need for analyzing optimal bequest wedges in Section 4 and Section 5.

3.1 Equilibrium Parental Behavior under Differentiability

In this section, we provide a marginal condition that characterizes equilibrium bequest behavior assuming differentiability of the policy function that describes offspring's savings. Recall that $b_{t+2}(\cdot, Q_{t+1})$ represents the offspring's equilibrium choice under price sequence Q_{t+1} as a function of the bequests he receives from his parent, b_{t+1}. In the rest of this section, we drop the dependence of b_{t+2} on Q_{t+1} in order to ease notation. Now consider a parent's problem. The parent chooses b_{t+1} subject to the flow budget constraints and the function $b_{t+2}(b_{t+1})$, defined by (2), which describes offspring saving decision. The parent's first-order optimality condition with respect to the bequest decision, b_{t+1}, is

$$u'(c_t) = \gamma \left(u'(c_{t+1}) \left[R_{t+1} - \frac{\partial b_{t+2}(b_{t+1})}{\partial b_{t+1}} \right] + \delta V_1(a_{t+2}, Q_{t+2}) R_{t+2} \frac{\partial b_{t+2}(b_{t+1})}{\partial b_{t+1}} \right), \tag{6}$$

where the derivatives are all evaluated at the equilibrium allocation.

Observe that the offspring's first-order optimality condition for b_{t+2} is given by

$$u'(c_{t+1}) = \beta\delta V_1(a_{t+2}, Q_{t+2})R_{t+2}. \tag{7}$$

Using (7) in the parental optimality condition (6), we get the following proposition which describes equilibrium parental bequeathing behavior under differentiability.

Proposition 2 *Suppose $b_{t+2}(\cdot)$ is differentiable. Equilibrium bequest behavior is characterized by*

$$u'(c_t) = \gamma\left(R_{t+1}u'(c_{t+1}) + \frac{\partial b_{t+2}(b_{t+1})}{\partial b_{t+1}}u'(c_{t+1})\left[-1+\frac{1}{\beta}\right]\right). \tag{8}$$

Equation (8) is the usual savings optimality condition, with an additional term on the right-hand side. The left-hand side is the marginal cost of increasing bequests, which equals the utility loss from forgone parental consumption. The first term on the right-hand side is the usual marginal benefit of increasing saving – the utility gain from increased consumption in the period during which returns to savings are received. There is a second term on the right-hand side, however. One can see that this term does not show up in the solution to the usual savings problems where $\beta = 1$, meaning the saver and the person receiving savings agree on what the receiver will do with the savings (an implication of the Envelope condition). This additional term summarizes how increasing parental transfers affects parental welfare by affecting the consumption levels of the offspring. It is a multiplication of two terms: the first term,

$$\frac{\partial b_{t+2}(b_{t+1})}{\partial b_{t+1}} > 0,$$

tells how the offspring's saving is affected by an increase in bequests. In general, this derivative is weakly positive since increasing transfers increases the period $t+1$ wealth of the offspring, which weakly increases his savings. As we show in Lemma 16 in Appendix, under the assumption of differentiability of $b_{t+2}(\cdot)$, this derivative is strictly positive.

The second term,

$$= u'(c_{t+1})\left[-1+\frac{1}{\beta}\right] > 0,$$

represents the utility value to the parent of increasing b_{t+2} marginally and is strictly positive whenever $\beta < 1$.

Intuitively, the parent knows that from his perspective the offspring is undersaving. So, parental welfare goes up if the parent can make the offspring increase his savings, which is possible by increasing bequests since $\frac{\partial b_{t+2}(b_{t+1})}{\partial b_{t+1}} > 0$. As a result, the additional term in (8) is positive: there is an additional marginal benefit of increasing transfers for the parent. It is this extra benefit of bequeathing that makes the parent behave according to (4). Observe that under differentiability the strict version of equation (4) holds.

4 Ramsey (Unconstrained Efficiency)

Defining efficiency in environments where agents do not agree about intertemporal trade-offs might be controversial. We start with a widely adopted benchmark. The *Ramsey* allocation is the solution to a fictitious social planner's consumption-saving problem where the planner discounts exponentially. The planner is unconstrained in the sense that he can choose bequests and savings for all generations.

This allocation has at least three desirable properties. First, it corresponds to the *equilibrium allocation* in absence of present-bias problems. Second, it is unique and simple to characterize, as it is the dynastic solution to the standard Ramsey-Cass-Koopmans optimal growth problem. In our model, this allocation also coincides with the optimal allocation according to the parents' preferences at all generations.[15]

We first characterize the Ramsey allocation. Then, we show that the optimal level of bequests does not satisfy the parent's optimality condition in laissez-faire equilibrium: parents bequeath too much relative to the Ramsey level. This implies that implementing the Ramsey allocation in the market requires government intervention. Finally, we provide an implementation of the Ramsey allocation in the market through linear taxes on savings and bequests for a special class of equilibria. The optimal tax on bequests is positive.

4.1 The Ramsey Allocation

We assume that society cares about the offspring only indirectly through the parents' altruism.[16] Under this assumption, the Ramsey allocation is unique, and given by the solution to

[15]By taking a long-term perspective and evaluating welfare according to the parent's preference, we are in fact following much of the literature on self-control. See DellaVigna and Malmendier (2004), Gruber and Koszegi (2004) and O'Donoghue and Rabin (2006), for example.

[16]Allowing society to care about future generations directly does not change any of our considerations, but it does add a well-known additional channel that calls for subsidizing wealth transfers. See Farhi and Werning (2010), Kaplow (1995), and Kaplow (2001). Since we want to analyze our mechanism in isolation, we assume that society cares about offspring only indirectly through parental welfare.

a fictitious social planner's consumption saving problem where the planner has altruism and discount factors γ and δ, respectively. Suppose t is a period of parenthood. The following Euler equations characterize the Ramsey levels of bequests and savings, which we denote with an asterisk throughout the paper: for all t,

$$\begin{aligned} u'(c_t^*) &= \gamma f'(k_{t+1}^*)u'(c_{t+1}^*), \\ u'(c_{t+1}^*) &= \delta f'(k_{t+2}^*)u'(c_{t+2}^*). \end{aligned} \tag{9}$$

4.2 The Ramsey Wedge

Now we turn to the implications of Proposition 1 for equilibrium wealth transfer behavior relative to the Ramsey allocation. Remember from (9) that in the Ramsey allocation we have

$$u'\left(c_t^*\right) = \gamma R_{t+1}^* u'\left(c_{t+1}^*\right).$$

Comparison of this condition with the equilibrium condition (4) implies that the equilibrium level of parental transfers does not satisfy the Ramsey condition for bequests. In the Ramsey allocation, the offspring's choice of savings is optimal from parents' perspective, which implies that the only marginal benefit of bequests comes from increased offspring consumption in the next period. On the other hand, in equilibrium, parents face offspring with present-bias problems, which implies there is an additional benefit to bequeathing: compensating for offspring's period $t+1$ undersaving behavior. Therefore, the parent will keep increasing his transfers beyond the Ramsey level. We summarize this result with the following corollary.

Corollary 3 *A parent with a Ramsey level of wealth bequeaths more in equilibrium compared to Ramsey level of bequests.*

Proof. Compare conditions (9) and (4) and strict concavity of u. ■

If we assume differentiability of policy functions, Corollary 3 translates into a positive bequest wedge. To see this, first define the *Ramsey bequest wedge* as:

$$BW_t^* = -u'(c_t^*) + \gamma u'(c_{t+1}^*)\left(R_{t+1}^* + \frac{\partial b_{t+2}(b_{t+1}^*, Q_{t+1}^*)}{\partial b_{t+1}}\left\{-1 + \frac{1}{\beta}\right\}\right).$$

Remember that parents bequeath according to (8). The Ramsey bequest wedge in period t measures the efficient distortion that the planner needs to create in the bequest decision of a parent in that period in order to make him choose the Ramsey level of bequests. A positive (negative) BW_t means that a parent facing the Ramsey interest rate, $R_{t+1}^* = f'(b_{t+1}^*)$, would like to increase (decrease) his bequests marginally above (below) the Ramsey level if there is no government intervention.

Corollary 4 *Suppose $b_{t+2}(\cdot)$ is differentiable. Then, $BW_t^* > 0$ for all parenthood periods t.*

Proof. Follows directly from (9), the definition of BW_t^*, and that $\frac{\partial b_{t+2}(b_{t+1}^*, Q_{t+1}^*)}{\partial b_{t+1}} > 0$, which follows from Lemma 16 in the Appendix. ■

4.3 Implementation: Ramsey Taxation of Bequests

In this section, we want to implement the Ramsey allocation through a linear tax system on life-cycle savings and parental wealth transfers. Let τ_{t+1} denote the linear tax rate on returns to period t savings, b_{t+1}. If t is a period of parenthood, then τ_{t+1} is a tax on bequests. Tax proceeds are rebated in a lump-sum manner in every period, so that the government balances its budget period by period. Letting T_t denote lump-sum taxes in period t,

$$T_t = R_t \tau_t b_t.$$

Let $\Upsilon := \{\tau_t, T_t\}_{t=0}^{\infty}$ be the sequence of taxes that the government chooses and commits to at the beginning of time and $\Upsilon_t := \{\tau_s, T_s\}_{s=t}^{\infty}$. Let $\Upsilon^* := \{\tau_t^*, T_t^*\}_{t=0}^{\infty}$ denote a tax system that implements the Ramsey allocation. We are interested in the Ramsey taxes on wealth transfers.

Letting $\Psi := (Q, \Upsilon)$ be the joint sequence of prices and taxes, let $\Psi_t := (Q_t, \Upsilon_t)$. Define $V_t(a_t, \Psi_t)$ as the problem of a parent with wealth level a_t in calendar year t facing Ψ_t, where the wealth level is $a_t := R_t b_t(1 - \tau_t) + T_t + w_t$. The parent's problem is given by

$$V(a_t, \Psi_t) = \max_{b_{t+1} \geq -B(\Psi_{t+1})} u(c_t) + \gamma \left[u\left(c_{t+1}(b_{t+1}, \Psi_{t+1})\right) + \delta V\left(a_{t+2}(b_{t+1}, \Psi_{t+1}), \Psi_{t+2}\right) \right],$$

subject to the budget constraints

$$\begin{aligned} c_t &= a_t - b_{t+1}, \\ c_{t+1}(b_{t+1}, \Psi_{t+1}) &= R_{t+1} b_{t+1}(1 - \tau_{t+1}) + T_{t+1} + w_{t+1} - b_{t+2}(b_{t+1}, \Psi_{t+1}), \\ a_{t+2}(b_{t+1}, \Psi_{t+1}) &= R_{t+2} b_{t+2}(b_{t+1}, \Psi_{t+1})(1 - \tau_{t+2}) + T_{t+2} + w_{t+2}, \end{aligned}$$

and the offspring's policy function is defined as

$$b_{t+2}(b_{t+1}, \Psi_{t+1}) = \arg \max_{\tilde{b}_{t+2} \geq -B(\Psi_{t+2}),} u(\tilde{c}_{t+2}) + \beta \delta V(\tilde{a}_{t+2}, \Psi_{t+2}),$$

subject to

$$\begin{aligned}\tilde{c}_{t+2} &= R_{t+1}b_{t+1}(1-\tau_{t+1})+T_{t+1}+w_{t+1}-\tilde{b}_{t+2}, \\ \tilde{a}_{t+2} &= R_{t+2}\tilde{b}_{t+2}(1-\tau_{t+2})+T_{t+2}+w_{t+2}.\end{aligned}$$

The natural debt limit under Ψ_t is given by

$$B(\Psi_t) := \sum_{s=t}^{\infty} \frac{w_s + T_s}{\Pi_{p=t}^{s} R_s(1-\tau_s)}.$$

Note that the offspring's optimal policy is also a function of taxes.

In general, an agent's problem at any age is not convex since each parent faces a constraint describing the offspring's policy, which may potentially break the convexity of the constraint set. Therefore, showing that the first-order optimality conditions of agents are satisfied by the Ramsey allocation under a tax system does not guarantee that the tax system implements the Ramsey allocation. As a result, Proposition 1 does not automatically imply that there is a linear tax system that implements the Ramsey allocation. Therefore, we restrict attention to Markov equilibria with policy functions that are linear in current wealth. The linearity of the policy functions guarantees that agents' constraint sets are convex, thus implying that their problems are concave. Hence, we have the following implementation result.

Proposition 5 *Suppose (Markov) equilibrium (with taxes) admits policies that are linear in current wealth. Then, there is a linear tax system that implements the Ramsey allocation. In this system, policies are strictly increasing, and optimal bequest taxes are strictly positive if and only if $\beta < 1$. They are given by*

$$\tau_{t+1}^{*} = \left(-1+\frac{1}{\beta}\right) M_{t+2}(\Psi_{t+1}^{*}) \frac{1}{R_{t+1}^{*}} > 0,$$

where

$$M_{t+2}(\Psi_{t+1}^{*}) = \frac{\partial b_{t+2}\left(b_{t+1}^{*}, \Psi_{t+1}^{*}\right)}{\partial b_{t+1}} > 0$$

is the coefficient of offspring's (linear) policy function under the Ramsey tax system and prices implied by the Ramsey allocation.

Proof. If policy functions are linear, they are differentiable hence strictly from Lemma 16 in the Appendix. Moreover, the linearity of the policies implies that each agent's problem is concave, which implies that, once feasibility is guaranteed, the parent's first-order optimality conditions are necessary and sufficient for the equilibrium. Then it is easy to derive the optimality condition for parental bequest choice under taxes, analogous to (8) :

$$u'(c_t) = \gamma u'(c_{t+1}) \left(R_{t+1}\left(1-\tau_{t+1}\right) + \frac{\partial b_{t+2}\left(b_{t+1}, \Psi_{t+1}\right)}{\partial b_{t+1}} \left\{ -1 + \frac{1}{\beta} \right\} \right).$$

Substituting in the Ramsey allocation and using (9) gives the result. ■

Next, we show that when the utility function is of the CEIS form, there is always a Markov equilibrium with policy functions that are linear in current wealth.[17]

Proposition 6 *Suppose period utility is of the CEIS form, meaning*

$$\begin{aligned} u(c) &= \frac{c^{1-\sigma}}{1-\sigma}, \ \textit{for } \sigma \in (0,1) \ \textit{and } \sigma > 1; \\ &= \log(c), \ \textit{for } \sigma = 1. \end{aligned}$$

Then, if an equilibrium with taxes exists, there is an equilibrium in which consumption in each period is a linear function of the net present value of wealth as of that period.

Proof. Relegated to the Appendix. ■

It is interesting to note that even though the government corrects offspring's saving behavior through saving taxes, parental transfers should still be taxed to achieve the Ramsey allocation. Here, one might ask: given that from the Ramsey perspective the offspring is saving the right amount (thanks to corrective taxes that offset the effect of offspring's present-bias problem), why does the parent still bequeath more than the Ramsey level? This occurs because, from the parent's perspective, the offspring is undersaving. The taxes that are levied on the offspring create a diversion between the parent's and the planner's perception of what is optimal for the offspring. The planner knows that the lump-sum taxes required to balance its budget in period $t+2$ equal $T_{t+2} = R_{t+2}\tau_{t+2}b_{t+2}$. Thus, the return to the savings of the offspring in any period $t+1$ is actually R_{t+2}, and the taxes are there only to drive the children to Ramsey behavior. The parent, on the other hand, takes lump-sum taxes as given and, hence, sees the return as $R_{t+2}(1-\tau_{t+2})$, and wants the child to save optimally according to this return. So, the Ramsey level of saving that the planner makes the offspring save is still too low from the parent's perspective at the interest rate $R_{t+2}(1-\tau_{t+2})$. As a result, the parent still has a motive to transfer more than the Ramsey level. To discourage this, we need a tax on bequests.

We know that if the government is allowed to command offspring to the Ramsey allocation, then there is no need to distort parents' decisions. A natural question that follows, then, is whether the bequest tax result is peculiar to the assumption that the government is restricted to use linear taxes to discipline the offspring's present-bias problem. The proof of Proposition

[17]For a special case of our model economy with partial equilibrium and constant prices, Phelps and Pollak (1968) and Laibson (1994) have shown existence of linear equilibria under CEIS utility.

