Optimal Sustainable Intergenerational Insurance*

Francesco Lancia Alessia Russo Tim Worrall

December 2020

Abstract

Optimal intergenerational insurance is examined in a stochastic overlapping generations endowment economy with limited enforcement of risk-sharing transfers. Transfers are chosen by a benevolent planner who maximizes the expected discounted utility of all generations while respecting the participation constraint of each generation. We show that the optimal sustainable intergenerational insurance is history dependent. The risk from a shock is unevenly spread into the future, generating heteroscedasticity and autocorrelation of consumption even in the long run. The optimum can be interpreted as a social security scheme characterized by a minimum welfare entitlement for the old and state-contingent entitlement thresholds.

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

JEL CODES: D64; E21; H55.

*Lancia: University of Salerno, Centre for Economic Policy Research and Center for Studies in Economics and Finance, Via Giovanni Paolo II, 132, 84084 - Fisciano (SA), Italy (e-mail: flancia@unisa.it). Russo: BI Norwegian Business School and Centre for Economic Policy Research, Nydalsveien 37, N-0484 Oslo, Norway (e-mail: alessia.russo@bi.no). Worrall: University of Edinburgh, 30-31 Buccleuch Place, Edinburgh, EH8 9JT, UK (e-mail: tim.worrall@ed.ac.uk). We thank Spiros Bougheas, Francesco Caselli, Gabrielle Demange, Martín Gonzalez-Eiras, Alexander Karaivanov, Paul Klein, Dirk Krueger, Sarolta Laczó, Espen Moen, Iacopo Morchio, Nicola Pavoni, José-Víctor Ríos-Rull, Karl Schlag, Kjetil Storesletten and Aleh Tsyvinski for helpful comments. The paper has benefited from presentations at several conferences and workshops including at the NBER Summer Institute on Macro Public Finance in Cambridge, Massachusetts, the SED Meeting in Edinburgh, the SAET Meeting 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 meetings in Manchester and the Vienna Macroeconomics Workshop. The third author acknowledges the support of the UK Economic and Social Research Council grant ES/L009633/1.

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] Confronted with such shocks, it is desirable to share risk across generations through a social security scheme or other forms of insurance. 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 sustainability is becoming increasingly relevant in many OECD countries as the relative standard of living of the younger generation has worsened in recent decades.[2] If this generational shift persists, then future generations may be less willing to contribute to social security arrangements than in the past. Therefore, a natural question to ask is, how should an optimal intergenerational insurance arrangement be structured when there is limited enforcement of risk-sharing transfers?

Despite its policy relevance, this question has not been fully addressed in the literature on intergenerational risk sharing. The normative approach in this literature investigates the optimal design of intergenerational insurance but neglects the limited enforcement of risk-sharing transfers by assuming that transfers are mandatory. Meanwhile, the positive approach highlights the political limits to intergenerational risk sharing, 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 there is limited enforcement of risk-sharing transfers. Limited enforcement is modeled 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. We say that any arrangement of risk-sharing transfers is sustainable if it satisfies the participation constraint of every generation. An optimal intergenerational insurance arrangement is determined by a benevolent planner who chooses sustainable transfers to maximize the expected discounted utility of all generations.

The model is simple and the economy is stationary. There is a representative agent in each generation and a single, non-storable consumption good. Agents live for two periods: young and old. The endowments of both the young and the old are stochastic. The shock to endowments is identically and independently distributed over time, and

[1] Andrew Glover, Jonathan Heathcote, Dirk Krueger and José-Víctor Ríos-Rull (2020a) find that the Financial Crisis of 2008 had a greater negative impact on the older generation while the young benefited from the fall in asset prices. Andrew Glover, Jonathan Heathcote, Dirk Krueger and José-Víctor Ríos-Rull (2020b) 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 the 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.

there may be aggregate as well as idiosyncratic risk. There is no population growth, no production, no altruism, and no asymmetry of information. 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, Peter 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) and the young will only make a transfer to the old if the promise made to them for their old age at least matches their lifetime expected utility from autarky and they anticipate that the promise will be honored by the next generation.

It is well-known (see, for example, S. Rao Aiyagari and Dan Peled, 1991; Subir Chattopadhyay and Piero Gottardi, 1999) that there are stationary transfers that improve upon autarky when endowments are such that the young wish to defer consumption to old age at a zero net interest rate (Proposition 1). Under this condition, we find that, when the first-best transfers cannot be sustained, there is a trade-off between efficiency and providing incentives for the young to participate. This trade-off is resolved by linking the utility that the young are promised for their old age to the past promise whenever their participation constraint binds. The resulting optimal sustainable intergenerational insurance arrangement is history dependent, even though the environment itself is stationary. The risk from an endowment shock is unevenly spread into the future, generating heteroscedasticity and autocorrelation of consumption.

To understand why there is history dependence, suppose that the first-best transfers would violate the participation constraint of the young in some state. To ensure that the current transfer made by the young is voluntary, either it is reduced, or the transfers promised to them when they are old are increased. Both changes are costly since a smaller current transfer reduces the amount of risk shared today, while increasing the utility 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, there is an optimal trade-off between reducing the current transfer and increasing the future promised utility, which depends both on the current endowment and the current promise. For example, consider some current endowment and current promise such that the future promise is higher than the current promise. If the same endowment state is repeated in the subsequent period, then the young in that period will be called upon to make a larger transfer, which in turn will require a higher promise of future utility to them as well. Thus, the transfer depends not only on the current endowment but also on the history of endowment states.

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 current transfer made by the young (or equivalently, the consumption of the young) and the future promised utility for their old age. Both policy functions depend on the endowment state and the current promise and are weakly increasing in the current promise, for a given endowment state (Lemmas 2 and 3 describe the properties of these two functions). There is a unique fixed point for the future promise that depends on the first-best transfer in that endowment state. For a given endowment state, the future promised utility is increased when the current promise is less than the corresponding fixed point, and it is decreased 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 is not binding. In that case, the future promise is *reset* to the largest value that maximizes the planner's payoff. This sets a minimum value for the future promise that is independent of the current endowment state and the history of endowment states. The optimum can be interpreted as a social security or state pension scheme with a minimum welfare entitlement for the old, which is determined by the resetting level, and state-contingent thresholds for welfare entitlements, which are determined by the fixed points.

The resetting property shows that the effect of a shock does not last forever and it is used to prove strong convergence to a unique non-degenerate invariant distribution of promised utilities (Proposition 4). Since the invariant distribution of promised utility is non-degenerate and the optimal insurance arrangement is history dependent, consumption fluctuates across states and over time, even in the long run. This is in stark contrast to the situation in which either transfers from the young to the old are enforced or there is no risk. In the former case, the promised utility is constant over time, except possibly in the initial period (Proposition 2). In the latter case, the promised utility is constant in the long run, although there may be a finite initial phase during which the promised utility is declining (Proposition 3). In either case, there is no inefficiency in the long run. Thus, *both* risk and limited enforcement are necessary for the long-run distribution of promised utility to be non-degenerate and for there to be inefficiency in the long run.

We use measures of entropy (see, for example, David Backus, Mikhail Chernov and Stanley Zin, 2014) and the bound on the variability of the implied yields introduced by Ian Martin and Stephen Ross (2019) to understand how risk is shared across generations. These risk measures are derived from the set of state prices and implied yields that correspond to the optimal sustainable intergenerational insurance. The implied yields increase with the current promise, indicating that generations born in a period with a higher promised utility bear greater risk. Moreover, the yield on all very long bonds converges to a long-run yield determined by the Perron root of the state price matrix, indicating that the exposure to a shock dies out as the time horizon becomes long enough. For some parameter values, the long-run yield is equal to the planner's discount factor (Proposition 5). We present an example with two endowment states. The solution for the two-state example can be derived by the use of a simple shooting algorithm without the need to solve a functional equation. We provide a closed-form solution for the bound on the variability of the implied yields (Proposition 6) and show that the invariant distribution of promised utility is a transformation of a geometric distribution with infinite but countable support (Proposition 7).

The existing literature on risk sharing in overlapping generations models has several strands. One strand considers public policies or other non-market mechanisms that improve risk sharing through a social security scheme (see, for example, Walter Enders and Harvey Lapan, 1982; Robert J. Shiller, 1999; Antonio Rangel and Richard Zeckhauser, 2000). In this strand of the literature, however, transfers are mandatory and attention is restricted to stationary transfers, in contrast to the voluntary and history-dependent transfers considered here. Our result on history dependence is foreshadowed in a mean-variance setting by Roger Gordon and Hal Varian (1988), who establish that any time-consistent optimal intergenerational risk-sharing agreement is non-stationary. Two closely related papers are Laurence Ball and Gregory Mankiw (2007) on risk sharing and Nobuhiro Kiyotaki and Shengxing Zhang (2018) on limited commitment. Ball and Mankiw (2007) consider how risk is allocated across generations in a complete-markets equilibrium in which all generations can trade contingent claims before they are born. They find that shocks are evenly spread across generations in an optimal allocation, meaning that consumption follows a random walk. This allocation cannot be sustainable because it implies that the participation constraint of some future generation is violated almost surely. In contrast, we show that shocks are unevenly spread across future generations to ensure that all participation constraints are met. Kiyotaki and Zheng (2018) study an overlapping generations model with limited commitment. However, in that paper, a worker trained by a firm cannot commit to staying with the current employer. Hence, the focus is on investment inefficiency, rather than intergenerational insurance.

A second strand of the literature provides simple necessary and sufficient criteria for Pareto optimality. Aiyagari and Peled (1991) derive a dominant root condition for interim optimality in an endowment economy with a finite state space. This approach has been extended by several authors (see, for example, Rodolfo Manuelli, 1990; Gabrielle Demange and Guy Laroque, 1999; Gaetano Bloise and Filippo L. Calciano, 2008). Pamela

Labadie (2004) shows how to interpret this characterization in terms of ex ante Pareto optimality. We provide a similar characterization in which the Pareto weights are determined endogenously by the participation constraints and the history of endowment states. This is discussed in Section 9.

A third strand of the literature studies political economy models of social security. Thomas Cooley and Jorge Soares (1999); Michele Boldrin and Aldo Rustichini (2000); Martín Gonzalez-Eiras and Dirk Niepelt (2008), for example, analyze settings in which agents vote on the tax and benefit rates for intergenerational transfers. In these papers, a social security scheme corresponds to the sub-game (or Markov) perfect equilibria of a repeated (or dynamic) game. Typically, it is assumed that voters in the working-age group are pivotal. In equilibrium, a social security scheme is supported by current workers in the expectation that future workers will do the same. The equilibria of these games are not necessarily Pareto optimal. In contrast, the approach presented here identifies constrained Pareto optimal intergenerational transfers that each generation unanimously agrees to respect.

The paper is methodologically related to the literature on risk sharing and limited commitment with infinitely-lived agents. Two polar cases are examined by this literature: two infinitely-lived agents (see, for example, Jonathan Thomas and Tim Worrall, 1988; Narayana Kocherlakota, 1996) and a continuum of infinitely-lived agents (see, for example, Jonathan Thomas and Tim Worrall, 2007; Dirk Krueger and Fabrizio Perri, 2011; Tobias 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 important gap by providing an analysis of optimal intergenerational insurance with limited commitment. We discuss how the results of the paper relate to this literature in Section 9.

The paper proceeds as follows. Section 1 sets out the model. Section 2 considers two benchmarks: one with enforcement of transfers from the young to the old and the other without risk. Section 3 provides the main results. Section 4 establishes convergence to an invariant distribution. Section 5 considers how risk is allocated at the invariant distribution. Section 6 studies a case with two endowment shocks and Section 7 examines comparative static results for that case. Section 8 shows how the model can be extended. Section 9 discusses the results in the context of the literature and Section 10 concludes. The 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. Each generation is composed of a single agent.[3] The agent is young in the first period of life 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 one that lives for just one period.

At each t, agents receive an endowment of a perishable consumption good that depends on the state of the world $s_t \in \mathcal{S} := \{1, 2, \ldots, S\}$ with $S \geq 2$. The endowments of the young and the old in state s_t are $e^y(s_t)$ and $e^o(s_t)$, and the aggregate endowment is $e(s_t) := e^y(s_t) + e^o(s_t)$. Endowments are finite and strictly positive. Denote the history of states up to and including date t by $s^t := (s_0, s_1, ..., s_t) \in \mathcal{S}^t$ and the probability of reaching history s^t by $\pi(s^t)$ where $\pi(s^{t-1})\pi(s_t \mid s^{t-1}) = \pi(s^{t-1}, s_t)$. We assume that states are identically and independently distributed (hereafter, i.i.d.). Hence, $\pi(s^t) = \pi(s_0) \cdot \ldots \cdot \pi(s_t)$ where $\pi(s_t)$ is the probability of state s_t and $\pi(s_{t+1} \mid s^t) = \pi(s_{t+1})$. All information about endowments and the probability distribution is public: there is complete information. Let $c^y(s^t)$ and $c^o(s^t)$ be the per-period consumption of the young and the old. There is no technology for saving or investment and hence, the aggregate endowment is consumed within the period, that is, $c^y(s^t) + c^o(s^t) = e(s_t)$. Endowments depend only on the current state whereas consumption can, in principle, depend on the history of states. In autarky, agents consume only their own endowments, that is, $c^y(s^{t-1}, s_t) = e^y(s_t)$ and $c^o(s^{t-1}, s_t) = e^o(s_t)$ for all t and (s^{t-1}, s_t).

Each generation is born after that period's uncertainty is resolved and when current endowments 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 endowment risk. The lifetime endowment utility of an agent born in state s_t is:

$$\hat{v}(s_t) := u(e^y(s_t)) + \beta \sum\nolimits_{s_{t+1}} \pi(s_{t+1}) u(e^o(s_{t+1})),$$

where $\beta \in (0, 1]$ is the generational discount factor and $u(\cdot)$ is the utility function, common to both the young and the old. Since endowments are positive and finite, $\hat{v}(s_t)$ is bounded.

[3]The assumption that there is a representative agent in each generation makes it possible to focus on intergenerational risk sharing. By doing so, however, we ignore questions about inequality within generations and its evolution over time.

Assumption 1. *The utility function* $u \colon \mathbb{R}_+ \to \mathbb{R} \cup \{-\infty\}$ *is strictly increasing, strictly concave, thrice continuously differentiable, it satisfies the Inada conditions and*

$$u(0) < \min_s \left\{ u(e^y(s)) - \beta \sum\nolimits_{s'} \pi(s') \left(u(e(s')) - u(e^o(s')) \right) \right\},$$

where s' *is the endowment state at the next date.*

The latter part of Assumption 1 is sufficient to guarantee that young agents have positive consumption at the optimal sustainable intergenerational insurance. Obviously, Assumption 1 is satisfied when $\lim_{c \to 0} u(c) = -\infty$, as is the case for the logarithmic utility function or for the CRRA utility function with a coefficient of relative risk aversion greater than one.

Let $\lambda(s)$ denote the ratio of marginal utilities of the young and the old in autarky in state s:

$$\lambda(s) := \frac{u_c(e^y(s))}{u_c(e^o(s))}.$$

If $\lambda(s) > 1$, then $e^y(s) < e^o(s)$ and the old are wealthier in that state. States are ordered so that $\lambda(S) \geq \lambda(S-1) \geq \cdots \geq \lambda(1)$. Thus, the relative endowment of the old, compared to the young, increases with the state.[4] Since $\lambda(s)$ varies across states, it is desirable to share risk across generations. In the absence of a storage technology and because the young are born after uncertainty is resolved, the only possibility for intergenerational insurance is through transfers between the young and the old. However, we require that all transfers must be voluntary. That is, agents make a transfer only if it is in their interest to do so. We assume that any generation that does not make a transfer when called upon to do so is punished by receiving no transfer when they reach old age. Therefore, for every history of shocks, the intergenerational insurance must provide all generations with at least the same lifetime utility they derive from consuming their endowments.

Benevolent Social Planner Consider the problem of a benevolent social planner who chooses an intergenerational insurance rule, that is, a function $\tau(s^t)$ that specifies the transfer between the young and the old for each history s^t. Since the aggregate endowment is consumed, $\tau(s^t) = e^y(s_t) - c^y(s^t)$ and $c^o(s^t) = e(s_t) - c^y(s^t)$. 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. Hence, the transfer from the young to the old

[4] When two states have the same value of $\lambda(\cdot)$, we use the convention that the states are ordered by the aggregate endowment, that is, higher states are associated with higher aggregate endowment. A special case is where states can be ordered so that the endowment of the old is increasing in s while that of the young is decreasing in s.

is always non-negative because the old would default if they were ever called upon to make a transfer. Therefore, $c^y(s^t) \leq e^y(s_t)$ for every history s^t and this non-negativity constraint can be expressed by requiring $c^y(s^t) \in \mathcal{Y}(s_t) := [0, e^y(s_t)]$ for every history s^t. For the young, the analogous participation constraint requires that any transfer made is compensated sufficiently by transfers received in old age, so that they are no worse off than reneging on the transfer today and receiving the corresponding autarkic utility. That is,

$$u(c^y(s^t)) + \beta \sum\nolimits_{s_{t+1}} \pi(s_{t+1}) u(e(s_{t+1}) - c^y(s^t, s_{t+1})) \geq \hat{v}(s_t) \qquad \forall\, s^t. \tag{1}$$

Let $\Lambda := \{\{c^y(s^t) \in \mathcal{Y}(s_t)\}_{t=0}^{\infty} \mid (1)\}$ denote the constraint set faced by the planner. Since utility is strictly concave and the constraints in (1) are linear in utility, Λ is convex and compact.

