DISCUSSION PAPER SERIES

DP16114
(v. 2)

The Cross-Sectional Implications of the Social Discount Rate

Maya Eden

MACROECONOMICS AND GROWTH

ISSN 0265-8003

The Cross-Sectional Implications of the Social Discount Rate

Maya Eden

Discussion Paper DP16114
First Published 05 May 2021
This Revision 21 July 2022

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

The Cross-Sectional Implications of the Social Discount Rate

Abstract

How should policy discount future returns? The standard approach to this normative question is to ask how much society should care about future generations. This paper establishes an alternative approach, based on the social desirability of age-based redistribution. The social discount rate is below the market interest rate only if it is desirable to increase the consumption of the young at the expense of the old. Along the balanced growth path, small deviations of the social discount rate from the market interest rate imply large welfare gains from redistributing consumption across age groups.

JEL Classification: D6, E61

Keywords: Overlapping Generations, Utilitarianism, age-based inequality, balanced growth path

Maya Eden - meden@brandeis.edu
Brandeis University and CEPR

Acknowledgements
I thank Geir Asheim, Matthias Doepke, Ben Eden, Mikhail Golosov, Paul Kelleher, Pete Klenow, Kevin Kuruc, Antony Millner and John Quiggin for helpful comments. This paper has also benefited from comments by seminar participants at the Chicago Fed, Iowa State University, the CEPR Macroeconomics and Growth Programme Online Meeting, and the Oxford Workshop on Global Priorities Research.

The Cross-Sectional Implications of the Social Discount Rate

Maya Eden*

May 24, 2022

Abstract

How should policy discount future returns? The standard approach to this normative question is to ask how much society should care about future generations. This paper establishes an alternative approach, based on the social desirability of age-based redistribution. The social discount rate is below the market interest rate only if it is desirable to increase the consumption of the young at the expense of the old. Along the balanced growth path, small deviations of the social discount rate from the market interest rate imply large welfare gains from redistributing consumption across age groups.

JEL Classification: D63, E61
Keywords: Overlapping generations, utilitarianism, prioritarianism, age-based inequality

*Affiliations: Brandeis University, CEPR, and the Global Priorities Institute, Oxford. Contact address: Maya Eden, Economics Department, Brandeis University, 415 South Street, Waltham, MA 02453; Email: meden@brandeis.edu. I thank Geir Asheim, Matthias Doepke, Ben Eden, Mikhail Golosov, Paul Kelleher, Pete Klenow, Kevin Kuruc, Antony Millner and John Quiggin for helpful comments. This paper has also benefited from comments by seminar participants at the Chicago Fed, Iowa State University, the CEPR Macroeconomics and Growth Programme Online Meeting, and the Oxford Workshop on Global Priorities Research.

1 Introduction

The social discount rate is the rate at which policy should discount future returns. There is substantial controversy about what is the appropriate benchmark for quantifying it.[1] In Nordhaus [2007], the annual social discount rate is 6%, consistent with long-run estimates of the real interest rate.[2] In contrast, Stern [2008] calibrates a social discount rate based on concerns for intergenerational equity, and obtains a social discount rate of 1.5%.

This paper provides another perspective on this debate by exploring the cross-sectional implications of the social discount rate. When generations are overlapping, the extent to which social preferences care about future generations simultaneously determines the social discount rate, and the marginal welfare weights of younger people compared to older ones. Along the balanced growth path, the distribution of consumption across age groups is socially optimal if and only if the social discount rate is equal to the market interest rate.

This result is useful because it allows us to check which social discount rates are consistent with our moral views about redistributing consumption across age groups. It turns out that even plausible deviations of the social discount rate from the market interest rate generate uneasy implications. For example, assuming a market interest rate of 6% (as in Nordhaus [2007]), the 1.5% social discount rate proposed by Stern [2008] implies that it must be socially desirable to reduce the consumption of a 70-year-old by $1 in order to increase the consumption of a 20 year-old by 10 cents. Such extreme ageist implications can be avoided only by choosing a social discount rate that is closer to the market interest rate.

The reason that Stern [2008] and others advocate for a low social discount rate is based on ethical considerations. If the social welfare function takes the

[1]See Greaves [2017] and Millner and Heal [2022] for recent reviews. See also Baumol [1968], Arrow et al. [2013], Gollier and Hammitt [2014], Kelleher [2017] and Drupp et al. [2018].

[2]It is important to note that there is a separate debate about what is the appropriate benchmark for quantifying the long-run interest rate, especially given the lack of assets with very long maturities. See Stern [2008] and Millner and Heal [2022] for further details.

discounted-utilitarian form, then the social discount rate reflects the social rate of pure time preference.[3] A high social rate of pure time preference is considered "ethically indefensible" (Ramsey [1928]), because it implies a morally-arbitrary preference for the well beings of people that are born sooner.

However, a low rate of pure time preference also has implications for how consumption should be distributed across age groups. To study these implications, I consider a standard neoclassical growth model with overlapping generations. A discounted-utilitarian planner can choose both how much to save and how to allocate consumption across people of different ages. With a low social rate of pure time preference (as in Stern [2007]), the optimal policy converges to a balanced growth path in which the consumption of 20 year-olds is 120% higher than its current equilibrium level, and the consumption of 70 year-olds is 20% *lower* than its current equilibrium level. Transitioning to this new balanced growth path requires capital accumulation, which is financed entirely by the current old: the optimal policy involves cutting the consumption of present-day 70 year-olds by 50% while increasing the consumption of 20 year-olds by 40%.

This analysis illustrates a tension between the intergenerational equity concerns emphasized in Ramsey [1928] and Stern [2008] and society's concern for its elderly. If one rejects the social desirability of these age-based transfers, then one must reject the discounted utilitarian framework that favors them. Of course, one need not reject either; perhaps massive redistribution from the old to the young really is the right thing to do. But, if not, then alternative normative frameworks should be deployed.[4]

[3] Specifically, the social discount rate is determined by the famous Ramsey [1928] equation: $r^s = \rho^s + \eta g$, where g is the growth rate of consumption; η is the elasticity of utilities with respect to consumption; and ρ^s is the pure rate of time preference in the planner's objective, $\sum_{t=0}^{\infty}\left(\frac{1}{1+\rho^s}\right)U_t$ (where U_t is the utility of generation t).

[4] For example, a prioritarian criterion (as in Adler [2019]) may generate a high social discount rate based on its aversion to inequality in utilities. In this framework, the reason for discounting future consumption is that future generations are better off than people alive today – and not that their well-being matters less. Because this framework can generate social discount rates that are equal to the market interest rate, it can imply that the optimal distribution of consumption across age groups is optimal.

This paper is related to a large literature on social discounting (see Millner and Heal [2022] for a recent review). The social discount rate is usually discussed in the context of non-overlapping generations models. Notable exceptions are Calvo and Obstfeld [1988], Quiggin [2012], Schneider et al. [2012], and Fleurbaey and Zuber [2015]. Calvo and Obstfeld [1988] consider a discounted-utilitarian social objective and an overlapping generations economy. They establish that, in order to implement the social optimum, transfers between age groups may be necessary. This paper adds to this literature by quantitatively exploring the relationship between social discounting and age-based transfers away from the optimal policy. The key insight here is that small deviations of the social discount rate from the market interest rate imply large differences in the marginal social welfare weights of different age groups. This quantitative insight is absent from previous work.

From a theoretical perspective, this paper is related to Quiggin [2012], who utilizes the overlapping generations structure to study the relationship between intertemporal social preferences and social preferences over the allocation of consumption across age groups. He establishes that, when people do not discount their own future utilities, then a social objective of maximizing the sum of current utilities is consistent with a social objective of maximizing the (undiscounted) sum of all current and future utilities. A tension between these two objectives arises only when individual preferences feature a positive rate of pure time preference (see also Schneider et al. [2012]).

This paper is also related to Farhi and Werning [2007] and Barrage [2018], who study the policy implications of a social discount rate that is lower than the market interest rate. These papers do not have any implications for the optimal distribution of consumption across age groups, which is the focus of this paper.[5] However, similar to this paper, they illustrate that deviations of the social discount rate from the market interest rate have far-reaching implications for optimal policy design.

[5] Farhi and Werning [2007] and Barrage [2018] consider a model of consecutive generations, in which only one generation is alive in each period. Consequently, there is no cross-sectional variation in age.

2 A decomposition of the social discount rates

Time is discrete and indexed $t = 0, ...$ In each time period, one generation is born, and lives for $1 < T < \infty$ periods. The assumption that $T > 1$ implies that generations are overlapping. For now, I assume that the size of each cohort is fixed.[6]

Let c_a^t denote generation t's consumption at age a. Let $\mathbf{c}^t = (c_1^t, ..., c_T^t)$ denote the consumption sequence of generation t, and let $\mathbf{c} = (\mathbf{c}^0, ...)$ denote the intergenerational consumption allocation.

The social preference relation is represented by a differentiable social welfare function, $W(\mathbf{c})$.[7] Given an allocation of consumption, $\mathbf{c}$, define the social marginal rate of substitution between c_a^t and $c_{a'}^{t'}$ as

$$MRS((a,t),(a',t')) = (\frac{\partial W(\mathbf{c})}{\partial c_a^t})/(\frac{\partial W(\mathbf{c})}{\partial c_{a'}^{t'}}) \tag{1}$$

Note that c_a^t is the consumption of generation t that takes place in period $t+a-1$, when that generation is aged a. If $t'+a' > t+a$, then the consumption $c_{a'}^{t'}$ takes place in a later date; in this case, the *social discount rate* between generation t in period $t+a-1$ and generation t' in period $t'+a'-1$ is

$$r^s_{a,t,a',t'} = MRS((a,t),(a',t'))^{\frac{1}{t'+a'-t-a}} - 1 \tag{2}$$

The social discount rate is the required average rate of return on a small investment at time $t+a-1$ which is financed by generation t, and benefits generation t' in period $t'+a'-1$.

