Intergenerational Insurance*

Francesco Lancia[1], Alessia Russo[2], and Tim Worrall[3]

[1] *Ca' Foscari University of Venice and the Centre for Economic Policy Research*

[2] *University of Padua and the Centre for Economic Policy Research*

[3] *University of Edinburgh*

[Original Version, December 2020, Revised, December 2021, June 2023 and December 2023]

Abstract

How should successive generations insure each other when the young can default on previously promised transfers to the old? This paper studies intergenerational insurance that maximizes the expected discounted utility of all generations subject to participation constraints for each generation. If complete insurance is unattainable, the optimal intergenerational insurance is history-dependent even when the environment is stationary. The risk from a generational shock is spread into the future, with periodic *resetting*. Interpreting intergenerational insurance in terms of public debt, the fiscal reaction function is nonlinear and the risk premium on debt is lower than the risk premium with complete insurance.

Keywords: Intergenerational insurance; limited commitment; risk sharing; stochastic overlapping generations; sustainable debt.

JEL CODES: D64; E21; H55.

* Original version circulated with the title "Optimal Sustainable Intergenerational Insurance". We thank the editor and two anonymous referees along with Spiros Bougheas, Francesco Caselli, Gabrielle Demange, Martín Gonzalez-Eiras, Sergey Foss, Alexander Karaivanov, Paul Klein, Dirk Krueger, Sarolta Laczó, Costas Milas, Espen Moen, Iacopo Morchio, Nicola Pavoni, José-Víctor Ríos-Rull, Karl Schlag, Kjetil Storesletten and Aleh Tsyvinski for helpful comments. The paper has also benefited from the comments of seminar participants at Cologne, the London School of Economics, New York University Abu Dhabi, Oslo and Warwick in addition to presentations at the NBER Summer Institute on Macro Public Finance, the SED Meeting in Edinburgh, the SAET Conference in Faro, the CSEF-IGIER Symposium on Economics and Institutions at Anacapri, the CEPR European Summer Symposium in International Macroeconomics at Tarragona, the EEA-ESEM Congress in Manchester, the Barcelona Graduate School of Economics Summer Forum and the Vienna Macroeconomics Workshop. Sergio Cappellini provided valuable research assistance. The first and second authors gratefully acknowledge the support of STARS@Unipd grant 482125 and the third author gratefully acknowledges the support of the UK Economic and Social Research Council grant ES/L009633/1.

Introduction

Countries face economic shocks that result in unequal exposure to risk across generations. The Financial Crisis of 2008 and the Covid-19 pandemic are two recent and notable examples.[1] Faced with such shocks, it is desirable to share risk across generations. However, full risk sharing is not sustainable if it commits future generations to transfers they would not wish to make once they are born. The issue of the sustainability of intergenerational insurance is becoming increasingly relevant in many advanced economies as the relative standard of living of the younger generation has worsened in recent decades.[2] If this generational shift persists, future generations may be less willing to contribute to insurance arrangements than in the past. Therefore, a natural question to ask is how an optimal intergenerational insurance arrangement should be structured when there is limited enforcement of risk-sharing transfers.

Despite its policy relevance, the literature on intergenerational insurance does not fully address this question. The normative approach in the literature investigates the optimal design of intergenerational insurance but assumes that transfers are mandatory, ignoring the issue of limited enforcement. Meanwhile, the positive approach highlights the political limits to intergenerational insurance while considering equilibrium allocations supported by a particular voting mechanism, which are not necessarily Pareto optimal.

In this paper, we examine optimal intergenerational insurance when subsequent generations can default on risk-sharing transfers promised to previous generations. We model the limited enforcement of transfers by assuming that transfers satisfy a participation constraint for each generation. This can be interpreted as requiring that the insurance arrangement be supported by each generation if put to a vote. An arrangement of risk-sharing transfers is *sustainable* if it satisfies the participation constraint of every generation. *Optimal* sustainable intergenerational insurance is determined by a benevolent social planner who chooses transfers to maximize the expected discounted utility of all generations subject to the participation constraints.

The model is simple. At each date, a new generation is born and lives for two periods. Each generation comprises a constant population of homogeneous agents with the population size normalized to one. Each agent receives an endowment of a single, nonstorable consumption good, both when young and old. Endowments are stochastic. Each generation is affected by an idiosyncratic shock (common to all agents within a generation) and an aggregate growth shock.

[1] Glover et al. (2020) find that the Financial Crisis of 2008 had a negative impact on the older generation, while the young benefited from the fall in asset prices. Glover et al. (2023) find that younger workers have been impacted to a greater extent by the response to the Covid-19 pandemic because they disproportionately work in sectors that have been particularly adversely affected, such as retail and hospitality.

[2] Part A of the Supplementary Appendix reports changes in the relative standard of living of the young and the old for six OECD countries using data from the Luxembourg Income Study Database.

We adopt the approach of Alvarez and Jermann (2001) and Krueger and Lustig (2010) and assume that preferences exhibit a constant coefficient of relative risk aversion (for simplicity, we concentrate on the case of logarithmic preferences) and that the idiosyncratic and growth shocks are independent and identically distributed. In this setting, the underlying economy is stationary. There are only two frictions. First, risk may not be allocated efficiently, even if the economy is dynamically efficient, because there is no market in which the young can share risk with previous generations (see, for example, Diamond, 1977). Second, the amount of risk that can be shared is limited because transfers between generations cannot be enforced. In particular, the old will not make a transfer to the young (since the old have no future). Conversely, the young may make a transfer to the old. However, the young will only do so if they receive promises for their old age that at least match their expected lifetime utility from autarky, and they anticipate that these promises will be honored by the next generation.

It is well known (see, for example, Aiyagari and Peled, 1991) that if endowments are such that the young wish to defer consumption to old age at a zero net interest rate, then there are stationary transfers that improve upon autarky (Proposition 2). Under this condition, and assuming that the first-best transfers cannot be sustained, there is a trade-off between efficiency and providing incentives for the young to make transfers to the old. This trade-off is resolved by linking the utility the young are promised for their old age to the promise made to the young of the previous generation. The resulting optimal sustainable intergenerational insurance arrangement is history dependent, even though the economic environment is stationary.

To understand why there is history dependence, suppose that the first-best transfers would violate the participation constraint of the young in some endowment state. To ensure that the current transfer made by the young is voluntary, either the current transfer is reduced below the first-best level, or the promised transfers for their old age are increased. Both changes are costly since a lower current transfer reduces the amount of risk shared today while increasing the transfers promised to the current young for their old age tightens the participation constraints of the next generation and reduces the risk that can be shared tomorrow. Therefore, an optimal trade-off exists between reducing the current transfer and increasing the future promise. This trade-off depends both on the current endowment and the current promise. For example, consider some current endowment and a current promise such that the future promise for the same endowment state is higher than the current promise. If the same endowment state is repeated in the subsequent period, then the young in that period are called upon to make a larger transfer, which in turn requires a higher promise of future utility to them as well. Thus, the transfer depends not only on the current endowment but also on the past promise, and hence, the history of endowment shocks.

The optimal sustainable intergenerational insurance is found by solving a functional equation derived from the planner's maximization problem. The solution is characterized by policy functions for the consumption of the young (or equivalently, the transfer made to the old) and the future promised utility for their old age in each endowment state. Both policy functions depend on the current endowment and the current promise. For a given endowment, the consumption of the young is weakly decreasing in the current promise, while the future promise is weakly increasing in the current promise (Lemmas 2 and 3). When the current endowment state is repeated, the policy function for the future promise has a unique fixed point, which (ignoring a boundary condition) equals the utility at the first-best outcome. Therefore, the future promise is higher than the current promise when it is less than the corresponding fixed point and lower than the current promise when the current promise is greater than the fixed point. When the promised utility is sufficiently low, there is some endowment state in which the participation constraint of the young does not bind. In that case, the future promise is *reset* to the largest value that maximizes the planner's payoff.

When a participation constraint binds, the risk affecting one generation is spread to future generations. The resetting property shows, however, that the effect of a shock does not last forever. Moreover, it implies strong convergence to a unique invariant distribution (Proposition 5). The invariant distribution exhibits history dependence, and consumption fluctuates across states and over time, even in the long run. This starkly contrasts to the situation under either full enforcement of transfers or no risk. In the former case, the promised utility is constant over time, except possibly in the initial period (Proposition 3). In the latter case, the promised utility is constant in the long run, although there may be an initial phase during which the promised utility falls (Proposition 4). In both cases, the allocation is efficient in the long run. Thus, *both* risk and limited enforcement are necessary for history dependence and inefficiency in the long run.

Transfers to the old can be interpreted in terms of debt. This interpretation allows us to address the dynamics of debt together with the issue of debt valuation and sustainability following the model-based approach introduced by Bohn (1995, 1998). With this interpretation, the planner issues one-period state-contingent bonds at the state price determined by the corresponding intertemporal marginal rate of substitution and balances the budget by taxing or subsidizing the young. Given the bond prices and taxes, the young buy the correct quantity of state-contingent bonds to finance their optimal old-age consumption. It is natural to measure debt relative to the endowment of the young when preferences are logarithmic. With debt measured in this way, there is a maximal debt limit and a debt policy function that determines the next-period debt as a function of the current debt and the next-period endowment share. This function is constant when debt is low but is nonlinear and strictly increasing when debt is above a critical threshold (Corollary 1). The debt policy function and the history of endowments determine the dynamics of debt. Debt rises or falls depending on the evolution of endowments, but eventually, resets to a minimum level, creating cycles of debt. The difference between debt and the revenue generated from issuing state-contingent bonds defines the fiscal reaction function that measures how the tax rate depends on debt. Absent enforcement frictions, the fiscal reaction function is linearly increasing in debt. However, with enforcement frictions, the fiscal reaction function is linear when debt is low but nonlinear when it is high. In particular, when debt is below the threshold, the amount of debt issued is independent of the current debt, while the price of the state-contingent bonds is linearly decreasing in debt. Thus, bond revenue falls with debt, and the tax rate rises linearly. Above the threshold, two factors affect the fiscal reaction function. The price of state-contingent bonds decreases with debt, while bond issuance increases with debt according to the nonlinear debt policy function. The combined effect of these two factors results in a nonlinear fiscal reaction function.

The model also provides implications for asset pricing and the dependence of asset prices on debt (Proposition 6). Since the idiosyncratic and growth shocks are independent and identically distributed, the implied conditional yields are the sum of a risk-adjusted component and a constant given by the log of the average growth rate. The price of state-contingent bonds decreases with debt, which implies that the conditional yields, including the risk-free rate, increase with debt. The discount factor of the planner and the average growth rate determine the yield on the long bond. However, the long-short spread may be positive or negative. The dynamics of debt imply that the long-short spread is positive when debt is low, and the young are poor because, in this case, debt will rise, leading to higher expected future yields. Likewise, the long-short spread is negative when debt is high, and the young are rich because debt will fall, leading to lower expected future yields.

The variability of yields and their decomposition into growth-adjusted and growth-dependent components is also significant for debt valuation. There is a linear decomposition of the risk premium on debt into a risk-adjusted component and a risk premium on aggregate risk (Proposition 7). The return on bonds increases with the endowment of the young next period, as does the marginal utility of consumption of the old next period. Thus, the return on bonds is positively correlated with the stochastic discount factor for a given debt, resulting in a risk premium on debt lower than the risk premium on aggregate risk. In the absence of enforcement frictions, this gap is zero. When there are enforcement frictions, debt is a hedge for the endowment risk, and this reduces the risk premium on debt. Consequently, for a fixed plan of future primary surpluses, higher debt can be sustained compared to a case where the future surpluses are discounted using the risk premium on aggregate risk. This gap between the risk premiums on aggregate risk and debt offers a potential resolution to the *debt valuation puzzle* posed by Jiang et al. (2021), who find that the value of U.S. debt exceeds the present value of future primary surpluses when discounted by the risk premium on aggregate risk.[3] Moreover, the risk premium on debt varies with debt. In particular, it rises or falls depending on whether the expected return on debt increases with debt at a faster or slower rate than the risk-free interest rate.

In an example with two endowment states, we provide a closed-form solution for the bound on the variability of the implied yields and show that the invariant distribution of debt is a transformation of a geometric distribution (Proposition 8). Numerically, the solution can be found using a shooting algorithm without the need to solve a functional equation. In this example, the risk premium increases with debt, leading to a reduction in the gap between the risk premium on aggregate risk and the risk premium on debt.

Literature. The paper builds on the literature on risk sharing in models with overlapping generations. In most of this literature, transfers are mandatory, and consideration is restricted to stationary transfers (see, for example, Shiller, 1999, Rangel and Zeckhauser, 2001), in contrast to the voluntary and history-dependent transfers considered here. Our result on history dependence is foreshadowed in a mean-variance setting by Gordon and Varian (1988), who establish that any time-consistent optimal intergenerational risk-sharing agreement is nonstationary. Ball and Mankiw (2007) analyze risk sharing when generations can trade contingent claims before they are born. They find that idiosyncratic shocks are spread equally across generations, and consumption follows a random walk, as in Hall (1978). Such an allocation is not sustainable since it violates the participation constraint of some future generation almost surely. In contrast, we show that although the effects of a shock can be persistent, they are unevenly spread across future generations, and resetting ensures that they cannot last forever.

By interpreting the transfer to the old as public debt, we complement the extensive literature on debt sustainability and the fiscal reaction function that began with Bohn (1995, 1998). Our result on the nonlinearity of the fiscal reaction function echoes the discussion of *fiscal fatigue*, which argues that the primary fiscal balance responds sluggishly to rising debt when debt is high because of the adverse implications of debt, such as the risk of default (see, for example, Mendoza and Ostry, 2008, Ghosh et al., 2013). Despite the absence of default in our model, enforcement constraints generate nonlinearity in the fiscal reaction function. Bhandari et al. (2017) also study optimal fiscal policy and debt dynamics but in a model with infinitely-lived and heterogeneous agents where markets are incomplete because of constraints on tax policy. Brunnermeier et al. (2023) provide a result similar to ours that the risk premium on debt is lower than the risk premium on aggregate risk. In their model, infinitely-lived agents must

[3] For an overview of the debt valuation and sustainability, see, for example, Reis (2022), Willems and Zettelmeyer (2022) and Jiang et al. (2023).

retain a fixed proportion of their idiosyncratic risk. Government debt serves as a hedge against idiosyncratic risk, and consequently, debt becomes a negative beta asset. The authors emphasize that debt can command a bubble premium, which may add to the safety of government debt. In contrast to Brunnermeier et al. (2023), our model has no bubble component, and the extent of risk sharing is determined endogenously, depending on the history of endowment shocks.

Methodologically, the paper relates to the literature on risk sharing and limited enforcement frictions with infinitely-lived agents. Two polar cases have been examined: one with two infinitely-lived agents (see, for example, Thomas and Worrall, 1988, Chari and Kehoe, 1990, Kocherlakota, 1996) and the other with a continuum of infinitely-lived agents (see, for example, Thomas and Worrall, 2007, Krueger and Perri, 2011, Broer, 2013). The overlapping generations model considered here has a continuum of agents, but only two agents are alive at any point in time. The model is not nested in either of the two infinitely-lived agent models but fills an essential gap in the literature by analyzing optimal intergenerational insurance with limited enforcement frictions. Here, we establish strong convergence to the invariant distribution, whereas Krueger and Perri (2011) and Broer (2013) consider the solution only at an invariant distribution and Thomas and Worrall (2007) discuss convergence only in a particular case.

Plan of paper. Section 1 sets out the model. Section 2 considers two benchmarks: one with full enforcement of transfers from the young to the old and the other without risk. Section 3 characterizes optimal sustainable intergenerational insurance and Section 4 establishes convergence to an invariant distribution on a countable ergodic set. Section 5 provides an interpretation of the optimum in terms of debt and derives the fiscal reaction function. Section 6 discusses the implications for asset pricing and Section 7 considers the valuation of debt. Section 8 presents an example with two endowment states. Section 9 concludes. The Online Appendix contains the proofs of the main results. Additional proofs and further details can be found in the Supplementary Appendix.

1 The Model

Time is discrete and indexed by $t = 0, 1, 2, \ldots, \infty$. The model consists of a pure exchange economy with an overlapping generations demographic structure. At each time t, a new generation is born and lives for two periods. The generation born at date t has a population of N_t homogeneous agents. We assume that there is no population growth and normalize $N_t = 1$, so it is as if each generation has a single agent.[4] Each agent is young in the first period of life

[4] The assumption that agents of the same generation are homogeneous makes it possible to focus on intergenerational risk sharing. However, it does mean that we ignore questions about inequality within generations and its evolution over time. Although we maintain the assumption of a constant population, the qualitative properties of the model are unchanged if there is a constant rate of population growth. Additionally, Part D of the Supplementary

and old in the second. The economy starts at $t = 0$ with an initial old agent and an initial young agent. Since time is infinite, the initial old agent is the only agent that lives for just one period.

At each time t, agents receive an endowment of a perishable consumption good. Endowments are finite and strictly positive. The young and the old endowments at time t are e_t^y and e_t^o with an aggregate endowment of $e_t = e_t^y + e_t^o$. The endowment *share* of the young is $s_t := e_t^y / e_t$ (the endowment share of the old is $1 - s_t$), and the gross *growth* rate of the aggregate endowment is $\gamma_t := e_t / e_{t-1}$. There is both idiosyncratic (share of the generation's endowment) risk and aggregate (growth) risk. The sequences of random variables $(s_t : t \geq 0)$ and $(\gamma_t : t \geq 0)$ take values in finite sets $\mathcal{I}$ and $\mathcal{J}$, respectively, where $|\mathcal{I}| = I \geq 2$ and $|\mathcal{J}| = J \geq 1$. The pair $\rho_t := (s_t, \gamma_t)$ taking values in $\mathcal{P} \subseteq \mathcal{I} \times \mathcal{J}$ follows a finite-state, aperiodic, time-homogeneous Markov process with the probability of transiting from ρ_t to state ρ_{t+1} next period given by $\varpi(\rho_t, \rho_{t+1})$.

Denote the history of endowment shares and growth rates up to and including time t by $s^t := (s_0, s_1, ..., s_t) \in \mathcal{I}^t$ and $\gamma^t := (\gamma_0, \gamma_1, ..., \gamma_t) \in \mathcal{J}^t$ and let $\rho^t := (\rho_0, \rho_1, ..., \rho_t) \in \mathcal{P}^t$. The distribution of ρ_0 is given by the function $\varpi(\rho_0)$ and the probability of reaching the history ρ^t is $\varpi(\rho^t) = \varpi(\rho^{t-1})\varpi(\rho_{t-1}, \rho_t)$. Hence, the aggregate endowment at time t is the random variable $e_t = \prod_{k=0}^t \gamma_k$ with $\gamma_0 = e_0$.

There is complete information. Endowments depend only on the current state, whereas consumption can, in principle, depend on the history of states. Denote the per-period consumption of the young by $C(\rho^t)$ and the corresponding consumption share by $c(\rho^t) = C(\rho^t)/e_t$. There is no technology to store the endowment from one period to the next, and hence, the aggregate endowment is consumed each period. Consequently, the per-period consumption of the old is $e_t - C(\rho^t)$ and the corresponding consumption share is $1 - c(\rho^t)$. In autarky, agents consume only their own endowments, that is, the consumption share of the young is s_t, and the consumption share of the old is $1 - s_t$ for all t and (ρ^{t-1}, ρ_t).

Each generation is born after that period's uncertainty is resolved when the growth rate of the economy and the endowment shares of the young and the old are known. Therefore, after birth, a generation only faces uncertainty in old age, and there is no insurance market in which the young can insure against their endowment risk. Let $\{C\} = \{C(\rho^t) : t \geq 0, \rho^t \in \mathcal{P}^t\}$ denote a history-contingent consumption stream of the young. Then, the lifetime utility gain over autarky for a generation born after the history ρ^t is:

$$U\left(\{C\}; \rho^t\right) := u(C(\rho^t)) - u(e_t^y) + \beta \sum\nolimits_{\rho_{t+1}} \varpi(\rho_t, \rho_{t+1}) \left(u(e_{t+1} - C(\rho^t, \rho_{t+1})) - u(e_{t+1}^o)\right),$$

Appendix examines the impact of a demographic cohort shock and shows how the effect of a shock can be amplified and persistent.

where $u(\cdot)$ is the per-period utility function, common to both the young and the old, and $\beta \in (0, 1]$ is the generational discount factor. We assume the per-period utility function is logarithmic, $u(\cdot) = \log(\cdot)$. Hence, the preferences of an agent can be expressed in terms of consumption and endowment shares. In particular, since $e_t^y = s_t e_t$ and $C(\rho^t) = c(\rho^t)e_t$, it follows that $u(C(\rho^t)) - u(e_t^y) = \log(c(\rho^t)) - \log(s_t)$ and $U(\{C\}; \rho^t) = U(\{c\}; \rho^t)$ where

$$U\left(\{c\}; \rho^t\right) := \log(c(\rho^t)) - \log(s_t) + \beta \sum\nolimits_{\rho_{t+1}} \varpi(\rho_t, \rho_{t+1}) \left(\log(1-c(\rho^t, \rho_{t+1})) - \log(1 - s_{t+1})\right).$$

We call the history-contingent stream of consumption shares $\{c\} = \{c(\rho^t) : t \geq 0, \rho^t \in \mathcal{P}^t\}$ an *intergenerational insurance rule* since it determines how consumption is allocated between the young and the old for any history ρ^t. Since storage is not possible and because the young are born after uncertainty is resolved, the only means of achieving intergenerational insurance is through transfers between the young and the old. We assume that there is a benevolent social planner who chooses an intergenerational insurance rule of history-contingent transfers to maximize a discounted sum of the expected utilities of all generations. Let the planner's expected discounted utility gain over autarky, conditional on the history ρ^t, be

$$V\left(\{c\}; \rho^t\right) := \frac{\beta}{\delta}\left(\log(1 - c(\rho^t)) - \log(1 - s_t)\right) + \mathbb{E}_t\left[\sum\nolimits_{j=t}^{\infty} \delta^{t-j} U\left(\{c\}; \rho^j\right)\right]$$

where $\mathbb{E}_t$ is the expectation over future histories at time t. The planner's discount factor is $\delta \in (0, 1)$, and the weight on the utility of the initial old is β/δ.[5] To maximize the discounted sum of expected lifetime utilities, the planner must respect the constraint that transfers are voluntary.[6] That is, the planner must respect the constraint that neither the old nor the young would be better off in autarky than adhering to the specified transfers for any history of shocks. For the old, this means they will not make a positive transfer to the young because there is no future benefit to offset such a transfer. Hence, the consumption of the young cannot exceed their endowment, or equivalently,

$$c(\rho^t) \leq s_t \quad \text{for all } t \geq 0 \text{ and } \rho^t \in \mathcal{P}^t. \tag{1}$$

The analogous participation constraint for the young requires that the conditional transfers promised for their old age sufficiently compensate for the transfer made when young so that they are no worse off than reneging on the transfer today and receiving the corresponding autarkic

[5] The assumption of geometric discounting for the planner is common (see, for example, Farhi and Werning, 2007). Using a weight of β/δ for the initial old preserves the same relative weights on the young and the old, including the initial old, in every period.