1 shows that as long as the offspring react to an increase in bequests by increasing their savings, parents will have the extra return to bequeathing and hence will bequeath too much if they are not taxed. Therefore, as long as the government policy leaves offspring's optimal saving policies monotone increasing in the bequests they receive, parents will bequeath too much if they are not taxed. We conclude that the optimality of the bequest tax is not peculiar to the implementation in which there are linear taxes on children's savings; bequest taxes would remain optimal for any tax system that does not eliminate the monotonicity of offspring's optimal saving policies in the amount of bequests they receive.

Finally, note that, implicit in the Markovianity assumption, we do not allow for conditional bequests. In particular, parents are not allowed to condition the payment of the bequest to a specific level of savings by the offspring. The lack of such - potentially welfare improving - arrangements might be justified by the fact that they are difficult to enforce in reality. Recall, indeed, that bequests payments - by definition - occur after parents die and saving decisions might be difficult to monitor, especially by third parties.

5 Constrained Efficiency

In the previous section, we analyzed how wealth transfer behavior in equilibrium compares to the allocation that would arise in absence of present-bias problems. Using this allocation as an efficiency benchmark implicitly assumes that the government has fiscal tools available to discipline present-bias problems of the offspring directly. In this section, we analyze whether it is still optimal to tax bequests when government does not have access to savings taxes to correct offspring's life-cycle saving behavior. Without the ability to affect the saving behavior of the offspring, the government has no way of implementing the Ramsey allocation. Therefore, we consider another efficiency concept where the planner is constrained like the parent in the sense that it cannot directly control offspring's decisions: the only way the planner can affect the saving behavior of the offspring is through the amount of bequests left. Also, much like the parent, the bequest choices of planner have to be Markov in current wealth. We call the solution to this planner's problem *the constrained efficient allocation.* We want to understand whether the equilibrium parental bequest behavior is constrained efficient.

Despite the fact that the parent and the planner agree on preferences and face similar constraints regarding offspring's saving choice, the equilibrium behavior of the parent does not coincide with the constrained efficient allocation and parents need to be taxed. They key observation is that the planner differs from the parents because the planner chooses bequests for all the parents, so it is large enough to affect prices by choosing allocations. Combined with the existence of present-bias problems, this creates a *price externality*, which implies that equilibrium parental transfers are too high compared to their constrained efficient levels. For the class of linear equilibria, we provide a (linear) bequest tax system that implements the constrained efficient allocation and show that the taxes on wealth transfers are positive.

5.1 The Constrained Efficient Allocation

We first define the constrained efficient allocation formally. Letting $W(f(b_t))$ denote the value to the planner that begins period t with $f(b_t)$ units of wealth, the constrained efficient allocation is given by the solution to the following planning problem:

$$W(f(b_t)) = \max_{b_{t+1}, b_{t+2}, c_t, c_{t+1}} u(c_t) + \gamma \left[u(c_{t+1}) + \delta W(f(b_{t+2})) \right],$$

subject to the budget constraints

$$\begin{aligned} c_t &= f(b_t) - b_{t+1}, \\ c_{t+1} &= f(b_{t+1}) - b_{t+2}, \end{aligned}$$

and the constraint that, taking R_{t+2}, w_{t+2} as given, offspring solve the following problem:

$$b_{t+2} \in \arg \max_{\tilde{b}_{t+2} \geq -B(Q_{t+2}),} u(\tilde{c}_{t+1}) + \beta\delta W(R_{t+2}\tilde{b}_{t+2} + w_{t+2}).$$

It is important to recall that the following pricing conditions are constraints in the planner's problem:

$$\begin{aligned} R_t &= f'(b_t), \\ w_t &= f(b_t) - f'(b_t)b_t. \end{aligned} \tag{10}$$

The planner faces incentive constraints because, like the parent, the planner has no direct way of affecting offspring's behavior. The key difference between the problem solved by the planner and the parent's problem is that the later takes prices as given, whereas the planner essentially chooses prices through choosing the allocation. Also, observe that the offspring in period $t+1$ knows that in the next period, the economy will be taken over by a planner. We define $b^{PL}_{t+2}(b_{t+1})$ as the planner's (correct) perception of offspring's period $t+1$ saving as a function of bequests received in the planner's problem.

Throughout the rest of the paper, we denote the constrained efficient allocation by double asterisk.

5.2 Comparing Equilibrium and Constrained Efficient Bequests

In this economy, the planner is like the parent in that it cannot choose children's behavior. Therefore, the planner also faces offspring who are undersaving, which implies that the planner shares parents' willingness to transfer more than the Ramsey level. Still, bequests should be taxed in order to correct a price externality. The idea is that the planner knows that the amount it bequeaths affects the interest rate the offspring faces in the next period, which affects their willingness to save. Increasing the level of intergenerational wealth transfers increases the offspring's willingness to save, which depresses the interest rate they face, which in turn decreases their willingness to save. Since an individual parent is too small to affect prices, parents do not realize the effect of their choice on the interest rates, which makes them transfer too much relative to the constrained efficient level. In technical terms, pricing equations (10) enter as constraints to the planning problem, whereas the parents take as given the prices that offspring face.

Proposition 7 establishes, formally, that parents bequeath too much relative to the constrained-efficient allocation if they are allowed to choose the bequest level in any parenthood period.

Proposition 7 *Suppose $\beta < 1$ and $F_{11} < 0$. Suppose further that W is differentiable. If the parents are allowed to choose bequests in some period t in place of the planner, then the equilibrium choice of bequests in period t is larger than the constrained efficient level. If $\beta = 1$ or $F_{11} = 0$, then the parent chooses the constrained efficient level of bequests.*

Proof. Relegated to the Appendix. ■

Proposition 7 is quite general in the sense that it does not require making any ad hoc differentiability assumptions on the policy functions. In order to guarantee strict monotonicity of the policy, we do, however, assume that the value function W is differentiable, which may not be true in general. However, since the utility function u is continuously differentiable, u is also locally Lipschitz continuous. As a consequence, it is not difficult to show that the value function W is locally Lipschitz continuous.[18] By Rademacher's theorem, W is hence almost everywhere differentiable. Proposition 7 can be shown only using the local Lipschitz continuity of W. The cost is a considerable increase in the technicality of the proof (details are available upon request; see also the brief discussion at the end of the present proof). In addition, it is possible to make the value function smooth by adding enough uncertainty to the problem, for instance, by adding income shocks, as it is done in Harris and Laibson (2001). So, in this

[18]Details are available upon request; see also Montrucchio and Pavoni (2010).

case assuming differentiability of W is without loss of generality. The proof of Proposition 7 is still quite technical, and hence it is relegated to the Appendix. In the next section, we show that, if we assume differentiability of the policy functions regarding offspring's saving behavior, Proposition 7 translates into a positive bequest wedge.

5.3 Constrained Efficient Wedge and Intuition

In this section, we assume differentiability of offspring's saving policy functions in the parent's and planner's problems, $b_{t+2}(\cdot)$ and $b^{PL}_{t+2}(\cdot)$ respectively. We define the constrained efficient bequest wedge by considering the following exercise. Suppose we allow a parent to deviate from the constrained efficient level of bequests in some period t, knowing that in the rest of the periods, parental transfers will be chosen by the planner. The parent takes the prices implied by the constrained efficient allocation, Q^{**}_t, as given and - when endowed with the constrained efficient level of assets b^{**}_t - solves

$$\max_{b_{t+1}} \; u(R^{**}_t b^{**}_t + w^{**}_t - b_{t+1}) + \gamma[u(R^{**}_{t+1} b_{t+1} + w^{**}_{t+1} - b_{t+2}) + \delta W(R^{**}_{t+2} b_{t+2} + w^{**}_{t+2})] \tag{11}$$

subject to the offspring's behavior, which taking Q^{**}_{t+1} as given, solves

$$\max_{b_{t+2}} u(R^{**}_{t+1} b_{t+1} + w^{**}_{t+1} - b_{t+2}) + \beta\delta W(R^{**}_{t+2} b_{t+2} + w^{**}_{t+2}).$$

The parent's bequest decision in this case solves the following optimality condition:

$$u'(c_t) = \gamma u'(c_{t+1}) \left(R^{**}_{t+1} + \frac{\partial b_{t+2}(b_{t+1}, Q^{**}_{t+1})}{\partial b_{t+1}} \left\{ -1 + \frac{1}{\beta} \right\} \right), \tag{12}$$

where the derivative of the policy is evaluated at the parental choice. Note that the price sequence Q^{**}_{t+1} is identical in both the parent's and the planner's problems, since a single parent deviating from the constrained efficient allocation cannot affect the prices.

The *constrained efficient bequest wedge* is defined as:

$$BW^{**}_t = -u'(c^{**}_t) + \gamma u'(c^{**}_{t+1}) \left(f'(b^{**}_{t+1}) + \frac{\partial b_{t+2}(b^{**}_{t+1}, Q^{**}_{t+1})}{\partial b_{t+1}} \left\{ -1 + \frac{1}{\beta} \right\} \right).$$

The constrained efficient bequest wedge in period t measures the constrained efficient distortion that the planner needs to create in the bequest decision of a parent in that period in order to ensure that constrained efficient level of bequests satisfies his optimality condition, (12). A positive (negative) BW^{**}_t means that a parent facing the constrained efficient interest rate, $f'(b^{**}_{t+1})$, would like to increase (decrease) his bequests marginally above the constrained efficient level. Our goal in this section is to sign the bequest wedge.

We now characterize the constrained efficient bequest behavior. Looking at the planning problem that defines the constrained efficient allocation reminds us that the planner is similar to a parent in the sense that he also has to take offspring's decisions as given. Replacing the constraint describing offspring saving with the associated policy function, $b^{PL}_{t+2}(\cdot)$, and taking the first-order optimality condition with respect to the bequest decision of the planner, we characterize the constrained efficient bequest decision as follows:

$$u'(c^{**}_t) = \gamma u'(c^{**}_{t+1}) \left(f'(b^{**}_{t+1}) + \frac{\partial b^{PL}_{t+2}(b^{**}_{t+1})}{\partial b_{t+1}} \left\{ -1 + \frac{1}{\beta} \right\} \right). \tag{13}$$

This condition says that much like the parent, a constrained planner also tends to transfer 'more' relative to the Ramsey allocation, and for the exact same intuitive reason: the offspring is undersaving, which implies an extra return on the planner's transfers, which induces him to transfer above the Ramsey level.

By comparing the planner's optimality condition for bequests, (13), with that of the parent's, (12), one realizes that they are very similar except for one important detail: parent's and the planner's perceptions of how bequests affect the offspring's saving choice, $\frac{\partial b_{t+2}}{\partial b_{t+1}}$ and $\frac{\partial b^{PL}_{t+2}}{\partial b_{t+1}}$, are different. They both know that, at b^{**}_{t+1}, the offspring chooses b_{t+2} according to

$$u'(c_{t+1}) = \beta\delta W'(R^{**}_{t+2} b_{t+2} + w^{**}_{t+2}) R^{**}_{t+2}. \tag{14}$$

However, the planner also knows that he can affect R_{t+2} through his b_{t+1} choice:

$$R_{t+2} = f'(b_{t+2}).$$

The planner calculates that increasing b_{t+1} increases the offspring's wealth next period, making him increase his savings, b_{t+2}. The parent also understands this effect. However, the planner also realizes that there is a second effect of increasing b_{t+1} on b_{t+2} that works in the opposite direction: when all the offspring increase their savings next period, this depresses the interest rate R_{t+2}, which tightens the incentive constraint describing offspring's saving choice, (14), and makes them decrease their savings, b_{t+2}. The parent does not realize this effect of increasing b_{t+1} on the interest rate since the parent - unlike the planner - is small and takes prices as given. This implies that the parent's perception of how much a marginal increase in bequests increases the offspring's savings is strictly higher than the planner's perception:

$$\frac{\partial b_{t+2}(b^{**}_{t+1}, Q^{**}_{t+1})}{\partial b_{t+1}} > \frac{\partial b^{PL}_{t+2}(b^{**}_{t+1})}{\partial b_{t+1}}. \tag{15}$$

Proposition 8 formalizes this intuition.

Proposition 8 *Assume that* $b_{t+2}(\cdot)$ *and* $b_{t+2}^{PL}(\cdot)$ *are differentiable. If both* $F_{11} < 0$ *and* $\beta < 1$ *then condition (15) holds.*

Proof. Relegated to the Appendix. ■

A direct implication of Proposition 8 is Corollary 9 which establishes the sign of the constrained efficient bequest wedge.

Corollary 9 *Assume that* $b_{t+2}(\cdot)$ *and* $b_{t+2}^{PL}(\cdot)$ *are differentiable. If both* $F_{11} < 0$ *and* $\beta < 1$ *then* $BW_t^{**} > 0$ *for all parenthood periods* t.

Proof. Follows directly from (15), (13), and the definition of BW_t^{**}. ■

A positive bequest wedge implies that the parent will keep increasing his bequests beyond the constrained efficient level. Therefore, if a government wants to make a parent choose the constrained efficient level of bequests, it needs to create a wedge between the parental marginal cost and benefit from transferring resources to the offspring.

5.4 Constrained Efficient Taxation

In this section, we want to implement the constrained efficient allocation through a simple tax system on parental wealth transfers. Let the linear tax rate on the returns to bequests of the period t parent be τ_{t+1} and lump-sum taxes be T_{t+1}. Tax proceeds are rebated in a lump-sum manner in every period, so that the government balances its budget period by period. Let $\Upsilon = \{\tau_t, T_t\}_{t\in\{0,2,4,..\}}$ be the sequence of bequest taxes that the government chooses and commits to at the beginning of time. The government chooses Υ to implement the constrained efficient allocation. Let Υ^{**} denote the tax system that implements the constrained efficient allocation. Observe that the offspring cannot be taxed.

Defining as before $\Psi_t := (Q_s, \Upsilon_s)_{s=t}^{\infty}$ and $a_t = R_t b_t + w_t$, the parent's problem under taxes is

$$V(a_t, \Psi_t) = \max_{b_{t+1}\geq -B(\Psi_{t+1})} u(c_t) + \gamma\left[u(c_{t+1}) + \delta V(a_{t+2}, \Psi_{t+2})\right],$$

subject to the budget constraints

$$\begin{aligned} c_t &= a_t - b_{t+1}, \\ c_{t+1}(b_{t+1}, \Psi_t) &= R_{t+1} b_{t+1}(1-\tau_{t+1}) + T_{t+1} + w_{t+1} - b_{t+2}(b_{t+1}, \Psi_t), \\ a_{t+2}(b_{t+1}, \Psi_t) &:= R_{t+2} b_{t+2}(b_{t+1}, \Psi_t) + w_{t+2}, \end{aligned}$$

where the offspring policy solves

$$b_{t+2}(b_{t+1},\Psi_t)=\arg\max_{\tilde{b}_{t+2}\geq -B(\Psi_{t+2})} u(R_{t+1}b_{t+1}(1-\tau_{t+1})+T_{t+1}+w_{t+1}-\tilde{b}_{t+2})+\beta\delta V(R_{t+2}\tilde{b}_{t+2}+w_{t+2},\Psi_{t+2}),$$

and in equilibrium

$$\begin{aligned} R_t &= f'(b_t), \\ w_t &= f(b_t)-f'(b_t)b_t. \end{aligned}$$

The government uses lump-sum taxes to rebate back what it receives through linear taxes on savings:

$$T_{t+1}=R_{t+1}\tau_{t+1}b_{t+1}.$$

Again, due to the existence of the constraint that describes offspring saving behavior, agents' problems are not necessarily concave. Therefore, showing that the first-order optimality conditions are satisfied by the constrained efficient allocation under a tax system does not guarantee that the tax system implements the constrained efficient allocation. Therefore, we restrict attention to Markov equilibria with policy functions that are linear in current wealth. This restriction guarantees that the agents' problems are concave. Remember that Proposition 6 guarantees that when utility function is of the CEIS form and equilibrium exists, then there exist at least one linear Markov equilibrium. Hence, we have the following implementation result.

