Intergenerational Disagreement and Optimal Taxation of Parental Transfers*

Nicola Pavoni† and Hakki Yazici‡

June, 2015

Abstract

We study optimal taxation of bequests and inter vivos transfers in a model where altruistic parents and their offspring disagree on intertemporal trade-offs. We show that the laissez-faire equilibrium is Pareto inefficient, and whenever offspring are impatient from their parents' perspective, optimal policy involves a positive tax on parental transfers. 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.

JEL classification: E21, H21, D91.

Keywords: Parental transfer taxation, intergenerational disagreement, altruism.

---

*A version of this paper previously circulated under the title "Present-Biased Preferences and Optimal Taxation of Parental Transfers." We thank the editor, Michelle Tertilt, and four anonymous referees for numerous and very insightful comments. We gratefully acknowledge the comments and suggestions of Per Krusell, Marco Ottaviani, Vincenzo Quadrini, seminar participants at Bogazici University, Erasmus University, Goethe University in Frankfurt, Istanbul Technical University, IIES, Meet the IGIER Scholars 2015 meetings, SED meetings in Ghent, University College London, University of Alicante, University of Bologna, University of Carlos III de Madrid, University of Konstanz, University of Surrey, University of Oxford, and the 8th CSEF-IGIER Symposium on Economics and Institutions. 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, FASS, Orta Mahalle, Universite Caddesi No:27 Tuzla, 34956 Istanbul.

1 Introduction

Transfer taxation has long been a highly controversial issue. Except for a brief period of time, the US 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. In virtually all traditional models of bequests, efficiency calls for either *zero or a negative tax* on parental transfers.[2] 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 is that different generations agree about intertemporal trade-offs. 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 (e.g., Bernheim (1989)). If the society cares only about the parent's welfare 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 redistribution or financing government expenditures. Following the same arguments, whenever society attaches direct welfare weight to future generations, parental transfers should actually be *subsidized* according to the altruistic model with intergenerational agreement.[3]

However, disagreements are common to most, if not all, parent-offspring relationships.

[1] Estate tax rates in the United States have been varying since the time they were introduced in 1916. For detailed information on the evolution of estate taxes in the United States, see Jacobson, Raub, and Johnson (2007). 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 has held steady at 40%. 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. For an overview on parental transfer taxation in OECD countries, see Cremer and Pestieau (2011).

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

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

Clarke, Preston, Raksin, and Bengtson (1999) study a wide range of disagreement patterns between older parents and adult children and develop a typology of disagreement issues. They consider a random sample of 496 parents (average age 62) and 641 children (average age 39) and ask about possible sources of disagreement.[4] More than 70% of the respondents report disagreements (about the same percentage among children and parents). The largest category of responses about conflict (38% among parents and 30% among children) is labeled as "Habits and Lifestyle Choices," and it includes sexual activity/orientation, living arrangements, quality of life, and allocation of resources and/or education. A conclusion the authors reach about this category of conflicts is that intertemporal allocation of resources is a common source of intergenerational disagreement. The following quote summarizes this point:

"There are also conflicts that seem to express a world view common to many in the older generation. This same father writes: 'He [39] wants all things like his generation of baby boomers, right now - new cars, new houses, vacations - all of it on one income and that a blue collar job income.' This is echoed elsewhere in another father's (60) comment over his daughter's (37) lack of 'concern for saving something for a rainy day.' Another father (71) reports that his son's (37) 'using credit cards to the limit' is an area of disagreement."

In this paper, we analyze optimal taxation of parental transfers in a world in which parents and offspring disagree on the intertemporal allocation of resources. We show that there is a genuine *efficiency* reason for government intervention in the market in the presence of intergenerational disagreement. In particular, we find that, whenever offspring are impatient from their parents' perspective, optimal intervention involves a positive tax on parental transfers.

We study transfer taxation using an intergenerational model in which bequests are motivated by altruism. The key principle for efficient bequest taxation can be fully grasped in the baseline model where agents live for two periods. In the first period, people make consumption-saving decisions, and in the second period, they choose how much to consume and bequeath to their offspring. Then, they are replaced by their offspring who go through the same life cycle. We model disagreement in a way that minimally departs from the standard model: a parent and an offspring agree on everything except for how the offspring should allocate his resources between his young and old age.

[4] The exact question was: *"No matter how well two people get along, there are times they disagree or get annoyed about something. In the last few years, what are some things on which you have differed, disagreed, or been disappointed about (even if not openly discussed) with your child (or parent)?"*

In this environment, we first characterize laissez-faire market equilibrium. We focus on Markov equilibria.[5] Using this characterization, we prove that equilibrium allocation is Pareto inefficient. The inefficiency stems from the fact that, as long as there is disagreement between parents and offspring, the latter do not fully internalize the consequences of their saving decisions on the former.

Having uncovered an efficiency reason to intervene in the market outcome, we next characterize optimal policy. We begin our analysis of optimal policy by focusing on a particular Pareto-efficient allocation that we call the "Ramsey" allocation. In this allocation, all the welfare weight is on the initial parent, and future generations are valued by society only indirectly via the initial parent's welfare. The Ramsey allocation is probably the most widely adopted benchmark in the literature. This choice of the benchmark efficient allocation is further motivated by the fact that in the standard model with no disagreements, a Ramsey government would find it optimal not to distort the equilibrium transfer decisions of parents at all. In this sense, any need for an intervention in the case with intergenerational disagreements will be coming from nowhere else but the existence of disagreements.

We find that, if children are less patient than what their parents want them to be, then parents bequeath too much in laissez-faire equilibrium relative to the Ramsey allocation, and it is optimal for the government to correct their behavior through bequest taxation. In short, the intuition for this result is as follows. Because of disagreement, in the laissez-faire equilibrium, the offspring save less than what their parents (and the planner) prefer. In particular, the parental welfare goes up if the parent can make the offspring increase his savings, which is possible by increasing bequests as long as the offspring's optimal saving policy is increasing in the amount of bequests received. In other terms, bequeathing has an additional benefit for the parents. This induces parents to transfer more than the Ramsey level.

In order to make children save the amount dictated by the Ramsey allocation, 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

[5]We cannot rule out the existence of multiple equilibria, but importantly, all our results are valid for all Markov equilibria.

change the gross return to their offsprings' savings. As a result, the parents still transfer too much to their offspring and, hence, should be taxed. Even though we find the case in which children are impatient from their parents' perspective more natural, we also analyze the case in which they act more patiently than what their parents would like, and we find that bequests should be subsidized in that case.

Next, we generalize our results to the whole Pareto frontier by analyzing optimal bequest taxation in the case in which society cares directly about future generations. The Pareto-optimal bequest tax can be decomposed between an "efficiency component" and an "intergenerational redistribution component". The efficiency component is the same as the Ramsey bequest tax: it is only present when there is intergenerational disagreement and calls for a tax whenever offspring are impatient from their parents' perspective. The second component, also present in models without intergenerational disagreements, represents a subsidy to parental transfers arising from intergenerational redistribution coming from a direct weight on future generations. The overall sign of the bequest tax depends on whether the efficiency wedge or the intergenerational redistribution wedge dominates. It is important, though, to stress that as long as offspring are impatient from their parents' perspective, the efficiency component is positive for any Pareto-efficient allocation and, hence, calls for a tax on parental transfers.

Our results extend to the case in which agents live for an arbitrary number of periods and parents coexist with their offspring. The optimal tax on inter vivos transfers obeys the same principles as the bequest tax and, thus, has the same sign. We also study the role of horizontal (cross-sectional) income heterogeneity by embedding our model into an optimal labor income taxation framework à la Mirrlees (1971). We find that although the optimality of a positive bequest tax is still dictated by the presence of intergenerational disagreements, the horizontal redistribution motive shapes the *curvature* of the tax schedule. In particular, when the income process is mean reverting across generations and society puts direct welfare weight on the offspring, the optimal bequest tax rate is typically increasing in the amount of bequests. The progressivity of bequest taxes stems from a mechanism similar to that in Farhi and Werning (2010), except that they find that optimality calls for progressive *subsidies* on bequests and not taxes. This is because they do not have intergenerational disagreement, which is the key driving force of bequest *taxes* in our environment.

In addition to its policy implications, this paper also makes a methodological contribution by deriving normative implications of models with intergenerational disagreements in the absence of differentiability assumptions. In the public finance literature, it is customary to compare equilibrium and efficient allocations by defining and measuring wedges that represent discrepancies between individual and social marginal returns to individual decisions. Our results, on the other hand, are presented over three levels of analysis. In addition to analyzing optimal wedges (which requires the assumption of differentiability) and optimal linear taxes (which requires further restrictive assumptions implying concavity of the agents' problems) we provide a nondifferentiable 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 disagreements for at least two reasons. First, it is well known that, in general, these models may not have equilibria with differentiable policy functions.[6] 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 is linked to a specific equilibrium (or equilibrium property such as differentiability or linearity of the policy).

In our implementation result, we show that when we focus our attention on linear Markov equilibria, the parental transfer wedge translates into a result on transfer taxation: efficient allocations can be implemented using linear wealth transfer taxes as long as the government has access to (linear) life-cycle saving taxation to offset offsprings' tendency to undersave. 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. It is important to note that the linear structure of taxes is not crucial for the optimality of bequest taxes. As long as the saving subsidy applied to offspring leaves their saving policy function strictly monotone in the amount of transfers they receive, parents have a motive to bequeath more than what is efficient, which implies it is optimal to tax bequests.

Related Literature. This paper is related to three strands of literature. The first is the literature on optimal taxation of bequests and inter vivos transfers.[7] Our contribution here is to provide a novel, pure efficiency argument for taxing parental transfers.[8] In addition to the

[6]Our model with intergenerational disagreements has an analytical structure that is quite similar to models with present-bias problems *à* la Laibson (1997), and it is well known that in these models, policy functions that describe people's life-cycle saving behavior may not be differentiable. See Harris and Laibson (2002).

[7]See Cremer and Pestieau (2011), Kaplow (2001), and Kopczuk (2009) for excellent surveys of the literature on optimal transfer taxation.

[8]It is, of course, possible to justify taxation of parental transfers based on equality grounds when other

altruistic model that is already discussed, a widely used model of bequests is the warm-glow (or "joy of giving") model. In this model too, the optimal bequest tax is zero or negative (i.e., a subsidy), depending on whether society cares about the offspring directly (e.g., Kopczuk (2009)). 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. According to this model, bequest taxes are simply a good way to finance positive government expenditures when lump-sum taxes are not 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.[9]

Our paper is also related to the literature on intergenerational disagreements. A seminal paper in this literature is Phelps and Pollak (1968), which analyzes the national saving rate in an environment in which each generation lives for a single period and is imperfectly altruistic: the rate at which each generation discounts the next generation's consumption relative to their own consumption is higher than the rate at which they discount consumption across any two subsequent future generations. Assuming that people have CEIS utility functions, returns to capital are linear and there is no depreciation, and focusing only on the equilibrium in which all generations save the same constant fraction of their income at all periods (i.e., a linear Markov equilibrium), the paper shows that equilibrium entails lower national saving compared with that in the Ramsey allocation.[10] Doepke and Zilibotti (2014) analyze an environment in which parents and children have preference disagreements, and parents can affect offspring's choices in two ways: by influencing their preferences via education and by imposing direct restrictions on their choice sets. They use this model to explain the variation in parenting styles across industrialized countries and over time. Finally, Doepke and Tertilt (2009) analyze a model where husbands and wives are imperfectly altruistic toward each other and disagree on the degree of their altruism toward their offspring. The authors use this model to explain the dramatic improvements in the legal rights of mar-

instruments of horizontal redistribution are limited. See, for instance, Piketty and Saez (2013).

[9]An obvious theoretical assumption - not yet carefully tested empirically - 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. See Kopczuk (2009) for a discussion of negative wealth externalities.

[10]Following Strotz (1955), Laibson (1997) applies this framework to individual consumption saving problem over the life cycle under self-control problems and also finds undersaving behavior. See O'Donoghue and Rabin (1999) for a more general application of this model to individual decision making.

ried women that occurred before the introduction of female suffrage. We contribute to this growing literature by analyzing the policy implications of intergenerational disagreements.

One possible interpretation of our positive model is that offspring agree with their parents regarding how much they *should* save, but they face self-control problems that prevent them from saving the right amount. Under this interpretation, the optimal tax problem is a paternalistic one in the sense that taxes are used to correct the 'wrong' saving behavior of the offspring. This is the focus of 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 the paper of Krusell, Kuruscu, and Smith (2010), which analyzes properties of linear taxes on life-cycle savings that implement the Ramsey allocation. Pavoni and Yazici (2015) also focus on optimal Ramsey taxation of life-cycle savings within a quasi-hyperbolic discounting model and allow for self-control problems that change with age. The current paper, on the other hand, does not assume that offsprings' true welfare coincides with that of their parents. The *offspring are truly more impatient than what parents want them to be*. In this environment, we show that, independent of which member society cares about, it is optimal to tax parental transfers. The optimality of bequest taxation does not stem from correcting people's mistakes (because there are none), but rather from externality that arises from intergenerational disagreements.[11] Since the analytical structures of the disagreement and self-control models are very similar, we also provide a methodological contribution to the self-control literature with our nondifferentiable analysis of optimal policy.

The rest of the paper is organized as follows. Section 2 introduces the baseline 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 provide a number of important generalizations of our result, including the Pareto-efficient taxation of inter vivos parental transfers. Section 6 concludes.

2 Model

The economy is populated by a continuum of a unit measure of dynasties that live for a countable infinity of periods, $t = 0, 1, ...$, where each agent within a dynasty is active for two

[11] We thank the editor for suggesting the intergenerational disagreement interpretation.

periods. In the first period of their lives, agents are young adults and make consumption saving decisions. In the second period, they become parents, decide how much to consume and bequeath, and die. The next period their offspring become young adults and go through the same life cycle. This is a model of non-overlapping generations.[12] People have one unit of time endowment that they supply inelastically to the market every period. Parents bequeath because they are altruistic.

The economy begins with an initial parent in period 0. Every subsequent even period is a parenthood (old adulthood) period, whereas every odd period is a young adulthood period. 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 $\delta, \gamma \in (0, 1)$, 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 parameter δ represents the discount factor that the parent thinks the offspring and all the future descendants should save according to during their young adulthood period. The parameter γ is the altruism factor.

We model disagreement between parents and offspring in a way that minimally departs from the standard dynastic framework: namely, the offspring agrees with the parent on everything except for how to allocate his wealth between young and old adulthood. Specifically, the offspring's preference is given by

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

with $\beta > 0$. In this formulation, as long as $\beta \neq 1$, the discount rate that the offspring uses between periods $t+1$ and $t+2$, $\beta\delta$, is different from what is appropriate from his parent's perspective, δ. Parents are sophisticated in the sense that they fully anticipate this discrepancy between their own and their offsprings' preferences.[13]

We do not make an assumption about whether β is smaller or larger than one from the

[12]In Section 5.3, we allow for a longer life cycle and show that our main results regarding bequest taxation are robust to such an extension. There, we also model periods in which parents and their offspring are alive together and analyze inter vivos transfer behavior and taxation. We show that the results regarding bequest taxation extend to inter vivos transfers.

[13]In Appendix D, we analyze what happens if parents naively believe that their offspring agree with them regarding intergenerational resource allocation and show that our main results carry over to this case as well.

outset. However, both the Clarke, Preston, Raksin, and Bengtson (1999) study and anecdotal evidence seem to suggest that it is more natural to consider the case in which $\beta < 1$. This is the case in which the offspring are impatient from the parents' perspective. It is important to note that Pareto inefficiency of laissez-faire equilibrium allocation does not depend on whether the offspring are more or less patient than what the parents would like them to be. We report how the sign of optimal transfer taxes depends on whether β is greater or less than unity throughout the paper.

The benchmark model has an alternative interpretation. Under this interpretation, even though the offspring's true preference coincides with that of his parents, that is, he evaluates his welfare according to $u(c_{t+1}) + \delta V_{t+2}$, the offspring faces self-control problems and saves according to $u(c_{t+1}) + \beta\delta V_{t+2}$. This interpretation is in line with the literature on self-control problems in the spirit of Laibson (1997). In the current paper, on the other hand, we assume that the offspring and the parents *truly disagree* in the sense that $u(c_{t+1}) + \beta\delta V_{t+2}$ represents not only the behavioral preference of the offspring but also his true preference.