[6] The assumption that the transfer is voluntary can be interpreted as requiring that the intergenerational insurance rule is supported by each generation if put to a vote.

lifetime utility. That is,

$$U\left(\{c\};\rho^t\right) \geq 0 \quad \text{for all } t \geq 0 \text{ and } \rho^t \in \mathcal{P}^t. \tag{2}$$

For expositional simplicity, let the initial state ρ_0 be given.[7] Hence, at $t = 0$, the planner chooses $\{c\}$ to maximize:

$$V\left(\{c\};\rho_0\right), \tag{3}$$

subject to the constraint set $\Lambda := \{\{c\} \mid (1) \text{ and } (2)\}$. Since utility is strictly concave, and the constraints in (2) are linear in utility, the planner's objective in equation (3) is concave and the constraint set Λ is convex and compact.

Definition 1. *An Intergenerational Insurance rule is sustainable if the history-dependent sequence* $\{c\} \in \Lambda$.

Definition 2. *A Sustainable Intergenerational Insurance rule is optimal if it maximizes the objective in equation* (3) *subject to the constraint that the initial old receive a utility from their consumption share of at least* $\bar{\omega}_0$*:*

$$\log(1 - c(\rho_0)) \geq \bar{\omega}_0. \tag{4}$$

We introduce constraint (4) with an exogenous initial target utility of $\bar{\omega}_0$ because it is useful when considering the evolution of the optimal sustainable intergenerational insurance rule in Section 3.[8] However, we will return to the case where the planner chooses the initial $\bar{\omega}_0$.

Since $U(\{C\};\rho^t) = U(\{c\};\rho^t)$ and utility is logarithmic, the objectives and constraints are equivalent whether consumption is expressed in levels or shares. That is, the economy with stochastic growth is equivalent to an economy with a constant endowment and consumption expressed as shares of the aggregate endowment. The growth rate of the consumption levels is simply the growth rate of the consumption shares multiplied by the growth rate of the aggregate endowment.

Remark 1. *For preferences that exhibit constant absolute risk aversion, this equivalence property is well-known to hold in models of idiosyncratic and aggregate risk with infinitely-lived agents (see, for example, Alvarez and Jermann, 2001, Krueger and Lustig, 2010). An analogous extension can be shown to hold here by defining growth-adjusted transition probabilities and*

[7] The analysis is easily generalized to any given initial distribution $\varpi(\rho_0)$.

[8] The initial target utility may also depend on the initial state. Varying $\bar{\omega}_0$ traces out the Pareto frontiers that trade-off the utility of the old against the planner's valuation of the expected discounted utility of all future generations.

discount factors to satisfy the following:

$$\hat{\varpi}(\rho_t, \rho_{t+1}) := \frac{\varpi(\rho_t, \rho_{t+1})(\gamma_{t+1})^{1-\alpha}}{\sum_{\rho_{t+1}} \varpi(\rho_t, \rho_{t+1})(\gamma_{t+1})^{1-\alpha}} \quad \text{and} \quad \frac{\hat{\beta}(\rho_t)}{\beta} = \frac{\hat{\delta}(\rho_t)}{\delta} := \sum_{\rho_{t+1}} \varpi(\rho_t, \rho_{t+1})(\gamma_{t+1})^{1-\alpha},$$

where α is the coefficient of relative risk aversion.

In what follows, we make the simplifying assumption that the shocks to endowment shares and growth rates are independent and are identically and independently distributed (hereafter, i.i.d.).

Assumption 1. *(i) The state ρ is i.i.d. with the probability given by $\varpi(\rho)$. (ii) The endowment share and the growth rate are independently distributed, that is, $\varpi(\rho) = \pi(s)\varsigma(\gamma)$ where $\pi(s)$ and $\varsigma(\gamma)$ are the marginal distributions for the endowment shares and the growth rates respectively.*

By Part (i) of Assumption 1, the economy is stationary. We make this assumption to emphasize that the history dependence we derive below follows from the participation constraints rather than from any feature of the economic environment itself.[9] Since the terms $U(\{c\};\rho^t)$ and $V(\{c\};\rho^t)$ depend on the growth rates γ_t and γ_{t+1} only via the transition function $\varpi(\rho_t, \rho_{t+1})$, it follows that the consumption shares in any optimal sustainable intergenerational insurance rule depend only on the history of endowment shares s^t.

Proposition 1. *Under Assumption 1, the consumption shares in any optimal sustainable intergenerational insurance rule depend only on the history s^t and are independent of the history of growth shocks γ^t.*

A similar result is well known from models with infinitely-lived agents (see, again, Alvarez and Jermann, 2001, Krueger and Lustig, 2010).[10] We will consider the asset pricing implications and contrast them to those from infinitely-lived models in Section 6.

Preliminaries. Since there are $I \geq 2$ states for the endowment share, order states such that $s(i) < s(i+1)$ for $i = 1, \dots, I-1$, so that, a higher state corresponds to a larger endowment share for the young. For convenience, we will refer to states $1, 2, \dots, I$ corresponding to the shares $s(1), s(2), \dots, s(I)$ and to simplify notation will sometimes express variables as a function of i rather than s.

[9] The assumption of i.i.d. shocks is standard in OLG models where a generation may cover 20-30 years.

[10] Under Assumption 1 and preferences exhibiting constant relative risk aversion, the discount factors defined in Remark 1 satisfy $\hat{\beta}/\beta = \hat{\delta}/\delta = \sum_\gamma \varsigma(\gamma)\gamma^{1-\alpha}$. If $\alpha \neq 1$, then the planner's objective is finite provided $\delta \sum_\gamma \varsigma(\gamma)\gamma^{1-\alpha} < 1$.

Under Assumption 1, the existence of a non-autarkic sustainable allocation can be addressed by considering small stationary transfers that depend only on the current endowment state. Denote the intertemporal marginal rate of substitution between the consumption share when young in state s and the consumption share when old in state r next period, evaluated at autarky, by $\hat{m}(s,r) := \beta s/(1-r)$ and let $\hat{q}(s,r) := \pi(r)\hat{m}(s,r)$. The terms $\hat{m}(s,r)$ and $\hat{q}(s,r)$ correspond to the stochastic discount factor and the state prices in an equilibrium model. Denote the $I \times I$ matrix of terms $\hat{q}(s,r)$ by $\hat{Q}$. A non-autarkic sustainable allocation exhausting the aggregate endowment and satisfying the participation constraints in (1) and (2) exists whenever the Perron root of $\hat{Q}$ is greater than one (see, for example, Aiyagari and Peled, 1991, Chattopadhyay and Gottardi, 1999). In this case, there exists a vector of strictly positive stationary transfers that improves the lifetime utility of the young in each state. Since the endowment states are independent, the matrix $\hat{Q}$ has rank one, and the Perron root is its trace. We assume that the trace of $\hat{Q}$ is larger than the harmonic mean of the growth factors, $\bar{\gamma} := (\sum_{\gamma} \varsigma(\gamma)\gamma^{-1})^{-1}$.

Assumption 2. $\sum_{s\in\mathcal{I}} \hat{q}(s,s) > \bar{\gamma}$.

If there is just one state with the young receiving a share s of the aggregate endowment and no growth, then Assumption 2 reduces to the standard Samuelson condition: $s > 1/(1+\beta)$. In this case, it is well known that there are Pareto-improving transfers from the young to the old. Assumption 2 is the generalization to the stochastic case and a natural assumption given that our focus is on transfers to the old.[11] Given Assumption 2, it follows that the constraint set Λ is nonempty.

Proposition 2. *Under Assumption 2, there exists a non-autarkic and stationary Sustainable Intergenerational Insurance rule.*

Furthermore, we assume:

Assumption 3. $s(1) \leq \delta/(\beta+\delta)$.

Assumption 3 provides a simple sufficient condition for the strong convergence result of Section 4. Since $\delta < 1$, Assumption 3 implies that $s(1) < 1/(1+\beta)$. That is, in the absence of growth, the state-wise Samuelson condition does not hold in every state, showing that our results do not depend on this property. In the terminology of Gale (1973), the economy can be viewed as a mix of Samuelson and classic cases.

[11] A simple sufficient condition for Assumption 2 to be satisfied is that the Frobenius lower bound, given by the minimum row sum of $\hat{Q}$, is greater than $\bar{\gamma}$. A row sum greater than $\bar{\gamma}$ implies that in autarky, the young would wish to save for their old age in each endowment state even if the net interest rate were zero.

2 Two Benchmarks

Before turning to the characterization of the optimal sustainable intergenerational insurance, it is helpful to consider two benchmark cases that illustrate the inefficiencies generated by the presence of limited enforcement and uncertainty. The first benchmark ignores the participation constraints of the young but not the participation constraints of the old. The second benchmark considers an economy without risk but requires that the planner respects the participation constraints of both the young and the old.

First Best. Suppose the planner ignores the participation constraints of the young and let $\Lambda^* := \{\{c\} \mid (1)\}$ denote the set of transfers without the constraints in (2).[12]

Definition 3. *An Intergenerational Insurance* $\{c\} \in \Lambda^*$ *is first best if it maximizes the objective function* (3) *subject to constraint* (4).

It is easy to verify that at the first-best optimum:

$$c^*(s^t) = \min\left\{\frac{\delta}{\beta+\delta}, s_t\right\} \quad \text{for all } t > 0 \text{ and } s^t \in \mathcal{S}^t. \tag{5}$$

Condition (5) shows that the consumption shares of the young are kept constant unless doing so would involve a transfer from the old to the young, in which case the consumption share is the autarky value.[13] That is, at the first best, the consumption share is independent of the history s^{t-1} and depends only on the current endowment share s_t when the nonnegativity constraint on the transfer binds. Under Assumption 3, there is always one state in which the participation constraint of the old holds with equality.

It can be seen from condition (5) that for states in which transfers are positive, the first-best consumption share of the young is independent of s. It is decreasing in β since a higher β puts more weight on the utility of the old who receive the transfer, and it is increasing in δ since a higher δ puts more weight on the utility of the young who make the transfer.

Let $\omega_{\min}(s) := \log(1-s)$ be the utility of the old at autarky and $\omega^* := \log(\beta/(\beta+\delta))$ be the utility of the old when the consumption share of the young is $\delta/(\beta+\delta)$. Then, $\omega^*(s) := \max\{\omega_{\min}(s), \omega^*\}$ is the utility of the old at the first-best solution when the endowment share of the young is s. Since s_0 is the endowment share of the young at the initial date, it follows

[12] Hereafter, the asterisk designates the first-best outcome. Note that the first best could be defined by assuming that the planner ignores the participation constraints of both the young and the old. The reason for presenting the first best as we do is to show that this allocation is stationary. Hence, any history dependence of the optimal sustainable intergenerational insurance rule derives from the imposition of the participation constraints of the young.

[13] Condition (5) is a special case of the familiar Arrow-Borch condition for optimal risk sharing modified to account for the constraint that transfers are only from the young to the old.

from Definition 2 that constraint (4) does not bind when $\bar{\omega}_0 \le \omega^*(s_0)$. In this case, the first-best consumption at $t = 0$ is $c^*(s_0)$, determined by condition (5) as in every other time $t > 0$. On the other hand, for $\bar{\omega}_0 > \omega^*(s_0)$, constraint (4) binds and $c^*(s_0) = 1 - \exp(\bar{\omega}_0)$. In this case, the initial transfer to the old is correspondingly higher than implied by condition (5).

Denote the planner's per-period payoff at the first-best allocation by $v^*(s) = \log(c^*(s)) + (\beta/\delta)\log(1 - c^*(s))$ and the expected discounted payoff to the planner by $V^*(s_0, \omega)$ when the initial endowment share is s_0 and the initial utility of the old is ω. The maximum utility the old can get occurs if they consume all of the endowment, so that $\omega_{\max} = \log(1) = 0$. Let $\Omega(s_0) = [\omega_{\min}(s_0), 0]$ be the set of possible utilities for the old at the initial state, $\bar{v}^* := \sum_s \pi(s)v^*(s)$ be the planner's expected per-period payoff at the first-best solution and $\bar{V}^* := \bar{v}^*/(1-\delta)$ be the corresponding continuation payoff. The first-best outcome is summarized in the following proposition.[14]

Proposition 3. *(i) The consumption share $c^*(s^t)$ is stationary and satisfies condition* (5) *for $t > 0$. For $t = 0$, $c^*(s_0)$ satisfies condition* (5) *for $\omega \le \omega^*(s_0)$ and $c^*(s_0) = 1 - \exp(\omega)$ for $\omega > \omega^*(s_0)$. (ii) The value function $V^*(s_0, \cdot)\colon \Omega(s_0) \to \mathbb{R}$ has $V^*(s_0, \omega) = v^*(s_0) + \delta\bar{V}^*$ for $\omega \le \omega^*(s_0)$ and $V^*(s_0, \omega) = (\beta/\delta)\omega + \log(1 - \exp(\omega)) + \delta\bar{V}^*$ for $\omega > \omega^*(s_0)$, where the derivative $V^*_\omega(s_0, \omega^*(s_0)) = \min\{0, (\beta/\delta) - ((1 - s_0)/s_0)\}$ with $\lim_{\omega\to 0} V^*_\omega(s_0, \omega) = -\infty$.*

The value function $V^*(s_0, \omega)$ is decreasing and concave in ω (strictly decreasing and strictly concave in ω for $\omega > \omega^*(s_0)$). The function "extends to the left" when the endowment share s_0 is higher.[15] If $\omega^*(s_0) > \omega_{\min}(s_0)$ (or equivalently, if $s_0 > \delta/(\beta + \delta)$), then $V^*(s_0, \omega)$ is independent of ω for $\omega \le \omega^*(s_0)$. Hence, in the absence of constraint (4), the planner would choose $\omega(s_0) = \omega^*(s_0)$ because this gives the highest utility to the initial old while maximizing the payoff to the planner. In this case, the allocation given by condition (5) holds in every period. In contrast, when $\bar{\omega}_0 > \omega^*(s_0)$, the consumption share of the young is lower than implied by condition (5), but only in the initial period. There is immediate convergence to the stationary first-best distribution in one period.

Since the payoff to the planner depends both on s and ω, the stationary distribution is a pair $(s, \omega^*(s))$, the endowment share and the corresponding utility promised to the old. We note for future reference that this stationary distribution has I values, one for each endowment state, with the probability of each pair given by $\pi(s)$.

Deterministic Economy. We now consider a deterministic economy with a constant growth rate γ and endowment share s. Unlike the previous benchmark, we assume that the planner

[14] The proof of Proposition 3 is omitted because it follows from standard arguments. Nonetheless, the properties of the function $V^*(s_0, \omega)$ are mirrored in Proposition 4 and Lemma 1, given below, which do respect the participation constraints of the young.

[15] That is, for $s > r$ where $\omega_{\min}(s) < \omega_{\min}(r)$, $V^*(s, \omega) = V^*(r, \omega)$ for $\omega \in \Omega(r)$.

respects the participation constraint of both the young and the old. Let $\hat{v} := \log(s) + \beta \log(1 - s)$ be the lifetime endowment utility. Assumption 2, together with the strict concavity of the utility function, implies that there is a unique $c_{\min} < s$, which is the lowest *stationary* consumption share of the young that satisfies the participation constraint with equality. The corresponding maximum utility of the old is $\omega_{\max} := \log(1 - c_{\min})$.[16] Analogously to condition (5), the first-best consumption share is $c^* = \delta/(\beta + \delta)$ and the corresponding utility of the old is $\omega^* := \log(\beta/(\beta + \delta))$. If δ is above a critical value, then $c^* > c_{\min}$ (or equivalently, $\omega^* < \omega_{\max}$) and the first-best consumption share is sustainable. Otherwise, the first-best consumption share is not sustainable.

Denote the consumption share of the young at time t by c_t and the corresponding utility of the old by $\omega_t = \log(1 - c_t)$. Consider the maximization problem in (3) with the participation constraints of the young given by $\log(c_t) + \beta \log(1 - c_{t+1}) \geq \hat{v}$. For $\bar{\omega}_0 \leq \omega^*$, constraint (4) does not bind and it is optimal to set $c_t = \max\{c^*, c_{\min}\}$ (or equivalently, $\omega_t = \min\{\omega^*, \omega_{\max}\}$) for all $t \geq 0$. On the other hand, consider the case where $\omega^* < \omega_{\max}$ and $\bar{\omega}_0 > \omega^*$. Then, at $t = 0$, c_0 must satisfy $\log(1 - c_0) \geq \bar{\omega}_0$, which requires that $c_0 < c^*$. Clearly, it is desirable to set c_0 such that $\log(1 - c_0) = \bar{\omega}_0$ and $c_1 = c^*$. However, setting $c_1 = c^*$ may violate the participation constraint of the young. In this case, c_1 has to be chosen to satisfy $\log(c_0) + \beta \log(1 - c_1) = \hat{v}$, which implies that $c_1 < c^*$. Repeating this argument for $t > 1$ shows that given c_t, the consumption share of the young at time $t + 1$ either satisfies $\log(c_t) + \beta \log(1 - c_{t+1}) = \hat{v}$ or $c_{t+1} = c^*$ if $\log(c_t) + \beta \log(1 - c^*) \geq \hat{v}$. Intuitively, if the consumption share of the young is low (or equivalently, if the utility of the old is large), then the planner would like to raise the consumption share of the young to c^* (or equivalently, reduce ω to ω^*) as fast as possible to improve welfare. However, if the consumption share of the next-period young is raised too much, then the lifetime utility of the current young falls, and their participation constraint is violated. That is, in the presence of limited enforcement, the consumption share of the young has to be raised gradually. It is useful to express this rule in terms of a policy function $g(\omega)$ for the promised utility next period, where

$$g(\omega) := \begin{cases} \omega^* & \text{for } \omega \in [\omega_{\min}, \omega^c], \\ \frac{1}{\beta}\left(\hat{v} - \log\left(1 - \exp(\omega)\right)\right) & \text{for } \omega \in (\omega^c, \omega_{\max}], \end{cases} \tag{6}$$

with $\omega_{\min} = \log(1 - s)$ and $\omega^c := \log(1 - \exp(\hat{v} - \beta\omega^*))$. It follows from the strict concavity of the utility function that $\omega^c > \omega^*$. The function $g(\omega)$ is increasing and convex in ω, as illustrated in Figure 1. The dynamic evolution of ω_t is straightforwardly derived from $g(\omega)$: for

[16] The maximum utility of the old can be found by solving $\log(1 - \exp(\omega_{\max})) + \beta\omega_{\max} = \hat{v}$. Equivalently, the minimum consumption is found by solving $\log(c_{\min}) + \beta \log(1 - c_{\min}) = \hat{v}$.

$\omega_t \in [\omega_{\min}, \omega^c]$, $\omega_{t+1} = \omega^*$ for all t; for $\omega_t \in (\omega^c, \omega_{\max}]$, ω_{t+1} declines monotonically. Since $\omega^c > \omega^*$, the process for ω_t converges to ω^*, attaining its long-run value in finite time.

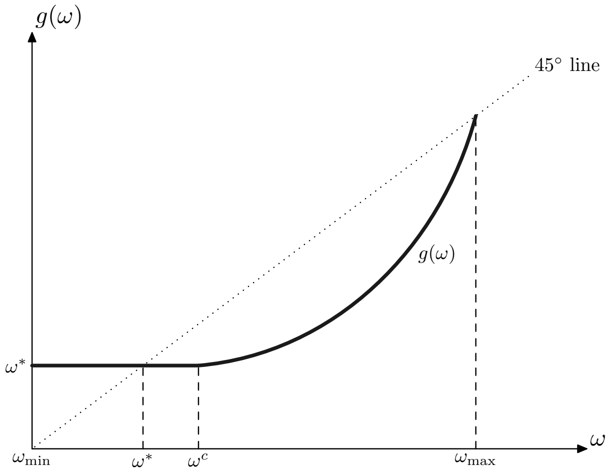

Figure 1: Policy Function in the Deterministic Case with $\omega_{\max} > \omega^*$.
Note: The solid line is the deterministic policy function $g(\omega)$. For any initial $\omega \in [\omega_{\min}, \omega_{\max})$, ω_t converges to ω^*.

Denote the per-period payoff to the planner with the first-best allocation in the absence of uncertainty by $v^* := \log(\delta/(\beta+\delta)) + (\beta/\delta)\omega^*$ and the expected discounted payoff to the planner for $\omega \in \Omega := [\omega_{\min}, \omega_{\max}]$ by $V(\omega)$. The optimal solution for the deterministic case with sustainable ω^* is summarized in the following proposition.

Proposition 4. *(i) If $\omega \in [\omega_{\min}, \omega^*]$, then the consumption share $c_t = \delta/(\beta + \delta)$ for $t \geq 0$. (ii) If $\omega \in (\omega^*, \omega_{\max}]$, then ω_{t+1} satisfies equation (6). There exists a finite T such that ω_t is monotonically decreasing for $t < T$ and $\omega_t = \omega^*$ for $t \geq T$. Likewise, c_t is monotonically increasing for $t < T$ and $c_t = c^*$ for $t \geq T$. (iii) The value function $V\colon \Omega \to \mathbb{R}$ is equal to $V(\omega) = v^*/(1 - \delta)$ for $\omega \in [\omega_{\min}, \omega^*]$ and is strictly decreasing and strictly concave for $\omega \in (\omega^*, \omega_{\max}]$ with* $\lim_{\omega \to \omega_{\max}} V_\omega(\omega) = -\infty$.

The optimal solution is either stationary or converges monotonically to a stationary point within finite time, with $c_T = c^*$ for T large enough. Hence, the long-run distribution of ω is degenerate and for the case where $c^* > c_{\min}$, it has a single mass point at $\{\omega^*\}$.

In the following sections, we show that when the first-best allocation violates a participation constraint of the young, and there is endowment risk, the optimal sustainable intergenerational insurance is history dependent even in the long run, and the ergodic set of utilities has more than I values. The benchmarks highlight that both limited enforcement of transfers and risk are necessary for this result.

3 Optimal Sustainable Intergenerational Insurance

In this section, we characterize the optimal intergenerational insurance rule under uncertainty when the planner respects the participation constraints of both the young and the old. Recall that shocks to growth rates and endowment shares are i.i.d. (Assumption 1) and that the optimal sustainable consumption shares depend only on the history of endowment share s^t (Proposition 1). We rule out the case in which the first-best outcome is sustainable and assume that the first-best allocation violates the participation constraint of the young in at least one state. Since the lifetime endowment utility of an agent is increasing in s, we assume that:

Assumption 4. $\log(c^*(I)) + \beta \sum_r \pi(r) \log(1 - c^*(r)) < \log(s(I)) + \beta \sum_r \pi(r) \log(1 - r)$.