Proposition 10 *Suppose Markov equilibrium optimal policies are linear in current wealth. Then, there is a linear tax system that implements the constrained efficient allocation. In this system, optimal bequest taxes* τ^{**}_{t+1} *solve:*

$$u'(c^{**}_t)=\gamma u'(c^{**}_{t+1})\left(R^{**}_{t+1}(1-\tau^{**}_{t+1})+M_{t+2}(\Psi^{**}_{t+1})\left\{-1+\frac{1}{\beta}\right\}\right),$$

where

$$M_{t+2}(\Psi^{**}_{t+1})=\frac{\partial b_{t+2}\left(b^{**}_{t+1},\Psi^{**}_{t+1}\right)}{\partial b_{t+1}}$$

is the coefficient of offspring's (linear) policy function under the constrained efficient tax system and prices implied by the constrained efficient allocation.

Proof. With linear policy functions, agents' problems are concave, which implies that the parent's first-order optimality conditions are necessary and sufficient to implement an allocation. The tax expression is obtained from the parent's optimality conditions evaluated at the constrained efficient allocation. ■

Proposition 10 defines the optimal bequest tax only implicitly. The sign of the wedge we establish in Corollary 9 intuitively suggests that the tax on bequests should be positive by comparing parents' perception of how offspring's saving behavior changes with bequests under laissez-faire equilibrium, $\frac{\partial b_{t+2}(b^{**}_{t+1},Q^{**}_{t+1})}{\partial b_{t+1}}$, with planner's perception of it in the planning problem, $\frac{\partial b^{PL}_{t+2}(b^{**}_{t+1})}{\partial b_{t+1}}$. However, establishing the sign of optimal bequest taxes requires more: one needs to compare parents' perception of how offspring's saving behavior changes with bequests under optimal taxes, $\frac{\partial b_{t+2}(b^{**}_{t+1},\Psi^{**}_{t+1})}{\partial b_{t+1}}$, to planner's perception, $\frac{\partial b^{PL}_{t+2}(b^{**}_{t+1})}{\partial b_{t+1}}$. The absence of wealth effects allows us to make this comparison and establish the sign of optimal bequest taxes for the case of logarithmic utility.

Proposition 11 *Suppose utility function is logarithmic, Markov equilibrium (with taxes) policies are linear in current wealth, and the constrained efficient policy $b^{PL}_{t+2}(\cdot)$ is differentiable. Then, the constrained efficient tax on bequests is strictly positive: $\tau^{**}_{t+1} > 0$.*

Proof. Relegated to the Appendix. ■

Observe that the concavity of the production function, together with the endogeneity of the interest rates, is crucial for the optimality of positive bequest taxes when government cannot correct children's life-cycle saving decisions. This is in contrast to the optimality of bequest taxation that we analyze in Section 4 where government can tax life-cycle savings and is able to implement the Ramsey allocation. There, taxing bequests is efficient also under partial equilibrium and with linear production function.

6 Quantitative Significance of Our Mechanism

Our main result states that, as long as children - the recipients of intergenerational transfers - face present-bias problems, parents transfer too much, and hence to restore efficiency, bequests should be taxed.

In this section we aim to shed some light on the quantitative importance of our mechanism for bequest taxation. Estate tax rates in the United States have been varying since the time they were introduced in 1916.[19] Recently, in 2013, after the expiration of the Bush tax cuts, the estate taxes increased from 35% to 55%. In the United Kingdom, the inheritance tax rate

[19]For detailed information on the evolution of estate taxes in the United States, see Jacobson, Raub, and Johnson (2007).

has held steady at 40%.[20] We want to understand whether our mechanism has a quantitative potential of providing a rationale for such observed levels of tax rates on bequests. To do so, we compute optimal bequest taxes that implement the Ramsey allocation for a parameterized version of our economy. We focus on *logarithmic utility* which allows us to find closed-form solutions for optimal bequest taxes. We also assume that time discount factor is equal to the altruism factor. The optimal bequest taxes are given by the following formula:[21]

$$1 - \tau_{t+1} = 1 - \delta + \beta\delta,$$

where δ and β represent the true discount factor and the degree of the present-bias problem that offspring face in the period in which they receive bequests, respectively.

To conduct a numerical analysis, we must choose particular values for the parameters of the model. Observe that the tax formula does not depend on the shape of the production function, F, or the depreciation rate, θ. So, we do not specify values for these parameters. The only parameters that are needed are the true yearly discount factor δ and the present-bias parameter β. We set $\delta = 0.96$ and $\beta = 0.81$ following Laibson, Repetto, and Tobacman (2007)'s benchmark estimates when utility is logarithmic. We then vary β between 0.5 and 0.9 to cover values of β that are estimated by Laibson, Repetto, and Tobacman (2007) in their robustness analysis. The following table summarizes the optimal bequest tax rates for different values of β.

β	0.5	0.6	0.7	0.81	0.9
bequest tax	48%	38.4%	28.8%	18.3%	9.6%

We conclude that the optimal taxes generated by our mechanism might be quantitatively significant and might have the potential of justifying the actual taxes on bequests.

7 Variations and Extensions

In this section, we discuss a variation and two extensions of our model that show the generality of our results. For the sake of brevity, throughout this section, we assume differentiability of

[20] Both countries have an exemption level below which wealth transfers go untaxed. In the United States, this level has decreased from 5 million US dollars to 1 million US dollars in 2013. In the United Kingdom, this amount is 325,000 pounds. For detailed information on the UK inheritance tax system, visit https://www.gov.uk/inheritance-tax.

[21] An added bonus of adopting a log-utility function is that this same formula gives the optimal bequest tax both in our benchmark model and in the more general model of Section 7.2 with multiple periods. Details on the derivation and computation of the optimal bequest tax for the log utility case are available upon request.

the policy functions and focus on Ramsey wedges.

7.1 Naive Agents

In the main body of the paper, we assume that agents are sophisticated in the sense that they are aware of their descendants' present-bias problems. In this section, we analyze whether the optimality of transfer taxation depends on the assumption of sophistication. To do so, *we assume that agents are naive, meaning that they are not aware of their descendant's present-bias problems.* We show that even naive parents bequeath too much in equilibrium relative to the Ramsey level and, hence, parental transfers should be taxed.

The problem of the naive parent is given by:

$$\bar{V}\left(a_t, Q_t\right) = \max_{b_{t+1} \geq -B(Q_{t+1})} u(c_t) + \gamma\left[u\left(\bar{c}_{t+1}(b_{t+1}, Q_{t+1})\right) + \delta \bar{V}\left(\bar{a}_{t+2}(b_{t+1}, Q_{t+1}), Q_{t+2}\right)\right],$$

subject to the budget constraints and the definition of wealth

$$\begin{aligned} c_t &= a_t - b_{t+1}, \\ \bar{c}_{t+1}(b_{t+1}, Q_{t+1}) &= R_{t+1} b_{t+1} + w_{t+1} - \bar{b}_{t+2}(b_{t+1}, Q_{t+1}), \\ \bar{a}_{t+2}(b_{t+1}, Q_{t+1}) &:= R_{t+2} \bar{b}_{t+2}(b_{t+1}, Q_{t+1}) + w_{t+2}, \end{aligned}$$

together with the condition defining the policy of the offspring:

$$\bar{b}_{t+2}(b_{t+1}, Q_{t+1}) = \arg \max_{\tilde{b}_{t+2} \geq -B(Q_{t+2})} u\left(R_{t+1} b_{t+1} + w_{t+1} - \tilde{b}_{t+2}\right) + \delta \bar{V}(R_{t+2} \tilde{b}_{t+2} + w_{t+2}, Q_{t+2}). \quad (16)$$

Observe that this problem is identical to the problem of the sophisticated parent, (1), except that in the condition defining the policy of the offspring, (16), the offspring discounts future with δ instead of $\beta\delta$. This reflects the assumption that parents naively believe that their children have no present-bias problems. We denote the naive parents' value function by $\bar{V}\left(a_t, Q_t\right)$. The naive value function gives the value of the solution to a standard dynamic programming problem with exponential discounters. It is well-known that since u is concave, this value function is *concave* as well. Note that the saving policy of the offspring in this problem is incorrect; it simply represents the *naive belief* of the parent about the saving policy of the offspring. To distinguish the parent's belief about the policy and the actual policy of the offspring in equilibrium, we denote the former by $\bar{b}_{t+2}(\cdot)$. Similarly, $\bar{c}_{t+1}(\cdot)$ and $\bar{a}_{t+2}(\cdot)$ refer to the naive belief of the parent about the offspring's period $t+1$ consumption and beginning of period $t+2$ wealth policies.

Now, we consider the saving decision of the offspring in equilibrium. The offspring solves

$$\max_{\tilde{b}_{t+2}\geq -B(Q_{t+2})} u\left(R_{t+1}b_{t+1}+w_{t+1}-\tilde{b}_{t+2}\right)+\beta\delta\bar{V}(R_{t+2}\tilde{b}_{t+2}+w_{t+2},Q_{t+2}).$$

Observe that the naive offspring also faces the naive value function $\bar{V}$ since his naive beliefs about how his descendants will allocate consumption over time is in line with that of his parent.

We are now ready to establish the optimality of bequest tax. The parent naively believes that the following first-order condition describes offspring saving behavior

$$u'\left(\bar{c}_{t+1}(b_{t+1},Q_{t+1})\right)=\delta R_{t+2}\bar{V}_1(\bar{a}_{t+2}(b_{t+1},Q_{t+1}),Q_{t+2}).$$

Taking the first-order optimality condition with respect to bequests in the naive parent's problem and using the naive optimality condition for offsprings above, we get that optimal bequest decision solves:

$$u'(c_t)=\gamma R_{t+1}u'\left(\bar{c}_{t+1}(b_{t+1},Q_{t+1})\right).$$

We define the *naive bequest wedge* as follows:

$$\overline{BW}_t=-u'(c_t^*)+\gamma R_{t+1}^*u'\left(\bar{c}_{t+1}(b_{t+1}^*,Q_{t+1}^*)\right).$$

where, $\bar{c}_{t+1}$ corresponds to what the naive parent believes the offspring will choose if he receives b_{t+1}^* as bequests and faces the price sequence implied by the Ramsey allocation, Q_{t+1}^*. We want to show that the naive bequest wedge is strictly positive, meaning given the Ramsey allocation, the naive parent would like to increase bequests. Remember that in the Ramsey allocation, we have

$$-u'(c_t^*)+\gamma R_{t+1}^*u'\left(c_{t+1}^*\right)=0.$$

Therefore, in order to establish the sign of $B\bar{W}_t$, we need to compare the Ramsey level of consumption of the offspring, c_{t+1}^*, with the parent's belief about how much the offspring will consume: $\bar{c}_{t+1}(b_{t+1}^*,Q_{t+1}^*)$.

Now, if the offspring chooses the Ramsey allocation in equilibrium, say thanks to the tax τ_{t+2}^*, then his first-order condition reads

$$u'\left(c_{t+1}^*\right)=\beta\delta R_{t+2}^*(1-\tau_{t+2}^*)\bar{V}_1(a_{t+2}^*,Q_{t+2}^*).$$

The parent - on the other hand - believes the offspring will behave according to

$$u'\left(\bar{c}_{t+1}(b_{t+1}^*,Q_{t+1}^*)\right)=\delta R_{t+2}^*(1-\tau_{t+2}^*)\bar{V}_1(\bar{a}_{t+2}(b_{t+1}^*,Q_{t+1}^*),Q_{t+2}^*).$$

Concavity of u and $\bar{V}$ then imply that

$$\bar{c}_{t+1}(b^*_{t+1}, Q^*_{t+1}) < c^*_{t+1}.$$

This ordering is intuitive: an agent with present-bias problems tends to consume more than an agent without present-bias problems, and the parent naively believes that the offspring has no present-bias problems. The direct implication of this is the following result:

Proposition 12 *Assume agents are naive and $\beta < 1$, then the naive bequest wedge is positive.*

Proof. From previous discussion, we have $\overline{BW}_t = -u'(c^*_t) + \gamma R^*_{t+1} u'\left(\bar{c}_{t+1}(b^*_{t+1}, Q^*_{t+1})\right) > -u'(c^*_t) + \gamma R^*_{t+1} u'\left(c^*_{t+1}\right) = 0.$ ■

Therefore, we conclude that even when agents are naive, parents bequeath too much relative the Ramsey level and hence should be taxed. The intuition for the bequest tax result when agents are naive is different from the case with sophisticated agents though. In the naive case, parents do not bequeath too much relative to the Ramsey level because they want to compensate for their offspring's undersavings. They bequeath too much because their perceived marginal return to bequeathing, $u'(\bar{c}_{t+1}(b^*_{t+1}, Q^*_{t+1}))$ is larger than the actual marginal return, $u'(c^*_{t+1})$.

One can further show that, if we restrict attention to linear equilibria, then we can implement the Ramsey allocation via taxes, and the optimal tax on bequests is given by

$$\tau^*_{t+1} = 1 - \frac{u'(c^*_{t+1})}{u'(\bar{c}_{t+1}(b^*_{t+1}, Q^*_{t+1}))} > 0.$$

7.2 Longer Life Cycle, Coexistence, and Inter Vivos Taxation

We assume that people live for two periods in our benchmark model. In this section, we extend our model by allowing each agent within a dynasty to be active for $I + 1$ periods: in the first I periods, agents make consumption saving decisions facing present-bias problems. In the last period of their lives, parents coexist with their offspring who are already in the first period of young adulthood. Parents decide how much to consume and transfer to their offspring. Transfers can be made in two ways: *inter vivos transfers* are received by the offspring during the coexistence period, and bequests are received at the beginning of the next period, after the parent dies. We show that the longer life cycle does not alter the main result, that is, bequests should be taxed. Furthermore, thanks to the coexistence period in the extended model, we are able to analyze parental inter vivos transfer behavior and establish the optimality of *taxing inter vivos transfers as well.*

Consider any calendar year t in which there is a parent who is in the last period of his life. His preference over dynastic allocation is given by

$$V_t = u(c_t^o) + \gamma \left[u(c_t) + \delta u(c_{t+1}) + .. + \delta^{I-1} u(c_{t+I-1}) + \delta^I V_{t+I} \right],$$

where V_t represents the dynastic welfare of the parent who is in the last period of his life in period t and V_{t+I} represents that of the offspring in his terminal period, $t+I$. The term c_t^o is the last period consumption of the parent and c_t is the consumption level of the offspring who is at age 1 in period t. Observe that c_t^o and c_t occur in the same period. To keep aggregate labor supply constant across periods, we assume that only the offspring has one unit of time endowment in the period of coexistence. Clearly, this assumption is not material for any of our results.

We adopt the quasi-hyperbolic framework. The offspring's preference in period i of his life is

$$u(c_{t+i-1}) + \beta_i \delta \left[\sum_{j=i+1}^{I} \delta^{j-(i+1)} u(c_{t+j-1}) + \delta^{I-i} V_{t+I} \right], \text{ for } 1 \leq i \leq I-1$$
$$u(c_{t+I-1}) + \beta_I \delta V_{t+I}.$$

When $\beta_i = 1$ for all i, people are fully time-consistent at all ages; there is no present-bias problem. Whenever $\beta_i < 1$ for some i, people face present-bias problems at age i. Observe that we are extending the hyperbolic discounting model by allowing for the existence (and severity) of the present-bias problem, β_i, to depend on age.[22] We extend the standard model of present-bias in order to show that our transfer taxation results do not depend on how the degree of present-bias evolves over the life cycle. As we will see, the only assumption needed in order to establish optimality of transfer taxation is that the offspring has a present-bias problem in the period he receives the transfer.

Let d_t and b_{t+1}^o denote the inter vivos transfers and bequests made by the parent who is in his last period of life in period t. Let b_{t+i} denote the offspring's age i saving level.