Definition 1. *An Intergenerational Insurance rule is sustainable if the history-dependent sequence* $\{c^y(s^t)\}_{t=0}^{\infty} \in \Lambda$.

The planner seeks to address the conflict between generations by choosing a sustainable intergenerational insurance rule that maximizes a weighted sum of expected utilities of all generations. We suppose that the planner discounts the expected utility of future generations by a factor $\delta \in (0,1)$ and weighs the utility of the initial old by β/δ. In general, δ may be different from β.[5]

Definition 2. *A Sustainable Intergenerational Insurance rule is optimal if it maximizes the weighted sum of the expected utilities of all generations given by:*

$$\frac{\beta}{\delta}\sum_{s_0}\pi(s_0)u(e(s_0)-c^y(s_0))+\mathbb{E}_0\left[\sum_t \delta^t \left(u(c^y(s^t))+\beta\sum_{s_{t+1}}\pi(s_{t+1})u(e(s_{t+1})-c^y(s^t,s_{t+1}))\right)\right], \tag{2}$$

where $\mathbb{E}_0$ *is the expectation over all histories of shocks, subject to the constraint*

$$\sum\nolimits_{s_0} \pi(s_0) u(e(s_0) - c^y(s_0)) \geq \omega, \tag{3}$$

for some given ω.

Let $V(\omega)$ denote the value function corresponding to a solution of the optimization problem in Definition 2. The function $V(\omega)$ traces out the Pareto frontier between the expected utility of the current old and the expected discounted utility of all future gen-

[5]Geometric discounting is commonly assumed for a planner in the literature (see, for example, Emmanuel Farhi and Iván Werning, 2007). Giving a weight of β/δ to the initial old means that the relative weights on the old and the young are the same in every period.

erations.[6] It is defined on a set $\Omega = [\omega_{\min}, \omega_{\max}]$ where $\omega_{\min} := \sum_{s_t} \pi(s_t) u(e^o(s_t))$ is the expected autarkic utility of the old and $\omega_{\max}$ is the highest feasible value of expected utility of the old consistent with the participation constraints ($\omega_{\max}$ is discussed in more detail below). Constraint (3) requires that the planner, in making the choice at $t = 0$ and before the state s_0 is known, chooses the consumption of the young, $c^y(s_0)$, such that the expected utility offered to the initial old is at least ω. Let

$$\omega_0 := \sum\nolimits_{s_0} \pi(s_0) u(e(s_0) - \tilde{c}^y(s_0)), \tag{4}$$

where the notation $\tilde{c}^y(s_0)$ is adopted to emphasize that it is part of the optimal solution. Clearly, for $\omega \leq \omega_0$, constraint (3) does not bind and $V(\omega) = V(\omega_0)$; whereas for $\omega > \omega_0$, constraint (3) binds and $V(\omega) < V(\omega_0)$.

Preliminaries The existence of a sustainable non-autarkic allocation can be addressed by considering small stationary transfers (that is, transfers depending only on the current endowment state). Denote the intertemporal marginal rate of substitution between consumption when young in state s and consumption when old in state s', evaluated at autarky, by $\hat{m}(s, s') := \beta u_c(e^o(s'))/u_c(e^y(s))$ and let $\hat{q}(s, s') := \pi(s')\hat{m}(s, s')$. The terms $\hat{m}(s, s')$ and $\hat{q}(s, s')$ correspond to the stochastic discount factor and the state prices in an equilibrium model. Denote the matrix of terms $\hat{q}(s, s')$ by $\hat{Q}$. A sustainable and non-autarkic allocation exhausting the aggregate endowment and satisfying $c^y(s^t) \in \mathcal{Y}(s_t)$ and (1) exists whenever the Perron root (the leading eigenvalue) of $\hat{Q}$ is greater than one (see, for example, Aiyagari and Peled, 1991; Chattopadhyay and Gottardi, 1999). In this case, there are non-negative stationary transfers that improve the lifetime utility of the young in each state. Since the endowment shocks are transitory, the matrix $\hat{Q}$ has rank one and the Perron root is equal to the trace of the matrix. We assume that this trace is larger than one.

Assumption 2.

$$\beta \sum\nolimits_s \pi(s) \lambda(s)^{-1} > 1.$$

If there is just one state, then Assumption 2 reduces to the standard Samuelson condition: $\beta u_c(e^o) > u_c(e^y)$. In this case, it is well known that there are Pareto-improving transfers from the young to the old (Paul Samuelson, 1958). Assumption 2 is the generalisation to the stochastic case and a natural assumption given that our focus is on transfers

[6] More precisely, the function $\tilde{V}(\omega) := V(\omega) - \omega$ can be viewed as a Pareto frontier that trades off the expected utility of the current old against the planner's valuation of the expected discounted utility of all other generations.

to the old. Given Assumption 2, it follows that the constraint set Λ is non-empty (a simple proof is given in Part B of the Supplementary Appendix).[7]

Proposition 1. *Under Assumption 2, there exists a non-autarkic and stationary Sustainable Intergenerational Insurance rule.*

Since $\lambda(s)$ is increasing in s, Assumption 2 implies that $\beta > \lambda(1)$, or equivalently, $\beta u_c(e^o(1)) > u_c(e^y(1))$. That is, the Samuelson condition is satisfied in state 1. We shall assume that the opposite is true in state S.

Assumption 3.

$$\lambda(S) \geq \frac{\beta}{\delta}.$$

Since $\delta < 1$, Assumption 3 implies that $\beta < \lambda(S)$, or equivalently, $\beta u_c(e^o(S)) < u_c(e^y(S))$. We make Assumption 3 for two reasons. First, it shows that the analysis below does not depend on the Samuelson condition applying in every state. Second, it provides a simple sufficient condition for the strong convergence result of Section 4.

2 Two Benchmarks

Before turning to the characterization of the optimal sustainable intergenerational insurance, it is useful to consider two benchmark cases that serve to illustrate the inefficiencies generated by the presence of limited enforcement and uncertainty. In the first, the planner can enforce transfers from the young to the old but not from the old to the young. That is, the planner respects the participation constraint of the old but ignores the participation constraint of the young. We refer to this as the first-best outcome. The second benchmark has no uncertainty but requires that the planner respects the participation constraints of both the young and the old.

[7] 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 one. That is, for each state s,

$$\sum\nolimits_{s'} \hat{q}(s,s') = \beta \sum\nolimits_{s'} \pi(s') \frac{u_c(e^o(s'))}{u_c(e^y(s))} > 1.$$

This implies that in autarky and in every state the young would, if they could, prefer to save for their old age, even at a zero net rate of interest.

First Best We assume there is uncertainty, $S \geq 2$, but suppose that the planner ignores the participation constraints of the young. Let $\Lambda^* := \{c^y(s^t) \in \mathcal{Y}(s_t)\}_{t=0}^{\infty}$ denote the set of transfers without the constraints in (1).[8]

Definition 3. *An Intergenerational Insurance* $\{c^y(s^t)\}_{t=0}^{\infty} \in \Lambda^*$ *is first best if it maximizes the objective function* (2) *subject to constraint* (3)*.*

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

$$\frac{u_c(c^{y*}(s^t))}{u_c(e(s_t) - c^{y*}(s^t))} = \max\left\{\frac{\beta}{\delta}, \lambda(s_t)\right\} \qquad \forall\, s^t. \tag{5}$$

Condition (5) shows that $c^{y*}(s^t)$ is independent of the history s^{t-1} and depends only on the current state s_t, that is, transfers are stationary. Let $\tau^*(s) = e^y(s) - c^{y*}(s)$ denote the first-best transfer conditional on state s. The transfer $\tau^*(s) = 0$ for states in which the participation constraint of the old binds, that is, for states in which $\beta/\delta \leq \lambda(s)$. Under Assumption 3, there is always one such state, and hence, the first-best transfer is not positive in every state. The transfer $\tau^*(s) > 0$ for states in which $\beta/\delta > \lambda(s)$. For such states, condition (5) shows that risk is shared efficiently across generations. Condition (5) is the familiar optimal risk-sharing condition from Karl Borch (1962), modified to account for the constraint that transfers are only from the young to the old. It is designated by Labadie (2004) as an *equal-treatment Pareto optimum* in which all generations are weighed equally by the planner. It can be seen from condition (5) that for states in which transfers are positive $\tau^*(s)$ is increasing in β, since a higher β means more weight is placed on the utility of the old to whom the transfer is made, whereas $\tau^*(s)$ is decreasing in δ, since a higher δ means more weight is put on the utility of the young who make the transfer.

Let $\omega^* := \sum_s \pi(s)u(e(s) - c^{y*}(s))$ denote the expected utility of the old in the first-best solution. From the definition in (4), it follows that $\omega_0 = \omega^*$. Now consider constraint (3). For $\omega \leq \omega^*$, constraint (3) does not bind and the first-best consumption $c^{y*}(s_0)$ is determined by condition (5), as in every other period $t > 0$. For $\omega > \omega^*$, constraint (3) binds and the initial transfers to the old are correspondingly higher while maintaining a constant ratio of marginal utilities across states. In particular, if $\omega > \omega^*$, then there is a $\nu_0 > 0$ (the multiplier associated with constraint (3) is $(\beta/\delta)\nu_0$) such that consumption

[8] We could define the first best to allow the planner to enforce transfers from the old to the young. The reason for presenting the first best as we do is to show that the optimal allocation is stationary and therefore, any history dependence in the optimal sustainable intergenerational insurance rule derives from the imposition of the participation constraints of the young. Hereafter, the asterisk designates the first-best outcome without the participation constraints of the young.

at $t = 0$ satisfies:

$$\frac{u_c(c^{y*}(s_0))}{u_c(e(s_0) - c^{y*}(s_0))} = \max\left\{\frac{\beta}{\delta}(1 + \nu_0), \lambda(s_0)\right\} \qquad \forall\, s_0, \tag{6}$$

and constraint (3) holds with equality.

Denote the per-period payoff to the planner with the first-best allocation by $v^* := \sum_s \pi(s)u(c^{y*}(s)) + (\beta/\delta)\omega^*$ and the expected discounted payoff to the planner for $\omega \in \Omega$ by $V^*(\omega)$. The first-best outcome is summarized in the following proposition.[9]

Proposition 2. *(i) The consumption of the young $c^{y*}(s^t)$ is stationary and satisfies condition* (5) *for $t > 0$ and condition* (6) *for $t = 0$. (ii) The value function $V^*: \Omega \to \mathbb{R}$ 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 $\omega_{\max} := \sum_s \pi(s)u(e(s))$ and $\lim_{\omega \to \omega_{\max}} V^*_\omega(\omega) = -\infty$.*

Note that when constraint (3) is imposed and $\omega > \omega^*$, the consumption of the young is lower than the first-best consumption given by condition (5), but only in the initial period. There is immediate convergence to the stationary distribution with a single mass point at $\{\omega^*\}$ after the initial period.

Deterministic Economy We now consider a deterministic economy with only one endowment state but in which any intergenerational transfer respects the participation constraint of the young as well as that of the old. In this case, Assumption 2 reduces to the standard Samuelson condition. This assumption together with the strict concavity of the utility function implies that there is a unique consumption $c^y_{\min} < e^y$, which is the lowest *stationary* consumption of the young that satisfies the participation constraint

$$u(c^y_{\min}) + \beta u(e - c^y_{\min}) = \hat{v} := u(e^y) + \beta u(e^o). \tag{7}$$

The corresponding utility of the old is $\omega_{\max} = u(e - c^y_{\min})$. Analogously to condition (5), the first-best consumption c^{y*} satisfies $u_c(c^{y*})/u_c(e - c^{y*}) = \max\{\beta/\delta, \lambda\}$ and the corresponding utility of the old is $\omega^* = u(e - c^{y*})$. Since $\beta/\delta > \beta > \lambda$, the participation constraint of the old is satisfied at c^{y*}. Whether the participation constraint of the young is satisfied at c^{y*} depends on the value of δ. If and only if δ is above a critical value, then $c^{y*} > c^y_{\min}$ and the first-best consumption is sustainable.

Denote the consumption of the young at t by c^y_t and the corresponding utility of the old by $\omega_t = u(e - c^y_t)$. Consider the solution to the maximization problem in (2)

[9] The proof of Proposition 2 is omitted because it follows from standard arguments. Nonetheless, the properties of the function $V^*(\omega)$ are mirrored in Proposition 3 and Lemma 1, given below, which do respect the participation constraints of the young.

with the participation constraints of the young given by $u(c_t^y) + \beta u(e - c_{t+1}^y) \geq \hat{v}$ for $t \geq 0$. The solution to this problem is $c_t^y = \max\{c^{y*}, c_{\min}^y\}$ for all t. Now consider ω in constraint (3). If $\omega \leq \omega^*$, then it is optimal to set $c_t^y = \max\{c^{y*}, c_{\min}^y\}$ for all t. On the other hand, consider a case in which δ is large enough such that the first-best consumption is sustainable and $\omega \in (\omega^*, \omega_{\max})$. In this case, at $t = 0$, c_0^y must satisfy $u(e - c_0^y) \geq \omega$, which requires that $c_0^y < c^{y*}$. Clearly, it is desirable to set c_0^y such that $u(e - c_0^y) = \omega$ and $c_1^y = c^{y*}$. However, setting $c_1^y = c^{y*}$ may violate the participation constraint of the young. In such a case, c_1^y has to be chosen to satisfy $u(c_0^y) + \beta u(e - c_1^y) = \hat{v}$, which implies that $c_1^y < c^{y*}$. Repeating this argument for $t > 1$ shows that given c_t^y, c_{t+1}^y either satisfies $u(c_t^y) + \beta u(e - c_{t+1}^y) = \hat{v}$ or $c_{t+1}^y = c^{y*}$ if $u(c_t^y) + \beta u(e - c^{y*}) \geq \hat{v}$. It is useful to express this rule in terms of a policy function that determines the next-period value of the utility of the old, ω', as a function of the current value ω:

$$\omega' = f(\omega) := \begin{cases} \omega^* & \text{for } \omega \in [\omega_{\min}, \omega^c] \\ \frac{1}{\beta}\left(\hat{v} - u\left(e - u^{-1}(\omega)\right)\right) & \text{for } \omega \in (\omega^c, \omega_{\max}] \end{cases}, \tag{8}$$

where $\omega_{\min} = u(e^o)$ and $\omega^c := u(e - u^{-1}(\hat{v} - \beta\omega^*))$. It follows from the strict concavity of the utility function that $\omega^c > \omega^*$. The function $f(\omega)$ is increasing and convex in ω as illustrated in Figure 1. The dynamic evolution of ω_t is straightforwardly derived from $f(\omega)$: for $\omega_t \in [\omega_{\min}, \omega^c]$, $\omega_{t+1} = \omega^*$ 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 within a finite period of time. Intuitively, if the utility of the old is large (or equivalently, the consumption of the young is low), then the planner would like to reduce ω to ω^* (or equivalently, raise the consumption of the young to c^{y*}) as fast as possible to improve welfare. But if the consumption of the next-period young is raised too much, it will violate the participation constraint of the current young. The presence of limited enforcement means that the consumption of the young has to be raised slowly.

Denote the per-period payoff to the planner with the first-best allocation in the absence of uncertainty by $v^* := u(c^{y*}) + (\beta/\delta)\omega^*$ and the expected discounted payoff to the planner for $\omega_t \in \Omega$ by $V(\omega_t)$. The optimal solution for the deterministic case with sustainable ω^* is summarized in Proposition 3.[10]

Proposition 3. *(i) If $\omega \in [\omega_{\min}, \omega^*]$, then the consumption of the young is $c_t^y = c^{y*}$ for $t \geq 0$, where $u_c(c^{y*})/u_c(e - c^{y*}) = \beta/\delta$. (ii) If $\omega \in (\omega^*, \omega_{\max}]$, then the utility of the old ω_{t+1} satisfies equation* (8)*. There exists a finite $\hat{t}$ such that ω_t is monotonically decreasing for $t < \hat{t}$ and $\omega_t = \omega^*$ for $t \geq \hat{t}$. Likewise, c_t^y is monotonically increasing for $t < \hat{t}$ and*

[10] The proof of Proposition 3 is provided in Part B of the Supplementary Appendix.

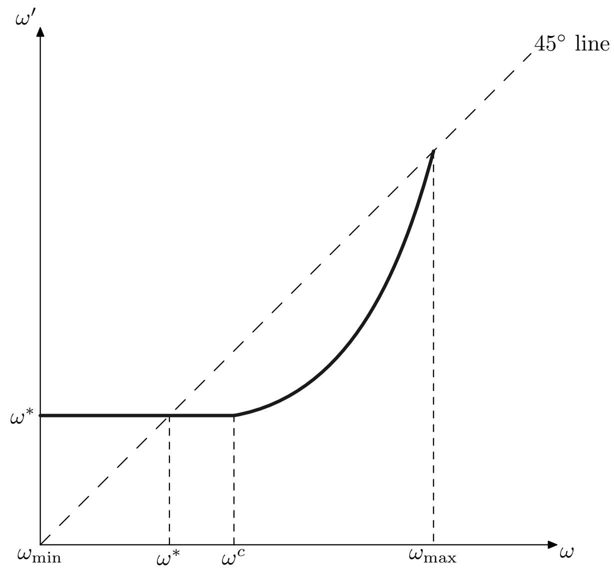


Figure 1: Policy Function in the Deterministic Case.