We reformulate the optimization problem described in Definition 2 recursively using the utility ω promised to the old as a state variable. Let ω_r denote the state-contingent utility promised to the current young for their old age when the endowment share of the young next period is r. Then, the planner's optimization problem is:

$$V(s, \omega) = \max_{\{c, (\omega_r)_{r \in \mathcal{I}}\} \in \Phi(s, \omega)} \frac{\beta}{\delta} \log(1 - c) + \log(c) + \delta \sum_r \pi(r) V(r, \omega_r), \tag{P1}$$

where $\Phi(s, \omega)$ is the constraint set given by the following inequalities:

$$\log(1 - c) \geq \omega, \tag{7}$$

$$c \leq s, \tag{8}$$

$$\omega_r \leq \omega_{\max}(r) \quad \text{for each } r \in \mathcal{I}, \tag{9}$$

$$\omega_r \geq \omega_{\min}(r) \quad \text{for each } r \in \mathcal{I}, \tag{10}$$

$$\log(c) + \beta \sum_r \pi(r) \omega_r \geq \log(s) + \beta \sum_r \pi(r) \log(1 - r). \tag{11}$$

The recursive formulation is similar to the promised-utility approach used in models with infinitely-lived agents (see, for example, Green, 1987, Spear and Srivastava, 1987, Thomas and Worrall, 1988, Atkeson and Lucas Jr., 1992). At each period, the planner chooses the consumption share of the young, c, and the state-contingent promise of utility, ω_r. The state variable ω embodies information about the history of shocks. Constraint (7) is the promise-keeping constraint, which requires the current old to receive at least what they were promised previously. It is analogous to constraint (4), but it is now required to hold in every period. Constraint (8) is the participation constraint for the old, which stipulates that the old do not transfer to the young. Constraints (9) and (10) require that the promise is feasible: $\omega_r \in \Omega(r) := [\omega_{\min}(r), \omega_{\max}(r)]$. Finally, constraint (11) requires that the consumption share of the young and the promises made to them for their old age at least match the expected lifetime utility they would receive in autarky.

It is easy to check that the constraint set $\Phi(s,\omega)$ is convex and compact. Denote the state vector by $x := (s,\omega)$ and let $f(x)$ and $g_r(x)$ for $r \in \mathcal{I}$ be the optimal consumption share of the young and the state-contingent promise of utility of the old next period. The compactness of the constraint set guarantees the existence of the optimal policies, and the strict concavity of the utility function guarantees uniqueness. The optimal allocation is solved recursively. Starting at date $t = 0$ with a given state s_0 and given $\omega_0 \in \Omega(s_0)$, solve the optimization problem P1 to obtain the policy functions $f(s_0,\omega_0)$ and $g_r(s_0,\omega_0)$ for $r \in \mathcal{I}$. For the second period, solve the maximization problem again using the endowment share realized at date 1, say $\hat{r}$, together with the utility promise from the first period, $g_{\hat{r}}(s_0,\omega_0)$, in equation (7). The process is then repeated for subsequent periods.

The function $V(s,\omega)$ cannot be found by standard contraction mapping arguments starting from an arbitrary value function because the value function associated with the autarkic allocation also satisfies the functional equation of problem P1. However, a similar iterative approach can be used to find the value function, starting from the first-best value functions $V^*(s,\omega)$ derived in Proposition 3. Following the arguments of Thomas and Worrall (1994), the limit of this iterative mapping is the optimal value function $V(s,\omega)$. Proposition 3 established that the first-best value function is nonincreasing, differentiable, and concave in ω, and the limit value function inherits these properties.

Lemma 1. *(i) The value function $V(s,\cdot)\colon\Omega(s) \to \mathbb{R}$ is nonincreasing, concave and continuously differentiable in ω, with $\omega_{\min}(s) < \omega_{\max}(s)$. (ii) For each $s \in \mathcal{I}$, there exists an $\omega^0(s) \in [\omega_{\min}(s),\omega^*(s)]$ such that $V(s,\omega)$ is strictly decreasing and strictly concave for $\omega > \omega^0(s)$. If $\omega^*(s) > \omega_{\min}(s)$, then $\omega^0(s) < \omega^*(s)$ for some state s and $\omega^0(s) > \omega_{\min}(s)$ for some (possibly different) state. For $\omega \in [\omega_{\min}(s),\omega^0(s)]$, $V_\omega(s,\omega) = 0$. If $\omega^*(s) = \omega_{\min}(s)$, then $\omega^0(s) = \omega^*(s)$ and $V_\omega(s,\omega^0(s)) \leq (\beta/\delta) - ((1-s)/s) \leq 0$. In either case, $\lim_{\omega\to\omega_{\max}(s)} V_\omega(s,\omega) = -(\beta/\delta)\lambda_{\max}(s)$, where $\lambda_{\max}(s) \in \mathbb{R}_+ \cup \{\infty\}$. (iii) The upper bounds satisfy $\omega_{\max}(s(i)) < \omega_{\max}(s(i-1)) < 0$. Similarly, $\omega^0(s(i)) \leq \omega^0(s(i-1))$ with strict inequality for at least one $i = 2,\dots,I$.*

The strict concavity of the objective function and the convexity of the constraint set guarantee the concavity of $V(s,\omega)$ in ω with $\omega^0(s) = \sup\{\omega \mid V_\omega(s,\omega) = 0\}$ if $V_\omega(s,\omega_{\min}(s)) = 0$ and $\omega^0(s) = \omega_{\min}(s)$ otherwise. Since the old will not transfer to the young voluntarily, $\omega_{\min}(s) = \log(1-s)$, the autarkic utility of the old, which is decreasing in s. The upper endpoints $\omega_{\max}(s)$ are determined by the system of equations $\log(1-\exp(\omega_{\max}(s)))+\beta\sum_r \pi(r)\omega_{\max}(r) = \log(s) + \beta\sum_r \pi(r)\log(1-r)$. It can be checked that there is a unique nontrivial solution with

$\omega_{\max}(s)$ decreasing with s and $\omega_{\min}(s) < \omega_{\max}(s) < 0$. Likewise, $\omega^0(s)$ is decreasing in s. Differentiability of $V(s,\omega)$ with respect to ω follows because the constraint set satisfies a linear independence constraint qualification when $\omega \in [\omega_{\min}(s), \omega_{\max}(s))$. The left-hand derivative of $V(s,\omega)$ with respect to ω, evaluated at $\omega_{\max}(s)$, is finite if $\omega_{\max}(s)$ is part of the ergodic set and is infinite otherwise.

Remark 2. *Recall that $\bar{\omega}_0$ is the exogenous target utility given in constraint (4). Given the definition of $\omega^0(s)$, the planner chooses the initial utility of the old such that $\omega_0 = \max\{\omega^0(s_0), \bar{\omega}_0\}$. If the planner is not subject to constraint (4) and can freely choose the initial utility, then the planner sets $\omega_0 = \omega^0(s_0)$. Note that each $\omega^0(s)$ is an optimal choice and depends on all of the model's primitives.*

Remark 3. *The optimal sustainable intergenerational insurance is not renegotiation-proof because, in the case of default, it would be possible to offer the promised utility $\omega^0(r)$ instead of $\omega_{\min}(r)$ without diminishing the planner's payoff. A renegotiation-proof outcome can then be derived by replacing constraint (11) with $\log(c) + \beta \sum_r \pi(r)\omega_r \geq \log(s) + \beta \sum_r \pi(r)\omega^0(r)$. Since $\omega^0(r)$ is determined as part of the solution and appears in the constraint, a fixed-point argument similar to that used by Thomas and Worrall (1994) is required to find the solution. Although imposing this tighter constraint restricts risk sharing, the structure of the constrained optimization problem is not affected. Therefore, we expect that the qualitative properties of the optimal solution are substantially unchanged.*

Optimal Policy Functions. We now turn to the properties of the policy functions $f(x)$ and $g_r(x)$. Given the differentiability of the value function, the first-order conditions for the programming problem P1 are:

$$f(x) = \min\left\{\frac{\delta(1+\mu(x))}{\beta(1+\lambda(x)) + \delta(1+\mu(x))}, s\right\} \tag{12}$$

$$V_\omega(r, g_r(x)) = -\frac{\beta}{\delta}\left(\mu(x) - \xi_r(x) + \eta_r(x)\right) \quad \text{for each } r \in \mathcal{I}, \tag{13}$$

where $(\beta/\delta)\lambda(x)$ is the multiplier corresponding to the promise-keeping constraint of equation (7), $\beta\pi(r)\xi_r(x)$ are the multipliers corresponding to the upper bound on the promised utility (9), $\beta\pi(r)\eta_r(x)$ are the multipliers corresponding to the lower bound on the promised utility (10), and $\mu(x)$ is the multiplier corresponding to the participation constraints of the young (11). Given the concavity of the programming problem, conditions (12) and (13) are both necessary and sufficient. There is also an envelope condition:

$$V_\omega(x) = -\frac{\beta}{\delta}\lambda(x). \tag{14}$$

Taken together, equations (13) and (14) imply the following updating property:

$$\lambda(x') = \mu(x) - \xi_r(x) + \eta_r(x), \tag{15}$$

where $x' = (r, g_r(x))$ is the next-period state variable. Equation (15) is easy to interpret. For simplicity, suppose that the boundary constraints on the promised utility do not bind, that is, $\xi_r(x) = \eta_r(x) = 0$. From equation (13), it follows that $\delta(1 + \mu(x))$ is the relative weight placed on the utility of the young and $\beta(1 + \lambda(x))$ is the relative weight placed on the utility of the old. The updating property in equation (15) shows that the relative weight placed on the utility of the old corresponds to the tightness of the participation constraint they faced when they were young.

The following two Lemmas describe the properties of the policy functions.[17]

Lemma 2. *(i) The policy function* $g_r(s, \cdot)\colon \Omega(s) \to [\omega^0(r), \omega_{\max}(r)]$ *is continuous and increasing in* ω *and strictly increasing for* $g_r(s, \omega) \in (\omega^0(r), \omega_{\max}(r))$*. (ii) For each* $r \in \mathcal{I}$ *and* $\omega \in (\omega_{\min}(s(i-1)), \omega_{\max}(s(i)))$*,* $g_r(s(i), \omega) \geq g_r(s(i-1), \omega)$ *with strict inequality for at least one* $i = 2, \ldots, I$*. For each* $s \in \mathcal{I}$*,* $g_{r(i)}(s, \omega) \leq g_{r(i-1)}(s, \omega)$ *with strict inequality for at least one* $i = 2, \ldots, I$*. (iii) For endowment state 1, there is a critical value* $\omega^c > \omega^0(1)$ *such that* $g_r(1, \omega) = \omega^0(r)$ *for* $\omega \in [\omega^0(1), \omega^c]$ *and* $r \in \mathcal{I}$*. (iv) For each* $s \in \mathcal{I}$*, there is a unique fixed point* $\omega^f(s) = \min\{\omega^*(s), \omega_{\max}(s)\}$ *of the mapping* $g_s(s, \omega)$ *with* $g_s(s, \omega) > \omega$ *for* $\omega < \omega^f(s)$ *and* $g_s(s, \omega) < \omega$ *for* $\omega > \omega^f(s)$*. For endowment state I,* $\omega^f(I) > \omega^0(I)$*.*

Lemma 3. *(i) The policy function* $f(s, \cdot)\colon \Omega(s) \to (0, s]$ *where* $f(s, \omega) = 1 - \exp(w)$ *for* $\omega \geq \omega^0(s)$ *and* $f(s, \omega) = 1 - \exp(\omega^0(s))$ *for* $\omega < \omega^0(s)$*. (ii)* $c^0(s) := f(s, \omega^0(s))$ *where* $c^0(s(i)) \geq c^0(s(i-1))$ *with strict inequality for at least one* $i = 2, \ldots, I$*. (iii) At the fixed point* $\omega^f(s)$*,* $f(s, \omega^f(s)) \leq c^*(s)$ *with equality for* $\omega^f(s) < \omega_{\max}(s)$*.*

The main properties of Lemmas 2 and 3 follow straightforwardly from the objective to share risk subject to the participation constraints. The policy function $g_r(s, \omega)$ is increasing in ω (Lemma 2(i)), whereas $f(s, \omega)$ is decreasing in ω (Lemma 3(i)). A higher promise to the current old means a lower consumption share for the current young and, for endowment states in which the participation constraint binds, this requires a higher future promise of utility for their old age as compensation. The consumption share of the young depends only indirectly on s when $\omega = \omega^0(s)$ or $\omega = \omega_{\max}(s)$ (Lemma 3(ii)), whereas $g_r(\omega, s)$ is increasing in s and decreasing in r (Lemma 2(ii)). The policy function $g_r(\omega, s)$ is increasing in s because a higher endowment share of the young today is associated with a larger risk-sharing transfer, which, if the participation constraint is binding, has to be compensated by a higher promise for tomorrow.

[17] To avoid the clumsy terminology of nondecreasing or weakly increasing, we describe a function as increasing if it is weakly increasing and highlight cases where a function is strictly increasing.

Likewise, the future promise is decreasing in r because a higher endowment share of the young tomorrow is associated with a higher consumption share when the participation constraint binds and, hence, a lower consumption share of the old tomorrow. Since the optimum is nontrivial and differs from the first best, there is at least one strict inequality in the relations of Lemma 2(ii), so that, $g_r(s(I), \omega) > g_r(s(1), \omega)$ and $g_{r(I)}(s, \omega) < g_{r(1)}(s, \omega)$.

Lemma 2(iii) shows that there is a range of ω above $\omega^0(1)$ such that the participation constraint of the young does not bind and hence, $g_r(1, \omega) = \omega^0(r)$ in this range. This is analogous to the deterministic case discussed in Section 2 where the policy function has an initial flat section (see, Figure 1). More generally, when the participation constraint of the young does not bind, it follows from equation (14) that $g_r(x) = \omega^0(r)$ and $x' = (r, \omega^0(r))$. In this case, we say that the promise is *reset*. The promise is reset to the value that gives the most to the current old while maximizing the payoff to the planner. Lemma 2(iii) shows that resetting, in particular, occurs in state 1 for any $\omega \in [\omega^0(1), \omega^c]$.

Lemmas 2(iv) and Lemma 3(iii) describe what happens when the same endowment share repeats in successive periods. Suppose for simplicity that $\eta_s(x) = \xi_s(x) = 0$ and $f(x) > s$. From equations (13) and (14), $\mu(s, \omega^f(s)) = \lambda(s, \omega^f(s))$ where $\omega^f(s)$ is the fixed point of $g_s(s, \omega)$. Using equation (12), this implies that the consumption share is first best and hence, $\omega^f(s) = \omega^*(s)$. Furthermore, $g_s(s, \omega) > \omega$ for $\omega < \omega^f(s)$ and $g_s(s, \omega) < \omega$ for $\omega > \omega^f(s)$. That is, when the same endowment share repeats, the promise falls if the previous promise was above the first best and rises if the previous promise was below the first best. It follows that the policy function $g_s(s, \omega) > \omega$ cuts the 45° line once from above. To understand this, consider some $\omega > \omega^f(s)$ and suppose, to the contrary, that $g_s(s, \omega) \geq \omega$. In this case, equations (13) and (14) imply that $\mu(s, \omega^f(s)) > \lambda(s, \omega^f(s))$, which in turn implies $\omega < \omega^*(s) = \omega^f(s)$ from equation (12), a contradiction. A similar argument shows that $g_s(s, \omega) > \omega$ for $\omega < \omega^f(s)$.[18]

The implications of Lemmas 2 and 3 can be illustrated by considering a particular *sample path* of the consumption share generated for a given history of endowment shares $s^T = (s_0, s_1, \ldots, s_T)$. The sample path of the consumption share is constructed iteratively from the policy functions $f(s, \omega)$ and $g_r(s, \omega)$ starting with $x_0 = (s_0, \omega_0)$ as follows: $c_t = f^t(s^t, x_0) := f(s_t, g^t(s^t, x_0))$, where $g^t(s^t, x_0) := g_{s_t}(s_{t-1}, g^{t-1}(s^{t-1}, x_0))$ and $g^0(s_0, x_0) = \omega_0$.

Figure 2 depicts such a sample path in a three-state example and illustrates three important properties.[19] First, the optimal sustainable consumption share fluctuates above and below the

[18] The argument can be extended to the case where the nonnegativity and upper bound constraints bind, and a complete proof of Lemma 2 is provided in the Online Appendix.

[19] The example has $\beta = \delta = 0.975$, $s(1) = 0.5$, $s(2) = 0.625$ and $s(3) = 0.8125$, with probabilities $\pi(1) = 0.5$, $\pi(2) = 0.25$ and $\pi(3) = 0.25$. The first best is $c^*(s) = 0.5$ for each state.

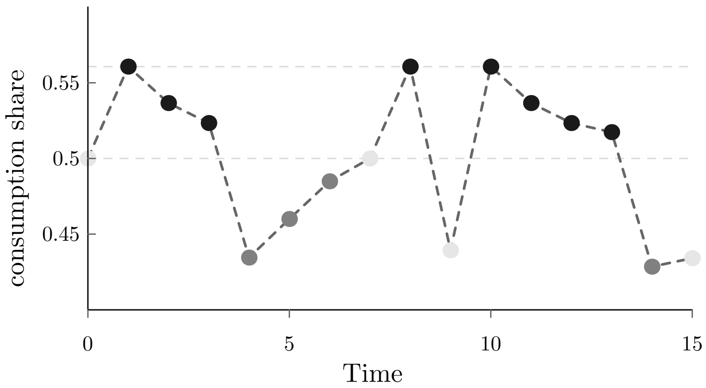

Figure 2: Sample Path of the Young Consumption Share.
Note: The illustration is for a case where $I = 3$ and $\beta = \delta$ (with first-best consumption share of $1/2$). The shade of the dots indicates the state s_t: light gray for $s_t = s(1)$, mid gray for $s_t = s(2)$ and dark gray for $s_t = s(3)$. The case illustrated has $s_0 = s(1)$ and $\omega_0 = \omega^0(1) = -\log(2)$.

first-best level of $1/2$.[20] Second, the path is history dependent. That is, the consumption share varies both with the current endowment state and the history of shocks. For example, the endowment share $s_t = s(3)$ occurs at $t = 8$ and $t = 13$, but the consumption share differs at the two dates. When state 3 occurs, the participation constraint of the young binds, and hence, a higher future utility must be promised to ensure that they are willing to share more of their current endowment. Subsequent realizations of state 3 exacerbate the situation because the young of the next generation must also deliver on past promises, meaning that the consumption share of the young falls when state 3 repeats. This property is evident in Figure 2 where c_t falls when state 3 repeats ($t = 2, 3$ and $t = 11, 12, 13$). This implies that the consumption share is not necessarily monotonic in the endowment. For example, the consumption share at $t = 4$, when the endowment share is $s_4 = s(2)$, is lower than the consumption share at $t = 9$, when the endowment share is $s_9 = s(1) < s(2)$. This non-monotonicity occurs because the promise made to the old for date $t = 4$ is higher than that made for date $t = 9$. Third, there are points in time when the consumption share returns to the same value in the same state. For example, this happens at $t = 7$, which has the same state (state 1) and same consumption as at $t = 0$. In this case, there is *resetting*. The path of the consumption share is the same following resetting if the same sequence of endowment shares occurs. Note that the definition of the resetting points is not unique. For example, there is resetting also at $t = 1, 8, 10$, when state 3 occurs after state 1. Before resetting occurs, the effect of a shock persists. However, once resetting occurs, the history of shocks is forgotten, and the subsequent sample path is identical when the same sequence of states occurs. That is, the sample paths between resettings are probabilistically

[20] By Lemma 1(ii), $\omega^0(s) \leq \omega^*(s)$. By Assumption 3, $\omega^*(1) = \omega_{\min}(1)$. Hence, $\omega^0(1) = \omega^*(1)$. Since $g_1(s, \omega)$ is increasing in ω, the promise is above the first-best level (or equivalently, the consumption share is below the first-best level) in state 1. From Lemma 2(iii), $\omega^0(I) < \omega^*(I)$ and therefore, for low values of ω, the promise is below the first-best level (or equivalently, the consumption share is above the first-best level) in state I.

identical. This property is used in the next section to establish convergence to a unique invariant distribution.

4 Convergence to the Invariant Distribution

This section considers the long-run distribution of the pair $x = (s, \omega)$. It shows that there is a unique and countable ergodic set E with cardinality $|E| > I$ and strong convergence to the corresponding invariant distribution. Let $\Omega = \cup_{r \in \mathcal{I}} \Omega(r)$ and $\mathcal{X} = \mathcal{I} \times \Omega$. The future evolution of x is a Markov chain defined by the transition function $P(x, A \times B) := \Pr\{x_{t+1} \in A \times B \mid x_t = x\} = \sum_{r \in A} \pi(r) \mathbb{1}_B g_r(x)$ where $A \subseteq \mathcal{I}$, $B \subseteq \Omega$ and $\mathbb{1}_B g_r(x) = 1$ if $g_r(x) \in B$ and zero otherwise. The chain starts from $x_0 = (s_0, \omega_0)$ with an initial promise $\omega_0 = \max\{\omega^0(s_0), \bar{\omega}_0\}$.

The monotonicity and resetting properties of Lemma 2 imply that starting from any x_t, a sequence of k consecutive state 1s (where the endowment share is $s(1)$) leads to $x_{t+k} = (1, \omega^0(1))$ for a finite k. This is because $g_1(1, \omega) < \omega$, so that repetition of state 1 leads to a decrease in ω and since $g_1(1, \omega) = \omega^0(1)$ for some $\omega > \omega^0(1)$, ω falls to $\omega^0(1)$ in finite time. In this case, we say that x is *reset* to $(1, \omega^0(1))$ at time $t + k$. Since the probability of the state 1 is $\pi(1) > 0$, the probability of a history of k consecutive state 1s is $\pi(1)^k > 0$. An immediate consequence is that Condition **M** of Stokey et al. (1989, page 348) is satisfied, and hence, there is strong convergence in the uniform metric to a unique invariant probability measure $\phi(X)$ for $X \in \mathcal{X}$.[21]

Since there is a positive probability that x is reset to $(1, \omega^0(1))$ in finite time, the Markov chain for x is *regenerative* and $(1, \omega^0(1))$ is a regeneration point (see, for example, Foss et al., 2018). For simplicity, suppose first that the process starts at $x_0 = (1, \omega^0(1))$. Recall that $g^t(s^t, x_0) = g_{s_t}(s_{t-1}, g^{t-1}(s^{t-1}, x_0))$ where $g^0(1, x_0) = \omega^0(1)$. Let $r_x := \min\{k \geq 1 \mid (s, g^k((s^{k-1}, s), x_0)) = x\}$ denote the first time that the process is equal to x starting from x_0. Then r_{x_0} is the first regeneration time, the first time after the initial period at which x_0 reoccurs. Any sample path of promises can be divided into different blocks, with each block starting at a regeneration time. This can be seen in Figure 2 where the first regeneration time occurs at $t = 7$. Although the blocks between regeneration points are not identical, the strong Markov property ensures that they are i.i.d. At each regeneration time, past shocks are forgotten, and the future evolution of x is probabilistically identical. The regeneration times are also i.i.d. and the expected regeneration time is $\varphi := \mathbb{E}_0[r_{x_0}]$, the same for any block. Moreover, φ is

[21] Condition **M** is satisfied because there is a $k \geq 1$ and an $\epsilon > 0$ such that the k-step transition function $P^k(x, \{(1, \omega^0(1))\}) > \epsilon$ for any $x \in \mathcal{X}$. In this case, $(1, \omega^0(1))$ is an atom of the Markov chain. Açikgöz (2018), Foss et al. (2018), and Zhu (2020) use similar arguments to establish strong convergence in the Aiyagari precautionary savings model with heterogeneous agents.

finite since all positive probability paths must have a sequence of endowment states leading to $x_0 = (1, \omega^0(1))$ as described above.