The parent, whose wealth level is $a_t = R_t b_t + w_t$, solves

$$V(a_t, Q_t) = \max_{b_{t+1}^o, d_t} u(c_t^o) + \gamma \left[\sum_{i=0}^{I-1} \delta^i u(c_{t+i}) + \delta^I V(a_{t+I}, Q_{t+I}) \right],$$

[22] If we were to take $\beta_i = \beta$ for all i, as previous papers have assumed, that would mean that the degree of present-bias problem is constant as people age, and all our results hold in this special case.

subject to the budget constraints[23]

$$
\begin{aligned}
c_t^o &= R_t b_t - b_{t+1}^o - d_t, \\
c_t(d_t, Q_t) &= d_t + w_t - b_{t+1}(d_t, Q_t), \\
c_{t+1}(d_t, b_{t+1}^o, Q_t) &= R_{t+1} b_{t+1}(d_t, Q_t) + w_{t+1} + R_{t+1} b_{t+1}^o - b_{t+2}(d_t, b_{t+1}^o, Q_t), \\
c_{t+i-1}(d_t, b_{t+1}^o, Q_t) &= R_{t+i-1} b_{t+i-1}(d_t, b_{t+1}^o, Q_t) + w_{t+i-1} - b_{t+i}(d_t, b_{t+1}^o, Q_t), \text{ for } 3 \leq i \leq I, \\
a_{t+I}(d_t, b_{t+1}^o, Q_t) &= R_{t+I} b_{t+I}(d_t, b_{t+1}^o, Q_t),
\end{aligned}
$$

and subject to the constraints defining the policy functions of future selves:[24]

$$
\begin{aligned}
b_{t+1}(d_t, Q_t) &= \arg\max \ u(c_t) + \beta_1 \delta \left[u(c_{t+1}) + ... + \delta^{I-1} u(c_{t+I-1}) + \delta^I V(a_{t+I}, Q_{t+I}) \right] \\
& \quad s.t. \\
b_{t+2}(d_t, b_{t+1}^o, Q_t) &= \arg\max \ \ u(c_{t+1}) + \beta_2 \delta \left[u(c_{t+2}) + ... + \delta^{I-2} u(c_{t+I-1}) + \delta^{I-1} V(a_{t+I}, Q_{t+I}) \right] \\
& \quad s.t. \\
& \quad ...
\end{aligned}
$$

$$
\begin{aligned}
& \quad ... \\
& \quad s.t. \\
b_{t+I-1}(d_t, b_{t+1}^o, Q_t) &= \arg\max \ u(c_{t+I-2}) + \beta_{I-1} \delta \left[u(c_{t+I-1}) + \delta V(a_{t+I}, Q_{t+I}) \right] \\
& \quad s.t \\
b_{t+I}(d_t, b_{t+1}^o, Q_t) &= \arg\max \ u(c_{t+I-1}) + \beta_I \delta V(a_{t+I}, Q_{t+I})).
\end{aligned}
$$

To better understand the notation and the nature of policy functions in the planning problem, first observe that when an agent of age i in calendar year $t+i-1$ is deciding b_{t+i}, he sees the direct effect of his choice on the next period self's saving choice. Clearly, agent i's saving choice indirectly affects his saving at age $n > i$, b_{t+n}, since b_{t+i} choice affects b_{t+i+1} choice, which then affects b_{t+i+2} choice, and so on. The dependence of b_{t+n} on b_{t+i} can be

[23]The real interest and wage rates are given by marginal products of capital and labor as in the benchmark model. The only difference is for the period right after coexistence, total capital stock in the economy is equal to the sum of the offspring's savings in the coexistence period and parental bequests

$$k_t = b_{t+1} + b_{t+1}^o.$$

[24]Again, the notation implicitly assumes policies are single valued. The usual caveat applies: whenever we have multiple solutions, the policies should be interpreted as selections from the policy correspondences.

described by a (nested) function $b_{t+n}(b_{t+n-1}(...b_{t+i+1}(b_{t+i}, Q_{t+i})...), Q_{t+n-1})$. Without creating a confusion, we can denote it as a single function of only (b_{t+i}, Q_{t+i}). Going backwards, the 'initial' choices in our recursion are d_t and b^o_{t+1}. As a consequence, the vector of functions $[b_{t+1}(d_t, Q_t), b_{t+2}(d_t, b^o_{t+1}, Q_t)..., b_{t+I}(d_t, b^o_{t+1}, Q_t)]$ denote the functions that describe how the offspring's savings choices over the life cycle depend on parental transfers they receive, (d_t, b^o_{t+1}). Observe that offspring's saving at age 1 in year t, denoted by b_{t+1}, only depends on inter vivos transfers she receives in that period but not on the level of bequests, b^o_{t+1}, since we assume that these people receive bequests only after their parents die, in year $t+1$.

We now derive a marginal condition that characterizes equilibrium inter vivos behavior assuming differentiability of the policy functions that describe offspring saving behavior at different ages.

Consider a parent's problem under laissez-faire of choosing d_t and b^o_{t+1} subject to the flow budget constraints and the offsprings' policy functions. Let $\frac{\partial b_{t+i}(d_t, b^o_{t+1}, Q_t)}{\partial d_t}$ represent how an increase in inter vivos transfers affects savings at age i calendar year $t+i-1$. For notational simplicity, we will write this partial derivative as $\frac{\partial b_{t+i}}{\partial d_t}$ whenever doing so does not create a confusion. The parent's first-order optimality condition with respect to the inter vivos decision is

$$u'(c^o_t) = \gamma \left(u'(c_t) \left[1 - \frac{\partial b_{t+1}}{\partial d_t} \right] + \sum_{i=1}^{I-1} \delta^{i-1} u'(c_{t+i}) \frac{\partial c_{t+i}}{\partial d_t} + \delta^I V_1(a_{t+I}, Q_{t+I}) R_{t+I} \frac{\partial b_{t+I}}{\partial d_t} \right), \quad (17)$$

where

$$\frac{\partial c_{t+i}}{\partial d_t} = \left[R_{t+i} \frac{\partial b_{t+i}}{\partial d_t} - \frac{\partial b_{t+i+1}}{\partial d_t} \right] \quad (18)$$

and, clearly, the derivatives are evaluated at the equilibrium allocation and prices. Consider the problem of the offspring in the first period of his adult life, in period t. When choosing his level of savings, b_{t+1}, this agent faces the policy functions $(b_{t+2}(b_{t+1}), ..., b_{t+I}(b_{t+1}))$. Taking first-order condition for his period t saving, we get

$$u'(c_t) = \beta_1 \delta \Big\{ \sum_{i=1}^{I-1} \delta^{i-1} u'(c_{t+i}) \frac{\partial c_{t+i}}{\partial b_{t+1}} + \delta^{I-1} V_1(a_{t+I}, Q_{t+I}) R_{t+I} \frac{\partial b_{t+I}}{\partial b_{t+1}} \Big\}, \quad (19)$$

where, again, the derivatives are evaluated at the equilibrium allocation and prices. Using (19) in parental optimality condition for inter vivos, (17), and the fact that the only way d_t affects decisions from $t+1$ onward is through its effect on period t decisions, meaning

$$\frac{\partial b_{t+i}}{\partial d_t} = \frac{\partial b_{t+i}}{\partial b_{t+1}} \frac{\partial b_{t+1}}{\partial d_t},$$

we get the following proposition which describes equilibrium parental inter vivos behavior under differentiability of policy functions.

Proposition 13 *Suppose the policy functions that describe offspring behavior over the life cycle are differentiable. The equilibrium inter vivos behavior is characterized by*

$$u'(c_t^o) = \gamma \left(u'(c_t) + \frac{\partial b_{t+1}(d_t, Q_t)}{\partial d_t} u'(c_t) \left\{ -1 + \frac{1}{\beta_1} \right\} \right). \tag{20}$$

Condition (20) is analogous to (8), the optimality condition for bequests in the benchmark model under differentiability. The left-hand side is the marginal cost of increasing inter vivos transfers, which equals the utility loss from forgone parental consumption. The first term on the right-hand side is the usual marginal benefit of increasing saving – the utility gain from increased consumption in the period during which returns to savings are received. The second term on the right-hand side summarizes how increasing inter vivos transfers affects parental welfare by affecting future consumption levels of the offspring. As long as $\beta_1 < 1$, meaning that the offspring faces present-bias problems in his first period of life, this term is positive: there is an additional marginal benefit of increasing transfers for the parent. Intuitively, the parent, being fully sophisticated, knows that from his perspective, self 1 is undersaving. So, parental welfare increases if the parent can make self 1 increase period 1 savings, which is possible by increasing inter vivos transfers since $\frac{\partial b_{t+1}(d_t)}{\partial d_t} > 0$.

To see the implication of Proposition 13 for Ramsey inter vivos taxation, first observe that in the Ramsey allocation we have

$$u'\left(c_t^{o*}\right) = \gamma u'\left(c_t^*\right). \tag{21}$$

Now define the *Ramsey inter vivos wedge* as:

$$IW_t^* = -u'(c_t^{o*}) + \gamma \left(u'(c_t^*) + \frac{\partial b_{t+1}(d_t^*, Q_t^*)}{\partial d_t} u'(c_t^*) \left\{ -1 + \frac{1}{\beta_1} \right\} \right).$$

The Ramsey inter vivos wedge in period t measures the distortion that the planner needs to create in the inter vivos decision of a parent in that period in order to make him choose the Ramsey level of inter vivos transfers. A positive (negative) IW_t means that a parent would like to increase (decrease) his inter vivos transfers marginally above (below) the Ramsey level if there is no government intervention.

Corollary 14 *Suppose the policy functions that describe offspring behavior over the life cycle are differentiable. Then, $IW_t^* > 0$ for all parenthood periods t.*

Proof. Follows directly from (21), the definition of IW_t^*, and that $\frac{\partial b_{t+1}(d_t^*,Q_t^*)}{\partial d_t} > 0$, which follows from Lemma 16 in the Appendix. ■

Corollary 14, which is analogous to Corollary 4 for the case of bequests in the main text, establishes that parents would increase their inter vivos transfers above the Ramsey level if there is no government intervention. One can further show that an analog to Proposition 5 also holds for the environment with multiperiod life cycle and coexistence: under the assumption that (Markov) equilibrium (with linear taxes) optimal policies are linear in current wealth, there is a linear tax system that implements the Ramsey allocation, and in this tax system, taxes on inter vivos is strictly positive.

The argument for positive bequest taxation in a multiperiod life-cycle environment is identical to the case for inter vivos taxation and therefore will be omitted for the sake of brevity. One can show optimality of bequest taxation by plugging the first-order optimality condition of offspring for period 2 savings, b_{t+2}, into the first-order optimality condition of bequests.

7.3 Long-Term Assets and Liquidity Constraints

The front-loading of consumption by offspring relative to what their parents and their long-term selves prefer lies at the heart of our transfer taxation results. A natural question then is: if the parents have access to assets with more than one-period maturities, can they force their offspring into the consumption patterns they want by carefully choosing the portfolio of these assets? This is an important question because if they can, then parents solve their children's present-bias problems without increasing transfers above the Ramsey level, which means there is no need to tax transfers based on present-bias problems.

First, observe that if the offspring are not liquidity constrained, meaning that they can borrow as much as they want within their natural borrowing limits, then the timing of transfers cannot constrain their consumption patterns at all. In this case, the strategy of using long-term assets is fruitless in disciplining offspring's saving behavior, and we are back at the benchmark environment without long-term assets: both bequests and inter vivos transfers should be taxed.[25] The real question, then, is what happens if parents have access to long-term assets *and* children face liquidity constraints? This is the question we pick up in this section. The main conclusion of this section is that *if the parents do not have a perfect portfolio of assets that allows them to target transfers to each and every period in the life*

[25]For a formal analysis of this claim, see Pavoni and Yazici (2012) section 11.E.

cycle of an offspring or if the liquidity constraints that offspring face in all periods are not tight enough, then parents will still transfer too much through some of those assets, and hence, these transfers should be taxed.

To see this, consider the case where offspring face liquidity constraints over their life cycle. Observe that we do not change the physical environment but do change the market structure in which people live. Thus, the Ramsey and constrained efficient allocations are exactly the same as before. We analyze the implementation of these allocations via taxation in an incomplete markets equilibrium with liquidity constraints and long-term assets. To make things simple, for now assume that people cannot borrow at all. Suppose there is an illiquid asset that the parent can use to transfer resources directly to period $t+2$. Let b^o_{t+2} denote this asset where the subscript refers to the period in which the child receives the bequest. Suppose the return to this asset is $R_{t+1}R_{t+2}$. In this environment, a parent has three different means of transferring resources to his offspring: inter vivos, d_t, liquid bequests, b^o_{t+1}, and illiquid bequests, b^o_{t+2}. To clarify the timing of transfers, let us write down the budget constraints:

$$\begin{aligned}
c^o_t &= R_t b_t - d_t - b^o_{t+1} - b^o_{t+2}, \\
c_t &= d_t + w_t - b_{t+1}, \\
c_{t+1} &= R_{t+1}b_{t+1} + w_{t+1} + R_{t+1}b^o_{t+1} - b_{t+2}, \\
c_{t+2} &= R_{t+2}b_{t+2} + w_{t+2} + R_{t+1}R_{t+2}b^o_{t+2} - b_{t+3}, \\
c_{t+i-1} &= R_{t+i-1}b_{t+i-1} + w_{t+i-1} - b_{t+i}, \text{ for } 4 \le i \le I.
\end{aligned}$$

Let $[b_{t+1}(d_t), b_{t+2}(d_t, b^o_{t+1}, b^o_{t+2})..., b_{t+I}(d_t, b^o_{t+1}, b^o_{t+2})]$ denote the functions that describe how the offspring's savings choices over the life cycle depend on parental transfers they receive, where we simplified notation by ignoring the dependence on the sequence of prices. Note that a change in illiquid bequests affects the offspring's saving decisions in periods $t+1$ and $t+2$, but not his saving decision in period t. This is due to our assumption that children learn the levels of both liquid and illiquid bequests in period t, after they make their period t savings decision.

For the sake of argument, suppose that people do not have any labor income, meaning $w_t = 0$ for all t. In this case, one can prove that the parents will use these three assets to keep children borrowing constrained in periods t and $t+1$. The intuition is roughly as follows. Suppose the child is not constrained in period t in equilibrium. This means his optimal saving level b_{t+1} is given by (19), meaning he saves according to his present-bias problem. Now, if the parent increases his bequests and decreases his inter vivos transfers in a way that keeps the total amount of transfers unchanged, the child will have to decrease his period t savings to keep his preferred allocation. If the parent keeps doing this back-loading of transfers, there will be a point at which the child will deplete his savings completely and will want to borrow, meaning that he will be borrowing constrained. From this point onward, back-loading of transfers strictly increases parental welfare, and the parent does this until the equilibrium level of the child's savings is exactly in line with that of the parent:

$$u'(c_t) = \delta\Big\{\sum_{i=1}^{I-1}\delta^{i-1}u'(c_{t+i})\frac{\partial c_{t+i}}{\partial b_{t+1}} + \delta^{I-1}V_1(a_{t+I}, Q_{t+I})R_{t+I}\frac{\partial b_{t+I}}{\partial b_{t+1}}\Big\}, \tag{22}$$

where the term $\frac{\partial c_{t+i}}{\partial b_{t+1}}$ is defined analogous to (18).

Identically, one can show that the parent can use liquid and illiquid bequests to keep the children borrowing constrained regarding period $t+1$ savings. In this case, the children's Euler equation reads:

$$u'(c_{t+1}) = \delta\Big\{\sum_{i=2}^{I-1}\delta^{i-2}u'(c_{t+i})\frac{\partial c_{t+i}}{\partial b_{t+2}} + \delta^{I-2}V_1(a_{t+I}, Q_{t+I})R_{t+I}\frac{\partial b_{t+I}}{\partial b_{t+2}}\Big\}. \tag{23}$$

So, by using the timing of transfers, the parent is able to force the offspring to the optimal allocation and get rid of the first and second period incentive constraints. This implies that the parent does not have any motive to transfer more than what is optimal, so the optimal tax on inter vivos and liquid bequests would be zero. For the case of inter vivos transfers, this can be seen by plugging (22) into the parental optimality condition for inter vivos transfers, equation (17). For bequests, it can be shown by plugging (23) into the parental optimality condition for bequests.