Note: The solid line is the deterministic policy function $f(\omega)$. For any initial $\omega \in (\omega_{\min}, \omega_{\max})$, ω_t converges to ω^*.

$c_t^y = c^{y*}$ *for* $t \geq \hat{t}$. *(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 where $c_t^y = c^{y*}$, within a finite period of time. Hence, the long-run distribution of ω is degenerate and for the case where $c^{y*} > c^y_{\min}$, it has a single mass point at $\{\omega^*\}$.

In the following sections, we show that in the case where the participation constraint of the young binds and there is risk, the long-run distribution of ω is non-degenerate. The benchmarks show that a combination of both limited enforcement of transfers and risk is necessary for the long-run distribution to be non-degenerate and for there to be inefficiency in the long run.

3 Optimal Policy Functions

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. To do so, we rule out the case in which the first-best outcome is sustainable and consider the case in which the first-best transfer violates the participation constraint of the young in at least one state.

Assumption 4. *There is at least one state $s \in \mathcal{S}$ such that*

$$u(c^{y*}(s)) + \beta \sum\nolimits_{s'} \pi(s')u(e(s') - c^{y*}(s')) < \hat{v}(s) = u(e^{y}(s)) + \beta\omega_{\min}.$$

We reformulate the optimization problem described in Definition 2 recursively. This is possible because the states are i.i.d. and all constraints are forward looking. Our characterization is similar to the promised-utility approach developed by Thomas and Worrall (1988), among others. For simplicity of notation, we often omit the t indexes and use primes to denote next-period variables. At each t, the expected utility $\omega \in \Omega$ promised to the current old embodies information about the history of shocks. The problem at each t is to determine the state-contingent consumption of the young, $c^{y}(s)$, and the state-contingent promise of expected utility, $\omega'(s)$, made to the current young for their old age.

The participation constraint of the young, constraint (1), can be rewritten as:

$$u(c^{y}(s)) + \beta\omega'(s) \geq u(e^{y}(s)) + \beta\omega_{\min} \qquad \forall\, s \in \mathcal{S}. \tag{9}$$

The participation constraint of the old is subsumed by the requirement that $c^{y}(s) \in \mathcal{Y}(s)$ for all s. In each period, the utility promises made to the young must be feasible:

$$\omega'(s) \leq \omega_{\max} \qquad \forall\, s \in \mathcal{S}; \tag{10}$$

and the expected utility of the current old must be at least that previously promised:

$$\sum\nolimits_{s} \pi(s)u(e(s) - c^{y}(s)) \geq \omega. \tag{11}$$

Constraint (11) is a promise-keeping constraint and is analogous to constraint (3), but is now required to hold in every period.[11] Let $\Phi := \{\{c^{y}(s) \in \mathcal{Y}(s), \omega'(s) \in \Omega\}_{s\in\mathcal{S}} \mid (9), (10) \text{ and } (11)\}$ denote the constraint set. Since utility is strictly concave and Ω is an interval, it is easily verified that Φ is convex and compact. The value function $V(\omega)$ satisfies the following functional equation:

$$V(\omega) = \max_{\{c^{y}(s),\omega'(s)\}_{s}\in\Phi} \left[\sum\nolimits_{s} \pi(s) \left(\tfrac{\beta}{\delta} u(e(s) - c^{y}(s)) + u(c^{y}(s)) + \delta V(\omega'(s))\right)\right]. \tag{12}$$

Denote the state vector by $x := (\omega, s)$ and the two stochastic policy functions that solve (12) by $c^{y}(x)$ and $f(x)$, which are the optimal consumption of the young and

[11] Note that $u^{-1}(\omega)$ is the promise offered to the old expressed in certainty-equivalent consumption.

the optimal promise of expected utility for their old age when the current old have a utility promise of ω and the current endowment state is s. The optimal allocation can be computed recursively, starting from an $\omega_0 \in \Omega$ (how ω_0 is determined is discussed below) and solving for the consumption of the young and the future promised utilities in the maximization problem of equation (12). The future promised utilities are then used to solve the dynamic program for the next period given the endowment state that is realized next period, and so on. The existence of the optimal allocation is guaranteed by the compactness of Ω and Φ. Uniqueness is guaranteed by the strict concavity of $u(\cdot)$.

The function $V(\omega)$ cannot be found by a standard contraction mapping argument starting from an arbitrary value function because the value function associated with the allocation in autarky also satisfies the functional equation (12). Nevertheless, a similar approach can be used to iterate the value function, starting from the first-best value function derived in Proposition 2. Following the arguments of Jonathan Thomas and Tim Worrall (1994), it can be shown that the limit of this iterative mapping is the optimal value function $V(\omega)$.[12] Proposition 2 established that the first-best value function is non-increasing, differentiable and concave, and the limit value function inherits these properties.[13]

Lemma 1. *The value function $V\colon \Omega \to \mathbb{R}$ is non-increasing, concave and continuously differentiable in ω, where $\Omega = [\omega_{\min}, \omega_{\max}]$ and $\omega_{\min} < \omega_{\max} < \sum_s \pi(s)u(e(s))$. There is an $\omega_0 \in (\omega_{\min}, \omega^*)$ such that $V(\omega)$ is constant for $\omega \leq \omega_0$ and is strictly decreasing and strictly concave for $\omega > \omega_0$ with $V_\omega(\omega_0) = 0$ and $\lim_{\omega \to \omega_{\max}} V_\omega(\omega) = -(\beta/\delta)\bar{\nu}$, where $\bar{\nu} \in \mathbb{R}_+ \cup \{\infty\}$.*

Concavity of the value function follows from the concavity of the objective function and the convexity of the constraint set. The lower endpoint $\omega_{\min}$ of the domain of $V(\omega)$ is the autarkic value since zero transfers are feasible. The upper endpoint $\omega_{\max}$ is determined by the choice of the consumption of the young and the promised utility that maximize the expected utility of the current old subject to constraints (9) and (10). This is itself a strictly concave programming problem and has a unique solution.[14] The latter part of Assumption 1 is sufficient to guarantee that $\omega_{\max} < \sum_s \pi(s)u(e(s))$. Differentiability follows because the constraint set satisfies a linear independence constraint qualification when $\omega \in [\omega_{\min}, \omega_{\max})$. The left-hand derivative of $V(\omega)$ evaluated at $\omega_{\max}$ can be finite

[12] It is this iteration, starting from the first-best value function, that is used to compute the optimal value function when we solve numerical examples in Sections 6 and 7.

[13] The proof of Lemma 1 is found in Part B of the Supplementary Appendix. The proof is standard though differentiability follows because a linear independence constraint qualification condition is satisfied.

[14] Finding $\omega_{\max}$ is straightforward since the value function $V(\omega)$ does not enter into the constraint set.

or infinite depending on whether or not $\omega_{\max}$ is part of the ergodic set, a question that is discussed in more detail below.

Given Lemma 1, the planner chooses an initial promise $\omega_0 = \sup\{\omega \mid V_\omega(\omega) = 0\}$. For $\omega < \omega_0$, constraint (11) does not bind and the promise to the initial old can be increased without reducing the planner's payoff. Therefore, attention can be restricted to promises $\omega \geq \omega_0$. The initial promise ω_0 is determined as part of the optimal solution and depends, in general, on all parameter values. However, Assumption 4 implies that the first-best outcome violates one of the participation constraints in (9) and hence, $\omega_0 < \omega^*$.

We now turn to the properties of the policy functions $f(x)$ and $c^y(x)$. Given the differentiability of the value function, the first-order conditions for the programming problem in equation (12) are:

$$\frac{u_c(c^y(x))}{u_c(e(s) - c^y(x))} = \frac{\beta}{\delta}\left(\frac{1 + \nu(\omega) + \eta(x)}{1 + \mu(x)}\right), \tag{13}$$

$$V_\omega(f(x)) = -\tfrac{\beta}{\delta}\left(\mu(x) - \zeta(x)\right), \tag{14}$$

where $\pi(s)\mu(x)$ are the multipliers associated with the participation constraints of the young (9), $\beta\pi(s)\zeta(x)$ are the multipliers on the upper bound for promised utility (10), $(\beta/\delta)\nu(\omega)$ is the multiplier associated with the promise-keeping constraint (11) and $(\beta/\delta)\pi(s)\eta(x)$ are the multipliers associated with the non-negativity constraints on transfers. Given the concavity of the programming problem, conditions (13) and (14) are necessary and sufficient. There is also an envelope condition:

$$V_\omega(\omega) = -\tfrac{\beta}{\delta}\nu(\omega). \tag{15}$$

Taken together, equations (14) and (15) imply the following updating property:

$$\nu(\omega') = \mu(x) - \zeta(x). \tag{16}$$

Equation (16) is easily interpreted. For simplicity, suppose that the upper bound constraint on promises and the non-negativity constraint on transfers do not bind, that is, $\zeta(x) = \eta(x) = 0$. From equation (13), it can be seen that $1 + \mu(x)$ is the relative weight placed on the utility of the young and $1 + \nu(\omega)$ is the relative weight placed on the utility of the old. The updating property in equation (16) 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 policy function for the future promise of expected utility is key to understand the evolution of the intergenerational insurance rule. It has the following properties:[15]

Lemma 2. *(i) The policy function* $f\colon \Omega \times \mathcal{S} \to [\omega_0, \omega_{\max}]$ *is continuous and increasing in* ω *and strictly increasing for* $f(\omega, s) \in (\omega_0, \omega_{\max})$. *(ii) There is at least one state* r *such that there is a critical value* $\omega^c(r) > \omega_0$ *with* $f(\omega, r) = \omega_0$ *for* $\omega \in [\omega_0, \omega^c(r)]$. *(iii) For each state* s*, there is a unique fixed point* $\omega^f(s)$ *of the mapping* $f(\omega, s)$ *with* $f(\omega, s) > \omega$ *for* $\omega < \omega^f(s)$ *and* $f(\omega, s) < \omega$ *for* $\omega > \omega^f(s)$. *For at least one state,* $\omega^f(s) > \omega_0$. *(iv) If the aggregate endowment is fixed, then* $f(\omega, s)$ *is decreasing in* s *and strictly decreasing for* $f(\omega, s) \in (\omega_0, \omega_{\max})$.

Figure 2 depicts an example, with three endowment states, of the policy function $f(\omega, s)$. The key properties of $f(\omega, s)$ are that, for a given s, it is continuous and increasing in ω, cuts the 45° line at most once from above (see state 1 in Figure 2) and there is a state such that $f(\omega, s)$ is constant in ω for some $\omega > \omega_0$ (see states 2 and 3 in Figure 2).

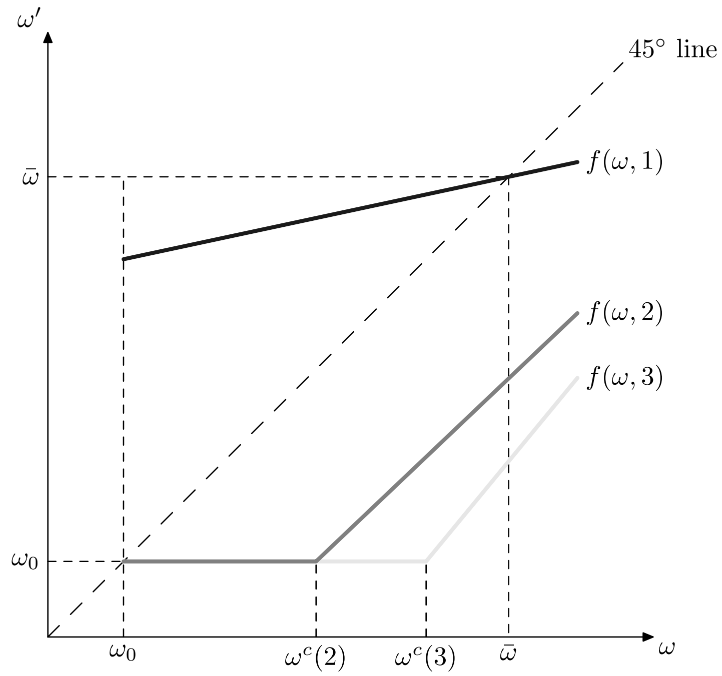


Figure 2: The Policy Function $f(\omega, s)$.

Note: The policy function $f(\omega, s)$ represents the future promise $\omega'(s)$ for $s \in \{1, 2, 3\}$ as a function of ω. The fixed points are $\omega^f(2) = \omega^f(3) = \omega_0$ and $\omega^f(1) = \bar{\omega}$, the largest fixed point.

Continuity of $f(\omega, s)$ in ω follows from the strict concavity of the programming problem. It is intuitive that $f(\omega, s)$ is increasing in ω: a higher promise to the current old means lower consumption for the current young and, for states in which the participa-

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

tion constraint is binding, lower consumption for the current young requires a higher future promised utility, namely higher $f(\omega, s)$, as compensation. This movement along the value function $V(\omega)$ lowers the payoff to the planner who will want to reduce the promise whenever the participation constraint permits. If the participation constraint of the young does not bind in state s, then it follows from the first-order conditions that:

$$f(\omega, s) = \omega_0.$$

That is, the promised utility is *reset* to its initial value whenever the participation constraint of the young does not bind. Once the promised utility is reset, it is as if the history of shocks is forgotten. Resetting occurs because not all participation constraints bind at ω_0, which follows from Assumption 3 that there are states in which there is no transfer at ω_0. For these states, the participation constraint of the young is strictly satisfied since $\omega_0 > \omega_{\min}$. This establishes property (*ii*) of Lemma 2. Part (*iii*) of Lemma 2 shows that the policy function $f(\omega, s)$ cuts the 45° line once from above and part (*iv*) shows that the promised utility is monotonic in s if there is no aggregate risk. We discuss these two properties after the following lemma that characterizes the policy function for the consumption of the young.

Lemma 3. *(i) The policy function $c^y \colon \Omega \times \mathcal{S} \to \mathbb{R}$ is continuous and decreasing in ω and strictly decreasing for $c^y(\omega, s) < e^y(s)$ and $f(\omega, s) < \omega_{\max}$. (ii) At the fixed point $\omega^f(s)$, $c^y(\omega^f(s), s) \leq \beta/\delta$ and with equality for $\omega^f(s) < \omega_{\max}$. (iii) If the aggregate endowment is fixed, then $c^y(\omega, s)$ is decreasing in s.*

As mentioned above, the consumption of the young is decreasing in ω. To understand part (*ii*) of Lemma 3 (and part (*iii*) of Lemma 2), suppose for simplicity that $\eta(x) = \zeta(x) = 0$. From equations (14) and (15), a fixed point of $f(x)$ corresponds to a stationary point of the updating condition (16), that is, $\mu(x) = \nu(\omega)$. Substituting this condition into (13) shows that the ratio of the marginal utilities equals β/δ and hence, consumption is at the first-best level: $c^y(\omega^f(s), s) = c^{y*}(s)$. Likewise, for $f(x) > \omega$, where the next-period promise is higher than today's promise, the consumption of the young is higher than the first-best consumption; and for $f(x) < \omega$, where the next-period promise is lower than today's promise, the consumption of the young is lower than the first-best consumption. To understand why $f(x)$ cuts the 45° line from above, consider some $\omega > \omega^f(s)$ and suppose, to the contrary, that $f(x) \geq \omega$. This would imply that the consumption of the young is no lower, and that the promised utility is higher, at ω than at $\omega^f(s)$. Since the participation constraint of the young is binding at ω, this would violate the participation constraint at $\omega^f(s)$. A similar argument can be made to show that $f(x) > \omega$ for $\omega < \omega^f(s)$.[16]

The policy functions $f(\omega, s)$ and $c^y(\omega, s)$ need not be monotonic in s. However, as shown in part *(iv)* of Lemma 2 and part *(iii)* of Lemma 3, both functions are decreasing in s when the aggregate endowment is fixed. This is intuitive because, absent aggregate risk, the ordering convention for $\lambda(s)$ implies that the autarkic utility is decreasing in s. Hence, if the participation constraint of the young is binding in two different endowment states, then either the consumption of the young or the future promise has to be lower in the higher state. The lemmas show that it is never optimal to have either increasing in the state.[17] Monotonicity of $f(\omega, s)$ in s can also be used to show that, absent aggregate risk, the fixed points $\omega^f(s)$ are ordered by state: $\omega^f(s) \geq \omega^f(r)$ for $s < r$, with strict inequality unless $\omega^f(s) = \omega^f(r) = \omega_0$ or $\omega^f(s) = \omega^f(r) = \omega_{\max}$.