Define the *social-individual discount rates* as

$$r^{si}_{a,a',t} = r^s_{a,t,a',t} \tag{3}$$

The social-individual discount rate, $r^{si}_{a,a',t}$, is defined so that society is roughly

[6] Allowing for population growth does not change the results, but complicates notation. In the quantitative exercise in section 5 I incorporate population growth.

[7] This assumption rules out the maximin welfare criterion advocated by Rawls [1974] (see also Asheim and Zuber [2013] in an intergenerational context).

indifferent with respect to taking ϵ units of consumption away from generation t at age a, and compensating it with $(1 + r^{si}_{a,a',t})^{a'-a}\epsilon$ additional units of consumption at age a'. This is the *social* rate of discount for an individual's own consumption; in principle, it may be different from the rate at which the individual discounts his own future consumption.

If, at some period τ, generation t is aged a and generation t' is aged a', then $MRS((a,t),(a',t'))$ captures the relative *distributional weights* of generations t and t' in period τ. In period τ, it is socially desirable to take one small unit of consumption from generation t and give it to generation t' if and only if $MRS((a,t),(a',t')) < 1$. It is useful to denote these cross-sectional relative distributional weights by

$$\mu^{\tau}_{a,a'} = MRS((a, \tau - a + 1), (a', \tau - a' + 1)) \tag{4}$$

The results in this paper build on the insight that, because generations are overlapping, the relative distributional weights, $\{\mu^{\tau}_{a,a'}\}_{a,a',\tau}$, and the social-individual discount rates, $\{r^{si}_{a,a',t}\}_{a,a',t}$, are sufficient for recovering the entire set of social discount rates, $\{r^{s}_{a,t,a',t'}\}_{a,a',t,t'}$.

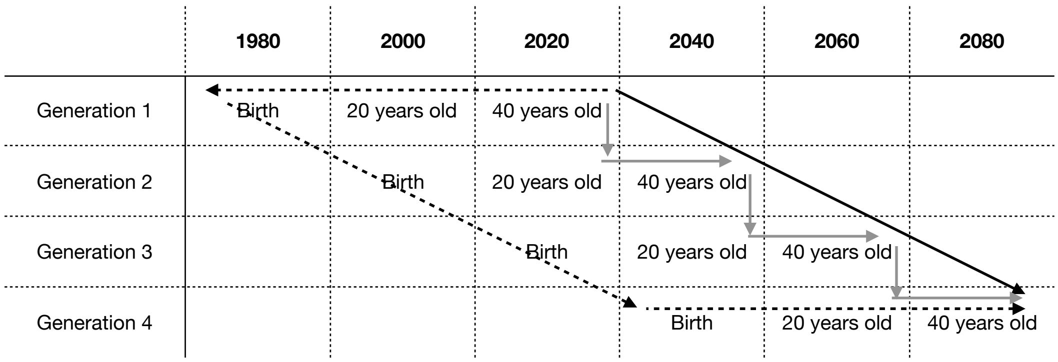
Proposition 1. *For every $t' > t$, the social discount rate $r^{s}_{a,t,a',t'}$ is given by*

$$(1 + r^{s}_{a,t,a',t'})^{t'+a'-(t+a)} = \left(\prod_{\tau=t+a-1}^{t'+a'-2} (1 + r^{si}_{1,2,\tau})\right)\left(\mu^{t+a-1}_{a,1}\mu^{t'+a'-1}_{2,a'}\prod_{\tau=t+a}^{t'+a'-2}\mu^{\tau}_{2,1}\right)$$

The proof of this proposition is in the appendix, together with other omitted proofs. This result decomposes the social discount rates into two terms. The first depends only on the social-individual discount rates. The second depends only on the relative distributional weights of different age groups.

Figure 1 illustrates the decomposition. The solid black line represents a dollar transfer between a 40 year-old in 2020, and a 40 year-old in 2080. Proposition 1 is based on the observation that this transfer can be implemented through a sequence of transfers that involve either transferring within-people, across time, or between people, within-time. This sequence is illustrated with

Figure 1: Decompositions of the social discount rate



Note: The solid back arrow represents a transfer between two 40-year-olds, one in 2020 and the other in 2080. The gray arrows represent the decomposition in Proposition 1, which is based on the insight that this transfer can be done through a sequence of transfers within-people and across-time, and transfers between people in a given time. The dashed arrows correspond to the decomposition in Fleurbaey and Zuber [2015].

the solid grey lines. In 2020, the dollar is transferred from the 40 year-old to a 20 year-old. Then, it is transferred from the 20 year-old in 2020 to the same person 20 years later, when he is 40. At that point, the dollar is transferred to a contemporaneous 20 year-old, and so on and so forth. This sequence of transfers is composed only of transfers between people in the same time period (vertical lines), and transfers within people across time (horizontal lines). The social desirability of the former depends on the distributional weights of different age groups, and the social desirability of latter depends on the social-individual discount rates.

The figure also illustrates the difference between the decomposition here and the decomposition in Fleurbaey and Zuber [2015] (section 6). Their decomposition builds on the observation that a transfer between two individuals can be implemented as a transfer between them at their respective births, plus transfers within each of their lifetimes. Based on their decomposition, Fleurbaey and Zuber [2015] conclude that the long-run social discount rate is determined by the marginal welfare gains from reallocating resources between two people at their respective births. The decomposition here shows that these welfare gains are determined by the social desirability of transferring resources across age groups: how much we care about the far future is related to how much we care about the current young.

Proposition 1 is useful because it reduces the question of social discounting across generations to two sub-questions. The first is, "how should society discount an individual's own future consumption?" Here, the common approach is to evoke the Pareto principle: if people discount their own future consumption at a certain rate, then society should respect their preferences and discount their future consumption at that rate as well. The second question is, "how should consumption be distributed across people alive today?" Here, the most common approach is utilitarian: a transfer between two individuals is desirable provided that it increases the total sum of their flow utilities.

3 An illustrative example

The usefulness of this decomposition can be illustrated with the following simple example. In this example, each generation is alive for two periods. The market interest rate is constant and equal to r. The equilibrium allocation is Allocation A.

Allocation A

Generation	$t=0$	$t=1$	$t=2$	$\ldots$	$t=n$	$t=n+1$
0	c_1^0	c_2^0				
1		c_1^1	c_2^1			
$\vdots$				$\ddots$		
n					c_1^n	c_2^n

If individuals from generation 0 can borrow and save at the market interest rate, then they choose their consumption sequences so that they are exactly indifferent with respect to saving an additional unit. It follows that individuals

Allocation B

Generation	$t=0$	$t=1$	$t=2$	$\dots$	$t=n$	$t=n+1$
0	c_1^0-1	$c_2^0+(1+r)$				
1		c_1^1	c_2^1			
$\vdots$				$\ddots$		
n					c_1^n	c_2^n

Allocation C

Generation	$t=0$	$t=1$	$t=2$	$\dots$	$t=n$	$t=n+1$
0	c_1^0-1	c_2^0				
1		$c_1^1+(1+r)\mu$	c_2^1			
$\vdots$				$\ddots$		
n					c_1^n	c_2^n

from generation 0 are indifferent between the equilibrium allocation, A, and an alternative allocation, B, in which their consumption is reduced by one (small) unit period 0, and increased by $1+r$ units in period 1.

If the social preference relation satisfies the standard Pareto condition, then it must be consistent with the Pareto indifference condition:[8] if all generations are indifferent between two allocations, then society must be indifferent between them as well. When there are no externalities, other generations care only about their own consumption sequences, which are the same in A and B. In this case, all generations are indifferent between the two allocations, and the

[8]The social preference relation satisfies the Pareto condition if it can be written as $W(\mathbf{c}) = \tilde{W}(\{U_t(\mathbf{c}^t)\}_{t=0}^{\infty})$, where $\tilde{W}$ is strictly increasing.

Allocation D

Generation	$t=0$	$t=1$	$t=2$	$\dots$	$t=n$	$t=n+1$
0	c_1^0-1	c_2^0				
1		c_1^1	$c_2^1+(1+r)^2\mu$			
$\vdots$				$\ddots$		
n					c_1^n	c_2^n

Allocation E

Generation	$t=0$	$t=1$	$t=2$	$\dots$	$t=n$	$t=n+1$
0	c_1^0-1	c_2^0				
1		c_1^1	c_2^1			
$\vdots$				$\ddots$		
n					$c_1^n+((1+r)\mu)^n$	c_2^n

Pareto indifference condition requires society to be indifferent as well. Thus,

$$\text{Allocation A} \;\sim\; \text{Allocation B}$$

where $\sim$ denotes the social indifference relation.[9]

This reasoning shows that, when people can frictionlessly borrow and save at the market interest rate and there are no externalities, then the social-individual discount rates are pinned down by the Pareto principle, and it holds that $r^{si}_{a,a',t}=r$ for all a,a' and t.

If, instead, people from generation 0 are borrowing constrained, then they

[9] $W(\mathbf{c}_A)=W(\mathbf{c}_B)$, where $\mathbf{c}_A$ and $\mathbf{c}_B$ are the consumption allocations in A and B, respectively.

strictly prefer allocation A over allocation B. In this case, the social-individual discount rates would be higher than the market interest rate ($r^{si}_{a,a',t} > r$). Saving constraints would have the opposite implication.