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 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 endowment of the initial parent in period 0, the feasibility for any $t \geq 0$ is

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

As evident from the previous feasibility condition, we assume there is one representative dynasty, which implies that in any calendar year, only one age group is alive. We could instead allow for members of different dynasties to be at different points in their life cycles. This would not change any of our results. 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). In the main body of the paper, we abstract from such "horizontal" distribution issues in order to isolate our mechanism. We show in Section 5.2 that the mechanism behind our results is robust to income heterogeneity.

3 Laissez Faire

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 onward.

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

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

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:[14]

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

$B(Q_t)$ is the "natural" (and never binding) borrowing limit defined by requiring consump-

[14]To save notation, we indicate the policy as a function. In case there are multiple solutions to the offspring's problem, $b_{t+2}(\cdot, Q_{t+1})$ should be intended as a selection from the policy correspondence.

tion 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

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

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, $b_{t+2}(\cdot, Q_{t+1})$, which describes how the offspring's saving choice changes as a function of parental bequests under a given price sequence. The parent is sophisticated in the sense that he correctly guesses his child's choice, and that is why he takes (2) into account as a constraint 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 value function and the policies are consistent with the parent's and offspring's problems (1) and (2); (ii) the prices satisfy (3); (iii) markets clear:* $b_t = k_t$ *for all* t.

Proposition 1 characterizes equilibrium parental transfer behavior. Proving Proposition 1 would be relatively easier if we could assume the differentiability of policy function, $b_{t+2}(\cdot, Q_{t+1})$, in bequests received. However, the dynastic intertemporal resource allocation problem with disagreements across dynasty members implies that agents play dynamic games, and it is well known (e.g., from the self-control literature) that in such environments, we cannot guarantee even the continuity of the policy functions even when we focus our attention on Markov equilibria.[15] To ensure that Proposition 1 describes a general feature

[15]See Morris and Postlewaite (1997) and Harris and Laibson (2002) for examples of economies with quasi-hyperbolic discounters where policy functions are discontinuous. Krusell and Smith (2003) show that even when we focus our attention on Markov equilibria, it is not possible to rule out the existence of discontinuous equilibria. Notice that the discontinuity in these models arises from a "disagreement" across an agent's multiple selves, whereas in our model the disagreement is across different generations.

of economies with intergenerational disagreements and is not an artifact of differentiability assumptions, we prove it without making any differentiability or continuity assumptions about the value or policy functions.

Proposition 1. *Suppose $\beta < 1$. Then, in equilibrium, in all parenthood periods t,*

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

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

$$u'(c_t) \leq R_{t+1}\gamma u'(c_{t+1}), \tag{5}$$

with strict inequality whenever the offspring's saving policy is strictly monotone in b_{t+1}.
If $\beta = 1$, then

$$u'(c_t) = R_{t+1}\gamma u'(c_{t+1}). \tag{6}$$

Proof. Relegated to Appendix A.1. □

We provide an intuition for Proposition 1 only for the $\beta < 1$ case; the intuition for the $\beta > 1$ case is symmetric. To get a better grasp on what the proposition says, first focus on the case in which the child and parent agree on intertemporal trade-offs, meaning $\beta = 1$. In that case, the parent chooses the level of transfers to equate the marginal cost of his forgone consumption (left-hand side of (6)) to the marginal benefit of his child's increased consumption in period $t+1$ (right-hand side of (6)). However, when $\beta < 1$, 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 $t+1$. The parent does this because bequeathing has an additional marginal benefit from the parent's perspective when the offspring is impatient. Intuitively, when $\beta < 1$, the offspring is undersaving from the parent's perspective, meaning that c_{t+1} is higher than that is desired by the parent while c_{t+2} is lower. As a result, in the eyes of the parent, a marginal unit saved by the offspring has a marginal cost $-u'(c_{t+1})$ that is lower than its marginal return $\delta R_{t+2}u'(c_{t+2})$. This implies that the parental welfare goes up if the parent can make the offspring increase his savings, which is possible by increasing bequests as long as the offspring's optimal saving policy is strictly increasing in the amount of bequests received. This is why increasing bequests carries an additional benefit. We will provide a sharper marginal characterization of equilibrium bequest behavior in Section 3.2, where we assume differentiability of the value and policy functions. This sharper characterization will also enable us to sharpen the intuition explained earlier. Before that, we prove in Section 3.1 that laissez-faire equilibrium is inefficient for the general case without assuming differentiability.

3.1 The Inefficiency of Laissez-Faire Equilibrium

Proposition 2 below states and proves that as long as $\beta \neq 1$ the laissez-faire equilibrium allocation is Pareto inefficient. The proof provides a resource-feasible perturbation of the equilibrium allocation that improves the welfare of some agents strictly without hurting others. The proof does not use any differentiability or continuity assumptions regarding the equilibrium value or policy functions. This is important because we know from Krusell and Smith (2003) that in models similar to ours, there can be many Markov equilibria, some of which might have discontinuous policy functions. By not making any assumptions on continuity or differentiability, our proof ensures that Pareto inefficiency of equilibrium is a general phenomenon that does not apply only to differentiable equilibria.

Proposition 2. *Suppose $\beta \neq 1$. Then, the laissez-faire equilibrium allocation is Pareto inefficient.*

Proof. Relegated to Appendix A.2. □

The intuition for the inefficiency of equilibrium is as follows. As is evident from equation (7), the parent's welfare, V_t, depends on how much the offspring saves. In this sense, there is a consumption externality. Obviously, this is true in the standard altruistic model without disagreements as well (see Bernheim (1989)). In the standard case, however, since the offspring fully agrees with the parent, in a sense the offspring internalizes the consequences of his saving on the parent, so the externality does not have a consequence. This can be seen by setting $\beta = 1$ in equation (7) :

$$V_t \quad = \quad u(c_t) + \gamma \Big[\underbrace{u(c_{t+1}) + \beta\delta V_{t+2}}_{\text{offspring welfare}} + (1-\beta)\delta V_{t+2} \Big] \tag{7}$$

When $\beta \neq 1$, the offspring does not fully internalize the consequences of his saving on parental welfare. The level of savings he chooses affects the term $(1-\beta)\delta V_{t+2}$ in (7), but this term is external to the offspring. This externality is the reason why the equilibrium is inefficient. In the case in which $\beta < 1$, the term that creates the externality asks for a higher saving rate. In this case, the planner can improve both agents' welfare by forcing the offspring to increase his savings by a small amount. This creates a second-order loss for the offspring, since he was already at his optimal allocation. It creates a first-order gain for the parent, since the equilibrium level of the offspring's saving is strictly suboptimal from the parent's perspective. Then, the planner simply transfers a suitable amount from the parent to the offspring to compensate for the second-order decline in the offspring's welfare, achieving the desired Pareto improvement.

It might be useful to relate the inefficiency result in our disagreement economy to what is obtained in an economy with self-control problems. According to the self-control interpretation, the offspring wants to save according to δ discounting but lacks self-control and ends up saving according to $\beta\delta$ discounting. In this context, increasing the offspring's savings by a small amount improves the welfare of both the offspring and the parent, implying that the original equilibrium cannot be efficient. In presence of self-control problems, the inefficiency arises because some agents cannot carry out actions that are optimal from everybody's perspective. In our model with disagreement instead, the $\beta\delta$ discounting represents the offspring's true preference, and thus, a perturbation that simply increases the offspring's saving rate cannot Pareto improve over equilibrium because it makes the offspring worse off. As we explained above, the inefficiency of laissez-faire equilibria in our disagreement economy comes from the existence of a *consumption externality* that arises from the presence of disagreement.

3.2 Equilibrium Parental Behavior under Differentiability

In this section, we provide a marginal condition that characterizes equilibrium bequest behavior, assuming differentiability of the policy functions that describe the 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}. Now consider a parent's problem. The parent chooses b_{t+1} subject to the flow budget constraints and the function $b_{t+2}(\cdot, Q_{t+1})$, defined by (2), which describes the offspring's saving decision. The parent's first-order optimality condition with respect to the bequest decision, b_{t+1}, is

$$u'(c_t) = \gamma \left(u'(c_{t+1}) \left[R_{t+1} - \frac{\partial b_{t+2}(b_{t+1}, Q_{t+1})}{\partial b_{t+1}} \right] + \delta V_1(a_{t+2}, Q_{t+2}) R_{t+2} \frac{\partial b_{t+2}(b_{t+1}, Q_{t+1})}{\partial b_{t+1}} \right), \tag{8}$$

where V_1 refers to the derivative of the value function with respect to its first argument and the derivatives are all evaluated at the equilibrium allocation.

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

$$u'(c_{t+1}) = \beta \delta V_1(a_{t+2}, Q_{t+2}) R_{t+2}. \tag{9}$$

Using (9) in the parental optimality condition (8), we get the following proposition, which describes equilibrium parental bequeathing behavior under differentiability.

Proposition 3. *Suppose $b_{t+2}(\cdot, Q_{t+1})$ is differentiable in b_{t+1}. Then, in any parenthood period t, the equilibrium bequest behavior is characterized by*

$$u'(c_t) = \gamma \left(R_{t+1} u'(c_{t+1}) + \frac{\partial b_{t+2}(b_{t+1}, Q_{t+1})}{\partial b_{t+1}} u'(c_{t+1}) \left[-1 + \frac{1}{\beta} \right] \right). \tag{10}$$

Equation (10) 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 changing the offspring's life-cycle consumption pattern. It is a multiplication of two terms: the first term,

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

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 12 in Appendix A.1, under the assumption of differentiability of $b_{t+2}(\cdot, Q_{t+1})$, this derivative is strictly positive.[16]

The second term,

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

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

Intuitively, when $\beta < 1$, the parent knows that from his perspective, the offspring is undersaving. More precisely, since the offspring's Euler equation implies that $-u'(c_{t+1}) + \beta\delta V_1(a_{t+2}, Q_{t+2}) R_{t+2} = 0$ (see equation (9)), the net return of a marginal unit of savings in the eyes of the parent is positive: $-u'(c_{t+1}) + \delta V_1(a_{t+2}, Q_{t+2}) R_{t+2} > 0$. So, parental welfare goes up if the parent can make the offspring increase his savings, which is possible by increasing bequests, since $\frac{\partial b_{t+2}(b_{t+1}, Q_{t+1})}{\partial b_{t+1}} > 0$. As a result, the additional term in (10) 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

In Proposition 2, we show that the laissez-faire equilibrium is unambiguously Pareto inefficient. The policy implications of this result might obviously depend on the particular Pareto-efficient allocation that the policy targets. In this section, we start our optimal bequest tax policy analysis by targeting a widely adopted benchmark Pareto-efficient allocation, namely the "Ramsey" allocation. The Ramsey allocation is the efficient allocation that puts all the weight on the initial generation parent. Thus, it is given by the solution to a fictitious social planner's consumption-saving problem where the planner discounts exponentially at rate δ between young and old adulthood and discounts future generations by altruism factor γ.

The Ramsey allocation has at least three desirable properties. First, it corresponds to the *equilibrium allocation* in the absence of conflict about intertemporal trade-offs. Thus, in the absence of disagreements, the optimal policy is simply not to distort bequests at all. In that sense, in the case of Ramsey allocation, the optimality of bequest tax under disagreements is coming purely from the existence of disagreements. Second, it is unique and simple to

[16] To be precise, Lemma 12 proves that as long as the value function is differentiable, the policy is strictly monotone. The differentiability of the value function is implied by the differentiability of the policy function by the implicit function theorem.

characterize as it is the dynastic solution to the standard Ramsey-Cass-Koopmans optimal growth problem. Finally, the saving behavior prescribed in the Ramsey allocation - saving according to discount factor δ - is optimal from the perpective of all the agents in the dynasty except for the young adult making the saving decision.[17]

We first characterize the Ramsey allocation. Then, we show that the Ramsey level of bequests does not satisfy the parent's optimality condition in the laissez-faire equilibrium: in particular, parents bequeath too much relative to the Ramsey level when offspring are impatient from parents' perspective. This implies that implementing the Ramsey allocation in the market requires government intervention. We show that, under differentiability assumptions on offsprings' saving policy function, this result translates into an optimal positive bequest wedge. 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. When $\beta < 1$, the optimal tax on bequests is positive.

In Section 5.1, we generalize our results to all the allocations on the Pareto frontier. There, we show that the optimal bequest wedge has a nice separable form between an efficiency component and an intergenerational redistribution component. The principle that governs the efficiency component of the wedge at any point on the Pareto frontier is identical to the one that determines the "Ramsey" bequest wedge that we study in the present section.

4.1 The Ramsey Allocation

The Ramsey allocation is unique and is given by the solution to a fictitious social planner's consumption-saving problem where the planner has altruism and discount factors γ and δ, respectively. The following Euler equations characterize the Ramsey levels of bequests and savings, which we denote with an asterisk. For all t even, we have:

$$u'(c_t^*) = \gamma f'(k_{t+1}^*)u'(c_{t+1}^*), \tag{11}$$

$$u'(c_{t+1}^*) = \delta f'(k_{t+2}^*)u'(c_{t+2}^*). \tag{12}$$

[17] Evaluating welfare from the perspective of the initial self is also the route taken in much of the related literature on self-control problems. See DellaVigna and Malmendier (2004), Gruber and Koszegi (2004), and O'Donoghue and Rabin (2006), for example.

4.2 The Ramsey Wedge

Now we turn to the implications of Proposition 1 for equilibrium wealth transfer behavior relative to the Ramsey allocation. A comparison of the Ramsey condition for bequests, (11), with the laissez-faire condition, (4), together with the pricing condition $R_{t+1} = f'(k_{t+1})$, implies that the equilibrium level of parental transfers does not satisfy the Ramsey condition for bequests in general. As we previously explained, when $\beta < 1$, in the laissez-faire equilibrium, whenever the offsprings' policy is strictly monotone, bequeathing has an additional benefit to the parents since their offspring are saving too little to begin with. As a result, a parent with a Ramsey level of wealth and facing Ramsey prices bequeaths more than the Ramsey level.

If we assume differentiability of policy functions, this argument translates into a positive bequest wedge for $\beta < 1$. To see this, first define the *Ramsey bequest wedge* as

$$BW_{t+1}^* := 1 - \left(\frac{u'(c_t^*)}{\gamma u'(c_{t+1}^*) f'(k_{t+1}^*)} - \frac{\partial b_{t+2}(k_{t+1}^*, Q_{t+1}^*)}{\partial b_{t+1}} \frac{\left[-1 + \frac{1}{\beta}\right]}{f'(k_{t+1}^*)} \right), \tag{13}$$

where Q_{t+1}^* corresponds to the price sequence implied by the Ramsey allocation, that is, $Q_{t+1}^* := \{R_{t+s}^*, w_{t+s}^*\}_{s=1}^{\infty} = \{f'(k_{t+s}^*), f(k_{t+s}^*) - f'(k_{t+s}^*)k_{t+s}^*\}_{s=0}^{\infty}$, and $b_{t+2}(\cdot, Q_{t+1}^*)$ is the equilibrium policy function of the offspring under Q_{t+1}^*. Recall that parents bequeath according to (10) in equilibrium. The Ramsey bequest wedge in period $t+1$ measures the efficient marginal distortion that the planner needs to create in the return to bequests, $R_{t+1}^* = f'(k_{t+1}^*)$, in order to make the parent choose the Ramsey level of bequests in equilibrium. A positive (negative) BW_{t+1}^* means that a parent facing the Ramsey interest rate, R_{t+1}^*, would like to increase (decrease) his bequests marginally above (below) the Ramsey level if there is no government intervention.