Now consider a starting point $x_0 = (i, \omega^0(i))$ for some initial state $s_0 = s(i)$. Given that $g_i(1, \omega^0(1)) = \omega^0(i)$ by Lemma 2(iii), a positive probability path that leads back to x_0 is constructed by a sequence of consecutive state 1s, as outlined above, followed by state i. Since the transition from state 1 to state i occurs with positive probability, $(i, \omega^0(i))$ is a regeneration point, and the blocks between these regeneration points are also probabilistically identical. As discussed in Remark 2, in the absence of constraint (4), the planner sets $\omega_0 = \omega^0(i)$ and the process starts in the ergodic set. However, if constraint (4) must be respected and $\bar{\omega}_0 > \omega^0(i)$, then $x_0 = (i, \bar{\omega}_0)$ and the process may start outside of the ergodic set. In this case, there is still a positive probability path back to a resetting point $(i, \omega^0(i))$. The only difference is that the first block in the regenerative process is different from subsequent blocks (which all start from $(i, \omega^0(i))$). However, this does not change the convergence properties of the process.

Let $R_x := \Pr(r_x < \infty)$ be the probability of attaining the pair $x = (s, \omega)$ in finite time starting from x_0. If $R_x > 0$, then x is said to be *accessible* from x_0. Since $x_0 = (i, \omega^0(i))$ has a positive probability mass and the set of endowment states $\mathcal{I}$ is finite and time is discrete, the associated set $E := \{x \mid R_x > 0\}$ is countable. Moreover, the set E is an equivalence class because every $x \in E$ is accessible from x_0, and there is a path from every accessible x back to x_0. Therefore, E is an absorbing set, that is, $P(x, E) = 1$ for all $x \in E$, and since no proper subset of E has this property, it is *ergodic* (see, for example, Stokey et al., 1989, chapter 11). Let φ_x denote the *expected return time* to x where $\varphi_{x_0} \equiv \varphi$. With φ finite, it follows that $R_x = 1$ and each φ_x is finite, that is, each $x \in E$ is positive recurrent.

Since the ergodic set E is countable, standard results on the convergence of positive recurrent Markov chains apply. To state these results, let P denote the transition matrix with elements $P(x, x') = \pi(r)\mathbb{1}_{\omega_r} g_r(x)$ where $x = (s, \omega)$ and $x' = (r, g_r(x))$. Similarly, let $P^k(x, x')$ be the elements of the corresponding k-period transition matrix.

Proposition 5. *(i) There is pointwise convergence to a unique and non-degenerate invariant distribution $\phi = \phi P$ where for each $x \in E$, $\phi(x) = \lim_{k\to\infty} P^k(\cdot, x) = \varphi_x^{-1}$. (ii) The invariant distribution is the limit of the iteration $\phi_{t+1}(x') = \sum_{x\in E} P(x, x')\phi_t(x)$ for any given $\phi_0(x)$. (iii) The cardinality $|E| > I$.*

Parts (i) and (ii) of Proposition 5 are standard and show convergence to a unique invariant distribution where the probability of each $x \in E$ is the inverse of the expected return time. The invariant distribution can be computed iteratively, given knowledge of the policy functions. In particular, for $s_0 = s(i)$, the invariant distribution can be computed starting from an initial distribution $\phi_0(x) = 1$ if $x = (i, \omega^0(i))$ and $\phi_0(x) = 0$ otherwise.[22] Part (iii) shows that the cardinality of the ergodic set is greater than I. That is, at the invariant distribution, there are multiple promised utilities associated with particular states. Hence, the history of endowment shocks affects the consumption allocation even in the long run. This result stands in contrast to the two benchmarks considered in Section 2. If transfers are enforced, or if there is no risk, then convergence is to an ergodic set with a cardinality equal to the cardinality of the set of endowment states.

Since Lemma 2(i) and (ii) show that $g_r(s, \omega)$ is increasing in s and ω, $g_r(I, \omega^*(I))$ is the largest promise that can be reached in state r starting with $x_0 = (i, \omega^0(i))$. If $g_r(I, \omega^*(I)) < \omega_{\max}(r)$, then any $x = (r, \omega)$ with $\omega \in (g_r(I, \omega^*(I)), \omega_{\max}(r))$ is not accessible from x_0. Therefore, such an x is transitory and is not part of the ergodic set. In Section 8, we compute the ergodic set and the invariant distribution in examples with $g_r(I, \omega^*(I)) < \omega_{\max}(r)$.[23]

Remark 4. *The convergence result and all the results of Section 3 apply when preferences exhibit constant relative risk aversion. They also hold for any concave utility function if the aggregate endowment is constant. Lemmas 2 and 3 and Proposition 5 remain valid if there is no growth but the aggregate endowment is state-dependent, except that the policy functions cannot be ordered by the endowment state (see, Lancia et al., 2022, for details).*

5 Debt

In this section, we reinterpret the optimal transfer to the old as debt. Suppose the planner issues one-period state-contingent bonds, which trade at the corresponding state prices. The planner uses the revenue generated by bond sales to fund the transfer to the old, balancing the budget by taxing or subsidizing the young. Given bond prices and taxes, the young buy the correct quantity of state-contingent bonds to finance their optimal old-age consumption. With this interpretation, the dynamics of debt and the fiscal reaction function can be examined.

The Debt Policy Function. It is convenient to measure debt relative to the endowment share of the current young. Then, the optimal debt $d(x)$ satisfies $\omega = \log(1 - s + sd(x))$, so that $d(x)$ is increasing in ω.[24] Let $d^0(s) := d(s, \omega^0(s)) \geq 0$ denote the minimum debt at the optimal solution when the endowment share of the young is s. Debt $d \in \mathcal{D} = [d_{\min}, d_{\max}]$ where the minimum debt $d_{\min} := \min_r d^0(r)$ and the maximum debt $d_{\max}$ is determined as the nontrivial

[22] The convergence results hold for any initial distribution $\phi_0(A)$ even if $A \not\subseteq E$ since eventually, once regeneration occurs, all subsequent promises belong to the ergodic set.

[23] The ergodic set and invariant distribution are difficult to characterize. In some cases, however, the invariant distribution is a transformation of a geometric distribution with a denumerable ergodic set, that is, $|E| = \aleph_0$.

[24] For brevity, in what follows, we often refer to $d(x)$ simply as outstanding debt without the caveat that it is measured relative to the endowment share of the young.

solution of $\log(1-d_{\max}) + \beta\sum_r \pi(r)(\log(1-r+rd_{\max}) - \log(1-r)) = 0$. We refer to $d_{\max}$ as the *debt limit* and $d_{\max} - d$ as the *fiscal space* (see, for example, Ghosh et al., 2013).[25] It follows straightforwardly that $d_{\max} < 1$, analogously to the result of Lemma 1 that $\omega_{\max}(s) < 0$. The *debt policy function* $b_r \colon \mathcal{D} \to \mathcal{D}$ determines the optimal debt next period when the current debt is d and the endowment share of the young next period is r. The properties of the debt policy functions are summarized in the following corollary to Lemmas 2 and 3.

Corollary 1. *(i) The debt policy function* $b_r \colon \mathcal{D} \to \mathcal{D}$ *is continuous in* d *with* $b_r(d) = d^0(r)$ *for* $d \le d^c$ *and* $b_r(d)$ *strictly increasing for* $d > d^c$. *The debt threshold* d^c *satisfies* $d^c = 1 - \exp(-\beta\sum_r \pi(r)(\log(1-r+rd^0(r)) - \log(1-r))) \in (d_{\min}, d_{\max})$ *with* $d_{\min} = 0$ *and* $d_{\max} < 1$. *(ii) For* $d \in \mathcal{D}$, $b_{r(i)}(d) \ge b_{r(i-1)}(d)$ *with strict inequality for at least one* $i = 2, \ldots, I$. *(iii) For each* $r \in \mathcal{I}$, *there is a unique fixed point* $d^f(r)$ *of the mapping* $b_r(d)$ *where* $d^f(r) = \min\{d^*(r), d_{\max}\}$, $d^*(r) = 1 - c^*(r)/r$ *is the first-best debt, and* $d^f(r(I)) > d^c$.

Corollary 1 reveals the benefits of measuring debt relative to the endowment share of the young. First, the debt policy functions depend on the current debt d but are independent of the current endowment share s. Second, there is a common threshold debt d^c, below which the debt policy function is flat and above which it is strictly increasing. For $d \le d^c$, the debt policy function $b_r(d) = d^0(r)$. Lemmas 2 and 3 show why the debt policy function is independent of s. When the participation constraint of the young binds, that is, when constraints (7) and (11) hold as equalities, the policy function for the promised utility $g_r(s,\omega)$ is an increasing function of $\log(s) - \log(1-\exp(\omega))$. With $\exp(\omega) = 1 - s + sd$, $g_r(s,\omega)$ is an increasing function of d, and hence, the debt policy function depends on the current debt and endowment state next period.[26]

Part (i) of Corollary 1 shows that the threshold d^c is determined by setting $b_r(d) = d^0(r)$ for each r. By Assumptions 3, $d_{\min} = 0$ and by Assumption 4, $d^c < d^*(I)$. Part (ii) shows that $b_r(d)$, and consequently, $rb_r(d)$, are increasing in r. Since the consumption share of the old decreases with r, the transfer to the old, $rb_r(d)$, is positively correlated with the marginal utility of consumption of the old. This positive correlation occurs because debt provides partial insurance. Note that the consumption share of the old decreases with r for a given debt d, while it increases with d for a fixed r. Therefore, in comparing two endowment states, the consumption share of the old may be higher when the young have a higher endowment share if

[25] The debt limit is different from the maximum sustainable debt (see, for example, Collard et al., 2015). The maximal sustainable debt focuses on the limit that external investors are willing to lend to a government, taking into account the probability of default. Typically, it is calculated using a fixed rule for government taxes and expenditure and a constant interest rate.

[26] For CRRA preferences with a coefficient of risk aversion greater than one, the same property applies with a different normalization of debt that depends on the coefficient of risk aversion.

the debt is sufficiently high. Part (iii) follows directly from Lemma 2(iv) and the fixed point of the mapping $b_r(d)$ corresponds to the first-best debt.

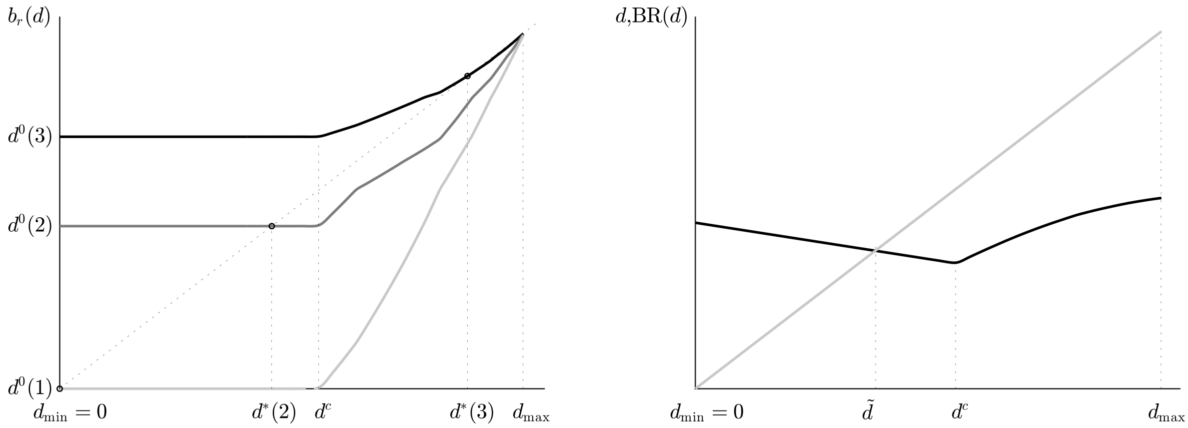

Figure 3: Panel A – Debt Dynamics. Panel B – Fiscal Reaction Function.

Note: The illustration is for the case $I = 3$ corresponding to the example in Figure 2. Panel A plots the optimal debt rule as a function of d. The light gray line is $b_1(d)$, the dark gray line is $b_2(d)$, and the black line is $b_3(d)$. The level $d^*(3)$ is the largest sustainable debt, and $d^*(1)$ is the lowest sustainable debt within the ergodic set. In Panel B, the fiscal reaction function is the difference between BR(d), the dark gray line, and d, the light gray line.

The Dynamics of Debt. The dynamics of debt are derived from the debt policy functions described in Corollary 1 and the history of endowment shares. Panel A of Figure 3 plots the debt policy functions corresponding to the same three-state example illustrated in Figure 2. For $d \leq d^c$, the debt policy function is independent of the current debt d and depends only on the endowment share of the young next period. In particular, $d^0(1) = d^*(1)$ and $d^0(2) = d^*(2)$, so that the consumption share is first best in states 1 and 2, whereas in state 3, the participation constraint binds, limiting the transfer from the young and hence, $d^0(3) < d^*(3)$. For $d > d^c$, debt falls when the endowment share of the young next period is $r(1)$ or $r(2)$. If, for example, there are enough consecutive occurrences of the endowment state 1, then debt falls to zero. Since such sequences occur with positive probability, debt is reset to zero periodically. If, on the other hand, the endowment share of the young next period is $r(3)$, then the debt rises for $d < d^*(3)$ but falls for $d > d^*(3)$. Thus, any debt $d > d^*(3)$ is transitory and cannot occur in the long run.[27] In summary, debt will rise or fall depending on the endowment share of the young next period and the current debt that encapsulates the history of endowment shares.

Fiscal Reaction Function. The fiscal reaction function shows how the tax rate depends on debt. Since the promised utility and debt are monotonically related, we abuse notation and rewrite the state space as $x = (s, d)$. With logarithmic preferences, the intertemporal marginal rate of substitution is $m(x, x') = \beta s(1 - d)/(1 - r(1 - b_r(d)))$, where $x = (s, d)$ is the current state and $x' = (r, b_r(d))$ is the next-period state. Since the endowment shares are i.i.d., the transition probability is $\pi(x, x') = \pi(r)$ and given debt d, the current young can be thought as

[27] In general, if $d^*(I) < d_{\max}$, then any $d \in [d^*(I), d_{\max})$ is transitory.

buying $rb_r(d)$ bonds contingent on a next-period endowment share of r at the state price of $q(x,x') = \pi(r)m(x,x')$. The bond purchases generate a bond revenue for the planner given by:

$$\mathrm{BR}(d) := \left(\frac{1}{s}\right)\sum\nolimits_r q(x,x')rb_r(d) = \beta\sum\nolimits_{r\in\mathcal{I}}\pi(r)\left(\frac{1-d}{1-r(1-b_r(d))}\right)rb_r(d).$$

The planner finances the current debt d by a combination of taxes (or subsidies) on the young and bond revenue $\mathrm{BR}(d)$. Hence, the budget constraint of the planner is:

$$\tau(d) = d - \mathrm{BR}(d), \tag{16}$$

where $\tau(d)$ is the tax rate on the young, measured as a share of their endowment. We refer to $\tau(d)$ as the *fiscal reaction function* and $s\tau(d)$ as the *primary fiscal balance*. A positive value of $s\tau(d)$ corresponds to a primary fiscal surplus, whereas a negative value of $s\tau(d)$ corresponds to a primary fiscal deficit.

Panel B of Figure 3 plots the outstanding debt d and the bond revenue $\mathrm{BR}(d)$ with the fiscal reaction function $\tau(d)$ given by the difference between the two lines. The properties of $\mathrm{BR}(d)$ are complex because $b_r(d)$ is increasing in d whereas the state price $q((s,d),(r,b_r(d)))$ is decreasing in both d and b_r. By Proposition 2, there are transfers next-period for any debt $d < d_{\max}$ and hence, $\mathrm{BR}(0)$ is strictly positive. Moreover, since $b_r(d)$ is constant for debt below the threshold level, $\mathrm{BR}(d)$ decreases linearly in d for $d < d^c$. Hence, the fiscal reaction function $\tau(d)$ increases linearly in d for $d < d^c$. There is an intersection point $\tilde{d}$ where $\tau(\tilde{d}) = 0$. For $d < \tilde{d}$, bond revenue exceeds the current debt, and the planner subsidizes the young, that is, there is a primary fiscal deficit. For $d > \tilde{d}$, bond revenue is insufficient to cover the current debt, and the planner taxes the young, that is, there is a primary fiscal surplus. For $d > d^c$, a rise in d, that is, a reduction in the fiscal space, leads to more bond issuance but the price of the bonds decreases. Thus, the net effect of a change in d on bond revenue is generally ambiguous. For the example illustrated in Panel B, the fiscal reaction function $\tau(d)$ is increasing in d but initially at a slower rate for debt above the threshold level and then at a higher rate when debt is large.

The situation depicted in Figure 3 contrasts with the two benchmarks discussed in Section 2. At the first best, the debt policy function is $b_r(d) = d^*(r)$ independent of d. Hence, the debt policy functions in Panel A of Figure 3 are horizontal lines with fixed points at $d^*(r)$. There are no dynamics of debt except in the initial period, although debt varies with the endowment share. Ignoring the nonnegativity constraint on transfers, the first-best bond revenue function is linearly decreasing in debt, resulting in a fiscal reaction function that is linearly increasing.[28]

[28] It can be shown that $\mathrm{BR}^*(d) = (a-1)(1-d)$ where $a = (1-\delta) + (\beta+\delta)\mathbb{E}_s s$ and $\mathbb{E}_s s$ is the expected endowment share. Hence, the fiscal reaction function is $\tau^*(d) = (1-a) + ad$. Since $\mathbb{E}_s s > \delta/(\beta+\delta)$, $a > 1$.

In the deterministic case, the debt policy function is a transformation of the policy function in Figure 1 with a critical debt $d^c = (\exp(\omega^c) - (1-s))/s$. If the initial debt is above d^c, debt falls, and once it reaches or falls below d^c, the debt next period equals the first-best level d^*. The dynamics of debt are transitory, with debt reaching the fixed point d^* in finite time. Along the transition path, debt falls, and the price of debt rises. These two offsetting effects make it possible for bond revenue to rise or fall during the transition.

The two benchmarks show that enforcement frictions lead to the nonlinearity of the fiscal reaction function. By showing how this arises within an optimizing framework, the paper contributes to the literature that examines and provides evidence of this nonlinearity (see, for example Mendoza and Ostry, 2008, Ghosh et al., 2013, among others).

6 Asset Pricing Implications

In this section, we examine the asset pricing implications of the model.[29] In an overlapping generations model, the growth-adjusted stochastic discount factor is given by the intertemporal marginal rate of substitution $m(x,x') := \beta u_c(1-c(x'))/u_c(c(x))$ where x is the current state, x' is the state next period, $u_c(c(x))$ is the marginal utility of the current young and $u_c(1-c(x'))$ is their marginal utility when old. This stochastic discount factor can be decomposed into two components:

$$m(x,x') = \underbrace{\delta\left(\frac{u_c(c(x'))}{u_c(c(x))}\right)}_{m_A(x,x')}\underbrace{\left(\frac{\beta}{\delta}\frac{u_c(1-c(x'))}{u_c(c(x'))}\right)}_{m_B(x,x')}. \tag{17}$$

The first component $m_A(x,x')$ represents risk sharing *across* two adjacent generations of the young and the second component $m_B(x,x')$ represents risk sharing *between* the young and the old at a given date. In a representative agent model, $m(x,x') = m_A(x,x')$ and the variability in the stochastic discount factor is determined by the variability of consumption, which in an endowment economy depends on the variability of the aggregate endowment. In contrast, in an overlapping generations model, if there is variability in the degree of risk sharing between the young and the old, then there is variability in $m_B(x,x')$, which interacts with the variability in $m_A(x,x')$ with consequent implications for asset pricing. In the optimal sustainable intergenerational insurance, the variability of $m_B(x,x')$ is determined by the first-order condition (12) and the updating rule (15). This variability depends on the current endowment share and the

[29] We follow several authors in addressing asset pricing in overlapping generations models (see, for example, Huberman, 1984, Huffman, 1986, Labadie, 1986) and Gârleanu and Panageas (2023) for a recent contribution.

outstanding debt. To simplify the discussion, we confine attention to states in the ergodic set.[30] We also assume that the bounds on debt do not bind. In this case, the first best exhibits complete insurance with the consumption share fully independent of the endowment state.[31]

Let Q denote the matrix of state prices $q(x,x') = \pi(r)m(x,x')$ where $x = (s,d)$ and $x' = (r, b_r(d))$, and let ϱ and ψ be the Perron root and corresponding eigenvector of Q. The Ross Recovery Theorem (Ross, 2015) shows that the k-period stochastic discount factor $m^k(x,x') = \varrho^k \psi(x)/\psi(x')$ where ϱ and $\psi(x)$ can be interpreted as the discount factor and inverse marginal utility of a pseudo-representative agent. Using the first-order condition (12) and the updating rule (15), $f(x')/(1-f(x')) = (\delta/\beta)(1+\mu(x'))/(1+\mu(x))$ where $f(x) = s(1-d)$ is the consumption share of the young and $\mu(x)$ is the multiplier on the corresponding participation constraint. To ease notation, let $\nu(x) := 1 + \mu(x)$ and $\nu_{\max} := \max_x \nu(x)$. Since we show below that $\varrho = \delta$, it follows from equation (17) that $\psi(x) = f(x)/\nu(x)$.[32] The unit price of a k-period discount bond in state x, $p^k(x)$, is given by the corresponding row sum of Q^k, the k-fold matrix power of Q. The corresponding yield is $y^k(x) := -(1/k)\log(p^k(x))$ and the yield on the long bond is $y^\infty(x) := \lim_{k\to\infty} y^k(x)$.

Martin and Ross (2019) shows that $|y^k(x) - y^\infty(x)| \leq (1/k)\Upsilon$ for $\Upsilon := \log(\psi_{\max}/\psi_{\min})$ where $\psi_{\max}$ and $\psi_{\min}$ are the maximum and minimum values of ψ. That is, Υ measures the range of the eigenvector and bounds the deviation of the yield from its long-run value. A low value of Υ means that the yield curve is relatively flat and that yields are not very sensitive to debt.[33]

The matrix Q is the growth-adjusted or de-trended state price matrix. Let $q^k_+(x,x')$ and $m^k_+(x,x')$ denote the unadjusted state prices and marginal rate of substitution conditional on state x when the state k-periods ahead is x' and the growth factor is γ. It can be checked that $q^k_+(x,x') = \varsigma(\gamma)\bar{\gamma}^{-k}(\bar{\gamma}/\gamma)q^k(x,x')$ and $m^k_+(x,x') = \bar{\gamma}^{-k}(\bar{\gamma}/\gamma)m^k(x,x')$ where $q^k(x,x') =$

[30] Limiting the analysis to the ergodic set is justified for two reasons. First, there is convergence to the ergodic set within finite time, as shown in Section 4. Second, absent constraint (4), the planner sets the initial debt to $d_{\min}$, which lies in the ergodic set. Furthermore, for simplicity and because it corresponds to our numerical procedures, we assume that the ergodic set is finite. Nevertheless, it is possible to adapt the arguments to the denumerable case or even more general state spaces (see, for example, Hansen and Scheinkman, 2009, Christensen, 2017).