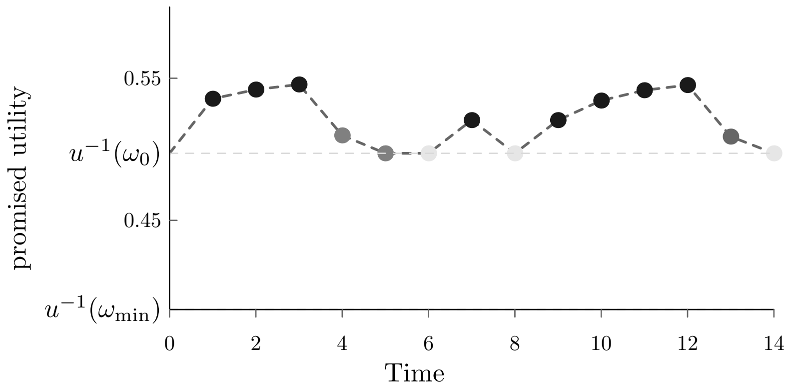


Figure 3: Sample Path of the Promised Utility.

Note: The shade of the dots indicates the state s_t: dark gray for $s_t = 1$, mid gray for $s_t = 2$ and light gray for $s_t = 3$. Promised utility is converted to certainty-equivalent consumption.

Figure 3 illustrates the evolution of promised utility for a given history of shocks, corresponding to the three-state example illustrated in Figure 2. The given history creates a particular sample path of the promise ω. Starting from the initial ω_0, the sample path for the history $s^T = (s_0, s_1, \ldots, s_T)$ is constructed iteratively from the policy function $f(\omega, s)$: $\omega_1 = f(\omega_0, s_0), \omega_2 = f(\omega_1, s_1), \ldots, \omega_{T+1} = f(\omega_T, s_T)$, where $s_t \in \{1, 2, 3\}$ for each t. Figure 3 depicts two important features. First, the path is history dependent. That is, promised utility, and therefore consumption, varies not only with the current

[16] The argument can be extended to the case where the non-negativity and upper bound constraints bind and the complete proof of Lemma 2 is given in the Appendix.

[17] The result regarding monotonicity of the policy function $f(\omega, s)$ in s can be extended to the case in which $e^y(1) \geq e^y(2) \geq \cdots \geq e^y(S)$ and $e(S) \geq e(S-1) \geq \cdots \geq e(1)$, that is, the endowment of the young is weakly decreasing in s but aggregate endowment is weakly increasing in s. By continuity, monotonicity of $f(\omega, s)$ in s is also preserved provided that aggregate risk is not too large. The convergence result discussed in the next section depends only on the monotonicity of the policy function $f(\omega, s)$ in ω and does not depend on whether it is monotonic in s.

endowment state but also with the history of endowment states. For example, state 1 occurs at both $t = 7$ and $t = 12$ but the promised utility differs between them. In particular, whenever state 1 occurs, the participation constraint of the young binds and a higher promised utility has to be offered to them to be willing to share more of their current relatively high endowment. Subsequent realizations of state 1 exacerbate the situation because the next generation of young must deliver on past promises as well. This can be seen in Figure 3 where ω_t increases when state 1 is repeated. Secondly, there are points in time at which promised utility is reset to ω_0. In the example, this happens whenever state 3 occurs and sometimes when state 2 occurs. Before resetting occurs, the effect of a shock persists. However, once resetting has occurred, the history of shocks is forgotten and the subsequent sample path is identical whenever the same sequence of states occurs. That is, the sample paths between resettings are probabilistically identical. This property is used in the next section to establish convergence to a unique invariant distribution.

The results of this section provide the basis for an interpretation of the optimal sustainable intergenerational insurance as a social security or state pension scheme. Following Farhi and Werning (2007), for example, the promise of expected utility can be interpreted as a welfare entitlement for the old. The optimum can be described as a two-tier insurance scheme. In the first tier, the young receive a minimum welfare entitlement for their old age, given by ω_0. Although ω_0 is determined as part of the overall solution, it is independent of past entitlements. This tier is more likely to apply when the young have a poor endowment shock and the current welfare entitlement is not too large. It applies to generations born at dates 4, 5, 7, and 13 in the example illustrated in Figure 3. In the second tier, the young receive a welfare entitlement that depends on the past entitlement relative to the state-contingent threshold given by the fixed point $\omega^f(s)$. These state-contingent thresholds $\omega^f(s) > \omega_0$ are computed directly from the primitives of the model. In this second tier, the next-period welfare entitlement increases when the current entitlement is below the corresponding threshold and decreases when the current entitlement is above the threshold. To satisfy the participation constraint of the young, the movement toward the threshold is partial, in a manner reminiscent of the deterministic case examined in Section 2. In Figure 3, for example, welfare entitlements increase when state 1 occurs since $\omega < \omega^f(1)$, while they decrease when state 2 or 3 occurs because $\omega > \omega^f(2) = \omega^f(3) = \omega_0$. The features of this two-tier structure are in line with the recent ILO report on social protection which recommends both a minimum welfare entitlement and regular adjustments "to secure an adequate level of income for all people of old age without overstretching the capacity of younger generations" (International Labour Organization, 2017, page 77).

4 Convergence to the Invariant Distribution

We now consider the long-run distribution of promised utilities and show that there is strong convergence to a unique invariant distribution. Define $\bar{\omega} := \max_s\{\omega^f(s)\}$ to be the largest of the unique fixed points of the S mappings $f(\omega, s)$. Either $\bar{\omega} < \omega_{\max}$ or $\bar{\omega} = \omega_{\max}$. If $\bar{\omega} < \omega_{\max}$, then any $\omega \in (\bar{\omega}, \omega_{\max}]$ is transitory and cannot be part of the invariant distribution. Since we previously argued that attention can be restricted to $\omega \geq \omega_0$, we will show that the invariant distribution is non-degenerate with support in the interval $[\omega_0, \bar{\omega}] \subset \Omega$. The convergence result relies on the resetting property and the monotonicity of the policy functions $f(\omega, s)$ in ω. Non-degeneracy of the distribution follows from Assumption 4 that the first best is not sustainable.

Denote the transition function for any ω and any set $A \subseteq [\omega_0, \bar{\omega}]$ by

$$P(\omega, A) = \Pr\{\omega_{t+1} \in A \mid \omega_t = \omega\} = \sum\nolimits_s \pi(s)\mathbb{1}_A f(\omega, s),$$

where $\mathbb{1}_A f(\omega, s) = 1$ if $f(\omega, s) \in A$ and zero otherwise. The corresponding n-step transition function is defined recursively by:

$$P^n(\omega, A) = \int P(\varpi, A)P^{n-1}(\omega, \mathrm{d}\varpi),$$

where $P^1(\omega, A) = P(\omega, A)$. An invariant distribution ϕ satisfies

$$\phi(A) = \int P(\omega, A)\phi(\mathrm{d}\omega) \quad \text{for all } A.$$

By part (ii) of Lemma 2, there is always some state such that $f(\omega, s) = \omega_0$ for $\omega \in (\omega_0, \omega^c(s))$. Therefore, there is some $N \geq 1$ and some $\epsilon > 0$ such that $P^N(\omega, \{\omega_0\}) > \epsilon$ for $\omega \in [\omega_0, \bar{\omega}]$. This is obvious when the policy function is ordered by state. Thus, for any initial value of ω, simply pick the highest state and consider the positive probability path in which this state is repeated. Eventually, ω_0 is reached. If the policy functions are not ordered by state, then a similar procedure is to take the sequence of states such that $f(\omega, s)$ is minimized at each stage along the path. This may be a sequence of different states but eventually there must be convergence with a positive probability of reaching ω_0. Standard results now apply. Since $P^N(\omega, \{\omega_0\}) > \epsilon$ for $N \geq 1$ and $\epsilon > 0$, Condition **M** of Nancy Stokey, Robert Lucas Jr. with Edward Prescott (1989, page 348) is satisfied. Hence, there is strong convergence in the uniform metric to a unique probability measure

$\phi(A)$ on $[\omega_0, \bar{\omega}]$.[18] Moreover, $\phi(A)$ can be computed iteratively using the standard adjoint operator.

Proposition 4. *For any given initial distribution $\phi_0(A)$ where $A \subseteq [\omega_0, \bar{\omega}]$, the sequence*

$$\phi_{t+1}(A) = \int P(\omega, A)\phi_t(\mathrm{d}\omega)$$

converges strongly to a unique invariant non-degenerate distribution $\phi(A)$, with a support $\operatorname{supp}\phi \subseteq [\omega_0, \bar{\omega}]$ *and a mass point at ω_0: $\phi(\{\omega_0\}) > 0$.*

The main idea of this result is that there is always a sequence of shocks such that eventually the participation constraint of the young does not bind. When this occurs, the promise to them is *reset* to ω_0. The process is regenerative (see, for example, Foss et al., 2018) and ω_0 is a regeneration point. Thus, whenever the process reaches that point, the probabilistic future evolution of promises is identical regardless of the point in time or the history of shocks that led up to the regeneration point. This can be seen in Figure 3 where there are regeneration points at dates 5, 6, 8 and 14. The cycles between regeneration points are not identical but they are i.i.d. and therefore, the distribution of future promises eventually converges to the invariant distribution, which is independent of the initial distribution. As argued above, the planner's payoff is maximized by choosing the initial promise ω_0 and therefore, the initial distribution satisfies $\phi_0(\{\omega_0\}) = 1$. Note, however, that Proposition 4 is more general and establishes convergence to the same invariant distribution for any initial distribution $\phi_0(A)$. Note also that, although there is eventually a stationary invariant distribution for the promises ω, at each point in time the promises are determined by the policy function $f(\omega, s)$ for a given ω. That is, the evolution of future promises depends on the history of endowment states since the last time the promise ω was reset to ω_0.

The non-degeneracy of the invariant distribution in Proposition 4 is to be contrasted with the two benchmarks considered in Section 2. If transfers are enforced, or if there is no risk, then there is convergence to a degenerate invariant distribution with unit mass at ω^*. If $\bar{\omega} < \omega_{\max}$, then it follows from part (*ii*) of Lemma 3 that $\bar{\omega} > \omega^*$. In this case, the invariant distribution has no mass point at $\bar{\omega}$ and the promise fluctuates above and below the first-best promise ω^* even in the long run, unlike the benchmark cases.

[18] Ömer Açikgöz (2018), Sergey Foss, Vsevolod Shneer, Jonathan Thomas and Tim Worrall (2018) and Shenghao Zhu (2020) use similar arguments to establish strong convergence in the case of a Bewley-Imrohoroglu-Huggett-Aiyagari precautionary savings model with heterogeneous agents.

It is not possible to give a more detailed characterization of the invariant distribution because it depends on the exact properties of the policy functions $f(\omega, s)$. Nonetheless, in Section 6 we consider an example with two endowment states and show that the invariant distribution ϕ is a transformation of a geometric distribution with a countable support and no mass at $\bar{\omega}$.

A straightforward corollary of Proposition 4 states that there is a stationary invariant distribution for $x = (\omega, s)$. That is, for any $\mathcal{X} = A \times B$ where $A \subseteq [\omega_0, \bar{\omega}]$ and $B \subseteq \mathcal{S}$, we can define a transition function $P(x, \mathcal{X})$ and show strong convergence to an invariant distribution, say $\varphi(\mathcal{X})$, where $\varphi(\mathcal{X}) = \phi(A)\pi(B)$ and $\pi(B) = \sum_{s \in B} \pi(s)$.[19] In the next section, we use the long-run convergence to the invariant distribution and the Markov property of the transition function to measure generational risk.

5 Measuring Generational Risk

We now assess how risk is shared between generations and across cohorts in the optimal sustainable intergenerational insurance described in Sections 3 and 4. To do this, we use conditional and mean entropy (see, for example, Backus, Chernov and Zin, 2014) and the bound on the variability of the implied yields introduced by Martin and Ross (2019) as measures of generational risk. An advantage of this approach is that it is closely tied with the implied state prices and yields and connects with the literature on the dominant root characterization of Pareto optimality.[20]

To proceed, write the stochastic discount factor (see, for example, Gur Huberman, 1984; Gregory W. Huffman, 1986; Pamela Labadie, 1986) as follows:

$$m(x, x') = \beta \frac{u_c(e(s') - c^y(x'))}{u_c(c^y(x))} = \delta \left(\frac{u_c(c^y(x'))}{u_c(c^y(x))} \right) \left(\frac{\beta}{\delta} \frac{u_c(e(s') - c^y(x'))}{u_c(c^y(x'))} \right), \tag{17}$$

where $x = (\omega, s)$ is the current state and $x' = (f(\omega, s), s')$ is the successor state next period. Equation (17) shows that the stochastic discount factor can be decomposed into three terms: the planner's discount factor and the two bracketed terms representing the

[19] The transition function is $P(x, \mathcal{X}) = \Pr\{(\omega', s') \in \mathcal{X} \mid x = (\omega, s)\} = \mathbb{1}_A f(\omega, s)\pi(B)$.

[20] Two alternative measures used to assess the divergence from first-best risk sharing are the insurance coefficient (see, for example, Greg Kaplan and Giovanni Violante, 2010) and a consumption equivalent welfare change (see, for example, Zheng Song, Kjetil Storesletten, Yikai Wang and Fabrizio Zilibotti, 2015). We discuss these alternatives in Part C of the Supplementary Appendix and compute their values at the invariant distribution for the numerical example considered in Section 7. It is shown that using the insurance coefficient or consumption equivalent welfare change lead to substantively similar conclusions to those presented below.

risk sharing across two adjacent generations and the risk sharing between generations at a given date.[21]

Using the stochastic discount factor, denote state prices by $q(x,x') := \pi(x,x')m(x,x')$ (where $\pi(x,x') = \pi(s')$ since shocks are i.i.d.) and risk-neutral probabilities by $\varrho(x,x') := q(x,x')/\sum_{x'} q(x,x')$. Let Q, Π and Γ denote the matrix of state prices, transition probabilities and risk-neutral probabilities. Conditional entropy, $L(x)$, is the Kullback-Liebler divergence between the rows of the matrices Π and Γ corresponding to state x. It is non-negative and provides an upper bound on the expected log excess returns of any asset conditional on state x. Mean entropy, $\bar{L}$, is the average of conditional entropy taken over the invariant distribution:

$$\bar{L} = \sum\nolimits_x \varphi(x)L(x) \quad \text{where} \quad L(x) = -\sum\nolimits_{x'} \pi(x,x') \log\left(\frac{\varrho(x,x')}{\pi(x,x')}\right).$$

Conditional and mean entropy over k periods, $L^k(x)$ and $\bar{L}^k$, are found by replacing $\pi(x,x')$ and $\varrho(x,x')$ in the formula above by the appropriate elements of matrix powers Π^k and Γ^k. Conditional entropy, $L^k(x)$, provides an upper bound on the expected log excess returns of any k-period asset, conditional on state x, while mean entropy per period, $\bar{L}^k/k$, provides an overall measure of risk and how it is spread over time.

To see the connection between entropy and implied yields, let $p^k(x)$ denote the price of a k-period discount bond, which is defined recursively by:

$$p^k(x) = \sum\nolimits_{x'} q(x,x')p^{k-1}(x'),$$

where $p^0(x) \equiv 1$. The continuously compounded yield on a k-period discount bond is $y^k(x) := -(1/k)\log(p^k(x))$, where $y^\infty := \lim_{k\to\infty} y^k$ and $\bar{y}^k := \sum_x \varphi(x)y^k(x)$ are the long-run and average yields. The mean entropy per period satisfies:

$$\frac{\bar{L}^k}{k} = y^\infty - \bar{y}^k, \tag{18}$$

where $y^\infty = -\log(\rho)$ and ρ is the Perron root of the matrix Q.[22] Since mean entropy is non-negative, the average yield cannot be greater than the long-run yield. Martin and Ross (2019) show that $|y^k(x) - y^\infty| \leq (1/k)\Upsilon$, where

$$\Upsilon := \log\left(\frac{\psi_{\max}}{\psi_{\min}}\right)$$

[21]Since shocks are transitory, there is no permanent component to the stochastic discount factor.
[22]See Part D of the Supplementary Appendix for details.

and $\psi_{\max}$ and $\psi_{\min}$ are the maximum and minimum values of ψ, the eigenvector of Q that corresponds to the Perron root.[23] Therefore, the deviation of the yield from its mean satisfies:

$$\Delta y^k(x) := y^k(x) - \bar{y}^k \in \left[\frac{\bar{L}^k - \Upsilon}{k}, \frac{\bar{L}^k + \Upsilon}{k}\right].$$

Since $y^k(x)$ cannot be greater than the average yield for all x, it follows that $\bar{L}^k \leq \Upsilon$. While entropy is a measure of the variability of the stochastic discount factor, which corresponds to the variability in the eigenvector ψ, the term Υ is a measure of the range of the variability in ψ and provides an upper bound for mean entropy.

Before considering these risk measures in general, we consider them for the first-best benchmark. In the first best, the second bracketed term of equation (17) is determined by equation (5). For simplicity, suppose that the non-negativity constraint on transfers does not bind. In that case, $m^*(s, s') = \delta u_c(c^{y*}(s'))/u_c(c^{y*}(s))$. This is akin to a representative agent model. Since the corresponding matrix of state prices has rank one, the Perron root is equal to the trace of the matrix, so that $\rho^* = \delta$. The elements of the corresponding eigenvector are $\psi^*(s) = 1/u_c(c^{y*}(s))$. Since $c^{y*}(s)$ is increasing in the aggregate endowment, $\psi^*_{\max}$ corresponds to an endowment state with the highest aggregate endowment and $\psi^*_{\min}$ to an endowment state with the lowest aggregate endowment. We can use the Ross Recovery Theorem (see, Steve Ross, 2015) to find:

$$L^{k*}(s) = \log\left(\sum\nolimits_s \pi(s)\frac{1}{\psi^*(s)}\right) - \sum\nolimits_s \pi(s)\log\left(\frac{1}{\psi^*(s)}\right) \;\; \forall s,\, k; \;\; \text{and} \;\; \Upsilon^* = \log\left(\frac{\psi^*_{\max}}{\psi^*_{\min}}\right).$$

Entropy in the first best is independent of the endowment state s and time horizon k because the endowment shocks are transitory. There is no differential in risk shared across endowment states or time. If there is no aggregate risk, then the first-best consumption is independent of the endowment state and the eigenvector can be normalized to the unit vector. In this case, $L^{k*}(s) = \Upsilon^* = 0$ and $y^{k*}(s) = -\log(\delta)$ for all s and k.