Corollary 4. *Suppose $b_{t+2}(\cdot, Q_{t+1})$ is differentiable in b_{t+1} and $\beta < 1$. Then, for any parenthood period t, $BW_{t+1}^* > 0$.*

Proof. Using equation (11) in the definition of BW_{t+1}^*, we get

$$BW_{t+1}^* = \frac{\partial b_{t+2}(k_{t+1}^*, Q_{t+1}^*)}{\partial b_{t+1}} \frac{\left[-1 + \frac{1}{\beta}\right]}{f'(k_{t+1}^*)}. \tag{14}$$

Lemma 12 in Appendix A.1 shows that when $b_{t+2}(\cdot, Q_{t+1})$ is differentiable in b_{t+1}, then $\frac{\partial b_{t+2}(k^*_{t+1}, Q^*_{t+1})}{\partial b_{t+1}} > 0$. $BW^*_{t+1} > 0$ then follows from $0 < \beta < 1$. $\square$

Both the Ramsey planner and the parents in equilibrium face a "technological" marginal return equal to $f'(k^*_{t+1})$ when bequeathing. We know, however, that the undersaving behavior of the offspring creates an additional return to bequeathing for the parents. In order to align the parent's return with the Ramsey return, the planner needs to reduce the return (interest rate) the parent faces by (14), which is equal to the additional return normalized by the Ramsey return. Notice that the bequest wedge is *strictly* positive, since, under the differentiability assumption, the offspring's saving is strictly increasing in the bequests received, which implies that the additional return to bequeathing is strictly positive.

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

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

subject to

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

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

Proof. The linearity of the policy functions 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. It is easy to derive the optimality condition for parental bequest choice under taxes, analogous to (10):

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

Substituting in the Ramsey allocation and using (11) gives the expression for the optimal bequest tax in the statement of Proposition 5. The fact that $M_{t+2}(\Psi_{t+1}^{*})$ is strictly positive follows from the differentiability of the offspring's policy function. □

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.[18]

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

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

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 Appendix A.3. □

It is interesting to note that even though the government corrects offsprings' saving behavior through saving taxes, parental transfers should still be taxed to achieve the Ramsey

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

allocation. Here, one might ask: given that from the Ramsey perspective, the offspring is saving the right amount (thanks to corrective taxes), why does the parent still bequeath more than the Ramsey level? This occurs because, from the parent's perspective, the offspring is still 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 government sets the lump-sum taxes to balance its budget in period $t+2$, meaning that it sets $T_{t+2} = R_{t+2}\tau_{t+2}b_{t+2}$. Thus, from the government's perspective, the return to the offspring's savings in period $t+1$ is actually R_{t+2}, and the taxes are there only to drive the offspring 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. This is why the offspring, who save at the Ramsey level under the optimal saving subsidy, are still undersaving from the parents' perspective. As a result, parents still have a motive to transfer more than the Ramsey level. To discourage this, the government finds it optimal to tax bequests. Another way of interpreting this result is as follows. By bequeathing an extra unit to the offspring, parents increase offsprings' savings, which increases the cost of financing the subsidy for the government. In this sense, bequests generate a negative "fiscal externality". The tax on bequests can be seen as a Pigouvian tax internalizing this externality.

If the government (or the parents) could command the offspring to the Ramsey allocation, then there would be no need to distort parents' bequest decision. A natural question that follows, then, is whether the bequest tax result is peculiar to the assumption that the government is restricted to using linear taxes to discipline offsprings' savings. The proof of Proposition 1 shows that as long as the offspring react to an increase in bequests by increasing their savings, parents do have an extra return to bequeathing and, hence, will bequeath too much if there is no government intervention. Therefore, as long as the government uses a kind of tax policy that leaves offsprings' equilibrium savings strictly 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 completely eliminate the monotonicity of offsprings' 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 bequest payments - by definition - occur after parents die and saving decisions might be difficult to monitor, especially by third parties. In this sense, the "Rotten Kid Theorem" of Becker (1974) does not hold in our environment by assumption.

4.4 A Back-of-the-Envelope Quantitative Analysis

Our main result states that, as long as offspring - the recipients of intergenerational transfers - are too impatient from their parents' perspective, parents transfer too much, and hence, to restore efficiency, bequests should be taxed.

In this section we aim at giving a rough idea of the quantitative importance of our mechanism for bequest taxation. 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.

Proposition 7. *Suppose $u(c) = \log(c)$. Then, the optimal tax on the bequest received by offspring in any period $t+1$ is given by*

$$\tau_{t+1}^{*} = \frac{\delta(1+\gamma)(1-\beta)}{1+\delta}.$$

Proof. Relegated to Appendix A.4. □

Observe that, under the logarithmic utility function assumption, the tax formula does not depend on the shape of the production function, F, or the depreciation rate, θ. Therefore, we do not need to specify values for these parameters. The only parameter values that are needed to calculate optimal taxes are the discount factor, δ, the altruism factor, γ, and the parameter, β, that represents the degree of disagreement. We assume perfect altruism in the sense that parents discount offsprings' consumption as much as they discount their own consumption, (i.e., $\gamma = \delta$). In this case, we have $\tau_{t+1}^{*} = \delta(1-\beta)$. In quantitative work, it is common to calibrate the discount factor using the annual real interest rate as a target. Under the assumption that the interest rate is constant at R and equilibrium allocation is stationary in the sense that consumption across young adulthood periods and parenthood periods are

Table 1: Optimal Bequest Taxes

β	0.5	0.6	0.7	0.8	0.9	0.95
Calibrated δ	0.571	0.545	0.524	0.506	0.491	0.484
Bequest tax	28.55%	21.80%	15.72%	10.12%	4.91%	2.42%

constant over time, one can show that β, δ, and R have to jointly satisfy

$$\frac{\delta^2\beta}{1-\delta+\beta\delta}R^2 = 1.^{19} \tag{15}$$

We take the annual real interest rate to be 3% and assume that each period in our model corresponds to 25 years, implying $R = 1.03^{25}$. Table 1 reports the calibrated values of δ in the second row and the corresponding optimal bequest tax rates in the third row for different values of β specified in the first row. A glance at Table 1 seems to reveal that the optimal taxes generated by our mechanism has the potential of being quantitatively significant.

5 Extensions

In this section, we consider several extensions of the baseline framework. First and foremost, we show that the optimality of creating a positive bequest wedge shown in Section 4 for the case of Ramsey allocation holds for other allocations on the Pareto frontier. Second, we extend our model by introducing horizontal inequality and show that bequest taxation is still optimal. Moreover, we show that optimal bequest taxes are typically progressive. Finally, we show that our normative results are also robust to extending people's life cycle to any finite periods and allowing parents and offspring to coexist in the same period. This also allows us to show that it is optimal to tax inter vivos transfers as well. The optimality

[19] Using the usual envelope condition, offsprings' equilibrium optimality condition for saving (9) becomes $u'(c_{t+1}) = R\beta\delta u'(c_{t+2})$. This condition, together with the parental optimality condition (10) evaluated at $\gamma = \delta$ and the steady-state requirement that $c_t = c_{t+2}$, implies that

$$\delta^2\beta R^2\Big(1 + \frac{\partial b_{t+2}(b_{t+1}, Q_{t+1})}{\partial b_{t+1}}\frac{-1+1/\beta}{R}\Big) = 1.$$

It follows from equation (44) in the proof of Proposition 7 in Appendix A.4 that, under $\gamma = \delta$, $\frac{\partial b_{t+2}(b_{t+1},Q_{t+1})}{\partial b_{t+1}} = R\frac{\delta\beta}{1-\delta+\delta\beta}$. Together these imply condition (15).

of transfer taxation in the presence of illiquid bequests is also discussed.

5.1 Intergenerational Redistribution: Pareto-Efficient Bequest Wedges

The Ramsey allocation is one of the many possible Pareto-efficient allocations. Specifically, it is the efficient allocation that arises when the planner only cares about the initial parent directly. In this section, we study the whole Pareto frontier by allowing the planner to care about all generations directly, and we characterize Pareto-efficient bequest wedges.

We begin by characterizing the Pareto frontier of the economy. Let $\mathbf{U} = \{U_n\}_{n=1}^{\infty}$ be any given sequence of utilities where $U_n \in \mathbb{R}$ represents the minimum level of utility that needs to be delivered to n^{th} generation agents (who are born in calendar time $2n-1$), starting with the agent born in period 1. Let $\mathcal{U}$ be the set of such utility sequences that are achievable by feasible allocations.[20]

We compute a Pareto-efficient allocation by solving a planning problem where we maximize the utility of one agent, initial parent, subject to delivering each other generation the utility promised to them by $\mathbf{U}$. By changing $\mathbf{U}$ in the set of all feasible utility sequences $\mathcal{U}$, the solution to this planning problem traces the whole Pareto frontier of the economy.

We use a simple recursive formulation of the planning problem to characterize the Pareto frontier of the economy. If k is the level of capital at the beginning of the parenthood period in a given generation, $W : K \times \mathcal{U} \to \mathbb{R}$, is the value function that denotes the maximal utility for current parents that can be achieved subject to providing future generations a sequence $\mathbf{U}$ of utilities.[21] The value function is given by the solution to the following problem:

$$\begin{aligned} W(k, \mathbf{U}) = \max_{y,k'} \quad & u(f(k) - y) + \gamma[u(f(y) - k') + \delta W(k', \mathbf{U}')] \\ s.t. \qquad & 0 \le y \le f(k), \quad 0 \le k' \le f(y), \quad \mathbf{U}' = C(\mathbf{U}); \\ & u(f(y) - k') + \beta\delta W(k', \mathbf{U}') \ge U. \end{aligned} \tag{16}$$

[20] A key feature restricting the set $\mathcal{U}$ is the domain of $u : \mathbb{R}_+ \to \mathbb{R}$. If $u(0) > -\infty$, to be achievable, a sequence of utilities $\mathbf{U}$ must admit a sequence of capital levels $\{k_t\}_{t=0}^{\infty}$ such that, for all t, we have $0 \le k_{t+1} \le f(k_t)$ and for generation n who are born in period $s = 2n-1$

$$U_n = u\left(f(k_s) - k_{s+1}\right) + \beta\delta\Big\{u\left(f(k_{s+1}) - k_{s+2}\right) + \sum_{t=s+1}^{\infty} \gamma^{t-s}\delta^{t-s-1}\Big[u\left(f(k_{2t-1}) - k_{2t}\right) + \delta u\left(f(k_{2t}) - k_{2t+1}\right)\Big]\Big\}.$$

[21] $K \equiv [0, \bar{k}]$ is the domain of the level of capital stock. It is well known that, with a positive rate of depreciation and Inada on f at infinity, there is a $\bar{k}$ such that $f(k) < k \; \forall k \ge \bar{k}$.

In the previous notation, $\mathbf{U}' = C(\mathbf{U})$ indicates the continuation sequence of $\mathbf{U}$. The choice variables y and k' represent, respectively, the bequest decision of the parent and the saving decision of the offspring. Constraint (16) guarantees that each new generation gets a level of utility U, where U refers to the first term of $\mathbf{U}$.

Since the problem is concave in the choice variables for each $(k, \mathbf{U})$, it can be shown that the value function W is concave in k for each $\mathbf{U}$. Since u is differentiable, an application of the Benveniste and Scheinkman lemma implies that W is differentiable in the first argument for all $\mathbf{U} \in \mathcal{U}$.[22] Formally, let $g : K \times \mathcal{U} \to R_+$ be the policy function for y and $h : K \times \mathcal{U} \to R_+$ represent the saving policy of the offspring. The envelope theorem implies that

$$W_1(k, \mathbf{U}) = f'(k)u'(f(k) - g(k, \mathbf{U})),$$

whenever $0 < g(k, \mathbf{U}) < f(k)$, where W_1 refers to the derivative of W with respect to the first argument.

Concave problems admit Kuhn-Tucker multipliers. If we denote by λ the multiplier associated with constraint (16), the first-order conditions for optimality are

$$y \text{ (bequest)} : \quad u'(f(k) - y) = (\gamma + \lambda) f'(y) u'(f(y) - k') \tag{17}$$

$$k' \text{ (savings)} : \quad u'(f(y) - k') = \delta \frac{\gamma + \lambda\beta}{\gamma + \lambda} f'(k') u'(f(k') - y'). \tag{18}$$

Once we solve for the policy functions g and h, we can construct the Pareto-efficient sequence of capital levels in the usual way. Let k_0^* be the initial level of capital and $\bar{\mathbf{U}} = \{\bar{U}_n\}_{n=1}^{\infty}$ be the sequence of utilities guaranteed to the offspring determining a particular point on the Pareto frontier, with continuations $\bar{\mathbf{U}}_m = C\left(C\left(\ldots(\bar{\mathbf{U}})\right)\right) = \{\bar{U}_n\}_{n=m}^{\infty}$, $m \geq 1$. The sequence of capital levels associated with this particular efficient allocation can be recovered, recursively, as follows: $k_1^* = g(k_0, \bar{\mathbf{U}}_1)$, $k_2^* = h(k_1^*, \bar{\mathbf{U}}_1)$, $k_3^* = g(k_2^*, \bar{\mathbf{U}}_2)$, $k_4^* = h(k_3^*, \bar{\mathbf{U}}_2)$, and so on. The capital sequence $\{k_t^*\}_{t=0}^{\infty}$, together with (17), allows us to recover the sequence of multipliers as follows. If t is an even period, then we have

$$u'(f(k_t^*) - k_{t+1}^*) = (\gamma + \lambda_{t+1}^*) f'(k_{t+1}^*) u'(f(k_{t+1}^*) - k_{t+2}^*), \tag{19}$$

where $\lambda_{t+1}^* \geq 0$ represents the multiplier on the utility promise constraint, (16), for agent

[22] See Stokey, Lucas, Jr., and Prescott (1989), Theorem 4.10.

born in period $t+1$. This will be the multiplier associated with constraint (16) in the recursive problem, for $k = k_t^*$ (t even), and $\mathbf{U} = \bar{\mathbf{U}}_{\frac{t}{2}+1}$, where the latter is the relevant continuation of utility sequence from $\bar{\mathbf{U}}$.

We generalize the definition of the Ramsey bequest wedge given by (13) for a Pareto-efficient allocation indexed by $\mathbf{U}$:

$$BW_{t+1}^*(\mathbf{U}) \equiv 1 - \left(\frac{u'(c_t^*)}{\gamma u'(c_{t+1}^*) f'(k_{t+1}^*)} - \frac{\partial b_{t+2}(k_{t+1}^*, Q_{t+1}^*)}{\partial b_{t+1}} \frac{\left[-1 + \frac{1}{\beta}\right]}{f'(k_{t+1}^*)} \right),$$

where Q_{t+1}^* is the price sequence generated from the Pareto-efficient allocation as before and $\frac{\partial b_{t+2}(k_{t+1}^*, Q_{t+1}^*)}{\partial b_{t+1}}$ represents the derivative of period $t+1$ offsprings' policy with respect to the bequests received at k_{t+1}^* and Q_{t+1}^*. Using (19) in the previous definition of $BW_{t+1}^*(\mathbf{U})$ gives

$$BW_{t+1}^*(\mathbf{U}) = -\frac{\lambda_{t+1}^*}{\gamma} + \frac{\partial b_{t+2}(k_{t+1}^*, Q_{t+1}^*)}{\partial b_{t+1}} \frac{\left[-1 + \frac{1}{\beta}\right]}{f'(k_{t+1}^*)}. \tag{20}$$

The bequest wedge characterized by (20) has two components. Consider the second term on the right-hand side of (20). This term represents the efficiency component of the bequest wedge that arises from intergenerational disagreements. A comparison of this component with the Ramsey bequest wedge (14) shows that the two expressions are essentially identical. The efficiency component is positive, meaning that it calls for a tax on bequests, as long as the offspring is impatient from the parent's perspective ($\beta < 1$). The intuition is the same as in the Ramsey case. In particular, observe that the sole reason for the efficiency component is the presence of disagreement: the wedge disappears when there is no disagreement ($\beta = 1$).

In the Ramsey case, the efficiency wedge is the only component of the wedge. In general, the Pareto-efficient bequest wedge has another component, though. The first term on the right-hand side of (20) represents this component, which comes from *intergenerational redistribution*. The intergenerational redistribution wedge is also present in the standard altruistic model of bequests without disagreements.[23] It is indeed easy to see from (20) that this component exists even when $\beta = 1$. The intergenerational redistribution wedge is negative and thus calls for a subsidy on bequests. It is strictly negative as long as the multiplier, λ_{t+1}^*, is strictly positive. This is true as long as the constraint (16) binds, which is equivalent to the planner putting a direct welfare weight on the offspring born in period $t+1$.

[23]See Kaplow (1995), Kaplow (2001), and Farhi and Werning (2010).