[31] Although it is restrictive to assume that the bounds on debt are non-binding, it simplifies the analysis, and we will note how results differ when the bounds bind.

[32] The multiplicative decomposition of $\psi(x)$ into the components $f(x)$ and $1/\nu(x)$ is reminiscent of a number of other asset pricing models (see, for example, Bansal and Lehmann, 1997).

[33] The bound Υ provides a measure of the variability of the yields. Two alternative measures used to assess how risk is shared are the insurance coefficient (see, for example, Kaplan and Violante, 2010) and a consumption equivalent welfare change (see, for example, Song et al., 2015). We discuss these alternatives in Part C of the Supplementary Appendix and show that these two measures share similar comparative static properties with the bound Υ.

$\pi^k(x,x')m^k(x,x')$.[34] Letting $y_+^k(x)$ be the yield on the k-period bond in the unadjusted case, we can establish the following proposition.

Proposition 6. *For each $x \in E$: (i) The yield on a k-period bond $y_+^k(x) = y^k(x) + \log(\bar{\gamma})$. (ii) The yield on the long bond $y_+^\infty(x) = y^\infty + \log(\bar{\gamma})$ with $y^\infty = -\log(\delta)$. (iii) The yield $y^k(x)$ is increasing in d for each s and k. (iv) The long-short spreads on yields satisfy $y^\infty - y^1(1, d^*(1)) > 0 > y^\infty - y^1(I, d^*(I))$. (v) The measure $\Upsilon = \log(\nu_{\max})$ where $\nu_{\max} = \nu(I, d^*(I))$.*

Part (i) of Proposition 6 shows that the difference between the yields in the growth-adjusted and unadjusted cases is simply the average growth rate as measured by $\log(\bar{\gamma})$, independently of the current state x or the time horizon k. This independence follows from Assumption 1 that the growth shocks are i.i.d., meaning that each generation faces the same growth risk. A similar result, that market risk premia are unaffected by market incompleteness, is established by Krueger and Lustig (2010) in a model with infinitely-lived agents and uninsurable idiosyncratic as well as aggregate risk. Part (ii) follows from the result of Martin and Ross (2019) that $y^\infty = -\log(\varrho)$, independently of x and the fact that $\varrho = \delta$ if the upper bound and nonnegativity constraints do not bind.[35] To understand Part (iii), note that the consumption share of the young is decreasing in d and that, since $b_r(d)$ is increasing in d from Corollary 1, the consumption share of the old next period is increasing in d. Consequently, the stochastic discount factor $m(x,x')$ decreases in d. Since the transition probabilities do not depend on d, the price of the one-period discount bond is decreasing in d, or, equivalently, its yield is increasing in d. Thus, an agent born into a generation with higher debt faces higher one-period yields. Since bond prices are linked recursively, this property holds for bonds of any maturity.

Part (iv) of Proposition 6 shows that the long-short spread $y^\infty - y^1(x)$ is positive when the young have a low endowment share and the debt is low. In this case, it follows from Section 5 that debt is likely to increase in the future with a corresponding increase in yields. Conversely, the spread is negative when the young have a high endowment share, and the debt is high, in which case, both debt and yields are likely to fall in the future. Part (v) shows that Υ is determined by the multiplier $\nu_{\max}$ on the participation constraint corresponding to the fixed point of the debt policy function when the endowment share of the young is the largest. That is, the bound on the variability of the yield curve is determined by the tightness of the participation constraint at the largest debt in the ergodic set.

---

[34] With stochastic growth, the Ross Recovery Theorem does not recover the true probability transition matrix $\pi^k(x,x')$. Instead, it recovers a transition matrix where probabilities are weighted by the relative growth factors (see, for example, Hansen and Scheinkman, 2012).

[35] If the upper bound constraint does not bind, then $\varrho \leq \delta$ and if the nonnegativity constraints do not bind, then $\varrho \geq \delta$.

To help understand the results of Proposition 6, consider the first-best and deterministic benchmarks of Section 2. In the first best, the debt policy functions are constants, and the yield curve is flat with $y^k_+(x) = -\log(\delta) + \log(\bar{\gamma})$ and $\Upsilon = 0$. Despite the flat yield curve, the risk premium on the aggregate risk is positive because the return on debt is high when the growth rate is high. Specifically, the expected return on a one-period bond is $\mathbb{E}_\gamma \gamma/\delta$, while the risk-free rate is $\bar{\gamma}/\delta$. Thus, the risk premium is $(\mathbb{E}_\gamma \gamma - \bar{\gamma})/\delta$, which is strictly positive when the growth shocks are non-degenerate. A Lucas tree or any other asset that pays a share of the aggregate endowment will carry this positive risk premium, so that the risk premium on aggregate risk corresponds to the risk premium on debt with complete insurance. In the deterministic case, the risk premium is zero. However, along the transition path, as debt falls, the yield $y^k(d)$ decreases to its long-run value of $y^\infty = -\log(\delta) + \log(\gamma)$, where γ is the deterministic growth rate. Thus, $\Upsilon > 0$ in the transition, even though there is no risk.[36]

7 Debt Valuation

The budget constraint in equation (16) can be iterated forward to show that current debt equals the present value of all future primary surpluses.[37] As pointed out by Bohn (1995), this present value depends on the risk premium on debt. In this section, we focus on the multiplicative risk premium on debt because it is the negative of the covariance of the stochastic discount factor and the return on debt and because it is independent of the endowment share. When there is a growth shock γ, the return on debt is $R_+(x, x') = rb_r(d)\gamma e/(s\text{BR}(d)e)$, where $s\text{BR}(d)e$ is the value of bonds issued today. The multiplicative risk premium is $\text{MRP}_+(d) = (\bar{R}_+(x) - R^f_+(x))/R^f_+(x)$ where $\bar{R}_+(x)$ is the expected return on debt and $R^f_+(x)$ is the risk-free rate on interest in state x. Denote the corresponding growth-adjusted values by $\text{MRP}(d)$, $\bar{R}(x)$ and $R^f(x)$. As we have shown in Section 6, the risk premium on debt with complete insurance equals the risk premium on aggregate risk and we denote the common multiplicative risk premium by MRP^*. The following proposition shows that the multiplicative risk premium has a linear decomposition that depends on the growth-adjusted multiplicative risk premium and the multiplicative risk premium with complete insurance.

Proposition 7. *The multiplicative risk premium* $MRP_+(d) = MRP(d) + \alpha(d)MRP^*$, *where* $\alpha(d) = \bar{R}(x)/R^f(x)$. *The components satisfy: (i)* $MRP^* = (\mathbb{E}_\gamma \gamma - \bar{\gamma})/\bar{\gamma} \geq 0$; *(ii)* $MRP(d) < 0$; *and (iii)* $0 < \alpha(d) < 1$.

[36] The ergodic set is degenerate at d^* in the deterministic case. Once debt reaches this level, the yield curve is flat.

[37] Jiang et al. (2023) define *fiscal capacity* as the present value of future surpluses. Since, in our model, debt is determined optimally, there is no mispricing or bubble component, and debt and fiscal capacity are equivalent in this sense. Other authors often use the term fiscal capacity more broadly to encompass both the debt limit and fiscal space.

The decomposition of $\text{MRP}_+(d)$ into components depending on $\text{MRP}(d)$ and MRP^* is analogous to the result of Proposition 6 that the conditional yield is the sum of a growth-adjusted yield and a component corresponding to the average growth rate, and similarly follows from Assumption 1 that the shocks to growth and the endowment share are independent of each other and i.i.d. Part (i) of Proposition 7 shows that MRP^* is nonnegative. As discussed in Section 6, MRP^* is strictly positive if the growth shocks are non-degenerate. To understand Part (ii), note that the return $R(x, x') = rb_r(d)/(s\text{BR}(d))$ is increasing in r, from Part (ii) of Corollary 1. Moreover, the consumption share of the old is decreasing in r, from Lemma 2, and hence, the stochastic discount factor $m(x, x')$ is increasing in r. Thus, the returns are high when the marginal utility of consumption of the old is high, resulting in a positive covariance term and, correspondingly, a negative growth-adjusted multiplicative risk premium. By comparison, with complete insurance, the stochastic discount factor is constant so that its covariance with the returns is zero, and hence, $\text{MRP}(d) = 0$. As noted in equation (17), the stochastic discount factor comprises two components that measure risk sharing across two adjacent generations of the young and risk sharing between the young and the old. The first component $m_A(x, x')$ is decreasing in r, whereas the second component $m_B(x, x')$ is increasing in r. In a representative agent model, only $m_A(x, x')$ is present, and high debt returns are associated with a low marginal utility of consumption of the young, generating a positive risk premium. In contrast, $m_B(x, x')$ dominates in the overlapping generations model, making debt a negative beta asset.

Part (iii) of Proposition 7 shows that $\alpha(d) < 1$, and hence, the gap $\text{MRP}^* - \text{MRP}_+(d) > 0$ for each d. That is, the multiplicative risk premium on debt is lower than the multiplicative risk premium on aggregate risk. Using U.S. data, Jiang et al. (2021) show that the observed value of debt is higher than the present value of future primary surpluses when discounted using the risk premium on aggregate risk, a debt valuation puzzle. Convenience yields, seigniorage and other service flow values have been offered as potential explanations for this puzzle. Our results suggest an additional explanation. In the presence of enforcement frictions, risk sharing is partial and debt serves as a hedge against idiosyncratic risk, lowering the risk premium and raising the value of debt.[38]

Part (iii) of Proposition 7 also shows that the gap $\text{MRP}^* - \text{MRP}_+(d)$ depends on d, evolving according to the dynamics of debt outlined in Section 5. For $d \leq d^c$, this gap is independent of d. For $d > d^c$, the effect of debt on the size of the gap is ambiguous. From Proposition 6, the risk-free interest rate increases with debt. Therefore, the gap rises or falls depending on whether

[38] Jiang et al. (2022) examine how to manufacture risk-free government debt. With the primary surplus disaggregated into tax and expenditure components, the risk premium on debt is a weighted average of the risk premiums on taxes and expenditure. Consequently, the risk premium on debt can be eliminated, but only at the cost of making taxes or expenditures less cyclical. Since we do not distinguish between taxes and expenditure, the risk premium on the primary surplus equals the risk premium on debt, and making debt risk-free may not be feasible or desirable.

the expected return on debt increases with debt at a faster or slower rate than the risk-free interest rate. Although the overall effect is ambiguous, in the example of Section 8, $\text{MRP}^* - \text{MRP}_+(d)$ decreases with d for $d > d^c$.

8 Two-State Example

Finding the optimal sustainable intergenerational insurance is complex because it involves solving the functional equation of problem P1. In this section, we present an example with $I = 2$ that can be solved using a simple shooting algorithm.[39] For this case, we solve for the invariant distribution and derive a closed-form solution for the Martin-Ross measure Υ.

Suppose there are two possible endowment shares for the young: $s(1) = \kappa - \epsilon(1-\pi)/\pi$ and $s(2) = \kappa + \epsilon$, where $\pi = \pi(1)$, $\kappa \geq 1/2$ and $\epsilon > 0$. The young are poor in state 1 and rich in state 2, and an increase in ϵ is a mean-preserving spread of the risk. By Assumptions 3 and 4, $d^*(2) > d^0(2) > d^c > d^0(1) = d^*(1) = 0$. By Corollary 1, the debt policy functions satisfy $b_2(d) > b_1(d)$. We make two additional assumptions.

Assumption 5. *(i) $d^*(2) < d_{\max}$; and (ii) $b_1(d^*(2)) < d^c$.*

Part (i) of Assumption 5 implies that the upper debt limit never binds. By Part (ii), debt is below d^c whenever state 1 occurs. In such a case, the history of endowment states is forgotten once state 1 occurs and the dynamics of debt depend only on the number of consecutive state 2s in the most recent history, starting from the resetting level $d^0(r)$. The longer is the sequence of state 2s, the larger is the level of debt, approaching $d^*(2)$ if state 2 is repeated infinitely often. The set of parameter values that satisfy Assumption 5, as well as Assumptions 2-4, is nonempty with the following belonging to this set.

Example 1. *$\delta = \beta = \exp(-1/75)$, $\pi = 1/2$, $\kappa = 3/5$, and $\epsilon = 1/10$.*

To simplify notation, let $d^{(n)}(s)$ be the debt in state s after n consecutive state 2s, where $d^{(0)}(s) = d^0(s)$ are the resetting levels and $\lim_{n\to\infty} d^{(n)}(2) = d^*(2)$. Under Assumption 5, the invariant distribution of debt is a transformation of a geometric distribution and the bound Υ has a closed-form solution.

Proposition 8. *Under Assumption 5: (i) The ergodic set $E = \{(s, d^{(n)}(s))_{n\geq 0, s=1,2}\}$ with a probability mass function $\phi(s, d^{(n)}(s)) = \phi(s, d^0(s))(1-\pi)^n$ for $n \geq 1$ where $\phi(1, d^0(1)) = \pi^2$*

[39] Part E of the Supplementary Appendix provides details of the shooting algorithm.

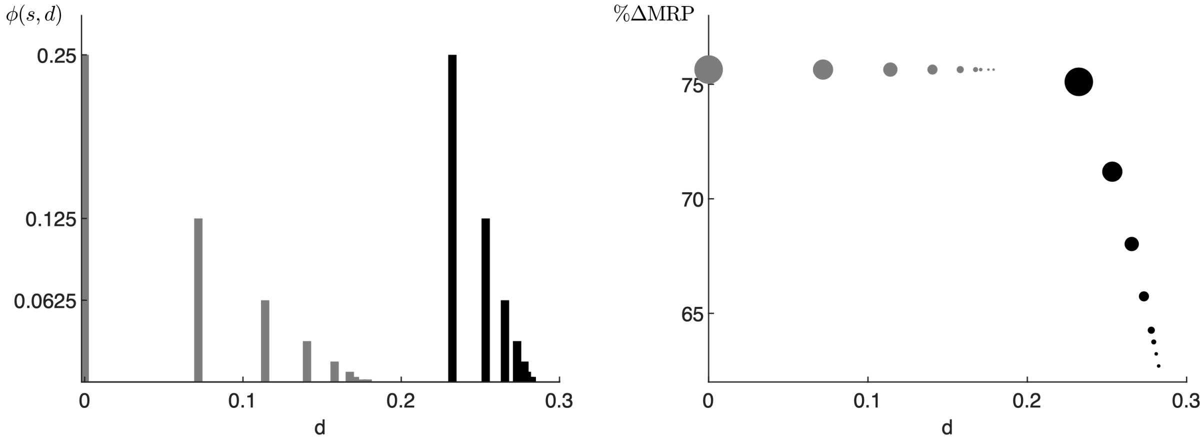

Figure 4: Panel A – Invariant Distribution Panel B – Multiplicative Risk Premium.

Note: Panel A plots the invariant distribution $\phi(s, d^{(n)}(s))$ for $s = 1$ (light gray bars) and $s = 2$ (dark gray bars) for the parameters of Example 1. Panel B plots $\%\Delta\text{MRP}(d) = (\text{MRP}^* - \text{MRP}_+(d))/\text{MRP}^*$ for the values of d in the ergodic set for the parameters of Example 1. Light gray dots correspond to state 1 and the dark gray dots to state 2. The size of each dot indicates the frequency of occurrence.

and $\phi(2, d^0(2)) = \pi(1-\pi)$. *(ii)* $\Upsilon = \log(\delta/\beta) - \log(\chi^{-1} - 1)$ *where*

$$\chi = \left(\frac{\delta}{\beta}\right)^{\frac{1-\pi}{\pi}} \left(\frac{\beta+\delta}{\delta}\right)^{\frac{1+\beta(1-\pi)}{\beta\pi}} (\kappa+\epsilon)^{\frac{1}{\beta\pi}} (1-\kappa-\epsilon)^{\frac{1-\pi}{\pi}} \left(1-\kappa+\epsilon\frac{1-\pi}{\pi}\right).$$

Since debt is reset to $d^0(s)$ after an occurrence of state 1, the invariant distribution of the pair (s, d) depends only on the number of consecutive state 2s. Therefore, the invariant distribution is a transformation of a geometric distribution. As stated in Part (i) of Proposition 8, the invariant distribution has a probability mass of $\phi(1, d^0(1)) = \pi^2$ and $\phi(2, d^0(2)) = \pi(1-\pi)$ at the regeneration states and zero probability mass at states $(s, b_s(d^*(2)))$. Furthermore, low debt levels occur only in state 1, while high levels occur only in state 2. Panel A of Figure 4 plots the invariant distribution for the parameter values of Example 1.

Part (ii) of Proposition 8 provides a closed-form solution for the bound Υ. By Proposition 6, the bound is strictly positive and determined by the tightness of the participation constraint of the young when $x = (2, d^*(2))$. Using this closed-form solution, it is easily checked that Υ decreases with the discount factors β or δ, that is, as either the agent or the planner becomes more patient. Moreover, Υ decreases with the average endowment share to the young, κ, and increases with risk, ϵ.[40]

Panel B of Figure 4 illustrates the impact of debt on the risk premium in a version of Example 1 with stochastic growth. In this example, the arithmetic mean growth rate is set to 4%

[40] Part C of the Supplementary Appendix presents the comparative static properties of Υ even for parameter values that violate Assumption 5.

and the corresponding multiplicative risk premium is approximately 5%. Proposition 7 shows that $\mathrm{MRP}^* > \mathrm{MRP}_+(d)$ and Panel B illustrates that the gap is constant when debt is low, but decreases with debt when debt is high. As noted in Section 7, the multiplicative risk premium may increase or decrease with debt for $d > d^c$, depending on the relative magnitude of the effect of debt on its return and the marginal utility of consumption of the old. In this example, the effect on the return dominates causing the risk premium to rise with debt. Since the risk premium on aggregate risk is independent of debt, a rise in debt narrows the gap between the risk premiums on aggregate risk and debt.

9 Conclusion

The paper has developed a theory of intergenerational insurance in a stochastic overlapping generations model with limited enforcement of risk-sharing transfers. Despite the stationarity of the underlying economic environment, the generational risk is spread across future generations in ways that cause transfers to be history dependent. There is periodic resetting, and the history of shocks is forgotten when this occurs. By interpreting intergenerational insurance in terms of debt, we provide a theory of the dynamics of debt that offers a new perspective on the fiscal reaction function and the sustainability and valuation of debt. With complete insurance, the fiscal reaction function is linear, and the risk premium on debt equals the risk premium on aggregate risk. When there are enforcement frictions, intergenerational insurance is incomplete, the fiscal reaction function is nonlinear, and the risk premium on debt is below the risk premium on aggregate risk.

The results suggest several potential directions for future research. First, the qualitative predictions about the dynamics of debt could be compared with historical data for advanced economies, for example, with a specific focus on the baby boom and subsequent generations. Second, the model has no heterogeneity within a generation. Enriching the demographic structure of the model, either by having more than two overlapping generations or allowing for heterogeneity within the same generation, would make it possible to address the interdependence between intergenerational and intragenerational insurance. Third, to study the interplay between self-insurance and intergenerational insurance, a technology that can transform endowments across dates could be added. Finally, incorporating a stochastic demand for public good provision would allow the study of the risk premia associated with the various components of the primary surplus.

References

Açikgöz, Ömer (2018) "On the Existence and Uniqueness of Stationary Equilibrium in Bewley Economies with Production," *Journal of Economic Theory*, 173, 18–25, 10.1016/j.jet.2017.10.006.

Aiyagari, S. Rao and Dan Peled (1991) "Dominant Root Characterization of Pareto Optimality and the Existence of Optimal Equilibria in Stochastic Overlapping Generations Models," *Journal of Economic Theory*, 54 (1), 69–83, 10.1016/0022-0531(91)90105-D.

Alvarez, Fernando and Urban J. Jermann (2001) "Quantitative Asset Pricing Implications of Endogenous Solvency Constraints," *Review of Financial Studies*, 14 (4), 1117–1151, 10.1093/rfs/14.4.1117.

Atkeson, Andrew and Robert E. Lucas Jr. (1992) "On Efficient Distribution With Private Information," *The Review of Economic Studies*, 59 (3), 427–453, 10.2307/2297858.

Ball, Laurence and N. Gregory Mankiw (2007) "Intergenerational Risk Sharing in the Spirit of Arrow, Debreu, and Rawls, with Applications to Social Security Design," *Journal of Political Economy*, 115 (4), 523–547, 10.1086/520646.

Bansal, Ravi and Bruce N. Lehmann (1997) "Growth-Optimal Portfolio Restrictions On Asset Pricing Models," *Macroeconomic Dynamics*, 1 (2), 333–354, 10.1017/S1365100597003039.

Bhandari, Anmol, David Evans, Mikhail Golosov, and Thomas J. Sargent (2017) "Fiscal Policy and Debt Management with Incomplete Markets," *The Quarterly Journal of Economics*, 132 (2), 617–663, 10.1093/qje/qjw041.

Bohn, Henning (1995) "The Sustainability of Budget Deficits in a Stochastic Economy," *Journal of Money, Credit and Banking*, 27 (1), 257–271, 10.2307/2077862.

——— (1998) "The Behavior of U. S. Public Debt and Deficits," *The Quarterly Journal of Economics*, 113 (3), 949–963, 10.1162/003355398555793.

Broer, Tobias (2013) "The Wrong Shape of Insurance? What Cross-Sectional Distributions Tell Us about Models of Consumption Smoothing," *American Economic Journal: Macroeconomics*, 5 (4), 107–140, 10.1257/mac.5.4.107.

Brunnermeier, Markus K, Sebastian A Merkel, and Yuliy Sannikov (2023) "Safe Assets: A Dynamic Trading Perspective," *Journal of Political Economy*, forthcoming.

Chari, V. V. and Patrick J. Kehoe (1990) "Sustainable Plans," *Journal of Political Economy*, 98 (4), 783–802, 10.1086/261706.

Chattopadhyay, Subir and Piero Gottardi (1999) "Stochastic OLG Models, Market Structure, and Optimality," *Journal of Economic Theory*, 89 (1), 21–67, 10.1006/jeth.1999.2543.

Christensen, Timothy M. (2017) "Nonparametric Stochastic Discount Factor Decomposition," *Econometrica*, 85 (5), 1501–1536, 10.3982/ECTA11600.

Collard, Fabrice, Michel Habib, and Jean-Charles Rochet (2015) "Sovereign Debt Sustainability in Advanced Economies," *Journal of the European Economic Association*, 13 (3), 381–420, 10.1111/jeea.12135.

Diamond, Peter (1977) "A Framework for Social Security Analysis," *Journal of Public Economics*, 8 (3), 275–298, 10.1016/0047-2727(77)90002-0.

Farhi, Emmanuel and Iván Werning (2007) "Inequality and Social Discounting," *Journal of Political Economy*, 115 (3), 365–402, 10.1086/518741.