The presence of externalities may also affect the social-individual discount rates. If the consumption of generation 0 imposes some externalities on other generations, then other generations may not be indifferent between A and B, even though their consumption sequences are the same in both. For example, consider the case in which consumption is associated with driving, which creates traffic, noise, and air pollution. In this case, generation 1 may have a strict preference for A over B, because it experiences the negative externalities from generation 0's consumption in period 1, but not in period 0. In this case (assuming that all other generations are indifferent), Pareto requires society to have a strict preference for A as well, implying that $r^{si}_{1,2,0} > r$.[10] How negative consumption externalities affect social-individual discount rates in periods $t > 0$ is less clear: generation $t-1$ prefers that generation t saves more (and consumes less in period t), while generation $t+1$ prefers that generation t saves less (and consumes less in period $t+1$). Of course, positive consumption externalities would have the opposite implications.

Before proceeding, it is worth noting that, while commonly assumed, the Pareto condition is somewhat controversial in an inter-temporal context.[11] Because of various forms of present-bias and dynamic inconsistency, it is possible to argue that people save too little for their own good. This implies that, while individuals may be indifferent between A and B, society should strictly prefer B. These paternalistic concerns imply social-individual discount rates that are lower than the market interest rate ($r^{si}_{a,a',t} < r$).

Next, define the scalar $\mu = \mu^1_{2,1}$ so that

$$\text{Allocation B} \sim \text{Allocation C}$$

[10]Note that $r^{si}_{1,2,0} < r$ is inconsistent with the Pareto condition, because both generation 0 and generation 1 strictly prefer A over an allocation in which generation 0 consumes 1 unit less in period 0, and an additional positive amount in period 1 that is less than $1 + r$. Hence, Pareto requires that generation 0 strictly prefers to save at the social discount rate, and thus $r^{si}_{1,2,0} > r$.

[11]See, for example, Caplin and Leahy [2004].

This parameter captures the desirability of redistributing consumption from older people to younger ones. For example, $\mu < 1$ implies that it is desirable to redistribute towards younger people; in this case, society is willing to take away $(1+r)$ units from an old person in order to give a younger person the smaller amount of $(1+r)\mu$.

The next step mimics the first step: if individuals from generation 1 can borrow and lend at the market interest rate, then they choose their consumption sequences so that they are indifferent with respect to saving another $(1+r)\mu$ small units at the market interest rate. This means that they are indifferent with respect to giving up $(1+r)\mu$ units in period 1 in exchange for an additional $(1+r)\left((1+r)\mu\right) = (1+r)^2\mu$ units in period 2. Consequently, they are indifferent between allocations C and D. Once again, the Pareto principle implies that

$$\text{Allocation C} \sim \text{Allocation D}$$

If the value of μ is constant throughout time, then repeating the same argument and using the transitivity of the social indifference relation implies that

$$\text{Allocation A} \sim \text{Allocation B} \sim ... \sim \text{Allocation E}$$

This example establishes that, if the relative distributional weights of different age groups are time invariant, and if the social preference relation is consistent with the Pareto principle, then $1 + r^s_{1,0,1,n} = (1+r)\mu$. This suggests a straightforward mapping between the social discount rates and the cross-sectional distributional weights.

4 Balanced growth path

This section explores the implications of Proposition 1 along a balanced growth path. As famously illustrated by Kaldor [1961] and later by Herrendorf et al. [2019], over long time horizons (several decades), the growth rate of output per-capita in the US has been roughly constant, and the real interest rate has been trend-less. These regularities are consistent with a balanced-growth path equilibrium in which the interest rate is constant and output per-capita grows at a constant rate.

Let y_a^t denote the income of generation t at age a. I assume that $y_a^t = (1+g)^t y_a^0$ for some $g > 0$. This allows for life-cycle variation in earnings, but assumes that the income of each age group grows at the same rate. Further, I assume that all generations have the same preferences over consumption sequences, which can be represented by a utility function, $U(\cdot)$. To be consistent with a balanced-growth-path equilibrium, I assume that U is homogeneous.[12]

Section 5 discusses how the market interest rate, r, is determined in the closed-economy equilibrium of a simple over-lapping generations model. However, for the results that follow, it is sufficient to assume that r is exogenously fixed and time-invariant. This can be interpreted as a small-open-economy assumption, or as a partial equilibrium model of a balanced growth path.

The consumption sequence of generation t solves the following constrained optimization problem:

$$\mathbf{c}^t = \arg\max_{\mathbf{c}^t} U(\mathbf{c}^t) \text{ s.t. } \sum_{a=1}^{T} \frac{c_a^t}{(1+r)^a} = \sum_{a=1}^{T} \frac{y_a^t}{(1+r)^a} \tag{5}$$

Each generation chooses its consumption stream optimally, subject to its intertemporal budget constraint. This optimization problem abstracts from any borrowing constraints, and assumes that people can always borrow and save at the market interest rate. The realistic possibility of borrowing constraints complicates the analysis, but does not affect the main results. I discuss this further in Appendix D.

I restrict attention to a particular class of social welfare functions which imply time-invariant distributional weights along the balanced growth path.

[12]In a production economy, there is an equilibrium balanced growth path if preferences are homothetic, total factor productivity grows at a constant rate, and the production function has constant returns to scale in capital and labor. Section 5 discusses the (closed-economy) equilibrium balanced growth path in a simple neoclassical growth model.

The social welfare function, W, is of the form

$$W(\mathbf{c}) = \sum_{t=0}^{\infty} \left(\frac{1}{1+\rho^s} \right)^t \phi(U(\mathbf{c}^t)) \tag{6}$$

where ϕ is some strictly increasing and homogeneous function, and $\rho^s > -1$.[13]

This functional form nests two important special cases. If $\phi(U(\mathbf{c}^t))$ is a measure of generation t's cardinal lifetime utility, then the social welfare function is a discounted sum of lifetime utilities (note that, as ϕ is strictly increasing, $\phi(U(\cdot))$ is a representation of individual preferences, which can be interpreted as a utility function).[14] If, instead, $U(\mathbf{c}^t)$ is a cardinal measure of generation t's lifetime utility (and not merely an ordinal representation of its preferences) and ϕ is concave, then the social welfare function is of the discounted-prioritarian form. The prioritarian objective differs from the utilitarian objective in its aversion to inequality in lifetime utilities.[15] In both cases, ρ^s is the social rate of pure time preference, or the rate at which the social objective discounts the well-beings of future generations.

The following proposition establishes that, along the balanced growth path, the social discount rate and the cross-sectional distributional weights are closely related.

Proposition 2. *Along the balanced growth path,*

1. *There exists $\mu > 0$ such that $\mu^t_{a',a} = \mu^{a'-a}$ for all $a, a' \leq T$ and t.*

[13] Given the assumptions that $g > 0$ and ϕ is strictly increasing, the infinite sum in equation 6 does not converge for $\rho \leq 0$ (and, depending on ϕ, may also not converge for some $\rho^s > 0$). For these parameters, the social welfare function can only be defined over an arbitrarily large, but finite, number of cohorts.

[14] In this case, the social welfare function can also be interpreted as rank-discounted utilitarian, as in Zuber and Asheim [2012]. According to this approach, the rate of pure time preference, ρ^s, depends on whether future generations are likely to be better or worse off than current generations. As I am restricting attention to a balanced growth path in which $g > 0$, a specification in which $\rho^s > 0$ can be viewed as contingent on the assumption that lifetime utilities grow at a positive rate. This criterion allows for the rate of pure time preference to change if the rate of growth becomes negative.

[15] For a discussion of the prioritarian welfare criterion, see, for example, Adler [2019], Chapter 3.1.

2. *For every age, a, and time periods, $t' > t$, it holds that $r^s_{a,t,a,t'} = r^s$, where*

$$1 + r^s = (1 + r)\mu$$

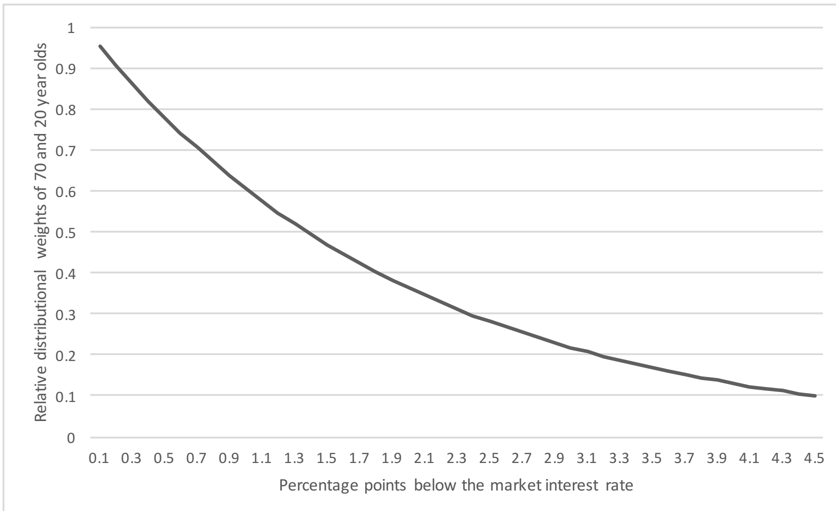
The rate r^s is the social discount rate applied to transfers between two people of the same age, alive in different periods. This proposition establishes that whenever $r^s \neq r$, then the cross-sectional distribution of consumption across age groups is suboptimal.