We now argue that even in this case, it is optimal to tax parental transfers: not inter vivos or bequests but illiquid bequests. Notice that the parent does not have a long-term asset that pays in period $t+3$. As a result, the offspring will save according to his present-bias problem in period $t+2$ and choose b_{t+3} to satisfy

$$u'(c_{t+2}) = \beta_3\delta\Big\{\sum_{i=3}^{I-1}\delta^{i-3}u'(c_{t+i})\frac{\partial c_{t+i}}{\partial b_{t+3}} + \delta^{I-3}V_1(a_{t+I}, Q_{t+I})R_{t+I}\frac{\partial b_{t+I}}{\partial b_{t+3}}\Big\}. \tag{24}$$

Now consider the optimality condition regarding illiquid bequests, which is given by

$$(b^o_{t+2}) : u'(c^o_t) = \gamma\delta\left(\begin{array}{c} u'(c_{t+1})\left[-\frac{db_{t+2}}{db^o_{t+2}}\right] + \delta u'(c_{t+2})\left[R_{t+2}\frac{db_{t+2}}{db^o_{t+2}} + R_{t+1}R_{t+2} - \frac{db_{t+3}}{db^o_{t+2}}\right] \\ + \sum_{i=3}^{I-1}\delta^{i-1}u'(c_{t+i})\frac{\partial c_{t+i}}{\partial b^o_{t+2}} + \delta^{I-1}V_1(a_{t+I}, Q_{t+I})R_{t+I}\frac{\partial b_{t+I}}{\partial b^o_{t+2}}\end{array}\right).$$

Observe that a change in illiquid bequests affects b_{t+3} both directly and indirectly through its effect on b_{t+2} :

$$\frac{db_{t+3}}{db^o_{t+2}} = \frac{\partial b_{t+3}}{\partial b^o_{t+2}} + \frac{\partial b_{t+3}}{\partial b_{t+2}}\frac{db_{t+2}}{db^o_{t+2}}. \tag{25}$$

Now, plugging (25) into the parent's first-order condition for illiquid bequests and rearranging, we get

$$u'(c^o_t) = \gamma\left(R_{t+1}R_{t+2}\delta^2 u'(c_{t+2}) + \frac{\partial b_{t+2}}{\partial b^o_{t+2}}\Delta_{t+1} + \frac{\partial b_{t+3}}{\partial b^o_{t+2}}\Delta_{t+2}\right), \tag{26}$$

where

$$\Delta_{t+1} = \left[-u'(c_{t+1}) + \delta\left(\sum_{i=2}^{I-1}\delta^{i-2}u'(c_{t+i})\frac{\partial c_{t+i}}{\partial b_{t+2}} + \delta^{I-2}V_1(a_{t+I}, Q_{t+I})R_{t+I}\frac{\partial b_{t+I}}{\partial b_{t+2}}\right)\right], \tag{27}$$

$$\Delta_{t+2} = \left[-u'(c_{t+2}) + \delta\left(\sum_{i=3}^{I}\delta^{i-3}u'(c_{t+i})\frac{\partial c_{t+i}}{\partial b_{t+3}} + \delta^{I-3}V_1(a_{t+I}, Q_{t+I})R_{t+I}\frac{\partial b_{t+I}}{\partial b_{t+3}}\right)\right]. \tag{28}$$

Using (23) in (27) and the offspring's optimality condition for period $t+2$ savings, (24), in (28), we get

$$\begin{aligned}\Delta_{t+1} &= 0,\\ \Delta_{t+2} &= \delta u'(c_{t+2})\left[-1 + \frac{1}{\beta_3}\right].\end{aligned}$$

Plugging these in (26), we get

$$u'(c^o_t) = \gamma\left(R_{t+1}R_{t+2}\delta^2 u'(c_{t+2}) + \frac{\partial b_{t+3}}{\partial b^o_{t+2}}\delta u'(c_{t+2})\left[-1 + \frac{1}{\beta_3}\right]\right).$$

The second term on the right-hand side is positive, which means that parents do too much illiquid bequeathing and therefore should be taxed.

Intuitively, since the parent cannot transfer resources to period $t+3$ (the fourth period of the offspring's life) directly, there is no way to discipline period $t+2$ savings of the offspring via the timing of transfers. Therefore, the parent still bequests too much, this time through the illiquid asset, and hence, the optimal tax on illiquid bequests is positive. If there were a perfect set of long-term assets, $\{b^o_{t+i}\}_{i=1}^I$, available to the parent, only in that case would the parent be able to control the offspring's consumption completely and would not need to transfer too much relative to the Ramsey allocation through any of these assets. Only in that case would the optimality of transfer taxation break. As long as the parent does not have access to such a perfect set of assets, the offspring will be undersaving from the parent's perspective in some period, and hence, the parent will transfer too much to that period, requiring a tax on transfers to that period.

Remember that we assumed that the offspring have no labor income of their own, which – together with the no borrowing constraint – allowed parents to control the levels of the first two periods of consumption. When offspring receive labor income over their life cycle, then, depending on the shape of the wage profile over the life cycle, it might not be possible for parents to drive children to behave according to the Euler equations (22) and (23) above even if parents have access to a perfect portfolio of assets. Alternatively, if the borrowing limits are not exactly zero but allow for some undersaving behavior for the offspring, then, depending on the wage profile over the life cycle, again parents may not be able to arrange children's first and second period consumption according to (22) and (23). In these cases, parents will have motives to transfer too much through inter vivos and liquid bequests, and the optimal taxes on these transfers should again be positive.

7.4 Age and Income Heterogeneity in the Population

In the main model, for expositional purposes, we assume that members of different dynasties are all at the same age. This implies that there is a single age group alive in each period (other than the coexistence period). It is straightforward to extend the model to allow for different cohorts to coexist, and obviously the main intuition behind the transfer taxation result does not depend on whether there are people of different ages in the economy.

Another simplifying assumption we make is that all the agents in the economy have the same productivity - the ability to turn their labor into earnings. Instead, we could have assumed that people differ in their productivity levels, which would translate into income inequality in equilibrium. Obviously, the existence of a parental tendency to transfer too much when facing offspring with present-bias problems does not depend on the income level of the parent or the offspring. Therefore, even though the optimal tax rates might depend on income levels, the result that transfers should be taxed remains intact when there is income heterogeneity.[26] So, even if we cannot observe productivity types, we can still argue that increasing transfer taxes from zero to a some positive number uniformly across the income distribution is going to benefit everyone in the economy. Another issue that comes to mind when there are different income groups is redistribution. In this paper, we are silent about redistribution on purpose, to highlight the efficiency enhancing role of transfer taxation. It would be easy to see, however, how our normative predictions do not change whenever redistribution can be performed via lumps-sum taxes. In other words, our results survive in an environment with

[26]We know that with logarithmic utility, optimal transfer tax rates are independent of income.

horizontal inequality as long as capital taxation is not shaped by redistributional purposes.

8 Conclusion

We study optimal taxation of parental transfers in a model with altruistic parents facing children with present-biased preferences. We compare the equilibrium allocation to two benchmark allocations that represent different notions of efficiency. We consider the Ramsey allocation, which might be an appropriate notion of efficiency when a government can tax life-cycle savings as well as bequests, and the constrained efficient allocation, which is an appropriate benchmark for a government that cannot use taxes to correct for the present-bias in agents' savings decisions. We show that in both environments parents transfer too much to their offspring relative to the social optimum, and hence, parental wealth transfers should be taxed. Interestingly, the optimality of a positive tax in the case where the government has limited tax instruments is motivated by the presence of a pure price externality.

We performed a simple simulation exercise and - for a reasonable set of parameters - we obtained bequest taxes between 10% and 48% depending on the degree of present-bias. These figures are similar to existing bequest and inter vivos taxes in the US and UK. We also showed that the optimality of positive taxes on transfers remains valid even when (i) we consider life cycles with arbitrarily long finite horizons, (ii) when parents can trade long maturity assets, and (iii) when parents are naive in the sense that they do not realize offspring's present-bias problems.

References

Ameriks, J., A. Caplin, J. Leahy, and T. Tyler (2007): "Measuring Self-Control Problems," *American Economic Review*, 97(3), 966–972.

Caspi, A., B. W. Roberts, R. W. Robins, and K. H. Trzesniewski (2003): "Personality Trait Development in Adulthood," in *Handbook of the Life Course*, ed. by J. Mortimer, and M. Shanahan, chap. 26, pp. 579–598. New York, NY: Kluwer Academic.

Cremer, H., and P. Pestieau (2010): "Wealth and Wealth Transfer Taxation: A Survey," International Studies Program Working Paper 10-17, Andrew Young School of Policy Studies.

DellaVigna, S. (2009): "Psychology and Economics: Evidence from the Field," *Journal of Economic Literature*, 47(2), 315–372.

DellaVigna, S., and U. Malmendier (2004): "Contract Design and Self-control: Theory and Evidence," *The Quarterly Journal of Economics*, 119(2), 353–402.

Farhi, E., and I. Werning (2010): "Progressive Estate Taxation," *The Quarterly Journal of Economics*, 125(2), 635–673.

Frederick, S., G. Loewenstein, and T. O'Donoghue (2002): "Time Discounting and Time Preference: A Critical Review," *Journal of Economic Literature*, 40(2), 351–401.

Green, L., A. F. Fry, and J. Myerson (1994): "Discounting of Delayed Rewards: A Life-Span Comparison," *Psychological Science*, 5, 33–36.

Green, L., J. Myerson, and P. Ostaszewski (1999): "Discounting of Delayed Rewards across the Life-Span: Age Differences in Individual Discounting Functions," *Behavioural Processes*, 46(1), 89–96.

Gruber, J., and B. Koszegi (2004): "Tax incidence when individuals are time-inconsistent: the case of cigarette excise taxes," *Journal of Public Economics*, 88(9-10), 1959–1987.

Harris, C. J., and D. Laibson (2001): "Dynamic Choices of Hyperbolic Consumers," *Econometrica*, 69(4), 935–957.

——— (2002): "Hyperbolic Discounting and Consumption," in *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress*, ed. by M. Dewatripont, L. P. Hansen, and S. Turnovsky, vol. 1, pp. 258–298. Cambridge: Cambridge University Press.

Jacobson, D. B., B. G. Raub, and B. W. Johnson (2007): "The Estate Tax: Ninety Years and Counting," Discussion paper, Internal Revenue Service, U.S. Department of the Treasury.

Kaplow, L. (1995): "A Note on Subsidizing Gifts," *Journal of Public Economics*, 58(3), 469–477.

——— (2001): "A Framework for Assessing Estate and Gift Taxation," in *Rethinking Estate and Gift Taxation*, ed. by W. Gale, J. Hines, and J. Slemrod. Brookings Institution Press.

Kopczuk, W. (2010): "Economics of Estate Taxation: A Brief Review of Theory and Evidence," NBER Working Paper 15741, National Bureau of Economic Research.

Krusell, P., B. Kuruscu, and A. J. Smith (2010): "Temptation and Taxation," *Econometrica*, 78(6), 2063–2084.

Krusell, P., and A. A. Smith (2003): "Consumption–Savings Decisions with Quasi–Geometric Discounting," *Econometrica*, 71(1), 365–375.

Laibson, D. (1994): "Self-Control and Saving," Discussion paper.

——— (1997): "Golden Eggs and Hyperbolic Discounting," *The Quarterly Journal of Economics*, 112(2), 443–477.

Laibson, D., A. Repetto, and J. Tobacman (2007): "Estimating Discount Functions with Consumption Choices over the Lifecycle," NBER Working Paper 13314, National Bureau of Economic Research.

Laitner, J. (1997): "Intergenerational and Interhousehold Economic Links," in *Handbook of Population and Family Economics*, ed. by M. R. Rosenzweig, and O. Stark, vol. 1. New York: Elsevier/North Holland.

Mirrlees, J. A. (1971): "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 38(114), 175–208.

Montrucchio, L., and N. Pavoni (2010): "Dynamic Contracts, Recursive Formulations, and Envelope Theorems," Discussion paper.

Morris, S., and A. Postlewaite (1997): "Observational Implications of Non-Exponential Discounting," Discussion paper.

O'Donoghue, T., and M. Rabin (2003): "Studying Optimal Paternalism, Illustrated by a Model of Sin Taxes," *American Economic Review*, 93(2), 186–191.

——— (2006): "Optimal Sin Taxes," *Journal of Public Economics*, 90(10-11), 1825–1849.

Pavoni, N., and H. Yazici (2012): "Optimal Life-cycle Capital Taxation under Self-Control Problems," IGIER Working Paper 467, IGIER.

Phelps, E. S., and R. A. Pollak (1968): "On Second-Best National Saving and Game-Equilibrium Growth," *The Review of Economic Studies*, 35(2), 185–199.

Piketty, T., and E. Saez (2013): "A Theory of Optimal Inheritance Taxation," *Econometrica*, forthcoming.

Read, D., and N. Read (2004): "Time Discounting over the Lifespan," *Organizational Behavior and Human Decision Processes*, 94(1), 22–32.

Strotz, R. H. (1955): "Myopia and Inconsistency in Dynamic Utility Maximization," *The Review of Economic Studies*, 23(3), 165–180.

A Appendix: Proofs

A.1 Proof of Proposition 1

Before starting the proof, we demonstrate two preliminary lemmas which we use later. To simplify notation, we indicate offspring policies disregarding price sequence indexing.

Lemma 15 *The policy of the offspring $b_{t+2}(\cdot)$ is increasing in the amount of bequests received and the value function $V(\cdot)$ is strictly increasing in wealth.*

Proof. It is easy to see that $V(\cdot)$ is strictly increasing. A higher amount of assets enlarges the constraint set of the parent with at least one allocation that strictly improves his welfare: the one in which he consumes all the extra wealth in period t.

The monotonicity of $b_{t+2}(\cdot)$ is shown as follows. From the definition of the policy for the offspring at b_{t+1}, for $\varepsilon > 0$, we have

$$\begin{aligned} & u(R_{t+1}b_{t+1} + w_{t+1} - b_{t+2}(b_{t+1})) - u(R_{t+1}b_{t+1} + w_{t+1} - b_{t+2}(b_{t+1} + \epsilon)) \\ \geq\ & \beta[V(R_{t+2}b_{t+2}(b_{t+1} + \varepsilon) + w_{t+2}, Q_{t+2}) - V(R_{t+2}b_{t+2}(b_{t+1}) + w_{t+2}, Q_{t+2})]. \end{aligned} \tag{29}$$

Using the definition of the policy for the offspring at $b_{t+1} + \varepsilon$, we have

$$\begin{aligned} & u(R_{t+1}(b_{t+1} + \epsilon) + w_{t+1} - b_{t+2}(b_{t+1})) - u(R_{t+1}(b_{t+1} + \epsilon) + w_{t+1} - b_{t+2}(b_{t+1} + \epsilon)) \\ \leq\ & \beta[R_{t+2}V(b_{t+2}(b_{t+1} + \varepsilon) + w_{t+2}, Q_{t+2}) - V(R_{t+2}b_{t+2}(b_{t+1}) + w_{t+2}, Q_{t+2})]. \end{aligned} \tag{30}$$

Combining (29) and (30), we get

$$\begin{aligned} & u(R_{t+1}b_{t+1} + w_{t+1} - b_{t+2}(b_{t+1})) - u(R_{t+1}b_{t+1} + w_{t+1} - b_{t+2}(b_{t+1} + \epsilon)) \\ \geq\ & u(R_{t+1}(b_{t+1} + \epsilon) + w_{t+1} - b_{t+2}(b_{t+1})) - u(R_{t+1}(b_{t+1} + \epsilon) + w_{t+1} - b_{t+2}(b_{t+1} + \epsilon)). \end{aligned} \tag{31}$$

Assume for the sake of contradiction that $b_{t+2}(b_{t+1} + \varepsilon) < b_{t+2}(b_{t+1})$. Combined with strict concavity of u, this contradicts with (31). Thus, it must be that $b_{t+2}(b_{t+1} + \varepsilon) \geq b_{t+2}(b_{t+1})$. ■

Lemma 16 *Suppose the value function $V(\cdot)$ is differentiable. Then, $b_{t+2}(\cdot)$ is strictly monotone.*

Proof. Recall that the solution must be interior. A necessary condition for optimality of offspring's savings is

$$u'(R_{t+1}b_{t+1} + w_{t+1} - b_{t+2}) = \beta\delta V_1(R_{t+2}b_{t+2} + w_{t+2}, Q_{t+2})R_{t+2}. \tag{32}$$

Now suppose the offspring receives a higher level of bequests from the parent, $b_{t+1} + x$. It would never be optimal for the offspring to use all of the increase in his wealth for current consumption since this does not satisfy his necessary condition for optimality:

$$u'(R_{t+1}(b_{t+1} + x) + w_{t+1} - b_{t+2}) < \beta\delta V_1(R_{t+2}b_{t+2} + w_{t+2}, Q_{t+2})R_{t+2},$$

which follows from (32) and strict concavity of the utility function. ■

Proof. (Core Proof of Proposition 1) In the proof of this proposition, we use monotonicity of $b_{t+2}(\cdot)$ and strict monotonicity of $V(\cdot)$, which we establish in Lemma 15.