Foss, Sergey, Vsevolod Shneer, Jonathan P. Thomas, and Tim Worrall (2018) "Stochastic Stability of Monotone Economies in Regenerative Environments," *Journal of Economic Theory*, 173, 334–360, 10.1016/j.jet.2017.11.004.

Gale, David (1973) "Pure Exchange Equilibrium of Dynamic Economic Models," *Journal of Economic Theory*, 6 (1), 12–36, 10.1016/0022-0531(73)90041-0.

Gârleanu, Nicolae and Stavros Panageas (2023) "Heterogeneity and Asset Prices: An Intergenerational Approach," *Journal of Political Economy*, 131 (4), 839–876, 10.1086/722224.

Ghosh, Atish R., Jun I. Kim, Enrique G. Mendoza, Jonathan D. Ostry, and Mahvash S. Qureshi (2013) "Fiscal Fatigue, Fiscal Space and Debt Sustainability in Advanced Economies," *The Economic Journal*, 123 (566), F4–F30, 10.1111/ecoj.12010.

Glover, Andrew, Jonathan Heathcote, Dirk Krueger, and José-Víctor Ríos-Rull (2020) "Intergenerational Redistribution in the Great Recession," *Journal of Political Economy*, 128 (10), 3730–3778, 10.1086/708820.

——— (2023) "Health versus Wealth: On the Distributional Effects of Controlling a Pandemic," *Journal of Monetary Economics*, 140, 34–59, 10.1016/j.jmoneco.2023.07.003.

Gordon, Roger H. and Hal R. Varian (1988) "Intergenerational Risk Sharing," *Journal of Public Economics*, 37 (2), 185–202, 10.1016/0047-2727(88)90070-9.

Green, Edward J (1987) "Lending and Smoothing of Uninsurable Income," in Prescott, Edward C. and Neil Wallace eds. *Contractual Arrangements for Intertemporal Trade*, 1 of Minnesota Studies in Macroeconomics, Chap. 1, 3–25: University of Minnesota Press, Minneapolis.

Hall, Robert E. (1978) "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy*, 86 (6), 971–987, 10.1086/260724.

Hansen, Lars Peter and José A. Scheinkman (2009) "Long-Term Risk: An Operator Approach," *Econometrica*, 77 (1), 177–234, 10.3982/ECTA6761.

——— (2012) "Recursive Utility in a Markov Environment with Stochastic Growth," *Proceedings of the National Academy of Sciences*, 109 (30), 11967–11972, 10.1073/pnas.1200237109.

Huberman, Gur (1984) "Capital Asset Pricing in an Overlapping Generations Model," *Journal of Economic Theory*, 33 (2), 232–248, 10.1016/0022-0531(84)90088-7.

Huffman, Gregory W. (1986) "The Representative Agent, Overlapping Generations, and Asset Pricing," *The Canadian Journal of Economics*, 19 (3), 511–521, 10.2307/135344.

Jiang, Zhengyang, Hanno Lustig, Stijn Van Nieuwerburgh, and Mindy Z. Xiaolan (2021) "The U.S. Public Debt Valuation Puzzle," Working Paper 26583, National Bureau of Economic Research, 10.3386/w26583, Original version December 2019.

——— (2022) "Manufacturing Risk-free Government Debt," Working Paper 27786, National Bureau of Economic Research, 10.3386/w27786, Original version September 2020.

——— (2023) "Fiscal Capacity: An Asset Pricing Perspective," *Annual Review of Financial Economics*, 15 (1), 197–219, 10.1146/annurev-financial-110921-103651.

Kaplan, Greg and Giovanni L. Violante (2010) "How Much Consumption Insurance beyond Self-Insurance?" *American Economic Journal: Macroeconomics*, 2 (4), 53–87, 10.1257/mac.2.4.53.

Kocherlakota, Narayana R. (1996) "Implications of Efficient Risk Sharing without Commitment," *Review of Economic Studies*, 63 (4), 595–609, 10.2307/2297795.

Krueger, Dirk and Hanno Lustig (2010) "When is Market Incompleteness Irrelevant for the Price of Aggregate Risk (and When is it Not)?" *Journal of Economic Theory*, 145 (1), 1–41, 10.1016/j.jet.2009.10.005.

Krueger, Dirk and Fabrizio Perri (2011) "Public versus Private Risk Sharing," *Journal of Economic Theory*, 146 (3), 920–956, 10.1016/j.jet.2010.08.013.

Labadie, Pamela (1986) "Comparative Dynamics and Risk Premia in an Overlapping Generations Model," *Review of Economic Studies*, 53 (1), 139–152, 10.2307/2297597.

Lancia, Francesco, Alessia Russo, and Tim Worrall (2022) "Optimal Sustainable Intergenerational Insurance," CEPR Press Discussion Paper No. 15540 (Revised), Centre for Economic Policy Research, London.

Martin, Ian and Stephen Ross (2019) "Notes on the Yield Curve," *Journal of Financial Economics*, 134 (3), 689–702, 10.1016/j.jfineco.2019.04.014.

Mendoza, Enrique G. and Jonathan D. Ostry (2008) "International Evidence on Fiscal Solvency: Is Fiscal Policy "Responsible"?" *Journal of Monetary Economics*, 55 (6), 1081–1093, 10.1016/j.jmoneco.2008.06.003.

Rangel, Antonio and Richard Zeckhauser (2001) "Can Market and Voting Institutions Generate Optimal Intergenerational Risk Sharing?" in Campbell, John Y. and Martin Feldstein eds. *Risk Aspects of Investment-Based Social Security Reform*, Chap. 4, 113–152: The University of Chicago Press, Chicago, 10.7208/chicago/9780226092560.003.0005.

Reis, Ricardo (2022) "Debt Revenue and the Sustainability of Public Debt," *Journal of Economic Perspectives*, 36 (4), 103–24, 10.1257/jep.36.4.103.

Ross, Steve (2015) "The Recovery Theorem," *Journal of Finance*, 70 (2), 615–648, 10.1111/jofi.12092.

Shiller, Robert J. (1999) “Social Security and Institutions for Intergenerational, Intragenerational, and International Risk-sharing,” *Carnegie-Rochester Conference Series on Public Policy*, 50, 165–204, 10.1016/S0167-2231(99)00026-3.

Song, Zheng, Kjetil Storesletten, Yikai Wang, and Fabrizio Zilibotti (2015) “Sharing High Growth across Generations: Pensions and Demographic Transition in China,” *American Economic Journal: Macroeconomics*, 7 (2), 1–39, 10.1257/mac.20130322.

Spear, Stephen and Sanjay Srivastava (1987) “On Repeated Moral Hazard with Discounting,” *Review of Economic Studies*, 54 (4), 599–617, 10.2307/2297484.

Stokey, Nancy L., Robert E. Lucas Jr., with Edward C. Prescott (1989) *Recursive Methods in Economic Dynamics*: Harvard University Press, Cambridge, Mass.

Thomas, Jonathan P. and Tim Worrall (1988) “Self-Enforcing Wage Contracts,” *Review of Economic Studies*, 55 (4), 541–554, 10.2307/2297404.

——— (1994) “Foreign Direct Investment and the Risk of Expropriation,” *Review of Economic Studies*, 61 (1), 81–108, 10.2307/2297878.

——— (2007) “Unemployment Insurance under Moral Hazard and Limited Commitment: Public versus Private Provision,” *Journal of Public Economic Theory*, 9 (1), 151–181, 10.1111/j.1467-9779.2007.00302.x.

Willems, Tim and Jeromin Zettelmeyer (2022) “Sovereign Debt Sustainability and Central Bank Credibility,” *Annual Review of Financial Economics*, 14 (1), 75–93, 10.1146/annurev-financial-112921-110812.

Zhu, Shenghao (2020) “Existence of Equilibrium in an Incomplete Market Model with Endogenous Labor Supply,” *International Economic Review*, 61 (3), 1115–1138, 10.1111/iere.12451.

Online Appendix (Proofs of Main Results)

This Appendix contains the proofs of the main results. Omitted proofs can be found in Part B of the Supplementary Appendix.

Proof of Lemma 2.

(i) Since the constraint set $\Phi(s,\omega)$ is convex and the objective function is strictly concave, the policy function $g_r(\omega,s)$ is single-valued and continuous in ω. Let $h_s(\omega) := -(\delta/\beta)V_\omega(s,\omega)$ where $h_s\colon \Omega(s) \to [\lambda_{\min}(s), \lambda_{\max}(s)]$ with $\lambda_{\min}(s) = \max\{0, (\delta/\beta)((1-s)/s) - 1\}$. Let $h_s^{-1}\colon [\lambda_{\min}(s), \lambda_{\max}(s)] \to \Omega(s)$ be its inverse. By the concavity of the frontier $V(s,\omega)$ in ω, $h_s^{-1}(\lambda)$ is strictly increasing in λ for $\lambda > \lambda_{\min}(s)$. Suppose first that $\omega \geq \omega^0(s)$. Hence, from (7), $f(s,\omega) = 1-\exp(\omega)$. Since $g_r(s,\omega) = \max\{\omega_{\min}(r), \min\{\omega_{\max}(r), h_r^{-1}(\mu(s,\omega))\}\}$, substituting into (11), there is a unique value (possibly zero) of μ that satisfies the constraint. If $\mu(s,\omega)=0$, then $g_r(s,\omega) = \omega^0(r)$ for each r. If $\mu(s,\omega) > 0$, then $\mu(s,\omega)$ is strictly increasing in ω since $f(s,\omega)$ is strictly decreasing in ω and $h_r^{-1}(\mu)$ is increasing in μ. Thus, $g_r(s,\omega)$ is strictly increasing in ω for $g_r(s,\omega) \in (\omega^0(r), \omega_{\max}(r))$. If $\omega < \omega^0(s)$, then $\lambda(s,\omega) = 0$ and hence, since $f(s,\omega)$ is independent of ω, $g_r(s,\omega)$ is also independent of ω.

(ii) Consider states $s(i) > s(i-1)$, $i = 2,\dots,I$. For brevity, write $g_r(i,\omega)$ for $g_r(s(i),\omega)$ and $g_i(s,\omega)$ for $g_{r(i)}(s,\omega)$ etc. We first show that $\mu(i,\omega) \geq \mu(i-1,\omega)$ for $\omega \in [\omega_{\min}(i-1), \omega_{\max}(i)]$ with a strict inequality unless $\mu(i,\omega) = \mu(i-1,\omega) = 0$. Suppose to the contrary that $\mu(i-1,\omega) \geq \mu(i,\omega) > 0$. It follows from (12) that $g_r(i-1,\omega) \geq g_r(i,\omega)$. Using (11) and $\hat{v}(s) = \log(s) + \beta\sum_r \pi(r)\log(1-r)$, gives

$$\log(f(i-1,\omega)) - \log(f(i,\omega)) = (\hat{v}(i-1) - \hat{v}(i)) + \beta\sum\nolimits_r \pi(r)\left(g_r(i,\omega) - g_r(i-1,\omega)\right).$$

Since $\hat{v}(i-1) - \hat{v}(i) < 0$ and $g_r(i,\omega) - g_r(i-1,\omega) \leq 0$, $f(i,\omega) > f(i-1,\omega)$ and $\log(1-f(i-1,\omega)) > \log(1-f(i,\omega)) \geq \omega$. Hence, $\lambda(i-1,\omega) = 0 \leq \lambda(i,\omega)$. However, since $\lambda(i,\omega) \geq \lambda(i-1,\omega)$ and $\mu(i-1,\omega) \geq \mu(i,\omega)$, it follows from (12) that $f(i-1,\omega) \geq f(i,\omega)$, a contradiction. Hence, if $\mu(i-1,\omega) = \mu(i,\omega) = 0$, then $g_r(i-1,\omega) = g_r(i,\omega) = \omega^0(r)$ independently of s. If, however, $\mu(i-1,\omega) > 0$, then it follows from (12) that $g_r(i-1,\omega) < g_r(i,\omega)$ for $\omega \in [\omega_{\min}(i-1), \omega_{\max}(i)]$. By Assumption 3, $\mu(1,\omega^0(1)) = 0$ and by Assumption 4, $\mu(I,\omega^0(I)) > 0$. Since $\mu(s,\omega)$ is increasing in ω, $\mu(I,\omega^0(I)) > 0$ and $\mu(I,\omega) > \mu(1,\omega)$ for $\omega \in (\omega^0(1), \omega_{\max}(I))$. Hence, from (13), $V_\omega(r, g_r(I,\omega)) < V_\omega(r, g_r(1,\omega))$ and therefore, from the strict concavity of $V(r,\omega)$ in ω for $\omega > \omega^0(1) \geq \omega^0(r)$ it follows that $g_r(I,\omega) > g_r(1,\omega)$. Next, if $g_i(x) \leq \omega^0(i-1)$ or $g_{i-1}(x) \geq \omega_{\max}(i)$, then $g_{i-1}(x) \geq g_i(x)$. Therefore, suppose $g_i(x), g_{i-1}(x) \in (\omega^0(i-1), \omega_{\max}(i))$. We first show that $V_\omega(i-1,\omega) \geq V_\omega(i,\omega)$ for $\omega \in (\omega^0(i-1), \omega_{\max}(i))$. For $\omega > \omega^0(i-1)$, it follows that $\lambda(i-1,\omega) > 0$ and since $\omega^0(i-1) \geq \omega^0(i)$,

$\lambda(i,\omega) > 0$. Therefore, $f(i,\omega) = f(i-1,\omega)$. In this case, it follows from above that $\mu(i,\omega) \geq \mu(i-1,\omega)$ with equality only if $\mu(i,\omega) = \mu(i-1,\omega) = 0$. Hence, it follows from (12) that $\lambda(i-1,\omega) \leq \lambda(i,\omega)$ with strict inequality if $\mu(i,\omega) > 0$. Using (14), it follows that $V_\omega(i-1,\omega) \geq V_\omega(i,\omega)$ with strict inequality if $\mu(i,\omega) > 0$. For $g_i(x), g_{i-1}(x) > \omega^0(i-1)$, $\eta_i(x) = \eta_{i-1}(x) = 0$ and for $g_i(x), g_{i-1}(x) < \omega_{\max}(i)$, $\xi_i(x) = \xi_{i-1}(x) = 0$. Hence, it follows from (13) that $V_\omega(i, g_i(s,\omega)) = V_\omega(i-1, g_{i-1}(s,\omega))$. Since $V_\omega(i-1,\omega) \geq V_\omega(i,\omega)$, it follows from the concavity of $V(\cdot,\omega)$ in ω that $g_{i-1}(s,\omega) \geq g_i(s,\omega)$. The inequality is strict if $V_\omega(i-1,\omega) > V_\omega(i,\omega)$ by the strict concavity of $V(\cdot,\omega)$ in ω. Since $\mu(I,\omega) > \mu(1,\omega)$ for $\omega \in (\omega^0(1), \omega_{\max}(I))$, $V_\omega(1,\omega) > V_\omega(I,\omega)$ and hence, $g_1(s,\omega) > g_I(s,\omega)$.

(iii) Since $\mu(1,\omega^0(1)) = 0$ and $f(1,\omega^0(1)) = s(1)$ it follows that $g_r(1,\omega^0(1)) = \omega^0(r)$ for each r. Since $\omega^0(r) > \omega_{\min}(r)$ for at least some r, it follows that (11) is strictly slack and there is some $\omega^c > \omega^0(1)$ such that (11) is non-binding with $g_r(1,\omega) = \omega^0(r)$ for each r and $\omega \in [\omega^0(1), \omega^c]$.

(iv) It follows from (12) that for $\omega = \omega^*(s) > \omega_{\min}(s)$, $\mu(s,\omega) = \lambda(s,\omega)$. In this case, $V_\omega(s,\omega^*(s)) = V_\omega(r, g_r(s,\omega^*(s)))$ for $g_r(s,\omega^*(s)) \in (\omega^0(r), \omega_{\max}(r))$, and, in particular, $g_s(s,\omega^*(s)) = \omega^*(s)$, so that $\omega^*(s)$ is a fixed point of the mapping $g_s(s,\omega)$. Equally, for $\omega < \omega^*(s)$, it follows from (12) that $\mu(s,\omega) > \lambda(s,\omega)$, so that from the concavity of the frontier, $g_s(s,\omega) > \omega^*(s)$. Likewise, for $\omega > \omega^*(s)$, it follows from (12) that $\mu(s,\omega) < \lambda(s,\omega)$, so that from the concavity of the frontier, $g_s(s,\omega) < \omega^*(s)$. If $\omega^*(s) = \omega_{\min}(s)$, then $f(s,\omega) = s$ and $\mu(s,\omega^*(s)) = 0$ by Assumption 2. Hence, $g_s(s,\omega^*(s)) = \omega^*(s)$. Since $\omega^0(s)$ is decreasing in s, it follows by Assumption 4 that $\omega^0(I) < \omega^f(I) \leq \omega^*$. ■

Proof of Lemma 3.

(i) For $\omega > \omega^0(s)$, $\lambda(s,\omega) > 0$ and therefore, it follows from (7) that $f(s,\omega) = 1 - \exp(w)$. For $\omega = \omega^0(s)$, either $\lambda(s,\omega^0(s)) > 0$ or $\lambda(s,\omega^0(s)) = 0$. In either case, it follows from (7) or the definition of $\omega^0(s)$ that $f(s,\omega^0(s)) = 1 - \exp(w^0(s))$. For $\omega < \omega^0(s)$, it follows that $\lambda(s,\omega) = 0$. From (12), let $\phi(s,\mu) = \min\{\delta(1+\mu)/(\beta+\delta(1+\mu)), s\}$ where $\phi(s,\mu)$ is increasing in μ with $\phi(s,0) = c^*(s)$. Recall that $h_r^{-1}(\mu)$, defined in the proof of Lemma 2, satisfies $V_\omega(r, h_r^{-1}(\mu)) = -(\beta/\delta)\mu$ where $g_r(s,\omega) = \max\{\omega_{\min}(r), \min\{\omega_{\max}(r), h_r^{-1}(\mu(s,\omega))\}\}$. Since $h_r^{-1}(\mu)$ is increasing in μ, it follows from (11) that when $f(s,\omega^0(s)) = 1 - \exp(w^0(s))$, there is a unique value of μ, say $\mu^0(s)$, that solves the constraint. Furthermore, $\omega^0(s) = \log(1 - \phi(s,\mu^0(s)))$.

(ii) Since $\hat{v}(i) > \hat{v}(i-1)$, it follows from Part (i) that $\mu^0(i) \geq \mu^0(i-1)$ with strict inequality if $\mu^0(i) > 0$. Therefore, since $\phi(s,\mu)$ is strictly increasing in μ and independent of s for $\mu > 0$,

$c^0(i) \geq c^0(i-1)$ with strict inequality if $\mu^0(i) > 0$. By Assumption 4, $\mu^0(I) > 0$ and by Assumption 3, $\mu^0(1) = 0$. Hence, $c^0(I) > c^0(1)$.

(iii) Lemma 2 establishes that at the fixed point, $\omega^f(s) = \min\{\omega_{\max}(s), \omega^*(s)\}$. Hence, $f(s, \omega^f(s)) \leq c^*(s)$ with equality for $\omega^f(s) < \omega_{\max}(s)$. ∎

Proof of Proposition 5. Using the properties of $g_r(x)$ from Lemma 2 and the argument in the text, it follows that there is an $k \geq 1$ and $\epsilon > 0$ such that $P^k(x, \{x_0\}) > \epsilon$ for each $x \in \mathcal{X}$ and any x_0. Hence, Condition **M** of Stokey et al. (1989, page 348) is satisfied. Therefore, Theorem 11.12 of Stokey et al. (1989) applies and there is strong convergence. Non-degeneracy with $|E| > I$ follows from Assumption 4. The finiteness of the return times follows from Lemma 2(iii) and the finiteness of $\mathcal{I}$. The relationship between the probability mass and the expected return times and the pointwise convergence is standard (see, for example, Theorems 10.2.3 and 13.1.2, Meyn and Tweedie, 2009). ∎

Proof of Proposition 6.

(i) Since $q_+^k(x, x') = \varsigma(\gamma)\bar{\gamma}^{-k}(\bar{\gamma}/\gamma)q^k(x, x')$, summing over x' and γ, the unadjusted bond prices are $p_+^k(x) = \bar{\gamma}^{-k}p^k(x)$. Hence, the yields satisfy $y_+^k(x) = y^k(x) + \log(\bar{\gamma})$.

(ii) It is a standard result (see, for example, Martin and Ross, 2019) that $\lim_{k\to\infty} y^k(x) = \mathbb{E}_\phi[\log(m(x, x'))] = \log(\varrho)$, where $\mathbb{E}_\phi$ is the expectation taken over the invariant distribution of x and ϱ is the Perron root of the matrix Q. Taking logs in equation (17), $\log(m(x, x')) = \log(\beta) + \log(c(x)) - \log(1 - c(x'))$. Using equations (12) and (15), assuming the non-negativity constraints and the upper bound constraint do not bind, gives $\log(c(x')) - \log(1 - c(x')) = -\log(\beta/\delta) + \log(\nu(x')) - \log(\nu(x))$, where $\nu(x) = 1 + \mu(x)$. Therefore, $\log(m(x, x')) = \log(\delta) + \log(c(x)) - \log(c(x')) + \log(\nu(x')) - \log(\nu(x))$. Taking expectations at the invariant distribution, $\mathbb{E}_\phi[\log(m(x, x'))] = \log(\delta)$. Hence, $\varrho = \delta$ and $\lim_{k\to\infty} y_+^k(x) = \log(\delta) + \log(\bar{\gamma})$.

(iii) Recall that $m(x, x') = m((s, d), (r, b_r(d))) = \beta s(1 - d)/(1 - r(1 - b_r(d)))$. Since $b_r(d)$ is increasing in d by Corollary 1, it follows that $m(x, x')$ is decreasing in d. The price of a one-period discount bond in state (s, d) is $p^1(s, d) = \sum_r \pi(r)m((s, d), (r, b_r(d)))$, which is also decreasing in d. Making the induction hypothesis that the price of a k-period discount bond is decreasing in d, $p^{k+1}(s, d) = \sum_r \pi(r)m((s, d), (r, b_r(d)))p^k(r, b_r(d))$. Since $p^k(s, d)$ and $m((s, d), (r, b_r(d)))$ are positive and decreasing in d, and $b_r(d)$ is increasing in d, it follows that $p^{k+1}(s, d)$ is decreasing d. Hence, the conditional yield $y^k(s, d) = -(1/k)\log(p^k(s, d))$ is increasing in d for each s and k.

(iv) From Corollary 1, the fixed points of the mappings of $b_r(d)$ are $d^*(r)$ when the upper bound constraint does not bind, and the consumption share is at the first best at these fixed points. Hence, $m((s, d^*(s)), (s, d^*(s))) = \delta$. By Lemma 2, the consumption share of the old decreases with r. Hence, $m((1, d^*(1)), (r, b_r(d^*(1)))) \geq \delta$ with a strict inequality for some r. Taking expectations, the bond price $p^1(1, d^*(1)) > \delta$ and consequently, the yield $y^1(1, d^*(1)) < -\log(\delta)$. Since $y^\infty = -\log(\delta)$, $y^\infty - y^1(1, d^*(1)) > 0$. Likewise, it can be checked that $m((I, d^*(I)), (r, b_r(d^*(I)))) \leq \delta$ with a strict inequality for some r, which shows that $y^\infty - y^1(I, d^*(I)) < 0$.