To fully appreciate the ethical implications of this result, consider the case in which $r^s < r$. By the above proposition, in this case, it must hold that $\mu < 1$. Hence, it is socially desirable to redistribute from old people to younger people. By the first clause of the proposition, it is socially desirable to reduce the consumption of someone aged a by 1 unit in order to increase the consumption of a newborn by μ^a units. In the limit $T \to \infty$, the age a can be chosen to be arbitrarily large; consequently, μ^a can be made arbitrarily small. In this case, it is socially desirable to take away \$1 from an extremely elderly person in order to give a young person basically nothing. This is almost a violation of the Pareto condition, as social welfare is "improved" by making an elderly person worse-off without any detectible gain to anyone else.

One might accept this problematic implication on the grounds that, in practice, the age distribution is bounded; consequently, the relative marginal social welfare weights of old and young people are bounded by μ^T. However, the quantitative implications are uneasy even for a realistic age distribution. For example, assume that the interest rate is 6% (as in Nordhaus [2007]) and that the social discount rate is 1.5% (as in Stern [2008]). These numbers imply that $\mu \approx 0.955$. Under these assumptions, it is welfare-improving to reduce the consumption of a 70 year-old by \$1 in order to increase the consumption of a 20 year-old by 10 cents.

Figure 2 plots the implied relative distributional weights of 70 year-olds and 20 year-olds, for a range of social discount rates that are below the market interest rate. The figure illustrates that, unless the social discount rate is very close to the market interest rate, there are substantial gains from re-

Figure 2: The relative distributional weights of 70-year olds and 20-year olds implied by different deviations of the social discount rate from the market interest rate



The x-axis corresponds to $100 * (r - r^s)$. Relative distributional weights are computed based on Proposition 2. By the first clause of the proposition, relative distributional weights are given by $\mu^{70-20} = \mu^{50}$. By the second clause of the proposition, $\mu \approx 1 + r^s - r$. The relative distributional weights are therefore computed as $(1 + r^s - r)^{50}$.

distributing across age groups. For example, a social discount rate that is 1.5 percentage points below the market interest rate implies that it is socially desirable to reduce the consumption of a 70 year-old by $1 in order to increase the consumption of a 20 year-old by 50 cents.

5 Optimal policy with a low social rate of pure time preference

So far, the analysis has been focused on social marginal rates of substitution along a given (and possibly suboptimal) balanced growth path. Of course, a social planner with a low pure rate of time preference may find it optimal to change the balanced growth path itself, and institute permanent changes in the allocation of consumption across age groups. In this section, I consider a standard neoclassical growth model with overlapping generations, and contrast its equilibrium balanced growth path with the optimal policy of a patient, discounted-utilitarian planner.

5.1 Neoclassical growth model with overlapping generations

The standard neoclassical growth model features an infinitely-lived, representative household. Here, I slightly modify the standard framework to allow for overlapping generations.

The individual's utility function, U, takes the time-separable, constant intertemporal elasticity of substitution form:

$$U(\mathbf{c}^t) = \sum_{a=1}^{T} \psi_a \frac{(c_a^t)^{1-\gamma}}{1-\gamma}$$

The parameter γ is the inverse of the intertemporal elasticity of substitution. It is usually assumed that ψ_a takes the form $\psi_a = \beta^a$ where $\beta \in (0,1)$ is the subjective discount factor. I introduce this more-flexible functional form so that ψ_a may also capture variation in consumption needs over the lifecycle.

Aggregate output is produced according to a standard Cobb-Douglas production function,

$$Y_t = K_t^{\alpha}(A_t L_t)^{1-\alpha}$$

where $L > 0$ is aggregate labor inputs, $K > 0$ is aggregate capital inputs, A is labor-augmenting productivity and $\alpha \in (0,1)$ is the capital intensity of production.

Labor is supplied inelastically; each person supplies $l_a \geq 0$ units of labor at age a. Population grows at a constant rate, n. The size of the cohort born at time t is $(1+n)^t$, and thus labor supply at time t is given by[16]

$$L_t = \sum_{a=1}^{T} l_a (1+n)^{t-a+1} \tag{7}$$

Output is produced by a representative, infinitely-lived firm, who owns the economy's capital stock. The firm's profits are given by

$$\Pi_t = Y_t - w_t L_t - I_t \tag{8}$$

where w_t is the wage rate and I_t is investment at time t. The firm is a price taker, and takes the wage rates and the market interest rates as given. It chooses investment and labor inputs to maximize the net present value of profits.

The firm's optimization problem is

$$\max_{\{K_{t+1}, I_t, L_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \frac{\Pi_t}{\prod_{\tau=0}^{t-1}(1+r_\tau)} \text{ s.t. } K_0 \text{ and} \tag{9}$$

$$K_{t+1} = (1-\delta)K_t + I_t \tag{10}$$

where $\delta \in (0,1)$ is the depreciation rate of capital.

Productivity growth is constant at the rate g:

$$A_t = (1+g)^t A_0$$

Individuals maximize their utilities subject to their budget constraints. The optimization problem of generation t is given by

$$\max_{c^t} U(c^t) \text{ s.t. } \sum_{a=1}^{T} \frac{c_a^t - y_a^t}{\prod_{\tau=t}^{t+a-1}(1+r_\tau)} = 0 \tag{11}$$

where y_a^t is the income that generation t receives at age a. For simplicity, I

[16] To simplify, I assume that the age distribution is already at its steady state level at time 0. This would be the case if, for example, time starts at some $t_0 < -T$ rather than at $t_0 = 0$.

abstract from borrowing constraints. As illustrated in Appendix D, in the presence of occasionally-binding borrowing constraints, the optimal transfers from old to young are larger than the ones suggested by this calibration exercise.

I assume that the shares of labor income and capital income that are paid to people of age a are constant across time, so that y_a^t takes the form

$$y_a^t = \frac{s_{a,l}}{(1+n)^t} w_t L_t + \frac{s_{a,\pi}}{(1+n)^t} \Pi_t \tag{12}$$

where $s_{a,l}, s_{a,\pi} \geq 0$.[17] For example, if there are no intergenerational transfers of labor income, then $s_{a,l}$ is proportional to l_a. In the presence of a social security scheme, $s_{a,l}$ would be lower for younger people and higher for older people. This framework also allows for various assumptions about the ownership structure of the representative firm. For example, it is possible to assume that all people alive own equal shares of the representative firm, or that ownership of the firm is inherited at a certain age.

Finally, the aggregate resource constraint at time t is given by:

$$\sum_{a=1}^{T} c_a^{t+a-1} (1+n)^{t+a-1} + I_t = K_t^{\alpha} (A_t L_t)^{1-\alpha} \tag{13}$$

An equilibrium is defined as $\{\Pi_t, r_t, w_t, L_t, K_t, I_t, Y_t, \pi_t, \{c_a^t, y_a^t\}_{a=1}^T\}_{t=0}^\infty$ which, given K_0, jointly solve the firm's optimization problem (expressions 8 and 9), the individual's optimization problems (expressions 11 and 12), and the aggregate resource constraints (expressions 7 and 13).

In what follows, I restrict attention to an equilibrium balanced growth path, along which output per-capita grows at the constant rate g. The following claim establishes that, in such an equilibrium, each age-specific consumption level, c_a^t, must grow at the rate g as well.

Claim 1. *In any balanced growth path equilibrium in which output per-capita grows at a constant rate g, the consumption levels, $\{c_a\}_{a=1}^T$, must grow at the rate g as well.*

[17] The aggregate resource constraint (equation 13) implies that the shares must add up to 1.

An implication of this claim is that, along a balanced growth path, individual consumption growth over time can be inferred from the cross-sectional distribution of consumption across age groups. This feature will be used for calibrating the parameters $\{\psi_a\}_{a=1}^T$.

5.2 Optimal policy

Consider a social objective of the discounted-utilitarian form:

$$W(\mathbf{c}) = \sum_{t=0}^{\infty} \left(\frac{1}{1+\rho^s}\right)^t U(\mathbf{c}^t)$$

The optimal policy is the solution to the optimization problem:

$$\max_{\mathbf{c},\{K_t\}_{t=1}^{\infty},\{I_t\}_{t=0}^{\infty}} W(\mathbf{c}) \text{ s.t. } K_0 \text{ and equations 10 and 13.}$$

In this optimization problem, the planner can choose how much each person consumes at each age, and how much the economy saves for future generations.

The equilibrium allocation may deviate from the planner's solution for two reasons. First, policymakers may have a different objective from the objective of the discounted-utilitarian planner. For example, current policies may be designed with the objective of appeasing voters, and may thus under-represent the interests of future generations.

Second, policymakers may face various constraints that are not modeled here. If labor supply is endogenous and taxation is distortionary, even a well-intentioned planner may be unable to implement the optimal policy. Such constraints are undoubtedly important; however, the purpose of the current exercise is not to characterize the optimal *feasible* policy. Rather, it is to illustrate the implications of a low social discount rate for the *desirable* reallocation of consumption across age groups – even if such a reallocation may not be feasible in practice.

5.3 Calibration

Table 1 summarizes the model's calibrated parameters. Most of the parameters are standard. The capital intensity parameters, α, was chosen to roughly match the capital income share, and the parameters n and g were chosen to roughly correspond to the average annual growth rates of population and GDP per-capita in the US, respectively.[18] The parameter γ was chosen as $\gamma = 2$, corresponding to an intertemporal elasticity of substitution of $1/2$. The capital depreciation rate, δ, was chosen to roughly match the average depreciation rate of capital in the US.

Table 1: Calibration parameters

	Value	Target
α	0.3	Capital income share
δ (annual)	0.05	Capital depreciation rate
$1/\gamma$	1/ 2	Intertemporal elasticity of substitution
n (annual)	0.05	Population growth rate
g (annual)	0.02	GDP per-capita growth rate
T	80	Maximum age
r (annual)	0.0617	Investment share of GDP
$\{\psi_a\}_{a=1}^{T}$	Figure 4	Consumption-age profile
ρ^s (annual)	0.001	Stern [2007]

[18] In particular, the average rate of population growth is around 1%, and the average rate of GDP growth was 3%. Along the balanced growth path, this implies that g is around 2%.