Assume for the sake of finding a contradiction that in equilibrium

$$-u'\left(R_t b_t + w_t - b_{t+1}\right) + \gamma u'\left(R_{t+1} b_{t+1} + w_{t+1} - b_{t+2}(b_{t+1})\right) > 0. \tag{33}$$

We want to show that there is a small positive $\varepsilon > 0$ such that

$$\begin{aligned}
&u\left(R_t b_t + w_t - b_{t+1} - \varepsilon\right) + \gamma\left[u\left(R_{t+1}(b_{t+1}+\varepsilon) + w_{t+1} - b_{t+2}(b_{t+1}+\varepsilon)\right) + V\left(R_{t+2} b_{t+2}(b_{t+1}+\varepsilon) + w_{t+2}, Q_{t+2}\right)\right] \\
&> u\left(R_t b_t + w_t - b_{t+1}\right) + \gamma\left[u\left(R_{t+1} b_{t+1} + w_{t+1} - b_{t+2}(b_{t+1})\right) + V\left(R_{t+2} b_{t+2}(b_{t+1}) + w_{t+2}, Q_{t+2}\right)\right].
\end{aligned}$$

Since u is a differentiable function, for ε sufficiently small, under the assumption (33), we have

$$\begin{aligned}
& u\left(R_t b_t + w_t - b_{t+1} - \varepsilon\right) + \gamma u\left(R_{t+1}(b_{t+1}+\varepsilon) + w_{t+1} - b_{t+2}(b_{t+1})\right) \\
> \;& u\left(R_t b_t + w_t - b_{t+1}\right) + \gamma u\left(R_{t+1} b_{t+1} + w_{t+1} - b_{t+2}(b_{t+1})\right).
\end{aligned}$$

Then, we have

$$\begin{aligned}
&u\left(R_t b_t + w_t - b_{t+1} - \varepsilon\right) + \gamma\left[u\left(R_{t+1}(b_{t+1}+\varepsilon) + w_{t+1} - b_{t+2}(b_{t+1}+\varepsilon)\right) + V\left(R_{t+2} b_{t+2}(b_{t+1}+\varepsilon) + w_{t+2}, Q_{t+2}\right)\right] \\
&\geq u\left(R_t b_t + w_t - b_{t+1} - \varepsilon\right) + \gamma\left[\begin{array}{c} u\left(R_{t+1}(b_{t+1}+\varepsilon) + w_{t+1} - b_{t+2}(b_{t+1})\right) + \beta V\left(R_{t+2} b_{t+2}(b_{t+1}) + w_{t+2}, Q_{t+2}\right) \\ +(1-\beta) V\left(R_{t+2} b_{t+2}(b_{t+1}+\varepsilon) + w_{t+2}, Q_{t+2}\right) \end{array}\right] \\
&> u\left(R_t b_t + w_t - b_{t+1}\right) + \gamma\left[\begin{array}{c} u\left(R_{t+1} b_{t+1} + w_{t+1} - b_{t+2}(b_{t+1})\right) + \beta V\left(R_{t+2} b_{t+2}(b_{t+1}) + w_{t+2}, Q_{t+2}\right) \\ +(1-\beta) V\left(R_{t+2} b_{t+2}(b_{t+1}+\varepsilon) + w_{t+2}, Q_{t+2}\right) \end{array}\right] \\
&\geq u\left(R_t b_t + w_t - b_{t+1}\right) + \gamma\left[\begin{array}{c} u\left(R_{t+1} b_{t+1} + w_{t+1} - b_{t+2}(b_{t+1})\right) + \beta V\left(R_{t+2} b_{t+2}(b_{t+1}) + w_{t+2}, Q_{t+2}\right) \\ +(1-\beta) V\left(R_{t+2} b_{t+2}(b_{t+1}) + w_{t+2}, Q_{t+2}\right), \end{array}\right]
\end{aligned}$$

where the first inequality follows from the definition of the policy $b_{t+2}(\cdot)$ and the fact that the function $V(\cdot)$ is strictly monotone, the second inequality follows from (33) and the differentiability of the utility function u as discussed above, and the last one from the monotonicity of both $b_{t+2}(\cdot)$ and $V(\cdot)$ and the fact that $\beta < 1$. It is now easy to see that the last row will have a strict inequality whenever the policy $b_{t+2}(\cdot)$ is strictly monotone. In this case, following the same line of proof we can show a contradiction to the weak inequality version of (33).

The case of $\beta = 1$ is trivial. ■

A.2 Proof of Proposition 6

Proof. Given any joint sequence of prices and taxes Ψ, let

$$\Gamma_s(b) = R_s(1-\tau_s)b + w_s + T_s + G_s$$

be the net present value of wealth as of the beginning of period s of an agent who saved b units in the previous period (of course, we only consider prices and taxes such that this sum converges), where G_s denotes the net present value of wages and lump-sum taxes from period $s+1$ onwards

$$G_s = \sum_{m=1}^{\infty} \frac{w_{s+m} + T_{s+m}}{\Pi_{n=1}^{m} R_{s+n}(1-\tau_{s+n})} = \frac{w_{s+1} + T_{s+1} + G_{s+1}}{R_{s+1}(1-\tau_{s+1})}.$$

We are going to construct an equilibrium in which agents' policies are linear in the current net present value of wealth. We do so in three steps.

Step 1.

We first guess that the value function of the parent has the form

$$V(R_s(1-\tau_s)b + w_s + T_s, \Psi_s) = \hat{V}(\Gamma_s(b), \Psi_s),$$

where $\hat{V}$ is homogeneous of degree $1-\sigma \leq 1$ in period s net present value of wealth, that is

$$\hat{V}(\lambda\Gamma_s(b), \Psi_s) = \lambda^{1-\sigma}\hat{V}(\Gamma_s(b), \Psi_s), \qquad \forall \quad \lambda > 0.$$

Step 2.

We now show that, given this guess about the value function, the consumption policy of period s is linear in $\Gamma_s(b)$ for each period s. In Step 3, we will verify that this policy indeed generates a value function that has homogeneity of degree $1-\sigma$ in $\Gamma_s(b)$.

We proceed by backward induction. Consider the problem of an offspring in period $s+1$.

Claim 1.

$$\hat{c}_{s+1} \text{ solves } \max u(c_{s+1}) + \beta\delta\hat{V}(\Gamma_{s+2}(b_{s+2}), \Psi_{s+2}) \text{ s.t. } c_{s+1} + b_{s+2} = \Gamma_{s+1}(b_{s+1}) - G_{s+1}$$

if and only if

$$\lambda\hat{c}_{s+1} \text{ solves } \max u(c_{s+1}) + \beta\delta\hat{V}(\Gamma_{s+2}(b_{s+2}), \Psi_{s+2}) \text{ s.t. } c_{s+1} + b_{s+2} = \lambda\Gamma_{s+1}(b_{s+1}) - G_{s+1}.$$

First we show that budget constraint is homogenous of degree 1 in consumption and net present value of wealth. To do so, let $\Gamma_{s+1}(\hat{b}_{s+1})$, be the period $s+1$ net present value of wealth when $\Gamma_s(b_s)$ is period s net present value of wealth and period s consumption choice is $\hat{c}_s$. Now we show that when period s wealth is $\lambda\Gamma_s(b_s)$ and the agent consumes $\lambda\hat{c}_s$ in period s, then period $s+1$ net present value of wealth will be $\lambda\Gamma_{s+1}(b_{s+1})$. First, observe that period s saving in the latter case is given by

$$\left[\lambda\Gamma_s(b_s) - G_s - \lambda\hat{c}_s\right].$$

Plugging this value in the definition of net present value of wealth for period $s+1$, we get:

$$\begin{aligned}
& R_{s+1}(1-\tau_{s+1})\left[\lambda\Gamma_s(b_s) - G_s - \lambda\hat{c}_s\right] + w_{s+1} + T_{s+1} + G_{s+1} \\
= \; & \lambda\Gamma_s(b_s)R_{s+1}(1-\tau_{s+1}) - \lambda\hat{c}_s R_{s+1}(1-\tau_{s+1}) - G_s R_{s+1}(1-\tau_{s+1}) + T_{s+1} + G_{s+1} \\
= \; & \lambda\left[\Gamma_s(b_{s-1})R_s(1-\tau_s) - \hat{c}_s R_s(1-\tau_s)\right] \\
= \; & \lambda\Gamma_{s+1}(\hat{b}_{s+1}).
\end{aligned}$$

Suppose $\hat{c}_{s+1}$ solves

$$\max \ u(c_{s+1}) + \beta\delta\hat{V}(\Gamma_{s+2}(b_{s+2}), \Psi_{s+2})$$

s.t. $c_{s+1} + b_{s+2} = \Gamma_{s+1}(b_{s+1}) - G_{s+1}$, and for the sake of contradiction suppose $\lambda\hat{c}_{s+1}$ does not solve

$$\sup \ u(c_{s+1}) + \beta\delta\hat{V}(\Gamma_{s+2}(b_{s+2}), \Psi_{s+2})$$

s.t. $c_{s+1} + b_{s+2} = \lambda\Gamma_{s+1}(b_{s+1}) - G_{s+1}$. Let v^* be the solution of the previous supremum problem and define

$$\kappa := v^* - \left[u(\lambda\hat{c}_{s+1}) + \beta\delta\hat{V}(\lambda\Gamma_{s+2}(\hat{b}_{s+2}), \Psi_{s+2})\right] > 0.$$

Under our assumptions, for each $\varepsilon > 0$ there is a feasible $\bar{c}^{\varepsilon}_{s+1}$ such that

$$u(\bar{c}^{\varepsilon}_{s+1}) + \beta\delta\hat{V}(\Gamma_{s+2}(\bar{b}^{\varepsilon}_{s+2}), \Psi_{s+2}) > u(\lambda\hat{c}_{s+1}) + \beta\delta\hat{V}(\lambda\Gamma_{s+2}(\hat{b}_{s+2}), \Psi_{s+2}) + \kappa - \varepsilon,$$

where $\bar{b}^{\varepsilon}_{s+2}$ is adjusted so as to maintain feasibility. By homogeneity of the utility and value functions, for all $\lambda > 0$ the previous statement is equivalent to

$$u\left(\frac{\bar{c}^{\varepsilon}_{s+1}}{\lambda}\right) + \beta\delta\hat{V}\left(\frac{\Gamma_{s+2}(\bar{b}^{\varepsilon}_{s+2})}{\lambda}, \Psi_{s+2}\right) > u(\hat{c}_{s+1}) + \beta\delta\hat{V}\left(\Gamma_{s+2}(\hat{b}_{s+2}), \Psi_{s+2}\right) + \frac{\kappa - \varepsilon}{\lambda}.$$

We also know that if $\bar{c}^{\varepsilon}_{s+1}$ and $\Gamma_{s+2}(\bar{b}^{\varepsilon}_{s+2})$ are feasible in the problem of the agent facing wealth $\lambda\Gamma_{s+1}(b_{s+1})$, so are $\frac{\bar{c}^{\varepsilon}_{s+1}}{\lambda}$ and $\frac{\Gamma_{s+2}(\bar{b}^{\varepsilon}_{s+2})}{\lambda}$ in the problem of the agent facing $\Gamma_{s+1}(b_{s+1})$. Setting $\varepsilon^* = \frac{\kappa}{2} > 0$, we obtain a contradiction since $\frac{\bar{c}^{\varepsilon^*}_{s+1}}{\lambda}$ and $\frac{\Gamma_{s+2}(\bar{b}^{\varepsilon^*}_{s+2})}{\lambda}$ are feasible and give strictly higher utility to the agent's problem in year s. The converse of the claim is shown symmetrically.

Now, we prove the second step in the backward induction.

Claim 2.

$$(\hat{c}_s, \hat{c}_{s+1}) \text{ solves } \max u(c_s) + \gamma\left[u(c_{s+1}) + \delta\hat{V}(\Gamma_{s+2}(b_{s+2}), \Psi_{s+2})\right]$$

$$\text{s.t. (1) } c_s + b_{s+1} = \Gamma_s(b_s) - G_s \text{ and (2) } c_{s+1} \text{ solves agent's problem in year } s+1$$

if and only if

$$(\lambda\hat{c}_s, \lambda\hat{c}_{s+1}) \text{ solves } \max u(c_s) + \gamma\left[u(c_{s+1}) + \delta\hat{V}(\Gamma_{s+2}(b_{s+2}), \Psi_{s+2})\right]$$

$$\text{s.t. (3) } c_s + b_{s+1} = \lambda\Gamma_s(b_s) - G_s \text{ and (4) } c_{s+1} \text{ solves agent's problem in year } s+1.$$

Suppose for the sake of contradiction that $(\hat{c}_s, \hat{c}_{s+1})$ solves the corresponding problem but $(\lambda\hat{c}_s, \lambda\hat{c}_{s+1})$ does not. Since the proof follows the same principle as that in the proof of Claim 1, to save notation, we now propose the proof assuming existence of a solution for both problems. This assumption is not needed as we have shown above. Then, there exists a $(\bar{c}_s, \bar{c}_{s+1})$ such that

$$u(\bar{c}_s) + \gamma\left[u(\bar{c}_{s+1}) + \delta\hat{V}(\Gamma_{s+2}(\bar{b}_{s+2}), \Psi_{s+2})\right] > u(\lambda\hat{c}_s) + \gamma\left[u(\lambda\hat{c}_{s+1}) + \delta\hat{V}(\lambda\Gamma_{s+2}(\hat{b}_{s+2}), \Psi_{s+2})\right],$$

and (3) and (4) are satisfied. By homogeneity of utility and value functions, we have

$$u\left(\frac{\bar{c}_s}{\lambda}\right)+\gamma\left[u\left(\frac{\bar{c}_{s+1}}{\lambda}\right)+\delta\hat{V}\left(\frac{\Gamma_{s+2}\left(\bar{b}_{s+2}\right)}{\lambda},\Psi_{s+2}\right)\right]>u\left(\hat{c}_s\right)+\gamma\left[u\left(\hat{c}_{s+1}\right)+\delta\hat{V}(\Gamma_{s+2}(\hat{b}_{s+2}),\Psi_{s+2})\right].$$

Furthermore, as we have shown in the first step of the induction, if $\bar{c}_{s+1}$ solves the problem in year $s+1$ under $\Gamma_{s+1}(\bar{b}_{s+1})$, then $\frac{\bar{c}_{s+1}}{\lambda}$ solves the same problem under $\frac{\Gamma_{s+1}(\bar{b}_{s+1})}{\lambda}$. This means that $\left(\frac{\bar{c}_s}{\lambda},\frac{\bar{c}_{s+1}}{\lambda}\right)$ is in the constraint set of the agent's problem in year s, which combined with the fact that it gives strictly higher welfare than the equilibrium allocation implies a contradiction.

Step 3.