Intuitively, the offspring's welfare enters the planner's objective through two channels: one indirectly through the parent's welfare and the other directly. Thus, the planner cares about the offspring more than the parent does, and for this reason the parent bequeaths too little from the planner's perspective. It is then optimal to subsidize parents. The magnitude of the intergenerational redistribution wedge depends on the Pareto-efficient allocation itself. In the formula in (20), this dependence is implicit as λ^*_{t+1} depends on the sequence of utility promises, $\mathbf{U}$, which indexes Pareto-efficient allocations. Of course, even though the laissez-faire equilibrium allocation is never efficient, it might be the case that bequests must be optimally taxed at a *zero rate*. This would happen if the intergenerational disagreement and intergenerational redistribution components of the tax formula exactly offset each other.

Ramsey allocation as Pareto Improvement: An example with Log utility. The Ramsey allocation and the taxes that achieve this allocation might as well improve the utility *to all agents* in the economy compared with the laissez-faire equilibrium. In this section, we focus on an example in which the instantaneous utility function is logarithmic and production in period t is given by $f(k_t) = Rk_t + w$, where $R > 0$ and $w \geq 0$ are given parameters of the production function. In Proposition 14 in Appendix A.4, we derive an analytic condition regarding when the Ramsey allocation Pareto improves over the laissez-faire equilibrium allocation that depends only on γ, δ, and β. Using this condition, and focusing for simplicity on the case in which people are perfectly altruistic ($\gamma = \delta$), the following proposition shows that for any level of disagreement between parents and offspring that satisfies $\beta \in (0, 1)$, the Ramsey allocation dominates the equilibrium allocation as the discount factor approaches 1.[24]

Proposition 8. *Suppose $u(c) = log(c)$ and $f(k) = Rk + w$, with $R > 0$ and $w \geq 0$. Suppose further that $\gamma = \delta$. Then, for any $\beta \in (0, 1)$, the Ramsey allocation Pareto improves over laissez-faire allocation as $\delta \to 1$.*

Proof. Relegated to the end of Appendix A.4. □

By definition, the Ramsey allocation makes the period 0 parent better off relative to the laissez-faire equilibrium. Regarding an agent who belongs to any future generation, moving

[24]We focus on the case with $\beta < 1$ in order to have well-defined lifetime utilities for all $\delta \leq 1$. A direct inspection of Proposition 14 in Appendix A.4 shows that for the case of $\gamma \neq \delta$ we obtain the same result when both γ and δ are sufficiently close to 1.

Table 2: Ramsey Allocation as Pareto Improvement

β	0.5	0.6	0.7	0.8	0.9	0.95
$\bar{\delta}$	0.758	0.745	0.734	0.724	0.718	0.711

from the laissez-faire equilibrium to the Ramsey allocation involves a trade-off. On the one hand, such a move forces the agent to save according to discount factor δ for his old age, and this hurts his welfare, since he saves according to $\beta\delta$ in equilibrium, which is his most preferred saving level. On the other hand, the Ramsey allocation forces his descendants to save according to δ as well. This creates a welfare improvement for the agent in consideration, since in equilibrium the descendants save according to $\beta\delta$ whereas he prefers them to save according to δ. As the discount factor δ goes to 1, the welfare improvement accumulates over all descendants without being discounted. In this case, the future benefit dominates the current cost and the Ramsey allocation increases the welfare of the agent in consideration. As a result, the Ramsey allocation Pareto improves over the laissez-faire equilibrium.

This intuition applies to less extreme cases as well. Using Proposition 14 in Appendix A.4, for each β, we can find a value of $\delta = \gamma$ above which the Ramsey allocation Pareto dominates the equilibrium allocation. Table 2 reports threshold δ values that correspond to several values of β. A comparison with the calibrated δ values in Table 1 shows that, under the parametric assumptions made about the utility and production functions in this section, the Ramsey allocation is not likely to Pareto improve over the laissez-faire equilibrium allocation. Notice that, even when the Ramsey allocation does not constitute a Pareto improvement, it might still be the case that there is a feasible allocation associated with a positive tax on bequests that improves welfare for all generations over the laissez-faire.

5.2 Horizontal Redistribution

An important simplifying assumption we have made so far in the paper is that all dynasties in the economy are identical. In this section, we relax this assumption and allow people to differ in their skill levels, which translates into income inequality in equilibrium.

A key issue that arises when the population is divided among different income groups is horizontal redistribution, that is, redistribution across different families. In this section, we incorporate Mirrlees (1971)'s model of horizontal redistribution into our model of bequest taxation and analyze how the normative predictions of the latter change under inequality and redistribution. First, we show that in two important cases, all the normative results derived in the model without heterogeneity continue to hold exactly. These are: (i) whenever redistribution can be performed via nondistortionary taxation (which corresponds to the case in which skills are publicly observable), and (ii) even if redistribution is limited because people's skills are private information, whenever we are interested in the Ramsey allocation of Section 4, that is, in the notation of Section 5.1, when $\lambda = 0$. Second, we show that, when the planner cares directly about the offspring's welfare (i.e., when $\lambda > 0$) and, in addition, skills are private information to the agents, horizontal redistribution introduces an element of progressivity into optimal bequest taxation. We show that under certain conditions (e.g., logarithmic utility), the optimal bequest tax is *progressive*. The progressivity of optimal bequest taxes can be interpretted to provide a justification for bequest tax exemptions observed in actual tax systems which constitute a specific form of progressivity.

To reduce the notational burden, we present our results using a three-period framework. In order to analyze efficient bequest taxation in the presence of imperfect horizontal redistribution, we follow the Mirrleesian tradition, assuming that people's labor incomes are publicly observable whereas their productivities and labor supplies are not.

A Mirrleesian Model Consider an economy that lasts three periods. At $t = 0$, skill level $s \geq 0$ is drawn from a set S according to a distribution H and parents choose labor $l_0 \in [0, 1]$ with flow utility

$$u(c_0) - v(l_0),$$

where the function v is increasing and convex, with $v'(0) = 0$ and $v'(1) = \infty$. To simplify the analysis, we assume that in periods 1 and 2, labor is supplied inelastically at $\bar{l}_t = 1$, as in the previous sections of the paper. Income in period $t = 0$ equals $y_0 = w_0 s l_0$, which is publicly observable, as is the real wage per unit of skill w_0. In periods $t = 1, 2$, all agents have a common labor income y_t, that is, the offspring do not inherit the skills of their parents. This is an extreme form of mean reversion in skills, which allows us to analyze optimal bequest taxation in an environment in which the bequest behavior is qualitatively in line with data (richer parents leaving higher bequests).

Throughout this section, we assume that the planner is Utilitarian and cares about all the dynasties equally. Given an allocation, define $z_t(s) := u(c_t(s))$ as the utility an agent from family s in period t receives from consuming $c_t(s)$. Since $u(\cdot)$ is one-to-one, a planner assigning a consumption allocation is the same as a planner assigning utility. We find it convenient to write down the planning problem where planner chooses utilities. Letting λ be the direct welfare weight the planner puts on the offspring, the planner's objective function is

$$\max_{\{z_t(\cdot)\}_{t=0}^{2}, y_0(\cdot)} \int_S \left\{ z_0(s) - v\left(\frac{y_0(s)}{w_0 s}\right) + (\gamma+\lambda)z_1(s) + \delta(\gamma+\lambda\beta)z_2(s) \right\} dH(s). \tag{21}$$

For each $t = 0, 1, 2$, define

$$C_t := \int_S c_t(s)dH(s) = \int_S g(z_t(s))dH(s), \quad L_0 := \int_S l_0(s)sdH(s), \quad L_1 = L_2 = 1,$$

where the function g represents the inverse of the utility function u (i.e., $g := u^{-1}$), and the variables C_t and L_t represent the aggregate amounts of consumption and labor in efficiency units in period t. Let K_t denote aggregate capital stock installed as of period t. As before, factors of production are priced competitively, and in particular, $w_t := F_L(K_t, L_t)$, where F_L refers to the partial derivative of F with respect to labor. Aggregate feasibility is then given by, for each $t = 0, 1, 2$,

$$C_t + K_{t+1} \leq F(K_t, L_t) + (1-\theta)K_t, \qquad \text{with} \quad K_0 \quad \text{given}. \tag{22}$$

We indicate by μ_t the multiplier associated with the feasibility constraint in period t.

We assume that the skill level, s, is privately observed by the parent. This implies that the planner also faces a set of incentive constraints in period 0, namely, for all $s, \hat{s} \in S$:

$$z_0(s) - v\left(\frac{y_0(s)}{sw_0}\right) + \gamma\left[z_1(s) + \delta z_2(s)\right] \geq z_0(\hat{s}) - v\left(\frac{y_0(\hat{s})}{sw_0}\right) + \gamma\left[z_1(\hat{s}) + \delta z_2(\hat{s})\right]. \tag{23}$$

Notice that since all information about skill types is revealed once and for all in period 0, there are only period 0 incentive constraints.

Given λ, the planner's problem then is to maximize (21) subject to (22) and (23). As explained in Section 5.1, by varying $\lambda \in [0, \infty)$, one can trace the Pareto frontier of the economy across generations.

For any $\lambda \geq 0$, we characterize the corresponding Pareto-efficient allocation by considering a perturbation of it as follows: for a given s, set $z_0^{\varepsilon}(s) = z_0^*(s) - \gamma\varepsilon$, $z_1^{\varepsilon}(s) = z_1^*(s) + \varepsilon$, and $z_2^\varepsilon(s) = z_2^*(s)$. Notice that since

$$z_0^\varepsilon(s) + \gamma\left[z_1^\varepsilon(s) + \delta z_2^\varepsilon(s)\right] = z_0^*(s) + \gamma\left[z_1^*(s) + \delta z_2^*(s)\right],$$

all such perturbations are by construction incentive compatible. This implies that for this set of perturbations, one can disregard the incentive constraints while solving the planner's problem. The efficient allocation has to solve the first-order optimality condition for the planner's problem over the choices of ε in a neighborhood of zero; that is, the following condition has to hold:

$$\lambda + \frac{\mu_0\gamma}{u'(c_0^*(s))} - \frac{\mu_0}{u'(c_1^*(s))f'(K_1^*)} = 0,$$

where $f'(K_1^*) := F_K(K_1^*, 1) + (1-\theta)$ is the return to capital in period 1. The previous condition can be rearranged as

$$\frac{\lambda}{\mu_0\gamma}u'(c_0^*(s)) = \frac{u'(c_0^*(s))}{\gamma f'(K_1^*)u'(c_1^*(s))} - 1. \tag{24}$$

Recall that the bequest wedge for a particular Pareto-efficient allocation is the wedge that the planner needs to create in order to make the parent choose that Pareto-efficient allocation in equilibrium. Following the definition of the bequest wedge introduced by (13), we define the bequest wedge for a family of type s under Pareto index λ to be[25]

$$BW^*(\lambda, s) \equiv 1 - \left(\frac{u'(c_0^*(s))}{\gamma f'(K_1^*)u'(c_1^*(s))} - \frac{\partial b_2(b_1^*(s), Q_1^*)}{\partial b_1}\frac{\left[-1+\frac{1}{\beta}\right]}{f'(K_1^*)}\right). \tag{25}$$

Plugging (24) into (25) we get

$$BW^*(\lambda, s) = -\frac{\lambda}{\mu_0\gamma}u'(c_0^*(s)) + \frac{\partial b_2(b_1^*(s), Q_1^*)}{\partial b_1}\frac{\left[-1+\frac{1}{\beta}\right]}{f'(K_1^*)}. \tag{26}$$

Similar to the bequest wedge characterized by (20) in Section 5.1, the bequest wedge characterized by (26) has two components. The first component arises from intergenerational redistribution and always implies a subsidy on bequests, whereas the second one is the efficiency component that is there to correct the externality arising from intergenerational

[25]Using the taxation principle, the wedge can also be written as a function of parent's income y_0.

disagreements, and it always implies a tax.

We first analyze the general case in which future generation is directly weighted in the planner's objective and skills are private information. The intergenerational redistribution component is increasing in s (a subsidy decreasing in parental income). Since the effort cost function v is increasing and convex, it is straightforward to show that incentive compatibility implies that $c_0^*(s)$ is increasing in s. This implies that the intergenerational redistribution component of the bequest wedge given in (26) calls for a subsidy decreasing with s. The mechanism behind the progressivity of the intergenerational redistribution component is the same as in Farhi and Werning (2010). Intuitively, as long as $\lambda > 0$, the planner's preference for c_1 relative to c_0 is stronger than that of the parents. This is easy to see by comparing the planner's objective function, (21), with incentive constraints, (23). This means that the planner will find it more effective to generate redistribution by reducing inequality in c_1 compared with reducing inequality in c_0. In turn, this implies that the optimal cross-sectional distribution of c_1 is more compressed than that of c_0. As a consequence, the efficient allocation displays increasing consumption (between periods $t = 0$ and $t = 1$) for agents with low s and decreasing consumption for agents with high s. Since all parents face the same interest rate, $R_1^* = f'(K_1^*)$, but must be content with different intertemporal patterns of consumption, a natural implementation of the efficient allocation would prescribe higher bequest taxes to rich parents (since otherwise they will have an incentive to deviate from the recommended consumption plan by increasing bequests) and lower bequest taxes to poor parents.

Even though the sign of the efficiency component of the bequest wedge in (26) is determinate, whether it increases or decreases with s is not. This is because a Pareto-efficient allocation does not specify a bequest distribution, $\{b_1^*(s)\}_{s \in S}$; it only specifies an aggregate capital stock, K_1^*. Since, in general, the derivative of the offspring policy may depend on the level of bequests received, the monotonicity properties of the efficiency wedge depend on the specifics of the implementation and are indeterminate. Under the CEIS utility function, however, offspring always save a constant fraction of the wealth transfers they receive from their parents. This implies that the derivative of the offspring saving policy is a constant that is independent of s, further implying that the efficiency component is flat. Thus, in this latter case, depending on which component dominates, the bequest wedge implies either a subsidy that is decreasing with s or a tax that is increasing with s.

In Proposition 15 in Appendix B, we provide a simple tax system that implements Pareto-efficient allocations in a market economy where people trade a risk-free bond. We show that, under the further restrictive assumption of logarithmic preferences, this tax system features progressive bequest taxation in the sense that the bequest tax is increasing in the amount of bequests.

There are two important cases in which the bequest wedge in (26) simplifies to the bequest wedge expressions derived earlier. First, *if skills are publicly observable*, then since $u'(c_0^*(s)) = \mu_0$ for all s (full redistribution), (26) reduces to

$$BW^*(\lambda, s) = -\frac{\lambda}{\gamma} + \frac{\partial b_2(K_1^*, Q_1^*)}{\partial b_1} \frac{\left[-1 + \frac{1}{\beta}\right]}{f'(K_1^*)},$$

where we assume the planner chooses $b_1^*(s) = K_1^*$ in the efficient allocation. This expression is identical to (20), which is the expression for the bequest wedge in Section 5.1 when there is no heterogeneity. Under the CEIS assumption, an equilibrium with linear strategies exists, and following the proof of Proposition 5, one can show that there is a corresponding *flat* optimal tax or subsidy.

Second, *if the planner only cares about the parent directly ($\lambda = 0$)*, then, even in the presence of income heterogeneity and imperfect redistribution, the bequest wedge given by (26) reduces to expression (14) in Section 4:

$$BW^*(0, s) = \frac{\partial b_2(b_1^*(s), Q_1^*)}{\partial b_1} \frac{\left[-1 + \frac{1}{\beta}\right]}{f'(K_1^*)} > 0.$$

Since future generations are not weighted directly, the bequest wedge consists only of the efficiency component and, hence, is unambiguously positive. Again under the CEIS assumption, Proposition 5 implies that there is a linear equilibrium in which the derivative of the offspring's saving policy is a constant, implying that there is a corresponding *flat* optimal bequest tax that implements this particular Pareto-efficient allocation.

5.3 Longer Life Cycle, Coexistence, and Inter Vivos Taxation

In our benchmark model, we assume that people live for two periods and parents and offspring do not coexist. 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. 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.