The only non-standard parameters are the age-specific discount factors, $\{\psi_a\}_{a=1}^T$. Note that the household's optimization problem yields the familiar Euler equations,

$$\psi_a(c_a^t)^{-\gamma} = \psi_{a+1}(c_{a+1}^t)^{-\gamma}(1 + r_{t+a-1}) \tag{14}$$

for every $a = 1, ..., T-1$. By Claim 1, along the balanced growth path,

$$c_{a+1}^t = (1+g)c_{a+1}^{t-1}$$

Thus, the Euler condition can be rewritten as:

$$\psi_a(c_a^t)^{-\gamma} = \psi_{a+1}((1+g)c_{a+1}^{t-1})^{-\gamma}(1+r)$$

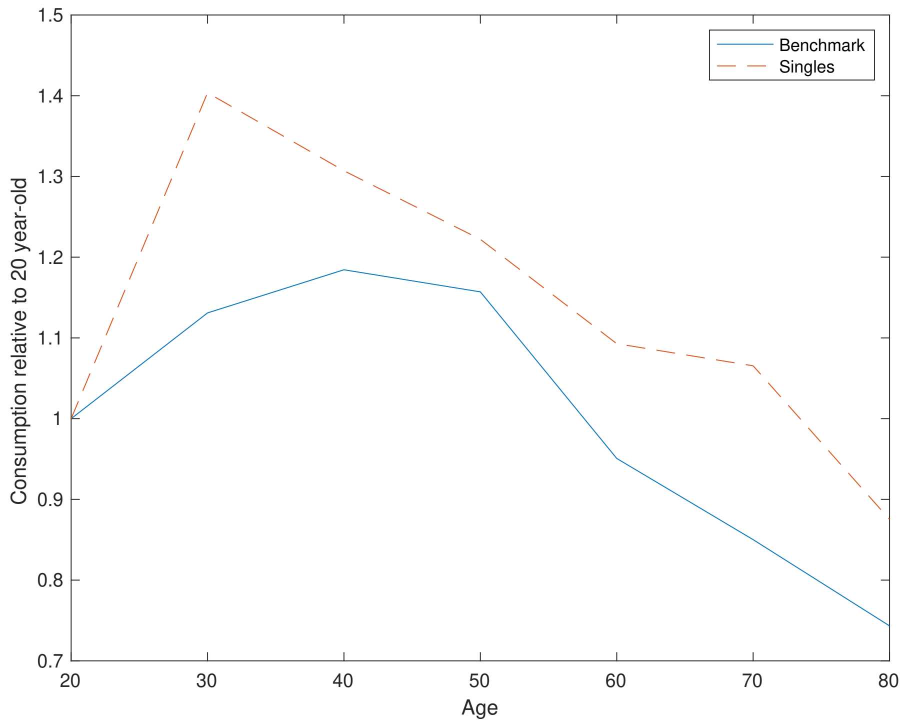
Note that generation t is aged a in the same period that generation $t-1$ is aged $a+1$. The ratios ψ_a/ψ_{a+1} can therefore be calibrated based on the cross-sectional distribution of consumption across age groups in a given period.

I estimate the cross-sectional distribution of consumption using the United States Consumer Expenditure Survey, 2019 (henceforce CEX). The CEX reports consumption expenditure at the household level rather than at the individual level. To create a per-capita measure, I divide household consumption by 2 for all households consisting of a married couple. This measurement counts the consumption of children as part of the consumption of their parents. As it is not obvious that this is the right approach, I also consider an alternative specification in which the parameters $\{\psi_a\}_{a=1}^T$ are calibrated based only on singles. The results are very similar.

I consider a period to be 10 years. I therefore divide the sample into 6 groups, corresponding to people in their 20s, 30s, 40s, 50s, 60s, 70s and 80s. I then estimate relative consumption levels by regressing the log of consumption expenditure per adult on age-group fixed-effects. The results are illustrated in Figure 3.

To estimate $\{\psi_a\}_{a=1}^T$ from the Euler equations, it is necessary to take a stance on the empirical counterpart of the real interest rate, r. I calibrate r

Figure 3: Consumption by age group



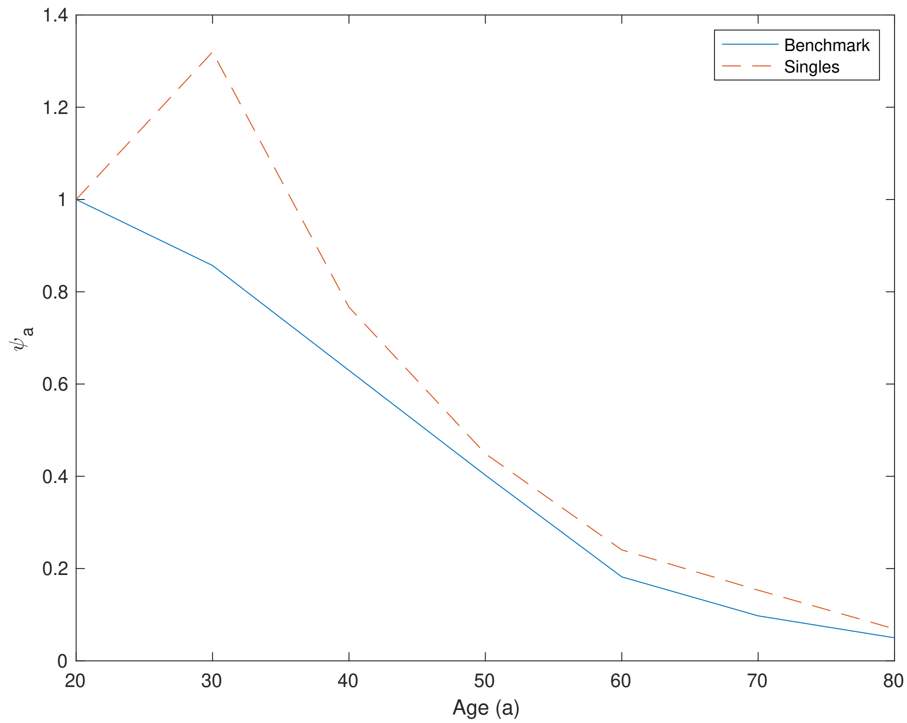
Note: Estimates are based on data from the CEX. Consumption is measured as household expenditure per-adult, where the number of adults in the household is assumed to be 2 for married couples and 1 otherwise. The "singles" specification includes only households that consist of a single person. Averages are computed using the CEX sample weights.

so that the steady-state investment share of GDP matches the average in the United States between 1947-2019, which was 22.4%.[19] The implied interest rate is around 6%, which is similar to the real interest rate used in Nordhaus [2007] and Weitzman [2007]. This interest rate is somewhat lower than the long-run rate of return on stocks but higher than the risk-free rate. The resulting estimates of $\{\psi_a\}_{a=1}^{T}$ are in Figure 4.

The investment-to-output ratio is also used for calibrating the equilibrium capital levels along the balanced growth path. Note that the firm's optimiza-

[19] See https://www.ceicdata.com/en/indicator/united-states/investment–nominal-gdp.

Figure 4: Calibrated values of $\{\psi_a\}_{a=1}^{T}$



tion problem implies the following no-arbitrage condition,

$$r_t = \frac{\partial Y_{t+1}}{\partial K_{t+1}} - \delta = \alpha \frac{Y_{t+1}}{K_{t+1}} - \delta$$

where the second equality follows from the Cobb-Douglas production structure. Given r, A_{t+1} and L_{t+1}, this identity can be used for backing out the capital stock, K_{t+1}, that is consistent with the balanced growth path equilibrium. I set the initial capital stock, K_0, so that the above relationship holds given the normalization $A_0 L_0 = 1$.

5.4 Optimal policy

I consider a social annual rate of pure time preference of 0.1%, as in Stern [2007].[20] This low rate of pure time preference is meant to capture a social objective that cares equally about all generations, and discounts the future only because of a small probability of extinction.

Figure 5 plots the optimal path of the capital stock per effective unit of labor $(K_t/((1+g)(1+n))^t)$, starting from the equilibrium balanced growth path. Given a low rate of pure time preference, the optimal policy converges to a balanced growth path in which capital stocks are over three times larger than their calibrated equilibrium levels.[21]

The main result is that the optimal policy requires substantial reallocation of consumption across age groups. This is illustrated in Figure 6. According to this welfare criterion, it is optimal to decrease aggregate consumption on impact so that the economy can shift resources towards investment in capital. However, this sacrifice is not shared equally across age groups: while it is optimal for people in their 80s to reduce their consumption by over 50%, people in their 20s should actually increase their consumption by 40%.