(v) By definition $\Upsilon = \log(\psi_{\max}/\psi_{\min})$ where $\psi_{\max}$ and $\psi_{\min}$ are the maximum and minimum values of the eigenvector of the matrix Q. Using (12) and (15) and assuming the non-negativity and upper bound constraints do not bind, $m_B(x, x') = v(x')/v(x)$. Since $m_A(x, x') = \delta f(x)/f(x')$, the eigenvector $\psi(x) = f(x)/v(x)$. Since $f(x') = \delta v(x)/(\beta v(x) + \delta v(x'))$, it follows that $\psi(x') = \delta/(\beta v(x)+\delta v(x'))$. The maximum value of $\psi(x')$ occurs when $v(x) = v(x') = 1$, in which case $\psi_{\max} = \delta/(\beta + \delta)$. The minimum value occurs when $v(x) = v(x') = v_{\max}$, in which case $\psi_{\min} = \delta/((\beta + \delta)v_{\max})$. Hence, $\Upsilon = \log(\psi_{\max}/\psi_{\min}) = \log(v_{\max})$. It is easily checked that $v(s, d)$ is increasing in d with $v(s, d^0(s))$ increasing in s, so that for $(s, d) \in E$, $v_{\max} = v(I, d^*(I))$. ∎

Proof of Proposition 7. With $R(x, x') = rb_r(d)/(s\mathrm{BR}(d))$, the expected return is $\bar{R}(x) = \sum_r \pi(r)rb_r(d)/(s\mathrm{BR}(d))$. The risk-free rate is $R^f(x) = (\sum_r q(x, x'))^{-1}$ where $q(x, x') = \pi(r)\beta s(1 - d)/(1 - r(1 - b_r(d))$. Therefore, $\bar{R}(x)/R^f(x)$ is independent of s. Since the risk-adjusted return on any asset is equal to the risk-free return, $\mathrm{MRP}(d) = -\operatorname{cov}(m(x, x'), R(x, x'))$ where $m(x, x') = q(x, x')/\pi(r)$. From Corollary 1, $b_r(d)$ is increasing in r and hence, $R(x, x')$ is increasing with r. From Lemma 2, old consumption $(1 - r(1 - b_r(d)))$ falls with r and hence, $m(x, x')$ is increasing with r. By Assumption 4, risk sharing is incomplete, and hence, the covariance term is positive and $\mathrm{MRP}(d) < 0$. That is, $\bar{R}(x)/R^f(x) < 1$. With growth shocks, $R_+(x, x') = R(x, x')\gamma$ and $q_+(x, x') = \varsigma(\gamma)q(x, x')/\gamma$. Hence, $\bar{R}_+(x) = \bar{R}(x)(\mathbb{E}_\gamma\gamma)$, $R^f_+(x) = R^f(x)\bar{\gamma}$, and

$$\mathrm{MRP}_+(d) = \frac{\bar{R}_+(x) - R^f_+(x)}{R^f_+(x)} = \left(\frac{\bar{R}(x)}{R^f(x)} - 1\right) + \left(\frac{\bar{R}(x)}{R^f(x)}\right)\left(\frac{\mathbb{E}_\gamma\gamma}{\bar{\gamma}} - 1\right).$$

Let $R^*_+(x, x')$ denote the returns with complete insurance. It is easy to check that

$$R^*_+(x, x') = \frac{\left(r - \frac{\delta}{\beta+\delta}\right)\gamma}{\delta\left(\sum_r \pi_r r - \frac{\delta}{\beta+\delta}\right)}.$$

The corresponding expected return is $\bar{R}^*_+(x) = (\mathbb{E}_\gamma\gamma)/\delta$. Likewise, the state price is $q^*_+(x, x') = \delta\varsigma(\gamma)\pi(r)/\gamma$, so that the risk-free return is $R^{f*}_+ = \bar{\gamma}/\delta$. Hence, the corresponding multiplicative risk premium is $\mathrm{MRP}^* = (\mathbb{E}_\gamma \gamma - \bar{\gamma})/\bar{\gamma}$. Since the arithmetic mean is larger than the harmonic mean, $\mathrm{MRP}^* > 0$. Substituting into the equation above gives $\mathrm{MRP}_+(d) = \mathrm{MRP}(d) + \alpha(d)\mathrm{MRP}^*$, where $\alpha(d) = \bar{R}(x)/R^f(x)$, as required. ∎

Proof of Proposition 8.

(i) Since the probability of endowment state 1 is π and debt is reset to the regeneration levels $d^0(s)$ after endowment state 1 has occurred, the probability that the state $(s, d^0(s))$ occurs is $\phi(s, d^0(s)) = \pi(s)\pi$, irrespective of the date or history. Therefore, T periods after such a resetting, the distribution function is:

$$\phi_T(s, d^{(n)}(s)) = \phi(s, d^0(s))(1-\pi)^n \quad \text{for} \quad n = 0, 1, 2, \ldots, T-1,$$

with $\phi_T(s, d^{(T)}(s)) = \pi(s)(1-\pi)^T$. Taking the limit as $T \to \infty$ gives the invariant distribution ϕ described in the text.

(ii) By Proposition 6, $\Upsilon = \log(\nu_{\max})$. The value of $\nu_{\max}$ can be found from the fixed point of the mapping $b_2(d)$, which occurs at $d = d^*(2)$. From the first-order condition (12), $\log(\nu_{\max}) = \log(\delta/\beta) + \log((s(1)(1 - b_1(d^*(2))))^{-1} - 1)$. Since the participation constraint is binding when $d = d^*(2)$ and $b_2(d^*(2)) = d^*(2)$, $b_1(d^*(2))$ can be found by solving:

$$\begin{aligned} \log(1 - d^*(2)) + \beta\left(\pi \log\left(1 - s(1) + s(1)b_1(d^*(2))\right) + (1-\pi)\log\left(1 - s(2) + s(2)d^*(2)\right)\right) \\ = \beta\left(\pi \log\left(1 - s(1)\right) + (1-\pi)\log\left(1 - s(2)\right)\right). \end{aligned}$$

Since $s(1) = \kappa - \epsilon(1-\pi)/\pi$ and $s(2) = \kappa + \epsilon$, setting $\chi = 1 - s(1)(1 - b_1(d^*(2)))$ and using $d^*(2) = 1 - \delta/(s(2)(\beta + \delta))$ gives the result in the text. ∎

References

Meyn, Sean and Richard L. Tweedie (2009) *Markov Chains and Stochastic Stability*, Cambridge Mathematical Library: Cambridge University Press, Cambridge, 2nd edition, 10.1017/CBO9780511626630.

Supplementary Appendix

This appendix presents supplementary material referenced in the paper. Part A provides evidence on the relative income of the young and the old for six OECD countries referred to in footnote 2 in the Introduction. Part B provides proofs of Propositions 2 and 4 from Sections 1 and 2 together with the proofs of Lemma 1 from Section 3 and Corollary 1 from Section 5. Parts C-E relate to the two-state example of Section 8. Part C examines two alternative welfare measures, the insurance coefficient and consumption-equivalent welfare change measure and compares some comparative static properties. Part D considers the impact of a demographic shock and shows how the effect of the shock is both amplified and persistent. Part E presents the shooting algorithm used to derive the optimal allocation in the two-state example. Part F describes the pseudo-code for the numerical algorithms used in the paper.

A Change in Relative Income of Young and Old

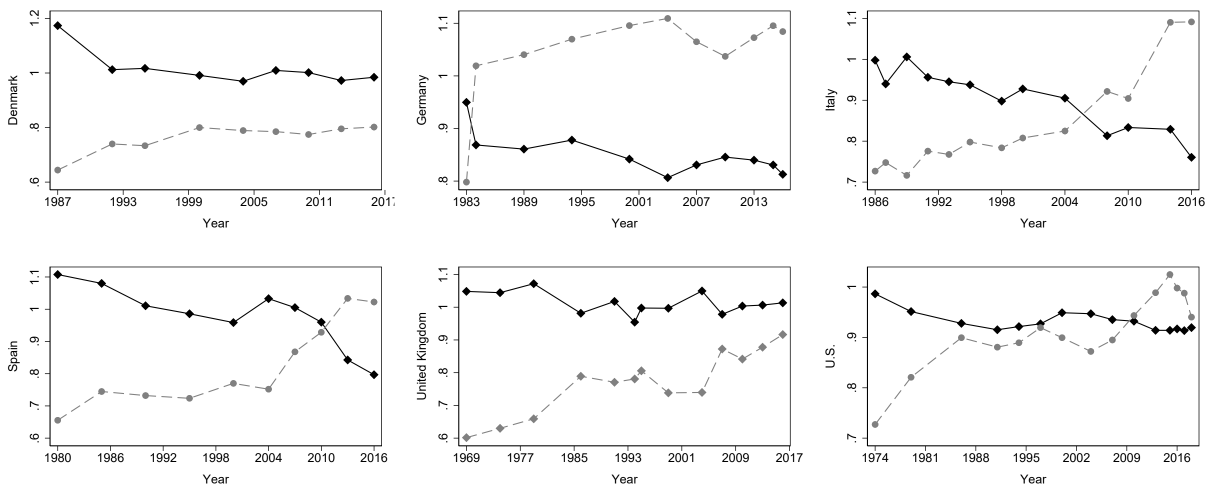

Figure A.1: Relative Income of Young and Old for six OECD Countries

Note: The solid line is the average net (of taxes and transfers) equivalized disposable income for individuals aged 25-34 divided by the average of the same measure for the whole population. The dotted line is the corresponding ratio for individuals aged 65-74.

Figure A.1 illustrates the average disposable income of individuals aged 25-34 (the young) and the average disposable income of individuals aged 65-74 (the old) relative to the national average over recent decades for Denmark, Germany, Italy, Spain, U.K. and U.S. (data periods are country specific). Data is taken from the Luxembourg Income Study Database available at `www.lisdatacenter.org`. In each country there has been an improvement in the average disposable income of the old compared to the average disposable income of the young over the sample period. For example, the average disposable income of the young in the U.S. has fallen from just below the national average to just above 90% of the national average during 1974-2018. Over the same period, the old have fared much better with their average disposal income rising from approximately 70% of the national average to become roughly equal to the national average. Moreover, the old overtook the young for the first time around the time of the financial crisis of 2008.

A similar pattern can be seen in Italy and Spain and a narrowing of the gap between the young and the old can also be observed in Denmark and the U.K. Germany is somewhat different with the old overtaking the young as early as the 1980s.

B Omitted Proofs

Proof of Proposition 2. Here, we consider the case without growth shocks, so that $\bar{\gamma} = 1$. The lifetime endowment utility of an agent born in a state with endowment share s is:

$$\hat{v}(s) := \log(s) + \beta \sum_r \pi(r) \log(1 - r).$$

Consider a small transfer $\mathrm{d}\tau(s)$ in state s from the young to the old. The problem of existence of a sustainable allocation can be answered by finding a vector of positive transfer $\mathrm{d}\tau$ such that there is a weak improvement over the lifetime endowment utility in all states and a strict improvement in at least one state. The change in the lifetime endowment utility induced by a vector $\mathrm{d}\tau$ is non-negative if

$$-s^{-1}\mathrm{d}\tau(s) + \beta \sum_r \pi(r)(1 - r)^{-1}\mathrm{d}\tau(r) \geq 0. \tag{B.1}$$

Rearranging (B.1) in terms of the marginal rates of substitution $\hat{m}(s, r)$, we have:

$$-\mathrm{d}\tau(s) + \sum_r \pi(r)\hat{m}(s, r)\mathrm{d}\tau(r) \geq 0.$$

The problem of existence can then be addressed by finding a vector $\mathrm{d}\tau > 0$ that solves:

$$\left(\hat{Q} - I\right) \mathrm{d}\tau \geq 0, \tag{B.2}$$

where I is the identity matrix and $\hat{Q}$ is the matrix of $\hat{q}(s, r) = \pi(r)\hat{m}(s, r)$. Equation (B.2) has a well-known solution. Using the Perron-Frobenius theorem, there exists a strictly positive solution for $\mathrm{d}\tau$, provided that the Perron root, that is, the largest eigenvalue of $\hat{Q}$, is greater than one. This is satisfied by Assumption 2, which guarantees the existence of positive transfers from the young to the old that improve the utility of each generation. ∎

Proof of Proposition 4. Let $\omega_0^c = \omega^*$ and define the critical utility ω_1^c by:

$$\log(1-\exp(\omega_1^c)) + \beta\omega^* = \hat{v} := \log(s) + \beta\log(1-s).$$

Define ω_n^c recursively by:

$$\log(1-\exp(\omega_n^c))) + \beta\omega_{n-1}^c = \hat{v} \quad \text{for } n = 2, 3, \ldots, \infty.$$

From the strict concavity of the logarithmic utility function $\omega_n^c > \omega_{n-1}^c$ and $\lim_{n\to\infty}\omega_n^c = \omega_{\max} = \log(s)$. Let $v^* = \log(1-\exp(\omega^*)) + (\beta/\delta)\omega^*$. With some abuse of notation, let $V_n(\omega)$ denote the value function when $\omega \in (\omega_{n-1}^c, \omega_n^c]$. Hence,

$$V_n(\omega) = \log(1-\exp(\omega)) + \frac{\beta}{\delta}\omega + \delta V_{n-1}\left(\frac{1}{\beta}\left(\hat{v} - \log(1-\exp(\omega))\right)\right).$$

For $\omega \leq \omega^*$, $\omega' = \omega^*$. Therefore, $V(\omega) = v^*/(1-\delta)$ for $\omega \in [\omega_{\min}, \omega^*]$. For $\omega \in (\omega^*, \omega_1^c]$,

$$V_1(\omega) = \log(1-\exp(\omega)) + \frac{\beta}{\delta}\omega + \frac{\delta}{1-\delta}v^*.$$

Differentiating the function $V_1(\omega)$ gives:

$$\frac{\mathrm{d}V_1(\omega)}{\mathrm{d}\omega} = \frac{\beta}{\delta} - \frac{\exp(\omega)}{1-\exp(\omega)}.$$

Let $h(\omega) := \exp(\omega)/(1-\exp(\omega))$. Since $\omega > \omega^*$, $h(\omega) > \beta/\delta$ and $\mathrm{d}V_1(\omega)/\mathrm{d}\omega < 0$. Note that $h(\omega^*) = \beta/\delta$ and therefore, in the limit as $\omega \to \omega^*$, $\mathrm{d}V_1(\omega)/\mathrm{d}\omega = 0$. Furthermore, the function $V_1(\omega)$ is strictly concave because $h(\omega)$ is increasing in ω. Using this result, we can proceed by induction and assume $V_{n-1}(\omega)$ is decreasing and strictly concave. Then, it is straightforward to establish that $V_n(\omega)$ is decreasing and strictly concave. Continuity follows since $\lim_{\omega\to\omega_n^c} V_{n+1}(\omega) = V_n(\omega_n^c)$. To establish differentiability, we need to demonstrate that:

$$\lim_{\omega\to\omega_n^c}\frac{\mathrm{d}V_{n+1}(\omega)}{\mathrm{d}\omega} = \frac{\mathrm{d}V_n(\omega_n^c)}{\mathrm{d}\omega}.$$

To show this, note that for $\omega \in (\omega_n^c, \omega_{n+1}^c)$:

$$\frac{\mathrm{d}V_{n+1}(\omega)}{\mathrm{d}\omega} = \frac{\beta}{\delta} - h(\omega)\left(1 - \frac{\delta}{\beta}\frac{\mathrm{d}V_n(\omega')}{\mathrm{d}\omega}\right).$$

Starting with $n = 1$, we have:

$$\lim_{\omega\to\omega_1^c} \frac{\mathrm{d}V_2(\omega)}{\mathrm{d}\omega} = \frac{\beta}{\delta} - h(\omega_1^c)\left(1 - \frac{\delta}{\beta}\lim_{\omega\to\omega_0^c}\frac{\mathrm{d}V_1(\omega)}{\mathrm{d}\omega}\right).$$

Since $\lim_{\omega\to\omega_0^c} \mathrm{d}V_1(\omega)/\mathrm{d}\omega = 0$, we have:

$$\lim_{\omega\to\omega_1^c} \frac{\mathrm{d}V_2(\omega)}{\mathrm{d}\omega} = \frac{\beta}{\delta} - h(\omega_1^c) = \frac{\mathrm{d}V_1(\omega_1^c)}{\mathrm{d}\omega}.$$

Therefore, make the recursive assumption that $\lim_{\omega\to\omega_{n-1}^c} \mathrm{d}V_n(\omega)/\mathrm{d}\omega = \mathrm{d}V_{n-1}(\omega_{n-1}^c)/\mathrm{d}\omega$. In general, we have:

$$\lim_{\omega\to\omega_n^c} \frac{\mathrm{d}V_{n+1}(\omega)}{\mathrm{d}\omega} = \frac{\beta}{\delta} - h(\omega_n^c)\left(1 - \frac{\delta}{\beta}\lim_{\omega\to\omega_{n-1}^c}\frac{\mathrm{d}V_n(\omega)}{\mathrm{d}\omega}\right)$$
$$\frac{\mathrm{d}V_n(\omega_n^c)}{\mathrm{d}\omega} = \frac{\beta}{\delta} - h(\omega_n^c)\left(1 - \frac{\delta}{\beta}\frac{\mathrm{d}V_{n-1}(\omega_{n-1}^c)}{\mathrm{d}\omega}\right).$$

By the recursive assumption, these two equations are equal. Hence, we conclude that $V(\omega)$ is differentiable. In particular, repeated substitution gives:

$$\frac{\mathrm{d}V_n(\omega_n^c)}{\mathrm{d}\omega} = \frac{\beta}{\delta} - \left(\frac{\delta}{\beta}\right)^{n-1}\prod_{j=1}^{n} h(\omega_j^c).$$

Since $h(\omega_j^c) \in [\beta/\delta, s/(1-s))$, taking the limit as $n \to \infty$, or equivalently, $\omega \to \omega_{\max}$, gives $\lim_{\omega\to\omega_{\max}} \mathrm{d}V(\omega)/\mathrm{d}\omega = -\infty$. Equation (6) follows from above given that $\omega' = \omega^*$ or satisfies $\log(1-\exp(\omega)) + \beta\omega' = \hat{v}$ if $\log(1-\exp(\omega)) + \beta\omega^* < \hat{v}$. ∎

Proof of Lemma 1.

(i) Given the participation constraint of the old, $\omega \geq \omega_{\min}(s) = \log(1-s)$. The largest feasible ω, $\omega_{\max}$, can be found by solving the set of equations $\log(1-\exp(\omega_{\max}(s))) + \beta\sum_r \pi(r)\omega_{\max}(r) = \hat{v}(s)$. There is a trivial solution where $\omega_{\max}(s) = \omega_{\min}(s)$ but by Proposition 2 there is also a non-trivial solution with $\omega_{\max}(s) > \omega_{\min}(s)$ for each s. Since the utility function is concave, this non-trivial solution is unique. Let $\Delta\varpi := \sum_r \pi(r)(\omega_{\max}(r) - \omega_{\min}(r))$. Then, $\omega_{\max}(s)$ can be found by solving the set of equations $\log(1-\exp(\omega_{\max}(s))) - \log(1-\exp(\omega_{\min}(s))) + \beta\Delta\varpi = 0$. Since $\Delta\varpi > 0$ by Proposition 2, it follows that $\omega_{\max}(s) > \omega_{\min}(s)$ for all s. A reduction in ω, enlarges the constraint set $\Phi(s,\omega)$ of Problem P1 and hence, $V(s,\omega)$ is non-increasing in ω. To show that $V(s,\omega)$ is concave in ω, consider the mapping T where

$$(TJ)(s,\omega) = \max_{\{c,(\omega_r)_{r\in\mathcal{I}}\}\in\Phi(s,\omega)} \frac{\beta}{\delta}\log(1-c) + \log(c) + \delta\sum_r \pi(r)J(r,\omega_r).$$

Consider $J = V^*$, the first-best frontier. Proposition 3 established that $V^*(s, \omega)$ is concave in ω. It follows from the definitions of T and V^* that $TV^*(s, \omega) \leq V^*(s, \omega)$ because $V^*(s, \omega) \leq v^*(s) + \delta\bar{V}^*$ and the mapping T adds the participation constraints (11). That is, $T^n V^*(s, \omega) \leq T^{n-1}V^*(s, \omega)$ for $n = 1$. Now, make the induction hypothesis that $T^n V^*(s, \omega) \leq T^{n-1}V^*(s, \omega)$ for $n \geq 2$ and apply the mapping T to the two functions $T^n V^*(s, \omega)$ and $T^{n-1}V^*(s, \omega)$. It is straightforward to show that $T^{n+1}V^*(s, \omega) \leq T^n V^*(s, \omega)$, because the constraint set is the same in both cases but, by the induction hypothesis, the objective is no greater in the former case. Hence, the sequence $T^n V^*(s, \omega)$ is non-increasing and converges. Let $V^\infty(s, \omega) := \lim_{n\to\infty} T^n V^*(s, \omega)$ be the pointwise limit of the mapping T. We have that V^∞ and V are both fixed points of T. Since the mapping is monotonic, $T^n(V^*) \geq T^n(V) = V$. Hence, $V^\infty \geq V$ but, since V is the maximum, we have that $V^\infty = V$. Starting from V^*, the objective function in the mapping T is concave because V^* and the utility function are concave. The constraint set $\Phi(s, \omega)$ is convex. Hence, $TV^*(s, \omega)$ is concave. By induction, $T^n V^*(s, \omega)$ and the limit function V are also concave. Differentiability follows because the linear independence constraint qualification is satisfied on the interior of the domain. There are $I + 1$ choice variables and $2I + 3$ constraints. Since V is concave, it is differentiable if the multipliers associated with the constraints are unique. The multipliers are unique if the linear independence constraint qualification is satisfied, that is, if the gradients of the binding constraints are linearly independent at the solution. Since $\omega_{\max}(s) > \omega_{\min}(s)$, the corresponding upper and lower bound constraints in (9) and (10) are not both active. Not all lower bound constraints in (10) are active because this would involve no transfers next period, which we know from Proposition 2 is not optimal. For $\omega < \omega_{\max}(s)$, not all upper bound constraints in (9) are active because this would imply that the participation constraint of the young (11) does not bind, in which case some ω_r can be lowered below $\omega_{\max}(r)$ to raise the planner's payoff. If $\omega > \omega_{\min}(s)$, then the non-negativity constraint (8) is not active. If $\omega = \omega_{\min}(s)$ and (8) binds, then the young participation constraint is not active. Hence, in either case, there are at most $I + 1$ active constraints. Since the marginal utility is strictly increasing, $\beta > 0$, and $\pi(s) > 0$ for each s, it can be checked that the matrix of active constraints has full rank. Hence, the multipliers are unique and $V(s, \omega)$ is differentiable in ω on $(\omega_{\min}(s), \omega_{\max}(s))$. Since $V(s, \omega)$ is concave and differentiable in ω, it is also continuously differentiable in ω.