There are two natural ways to extend our baseline model to a multiperiod setup, and both deliver qualitatively the same normative predictions. First, one can envision agents as standard exponential discounters, perhaps with varying discount factors over the life cycle, so that the discount factor between their own consumption at age i and $i+1$ is $\beta_i \delta \leq \delta$. In this formulation, agents are time consistent. Second, we can assume that agents face self-control problems and have time-inconsistent preferences. In this model, agents disagree with their future selves over intertemporal trade-offs as well. That is, at any age i, people discount the future with $\beta_i \delta$ and want their future selves to discount with δ. To clarify that it is the intergenerational disagreement that implies the optimality of transfer taxation and not necessarily self-control problems, in the rest of the section, we focus on the time-consistent extension.[26]

[26] In an earlier working paper version, Pavoni and Yazici (2014), we focus on the time-inconsistent extension where, in addition to intergenerational disagreement, people disagree with their future selves as well. We find that our results survive this double degree of disagreement as well.

The offspring's preference in period i of his life can be expressed recursively as

$$\begin{aligned} V^c_{t+i-1} &= u(c_{t+i-1}) + \beta_i \delta V^c_{t+i} \quad \text{for} \quad i = 1, \ldots, I-1, \text{ and} \\ V^c_{t+I-1} &= u(c_{t+I-1}) + \beta_I \delta V_{t+I.} \end{aligned}$$

Observe that we allow for the disagreement parameter, β, to depend on i. When $\beta_i = 1$ for all i, there is no disagreement between generations about intertemporal trade-offs. Whenever $\beta_i < 1$ for some i, people at age i disagree with their parents. If we set $\beta_i = \beta$ for all i, that would mean that the degree of intergenerational disagreement is constant over age. We allow for a time-varying degree of disagreement in order to show that our transfer taxation results do not depend on how the degree of intergenerational disagreement evolves over the life cycle. We show that, *as long as the parent and the offspring disagree on an intertemporal trade-off either during the period the offspring receives transfers or in any subsequent period (or both), it is optimal to distort parental transfer behaviour.*

Let d_t and b^o_{t+1} 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

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

subject to the budget constraints[27]

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

and subject to the constraints defining the life-cycle saving policy functions of the offspring:

[27] 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 offsprings' savings in the coexistence period and parental bequests $k_{t+1} = b_{t+1} + b^o_{t+1}$.

that is, $[b_{t+1}(d_t, Q_t), b_{t+2}(d_t, b^o_{t+1}, Q_t)..., b_{t+I}(d_t, b^o_{t+1}, Q_t)]$ solves[28]

$$\max_{\{\hat{b}_{t+i}\}_{i=1}^{I}} u(c_t) + \sum_{i=1}^{I-1} \bar{\beta}_i \delta^i u(c_{t+i}) + \bar{\beta}_I \delta^I V(a_{t+I}, Q_{t+I}),$$

where $\bar{\beta}_i = \prod_{n=1}^{i} \beta_n$.

Notice that since each agent is time consistent, we can assume that the offspring choose their lifetime savings once at the beginning of their life cycle. The vector of functions $[b_{t+1}(d_t, Q_t), b_{t+2}(d_t, b^o_{t+1}, Q_t)..., b_{t+I}(d_t, b^o_{t+1}, Q_t)]$ describes how offsprings' savings choices over the life cycle depend on the parental transfers they receive, (d_t, b^o_{t+1}), and prices, Q_t. Observe that offsprings' saving at age 1 in year t, denoted by b_{t+1}, only depends on the inter vivos transfers they receive 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 of choosing d_t and b^o_{t+1} subject to the flow budget constraints and the offspring's 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 confusion. The parent's first-order optimality condition with respect to the inter vivos decision is

$$u'(c^o_t) = \gamma \left(u'(c_t) \left[1 - \frac{\partial b_{t+1}}{\partial d_t} \right] + \sum_{i=1}^{I-1} \delta^i u'(c_{t+i}) \frac{\partial c_{t+i}}{\partial d_t} + \delta^I V_1(a_{t+I}, Q_{t+I}) R_{t+I} \frac{\partial b_{t+I}}{\partial d_t} \right), \quad (27)$$

where

$$\frac{\partial c_{t+i}}{\partial d_t} = \left[R_{t+i} \frac{\partial b_{t+i}}{\partial d_t} - \frac{\partial b_{t+i+1}}{\partial d_t} \right], \quad (28)$$

and, clearly, the derivatives are evaluated at the equilibrium allocation and price sequence.

It follows from the time-consistent problem of the offspring that the following Euler

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

equations characterize his optimal saving decisions:

$$\begin{aligned} u'(c_{t+i}) &= \beta_{i+1}\delta R_{t+i+1}u'(c_{t+i+1}), \quad \text{for } i = 0, 1, ..., I-2, \text{ and} \\ u'(c_{t+I-1}) &= \beta_I \delta R_{t+I} V_1(a_{t+I}, Q_{t+I}), \end{aligned} \tag{29}$$

where, again, the derivatives are evaluated at the equilibrium allocation and price sequence. Using (29) in the parental optimality condition for the inter vivos decision, (27), 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 9. *Suppose the policy functions that describe offspring behavior over the life cycle are differentiable. Then, equilibrium inter vivos behavior in any parenthood period t is characterized by*

$$u'(c_t^o) = \gamma \left(u'(c_t) + \Delta_t \right), \tag{30}$$

where

$$\Delta_t := \sum_{i=0}^{I-1} \delta^i u'(c_{t+i}) \left[-1 + \frac{1}{\beta_{i+1}} \right] \frac{\partial b_{t+i+1}(d_t, b_{t+1}^o, Q_t)}{\partial d_t}. \tag{31}$$

Condition (30) is analogous to (10), 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 transfers. The second term on the right-hand side, Δ_t, summarizes the extra benefit from increasing transfers that exists because of the presence of intergenerational disagreement. It is analogous to the second term on the right-hand side of (10). The main difference is that, since the parent and the offspring (potentially) disagree on savings at different ages, it is summed over all ages of disagreement. Intuitively, the offspring is saving less than optimal from the parent's perspective at any age i where $\beta_i < 1$. As a result, parental welfare increases if the parent can make the offspring increase period i savings, which is possible by increasing inter vivos transfers as long as $\frac{\partial b_{t+i+1}}{\partial d_t} > 0$.

To see the implication of Proposition 9 for Ramsey inter vivos taxation, first observe that in the Ramsey allocation we have

$$u'\left(c_t^{o*}\right) = \gamma u'\left(c_t^*\right). \tag{32}$$

Comparing (30) with (32) leads to the *Ramsey inter vivos wedge* as

$$IW_t^* \equiv 1 - \left(\frac{u'(c_t^{o*})}{\gamma u'(c_t^*)} - \frac{\Delta_t^*}{u'(c_t^*)}\right),$$

where Δ_t^* is Δ_t evaluated at the Ramsey allocation. The Ramsey inter vivos wedge in period t measures the distortion that the planner needs to create in the return to the inter vivos transfer of a parent in period t in order to make him choose the Ramsey level of 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 10. *Suppose the policy functions that describe offspring behavior over the life cycle are differentiable and $\beta_i \leq 1$ for all $i \in \{1, 2, ..., I\}$ with at least one inequality holding as strict. Then, $IW_t^* > 0$.*

Proof. Under the assumption of differentiability of policy functions, $\frac{\partial b_{t+1}(d_t^*, b_{t+1}^{o*}, Q_t^*)}{\partial d_t} > 0$, which follows from a simple extension of Lemma 12 in Appendix A.1. From (31), we get $\Delta_t^* > 0$ under the assumption that $\beta_i \leq 1$ for all i and $\beta_i < 1$ for at least one i. Plugging (32) in the definition of IW_t^*, we find that $IW_t^* = \frac{\Delta_t^*}{u'(c_t^*)} > 0$. □

Corollary 10, 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 transfers are strictly positive.[29]

The argument for positive taxation of bequests in a multiperiod life-cycle environment is identical to the argument for inter vivos taxation and is therefore omitted for the sake of

[29]One might argue that since inter vivos transfers are received while the parent is still alive, the parent might have more leverage in affecting offspring saving behavior in that period. If the parent is able to ensure that the offspring saves exactly according to parental preference, then they would not have a reason to do excessive inter vivos transfers, and hence, there would be no need to tax inter vivos transfers. The "Rotten Kid Theorem" of Becker (1974) would be reestablished in that case.

brevity. One can compute the optimal bequest wedge by plugging the first-order optimality conditions of the offspring, (29), into the first-order optimality condition describing the bequest behavior. The optimal bequest wedge, given by

$$BW^*_{t+1} = \frac{\Delta^*_{t+1}}{\delta f'(k^*_{t+1}) u'(c^*_{t+1})},$$

where

$$\Delta^*_{t+1} = \sum_{i=1}^{I-1} \delta^i u'(c^*_{t+i}) \left[-1 + \frac{1}{\beta_{i+1}}\right] \frac{\partial b_{t+i+1}(d^*_t, b^{o*}_{t+1}, Q^*_t)}{\partial b^o_{t+1}},$$

is positive as long as $\beta_i \leq 1$ for all $i \in \{2, ..., I\}$ with at least one inequality holding as strict. For $I = 2$, the wedge reduces to (14).

5.4 Long-Term Assets and Liquidity Constraints

The front-loading of consumption by offspring relative to what their parents 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?

First, observe that if the offspring do not face liquidity constraints, 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.[30] 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. As we will see below, we find that, *as long as the parents do not have access to a portfolio of assets that allows them to target transfers to each and every period of the offsprings' life cycle, transfer taxation remains optimal.*

We use the multiperiod environment laid out in Section 5.3 for our analysis. To make things simple, suppose people are not allowed to borrow at all. Suppose that in addition to inter vivos transfers and bequests, parents can use an illiquid asset to transfer resources directly to period $t+2$. Let b^o_{t+2} denote the amount of this illiquid bequest where the subscript refers to the period in which the child receives the bequest. Suppose the return to this asset

[30] For a formal analysis of this claim, see Pavoni and Yazici (2015) Section 11.E.

is $R_{t+1}R_{t+2}$. The budget constraints then are

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

Let $[b_{t+1}(d_t, Q_t), b_{t+2}(d_t, b_{t+1}^o, b_{t+2}^o, Q_t)..., b_{t+I}(d_t, b_{t+1}^o, b_{t+2}^o, Q_t)]$ denote the functions that describe how the offsprings' savings choices over the life cycle depend on parental transfers they receive.[31]

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 show that parents use the transfers to keep offspring borrowing constrained in periods t and $t+1$. The intuition is simple. Suppose the offspring is not constrained in period t in equilibrium. This means his optimal saving level b_{t+1} is given by (29). Now, if the parent increases bequests and decreases inter vivos transfers in a way that keeps the total amount of transfers unchanged, the offspring will have to decrease his period t savings to keep his preferred allocation. If the parent continues to backload transfers, there will be a point at which the offspring will completely deplete his savings and will become borrowing constrained. From this point onward, back-loading of transfers strictly increases parental welfare, and the parent does this until the equilibrium saving behavior of the offspring is exactly in line with that of the parent:[32]

$$u'(c_t) = \delta\Big\{\sum_{i=1}^{I-1}\delta^{i-1}u'(c_{t+i})\frac{\partial c_{t+i}}{\partial b_{t+1}} + \delta^{I-1}V_1(a_{t+I}, Q_{t+I})R_{t+I}\frac{\partial b_{t+I}}{\partial b_{t+1}}\Big\}, \tag{33}$$

where the term $\frac{\partial c_{t+i}}{\partial b_{t+1}}$ is defined analogous to (28). Similarly, the parent uses liquid and

[31]Notice 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 arises from our assumption that children learn the levels of both liquid and illiquid bequests in period t, after they make their period t savings decision.

[32]Notice that when $w_t > 0$ and/or the offspring can borrow a positive amount, using the timing of transfers to make the offspring save according to (33) might require leaving negative bequests.

illiquid bequests to align offspring saving behavior with his own in period $t+1$:

$$u'(c_{t+1}) = \delta \Big\{ \sum_{i=2}^{I-1} \delta^{i-2} u'(c_{t+i}) \frac{\partial c_{t+i}}{\partial b_{t+2}} + \delta^{I-2} V_1(a_{t+I}, Q_{t+I}) R_{t+I} \frac{\partial b_{t+I}}{\partial b_{t+2}} \Big\}. \tag{34}$$

The fact that the parent can control the offspring's savings behavior in a period using the timing of transfers implies that effectively there is no disagreement regarding that period's savings. As a result, the parent does not have a motive to transfer too much, and hence, there is no need to distort transfers. For inter vivos transfers, this can be seen by plugging (33) into the parental optimality condition for inter vivos transfers, equation (27). For bequests, it can be shown by plugging (34) 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 transfers 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 saves according to his own preference in periods $t+2$ onward, meaning that for each $i = 2, ..., I-1$, the offspring's saving satisfies (29). Plugging (33), (34), and (29) into the parent's optimality condition regarding illiquid bequests, we get[33]

$$u'(c_t^o) = \gamma \left(R_{t+1} R_{t+2} \delta^2 u'(c_{t+2}) + \Delta_{t+2} \right), \tag{35}$$

where

$$\Delta_{t+2} = \sum_{i=2}^{I-1} \delta^i u'(c_{t+i}) \left[-1 + \frac{1}{\beta_{i+1}} \right] \frac{\partial b_{t+i+1}(d_t, b_{t+1}^o, Q_t)}{\partial b_{t+2}^o}.$$

As long as $\beta_i \leq 1$ for all $i \in \{3, ..., I\}$ and $\beta_i < 1$ for some i, Δ_{t+2} is strictly positive, which means that parents do too much illiquid bequeathing and therefore should be taxed.

Intuitively, since the parent cannot directly transfer resources to period $t+3$ (the fourth period of the offspring's life), 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, $\{b_{t+i}^o\}_{i=1}^I$, 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.

[33] For an exact derivation of (35), see Appendix C.

6 Conclusion

We study the optimal taxation of parental transfers in a model where altruistic parents and their offspring disagree on intertemporal trade-offs. We prove that laissez-faire equilibrium is inefficient in a Pareto sense because of intergenerational disagreement. We focus on the Ramsey allocation, the allocation that maximizes the welfare of the initial parent, as our benchmark Pareto efficient allocation. We show that, if offspring are impatient from parents' perspective, then parents bequeath too much relative to Ramsey level, and hence, it is optimal to tax bequests. We then consider policies that target any point on the Pareto frontier of the economy. We find that intergenerational disagreement again calls for a taxation of bequests. In this case, there is another component of bequest taxation that comes from intergenerational redistribution, and this component always asks for a subsidy on bequests. The sign of the tax depends on which component dominates. We also consider an economy with income inequality across families in which the government wants to redistribute from the rich to the poor. We find that if labor income taxes are distortionary and society puts direct weight on future agents, then bequest taxes are progressive.

Finally, we show that the optimality of positive taxes on transfers remains valid even when (i) we consider life cycles with arbitrarily long finite horizons and (ii) parents can trade long maturity assets. In Appendix D we also show that our result survives even when parents are naive in the sense that they do not realize the existence of intergenerational disagreements.

References

BECKER, G. S. (1974): "A Theory of Social Interactions," *Journal of Political Economy*, 82(6), 1063–93.

BERNHEIM, B. D. (1989): "Intergenerational Altruism, Dynastic Equilibria and Social Welfare," *The Review of Economic Studies*, 56(1), pp. 119–128.

CLARKE, E. J., M. PRESTON, J. RAKSIN, AND V. L. BENGTSON (1999): "Types of Conflicts and Tensions Between Older Parents and Adult Children," *The Gerontologist*, 39(3), pp. 261–270.

CREMER, H., AND P. PESTIEAU (2011): "Wealth and Wealth Transfer Taxation: A Survey," in *The Elgar Guide to Tax Systems*, ed. by E. Albi, and J. Martinez-Vasquez, pp. 183–217. Northampton, MA: Edward Elgar.

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

DOEPKE, M., AND M. TERTILT (2009): "Women's Liberation: What's in It for Men?," *The Quarterly Journal of Economics*, 124(4), 1541–1591.

DOEPKE, M., AND F. ZILIBOTTI (2014): "Parenting with style: altruism and paternalism in intergenerational preference transmission," UBSCENTER - Working Papers 008, UBS International Center of Economics in Society - Department of Economics - University of Zurich.

FARHI, E., AND I. WERNING (2010): "Progressive Estate Taxation," *The Quarterly Journal of Economics*, 125(2), 635–673.