Now, we verify that, under consumption policies that are linear in the current wealth, the value function is in fact homogeneous of degree $1-\sigma$ in Γ, as assumed:

$$\begin{aligned}
V(\lambda\Gamma_s(b),\Psi_s) &= u(\lambda\hat{c}_s)+\gamma\left[u(\lambda\hat{c}_{s+1})+\delta V(\lambda\Gamma_{s+2}(\hat{b}_{s+2}),\Psi_{s+2})\right]\\
&= \frac{(\lambda\hat{c}_s)^{1-\sigma}}{1-\sigma}\gamma\left[\frac{(\lambda\hat{c}_{s+1})^{1-\sigma}}{1-\sigma}+\lambda^{1-\sigma}\delta V(\Gamma_{s+2}(\hat{b}_{s+2}),\Psi_{s+2})\right]\\
&= \lambda^{1-\sigma}\left\{u(\hat{c}_s)+\gamma\left[u(\hat{c}_{s+1})+\delta V(\Gamma_{s+2}(\hat{b}_{s+2}),\Psi_{s+2})\right]\right\}\\
&= \lambda^{1-\sigma}V(\Gamma_s(b),\Psi_s).
\end{aligned}$$

To complete the proof, observe that we have shown that consumption defined as a function of net present value of wealth, prices, and taxes, denote it by $c_s(\Gamma_s(b),\Psi_s)$, satisfies the following homogeneity of degree one in wealth:

$$c_s(\lambda\Gamma_s(b),\Psi_s)=\lambda c_s(\Gamma_s(b),\Psi_s).$$

In particular this implies

$$c_s(\Gamma_s(b),\Psi_s)=\Gamma_s(b)c_s(1,\Psi_s),$$

which means consumption is a linear function of wealth with a constant multiplier of $c_s(1,\Psi_s)$.

Using the homogeneity of the value function, we can show that the value function has the following simple form as a function of wealth:

$$V(\Gamma_s(b),\Psi_s)=\Gamma_s(b)^{1-\sigma}V(1,\Psi_s).$$

This ends the proof. Note in particular, that since we are driving conditions for an equilibrium, no verification stage is needed. ■

A.3 Proof of Proposition 7

We first define the equilibrium of an economy where period t bequest choice is made by individual parents and in all the other periods bequests are chosen by the planner. The aim of this exercise is to see if parents deviate from the constrained efficient level of bequests if allowed to do so in a period t.

Parent's Problem in Period t :

Taking $\hat{R}_{t+1}, \hat{w}_{t+1}, \hat{R}_{t+2}, \hat{w}_{t+2}$ as given, the parent solves

$$\max_{b_{t+1}} u(f(b_t^{**}) - b_{t+1}) + \gamma[u(\hat{R}_{t+1}b_{t+1} + \hat{w}_{t+1} - b_{t+2}(b_{t+1})) + \delta W(\hat{R}_{t+2}b_{t+2}(b_{t+1}) + \hat{w}_{t+2})]$$

s.t. taking $\hat{R}_{t+2}, \hat{w}_{t+2}$ as given, and believing that the offspring - facing the equilibrium prices - solves

$$\max_{b_{t+2}} u(\hat{R}_{t+1}b_{t+1} + \hat{w}_{t+1} - b_{t+2}) + \beta\delta W(\hat{R}_{t+2}b_{t+2} + \hat{w}_{t+2});$$

from which the parent derives the policy $b_{t+2}(\cdot)$. Let the parent's and offspring's choices be $\hat{b}_{t+1}$, and $\hat{b}_{t+2} = b_{t+2}(\hat{b}_{t+1})$. The prices are given by marginal products in equilibrium; that is, for $s = t+1, t+2$,

$$\hat{R}_s = f'(\hat{b}_s) \text{ and } \hat{w}_s = f(\hat{b}_s) - f'(\hat{b}_s)\hat{b}_s. \tag{34}$$

The planner's problem at the same continuation is exactly the same except the planner does not take prices as given, so the pricing conditions (34) are a constraint to his problem.

Before starting the main core of the proof, we need to show two preliminary lemmas.

Lemma 17 *The offspring's saving as a function of bequests received in the planner's problem,* $b_{t+2}^{PL}(\cdot)$ *and the planner's value function* $W(\cdot)$ *are increasing in their arguments, with* $W(\cdot)$ *strictly increasing. If* $W(\cdot)$ *is differentiable, then,* $b_{t+2}^{PL}(\cdot)$ *is strictly monotone.*

Proof. The weak monotonicity of the policy function and strict monotonicity of the value function can be shown by following exactly the same line of proof of Lemma 15. The strict monotonicity of the policy function under differentiability of W is shown by following exactly the same line of proof of Lemma 16. ■

Lemma 18 *Suppose the policy functions* $b_{t+2}(\cdot)$ *and* $b_{t+2}^{PL}(\cdot)$ *are strictly monotone and continuous. Define* $b_{t+1}^{\varepsilon} := b_{t+1}^{**} + \varepsilon$. *Then, for all* $\varepsilon > 0$ *sufficiently small, there exists* $\bar{b}_{t+1}^{\varepsilon}$ *such that*

$$R_{t+2}^{**}b_{t+2}(\bar{b}_{t+1}^{\varepsilon}) + w_{t+2}^{**} = f\left(b_{t+2}^{PL}(b_{t+1}^{\varepsilon})\right).$$

Proof. For ease of notation, first let $g_1(b_{t+1}) = R_{t+2}^{**}b_{t+2}(b_{t+1}) + w_{t+2}^{**}$ and $g_2(b_{t+1}) = f\left(b_{t+2}^{PL}(b_{t+1})\right)$. Now pick $\eta > 0$. Observe that $g_1(b_{t+1}^{**} + \eta) - g_1(b_{t+1}^{**}) > 0$ since g_1 is strictly monotone. By continuity of g_2 at b_{t+1}^{**}, there is a $\delta > 0$ such that for all b_{t+1} with $|b_{t+1} - b_{t+1}^{**}| < \delta$, we have

$$|g_2(b_{t+1}) - g_2(b_{t+1}^{**})| < g_1(b_{t+1}^{**} + \eta) - g_1(b_{t+1}^{**}).$$

Now choose any $\varepsilon > 0$ such that $|b^{\varepsilon}_{t+1} - b^{**}_{t+1}| < \delta$. By continuity and strict monotonicity of g_2, we have

$$g_2(b^{\varepsilon}_{t+1}) - g_2(b^{**}_{t+1}) < g_1(b^{**}_{t+1} + \eta) - g_1(b^{**}_{t+1}).$$

Remember that $g_1(b^{**}_{t+1}) = g_2(b^{**}_{t+1})$, and therefore

$$g_2(b^{\varepsilon}_{t+1}) < g_1(b^{**}_{t+1} + \eta).$$

Observe that g_1 is a continuous function on $[b^{**}_{t+1}, b^{**}_{t+1} + \eta]$, with

$$g_1(b^{**}_{t+1}) = g_2(b^{**}_{t+1}) < g_2(b^{\varepsilon}_{t+1}) < g_1(b^{**}_{t+1} + \eta).$$

Then, by Intermediate Value Theorem, there exists $\bar{b}^{\varepsilon}_{t+1} \in (b^{**}_{t+1}, b^{**}_{t+1} + \eta)$ such that

$$g_1(\bar{b}^{\varepsilon}_{t+1}) = g_2(b^{\varepsilon}_{t+1}).$$

■

Proof. (Core Proof of Proposition 7) Now we are ready to prove Proposition 7. Recall we need to show that the parent's choice is larger than that of the planner: $\hat{b}_{t+1} \geq b^{**}_{t+1}$. Suppose for the sake of contradiction that $\hat{b}_{t+1} < b^{**}_{t+1}$. We will show that we can find a $\bar{b}_{t+1}$ that strictly improves parent's welfare thereby giving us the contradiction we desire.

Note that W is continuous by assumption (as it is differentiable) and the budget set of the offspring can be considered as a compact valued correspondence (these are standard arguments, the correspondence is bounded below the the natural borrowing limit, and we can bound above the amount of savings since consumption will never be negative). Then, by Maximum Theorem, the policies b_{t+2} and b^{PL}_{t+2} are continuous functions. They are also strictly monotone by Lemma 17. Then, by Lemma 18, we can choose $\bar{b}_{t+1}$ such that

$$\hat{R}_{t+2} b_{t+2}(\bar{b}_{t+1}) + \hat{w}_{t+2} = f\left(b^{PL}_{t+2}(b^{**}_{t+1})\right).$$

Observe that $b^{PL}_{t+2}(b^{**}_{t+1}) > b_{t+2}(\bar{b}_{t+1})$. This follows from strict concavity of f and the fact that $b^{PL}_{t+2}(\hat{b}_{t+1}) < b^{PL}_{t+2}(b^{**}_{t+1})$, where the latter follows from strict monotonicity of the policy.

We now show that $\bar{b}_{t+1}$ improves over $\hat{b}_{t+1}$ for the parent, giving a contradiction:

$$\begin{aligned} & u\left(f(b^{**}_t) - \bar{b}_{t+1}\right) + \gamma\left[u\left(\hat{R}_{t+1}\bar{b}_{t+1} + \hat{w}_{t+1} - b_{t+2}(\bar{b}_{t+1})\right) + \delta W\left(\hat{R}_{t+2} b_{t+2}(\bar{b}_{t+1}) + \hat{w}_{t+2}\right)\right] \\ > \; & u\left(f(b^{**}_t) - \hat{b}_{t+1}\right) + \gamma\left[u\left(\hat{R}_{t+1}\hat{b}_{t+1} + \hat{w}_{t+1} - b_{t+2}(\hat{b}_{t+1})\right) + \delta W\left(\hat{R}_{t+2} b_{t+2}(\hat{b}_{t+1}) + \hat{w}_{t+2}\right)\right]. \end{aligned}$$

To show this, it is enough to show

$$\begin{aligned} & u\left(f(b^{**}_t) - \bar{b}_{t+1}\right) + \gamma\left[u\left(\hat{R}_{t+1}\bar{b}_{t+1} + \hat{w}_{t+1} - b_{t+2}(\bar{b}_{t+1})\right) + \delta W\left(\hat{R}_{t+2} b_{t+2}(\bar{b}_{t+1}) + \hat{w}_{t+2}\right)\right] \\ > \; & u\left(f(b^{**}_t) - b^{**}_{t+1}\right) + \gamma\left[u\left(f(b^{**}_{t+1}) - b^{PL}_{t+2}(b^{**}_{t+1})\right) + \delta W\left(f(b^{PL}_{t+2}(b^{**}_{t+1}))\right)\right], \end{aligned} \tag{35}$$

since

$$
\begin{aligned}
& u\left(f(b_t^{**}) - b_{t+1}^{**}\right) + \gamma\left[u\left(f(b_{t+1}^{**}) - b_{t+2}^{PL}(b_{t+1}^{**})\right) + \delta W\left(f\left(b_{t+2}^{PL}(b_{t+1}^{**})\right)\right)\right] \\
\geq\ & u\left(f(b_t^{**}) - \hat{b}_{t+1}\right) + \gamma\left[u\left(f(\hat{b}_{t+1}) - b_{t+1}^{PL}(\hat{b}_{t+1})\right) + \delta W\left(f\left(b_{t+1}^{PL}(\hat{b}_{t+1})\right)\right)\right] \\
=\ & u\left(f(b_t^{**}) - \hat{b}_{t+1}\right) + \gamma\left[u\left(\hat{R}_{t+1}\hat{b}_{t+1} + \hat{w}_{t+1} - b_{t+2}(\hat{b}_{t+1})\right) + \delta W\left(\hat{R}_{t+2}b_{t+2}(\hat{b}_{t+1}) + \hat{w}_{t+2}\right)\right],
\end{aligned}
$$

where the inequality follows from the fact that the constrained efficient allocation solves the planning problem and the equality follows from the fact that when the planner chooses $\hat{b}_{t+1}$, the offspring face the equilibrium price $\hat{R}_{t+2}, \hat{w}_{t+2}$, and hence make the same choice, $b_{t+2}^{PL}(\hat{b}_{t+1}) = b_{t+2}(\hat{b}_{t+1})$.

By construction of $\bar{b}_{t+1}$, showing (35) reduces to showing

$$
\begin{aligned}
& u\left(f(b_t^{**}) - \bar{b}_{t+1}\right) + \gamma u\left(\hat{R}_{t+1}\bar{b}_{t+1} + \hat{w}_{t+1} - b_{t+2}(\bar{b}_{t+1})\right) \\
>\ & u\left(f(b_t^{**}) - b_{t+1}^{**}\right) + \gamma u\left(f(b_{t+1}^{**}) - b_{t+2}^{PL}(b_{t+1}^{**})\right).
\end{aligned} \tag{36}
$$

Now observe that

$$
\hat{R}_{t+2}b_{t+2}(\bar{b}_{t+1}) + \hat{w}_{t+2} = f\left(b_{t+2}^{PL}(b_{t+1}^{**})\right) > f(b_{t+2}(\hat{b}_{t+1})) = \hat{R}_{t+2}b_{t+2}(\hat{b}_{t+1}) + \hat{w}_{t+2},
$$

where the first equality is by construction and the inequality is true since $b_{t+2}^{PL}(b_{t+1}^{**}) > b_{t+2}(\hat{b}_{t+1})$. But this implies that

$$
b_{t+2}(\bar{b}_{t+1}) > b_{t+2}(\hat{b}_{t+1}),
$$

and that $\bar{b}_{t+1} > \hat{b}_{t+1}$ since the policy function is monotone, which implies that

$$
\hat{R}_{t+1}\bar{b}_{t+1} + \hat{w}_{t+1} > f(\bar{b}_{t+1}).
$$

This means, to show (36), it is enough to show that

$$
\begin{aligned}
& u\left(f(b_t^{**}) - \bar{b}_{t+1}\right) + \gamma u\left(f(\bar{b}_{t+1}) - b_{t+2}(\bar{b}_{t+1})\right) \\
\geq\ & u\left(f(b_t^{**}) - b_{t+1}^{**}\right) + \gamma u\left(f(b_{t+1}^{**}) - b_{t+2}^{PL}(b_{t+1}^{**})\right),
\end{aligned}
$$

which is true if we can show

$$
\begin{aligned}
& u\left(f(b_t^{**}) - \bar{b}_{t+1}\right) + \gamma u\left(f(\bar{b}_{t+1}) - b_{t+2}(\bar{b}_{t+1})\right) \\
\geq\ & u\left(f(b_t^{**}) - b_{t+1}^{**}\right) + \gamma u\left(f(b_{t+1}^{**}) - b_{t+2}(\bar{b}_{t+1})\right).
\end{aligned} \tag{37}
$$

Remember that in the constrained efficient allocation, we have

$$
u'\left(f(b_t^{**}) - b_{t+1}^{**}\right) \geq \gamma u'\left(f(b_{t+1}^{**}) - b_{t+2}^{PL}(b_{t+1}^{**})\right),
$$

which implies (from the inequality $b_{t+2}^{PL}(b_{t+1}^{**}) > b_{t+2}(\bar{b}_{t+1})$ we have observed after equation (A.3))

$$
u'\left(f(b_t^{**}) - b_{t+1}^{**}\right) > \gamma u'\left(f(b_{t+1}^{**}) - b_{t+2}(\bar{b}_{t+1})\right). \tag{38}
$$

Then, (38) implies that showing (37) reduces to showing $\bar{b}_{t+1} < b^{**}_{t+1}$.