If only partial risk sharing can be sustained, then $L^k(x)$ and $y^k(x)$ depend on the time horizon k and the state x, even in the absence of aggregate risk. In general, this dependence might be quite complex. Therefore, the next two sections explore an example

[23] For simplicity, and because it corresponds to our numerical procedure, we assume here that the invariant distribution of x is discrete and its sample space finite. The arguments in the text can be adapted for the case where the invariant distribution is not finite. For example, if the sample space of φ is not finite, then, following Timothy M. Christensen (2017) and Lars Peter Hansen and José A. Scheinkman (2009), define the pricing operator $\mathbb{M}d(x) = \mathbb{E}[m(x, x')d(x') \mid x]$ for a payoff function $d(x)$. In this case, instead of an eigenvector there is an eigenfunction $\psi(x)$ that is a solution to the Perron-Frobenius eigenfunction problem: $\mathbb{M}\psi(x) = \rho\psi(x)$.

with two endowment states to derive theoretical, numerical and comparative static results. Nevertheless, there are some properties that hold in general.

Proposition 5. *In the optimal sustainable intergenerational insurance: (i) The deviation of the conditional yield from its mean, $\Delta y^k(\omega, s)$, is increasing in ω for each k and s. (ii) For each x, $\lim_{k\to\infty} y^k(x) = -\log(\rho)$, where ρ is the Perron root of Q. (iii) If $\bar{\omega} < \omega_{\max}$, then $\rho \leq \delta$, with equality if the non-negativity constraint on transfers is not binding. (iv) Absent aggregate risk, if $\bar{\omega} < \omega_{\max}$, then $y^1(\bar{\omega}, 1) > -\log(\delta) > y^1(\omega_0, S)$.*

To understand Proposition 5 recall that the consumption of the young is decreasing in ω while the future promise is increasing in ω. Therefore, the stochastic discount factor $m((\omega, s), (f(\omega, s), s'))$ is decreasing in ω.[24] Taking the expectation over s', the price of the one-period bond decreases with ω or equivalently, the one-period yield, $y^1(\omega, s)$, increases with ω. Thus, an agent born into a generation where the promise is higher faces higher one-period yields. Part (i) shows this is true for bonds of any length. Part (ii) is a standard result that all yields converge in the long run. That is, the optimal exposure to the risk from a shock vanishes in the very long run. Part (iii) shows that if the upper bound constraint on promises and the non-negativity constraint on transfers are not binding, then the long-run yield is the same as in the first best and determined by the planner's discount factor. Finally, part (iv) shows that the yield is high when the young have a high endowment and the promise is large.

6 Two Endowment States

Finding the optimal sustainable intergenerational insurance is complex because it requires solving the functional equation (12). In this section, we present an example with two endowment states that can be solved independently of the solution of the value function $V(\omega)$. For this case, a full characterization of the optimal dynamics of consumption and the invariant distribution of promises is provided together with solutions for the generational risk measures outlined in the previous section. Assumptions 1-4 are maintained but we concentrate on a case with CRRA utility: $u(c) = (c^{1-\gamma} - 1)/(1-\gamma)$, where γ is the coefficient of relative risk aversion and $u(c) = \log(c)$ in the limit as $\gamma \to 1$. The two endowment states $s \in \{1, 2\}$ occur with probability π and $1-\pi$. There is no aggregate risk and the aggregate endowment is normalized to unity in both states. The endowments of the young are $e^y(1) = \kappa + \sigma$ and $e^y(2) = \kappa - \sigma\pi/(1-\pi)$, where $\kappa \in (1/2, 1)$ is the

[24] If consumption is ordered by endowment state (for example, if aggregate endowment is constant), then $m((\omega, s), (f(\omega, s), s'))$ is decreasing in s'. However, the dependence on s is not clear cut because a higher endowment state (lower endowment of the young) means a lower consumption of the young but also a lower future promise, leading to a lower consumption of the old in the subsequent period.

average endowment of the young (the average endowment of the old is $1-\kappa$) and $\sigma > 0$ is the variability of endowments. That is, the young have a good endowment shock in state 1 and a poor endowment shock in state 2. An increase in σ is a mean-preserving spread of endowment risk.[25]

By Assumptions 3 and 4, the promised utility satisfies $f(\omega_0, 1) > f(\omega_0, 2) = \omega_0$. From Lemma 2, $f(\omega, 1) > f(\omega, 2)$, and the largest fixed point $\bar{\omega} = \omega^f(1)$ is unique. We make two additional assumptions. First, that $\bar{\omega} < \omega_{\max}$: that is, the upper bound constraint (10) is not binding. Second, that $f(\omega, 2) = \omega_0$ for $\omega \in [\omega_0, \bar{\omega}]$: that is, the participation constraint of the young never binds when the young have the poor endowment shock, and consequently, the multiplier $\mu(\omega, 2) = 0$ for $\omega \in [\omega_0, \bar{\omega}]$. In this case, the promised utility is optimally reset to ω_0 whenever state 2 occurs. When the young have the good endowment shock, the promised utility increases, approaching $\bar{\omega}$ if state 1 is repeated infinitely often. Consequently, the history of states is forgotten whenever state 2 occurs and the future promise depends only on the number of consecutive state 1s in the most recent history. It can be checked that there is a non-empty set of parameter values that satisfies these two additional assumptions. For example, they are valid for the following parameter values.

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

We refer to Example 1 as our baseline or canonical case and all figures in this section are drawn for this case.[26]

Using the two additional assumptions, the next two propositions establish the properties of the optimum consumption and generational risk measures (Proposition 6) and derive the invariant distribution and moment conditions for consumption at the invariant distribution (Proposition 7).

Proposition 6. *Suppose $f(\omega, 2) = \omega_0$ for $\omega \in [\omega_0, \bar{\omega}]$ and $\bar{\omega} < \omega_{\max}$, then: (i) $c^y(\omega, 1)$ and $c^y(\omega, 2)$ are decreasing in ω, $c^y(\omega, 1) > c^y(\omega, 2)$ and $c^y(\bar{\omega}, 1) = c^{y*}(1) = c^{y*}(2) = c^y(\omega_0, 2)$. (ii) The Perron root of Q is $\rho = \delta$, the corresponding eigenvector ψ satisfies*

$$\psi(\omega, 1) = 1/(u_c(c^y(\omega, 1))v(\omega)) \quad \textit{and} \quad \psi(\omega, 2) = 1/u_c(c^y(\omega, 2))$$

[25] Assumptions 2-4 impose some restrictions on the parameter set. For example, with $\gamma = 1$ and $\pi = 1/2$, Assumption 2 is satisfied for $\beta > ((1-\kappa)^2 - \sigma^2)/(\kappa(1-\kappa) + \sigma^2)$; Assumption 3 is satisfied for $\sigma \geq \kappa - 1/2$; and Assumption 4 is satisfied for $\beta \leq -\log(2(\kappa+\sigma))/\log(2\sqrt{((1-\kappa)^2 - \sigma^2)})$.

[26] The value of δ in this example corresponds to a long-run interest rate of 1⅓%. Numerical calculations for this section are based on solving a simple shooting algorithm. See Part E of the Supplementary Appendix for details.

and the upper bound on mean entropy is $\Upsilon = \log(\upsilon(\bar{\omega}))$ *where* $\upsilon(\omega) := 1 + \mu(\omega, 1)$ *and* $\log(\upsilon(\bar{\omega}))$ *is determined from the participation constraint of the young in state 1 and the promise-keeping constraint. (iii) Conditional entropy* $L(\omega, 1)$ *and* $L(\omega, 2)$ *are given by:*

$$\begin{aligned} L(\omega, 1) &= \log\left(\tfrac{\pi}{\psi(\omega',1)} + \tfrac{1-\pi}{\psi(\omega',2)}\right) - \left(\pi \log\left(\tfrac{1}{\psi(\omega',1)}\right) + (1-\pi)\log\left(\tfrac{1}{\psi(\omega',2)}\right)\right), \\ L(\omega, 2) &= \log\left(\tfrac{\pi}{\psi(\omega_0,1)} + \tfrac{1-\pi}{\psi(\omega_0,2)}\right) - \left(\pi \log\left(\tfrac{1}{\psi(\omega_0,1)}\right) + (1-\pi)\log\left(\tfrac{1}{\psi(\omega_0,2)}\right)\right), \end{aligned} \tag{19}$$

where $\omega' = f(\omega, 1)$ *and* $L(\omega, 2)$ *is independent of* ω.

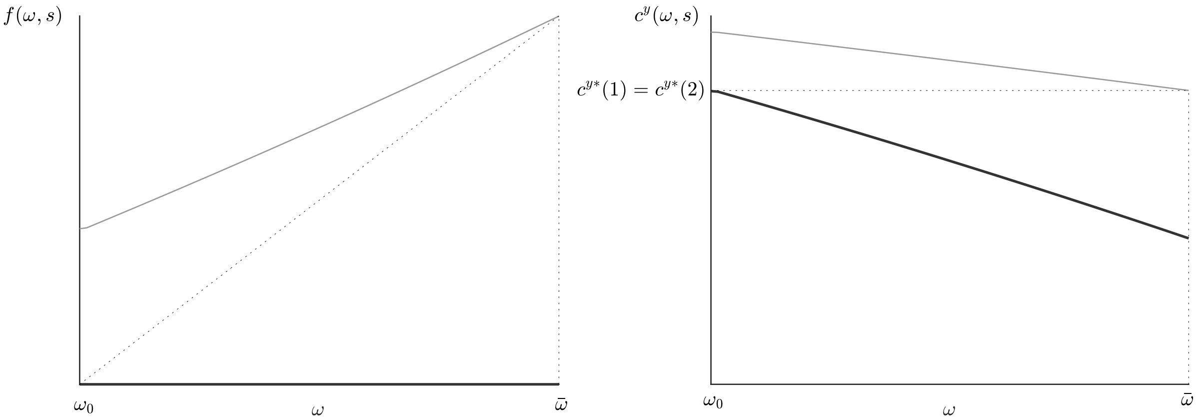


Figure 4: Panel A – Promised Utility. Panel B – Young Consumption.

Note: In Panel A, the light gray line is $f(\omega, 1)$ and the dark gray line is $f(\omega, 2)$. The dashed line is the 45° line. In Panel B, the light gray line is $c^y(\omega, 1)$ and the dark gray line is $c^y(\omega, 2)$.

The future promised utility and the consumption of the young are plotted in Figure 4 as a function of the current promise ω for each of the two endowment states. The policy function $f(\omega, 1)$ is monotonically increasing in ω with a fixed point at $\bar{\omega}$, while the policy function $f(\omega, 2) = \omega_0$ for $\omega \leq \bar{\omega}$. As stated in part (i) of Proposition 6, $c^y(\omega, s)$ is monotonically decreasing in ω and ordered by the state: $c^y(\omega, 1) > c^y(\omega, 2)$. Since $\bar{\omega}$ and ω_0 are fixed points of $f(\omega, 1)$ and $f(\omega, 2)$, it follows from Lemma 3 that $c^y(\bar{\omega}, 1) = c^{y*}(1)$ and $c^y(\omega_0, 2) = c^{y*}(2)$. With no aggregate risk, the first-best consumption is independent of the state, that is, $c^{y*}(1) = c^{y*}(2)$, and hence, $c^y(\bar{\omega}, 1) = c^y(\omega_0, 2)$.

The components of the eigenvector, $\psi(\omega, 1)$ and $\psi(\omega, 2)$, are plotted in Panel A of Figure 5. From part (i) of Proposition 6, it is easily checked that both components are decreasing in ω and $\psi(\omega, 2) > \psi(\omega, 1)$. Moreover, $\psi_{\max} = \psi(\omega_0, 2) = 1/u_c(c^{y*}(2))$ and $\psi_{\min} = \psi(\bar{\omega}, 1) = 1/(u_c(c^{y*}(1))\upsilon(\bar{\omega}))$, where $\upsilon(\bar{\omega}) = 1 + \mu(\bar{\omega}, 1)$. Since $c^{y*}(1) = c^{y*}(2)$ when there is no aggregate risk, the upper bound on mean entropy is $\Upsilon = \log(\psi_{\max}/\psi_{\min}) = \log(\upsilon(\bar{\omega}))$. That is, the bound Υ depends only on the the tightness of the participation constraint of the young in state 1 at $\bar{\omega}$ and can be derived entirely from the primitives of the model, without the need for numerical approximation, by using the promise-keeping constraint and the participation constraint of the young, both of which bind in state $(\bar{\omega}, 1)$.[27]

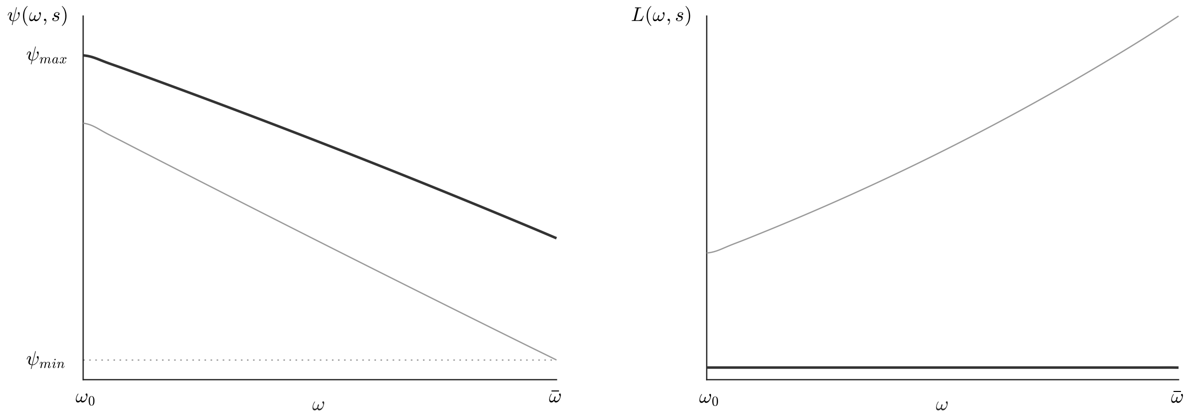


Figure 5: Panel A – Eigenvector. Panel B – Conditional Entropy.

Note: In Panel A, the light gray line is $\psi(\omega, 1)$ and the dark gray line is $\psi(\omega, 2)$. In Panel B, the light gray line is $L(\omega, 1)$ and the dark gray line is $L(\omega, 2)$.

While $\Upsilon = \log(\upsilon(\bar{\omega}))$, more generally the term $\log(\upsilon(\omega))$ provides a measure of the deviation of the ratio of marginal utilities from first-best risk sharing. Let $g(\omega, s)$ denote the ratio of marginal utilities of the young and the old in state (ω, s). It is increasing in ω and equals β/δ in the first best (when the non-negativity constraints on transfers do not bind). It is easy to verify that $g(\omega, 2) \geq \beta/\delta \geq g(\omega, 1)$ with $g(\bar{\omega}, 1) = \beta/\delta = g(\omega_0, 2)$ and $g(\omega_0, 1) < \beta/\delta < g(\bar{\omega}, 2)$. From condition (13), we have:

$$\log\left(\upsilon(\omega)\right) = \left(\log\left(g(\omega,2)\right) - \log\left(\frac{\beta}{\delta}\right)\right) + \left(\log\left(\frac{\beta}{\delta}\right) - \log\left(g(\omega,1)\right)\right).$$

As ω increases, the deviation of $g(\omega, s)$ from the first best decreases in state 1 and increases in state 2. Since $\upsilon(\omega)$ increases in ω, the aggregate deviation from the first best increases with ω. In this sense, an agent born in a period with a higher ω bears more risk. The measure Υ provides the upper bound on this risk.

Part (*iii*) of Proposition 6 gives the formula for the one-period conditional entropy. It provides a measure of the risk that the young face in their old age conditional on the state in which they are born. Both $L(\omega, 1)$ and $L(\omega, 2)$ are plotted in Panel B of Figure 5. Entropy is independent of ω in state 2 because the future promise is always reset to ω_0 irrespective of the current promise. Entropy is increasing in ω in state 1 in this example because the future promise increases with ω as does the variability of old-age consumption.

[27]The exact formula for Υ in terms of the parameters is given in the proof of Proposition 6 in the Appendix and the its comparative static properties are described in Section 7.