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 (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.

KOCHERLAKOTA, N. R. (2005): "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation," *Econometrica*, 73(5), 1587–1621.

KOPCZUK, W. (2009): "Economics of Estate Taxation: A Brief Review of Theory and Evidence," *Tax Law Review*, 63(1), 139–157.

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.

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.

MORRIS, S., AND A. POSTLEWAITE (1997): "Observational Implications of Non-Exponential Discounting," Discussion paper.

O'DONOGHUE, T., AND M. RABIN (1999): "Doing It Now or Later," *American Economic Review*, 89(1), 103–124.

——— (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 (2014): "Present-Bias Preferences and Optimal Taxation of Parental Transfers," Working paper.

——— (2015): "Optimal Life-cycle Capital Taxation under Self-Control Problems," *The Economic Journal, forthcoming*.

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*, 81(5), 1851–1886.

STOKEY, N. L., R. E. LUCAS, JR., AND E. C. PRESCOTT (1989): *Recursive Methods in Economic Dynamics*. Harvard University Press, Cambridge, MA.

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 that we use in the proof later on.

Lemma 11. *The policy of the offspring $b_{t+2}(\cdot, Q_{t+1})$ is increasing in the amount of bequests received, and the value function $V(\cdot, Q_t)$ is strictly increasing in wealth.*

Proof. It is easy to see that $V(\cdot, Q_t)$ 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 $b_{t+2}(\cdot, Q_{t+1})$ is shown as follows. The definition of the offspring's saving policy at b_{t+1} implies that, for $\varepsilon > 0$, we have

$$\begin{aligned}
& u(R_{t+1}b_{t+1} + w_{t+1} - b_{t+2}(b_{t+1}, Q_{t+1})) - u(R_{t+1}b_{t+1} + w_{t+1} - b_{t+2}(b_{t+1} + \varepsilon, Q_{t+1})) \\
\geq \; & \beta\delta[V(R_{t+2}b_{t+2}(b_{t+1} + \varepsilon, Q_{t+1}) + w_{t+2}, Q_{t+2}) - V(R_{t+2}b_{t+2}(b_{t+1}, Q_{t+1}) + w_{t+2}, Q_{t+2})].
\end{aligned} \tag{36}$$

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

$$\begin{aligned}
u(R_{t+1}(b_{t+1} + \epsilon) + w_{t+1} - b_{t+2}(b_{t+1}, Q_{t+1})) - u(R_{t+1}(b_{t+1} + \varepsilon) + w_{t+1} - b_{t+2}(b_{t+1} + \varepsilon, Q_{t+1})) \\
\leq \; \beta\delta[R_{t+2}V(b_{t+2}(b_{t+1} + \varepsilon, Q_{t+1}) + w_{t+2}, Q_{t+2}) - V(R_{t+2}b_{t+2}(b_{t+1}, Q_{t+1}) + w_{t+2}, Q_{t+2})].
\end{aligned} \tag{37}$$

Combining (36) and (37), we get

$$\begin{aligned}
& u(R_{t+1}b_{t+1} + w_{t+1} - b_{t+2}(b_{t+1}, Q_{t+1})) - u(R_{t+1}b_{t+1} + w_{t+1} - b_{t+2}(b_{t+1} + \varepsilon, Q_{t+1})) \\
\geq \; & u(R_{t+1}(b_{t+1} + \varepsilon) + w_{t+1} - b_{t+2}(b_{t+1}, Q_{t+1})) - u(R_{t+1}(b_{t+1} + \varepsilon) + w_{t+1} - b_{t+2}(b_{t+1} + \varepsilon, Q_{t+1})).
\end{aligned} \tag{38}$$

Assume for the sake of contradiction that $b_{t+2}(b_{t+1} + \varepsilon, Q_{t+1}) < b_{t+2}(b_{t+1}, Q_{t+1})$. Combined with strict concavity of u, this contradicts with (38). Thus, it must be that $b_{t+2}(b_{t+1} + \varepsilon, Q_{t+1}) \geq b_{t+2}(b_{t+1}, Q_{t+1})$. □

Lemma 12. *Suppose the value function $V(\cdot, Q_t)$ is differentiable. Then, $b_{t+2}(\cdot, Q_{t+1})$ is strictly monotone in the amount of bequests received.*

Proof. Because of the Inada assumption on the utility function, the solution to the offspring's problem stated in (2) must be interior. A necessary condition for the optimality of the offspring's savings is then

$$u'(R_{t+1}b_{t+1} + w_{t+1} - b_{t+2}) = \beta\delta R_{t+1}V_1(R_{t+2}b_{t+2} + w_{t+2}, Q_{t+2})R_{t+2}. \tag{39}$$

Now suppose the offspring receives a higher level of bequests from the parent, $b_{t+1} + x$, where $x > 0$. It would never be optimal for the offspring to use all of the increase in his wealth for current consumption, since this would not satisfy his necessary condition for optimality:

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

which follows from (39) and the strict concavity of the utility function. □

Proof. **(Core Proof of Proposition 1)** In the proof of this proposition, we use monotonicity of $b_{t+2}(\cdot, Q_{t+1})$ and strict monotonicity of $V(\cdot, Q_t)$, which we establish in Lemma 11.

Assume for the sake of finding a contradiction that in equilibrium

$$-u'(R_t b_t + w_t - b_{t+1}) + \gamma R_{t+1} u'(R_{t+1} b_{t+1} + w_{t+1} - b_{t+2}(b_{t+1}, Q_{t+1})) > 0. \tag{40}$$

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

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

Since u is a differentiable function, for ε sufficiently small, under the assumption (40), we have

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

Then, we have

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

where the first inequality follows from the definition of the policy $b_{t+2}(\cdot, Q_{t+1})$ and the monotonicity of both $b_{t+2}(\cdot, Q_{t+1})$ and $V(\cdot, Q_t)$, the second inequality follows from (40) and the differentiability of the utility function u as discussed earlier, and the last one from the monotonicity of both $b_{t+2}(\cdot, Q_{t+1})$ and $V(\cdot, Q_t)$ and the fact that $\beta < 1$. It is now easy to see that the last row will have a strict inequality whenever the policy $b_{t+2}(\cdot, Q_{t+1})$ is strictly monotone, since V is actually strictly increasing. In this case, following the same line of proof, we can show a contradiction to the weak inequality version of (40).

The case of $\beta > 1$ can be shown in an identical way. The case of $\beta = 1$ is trivial. □

A.2 Proof of Proposition 2

We start the proof with a preliminary lemma.

Lemma 13. *Let $\{c_t, k_t\}_{t=0}^{\infty}$ be an equilibrium allocation. Let c_{s+1} and c_{s+2} be the consumption of an agent in his two years of life (with c_{s+1} consumption when young and c_{s+2} when a parent), and let b_{s+2} be the agent's* *level of saving from period $s+1$ to $s+2$, which is equal to the equilibrium the level of capital generated in that period, k_{s+2}. Then, the following condition has to hold in equilibrium:*

$$-u'(c_{s+1}) + \beta\delta f'(k_{s+2})u'(c_{s+2}) = 0.$$

Proof. Recall that in the equilibrium allocation, offspring's decisions are interior due to the Inada condition. Notice also that the offspring is choosing both b_{s+2} and b_{s+3}. In particular, the offspring can always consider the usual Euler perturbation. The optimal choice of b_{s+2} must be such that there is no feasible deviation around b_{t+2} (keeping both b_{s+1} and b_{s+3} unchanged) that improves over the equilibrium level of b_{t+2}. The offspring's optimal decision must be such that the problem

$$\max_{\hat{x}\in O} u\left(R_{s+1}b_{s+1} + w_{s+1} - b_{s+2} - \hat{x}\right) + \delta\beta u\left(R_{s+2}(b_{s+2}+\hat{x}) + w_{s+2} - b_{s+3}\right),$$

where O is an open interval around zero, has the solution $\hat{x}^* = 0$. Since $u(\cdot)$ is differentiable, and recalling that $R_{s+2} = f'(k_{s+2})$ and $b_{s+2} = k_{s+2}$, the necessary condition for $\hat{x}^* = 0$ to be optimal is the first-order condition displayed in the statement of the lemma. □

We will prove the statement for the case $\beta < 1$. The proof of the case of $\beta > 1$ is symmetric and will be sketched next. For notational simplicity, we will make the perturbation for consumption in periods 0,1, and 2. Define a perturbation by $\hat{c}_0 = c_0 - \varepsilon$, $\hat{c}_1 = c_1 + f(k_1+\varepsilon) - f(k_1) - \eta$, and $\hat{c}_2 = c_2 + f(k_2+\eta) - f(k_2)$, where $\varepsilon, \eta > 0$.

Proof. **(Core Proof of Proposition 2)** We will show that there exist ε and η such that this perturbation makes both the initial parent and the agent who is born in period 1 strictly better off. The welfare of future generations will not be altered, since the perturbation only involves periods 0,1, and 2. First, we show that the offspring's welfare increases. We aim at showing that $u(c_1 + f(k_1+\varepsilon) - f(k_1) - \eta) + \delta\beta u(c_2 + f(k_2+\eta) - f(k_2)) - [u(c_1) + \delta\beta u(c_2)] > 0$. Letting $\varepsilon = \xi\eta$ with $\xi > 0$ to be chosen, the same condition is equivalent to itself divided by the positive number η. So, for $\eta > 0$ we have

$$\begin{aligned} & u\left(c_1 + f(k_1+\xi\eta) - f(k_1) - \eta\right) + \delta\beta u\left(c_2 + f(k_2+\eta) - f(k_2)\right) - [u(c_1) + \delta\beta u(c_2)] > 0 \\ \iff \quad & \frac{u\left(c_1 + f(k_1+\xi\eta) - f(k_1) - \eta\right) + \delta\beta u\left(c_2 + f(k_2+\eta) - f(k_2)\right) - [u(c_1) + \delta\beta u(c_2)]}{\eta} > 0. \end{aligned}$$

Since $u(\cdot)$ and $f(\cdot)$ are differentiable, we have

$$\begin{aligned} & u\left(c_1 + f(k_1+\xi\eta) - f(k_1) - \eta\right) + \delta\beta u\left(c_2 + f(k_2+\eta) - f(k_2)\right) - [u(c_1) + \delta\beta u(c_2)] \\ = \quad & \left[(\xi f'(k_1) - 1)u'(c_1) + \delta\beta f'(k_2)u'(c_2)\right]\eta + \mathcal{O}(\eta)\eta, \end{aligned}$$

where, by definition of the remainder $\mathcal{O}(\cdot)$ (as the right derivative exists),

$$\lim_{\eta\to 0^+, \eta>0} \quad \mathcal{O}(\eta) = 0.$$

Hence, for $\eta^* > 0$ small enough we have

$$\frac{\left[(\xi f'(k_1) - 1)u'(c_1) + \delta\beta f'(k_2)u'(c_2)\right]\eta^*}{\eta^*} + \mathcal{O}(\eta^*) = \xi f'(k_1)u'(c_1) + \mathcal{O}(\eta^*) > 0,$$

where we use Lemma 13.[34]

Next, we compute the change in welfare for the parent. Recall, $\varepsilon = \xi\eta$. Set

$$l := \max\left\{u'(c_0) - \gamma\delta\beta f'(k_1)f'(k_2)u'(c_2),\ (1-\beta)f'(k_2)\gamma\delta u'(c_2)\right\} > 0,$$

$$\xi := \frac{(1-\beta)\gamma\delta f'(k_2)u'(c_2)}{1+l} \in (0,1).$$

Then, since $u(\cdot)$ and $f(\cdot)$ are differentiable, following the same argument as earlier, we can find a small enough η such that

$$\frac{u(c_0 - \xi\eta) + \gamma u\left(c_1 + f(k_1 + \xi\eta) - f(k_1) - \eta\right) + \gamma\delta u\left(c_2 + f(k_2 + \eta) - f(k_2)\right) - [u(c_0) + \gamma u(c_1) + \gamma\delta u(c_2)]}{\eta} > 0$$

if and only if

$$-\xi u'(c_0) + \gamma(\xi f'(k_1) - 1)u'(c_1) + \gamma\delta f'(k_2)u'(c_2) > 0.$$

But the latter inequality must be true since we have

$$\begin{aligned}
&-\xi u'(c_0) + \gamma(\xi f'(k_1) - 1)u'(c_1) + \gamma\delta f'(k_2)u'(c_2) \\
&= -\xi\left[u'(c_0) - \gamma\beta\delta f'(k_1)f'(k_2)u'(c_2)\right] + (1-\beta)\gamma\delta f'(k_2)u'(c_2) > 0.
\end{aligned}$$

The equality between the first and second rows uses Lemma 13 to replace $u'(c_1)$ for $\beta\delta f'(k_2)u'(c_2)$ and rearranging terms, whereas the last inequality comes from our definition of ξ and l, which directly implies $\xi\left[u'(c_0) - \gamma\beta\delta f'(k_1)f'(k_2)u'(c_2)\right] < (1-\beta)\gamma\delta f'(k_2)u'(c_2)$.

For the case $\beta > 1$, define a perturbation by $\hat{c}_0 = c_0 - \varepsilon$, $\hat{c}_1 = c_1 + f(k_1 + \varepsilon) - f(k_1) + \eta$, and $\hat{c}_2 = c_2 + f(k_2 - \eta) - f(k_2)$. Following the same steps as earlier, for small $\eta > 0$, the change in offspring welfare again reduces to (disregarding the remainder $\mathcal{O}(\eta)$ since it goes to 0 as η goes to 0) $u'(c_1)\xi f'(k_1)$, which is strictly positive. The change in parental welfare reduces to

$$-\xi\left[u'(c_0) - \gamma\beta\delta f'(k_1)f'(k_2)u'(c_2)\right] + (\beta-1)\gamma\delta f'(k_2)u'(c_2) > 0,$$

which is strictly positive, since both of the terms in this expression are positive and at least one of them is strictly positive. □

A.3 CEIS utility functions

Proof of Proposition 6

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

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

[34]To see in more detail why we can guarantee the existence of such η^*, recall that the definition of $\mathcal{O}(\cdot)$ implies that for all $\zeta > 0$ we can find a $\psi > 0$ such that for all $\eta < \psi$ we have $|\mathcal{O}(\eta^*)| < \zeta$. Thus, since $\xi f'(k_1)u'(c_1) > 0$, we can find $0 < \zeta^* < \xi f'(k_1)u'(c_1)$ such that for all $0 < \eta^* < \psi^*$ the requirement is satisfied.

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 $s+1$ onward:

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

We will 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 $1-\rho \leq 1$ in period s net present value of wealth, that is,

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

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 $1-\rho$ 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 the budget constraint is homogeneous of degree 1 in consumption and the 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 the period s net present value of wealth and the period s consumption choice is $\hat{c}_s$. Now we show that when the period s wealth is $\lambda\Gamma_s(b_s)$ and the agent consumes $\lambda\hat{c}_s$ in period s, then the 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

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

Plugging this value into the definition of the 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.

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

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

if and only if

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

s.t. (3) $c_s + b_{s+1} = \lambda\Gamma_s(b_s) - G_s$ and (4) c_{s+1} 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 the existence of a solution for both problems. This assumption is not needed, as we have shown. 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 the 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 $1-\rho$ in Γ, as assumed:

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

To complete the proof, observe that we have shown that consumption defined as a function of the net present value of wealth, prices, and taxes, denoted 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:

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

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

A.4 The Logarithmic Case

Proof of Proposition 7

Proof. The proof proceeds in two steps. In the first step, for logarithmic utility, we compute a closed form solution for the equilibrium linear consumption policy as a function of the net present value of wealth and tax-price sequence. In the second step, we use the consumption policy computed in step 1 to compute M^*_{t+2} and plug that into the tax formula in Proposition 5.