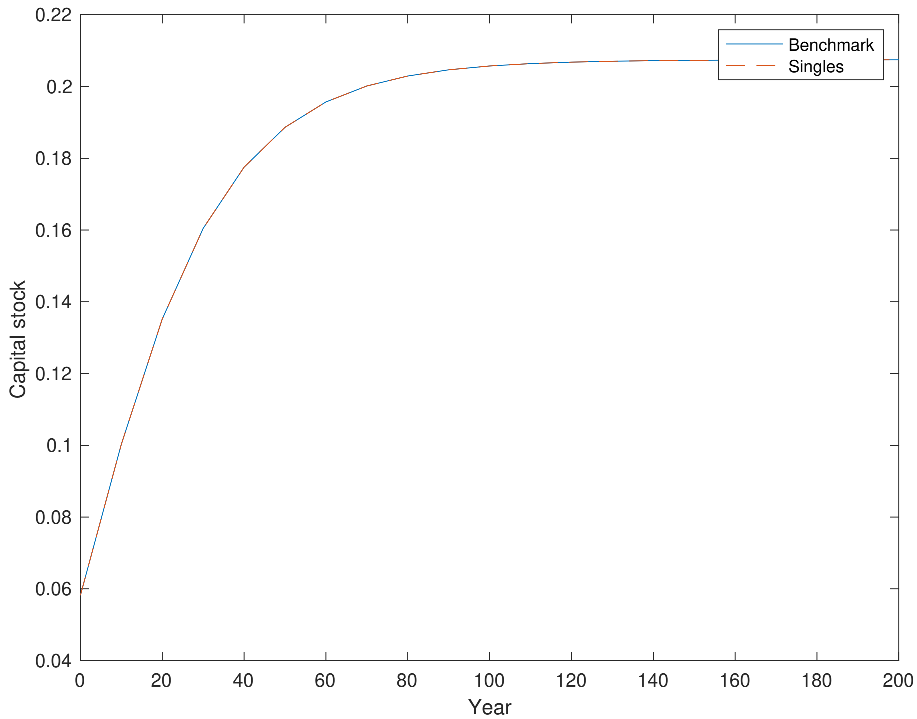
The consumption gains along the balanced growth path are also not shared equally. According to this welfare criterion, younger people should have higher consumption along the balanced growth path, while older people's consumptions should be significantly lower. For example, along the optimal balanced growth path, the consumption of 20 year-olds is 120% higher than it is under the current equilibrium, while the consumption of 80 year-olds is 40% lower.

This calibration illustrates that a discounted-utilitarian objective with a low rate of pure time preference supports massive transfers from older people to younger ones. The reason goes back to the insight of Proposition 1: in an overlapping generation model, caring more about the future means caring more about the young.

[20] Note that, because the length of the period in the simulation is 10 years, this corresponds to $\rho^s = 0.01$.

[21] The implication of a low social discount rate for the social desirability of massive capital accumulation is also emphasized in Barrage [2018].

Figure 5: Optimal capital stocks per effective unit of labor



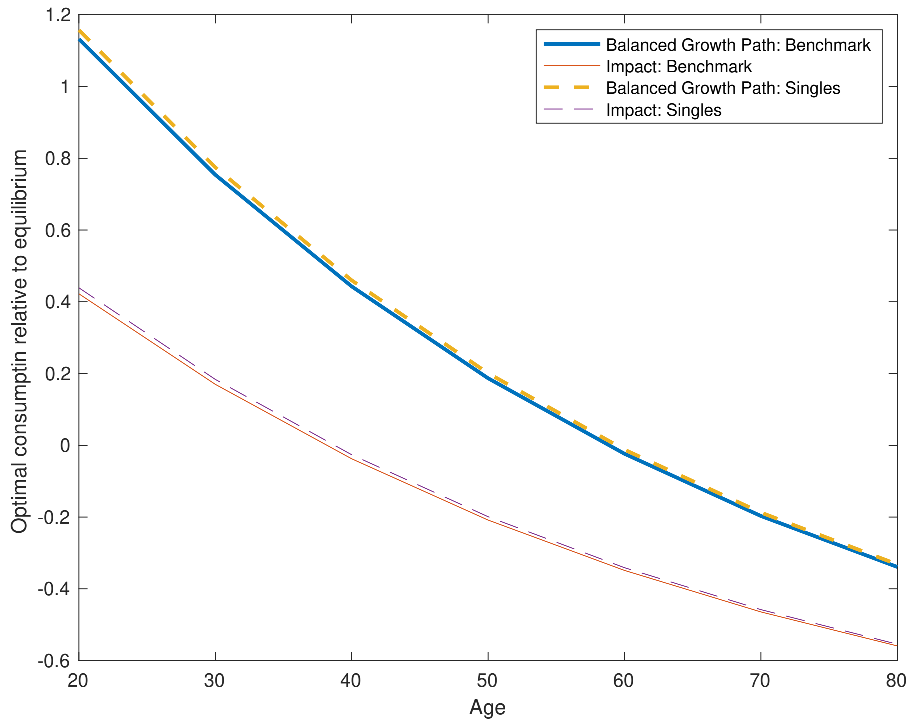
Note: The capital stock per effective unit of labor is defined as the ratio $K_t/((1+g)(1+n))^t$. It is constant along a balanced growth path.

6 Conclusion

This paper establishes an equivalence between two normative questions. The first is, how should policy discount future returns? This question is relevant for evaluating the optimal scope of public investment. In a discounted-utilitarian framework, the answer to this question depends on the social rate of pure time preference – the rate at which society should discount the utilities of future generations.

The second question is, how should society distribute resources across people of different age groups? This question became particularly contentious during the COVID-19 pandemic (see, for example, Hall et al. [2020]). Contain-

Figure 6: Optimal consumption relative to equilibrium consumption



ment measures disproportionately benefitted the elderly, who were at higher risk from the virus. However, some of the costs were born by children and working-age adults, who suffered serious disruptions. This raised the question of how to tradeoff benefits to the elderly with costs to younger people. This question also comes up during normal times, when policymakers face budgetary tradeoffs between programs that benefit the elderly (such as social security and medicaid) and programs that benefit younger people (such as childcare subsidies and playgrounds).

This paper establishes that the social discount rate is lower than the market interest rate if and only if it is socially desirable to increase the consumption of the young at the expense of the elderly. For example, the discounted-utilitarian objective in Stern [2008] implies a low social discount rate, because it features a low rate of pure time preference. Consequently, it also implies the desirability of massive transfers from older people to younger ones.

There are several possibilities forward. First, we can maintain that the social discount rate is lower than the market interest rate, and accept that it is socially desirable to transfer resources from the current old to the current young. Quantitatively, even a small deviation of the social discount rate from the market interest rate implies large welfare gains from redistribution across age groups.

A second possibility is to maintain that the current distribution of resources across age groups is close to optimal, and accept that the social discount rate is close to the market interest rate. Within the discounted-utilitarian framework, a social discount rate that is close to the market interest rate requires discounting the lifetime utilities of future generations – a conclusion which Ramsey [1928] and many others find ethically unacceptable, but which others defend.[22]

A third possibility is to forgo the assumption that the social objective takes a discounted-utilitarian form. For example, a prioritarian objective that is averse to inequality in lifetime utilities may imply a high social discount rate along the balanced growth path, even if it cares equally about all generations.

Finally, it is worth cautioning that the quantitative relationship between social discounting and the desirability of age-based redistribution changes as we move away from the balanced growth path. A growth slowdown, as suggested by Jones [2021], may imply long-run social discount rates that are below the market interest rate, even if there are no welfare gains from current redistribution between old and young. Furthermore, as pointed out by Stern [2008], the social discount rate is a useful statistic only for evaluating the social desirability of marginal inter-temporal changes along a given growth trajectory. In some instances, current saving may affect the growth path itself; for example, uncontrolled climate change may set humanity off on a different growth path. In this case, social discount rates along the previous balanced-growth path are of limited practical use.

[22] See, for example, Lloyd [2021].

References

Matthew D. Adler. *Measuring Social Welfare: an Introduction*. Oxford University Press, 2019.

Kenneth Arrow, Maureen Cropper, Christian Gollier, Ben Groom, Geoffrey Heal, Richard Newell, and William Nordhaus et al. Determining benefits and costs for future generations. *Science*, 341(6144):349–350, 2013.

Geir B. Asheim and Stéphane Zuber. A complete and strongly anonymous leximin relation on infinite streams. *Social Choice and Welfare*, 41(4):819–834, 2013.

Lint Barrage. Be careful what you calibrate for: social discounting in general equilibrium. *Journal of Public Economics*, 160:33–49, 2018.

William J. Baumol. On the social rate of discount. *American Economic Review*, 58(4):788–802, 1968.

Guillermo A. Calvo and Maurice Obstfeld. Optimal time-consistent fiscal policy with finite lifetimes. *Econometrica*, 56(2), March 1988.

Andrew Caplin and John Leahy. The social discount rate. *Journal of political economy*, 112(6):1257–1268, 2004.

Moritz A. Drupp, Mark C. Freeman, Ben Groom, and Frikk Nesje. Discounting disentangled. *American Economic Journal: Economic Policy*, 10(4):109–134, 2018.

Emmanuel Farhi and Ivan Werning. Inequality and social discounting. *Journal of political economy*, 115(3):365–402, 2007.

Marc Fleurbaey and Stéphane Zuber. Discounting, risk and inequality: A general approach. *Journal of Public Economics*, 128:34–49, 2015.

Christian Gollier and James K. Hammitt. The long-run discount rate controversy. *Annual Review of Resource Economics*, 6:273–295, 2014.

Hilary Greaves. Discounting for public policy: A survey. *Economics & Philosophy*, 33(3):391–439, 2017.

Robert E. Hall, Charles I. Jones, and Peter J. Klenow. Trading off consumption and covid-19 deaths. *Federal Reserve Bank of Minneapolis Quarterly Review*, 42(1):1–14, 2020.

Berthold Herrendorf, Richard Rogerson, and Akos Valentinyi. Growth and the kaldor facts. *Federal Reserve Bank of St. Louis Review*, 101(4):259–276, 2019.

Charles I. Jones. The past and future of economic growth: A semi-endogenous perspective. *Annual Review of Economics*, 2021. Forthcoming.

Nicholas Kaldor. Capital accumulation and economic growth. In Friedrich A. Lutz and D. C. Hague, editors, *The Theory of Capital*. St. Martin's Press, 1961.