We now turn to the properties of the optimum at the invariant distribution. To simplify notation, let $\omega_0^{(n)}$ denote the promised utility after n consecutive state 1s, starting from an initial promise ω_0.[28]

Proposition 7. *Suppose $f(\omega, 2) = \omega_0$ for $\omega \in [\omega_0, \bar{\omega}]$ and $\bar{\omega} < \omega_{\max}$, then at the invariant distribution: (i) $\phi(\{\omega_0^{(n)}\}) = (1-\pi)\pi^n$ for $n = 0, 1, \ldots, \infty$. (ii) Mean entropy is given by:*

$$\bar{L} = \sum\nolimits_{n=0}^{\infty} (1-\pi)\pi^{n+1} L(\omega_0^{(n)}, 1) + (1-\pi)L(2). \tag{20}$$

(iii) The logarithm of the ratio of marginal utilities is heteroscedastic with the endowment: $\operatorname{var}(\log(g(\omega, 2))) > \operatorname{var}(\log(g(\omega, 1)))$. *(iv) The auto-covariance of the promised utility over two adjacent periods is positive:* $\operatorname{cov}(\omega_t, \omega_{t+1}) > 0$. *(v) The auto-covariance of the consumption of the young conditional on the endowment state is non-negative with* $\operatorname{cov}(c_t^y, c_{t+1}^y \mid s_t = 1) > 0$ *and* $\operatorname{cov}(c_t^y, c_{t+1}^y \mid s_t = 2) = 0$.

Part (i) of Proposition 7 follows directly from the resetting property. Since the promised utility is reset to ω_0 whenever state 2 occurs and $\bar{\omega} < \omega_{\max}$, the invariant distribution of ω depends only on the number of consecutive state 1s. Since the distribution of n consecutive state 1s is geometric, the invariant distribution of ω is countable and determined from the geometric distribution by a change in variables. The invariant distribution has a probability mass of $1 - \pi$ at ω_0 and no probability mass at $\bar{\omega}$. The expected time to resetting is $1/(1 - \pi)$. The invariant distribution of the state x is also easily calculated from ϕ because $\varphi(\omega, s) = \pi(s)\phi(\omega)$. Using the invariant distribution $\varphi(\omega, s)$, mean entropy is given by equation (20) in part (ii) of the proposition.

Part (iii) of Proposition 7 shows that the consumption allocation is heteroscedastic. For Example 1, the variance of the consumption of the young is lower when their endowment is higher. The results in parts (iv) and (v) of Proposition 7 are illustrated in Figure 6, which plots the joint distribution of the promised utility (Panel A) and the joint distribution of the consumption of the young (Panel B) for two adjacent generations. It follows that $\operatorname{cov}(\omega_t, \omega_{t+1}) > 0$ because $f(\omega, s)$ is increasing in the current promise.[29] The properties of the auto-covariance of promised utility are reflected in the auto-covariance of individual consumption. Conditional on state 2, the auto-covariance between the consumption of the young in two adjacent generations is zero because the promise, and hence

[28] Formally, $\omega_0^{(n)} = f_1^{(n)}(\omega_0)$, $f_1(\omega) := f(\omega, 1)$ and $f_1^{(n)}$ is the n-fold composition of f_1 with $f_1^{(0)}(\omega) = \omega$. Since $\bar{\omega}$ is the fixed point of the policy function $f(\omega, 1)$, $\lim_{n\to\infty} \omega_0^{(n)} = \bar{\omega}$.

[29] At the invariant distribution, $\operatorname{cov}(\omega_t, \omega_{t+k})$ depends only on the time horizon k and is positive. As $k \to \infty$, the conditional expectation of ω_{t+k} converges to the mean of the invariant distribution, which is constant and independent of k. Hence, $\lim_{k\to\infty} \operatorname{cov}(\omega_t, \omega_{t+k}) = 0$.

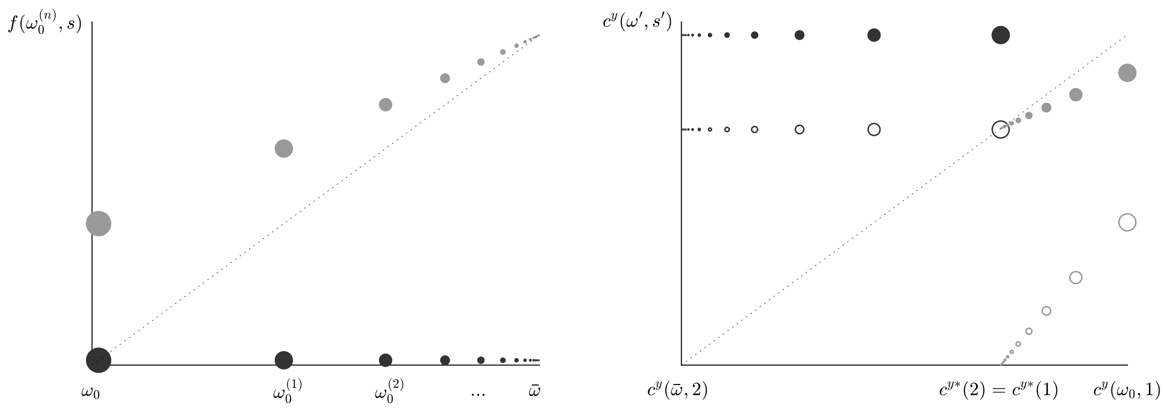


Figure 6: Panel A – Joint Distribution of ω. Panel B – Joint Distribution of $c^y(\omega, s)$.
Note: In both panels, the size of the bubble represents the frequency of occurrence. Light gray dots correspond to state 1 and dark gray dots to state 2 in the current period. The dashed line is the 45° line. In Panel B, a solid dot indicates that the endowment state is the same in both periods and an open dot means transition to the other endowment state.

next-period consumption, is always reset whenever state 2 occurs. On the other hand, conditional on state 1, the expected consumption of the young next period is increasing in the consumption of the current young.[30] The unconditional auto-covariance of consumption, $\text{cov}(c_t^y, c_{t+1}^y)$, is typically negative, as we illustrate in the next section. This is because consumption is high in the period after resetting but consumption is generally low in state 2.[31] The comparison to a model with infinitely-lived agents is discussed in Section 9, but it is worth noting that in an equivalent two-state case with two infinitely-lived agents, the conditional and unconditional auto-covariance of consumption across adjacent periods is zero.

7 Comparative Statics

In this section, we continue the two-state example of Section 6 and examine how the generational risk measures and the autocorrelation of consumption of the young across two adjacent generations respond to comparative static changes of endowment parameters and discount factors and how they depend on the time horizon.[32] For all comparative

[30] Absent aggregate risk, the same argument applies to the consumption of the old: $\text{cov}(c_t^o, c_{t+1}^o \mid 1) > 0$.

[31] By the law of total covariance, $\text{cov}(c_t^y, c_{t+1}^y) = \mathbb{E}_s[\text{cov}(c_t^y, c_{t+1}^y \mid s_t)] + \text{cov}(\mathbb{E}_\varphi[c_y^y \mid s_t], \mathbb{E}_\varphi[c_{t+1}^y \mid s_t])$. The first term is positive by part (v) of Proposition 7, but the second term is negative because $\mathbb{E}_\varphi[c_t^y \mid s_t = 1] > \mathbb{E}_\varphi[c_t^y \mid s_t = 2]$ and $\mathbb{E}_\varphi[c_{t+1}^y \mid s_t = 1] < \mathbb{E}_\varphi[c_{t+1}^y \mid s_t = 2]$.

[32] For the purposes of comparison, we use autocorrelation instead of auto-covariance. The conditional autocorrelation is given by $\text{corr}(c_t^y, c_{t+1}^y \mid s_t) := \text{cov}(c_t^y, c_{t+1}^y \mid s_t)/\sqrt{(\text{var}(c_t^y \mid s_t)\,\text{var}(c_{t+1}^y \mid s_t))}$ and the unconditional autocorrelation is given by $\text{corr}(c_t^y, c_{t+1}^y) := \text{cov}(c_t^y, c_{t+1}^y)/\,\text{var}(c_t^y)$.

statics, we change the value of the parameter of interest holding all other parameters at the values in the canonical case of Example 1.[33]

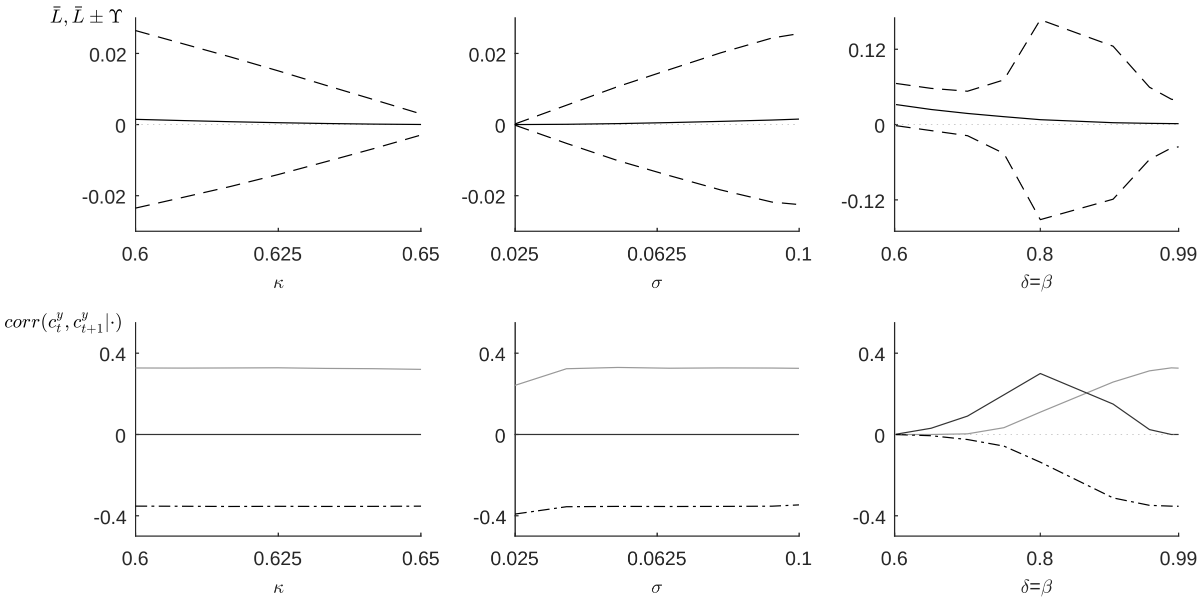


Figure 7: Comparative Statics at the Invariant Distribution.

Note: In the top row, the solid line is $\bar{L}$ and the dashed lines are $\bar{L} \pm \Upsilon$ (where Υ is scaled by 1/10). In the bottom row, the light gray line is the autocorrelation of consumption of the young between adjacent periods conditional on state 1, $\text{corr}(c_t^y, c_{t+1}^y \mid s_t = 1)$, the dark gray line is the corresponding autocorrelation conditional on state 2, $\text{corr}(c_t^y, c_{t+1}^y \mid s_t = 2)$, and the dashed line is the unconditional autocorrelation, $\text{corr}(c_t^y, c_{t+1}^y)$.

Changing the Endowment The effect of changes in κ and σ are illustrated in the first two columns of Figure 7. The top row illustrates the effect on $\bar{L}$ and $\bar{L} \pm \Upsilon$ and the bottom row illustrates the effect on the autocorrelation of consumption. A larger κ corresponds to a larger average endowment share to the young, while a smaller σ corresponds to reduced idiosyncratic uncertainty. Increasing κ, or reducing σ, increases risk sharing as measured by a reduction in $\bar{L}$ and Υ. For κ above some critical value, or σ below a critical value, the first best is sustainable in the long run at the invariant distribution, in which case $\bar{L} = \Upsilon = 0$.[34] For the range of κ and σ illustrated, the premise of Propositions 6 and 7 hold, that is, $f(\omega, 2) = \omega_0$ for $\omega \leq \bar{\omega}$ and $\bar{\omega} < \omega_{\max}$. The implications of this premise can be seen in the bottom row of Figure 7, where the autocorrelation of consumption conditional on state 2 is zero. The corresponding autocorrelation conditional on state 1

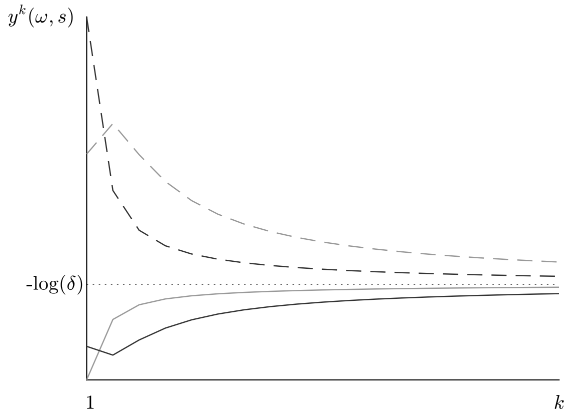
[33] In all cases, the invariant distribution is geometric, except when discount factors are changed. When the invariant distribution is not geometric, we can no longer rely on the shooting algorithm used in Section 6. In this case, we implement an algorithm based on a value function iteration method (see, Part F of the Supplementary Appendix for a description). Although we consider an example with two endowment states here, the value function iteration method can be applied when there are more than two states.

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

is positive, while the unconditional autocorrelation is negative. Neither is very sensitive to changes in κ and σ. Although the auto-covariance of consumption tends to zero as the consumption tends to the first best (for κ large enough, or σ small enough), the variance of consumption decreases at a similar rate, so that there is little change in the autocorrelation coefficient.

Changing the Discount Factors The final column of Figure 7 illustrates the effect of changes in the discount factor (holding $\beta = \delta$). The effect on the generational risk measures and the autocorrelation of consumption is non-monotonic. This occurs because the assumptions $f(\omega, 2) = \omega_0$ for $\omega \leq \bar{\omega}$ and $\omega < \omega_{\max}$ do not hold when the discount factor is sufficiently small. For high values of the discount factor, the invariant distribution has geometric probabilities as described in part (i) of Proposition 7. As the discount factor is decreased, either the current transfer is reduced, or the future promise increased, to satisfy the participation constraint of the young in state 1. This change spreads out the distribution of ω, increasing $\bar{\omega}$ but reducing ω_0 and $\omega_{\max}$. The reduced risk sharing and the increased spread of promised utility are reflected in an increase of $\bar{L}$ and Υ. The variance of consumption is increased, but so too is the absolute value of the auto-covariance of consumption, with an overall reduction in the absolute value of the autocorrelation (unconditional as well as conditional on state 1). As the discount factor is reduced further, both the upper bound constraint becomes binding, that is, $\bar{\omega} = \omega_{\max}$, and $f(\omega, 2) > \omega_0$ for high values of ω. Reversion to ω_0 occurs less frequently and as the discount factor falls, the invariant distribution has a positive probability mass at $\omega_{\max}$. Although ω_0 falls with the discount factor, the range $\bar{\omega} - \omega_0$ decreases, meaning that although $\bar{L}$ increases, the bound Υ decreases. For high values of ω, the future promise is strictly increasing in ω even in state 2 and therefore, the autocorrelation of consumption conditional on state 2 is positive. For a low enough discount factor ($\beta = 3/5$ in the figure), autarky is the only sustainable allocation: the invariant distribution tends to a degenerate distribution with a unit mass on $\omega_{\min}$ and the autocorrelation of consumption approaches zero because the endowments are serially uncorrelated.

Horizon Dependence Panel A of Figure 8 plots the yield curves $y^k(\omega_0, s)$ and $y^k(\bar{\omega}, s)$ for each s. Part (iv) of Proposition 5 demonstrated that $y^k(\bar{\omega}, 1) > -\log(\delta) > y^k(\omega_0, 2)$ for $k = 1$. The Figure shows that the same result holds for each k and furthermore, that $y^k(\bar{\omega}, 2) > -\log(\delta) > y^k(\omega_0, 1)$. Moreover, all yields converge to the long-run yield $y^\infty = -\log(\delta)$ as $k \to \infty$. Panel B of Figure 8 plots the mean entropy per period, $\bar{L}^k/k$, together with the bound Υ. As shown in equation (18), the mean entropy per period equals the deviation of the average yield from the long-run yield. The dashed lines provide bounds on $\Delta y^k(x)$ for each x and gives an indication of the maximum variability of the possible yield curves. The mean entropy per period and the bounds converge to



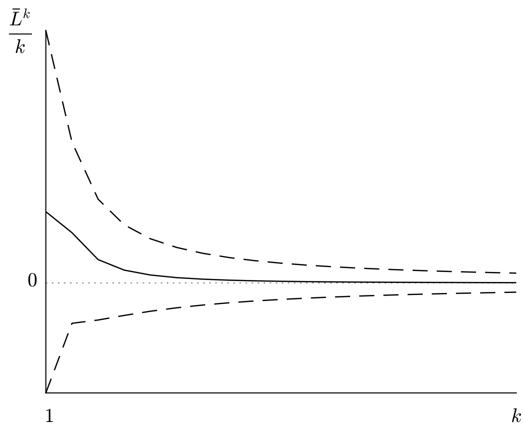


Figure 8: Panel A – Yield Curves. Panel B – Risk Measures.