Step 1. We use the guess and verify method to compute the 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 this 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 the 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

$$\hat{V}(\Gamma_t(b_t), \Psi_t) = \max_{C_t} \; u(C_t\Gamma_t(b_t)) + \gamma[u(C_{t+1}\Gamma_{t+1}(b_{t+1})) + \delta\hat{V}(\Gamma_{t+2}(b_{t+2}), \Psi_{t+2})] \tag{41}$$

$$s.t.$$

$$u'(C_{t+1}\Gamma_{t+1}(b_{t+1})) = \delta\beta\hat{V}_1(\Gamma_{t+2}(b_{t+2}), \Psi_{t+2}) R_{t+2}(1-\tau_{t+2}). \tag{42}$$

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

$$\begin{aligned}
\Gamma_{t+1}(b_{t+1}) &= R_{t+1}(1-\tau_{t+1})b_{t+1} + w_{t+1} + T_{t+1} + G_{t+1} \\
&= R_{t+1}(1-\tau_{t+1})\left[R_t(1-\tau_t)b_t + w_t + T_t - C_t\Gamma_t(b_t)\right] + w_{t+1} + T_{t+1} + G_{t+1} \\
&= R_{t+1}(1-\tau_{t+1})\left[R_t(1-\tau_t)b_t + w_t + T_t - C_t\Gamma_t(b_t) + \frac{w_{t+1} + T_{t+1} + G_{t+1}}{R_{t+1}(1-\tau_{t+1})}\right] \\
&= R_{t+1}(1-\tau_{t+1})\left[R_t(1-\tau_t)b_t + T_t - C_t\Gamma_t(b_t) + G_t\right] \\
&= R_{t+1}(1-\tau_{t+1})\Gamma_t(b_t)\left[1 - C_t\right].
\end{aligned} \tag{43}$$

Plugging the value function guess into the constraint of the planning problem, (42), and using $u(\cdot) = \log$, we get

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

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

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

or

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

Taking the first-order condition with respect to C_t in the parent's problem and again using (43), we get

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

Now we verify the value function to compute D :

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

which, using (43) and comparing the coefficients of $\log(\Gamma_t\,(b_t))$ on both sides of the previous equation, implies

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

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

$$\begin{aligned} M^*_{t+2} \equiv \frac{db_{t+2}(b_{t+1},\Psi_{t+1})}{db_{t+1}} &= R_{t+1}(1-\tau_{t+1}) - C_{t+1}R_{t+1}(1-\tau_{t+1}) \\ &= R_{t+1}(1-\tau_{t+1})\frac{\delta\beta D}{1+\delta\beta D}. \end{aligned} \tag{44}$$

Plugging this M^*_{t+2} value into the tax formula in Proposition 5 and doing some simple manipulation, we get the result.

□

Proof of Proposition 8 We first prove Proposition 14, which is of independent interest and is used for our computations in Table 2.

Proposition 14. *Suppose $u(c) = log(c)$ and for $t \geq 0$, $f_t(k) = R_t k + w_t$, with $R_t > 0$. The Ramsey allocation Pareto improves over the laissez-faire market equilibrium allocation if and only if*

$$\log\left(\frac{(1-\gamma\delta)(1+\gamma-\delta+\beta\delta)}{(1+\gamma)(1-\gamma\delta\beta-\delta+\beta\delta)}\right) + \log A + \log(1-\delta+\beta\delta)B - log(\beta)C \geq 0, \tag{45}$$

where

$$A = 1 + \frac{\beta\delta}{1-\gamma\delta}(1+\gamma),$$
$$B = 1 + \frac{\beta\delta}{1-\gamma\delta}\left(\frac{1+\gamma}{1-\gamma\delta} + \gamma\right),$$
$$C = \frac{\beta\delta}{1-\gamma\delta}\left(\frac{1+\gamma}{1-\gamma\delta}\right).$$

Proof. Step 1. We first compute the Ramsey allocation. The Euler equations that characterize the

Ramsey equilibrium are: for $t = 0, 2, 4, ...$

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

Since utility is logarithmic, this implies that for $t = 0, 2, 4, ...$

$$c_{t+1}^* = c_t^* x_{t+1}, \qquad c_{t+2}^* = c_{t+1}^* y_{t+2},$$

where

$$x_{t+1} = \gamma R_{t+1}, \qquad y_{t+2} = \delta R_{t+2}.$$

Equivalently, we can write

$$\begin{aligned} c_1^* &= c_0^* x_1, \\ c_2^* &= c_0^* x_1 y_2, \\ c_3^* &= c_0^* x_1 y_2 x_3, \\ c_4^* &= c_0^* x_1 y_2 x_3 y_4, \\ &\ldots \end{aligned}$$

Using this in the budget constraint, we get

$$c_0^*(1 + \frac{x_1}{R_1} + \frac{x_1 y_2}{R_1 R_2} + \frac{x_1 y_2 x_3}{R_1 R_2 R_3} + \frac{x_1 y_2 x_3 y_4}{R_1 R_2 R_3 R_4} + ...) = R_0 b_0 + w_0 + \frac{w_1}{R_1} + \frac{w_2}{R_1 R_2} + ... \equiv \Gamma,$$

which after plugging in values of x_{t+1} and y_{t+2} gives

$$c_0^*(1 + \gamma + \gamma\delta + \gamma^2\delta + \gamma^2\delta^2 + ...) = \Gamma,$$

implying

$$c_0^* = \Gamma \frac{1 - \gamma\delta}{1 + \gamma}.$$

Step 2. Next, we compute the laissez-faire equilibrium allocation, which we denote below by a hat. The Euler equations that characterize equilibrium are: for $t = 0, 2, 4, ...$

$$\begin{aligned} u'(\hat{c}_t) &= \frac{\gamma R}{1 - \delta + \beta\delta} u'(\hat{c}_{t+1}), \\ u'(\hat{c}_{t+1}) &= \beta\delta R u'(\hat{c}_{t+2}). \end{aligned}$$

Since utility is logarithmic, this implies that for $t = 0, 2, 4, ...$

$$\begin{aligned} \hat{c}_{t+1} &= \hat{c}_t \hat{x}_{t+1}, \\ \hat{c}_{t+2} &= \hat{c}_{t+1} \hat{y}_{t+1}, \end{aligned}$$

where

$$\hat{x}_{t+1} = \frac{\gamma R}{1-\delta+\beta\delta},$$
$$\hat{y}_{t+2} = \beta\delta R.$$

Following what we did while computing the Ramsey allocation, we find that

$$\hat{c}_0 = \Gamma\frac{1-\delta+\beta\delta-\beta\gamma\delta}{1-\delta+\beta\delta+\gamma}.$$

Step 3. Now we compare the welfare of all generations under the two allocations and find the restrictions on parameters under which Ramsey improves welfare for all.

Notice that the Ramsey allocation strictly increases the initial old's welfare over laissez-faire since by definition the Ramsey allocation is the unique allocation that maximizes the initial old's welfare. We need to find when it improves over laissez-faire for the rest of the agents.

Let us first check the agent who is born in period 1. Denote his utility under the Ramsey allocation by U^1_{RA} and that under the laissez-faire equilibrium by U^1_{LF} :

$$\begin{aligned} U^1_{RA} &= \log(c_0^* x_1) + \beta\delta[\log(c_0^* x_1 y_2) + \gamma\log(c_0^* x_1 y_2 x_3) + \gamma\delta\log(c_0^* x_1 y_2 x_3 y_4) + ...] \\ U^1_{LF} &= \log(\hat{c}_0\hat{x}_1) + \beta\delta[\log(\hat{c}_0\hat{x}_1\hat{y}_2) + \gamma\log(\hat{c}_0\hat{x}_1\hat{y}_2\hat{x}_3) + \gamma\delta\log(\hat{c}_0\hat{x}_1\hat{y}_2\hat{x}_3\hat{y}_4) + ...] \end{aligned}$$

Agent 1 is better off under the Ramsey allocation if and only if $U^1_{RA} \geq U^1_{LF}$. But this difference can be computed as follows:

$$U^1_{RA} - U^1_{LF} = [\log(c_0^*) - \log(\hat{c}_0)]A + [\log(\gamma) - \log(\frac{\gamma}{1-\delta+\beta\delta})]B + [\log(\delta) - \log(\beta\delta)]C, \tag{46}$$

where

$$A = 1 + \beta\delta[1+\gamma+\gamma\delta+\gamma^2\delta+\gamma^2\delta^2+...],$$
$$B = 1 + \beta\delta[1+2\gamma+2\gamma\delta+3\gamma^2\delta+3\gamma^2\delta^2+...],$$
$$C = \beta\delta[1+\gamma+2\gamma\delta+2\gamma^2\delta+3\gamma^2\delta^2+...].$$

With some algebra, we can show that

$$A = 1 + \frac{\beta\delta}{1-\gamma\delta}(1+\gamma),$$
$$B = 1 + \frac{\beta\delta}{1-\gamma\delta}(\frac{1+\gamma}{1-\gamma\delta}+\gamma),$$
$$C = \frac{\beta\delta}{1-\gamma\delta}(\frac{1+\gamma}{1-\gamma\delta}).$$

Plugging in the values of c_0^* and $\hat{c}_0$, condition (46) becomes

$$U^1_{RA} - U^1_{LF} = A\log D + B\log(1-\delta+\beta\delta) - C\log\beta, \tag{47}$$

where $D = \frac{(1-\gamma\delta)(1+\gamma-\delta+\beta\delta)}{(1+\gamma)(1-\gamma\delta\beta-\delta+\beta\delta)}$. Since neither the Ramsey nor the laissez-faire allocations are assumed to be stationary, the fact that agent 1 is better off under the Ramsey allocation does not imply that future generations are better off too. Now, we show that condition (47) guarantees that agents born in future generations are better off under the Ramsey allocation as well. Let us first look at the agent born in period 3.

$$\begin{aligned} U_{RA}^3 &= \log(c_0^* x_1 y_2 x_3) + \beta\delta[\log(c_0^* x_1 y_2 x_3 y_4) + \gamma \log(c_0^* x_1 y_2 x_3 y_4 x_5) + ...] \\ &= U_{RA}^1 + \log(y_2 x_3) + \beta\delta[\log(x_3 y_4) + \gamma \log(x_4 y_5) + ...]. \end{aligned}$$

Similarly,

$$U_{LF}^3 = U_{LF}^1 + \log(\hat{y}_2 \hat{x}_3) + \beta\delta[\log(\hat{x}_3 \hat{y}_4) + \gamma \log(\hat{x}_4 \hat{y}_5) + ...]$$

After some algebra, we reach

$$\begin{aligned} U_{RA}^3 - U_{LF}^3 = U_{RA}^1 - U_{LF}^1 + [log(\gamma\delta) - log(\frac{\gamma\delta\beta}{1-\delta+\beta\delta})]A \\ = U_{RA}^1 - U_{LF}^1 - log(\frac{\beta}{1-\delta+\beta\delta})A \\ > U^1(x,y) - U^1(\hat{x},\hat{y}), \end{aligned}$$

since $\frac{\beta}{1-\delta+\beta\delta} < 1$. Thus, we have shown that condition (47) guarantees that agent 3 is better off under the Ramsey allocation.

In a similar way, one can show that for any t odd, $U^t(x,y) - U^t(\hat{x},\hat{y}) > U^1(x,y) - U^1(\hat{x},\hat{y})$, ending the proof.

□

Core Proof of Proposition 8

Proof. The statement of the proposition 8 considers a special case of Proposition 14 with $R_t = R > 0$, $w_t = w \geq 0$, and $\gamma = \delta$. When $\gamma = \delta$, the condition in Proposition 14 for the Ramsey allocation to be Pareto improving becomes

$$\log(1-\delta+\beta\delta)[1 + \frac{\beta\delta}{1-\delta}(1 + \frac{\delta}{1-\delta^2})] - \log(\beta)\frac{\beta\delta}{1-\delta}\frac{1}{1-\delta^2} \geq 0.$$

At no $\beta \in (0,1)$, a limit exists for the prior expression as $\delta \to 1$. But the left- and right-hand limits exist. Since we are interested in the value as δ approaches 1 from below, the relevant limit is the left-hand one, which is

$$\lim_{\delta \to 1^-} \log(\frac{1}{1-\delta+\beta\delta})[1 + \frac{\beta\delta}{1-\delta}(1 + \frac{\delta}{1-\delta^2})] + \log(\beta)\frac{\beta\delta}{1-\delta}\frac{1}{1-\delta^2} = \infty.$$

Thus, we have proved the proposition. □

B Horizontal Redistribution: Implementation

In this section, we provide a tax system that implements the Mirrleesian allocation for any $\lambda \geq 0$ in a market setup where people buy and sell bonds. If $\lambda = 0$, bequest taxes are flat. When $\lambda > 1$, bequest taxes are progressive as long as the CEIS coefficient is weakly less than 1.

In the implementation exercise, the parent's budget constraint is

$$c_0 + b_1 = R_0 b_0 + y_0 - T_b(b_1) + T_0(y_0),$$

where labor income tax T_0 is a function of the parent's declaration and - by the taxation principle - can always be written in terms of the (observable) period 0 income y_0. The function $T_b(\cdot)$ represents the bequest tax that is allowed to depend nonlinearly on the level of bequests b_1 and, for notational simplicity, is assumed to be paid by the parent.[35] The offspring's budget constraints are given by

$$\begin{aligned} c_1 + b_2 &= R_1(1 - \tau_1(y_0))b_1 + y_1 + T_1(y_0), \\ c_2 &= R_2(1 - \tau_2(y_0))b_2 + y_2 + T_2(y_0), \end{aligned}$$

where, for $t = 1, 2$, τ_t, and T_t represent wealth (or inheritance) taxes and labor income taxes, respectively. As indicated, both labor and wealth taxes are allowed to depend on the initial income level y_0 of the dynasty. We denote a tax scheme that implements the efficient allocation by

$$\left(\{T_t^*(\cdot)\}_{t=0,1,2}, T_b^*(\cdot), \tau_1^*(\cdot), \tau_2^*(\cdot) \right).$$

In this tax system, both T_b and τ_1 are taxes on bequests, and we can obviously dispense with one of them. In what follows, we set $\tau_1 \equiv 0$.

Proposition 15. *Assume CEIS preferences and $\beta < 1$. Then, for any $\lambda \geq 0$, there is a tax scheme that implements the corresponding Pareto-efficient allocation. Moreover, (i) if $\lambda = 0$, then the bequest tax function $T_b^*(\cdot)$ is linear in b_1. (ii) If $\lambda > 0$, then the bequest tax is progressive, that is, $T_b^*(\cdot)$ convex, whenever $\rho \leq 1$.*

Proof. In the proof of the proposition, we write all taxes that are functions of y_0 as functions of skill types, s. Since $y_0^*(s)$ increases with s, and hence, there is be a one-to-one relationship between y_0 and s, this is without loss of generality.

First, consider the problem of the parent. A parent who is endowed with asset level $b_0 = K_0$ and skill level s, and declaring to be of skill σ, solves

$$\begin{aligned} V(\sigma|s) := \max_{\{c_t, b_t\}_{t=0}^2} \quad & u(c_0) - v\left(\frac{y_0(\sigma)}{w_0 s}\right) + \gamma\left[u(c_1) + \delta u(c_2)\right] \\ s.t. \quad & c_0 + b_1 = R_0 b_0 + y_0(\sigma) - T_b(b_1) + T_0(\sigma); \\ & b_2 \in \arg\max_{\hat{b}} \; u(\hat{c}_1) + \delta\beta u(\hat{c}_2); \\ & s.t.\; \hat{c}_1 + \hat{b} = R_1(1 - \tau_1(\sigma))b_1 + y_1 + T_1(\sigma); \\ & \qquad \hat{c}_2 = R_2(1 - \tau_2(\sigma))\hat{b} + y_2 + T_2(\sigma). \end{aligned} \tag{48}$$

The definition of the offspring's problem is given by (48) and therefore is not stated separately.

[35] As discussed in Farhi and Werning (2010), it is easy to see that any given bequest tax can be replaced by a tax on the offspring (inheritance tax) with the same characteristics.

Definition. *We say that a system of taxes implements a Pareto-efficient allocation* $\left(\{c_t^*(s)\}_{t=0,1,2}, y_0^*(s)\right)_{s\in S}$ *if there exists a price system,* $(R_t^*, w_t^*)_{t=0,1,2}$*, such that, at this price system and taxes, for each* s *we have* $V(s|s) \geq V(\sigma|s)$ *for all* σ *and for each* s *the solution to* $V(s|s)$ *is* $(\{c_t^*(s)\}_{t=0,1,2}, y_0^*(s))$ *and all markets clear.*

The core of the implementation exercise follows the same logic as in Kocherlakota (2005). We will show that *for all* s, a dynasty whose parent declares σ finds it optimal to choose the allocation prescribed by the efficient allocation to type σ families, $\{c_t^*(\sigma)\}_{t=0,1,2}$ (and, of course, the parent is forced to generate income $y_0^*(\sigma)$). Once this key fact is shown, the inequality $V(s|s) \geq V(\sigma|s)$ is a direct consequence of the incentive compatibility property of the Mirrleesian allocation.