J. Paul Kelleher. Pure time preference in intertemporal welfare economics. *Economics & Philosophy*, 33(3):441–473, 2017.

Harry R. Lloyd. Time discounting, consistency, and special obligations: a defence of robust temporalism. Global Priorities Institute, working paper no. 11-2021, 2021.

Antony Millner and Geoffrey Heal. Choosing the future: Markets, ethics, and rapprochement in social discounting. *Journal of Economic Literature*, 2022. Forthcoming.

William D. Nordhaus. A review of the stern review on the economics of climate change. *Journal of economic literature*, 45(3):686–702, 2007.

John Quiggin. Equity between overlapping generations. *Journal of Public Economic Theory*, 14(2):273–283, 2012.

Frank Plumpton Ramsey. A mathematical theory of saving. *The Economic Journal*, 38(152):543–559, 1928.

John Rawls. Some reasons for the maximin criterion. *The American Economic Review*, 64(2):141–146, 1974.

Maik T. Schneider, Christian P. Traeger, and Ralph Winkler. Trading off generations: Equity, discounting, and climate change. *European Economic Review*, 56(8):1621–1644, 2012.

Nicholas Stern. *The economics of climate change: the Stern review*. Cambridge University press, 2007.

Nicholas Stern. The economics of climate change. *American Economic Review*, 98(2):1–37, 2008.

Martin L. Weitzman. A review of the stern review on the economics of climate change. *Journal of economic literature*, 45(3):703–724, 2007.

Stéphane Zuber and Geir B. Asheim. Justifying social discounting: the rank-discounted utilitarian approach. *Journal of Economic Theory*, 147(4):1572–1601, 2012.

A Proof of Proposition 1

$$MRS((a,t),(a',t')) = \frac{\frac{\partial W(\mathbf{c})}{\partial c_a^t}}{\frac{\partial W(\mathbf{c})}{\partial c_{a'}^{t'}}} = \frac{\frac{\partial W(\mathbf{c})}{\partial c_a^t}}{\frac{\partial W(\mathbf{c})}{\partial c_1^{t+a-1}}} \frac{\frac{\partial W(\mathbf{c})}{\partial c_1^{t+a-1}}}{\frac{\partial W(\mathbf{c})}{\partial c_2^{t+a-1}}} \frac{\frac{\partial W(\mathbf{c})}{\partial c_2^{t+a-1}}}{\frac{\partial W(\mathbf{c})}{\partial c_1^{t+a}}} \cdots \frac{\frac{\partial W(\mathbf{c})}{\partial c_1^{t'-a'-2}}}{\frac{\partial W(\mathbf{c})}{\partial c_2^{t'-a'-2}}} \frac{\frac{\partial W(\mathbf{c})}{\partial c_2^{t'-a'-2}}}{\frac{\partial W(\mathbf{c})}{\partial c_{a'}^{t'}}} =$$

$$\mu_{a,1}^{t+a-1}(1+r_{1,2,t+a-1}^{si})\mu_{2,1}^{t+a} \cdots (1+r_{1,2,t'+a'-2}^{si})\mu_{2,a'}^{t'+a'-1}$$

Rearranging the terms yields the desired expression.

B Proof of Proposition 2

The derivative of equation 6 with respect to c_a^t is

$$\frac{\partial W(\mathbf{c})}{\partial c_a^t} = \left(\frac{1}{1+\rho^s}\right)^t \phi'(U(\mathbf{c}^t)) \frac{\partial U(\mathbf{c}^t)}{\partial c_a^t} \tag{15}$$

Along the balanced growth path, $c_a^t = (1+g)^t c_a^0$. As U is homogeneous, there exists some η such that $U(\mathbf{c}^t) = U((1+g)^t \mathbf{c}^0) = (1+g)^{\eta t} U(\mathbf{c}^0)$. For this η, it holds that

$$\frac{\partial U(\mathbf{c}^t)}{\partial c_a^t} = (1+g)^{(\eta-1)t} \frac{\partial U(\mathbf{c}^0)}{\partial c_a^0} \tag{16}$$

Because ϕ is homogeneous, so is ϕ', and hence there exists ζ such that $\phi'(U(\mathbf{c}^t)) = \phi'((1+g)^{t\eta} U(\mathbf{c}^0)) = (1+g)^{t\eta\zeta} \phi'(U(\mathbf{c}^0))$. Substituting yields

$$\frac{\partial W(\mathbf{c})}{\partial c_a^t} = \left(\frac{1}{1+\rho^s}\right)^t (1+g)^{t\eta\zeta} \phi'(U(\mathbf{c}^0))(1+g)^{(\eta-1)t} \frac{\partial U(\mathbf{c}^0)}{\partial c_a^0} \tag{17}$$

By the Euler equation (for generation 0),

$$\frac{\partial U(\mathbf{c}^0)}{\partial c_a^0}(1+r)^{a-1} = \frac{\partial U(\mathbf{c}^0)}{\partial c_1^0} \tag{18}$$

Substituting yields

$$\frac{\partial W(\mathbf{c})}{\partial c_a^t} = \left(\frac{1}{1+\rho^s}\right)^t (1+g)^{t\eta\zeta} \phi'(U(\mathbf{c}^0))(1+g)^{(\eta-1)t} \left(\frac{1}{1+r}\right)^{a-1} \frac{\partial U(\mathbf{c}^0)}{\partial c_1^0} \tag{19}$$

Generation t is aged a in period $\tau = t + a - 1$. Substituting yields

$$\frac{\partial W(\mathbf{c})}{\partial c_a^t} = \left(\frac{1}{1+\rho^s}\right)^{\tau-a+1} (1+g)^{(\tau-a+1)\eta\zeta} \phi'(U(\mathbf{c}^0))(1+g)^{(\eta-1)(\tau-a+1)} \left(\frac{1}{1+r}\right)^{a-1} \frac{\partial U(\mathbf{c}^0)}{\partial c_1^0}$$

In this period, the people aged $\tilde{a}$ are from generation $\tilde{t}$, where $\tilde{t} = \tau - \tilde{a} + 1$. Substituting into the above expression implies that

$$\frac{\partial W(\mathbf{c})}{\partial c_{\tilde{a}}^{\tilde{t}}} = \left(\frac{1}{1+\rho^s}\right)^{\tau-\tilde{a}+1} (1+g)^{(\tau-\tilde{a}+1)\eta\zeta} \phi'(U(\mathbf{c}^0))(1+g)^{(\eta-1)(\tau-a+1)} \left(\frac{1}{1+r}\right)^{\tilde{a}-1} \frac{\partial U(\mathbf{c}^0)}{\partial c_1^0}$$

and hence

$$\mu_{a,\tilde{a}}^t = \frac{\frac{\partial W(\mathbf{c})}{\partial c_a^t}}{\frac{\partial W(\mathbf{c})}{\partial c_{\tilde{a}}^{\tilde{t}}}} = \left(\frac{1}{1+\rho^s}\right)^{\tilde{a}-a} (1+g)^{(\tilde{a}-a)\eta\zeta} (1+g)^{(\eta-1)(\tilde{a}-a)} \left(\frac{1}{1+r}\right)^{a-\tilde{a}} \tag{20}$$

$$= \left(\left(\left(\frac{1+r}{1+\rho^s} \right) (1+g)^{\eta-1+\eta\zeta} \right)^{-1} \right)^{a-\tilde{a}}$$

setting

$$\mu = \left(\left(\frac{1+r}{1+\rho^s} \right) (1+g)^{\eta-1+\eta\zeta} \right)^{-1}$$

concludes the proof of the first part of the proposition.

To prove the second part of the proposition, note that, by equation 19,

$$\frac{\frac{\partial W(\mathbf{c})}{\partial c_a^t}}{\frac{\partial W(\mathbf{c})}{\partial c_a^{t'}}} = \frac{\left(\frac{1}{1+\rho^s}\right)^t (1+g)^{t\eta\zeta}(1+g)^{(\eta-1)t}}{\left(\frac{1}{1+\rho^s}\right)^{t'} (1+g)^{t'\eta\zeta}(1+g)^{(\eta-1)t'}} = \frac{(\mu(1+r))^{-t}}{(\mu(1+r))^{-t'}} = (\mu(1+r))^{t'-t} \tag{21}$$

concluding the proof.

C Proof of Claim 1

It is useful to normalize $A_0 L_0 = 1$. Note that, to satisfy the aggregate production function, if output per-capita grows at the rate g, then the per-capita capital stock must grow at the rate g as well. Thus, the equilibrium balanced growth path features some $y = Y_t/((1+g)(1+n))^t$, $i = I_t/((1+g)(1+n))^t$, $k = K_t/((1+g)(1+n))^t$, r and $\pi = \Pi_t/((1+g)(1+n))^t$ such that:

$$y = k^\alpha$$

$$(1+g)(1+n)k = (1-\delta)k + i$$

$$\pi = \alpha y - i$$

$$r = \alpha \frac{y}{k} - \delta$$

The first expression is the aggregate production function, and the second expression is the capital accumulation equation (note that, along the balanced growth path, $K_{t+1} = (1+g)(1+n)K_t$).

The third expression is the firm's profits. As, in equilibrium, $(1-\alpha)Y = w_t L_t$, the firm's profits are $\alpha Y - I$. The fourth expression is the no-arbitrage condition, which can be obtained by taking the firm's first-order condition with respect to investment, and using the Cobb-Douglas production structure (so that $MPK = \alpha Y/K$).

This system implies that the interest rate, r is constant over time and that π and y are constant over time as well.