Note: In Panel A, the light gray solid line is $y^k(\omega_0, 1)$, the light gray dotted line is $y^k(\bar{\omega}, 1)$, the dark gray solid line is $y^k(\omega_0, 2)$, and the dark gray dotted line is $y^k(\bar{\omega}, 2)$. In Panel B, the solid line is $\bar{L}^k/k$ and the dashed lines are $(\bar{L}^k - \Upsilon)/k$ and $(\bar{L}^k + \Upsilon)/k$.

zero with the horizon k showing that the influence of past shocks dies out over time. Similarly, the autocorrelation function, $\text{corr}(c_t^y, c_{t+k}^y)$, is monotonically declining in k and tends to zero with the horizon, implying that there is little persistence in consumption between generations born far apart from each other.

8 Extensions

We have kept the model as simple as possible for reasons of readability and tractability. In particular, to emphasize that the dynamics of the model derive from the participation constraints themselves and not from the underlying endowment process, we have considered an economic environment that is stationary. Nevertheless, the results and analysis are extendable in a number of directions and we briefly discuss five of them.

Heterogeneous Preferences The assumption in the basic model that the young and the old have a common utility function is not essential. Allowing for heterogeneity of preferences complicates the notation but does not change the results. As a special case, consider the situation in which the young have risk-neutral preferences. In this case, the ordering of the states $\lambda(s) \geq \lambda(q)$ for $s > q$ implies that the endowment of the old is monotonic in s: $e^o(s) \geq e^o(q)$. It can then be verified that the policy functions for the future promised utility and the consumption of the old are also monotonic in s: $f(\omega, s) \leq f(\omega, q)$ and $c^o(\omega, s) \geq c^o(\omega, q)$ for $s > q$.

Savings We have assumed that the only method of insurance is through intergenerational transfers. Now suppose that the young have access to a linear storage technology that delivers R units of endowment when they are old for every unit stored. Suppose that access to storage is not available in autarky but only to the young who do not default on the transfers they are called on to make. It is clear that storage is not used if the gross rate of return R is too low. In other words, there is an $\bar{R}$ such that for $R < \bar{R}$, storage is not used even if available. Given the solution to the optimal intergenerational insurance rule, $\bar{R}^{-1} = \max_x\{p^1(x)\}$. In the two-state case of Section 6, the maximum is attained when $x = (\omega_0, 1)$. With the parameter values of Example 1, $\bar{R} \approx 1.08171$, demonstrating that even when storage has a positive net return, the possibility of storage may have no impact on the optimal sustainable intergenerational insurance.

The situation is only slightly different if individuals have access to the same storage technology, with gross rate of return R, in autarky. Let a denote the amount stored, then the lifetime autarkic utility is:

$$\hat{v}(s; R) := \max_{a \geq 0}\ u(e^y(s) - a) + \beta \sum\nolimits_{s'} \pi(s')u(e^o(s') + Ra),$$

Let $a(s; R)$ denote the optimum amount stored. The function $\hat{v}(s; R)$ is increasing in R and $a(s; R)$ is weakly increasing in R (strictly if $a(s; R) > 0$). Hence, there is a critical value $\hat{R}$ such that for $R < \hat{R}$, $a(s; R) = 0$ for each s. In particular, $\hat{R}^{-1} = \max_s \sum_{s'} \hat{q}(s, s')$, where $\hat{q}(s, s')$ is the state price evaluated in autarky. We can show that $\hat{R} < 1$. To see this, note that Assumption 2 requires that the Perron root of the matrix $\hat{Q}$ is greater than one. This can only occur if at least one of the row sums of $\hat{Q}$ is greater than one: $\sum_{s'} \hat{q}(s, s') > 1$ for some state s (or equivalently, $\hat{R} < 1$).[35] Nevertheless, for $R \leq \hat{R}$, there is no storage in autarky and provided that $R \leq \min\{\hat{R}, \bar{R}\}$, the possibility of storage has no effect on the optimal sustainable intergenerational insurance. Of course, the interesting case is when the storage possibility does change the optimal sustainable intergenerational insurance, but this is left for future research.

Growth It has been assumed that the distribution of endowments is the same at all dates. This can be generalized to allow for stochastic growth of the endowment along the lines set out in Fernando Alvarez and Urban J. Jermann (2001) for a limited commitment model with infinitely-lived agents and in Dirk Krueger and Hanno Lustig (2010) for a Bewley economy. To do this, decompose the endowment state into an idiosyncratic shock ς and an aggregate growth shock θ so that $s = (\varsigma, \theta)$. Suppose that the idiosyncratic and aggregate components are i.i.d. and independent of each other, so that $\pi(s) = \pi(\varsigma)\pi(\theta)$ where $\pi(\varsigma)$ and $\pi(\theta)$ are the probabilities of the two shocks. Let $\alpha(\varsigma) = e^y_t/e_t$ denote the endowment share of the young and suppose that it depends only on ς. Similarly,

[35] For the two-state case of Section 6, and setting $\gamma = 1$ and $\pi = 1/2$, it can be verified that $\hat{R} = \beta^{-1}((1-\kappa)^2 - \sigma^2)/((1-\kappa)(\kappa+\sigma))$. Assumption 2 is satisfied for $\beta \geq ((1-\kappa)^2 - \sigma^2)/(\kappa(1-\kappa) + \sigma^2))$. Hence, $\hat{R} \leq (\kappa(1-\kappa) + \sigma^2)/((1-\kappa)(\kappa+\sigma)) < 1$. With $\kappa = 3/5$ and $\sigma = 1/10$, we have $\hat{R} \leq 25/28$.

let $\chi(\theta) = e_t/e_{t-1}$ denote the growth rate of the aggregate endowment and suppose that it depends only on θ at time t. Furthermore, suppose that the preferences of both agents exhibit CRRA with relative risk aversion coefficient γ. It is then possible to rewrite the planner's problem to be stationary as follows: normalize consumption at t by dividing by e_t, normalize utility variables ω, V and $\hat{v}$, at t by dividing by $e_t^{1-\gamma}$ and normalize the discount factors by multiplying by $\sum_\theta \pi(\theta)\chi(\theta)^{1-\gamma}$. The modified planner's problem with these normalized variables is identical to the stationary problem described in Section 3.[36] Hence, the solution with a stochastic growth component is obtained by a simple reinterpretation of the variables.

Altruism To incorporate altruism into the model, we consider "warm glow" preferences (see, for example, James Andreoni, 1989) where the old attach a weight of $\xi > 0$ to the utility of the young.[37] In this case, the lifetime utility of an individual born after the history s^t is:

$$u(c^y(s^t)) + \beta \sum\nolimits_{s_{t+1}} \pi(s_{t+1}) \left(u(c^o(s^{t+1})) + \xi u(c^y(s^{t+1})) \right).$$

With these preferences, the participation constraints of the old and the young are:

$$\begin{aligned} u(e(s) - c^y(s)) + \xi u(c^y(s)) \geq u(e^o(s)) + \xi u(e^y(s)) \quad \forall\, s \in \mathcal{S}, \\ u(c^y(s)) + \beta\omega'(s) \geq u(e^y(s)) + \beta \sum\nolimits_{s'} \pi(s') \left(u(e^o(s')) + \xi u(e^y(s')) \right) \quad \forall\, s \in \mathcal{S}, \end{aligned}$$

and the promise-keeping constraint is:

$$\sum\nolimits_s \pi(s) \left(u(e(s) - c^y(s)) + \xi u(c^y(s)) \right) \geq \omega.$$

The analysis of Section 3 can be applied *mutatis mutandis*. The first-order conditions for the optimal sustainable intergenerational insurance are given by equation (14) and

$$\frac{u_c(c^y(x))}{u_c(e(s) - c^y(x))} = \frac{\beta}{\delta} \left(\frac{1 + \nu(\omega) + \eta(x)}{1 + \mu(x) + \xi\frac{\beta}{\delta}(1 + \nu(\omega) + \eta(x))} \right), \tag{21}$$

where multipliers are as specified in Section 3. Condition (21) reduces to condition (13) when $\xi = 0$. It is easy to verify that Lemmas 2 and 3 continue to hold when $\xi > 0$, provided that ξ is not too large.

[36] The growth rate cannot be too high. The modified discount factor must satisfy $\delta \sum_\theta \pi(\theta)\chi(\theta)^{1-\gamma} < 1$ for the planner's problem to be finite.

[37] It is simple to include a bequest from the old. The optimal solution determines the transfer net of the bequest. Although there may be bequests in autarky, provided $\xi > 0$ is small enough, the bequest at the optimum is zero.

Renegotiation-proofness The optimal sustainable intergenerational insurance is not renegotiation-proof. The reason is that the young receive their autarkic utility when they default. In the case of default, it would be possible to offer the promised utility ω_0 instead of $\omega_{\min}$ without diminishing the planner's payoff. Hence, the constrained efficient solution characterized in Section 3 is not renegotiation-proof despite belonging to the Pareto frontier of the set of all equilibrium payoffs. A simple modification of the participation constraint is needed to find a renegotiation-proof outcome. Replace the participation constraint of the young in equation (9) by:

$$u\left(c^y(s)\right) + \beta\omega'(s) \geq u\left(e^y(s)\right) + \beta\omega_0. \tag{22}$$

Since ω_0 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.[38] Although imposing the tighter constraint (22) restricts risk sharing, the qualitative properties of the optimal solution are substantially unchanged.

9 Discussion

Comparison with infinitely-lived agents It is worthwhile contrasting our results to those on risk sharing and limited commitment with infinitely-lived agents. The case of two infinitely-lived agents with endowment risk has been considered by Thomas and Worrall (1988) and Kocherlakota (1996). A common feature of that model with the overlapping generations model considered here is that only one agent is ever constrained at any point in time, namely, the agent making the transfer or facing a non-negativity constraint. To illustrate the contrast between the two models, suppose there are just two i.i.d. endowment states: in state 1, one agent has a higher endowment than the other, while in state 2 the situation is reversed. In the infinitely-live case, there is convergence to the invariant distribution immediately after both states have occurred. If there is partial insurance at the invariant distribution, then there are two different ratios of marginal utilities associated with the two endowment states. Since consumption is determined by the endowment state, the conditional and unconditional autocorrelation of consumption between two adjacent dates is zero and there is no persistence in consumption. In the overlapping generations model, if there is partial insurance, then convergence occurs but not within a finite period of time. With two endowment states, each generation faces only two potential ratios of marginal utilities, but these ratios differ from one generation to the

[38]Edward C. Prescott and José-Víctor Ríos-Rull (2005) consider a similar condition in a deterministic overlapping generations model, which they refer to as a no-restarting condition.

next depending on the previous promise. Hence, as we showed in Section 6, consumption is autocorrelated even in the long run.

The optimal solution in the overlapping generations model is closer to that in models of risk sharing and limited commitment with a continuum of infinitely-lived agents (see, for example, Thomas and Worrall, 2007; Krueger and Perri, 2011; Broer, 2013). In those models, agents have high or low income (employed or unemployed). If there is partial insurance, then there is a finite set of possible transfers from the employed to the unemployed at the invariant distribution. There are three different ratios of marginal utilities, one each for constrained agents whether employed or unemployed and one for unconstrained and unemployed agents. Each ratio determines a value of consumption that depends only on the employment state for constrained agents but varies with the spell of unemployment for unconstrained and unemployed agents. To maintain a constant growth rate of marginal utilities, the consumption of an unconstrained and unemployed agent varies over time. This is in contrast to the overlapping generations model considered here, where consumption changes when the same state reoccurs and the young are constrained. Another point of contrast is that we solve for the optimal sustainable intergenerational insurance for any given promise and 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 special case.

Pareto Optimality Two concepts of Pareto optimality are widely used in stochastic overlapping generations models: *ex ante* optimality and *interim* optimality. In an ex ante Pareto optimum, it is impossible to increase the expected lifetime utility of one generation without reducing that of another generation. In the interim case, generations are distinguished by the state as well as the date of birth and at a Pareto optimum it is not possible to increase the lifetime utility of any generation without decreasing the lifetime utility of another generation at a different date or in a different state. These two concepts are often referred to as equal-treatment Pareto optimality and conditional Pareto optimality (see, for example, Aiyagari and Peled, 1991; Labadie, 2004).[39]

The optimal sustainable intergenerational insurance we have characterized is conditionally Pareto optimal. The ratio of marginal utilities varies across endowment states because the participation constraint of the young is sometimes binding. This means that there is only partial sharing of endowment risk, unlike in an equal-treatment Pareto op-

[39] The terms interim and conditional Pareto optimality are not used with complete consistency in the literature. A distinction is sometimes drawn depending upon whether allocations considered are time-invariant or non-stationary (see, for eample, Gabrielle Demange, 2002).

timum. The ratio of marginal utilities is equal across all endowment states only when the participation constraint of the young does not bind at all, that is, when the optimal allocation is first best. In this case, the solution achieves equal treatment. Nevertheless, even when the first best cannot be sustained, the optimal sustainable intergenerational insurance has a built-in equal-treatment property. All generations born in the same state $x = (\omega, s)$ receive the same conditional allocation.

The Pareto weight that each agent receives at the optimum depends on the date of birth, the promised utility and the endowment state at that date. By assumption, the utility of a generation born at time t is discounted geometrically by a weight of δ^t. To understand the Pareto weight for each agent, reconsider condition (13) and, for the purposes of discussion, ignore the non-negativity constraint on transfers and the upper bound constraint on promised utility. In determining the optimum, the Pareto weight on the utility of the old born at t can be interpreted to be $\delta^t\beta(1 + \nu(\omega))$, where ω is the promise they received when they were young. The young they interact with at $t+1$ have a Pareto weight of $\delta^{t+1}(1 + \mu(\omega, s))$, which reflects the promise made to the old and the endowment state. By the updating property of equation (16), $\nu(\omega') = \mu(\omega, s)$. If $\mu(\omega, s) < \nu(\omega)$ (for example, if $\nu(\omega) > 0$ and the participation constraint of the young does not bind), then $\omega' < \omega$ and the young are given a lower weight than the old (since the young are born later, they always receive a lower weight by the factor δ). If, on the other hand, $\mu(\omega, s) > \nu(\omega)$, that is, when the participation constraint of the young binds sufficiently that they are promised more than the current old, then $\omega' > \omega$ and the young are given a greater weight than the old. Therefore, a generation born at t and receiving a promise of ω' for old age can be considered to have a Pareto weight $\delta^t(1 + \nu(\omega'))$. The optimal intergenerational insurance is conditionally Pareto optimal, each generation with the same promise and endowment is treated equally, but the Pareto weights are determined endogenously by the participation constraints and the history of shocks.

10 Conclusion and Future Research

A theory of intergenerational insurance is developed in a stochastic overlapping generations model in which risk-sharing transfers are voluntary. The model implies that (i) generational risk is spread across future generations in ways that create history dependence of transfers with periodic resetting, at which time the history of shocks is forgotten; (ii) there is heteroscedasticity of consumption conditional on the endowment and autocorrelation of consumption between adjacent generations, even in the long run.

The model is parsimonious in its assumptions to highlight the role played by the limited enforcement of risk-sharing transfers. In particular, it assumes that the endowment process is stationary. Despite the stationarity of the environment, the limited enforcement of transfers can generate heteroscedasticity and autocorrelation of consumption with the effect of a shock persisting for many periods. An alternative approach is to allow for rich dynamics in the endowment or earnings process (see, for example, Mariacristina De Nardi, Giulio Fella and Gonzalo Paz-Pardo, 2020) and to study the implication for consumption without the need to posit any enforcement friction. An important avenue for future research is to combine these two approaches to help understand the role of both earning dynamics and enforcement in determining consumption allocations.

Although we have shown how the basic model can be extended in various directions, the model ignores important aspects of intergenerational insurance. Firstly, it cannot examine the interdependence between intergenerational insurance and intra-generational insurance since there is only a representative agent in each generation. Secondly, although we showed how storage could be incorporated into the model, we only considered cases in which storage is not used at the optimum. Therefore, there are two natural and important directions for future research: (*i*) enriching the demographic structure, either by having more than two overlapping generations or by allowing for heterogeneous agents within the same cohort and (*ii*) introducing a storage or production technology that is used at the optimum.

References

Açikgöz, Ömer. 2018. "On the Existence and Uniqueness of Stationary Equilibrium in Bewley Economies with Production." *Journal of Economic Theory*, 173: 18–25.

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.

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

Andreoni, James. 1989. "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence." *Journal of Political Economy*, 97(6): 1447–1458.

Backus, David, Mikhail Chernov, and Stanley Zin. 2014. "Sources of Entropy in Representative Agent Models." *Journal of Finance*, 69(1): 51–99.

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.

Bloise, Gaetano, and Filippo L. Calciano. 2008. “A Characterization of Inefficiency in Stochastic Overlapping Generations Economies.” *Journal of Economic Theory*, 143(1): 442–468.

Boldrin, Michele, and Aldo Rustichini. 2000. “Political Equilibria with Social Security.” *Review of Economic Dynamics*, 3(1): 41–78.

Borch, Karl. 1962. “Equilibrium in a Reinsurance Market.” *Econometrica*, 30(3): 424–444.

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.

Chattopadhyay, Subir, and Piero Gottardi. 1999. “Stochastic OLG Models, Market Structure, and Optimality.” *Journal of Economic Theory*, 89(1): 21–67.

Christensen, Timothy M. 2017. “Nonparametric Stochastic Discount Factor Decomposition.” *Econometrica*, 85(5): 1501–1536.

Cooley, Thomas, and Jorge Soares. 1999. “A Positive Theory of Social Security Based on Reputation.” *Journal of Political Economy*, 107(1): 135–160.