(ii) If constraint (7) binds, then the frontier is strictly downward sloping. Strict concavity of V when V is strictly downward sloping follows since TV is strictly concave when the frontier is strictly downward sloping because of the strict concavity of the logarithmic utility function and the concavity of V. If $\omega^0(s) > \omega^*(s)$, then it would be possible to lower $\omega^0(s)$, increase consumption of the current young keeping all future promises the same without violating any constraints and increase the planner's utility. Assumption 4 guarantees that $\omega^0(s) < \omega^*(s)$ for at least one state s. If $\omega_{\min}(s) = \omega^*(s)$, then $\omega^0(s) = \omega^*(s)$ and hence, constraint (11) does not bind. Therefore, $\mu(s, \omega_{\min}(s)) = 0$. In this case, from equation (13), either $V_\omega(r, g_r(s, \omega)) = 0$ for $\omega = \omega_{\min}(s)$ or one of the constraints (9) or (10) binds and $g_r(s, \omega)$ is independent of ω. Therefore, in either case, $V_\omega(s, \omega^0(s)) = (\beta/\delta) - (\exp(\omega_{\min})/(1 - \exp(\omega_{\min}))) = (\beta/\delta) - ((1 - s)/s) \leq 0$. If $\omega_{\min}(s) < \omega^*(s)$, then it cannot be that $\omega^0(s) = \omega_{\min}(s)$ because this implies $c(s, \omega^0(s)) = s$, which, in turn, implies $\omega^0(s) \geq \omega^*(s)$ from equation (7), a contradiction. The multiplier $\lambda(s, \omega) \geq 0$ and since $V(s, \omega)$ is concave in ω, $\lambda(s, \omega)$ is increasing in ω. Let $\lambda_{\max}(s) := \lim_{\omega \to \omega_{\max}} \lambda(s, \omega)$, then $\lim_{\omega \to \omega_{\max}} V_\omega(s, \omega) = -(\beta/\delta)\lambda_{\max}(s)$, where $\lambda_{\max}(s) \in \mathbb{R}_+ \cup \{\infty\}$.

(iii) Since $\omega_{\min}(s) = \log(1 - s)$, it follows that $\omega_{\min}(s(i)) < \omega_{\min}(s(i - 1))$ for $i = 2, \ldots, I$. It follows from the proof of Part (i) that $\log(1 - \exp(\omega_{\max}(s))) - \log(1 - \exp(\omega_{\min}(s)))$ is independent of s. Therefore, $\omega_{\max}(s(i - 1)) > \omega_{\max}(s(i))$. Equally, $\Delta\varpi$ is bounded above since the endowment share is less than one in each state. Hence, $\omega_{\max}(s) < 0$, otherwise the constraint of the young cannot be satisfied. Since $\omega^0(s) = \log(1 - c^0(s))$, it follows from Lemma 3(ii) that $\omega^0(s(i)) \leq \omega^0(s(i - 1))$ with strict inequality for at least one i, and hence, $\omega^0(I) < \omega^0(1)$. ∎

Proof of Corollary 1.

(i) Since consumption $c = s(1 - d)$ for a fixed debt d, the participation constraint (11) can be rewritten as $\log(1 - d) + \beta \sum_r \pi(r)(\omega_r - \log(1 - r)) \geq 0$. Hence, the multiplier on this constraint depends only on d. If the constraint does not bind, then $g_r(s, \omega) = \omega^0(r)$ from Lemma 2 and consequently, $b_r(d) = d(r, \omega^0(r)) = d^0(r)$. There is a critical value of debt, d^c, below which the constraint does not bind. Since $b_r(d) = d^0(r)$ for $d \leq d^c$, it follows that $d^c = 1 - \exp(-\beta \sum_r \pi(r)(\log(1 - r + rd^0(r)) - \log(1 - r))) \in (d_{\min}, d_{\max})$. By Assumption 3, $d^0(1) = 0$ and hence, $d_{\min} = 0$. Since $d_{\max}$ is the largest non-trivial solution of $\log(1 - d_{\max}) + \beta \sum_r \pi(r)(\log(1 - r + rd_{\max}) - \log(1 - r)) = 0$, it follows that $d_{\max} < 1$. Let $\tilde{h}_s(d) := -(\delta/\beta)V_\omega(s, \log(1 - s + sd))$. Then, for $d \in [d^c, d_{\max})$, $b_r(d) = \tilde{h}_r^{-1}(\mu(d))$. Since $\mu(d)$ is strictly increasing in d and $\tilde{h}_r^{-1}(\mu)$ is strictly increasing in μ, it follows that $b_r(d)$ is strictly increasing in d for $d > d^c$.

(ii) Consider the value function

$$\hat{V}(s, d) = \max_{(b_r)_{r \in \mathcal{I}}} (\beta/\delta) \log(1 - s + sd) + \log(s(1 - d)) + \delta \sum_r \pi(r)\hat{V}(r, b_r)$$

subject to the constraint

$$\log(1 - d) + \beta \sum_r \pi(r) \left(\log(1 - r + rb_r) - \log(1 - r)\right) \geq 0.$$

It follows from this maximization that $\hat{V}_d(s', d) > \hat{V}_d(s, d)$ for $s' > s$. We want to show $b_{r'}(d) \geq b_r(d)$ for $r' > r$. Suppose to the contrary that $b_{r'}(d) < b_r(d)$. It follows from the first-order condition that $\hat{V}_d(r', b_{r'}(d)) \leq \hat{V}_d(r, b_r(d))$ with equality only if the multiplier $\mu(d)$ is not binding. But since $\hat{V}_d(r, b_{r'}(d)) < \hat{V}_d(r', b_{r'}(d))$ it follows that $\hat{V}_d(r, b_{r'}(d)) < \hat{V}_d(r, b_r(d))$, which from the concavity of $\hat{V}(r, d)$ in d, implies $b_{r'}(d) > b_r(d)$, a contradiction.

(iii) Lemma 2(iv) and Assumption 4 guarantee that $d^f(r(I)) > d^c$. ■

C Risk Measures and Comparative Statics

In this Appendix, we continue the two-state example of Section 8 and examine how the bound on the residual risk Υ and alternative measures of risk sharing respond to comparative statics changes of endowment parameters and discount factors. For all comparative statics, we change the value of the parameter of interest holding all other parameters at the values given in Example 1.[41]

The top row of Figure C.1 plots the bound Υ against κ, ϵ and δ, holding $\beta = \delta$ for relevant values of the parameters. A larger κ corresponds to a larger average endowment share to the young, while a smaller ϵ corresponds to reduced uncertainty. Increasing κ, or reducing ϵ, raises risk sharing as measured by a reduction in Υ. For κ above a critical value, or ϵ below a critical value, the first best is sustainable at the invariant distribution, in which case $\Upsilon = 0$.[42] The effect of changes in the discount factor on Υ is non-monotonic when we consider sufficiently low values of δ for which Assumption 5 does not hold. For high values of the discount factor, the invariant distribution has geometric probabilities as described in Part (i) of Proposition 8. As the discount factor falls, either the current transfer is reduced, or the newly state-contingent bonds increased, to satisfy the participation constraint of the young in state 2. This change reduces $d_{\max}$. However, since the change also raises $d^*(2)$, it effectively enlarges the ergodic set, resulting in an increase of risk, as reflected in the rise of Υ. As the discount factor is reduced further, the upper bound constraint on debt becomes binding and it is no longer true that resetting to the regeneration debt levels takes place any time state 1 occurs. Reversion to $d^0(r)$ occurs less frequently and the invariant distribution has now a positive probability mass at $d_{\max}$. The range $d_{\max} - d_{\min}$ decreases, implying that the bound Υ falls.

As alternative measures of risk sharing, we consider the *insurance coefficient* and the *consumption equivalent welfare change*, conditional on state $x = (s, d)$. The insurance coefficient

[41] In all cases, the invariant distribution is geometric, except when discount factors are changed. When the invariant distribution is not geometric, we can no longer rely on the shooting algorithm used in Section 8. In this case, we implement an algorithm based on a value function iteration method (see, Part F of the Supplementary Appendix for a description).

[42] The critical values are $\kappa \approx 0.6565$ and $\sigma \approx 0.0243$.

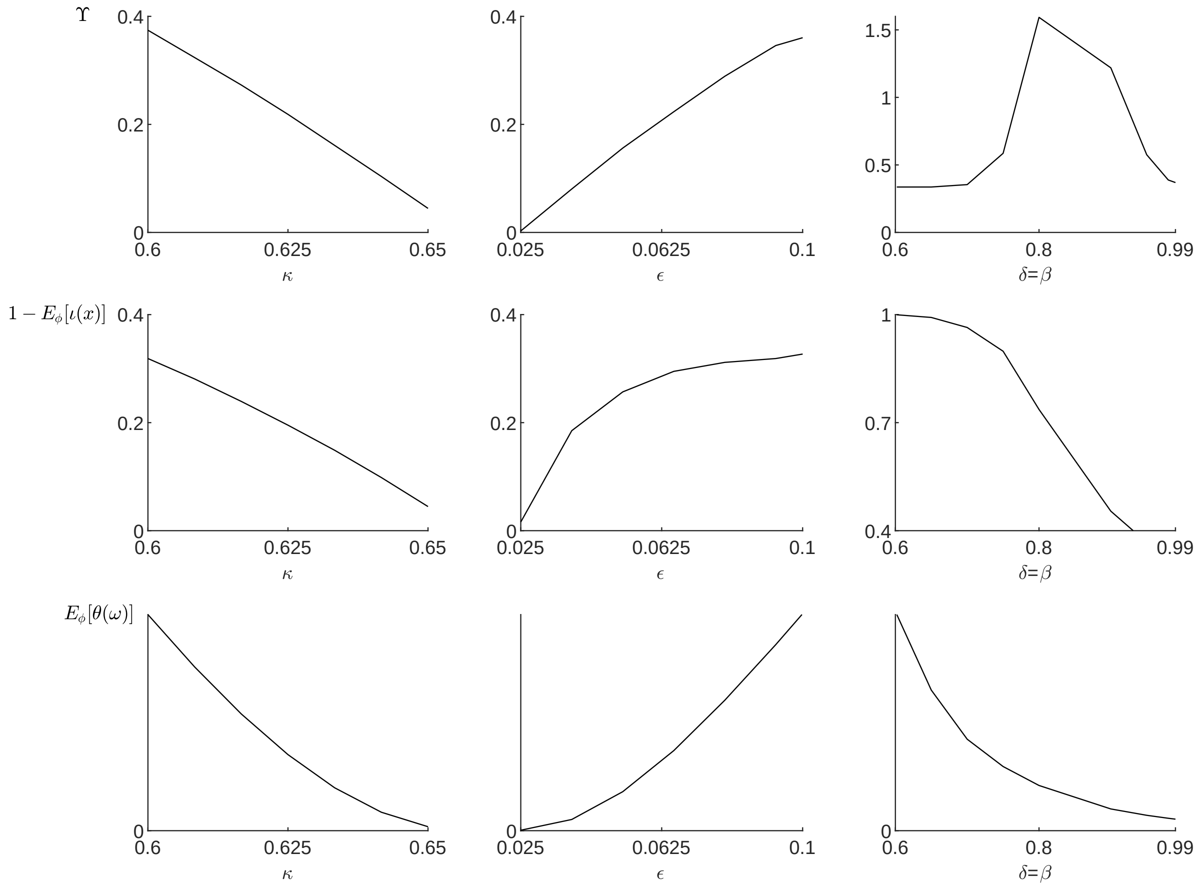

Figure C.1: Comparative Statics on the Bound Υ, the Insurance Coefficient and the Consumption Equivalent Welfare Change

Note: The top row illustrates the bound Υ. The middle row illustrates the average insurance coefficient $\mathbb{E}_\phi[\iota(x)]$. The bottom row illustrates the average consumption equivalent welfare change $\mathbb{E}_\phi[\theta(\omega)]$.

$\iota(x)$ is the fraction of the variance of the endowment shock that does not translate into a corresponding change in consumption. With i.i.d. shocks, the insurance coefficient is:

$$\iota(x) = 1 - \frac{\text{cov}\left(\log\left(c(r, b_r(d))\right), \log(r)\right)}{\text{var}\left(\log(r)\right)}.$$

where r is the endowment shock next period. At the first best, and provided that the boundary constraints on debt do not bind, the consumption share of the young is independent of the state x and the insurance coefficient is one. The expected value of the one minus the insurance coefficient, evaluated at the invariant distribution, is plotted in the middle row of Figure C.1 against κ, ϵ and δ. This measure is smaller when more risk is shared. The consumption equivalent welfare change relative to the first best for a given $\omega := \sum_s \pi(s)\omega_s$ is measured by solving the following equation in terms of θ:

$$\frac{1}{1-\delta}\left(\mathbb{E}[u(c^*(1-\theta))] + \frac{\beta}{\delta}\mathbb{E}[u((1-c^*)(1-\theta))]\right) = \bar{V}(\omega),$$

where $\bar{V}(\omega) = \sum_s \pi(s)V(s,\omega_s)$. The solution $\theta(\omega)$ measures the proportion by which the first-best consumption needs to be reduced to match the optimal solution for each ω. The expected value of $\theta(\omega)$, evaluated at the invariant distribution, is plotted in the bottom row of Figure C.1 against κ, ϵ and δ. The consumption equivalent welfare change is smaller when more risk is shared. The long-run welfare loss measure is the average of $\theta(\omega)$ at the invariant distribution of ω. The comparative statics on the bound Υ are similar to those of the insurance coefficient and the consumption equivalent welfare. The amount of risk shared at the optimal solution increases with κ and δ (provided that the upper bound on debt does not bind) but falls with ϵ.

D Demographic Shock

In this appendix, we illustrate how the economy responds to the arrival of unexpected demographic shock. Consider a one-off increase in fertility at date $T+1$. That is, suppose the cohort size is $N_t = 1$ for $t \leq T$ and $N_t = 1+\varepsilon$ for $t > T$. The ratio of young to old is N_t/N_{t-1}, which equals $1+\varepsilon$ for $t = T+1$ and equals one otherwise. Moreover, suppose that the endowment is proportional to the population size, so that the total endowment is $e_t^y N_t + e_t^o N_{t-1}$, and that the planner's weights placed on the utility of the young and the old are adjusted by the group sizes. Figure D.1 shows the impulse response function to a demographic shock of $\varepsilon = 0.05$ for the parameters of Example 1. The initial increase in the birth rate implies a larger weight on the young and a larger amount of total endowment to be shared at date $T+1$. Under full enforcement (light gray line), the consumption share of the young rises because more weight is placed on their consumption. However, the effect lasts only for one period and the consumption share reverts to its long-run value in the next period. With limited enforcement, the effect is amplified and persists for several periods. This is because a positive fertility shock relaxes the current participation constraint, leading to both increased consumption for the current young and reduced future promises. As a result, it also loosens participation constraints in the future, thereby prolonging the impact of the temporary demographic shock across several periods.

This positive demographic shock is equivalent to an exogenous unexpected decrease in the fiscal burden on the young and therefore, it is similar to the effects caused by shocks to expenditures or, in an economy with prices, by inflationary shocks, which erode the value of debt.

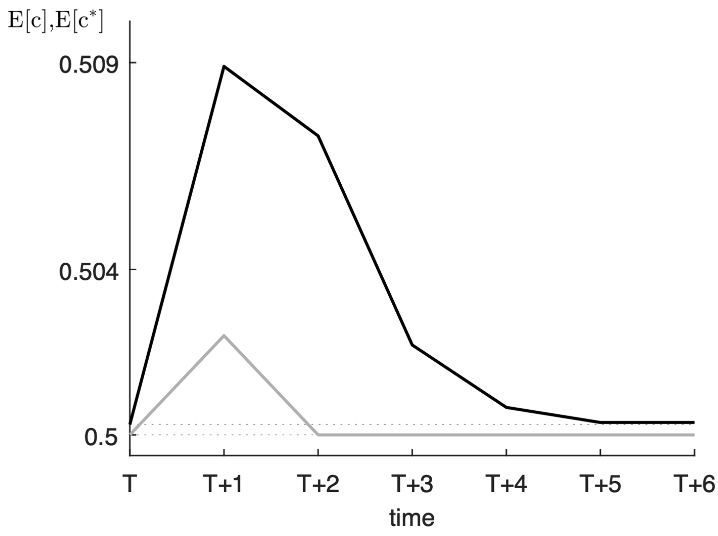

Figure D.1: Impulse Response Functions for a Demographic Shock.

Note: The figure plots the average consumption shares of the young in response to a demographic shock both for the limited enforcement (dark gray) and full enforcement case (light gray). The average is computed starting from the invariant distribution and recomputing the average for all possible sample paths.

E Shooting Algorithm

In the two-state economy in Section 8, under Assumption 5, the optimal consumption depends only on the number of previous state 2s. Let $c^{(n)}(s)$ denote the consumption after n consecutive state 2s and let $\mu^{(n)}$ denote the corresponding multiplier on the state 2 participation constraint. It follows from the first-order condition (12) and the updating rule (15) that

$$c^{(n)}(1) = \frac{\delta}{\beta\nu^{(n-1)} + \delta} \quad \text{and} \quad c^{(n)}(2) = \frac{\delta\nu^{(n)}}{\beta\nu^{(n-1)} + \delta\nu^{(n)}} \quad \text{for } n = 0, 1, 2 \ldots, \infty,$$

where $\nu^{(n)} = 1 + \mu^{(n)}$ and $\nu^{(-1)} = 1$. Let $\nu^{(\infty)} = \lim_{n\to\infty} \nu^{(n)}$. The participation constraint of the young binds in state 2. Hence,

$$\log\left(\frac{\delta\nu^{(n)}}{\beta\nu^{(n-1)}+\delta\nu^{(n)}}\right) + \beta\left(\pi \log\left(\frac{\beta\nu^{(n)}}{\beta\nu^{(n)}+\delta}\right) + (1-\pi)\log\left(\frac{\beta\nu^{(n)}}{\beta\nu^{(n)}+\delta\nu^{(n+1)}}\right)\right) = \upsilon(2). \tag{E.1}$$

Since equation (E.1) holds in the limit

$$\log\left(\frac{\delta}{\beta+\delta}\right) + \beta\left(\pi \log\left(\frac{\beta\nu^{(\infty)}}{\beta\nu^{(\infty)}+\delta}\right) + (1-\pi)\log\left(\frac{\beta}{\beta+\delta}\right)\right) = \upsilon(2). \tag{E.2}$$

Since $v(2) = \log(s(2)) + \beta\pi \log(1 - s(1)) + (1 - \pi) \log(1 - s(2)))$, equation (E.2) can be solved to give:

$$\nu^{(\infty)} = \frac{\delta}{\beta}\left(-1+\left(\left(\frac{\delta}{\beta}\right)^{\frac{1-\pi}{\pi}}\left(\frac{\beta+\delta}{\delta}\right)^{\frac{1+\beta(1-\pi)}{\beta\pi}}(s(2))^{\frac{1}{\beta\pi}}(1-s(2))^{\frac{1-\pi}{\pi}}(1-s(1))\right)^{-1}\right)^{-1}. \tag{E.3}$$

Using equations (E.1) and (E.2), gives a second-order difference equation for $\nu^{(n)}$:

$$\nu^{(n+1)} = \frac{\beta}{\delta}\nu^{(n)}\left(-1+\left(\frac{\beta\nu^{(n)}}{\beta\nu^{(n)}+\delta}\right)^{\frac{\pi}{1-\pi}}\left(\frac{\beta\nu^{(\infty)}+\delta}{\beta\nu^{(\infty)}}\right)^{\frac{\pi}{1-\pi}}\left(\frac{\beta+\delta}{\delta}\right)^{\frac{1}{\beta(1-\pi)}}\left(\frac{\beta+\delta}{\beta}\right)\left(1+\frac{\beta}{\delta}\frac{\nu^{(n-1)}}{\nu^{(n)}}\right)^{-\frac{1}{\beta(1-\pi)}}\right). \tag{E.4}$$

It can be shown that the second-order difference equation in (E.4) has a unique saddle path solution. Since $\nu^{(-1)} = 1$, the solution can be found by a *forward shooting* algorithm to search for $\nu^{(0)}$ such that the absolute difference between $\nu^{(\infty)}$ (given in (E.3)) and $\nu^{(N+1)}$ (given in (E.4)) is sufficiently close to zero for N sufficiently large.

F Pseudo-code for Numerical Algorithms

Algorithms are implemented in MATLAB®. At each iteration, the optimization uses the non-linear programming solver command `fsolve` in Algorithm 1 and command `fmincon` in Algorithm 2. Value function interpolation uses the spline method of the `interp1` command. In a typical example, the value function converges within 15 iterations.

Algorithm 1: Shooting Algorithm

procedure ▷ Find $\nu^{(0)} = 1 + \mu^{(0)}$ in two state economy (Section 8)
- target ← $\nu^{(\infty)}$ ▷ Use equation (E.3) in Appendix E
- tolerance ← $\epsilon > 0$ ▷ $\epsilon = 10^{-10}$
- **repeat**
  - initialization ← $\nu_0^{(0)} > 0$
  - Compute $\nu_0^{(N)}$ for $N = 20$ ▷ Use equation (E.4) in Appendix E
  - $d \leftarrow d(\nu_0^{(N)}, \nu^{(\infty)})$ ▷ $d(\nu_0^{(N)}, \nu^{(\infty)}) = |\nu_0^{(N)} - \nu^{(\infty)}|$
- **until** $d < \epsilon$
- $\nu^{(0)} \leftarrow \nu_0^{(0)}$

end procedure

Algorithm 2: Find Value and Policy Functions

```
procedure                                  ▷ Find solution to functional equation (P1)
    Ω ← [ω_min, ω_max]                     ▷ ω_min and ω_max computed
    gridpoints ← gp                        ▷ Discretize Ω: gp = 200 Chebyshev interpolation points
    tolerance ← ϵ > 0                      ▷ ϵ = 10^{-6}
    J ← V*                                 ▷ V* is first best
    repeat
        Compute TJ from J                  ▷ Use equation (P1) and interpolate
        d ← d(TJ, J)                       ▷ d(TJ, J) = max_ω |TJ(ω) − J(ω)|
        J ← TJ
    until d < ϵ
    V ← J
    Compute g_r(s, ω) and f(s, ω)          ▷ Using the function V just computed.
end procedure
```

Algorithm 3: Computing the Invariant Distribution

```
procedure                                  ▷ Find invariant distribution for x = (s, ω) ∈ X ⊂ ℝ^{nI×1}
    initialization ← a_0 = e(1/nI)         ▷ e = (1, 1, ..., 1) ∈ ℝ^{nI×1}
    Compute a = Πa_0                       ▷ Use the transition probability Π ⊂ ℝ^{nI×nI}
    tolerance ← ϵ > 0                      ▷ ϵ = 10^{-8}
    repeat
        Compute a = Πa                     ▷ a is eigenvector associated with 1
        d ← d(Πa, a)                       ▷ d(Πa, a) = max_x |Πa(x) − a(x)|
        a ← Πa
    until d < ϵ
    ϕ ← a/∑_x a(x)                         ▷ ϕ is normalized invariant distribution
end procedure
```