Now we prove that $\bar{b}_{t+1} \leq b^{**}_{t+1}$ and complete the proof. By the optimality of the policy function for the offspring from the parent's perspective at $\bar{b}_{t+1}$ and differentiability of W, we have that

$$u'\left(\hat{R}_{t+1}\bar{b}_{t+1} + \hat{w}_{t+1} - b_{t+2}(\bar{b}_{t+1})\right) = \beta\delta W'\left(\hat{R}_{t+2}b_{t+2}(\bar{b}_{t+1}) + \hat{w}_{t+2}\right)\hat{R}_{t+2}. \tag{39}$$

By the optimality of the policy function for the offspring from the planner's perspective at b^{**}_{t+1}, we have that

$$u'\left(f(b^{**}_{t+1}) - b^{PL}_{t+2}(b^{**}_{t+1})\right) = \beta\delta W'\left(f(b^{PL}_{t+2}(b^{**}_{t+1}))\right)f'(b^{PL}_{t+2}(b^{**}_{t+1})). \tag{40}$$

Comparison of (39) and (40) together with strict concavity of f - hence $f'(b^{PL}_{t+2}(b^{**}_{t+1})) < \hat{R}_{t+2}$ - and recalling that in the two conditions $W'(\cdot)$ is evaluated at the same point, gives $\bar{b}_{t+1} < b^{**}_{t+1}$. ■

Remark. *Note that we could have dispensed with the differentiability assumption on W. This would create two modifications in the proof of the proposition. First, without differentiability of W, we would not be able to show strict monotonicity of the policy functions (see Lemma 17), which is used in establishing Lemma 18. Hence, the range condition that we prove in Lemma 18 would become an assumption to the proposition. Second, in the last part of the proof, we would use the fact that W is locally Lipschitz continuous. In equations* (39) *and* (40)*, we would have to use right and left derivatives of W instead of the (both sides) differential W'. That would allow us to show the weak inequality $\bar{b}_{t+1} \leq b^{**}_{t+1}$, which is sufficient to complete the proof.*

A.4 Proof of Proposition 8

Proof. First, note that when the parent chooses the constrained efficient level of bequests, b^{**}_{t+1}, in problem (11), then the offspring also chooses the constrained efficient level of savings since he faces constrained efficient level of prices, $R^{**}_{t+2}, w^{**}_{t+2}$:

$$b_{t+2}(b^{**}_{t+1}) = b^{PL}_{t+2}(b^{**}_{t+1}).$$

Let $b^{\varepsilon}_{t+1} := b^{**}_{t+1} + \varepsilon$. For all ε, set $\bar{b}^{\varepsilon}_{t+1}$ such that

$$R^{**}_{t+2}b_{t+2}(\bar{b}^{\varepsilon}_{t+1}) + w^{**}_{t+2} = f\left(b^{PL}_{t+2}(b^{\varepsilon}_{t+1})\right). \tag{41}$$

Note that under the assumption that $b_{t+2}(.)$ and $b^{PL}_{t+2}(.)$ are differentiable, both policies are strictly monotone and continuous; hence such a $\bar{b}^{\varepsilon}_{t+1}$ exists. At b^{**}_{t+1}, we have $b^{PL}_{t+2}(b^{**}_{t+1}) = b_{t+2}(b^{**}_{t+1})$, and hence, $\bar{b}^{0}_{t+1} = b^{**}_{t+1}$. Differentiating (41) with respect to ε we have

$$R^{**}_{t+2}\frac{db_{t+2}(\bar{b}^{\varepsilon}_{t+1})}{db_{t+1}}\frac{d\bar{b}^{\varepsilon}_{t+1}}{d\varepsilon} = \frac{db^{PL}_{t+2}(b^{\varepsilon}_{t+1})}{db_{t+1}}f'\left(b^{PL}_{t+2}(b^{\varepsilon}_{t+1})\right),$$

which gives

$$\frac{d\bar{b}^{\varepsilon}}{d\varepsilon} = \frac{\frac{db^{PL}_{t+2}(b^{\varepsilon}_{t+1})}{db_{t+1}}f'\left(b^{PL}_{t+2}(b^{\varepsilon}_{t+1})\right)}{R^{**}_{t+2}\frac{db_{t+2}(\bar{b}^{\varepsilon}_{t+1})}{db_{t+1}}}. \tag{42}$$

Recall that from the planner's perspective, the offspring's optimality condition for savings is

$$u'\left(f(b^{\varepsilon}_{t+1}) - b^{PL}_{t+2}(b^{\varepsilon}_{t+1})\right) = \delta\beta W'\left(f\left(b^{PL}_{t+2}(b^{\varepsilon}_{t+1})\right)\right) f'\left(b^{PL}_{t+2}(b^{\varepsilon}_{t+1})\right), \tag{43}$$

while their optimality condition from the parent's perspective is

$$u'\left(f(\bar{b}^{\varepsilon}_{t+1}) - b_{t+2}(\bar{b}^{\varepsilon}_{t+1})\right) = \delta\beta W'\left(R^{**}_{t+2} b_{t+2}(\bar{b}^{\varepsilon}_{t+1}) + w^{**}_{t+2}\right) R^{**}_{t+2}. \tag{44}$$

Since the argument in the function W is the same by construction, and W is strictly monotone, equations (43) and (44) imply that for all ε we have

$$u'\left(f(b^{\varepsilon}_{t+1}) - b^{PL}_{t+2}(b^{\varepsilon}_{t+1})\right) R^{**}_{t+2} = u'\left(f(\bar{b}^{\varepsilon}_{t+1}) - b_{t+2}(\bar{b}^{\varepsilon}_{t+1})\right) f'\left(b^{PL}_{t+2}(b^{\varepsilon}_{t+1})\right). \tag{45}$$

Differentiating (45) with respect to ε, and evaluating it at $\varepsilon = 0$, we have

$$\begin{aligned}
&\left[f'\left(b^{**}_{t+1}\right) - \frac{db^{PL}_{t+2}(b^{**}_{t+1})}{db_{t+1}}\right] u''\left(c^{**}_{t+1}\right) R^{**}_{t+2} \\
=\; &\left[f'\left(b^{**}_{t+1}\right) - \frac{db_{t+2}(b^{**}_{t+1})}{db_{t+1}}\right]\left[\frac{d\bar{b}^{0}_{t+1}}{d\varepsilon}\right] u''\left(c^{**}_{t+1}\right) f'\left(b^{PL}_{t+2}(b^{**}_{t+1})\right) + u'\left(c^{**}_{t+1}\right) f''\left(b^{PL}_{t+2}(b^{**}_{t+1})\right) \frac{db^{PL}_{t+2}(b^{**}_{t+1})}{db_t} \\
=\; &\left[f'\left(b^{**}_{t+1}\right) \frac{\frac{db^{PL}_{t+2}(b^{**}_{t+1})}{db_{t+1}}}{\frac{db_{t+2}(b^{**}_{t+1})}{db_{t+1}}} - \frac{db^{PL}_{t+2}(b^{**}_{t+1})}{db_{t+1}}\right] u''\left(c^{**}_{t+1}\right) R^{**}_{t+2} + u'\left(c^{**}_{t+1}\right) f''\left(b^{PL}_{t+2}(b^{**}_{t+1})\right) \frac{db^{PL}_{t+2}(b^{**}_{t+1})}{db_{t+1}},
\end{aligned}$$

where from the second to the third line we used (42) and rearranged terms using the fact that for the equilibrium level of bequests we have $f'\left(b^{PL}_{t+2}(b^{**}_{t+1})\right) = R^{**}_{t+2} > 0$.

Now, comparing the first line with the last one, using strict monotonicity of the policies and strict concavity of the production function, strict monotonicity and strict concavity of the utility function, we have

$$\left[f'\left(b^{**}_{t+1}\right) - \frac{db^{PL}_{t+2}(b^{**}_{t+1})}{db_{t+1}}\right] > \left[f'\left(b^{**}_{t+1}\right) \frac{\frac{db^{PL}_{t+2}(b^{**}_{t+1})}{db_{t+1}}}{\frac{db_{t+2}(b^{**}_{t+1})}{db_{t+1}}} - \frac{db^{PL}_{t+2}(b^{**}_{t+1})}{db_{t+1}}\right],$$

which implies the desired result: $\frac{\frac{db^{PL}_{t+2}(b^{**}_{t+1})}{db_{t+1}}}{\frac{db_{t+2}(b^{**}_{t+1})}{db_{t+1}}} < 1$. ■

A.5 Proof of Proposition 11

Proof. The proof proceeds in two steps. In the first step, for logarithmic utility, we compute closed form solution for equilibrium linear consumption policy as a function of net present value of wealth and tax-price sequence. In the second step, using the consumption policy computed in step one, we compare $\frac{db_{t+2}(b^{**}_{t+1},\Psi^{**}_{t+1})}{db_{t+1}}$ and $\frac{db_{t+2}(b^{**}_{t+1},Q^{**}_{t+1})}{db_{t+1}}$, which, as we will see, will enable us to establish the tax result.

Step 1. We use guess and verify method to compute value and policy functions. First, remember from the proof of Proposition 6 that, given any joint sequence of taxes and prices Ψ, we can write the parent's value function as a function of his current net present value of wealth $\hat{V}(\Gamma_t(b_t), \Psi_t)$, where $\Gamma_t(b_t)$ represents the current net present value of wealth of a parent who saved b_t units during his young adulthood in period $t-1$. (Observe that in fact $\Gamma_t(b_t)$ also depends on the tax-price sequence Ψ_t, however, we omit to make this dependence explicit in order to ease notation). Now, we guess that the value function has the following form:

$$\hat{V}(\Gamma_t(b_t), \Psi_t) = D\log(\Gamma_t(b_t)) + B(\Psi_t),$$

where D is the constant of the parent's value function.

By assumption we are interested in equilibria where policies are linear in net present value of wealth. Therefore, let consumption in period t under wealth $\Gamma_t(b_t)$ be given by

$$c_t(\Gamma_t(b_t), \Psi_t) = C_t(\Psi_t)\Gamma_t(b_t),$$

where $C_t(\Psi_t)$ is the fraction of wealth consumed by agent under Ψ_t. In what follows, we omit the dependence of C_t on Ψ_t in order to ease notation. Using linearity of the policy functions, we can rewrite the parent's problem as:

$$\begin{aligned} \hat{V}(\Gamma_t(b_t), \Psi_t) = \max_{C_t} \quad & u(C_t\Gamma_t(b_t)) + \delta u(C_{t+1}\Gamma_{t+1}(b_{t+1})) + \delta^2\hat{V}(\Gamma_{t+2}(b_{t+2}), \Psi_{t+2}) && (46) \\ & s.t. \\ u'(C_{t+1}\Gamma_{t+1}(b_{t+1})) = \quad & \delta\beta\hat{V}_1(\Gamma_{t+2}(b_{t+2}), \Psi_{t+2}) R_{t+2}(1-\tau_{t+2}) && (47) \end{aligned}$$

Note that net present value of wealth in two consecutive periods are linked as follows:

$$\begin{aligned} \Gamma_{t+1}(b_{t+1}) &= R_{t+1}(1-\tau_{t+1})b_{t+1} + w_{t+1} + T_{t+1} + G_{t+1} \\ &= R_{t+1}(1-\tau_{t+1})\left[R_t(1-\tau_t)b_t + w_t + T_t - C_t\Gamma_t(b_t)\right] + w_{t+1} + T_{t+1} + G_{t+1} \\ &= R_{t+1}(1-\tau_{t+1})\left[R_t(1-\tau_t)b_t + w_t + T_t - C_t\Gamma_t(b_t) + \frac{w_{t+1} + T_{t+1} + G_{t+1}}{R_{t+1}(1-\tau_{t+1})}\right] \\ &= R_{t+1}(1-\tau_{t+1})\left[R_t(1-\tau_t)b_t + T_t - C_t\Gamma_t(b_t) + G_t\right] \\ &= R_{t+1}(1-\tau_{t+1})\Gamma_t(b_t)\left[1 - C_t\right]. \qquad (48) \end{aligned}$$

Plugging the value function guess in the constraint of the planning problem, (47), and using $u(\cdot) = \log$, we get:

$$\begin{aligned} (C_{t+1}\Gamma_{t+1}(b_{t+1}))^{-1} &= \frac{\delta\beta R_{t+2}(1-\tau_{t+2})D}{\Gamma_{t+2}(b_{t+2})} \\ &= \frac{\delta\beta R_{t+2}(1-\tau_{t+2})D}{R_{t+2}(1-\tau_{t+2})(1-C_{t+1})\Gamma_{t+1}(b_{t+1})}, \end{aligned}$$

where the second equality follows from the relationship between consecutive wealth levels that we just established. This implies

$$(C_{t+1})^{-1} = \frac{\delta\beta D}{(1 - C_{t+1})}$$

or

$$C_{t+1}(D) = \frac{1}{1 + \delta\beta D}.$$

Taking first-order condition with respect to bequests in the parent's problem and plugging in the $C_{t+1}(D)$ from above, we get:

$$C_t(D) = \frac{1}{1 + \delta + \delta^2 D}.$$

Now verify the value function to compute D :

$$D \log(\Gamma_t (b_t)) + B(\Psi_t) = \log (C_t(D)\Gamma_t (b_t)) + \delta \log(C_{t+1}(D)\Gamma_{t+1} (b_{t+1})) + \delta^2 \{D \log(\Gamma_{t+2} (b_{t+2})) + B(\Psi_{t+2})\},$$

which, using (48) and comparing the coefficients of $\log(\Gamma_t (b_t))$ on both sides of the above equation, implies

$$D = 1 + \delta + \delta^2 D,$$

and hence

$$D = \frac{1}{1 - \delta}.$$

Then, plugging D in $C_{t+1}(D)$, we get

$$C_{t+1}(D) = \frac{1 - \delta}{1 - \delta + \delta\beta}.$$

Remembering that $b_{t+2}(b_{t+1}, \Psi_{t+1}) = R_{t+1}(1 - \tau_{t+1})b_{t+1} + w_{t+1} + T_{t+1} - C_{t+1}\Gamma_{t+1}(b_{t+1})$, we get

$$\begin{aligned} \frac{db_{t+2}(b_{t+1}, \Psi_{t+1})}{db_{t+1}} &= R_{t+1}(1 - \tau_{t+1}) - C_{t+1}R_{t+1}(1 - \tau_{t+1}) \\ &= R_{t+1}(1 - \tau_{t+1})\frac{\delta\beta}{1 - \delta + \delta\beta}. \end{aligned}$$

Step 2. Parents choosing the constrained efficient level of bequests in equilibrium under taxes means that

$$u'(c_t^{**}) = \gamma u'(c_{t+1}^{**}) \left(R_{t+1}^{**}(1 - \tau_{t+1}^{**}) + \frac{db_{t+2}(b_{t+1}^{**}, \Psi_{t+1}^{**})}{db_{t+1}} \left\{ -1 + \frac{1}{\beta} \right\} \right). \tag{49}$$

Now we use our finding from step one:

$$\frac{db_{t+2}(b_{t+1}^{**}, \Psi_{t+1}^{**})}{db_{t+1}} = R_{t+1}^{**}(1 - \tau_{t+1}^{**})\frac{\beta\delta}{1 - \delta + \beta\delta}$$

and

$$\frac{db_{t+2}(b_{t+1}^{**}, Q_{t+1}^{**})}{db_{t+1}} = R_{t+1}^{**}\frac{\beta\delta}{1 - \delta + \beta\delta},$$

implying

$$\frac{db_{t+2}(b^{**}_{t+1}, \Psi^{**}_{t+1})}{db_{t+1}} = \frac{db_{t+2}(b^{**}_{t+1}, Q^{**}_{t+1})}{db_{t+1}}(1 - \tau^{**}_{t+1}). \tag{50}$$

Using equation (50) in equation (49) gives:

$$u'(c^{**}_t) = \gamma u'(c^{**}_{t+1})(1 - \tau^{**}_{t+1}) \left(R^{**}_{t+1} + \frac{db_{t+2}(b^{**}_{t+1}, Q^{**}_{t+1})}{db_{t+1}} \left\{ -1 + \frac{1}{\beta} \right\} \right). \tag{51}$$

A comparison of equation (51) with the condition for constrained efficient level of bequests under differentiability, which is given by

$$u'(c^{**}_t) = \gamma u'(c^{**}_{t+1}) \left(R^{**}_{t+1} + \frac{db^{PL}_{t+2}(b^{**}_{t+1})}{db_{t+1}} \left\{ -1 + \frac{1}{\beta} \right\} \right),$$

gives the tax formula to be

$$\tau^{**}_{t+1} = 1 - \frac{R^{**}_{t+1} + \frac{db^{PL}_{t+2}(b^{**}_{t+1})}{db_{t+1}} \left[-1 + \frac{1}{\beta} \right]}{R^{**}_{t+1} + \frac{db_{t+2}(b^{**}_{t+1}, Q^{**}_{t+1})}{db_{t+1}} \left[-1 + \frac{1}{\beta} \right]}.$$

Proposition 8 then implies that the tax is strictly positive. ■