Note that the firm's first order condition implies that labor income is a constant share of output:

$$w_t L_t = (1-\alpha)Y_t$$

It thus follows that the intertemporal budget constraint of generation t can be written as:

$$\sum_{a=1}^{T} \frac{c_a^t}{(1+r)^a} = \sum_{a=1}^{T} \frac{(1+g)^{t+a-1}(1+n)^a \left(s_{a,l}(1-\alpha)y + s_{a,\pi}\pi\right)}{(1+r)^a}$$

$$= (1+g)^{t-1} \sum_{a=1}^{T} \frac{(1+g)^a(1+n)^a \left(s_{a,l}(1-\alpha)y + s_{a,\pi}\pi\right)}{(1+r)^a}$$

It follows that the net-present value of income grows that the rate g. Because preferences are homothetic and the interest rate is constant across time, the optimal consumption plan of generation t is proportional to the optimal consumption plan of generation 0. It follows that $c_a^t = (1+g)^t c_a^0$.

D Borrowing constraints

To understand how the presence of borrowing constraints changes the relationship between the social discount rate and the cross-sectional distributional weights along the balanced growth path, I modify the model in section 4 as follows. Assume that people can buy and sell bonds (b_a^t) at the market interest rate, r, but that borrowing constraints are occasionally binding. The consumption sequence of generation t solves the following constrained optimization problem:

$$\mathbf{c}^t = \arg \max_{\mathbf{c}^t, \{b_a^t\}_{a=1}^T} U(\mathbf{c}^t) \text{ s.t. } c_a^t + b_a^t = y_a^t + (1+r)b_{a-1}^t \text{ and } b_a^t \geq \kappa_a^t \tag{22}$$

where $b_0^t = 0$ and $\{\kappa_a^t\}_{a=1}^T$ are the age-specific borrowing limits (it makes sense to assume that $\kappa_T^t = 0$, so that individuals cannot borrow in their last period of life). I assume that borrowing limits grow at the rate of gdp-per capita, so that $b_a^t = (1+g)^t b_a^0$. This will hold, for example, whenever borrowing limits are proportional to people's current or expected incomes (which are assumed to increase at the rate g).

I establish the following Proposition, which is analogous to Proposition 2 in the presence of borrowing constraints:

Proposition 3. *Along the balanced growth path, there exists $\mu > 0$ such that*

1. *for every t and $a' > a$, it holds that $\mu_{a',a}^t \leq \mu^{a'-a}$, and*
2. *For every age, a, and time periods, $t' > t$, it holds that $r_{a,t,a,t'}^s = r^s$, where*

$$1 + r^s = (1+r)\mu$$

The only difference between this result and Proposition 2 is in the first clause. In Proposition 2, it holds that $\mu_{a',a}^t = \mu^{a'-a}$, whereas, here, it is required only that $\mu_{a',a}^t \leq \mu^{a'-a}$ (when $a' > a$). This means that the relative distributional weights depicted in Figure 2 are upper-bounds. For example, if the market interest rate is 6% and the social discount rate is 1.5%, then, by the second clause, it follows that $\mu \approx 0.955$. By the first clause, society must have a weak preference for reducing the consumption of a 70 year-old by \$1 in order to increase the consumption of a 20 year-old by 10 cents. The social preference for this transfer may be strict if borrowing constraints are binding.

Proof. The proof of this modified proposition closely follows the steps of the proof of Proposition 2. Note that the first order condition with respect to b_a^t yields:

$$\frac{\partial U(\mathbf{c}^t)}{\partial c_a^t} = (1+r)\frac{\partial U(\mathbf{c}^t)}{\partial c_{a+1}^t} + \lambda_a^t$$

where $\lambda_a^t \geq 0$ is the Lagrange multiplier on the borrowing constraint. Dividing through by $\frac{\partial U(\mathbf{c}^t)}{\partial c_{a+1}^t}$ yields

$$\frac{\frac{\partial U(\mathbf{c}^t)}{\partial c_a^t}}{\frac{\partial U(\mathbf{c}^t)}{\partial c_{a+1}^t}} = 1 + r + \frac{\lambda_a^t}{\frac{\partial U(\mathbf{c}^t)}{\partial c_{a+1}^t}}$$

Note that, if $\{\tilde{c}_a^0\}_{a=1}^T$ is feasible for generation 0, then $\{(1+g)^t \tilde{c}_a^0\}_{a=1}^T$ is feasible for generation t (given the assumptions that y_a^t and κ_a^t grow at the rate g). Thus, given that preferences are homothetic, if $\{c_a^0\}_{a=1}^T$ is the optimal feasible consumption sequence for generation 0, then $\{(1+g)^t c_a^0\}_{a=1}^T$ is the optimal feasible consumption sequence of generation t. It follows that the ratio on the left-hand-side is constant across time, and hence $\lambda_a^t / \frac{\partial U(\mathbf{c}^t)}{\partial c_{a+1}^t}$ is constant across time.

Denote

$$r_a = r + \frac{\lambda_a^t}{\frac{\partial U(\mathbf{c}^t)}{\partial c_{a+1}^t}}$$

In the presence of borrowing constraints, the Euler condition in equation 18 becomes

$$\frac{\partial U(\mathbf{c}^0)}{\partial c_a^0} \prod_{\hat{a}=1}^{a-1} (1 + r_a) = \frac{\partial U(\mathbf{c}^0)}{\partial c_1^0} \tag{23}$$

Following the steps of the proof of Proposition 2 and replacing $(1+r)^{a-1}$ with the above identity implies that, for every t and $a \geq \tilde{a}$,

$$\mu_{a,\tilde{a}}^t = \frac{\frac{\partial W(\mathbf{c})}{\partial c_a^t}}{\frac{\partial W(\mathbf{c})}{\partial c_{\tilde{a}}^t}} = \left(\frac{1}{1+\rho^s}\right)^{\tilde{a}-a} (1+g)^{(\tilde{a}-a)\eta\zeta} (1+g)^{(\eta-1)(\tilde{a}-a)} \frac{1}{\prod_{\hat{a}=\tilde{a}}^{a-1} (1 + r_{\hat{a}})} \tag{24}$$

$$\leq \left(\left(\left(\frac{1+r}{1+\rho^s} \right) (1+g)^{\eta-1+\eta\zeta} \right)^{-1} \right)^{a-\tilde{a}}$$

where the inequality follows from the fact that $r_a \geq r$.

As in the proof of proposition 2, let

$$\mu = \left(\left(\frac{1+r}{1+\rho^s}\right)(1+g)^{\eta-1+\eta\zeta}\right)^{-1}$$

it follows that, for every t and $a \geq \tilde{a}$,

$$\mu^t_{a,\tilde{a}} \leq \mu^{a-\tilde{a}}$$

concluding the proof of the first clause of the proposition.

Note that, given this definition of μ, the identities in equation 21 continue to hold, thus establishing the second clause of the proposition. $\square$

Implications for section 5. The calibration in section 5 may also be affected by the presence of binding borrowing constraints. Note that the assumption that borrowing constraints are not binding is used only in the calibration of the parameters $\{\psi_a\}_{a=1}^T$ (the interest rate, r, is calibrated to match the steady state investment share of GDP, which is an equilibrium condition that does not rely on the Euler equation).

If, instead, there are binding borrowing constraints, then the Euler conditions along the balanced growth path are given by

$$\psi_a c_a^{-\gamma} = \psi_{a+1}((1+g)c_{a+1})^{-\gamma}(1+r_a) \Rightarrow \frac{\psi_{a+1}}{\psi_a} = \left(\frac{(1+g)c_{a+1}}{c_a}\right)^{\gamma}\left(\frac{1}{1+r_a}\right)$$

where $r_a \geq r$. The calibrated parameters, $\{\hat{\psi}_a\}_{a=1}^T$, satisfy

$$\frac{\hat{\psi}_{a+1}}{\hat{\psi}_a} = \left(\frac{(1+g)c_{a+1}}{c_a}\right)^{\gamma}\left(\frac{1}{1+r}\right) = \left(\frac{\psi_{a+1}}{\psi_a}\right)\left(\frac{1+r_a}{1+r}\right)$$

As $r_a \geq r$, it follows that

$$\frac{\hat{\psi}_{a+1}}{\hat{\psi}_a} \geq \frac{\psi_{a+1}}{\psi_a}$$

Note that the discounted-utilitarian objective considered in section 5 can be written as

$$\sum_{t=0}^{\infty}\left(\frac{1}{1+\rho^s}\right)^t\sum_{a=1}^{T}\psi_a\frac{(c_a^t)^{1-\gamma}}{1-\gamma}$$

Optimal cross-sectional allocation of consumption in each period implies that

$$\psi_a(c_a^t)^{-\gamma}=\left(\frac{1}{1+\rho^s}\right)\psi_{a+1}(c_{a+1}^{t-1})^{-\gamma}\Rightarrow\frac{c_{a+1}^{t-1}}{c_a^t}=\left(\left(\frac{1}{1+\rho^s}\right)\frac{\psi_{a+1}}{\psi_a}\right)^{\frac{1}{\gamma}}$$

The calibrated optimal consumption levels, $\hat{c}_a^t$ and $\hat{c}_{a+1}^{t-1}$, therefore satisfy

$$\frac{\hat{c}_{a+1}^{t-1}}{\hat{c}_a^t}=\left(\left(\frac{1}{1+\rho^s}\right)\frac{\hat{\psi}_{a+1}}{\hat{\psi}_a}\right)^{\frac{1}{\gamma}}\geq\left(\left(\frac{1}{1+\rho^s}\right)\frac{\psi_{a+1}}{\psi_a}\right)^{\frac{1}{\gamma}}=\frac{c_{a+1}^{t-1}}{c_a^t}$$

Consequently, if there are borrowing binding constraints, then the calibration *understates* the optimal transfers from older people to younger people.