Now, we construct the equilibrium price system and taxes. First, set $R_0^* = F_K(K_0^*, L_0^*) + (1-\theta)$ and $w_0^* = F_L(K_0^*, L_0^*)$ and for $t = 1, 2$, $R_t^* = F_K(K_t^*, 1) + (1-\theta)$ and $w_t^* = F_L(K_t^*, 1)$. The taxes that apply to the offspring are constructed as follows. For a given tax system and facing equilibrium prices, the first-order optimality condition of the offspring's problem in (48) is given by

$$u'(c_1(s)) = (1-\tau_2(s))\beta\delta R_2^* u'(c_2(s)).$$

Since the offspring's problem is strictly concave, this first-order optimality condition, together with the budget constraints, is necessary and sufficient for the solution. We choose $\tau_2^*(s)$ to ensure that the efficient allocation satisfies the offspring's optimality condition:

$$\tau_2^*(s) = 1 - \frac{u'(c_1^*(s))}{\beta\delta R_2^* u'(c_2^*(s))}. \tag{49}$$

We choose the labor income taxes that the offspring faces such that the s component of the efficient allocation is just affordable by the offspring whose parent has reported to be of type s. That is, we set $T_1^*(s)$ and $T_2^*(s)$ such that

$$c_1^*(s) + \frac{c_2^*(s)}{R_2(1-\tau_2^*(s))} = y_1 + T_1^*(s) + \frac{y_2 + T_2^*(s)}{R_2^*(1-\tau_2^*(s))} + R_1^* b_1^*(s).$$

The previous equation also pins down a value for $b_1^*(s)$ for given values of $T_1^*(s)$ and $T_2^*(s)$. Notice that if we change the values of taxes the corresponding value of $b_1^*(s)$ will also change. In other words, there is an indeterminacy regarding the choice of labor income taxes for periods 1 and 2. This set of taxes for the offspring ensures that the offspring whose parent reported to be of type s chooses $(c_1^*(s), c_2^*(s))$.

Before proceeding to construct the parent's taxes, we first compute the saving policy that comes out of the offspring's problem, (48), under these taxes. This is needed since the parent's taxes, which we will define later, depend on the derivative of the offspring's saving policy function. Recall from Proposition 6 that - with CEIS utility - the policy of the offspring is linear and increasing in wealth. We now compute this policy explicitly.[36] Let $M_\tau(s) := \frac{\partial b_2(\cdot,\psi(s))}{\partial b_1}$ be the derivative of the policy where $\psi(s)$ represents the joint sequence of prices $(R_t, w_t)_{t=1,2}$ and taxes $(\hat{T}_1(s), \hat{T}_2(s), \tau_2(s))$ that the offspring faces. Notice that M is allowed to depend on s because the taxes the offspring faces potentially depend on s. Plugging the period budget constraints of the offspring into their first-order optimality

[36]Note that under the previous offspring problem, with CEIS utility and a budget linear in b_1, the proof of linearity of the policy is immediate.

condition for saving, for a given sequence of taxes and prices, the policy of the offspring solves

$$\begin{aligned} u'(c_1) &= \delta\beta R_1(1-\tau_2)u'(c_2) \\ &\Longleftrightarrow \\ (R_1(1-\tau_1)b_1+y_1+T_1-(A_\tau+M_\tau b_1))^{-\rho} &= \delta\beta R_2(1-\tau_2)\left(R_2(1-\tau_2)(A_\tau+M_\tau b_1)+y_2+T_2\right)^{-\rho} \\ &\Longleftrightarrow \\ (R_1(1-\tau_1)b_1+y_1+T_1-(A_\tau+M_\tau b_1)) &= (\delta\beta R_2(1-\tau_2))^{-\frac{1}{\rho}}\left(R_2(1-\tau_2)(A_\tau+M_\tau b_1)+y_2+T_2\right), \end{aligned}$$

where A_τ and M_τ are the constants defining the policy and $\frac{1}{\rho}$ is the intertemporal elasticity of substitution. Since the previous relationship must hold for all b_1, taking the derivative with respect to b_1 on both sides, and evaluated at the efficient tax and price system $\tau_1^*(s), \tau_2^*(s), R_1^*, R_2^*$, we have

$$M_\tau^*(s) = \frac{(1-\tau_1^*(s))R_1^*}{(\delta\beta)^{-\frac{1}{\rho}}\left(R_2^*(1-\tau_2^*(s))\right)^{1-\frac{1}{\rho}}+1}. \tag{50}$$

Now we construct the taxes on parents. The first-order optimality condition of a parent announcing s is

$$u'(c_0(s))\left[1+\frac{\partial T_b(b_1(s))}{\partial b_1}\right] = \gamma R_1^* u'(c_1(s))\left\{1-\tau_1(s)+\frac{M_\tau(s)}{R_1^*}\left[\frac{1}{\beta}-1\right]\right\}.$$

Since the parent's problem is concave, this optimality condition evaluated at the correct offspring policy derivative, given by (50), together with the parent's first-period budget constraint, is necessary and sufficient for a solution. Therefore, we set the T_b function so that the efficient allocation satisfies the parental bequest optimality condition. This implies that

$$\frac{\partial T_b^*(b_1^*(s))}{\partial b_1} = \frac{\gamma R_1^* u'(c_1^*(s))}{u'(c_0^*(s))}\left\{1+\frac{M_\tau^*(s)}{R_1^*}\left[\frac{1}{\beta}-1\right]\right\}-1, \tag{51}$$

where we set the redundant $\tau_1^*(s)=0$.

Now we define the labor tax for the parent so that the parental period 0 budget constraint holds with equality under the effiicient allocation, meaning we choose $T_0^*(s)$ to satisfy

$$c_0^*(s)+b_1^*(s) = R_0^* b_0 + y_0^*(s) - T_b(b_1^*(s)) + T_0^*(s).$$

Under these taxes, the parents who report to be of type s choose the allocation that is intended for family type s. Finally, notice that the actual level of skills does not enter into the intertemporal budget constraints or the first-order optimality conditions of agents. This implies that *for all s*, a parent declaring σ (possibly different from s) will find it optimal to choose the Mirrlees consumption plan intended for σ type family, $c_t^*(\sigma)$, $t=0,1,2$. Since the Mirrlees allocation is incentive compatible the parents will actually tell the truth.

We are done with proving that the taxes defined implement the efficient allocation. Next we derive the properties of taxes.

Properties of Optimal Bequest Taxes

We first compute the wealth tax on the offspring. To do so, we need to first derive the planner's optimality condition for the offspring's savings. Consider the perturbation $z_0^\varepsilon(s) = z_0^*(s)$, $z_1^\varepsilon(s) = z_1^*(s)+\varepsilon$, and $z_2^\varepsilon(s) = z_2^*(s) - \frac{\varepsilon}{\delta}$. Since the efficient allocation, denoted by the starred allocation, solves the planning problem, $\varepsilon = 0$ must be the optimal level of perturbation, implying that

$$\lambda(1-\beta) - \frac{\mu_1}{u'(c_1^*(s))} + \frac{\mu_1}{\delta R_2^* u'(c_2^*(s))} = 0, \tag{52}$$

which can be rearranged as

$$\frac{\beta-1}{\beta}\left(\frac{\lambda u'(c_1^*(s))}{\mu_1} - 1\right) = \frac{u'(c_1^*(s))}{\delta\beta R_2^* u'(c_2^*(s))} - 1. \tag{53}$$

Evaluating (53) in the optimal wealth tax expression given by (49), we get

$$\tau_2^*(s) = \frac{1-\beta}{\beta}\left(\frac{\lambda u'(c_1^*(s))}{\mu_1} - 1\right). \tag{54}$$

Evaluating (24) in the optimal bequest tax expression given by (51), we get

$$\frac{\partial T_b^*(b_1^*(s))}{\partial b_1} = \frac{\left\{1 + \frac{M_\tau^*(s)}{R_1^*}\left[\frac{1}{\beta} - 1\right]\right\}}{1 + \frac{\lambda}{\mu_0\gamma} u'(c_0^*(s))} - 1. \tag{55}$$

Case (i): $\lambda = 0$ (Ramsey allocation): From condition (54), $\tau_2^*(s) = 1 - 1/\beta$, which is independent of s. This implies that $M_\tau^*(s)$, given by (50), is independent of s. The optimal bequest tax given by (55) reduces to

$$\frac{\partial T_b^*(b_1^*(s))}{\partial b_1} = \frac{M_\tau^*}{R_1^*}\left[\frac{1}{\beta} - 1\right] > 0.$$

In words, the marginal tax on bequests is strictly positive and is flat in the amount bequeathed. In fact, this tax expression is identical to the bequest tax of Proposition 5 in the baseline model without horizontal heterogeneity.

Case (ii) $\lambda > 0$: Notice first that when $\lambda > 0$, (54) implies that τ_2^* depends on s. Since $c_1^*(s)$ increases with s and $0 < \beta < 1$, $\tau_2^*(s)$ *decreases with* s. If $\rho \leq 1$, this implies that $M_\tau^*(s)$ is increasing in s. Since $c_0^*(s)$ is increasing in s, in this case we get that the right-hand side of equation (55) is increasing in s. If we show that $b_1^*(s)$ increases with s, we get a convex T_b^* as claimed. To do so, we need to show that in our implementation of the Pareto-efficient allocation, $b_1^*(s)$ increases with s. From the offspring budget constraint at the efficient allocation and taxes, and recalling that $\hat{T}_t^*(s) = \tau_t^*(s) R_t^* b_t^*(s)$, for all s we have: $c_1^*(s) + b_2^*(s) = R_1^* b_1^*(s) + y_1$ and $c_2^*(s) = R_2^* b_2^*(s) + y_2$, which combined together generates

$$c_1^*(s) + \frac{c_2^*(s)}{R_2^*} = y_1 + \frac{y_2}{R_2^*} + R_1^* b_1^*(s).$$

Since, $c_1^*(s)$ and $c_2^*(s)$ are increasing in s then $b_1^*(s)$ must also increase with s. In fact, the value of $b_1^*(s)$ is fully pinned down from this equation for all s.

□

C The Optimality Condition for Illiquid Bequests

In this section, we explicitly derive equation (35) in the main text, which is the optimality condition for illiquid bequests. The first-order condition for illiquid bequests is given by

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

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

$$\frac{db_{t+3}}{db_{t+2}^o} = \frac{\partial b_{t+3}}{\partial b_{t+2}^o} + \frac{\partial b_{t+3}}{\partial b_{t+2}} \frac{db_{t+2}}{db_{t+2}^o}. \tag{56}$$

Now, plugging (56) into the parent's first-order condition for illiquid bequests and rearranging, we get

$$u'(c_t^o) = \gamma \left(R_{t+1}R_{t+2}\delta^2 u'(c_{t+2}) + \delta \frac{\partial b_{t+2}}{\partial b_{t+2}^o} \Xi_{t+1} + \delta^2 \frac{\partial b_{t+3}}{\partial b_{t+2}^o} \Xi_{t+2} \right), \tag{57}$$

where

$$\Xi_{t+1} = \left[-u'(c_{t+1}) + \delta \left(\sum_{i=2}^{I-1} \delta^{i-2} u'(c_{t+i}) \frac{\partial c_{t+i}}{\partial b_{t+2}} + \delta^{I-2} V_1(a_{t+I}, Q_{t+I}) R_{t+I} \frac{\partial b_{t+I}}{\partial b_{t+2}} \right) \right], \tag{58}$$

$$\Xi_{t+2} = \left[-u'(c_{t+2}) + \delta \left(\sum_{i=3}^{I-1} \delta^{i-3} u'(c_{t+i}) \frac{\partial c_{t+i}}{\partial b_{t+3}} + \delta^{I-3} V_1(a_{t+I}, Q_{t+I}) R_{t+I} \frac{\partial b_{t+I}}{\partial b_{t+3}} \right) \right]. \tag{59}$$

Using (34) in (58), we get $\Xi_{t+1} = 0$. Using (29), which describes the offspring's optimality condition for saving during periods $i = 2, ..., I-1$, in (59), we get

$$\Xi_{t+2} = \sum_{t=2}^{I-1} \delta^{i-2} u'(c_{t+i}) \left[-1 + \frac{1}{\beta_{i+1}} \right] \frac{\partial b_{t+i+1}}{\partial b_{t+3}}.$$

Pluggin Ξ_{t+1} and Ξ_{t+2} into (57), we reach equation (35) in the text. □

D Naivete

In the main body of the paper, we assume that parents are sophisticated in the sense that they are aware of the fact that their descendants disagree with them about how much to save for their old age. 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 believe (incorrectly) that their descendants fully agree with themselves* regarding the intertemporal allocation of resources. We focus on the Ramsey wedge, but the results can be readily extended to other Pareto-efficient allocations following the approach in Section 5.1. We show that even naive parents bequeath too much in equilibrium relative to the Ramsey level, and hence it is optimal to tax parental transfers.[37]

[37] It might be worthwhile to note that the spirit of government intervention in this section is in some sense "paternalistic" in that it is aiming to correct a bias emerging from bounded rationality. No "paternalistic"

The problem of a naive parent in period t is given by

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

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:

$$\bar{b}_{t+2}(b_{t+1}, Q_{t+1}) = \arg \max_{\tilde{b}_{t+2} \geq -B(Q_{t+2})} u\left(R_{t+1} b_{t+1} + w_{t+1} - \tilde{b}_{t+2}\right) + \delta \bar{V}(R_{t+2} \tilde{b}_{t+2} + w_{t+2}, Q_{t+2}). \quad (60)$$

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, (60), the offspring discounts the future with δ instead of $\beta\delta$. This reflects the assumption that parents naively believe that their descendants are in full harmony with them regarding intertemporal preferences. We denote the naive parents' value function by $\bar{V}(a_t, Q_t)$. 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. Notice that the saving policy function of the offspring given in (60) is incorrect; it 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 true saving behavior 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 the bequest tax. The parent naively believes that the offspring solves (60), and hence, the following first-order condition describes the offspring's saving behavior:

$$u'(\bar{c}_{t+1}(b_{t+1}, Q_{t+1})) = \delta R_{t+2} \bar{V}_1(\bar{a}_{t+2}(b_{t+1}, Q_{t+1}), Q_{t+2}). \quad (61)$$

Taking the first-order optimality condition with respect to bequests in the naive parent's problem and substituting (61) in, we find that the optimality condition for bequests in equilibrium is given by

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

policy has been proposed so far in the paper.

We therefore define the *naive Ramsey bequest wedge* as follows:

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

where, $\bar{c}_{t+1}(b_{t+1}^*, Q_{t+1}^*)$ 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 that, given the Ramsey allocation, the naive parent would like to increase bequests. Remember that in the Ramsey allocation, we have

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

Plugging this condition into the definition of the naive bequest wedge implies that

$$\overline{BW}_t = 1 - \frac{u'(c_{t+1}^*)}{u'\left(\bar{c}_{t+1}(b_{t+1}^*, Q_{t+1}^*)\right)}. \tag{62}$$

Therefore, in order to establish the sign of $\overline{BW}_t$, 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 existence of an efficient wedge, τ_{t+2}^*, then his first-order condition reads as

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

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 implies that

$$\bar{c}_{t+1}(b_{t+1}^*, Q_{t+1}^*) < c_{t+1}^*.$$

This ordering is intuitive: an agent with a lower discount factor, $\beta\delta$, tends to consume more in the current period than an agent with a higher discount factor, δ, and the parent naively believes that the offspring has the higher discount factor. Therefore, we have the following result:

Proposition 16. *Assume agents are naive and $\beta < 1$; then the naive Ramsey bequest wedge is positive.*

Therefore, we conclude that even when agents are naive, parents bequeath too much relative to the Ramsey level, and hence, it is optimal to tax bequests. 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 bequeath too much relative to the Ramsey level not because they want to compensate for their offsprings' undersavings but 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 the naive bequest wedge in (62).