Demange, Gabrielle. 2002. “On Optimality in Intergenerational Risk Sharing.” *Economic Theory*, 20(1): 1–27.

Demange, Gabrielle, and Guy Laroque. 1999. “Social Security and Demographic Shocks.” *Econometrica*, 67(3): 527–542.

De Nardi, Mariacristina, Giulio Fella, and Gonzalo Paz-Pardo. 2020. “Nonlinear Household Earnings Dynamics, Self-Insurance, and Welfare.” *Journal of the European Economic Association*, 18(2): 890—926.

Diamond, Peter. 1977. “A Framework for Social Security Analysis.” *Journal of Public Economics*, 8(3): 275–298.

Enders, Walter, and Harvey Lapan. 1982. “Social Security Taxation and Intergenerational Risk Sharing.” *International Economic Review*, 23(3): 647–58.

Farhi, Emmanuel, and Iván Werning. 2007. “Inequality and Social Discounting.” *Journal of Political Economy*, 115(3): 365–402.

Foss, Sergey, Vsevolod Shneer, Jonathan Thomas, and Tim Worrall. 2018. “Stochastic Stability of Monotone Economies in Regenerative Environments.” *Journal of Economic Theory*, 173: 334–360.

Glover, Andrew, Jonathan Heathcoate, Dirk Kruger, and José-Víctor Ríos-Rull. 2020a. “Intergenerational Redistribution in the Great Recession.” *Journal of Political Economy*, 128(10): 3730–3778.

Glover, Andrew, Jonathan Heathcoate, Dirk Kruger, and José-Víctor Ríos-Rull. 2020b. "Health versus Wealth: On the Distributional Effects of Controlling a Pandemic." Centre for Economic Policy Research Discussion Paper 14606.

Gonzalez-Eiras, Martín, and Dirk Niepelt. 2008. "The Future of Social Security." *Journal of Monetary Economics*, 55(2): 197–218.

Gordon, Roger, and Hal Varian. 1988. "Intergenerational Risk Sharing." *Journal of Public Economics*, 37(2): 185–202.

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

Huberman, Gur. 1984. "Capital Asset Pricing in an Overlapping Generations Model." *Journal of Economic Theory*, 33(2): 232–248.

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

International Labour Organization. 2017. *World Social Protection Report 2017–19: Universal Social Protection to achieve the Sustainable Development Goals.* Geneva: International Labour Office.

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

Kiyotaki, Nobuhiro, and Shengxing Zhang. 2018. "Intangibles, Inequality and Stagnation." Mimeo.

Kocherlakota, Narayana. 1996. "Implications of Efficient Risk Sharing without Commitment." *Review of Economic Studies*, 63(4): 595–610.

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.

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

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

Labadie, Pamela. 2004. "Aggregate Risk Sharing and Equivalent Financial Mechanisms in an Endowment Economy of Incomplete Participation." *Economic Theory*, 24(4): 789–809.

Manuelli, Rodolfo. 1990. "Existence and Optimality of Currency Equilibrium in Stochastic Overlapping Generations Models: The Pure Endowment Case." *Journal of Economic Theory*, 51(2): 268–294.

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

Prescott, Edward C., and José-Víctor Ríos-Rull. 2005. "On Equilibrium for Overlapping Generations Organizations." *International Economic Review*, 46(4): 1065–1080.

Rangel, Antonio, and Richard Zeckhauser. 2000. "Can Market and Voting Institutions Generate Optimal Intergenerational Risk Sharing?" In *Risk Aspects of Investment-Based Social Security Reform, National Bureau of Economic Research Conference Report*, edited by John Y. Campbell and Martin Feldstein, 113–152. Chicago: The University of Chicago Press.

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

Samuelson, Paul. 1958. "An Exact Consumption-Loan Model of Interest with or without the Social Contrivance of Money." *Journal of Political Economy*, 66(6): 467–482.

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.

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.

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

Thomas, Jonathan, and Tim Worrall. 1988. "Self-Enforcing Wage Contracts." *Review of Economic Studies*, 55(4): 541–554.

Thomas, Jonathan, and Tim Worrall. 1994. "Foreign Direct Investment and the Risk of Expropriation." *Review of Economic Studies*, 61(1): 81–108.

Thomas, Jonathan, and Tim Worrall. 2007. "Unemployment Insurance under Moral Hazard and Limited Commitment: Public versus Private Provision." *Journal of Public Economic Theory*, 9(1): 151–181.

Zhu, Shenghao. 2020. "Existence of Equilibrium in an Incomplete-Market Model with Endogenous Labor Supply." *International Economic Review*, 61(3): 1115–1138.

Appendix

Proof of Lemma 2. It is established in the proof of Lemma 1 in Part B of the Supplementary Appendix that the participation constraints of the young and old cannot bind simultaneously in a given endowment state. For convenience, define:

$$h(\omega) := -\tfrac{\delta}{\beta}V_\omega(\omega); \quad g^c(\omega) := \tfrac{\beta}{\delta}(1 + h(\omega)); \quad v(\omega, \mu; e) := u(y(\omega, \mu; e)) + \beta h^{-1}(\mu - \zeta);$$

where $y(\omega, \mu; e)$ and $\vartheta(\omega, e, \hat{v})$ are defined implicitly by:

$$\frac{g^c(\omega)}{1+\mu} = \frac{u_c(y(\omega, \mu; e))}{u_c(e - y(\omega, \mu; e))}; \;\; v(\omega, \vartheta(\omega; e, \hat{v}); e) = \hat{v}; \;\; \text{and} \;\; g(\omega, \mu; e) := \frac{u_c(y(\omega, \mu; e))}{u_c(e - y(\omega, \mu; e))}.$$

The term $y(\omega, \mu; e)$ is the consumption of the young given ω, the multiplier μ and the aggregate endowment e. Likewise, $v(\omega, \mu; e)$ is the lifetime utility of the young and $\vartheta(\omega, e, \hat{v})$ is the value of μ when the participation constraint of the young is binding. Recall that $\omega_0 = \sup\{\omega \mid V_\omega(\omega) = 0\}$. The function $h \colon \Omega \to [0, \bar{\nu}]$ is strictly increasing for $\omega > \omega_0$ with $h(\omega_{\max}) = \bar{\nu}$ and $h(\omega) = 0$ for $\omega \le \omega_0$. From equation (14), $\omega' = h^{-1}(\mu - \zeta)$. The function y is continuous by the implicit function theorem because the derivative u_c and the function h are continuous. It can be checked that y is increasing in e (and $\partial y/\partial e < 1$), increasing in μ and decreasing in ω. Recall that $\zeta > 0$ only if $\omega' = \omega_{\max}$ and hence, $\mu \ge \bar{\nu}$. For $\mu = 0$, and hence, $\zeta = 0$, $h^{-1}(0) = \omega_0$, $y(\omega_0, 0; e) = c^{y*}(s)$, and $v(\omega_0, 0; e) = u(c^{y*}(s)) + \beta\omega_0$. It follows from the properties of $v(\omega, \mu; e)$ that ϑ is increasing in ω (weakly because the solution may be $\mu = 0$), decreasing in e and increasing in $\hat{v}$.

(i) Since the constraint set Φ is convex and the objective function is strictly concave, the policy function $f(\omega, s)$ is single-valued and continuous in ω. It is also non-decreasing in ω. It follows from the definitions above that $\omega' = f(\omega, s) = \min\{h^{-1}(\vartheta(\omega; e, \hat{v})), \bar{\omega}\}$, where $\bar{\omega} = \max_s \omega^f(s)$ is the largest fixed point of the mappings $f(\omega, s)$. For $\vartheta = 0$, $f(\omega, s) = \omega_0$. For $\vartheta > 0$, $h^{-1}(\vartheta(\omega; e, \hat{v}))$ is strictly increasing in ω and hence, $f(\omega, s)$ is strictly increasing in ω provided $f(\omega, s) < \omega_{\max}$. If $\zeta(\omega, s) > 0$ (or equivalently, $h^{-1}(\vartheta(\omega; e, \hat{v})) > \omega_{\max}$), then $f(\omega, s) = \omega_{\max}$.

(ii) The value of the critical threshold, $\omega^c(e, \hat{v})$, above which μ is positive, is determined by $v(\omega^c(e, \hat{v}), 0; e) = \hat{v}$. Thus, $\omega^c(e, \hat{v})$ is increasing in e and decreasing in $\hat{v}$. We next show that there is some state r such that $\omega^c(e(r), \hat{v}(r)) > \omega_0$. Taking $r = S$, it can be shown that $\mu(\omega_0, S) = 0$. To see this, suppose to the contrary that $\mu(\omega_0, S) > 0$. Then, $\eta(\omega_0, S) = 0$ and $g(\omega_0, \mu(\omega_0, S), e(S)) < \beta/\delta = g^c(\omega_0)$. Since there is no transfer from the old, $y(\omega_0, \mu(\omega_0, S), e(S)) \leq e^y(S)$ and hence, $g(\omega_0, \mu(\omega_0, S), e(S)) \geq \lambda(S)$. By Assumption 3, $\lambda(S) \geq \beta/\delta$, which gives a contradiction. Since $\omega_0 > \omega_{\min}$, $v(\omega_0, 0; e(S)) = u(e^y(S)) + \beta\omega_0 > u(e^y(S)) + \beta\omega_{\min} = \hat{v}(S)$. Finally, since $v(\omega, 0; e)$ is continuous and decreasing in ω and $v(\omega^c(e(S), \hat{v}(S)), 0; e(S)) = \hat{v}(S)$, it follows that $\omega^c(e(S), \hat{v}(S)) > \omega_0$, as required.

(iii) Existence of a fixed point $\omega^f(s)$ of the mapping $f(\omega, s)$ follows from the standard fixed point theorem given the continuity and monotonicity of $f(\omega, s)$ in ω. Part (ii) shows that there is at least one state, namely $s = S$, for which $\mu(\omega_0, s) = 0$. By Lemma 1, $\omega_0 < \omega^*$ and hence, at least one of the participation constraints of the young is binding at $\omega = \omega_0$. Thus, there is at least one state $r \in \mathcal{S}$ such that $\mu(\omega_0, r) > 0$. It follows that the set of states can be partitioned into two non-empty subsets, $\hat{\mathcal{S}}$ and its complement with $\mu(\omega_0, s) = 0$ for $s \in \hat{\mathcal{S}}$ and $\mu(\omega_0, r) > 0$ for $r \notin \hat{\mathcal{S}}$.

For states $s \in \hat{\mathcal{S}}$, $\mu(\omega_0, s) = \zeta(\omega_0, s) = 0$ and hence, $f(\omega_0, s) = h^{-1}(0) = \omega_0$. That is, ω_0 is a fixed point of the mapping $f(\omega, s)$. Since $f(\omega, s) \geq \omega_0$ (see, part (i)), there can be no $\omega^f(s) < \omega_0$. Now consider a state $s \notin \hat{\mathcal{S}}$ where $\mu(\omega_0, s) > 0$. Since $\nu(\omega_0) = 0$ and $\eta(\omega_0, s) = 0$ by complementary slackness, it follows that $f(\omega_0, s) > \omega_0$ (with $f(\omega_0, s) = \omega_{\max} > \omega_0$ if $\zeta(\omega_0, s) > 0$). That is, in any state where $\mu(\omega_0, s) > 0$, any fixed point satisfies $\omega^f(s) > \omega_0$. First, note that for $\omega^f(s) > \omega_0$, $\mu(\omega^f(s), s) = \nu(\omega^f(s)) + \zeta(\omega^f(s), s)$. If $\zeta(\omega^f(s), s) > 0$, then $\nu(\omega^f(s)) = \bar{\nu}$ and $\omega^f(s) = \omega_{\max}$. Then, from condition (13), $c^y(\omega_{\max}, s) < c^{y*}(s)$ and $u(c^y(\omega_{\max}, s)) = u(e^y(s)) - \beta(\omega_{\max} - \omega_{\min})$. If $\zeta(\omega^f(s), s) = 0$, then $\mu(\omega^f(s), s) = \nu(\omega^f(s))$ and hence, from condition (13), $c^y(\omega^f(s), s) = c^{y*}(s)$. Hence, $\omega^f(s) = \omega_{\min} + \beta^{-1}(u(e^y(s)) - u(c^{y*}(s)))$. Taking the cases $s \in \hat{\mathcal{S}}$ and $s \notin \hat{\mathcal{S}}$ together, we obtain:

$$\omega^f(s) = \min\left\{\max\left\{\omega_0, \omega_{\min} + \beta^{-1}\left(u(e^y(s)) - u(c^{y*}(s))\right)\right\}, \omega_{\max}\right\}. \tag{23}$$

From Proposition 2, $c^{y*}(s)$ is unique and hence, from equation (23) it follows that $\omega^f(s)$ is unique. There may, of course, be multiple states with the same fixed point.

(iv) Recall that $\omega'(\omega, e, \hat{v}) = \min\{h^{-1}(\vartheta(\omega, e, \hat{v})), \bar{\omega}\}$. It follows from the properties of h and ϑ derived above that $\omega'(\omega, e, \hat{v})$ is decreasing in e and increasing in $\hat{v}$. To determine how $f(\omega, s)$ depends on s, we need to know how $e(s)$ and $\hat{v}(s)$ depend on s. When e is fixed, it follows from the convention on $\lambda(s)$ that $e^y(1) \geq e^y(2) \geq \cdots \geq e^y(S)$. Hence, $\hat{v}(s)$ is decreasing in s. Moreover, for distinct states s and r with $s < r$ and $e^y(s) > e^y(r)$, we have $\hat{v}(s) > \hat{v}(r)$. Thus, $f(\omega, s) = \min\{h^{-1}(\vartheta(\omega, e, \hat{v}(s))), \bar{\omega}\}$ is decreasing in s. Moreover, for distinct s and r such that both $f(\omega, s) \in (\omega_0, \omega_{\max})$ and $f(\omega, r) \in (\omega_0, \omega_{\max})$, $f(\omega, s) >$

$f(\omega, r)$ for $e^y(s) > e^y(r)$. We can order fixed points: $\omega^f(s) \geq \omega^f(r)$ for every $s < r$, with strict inequality unless $\omega^f(s) = \omega^f(r) = \omega_0$ or $\omega^f(s) = \omega^f(r) = \omega_{\max}$. ∎

Proof of Lemma 3.

(i) It is established in part (i) of Lemma 2 that the function $y(\omega, \mu; e)$ is strictly decreasing in ω for a fixed μ and e, provided that the non-negativity constraint on transfers does not bind. If $\mu(\omega, s) > 0$, then $u(c^y(\omega, s)) + \beta f(\omega, s) = \hat{v}(s)$. Since Lemma 2 establishes that $f(\omega, s)$ is strictly increasing in ω for $f(\omega, s) \in (\omega_0, \omega_{\max})$, it follows that $c^y(\omega, s)$ is strictly decreasing in ω for $c^y(\omega, s) < e^y(s)$ and $f(\omega, s) < \omega_{\max}$.

(ii) This is shown in the proof of part (iii) of Lemma 2.

(iii) For a fixed aggregate endowment, the lifetime endowment utility $\hat{v}(s)$ is decreasing in s. If the participation constraint of the young is binding, then from Lemma 2, $\vartheta(\omega; e, \hat{v})$ is increasing in $\hat{v}$. Since $y(\omega, \vartheta(\omega; e, \hat{v}); e)$ is increasing in ϑ, $c^y(\omega, s)$ is increasing in $\hat{v}$ and hence, decreasing in s. It is constant in s if the participation constraint of neither the young nor the old is binding. ∎

Proof of Proposition 4. Since there is an $N \geq 1$ and $\epsilon > 0$ such that $P^N(\omega, \{\omega_0\}) > \epsilon$ for all $\omega \in [\omega_0, \bar{\omega}]$, it follows that Condition **M** of Stokey, Lucas Jr. and Prescott (1989, page 348) is satisfied. Then, Theorem 11.12 of Stokey, Lucas Jr. and Prescott (1989) is used to establish strong convergence. Non-degeneracy follows from Assumption 4 and existence of a mass point at ω_0 from Lemma 2. ∎

Proof of Proposition 5.

(i) To simplify notation, let $m(\omega, s, s') := m((\omega, s), (f(\omega, s), s'))$. It follows from equation (17) and the monotonicity of $f(\omega, s)$ in ω that $m(\omega, s, s')$ is decreasing in ω. The price of a one-period discount bond in state (ω, s) is $p^1(\omega, s) = \sum_{s'} \pi(s') m(\omega, s, s')$, which is decreasing in ω. Making the induction hypothesis that the price of a k-period discount bond is decreasing in ω, $p^{k+1}(\omega, s) = \sum_{s'} \pi(s') m(\omega, s, s') p^k(f(\omega, s), s')$. Since $p^k(\omega, s)$ and $m(\omega, s, s')$ are positive and decreasing in ω, and $f(\omega, s)$ is increasing in ω, it follows that $p^{k+1}(\omega, s)$ is decreasing in ω. Hence, the conditional yield $y^k(\omega, s) = -\log(p^k(\omega, s))$ and the deviation from the mean $\Delta y^k(\omega, s)$ are increasing in ω.

(ii) This is a standard result (see, for example, Martin and Ross, 2019).

(iii) It follows from part (ii) that $\lim_{k \to \infty} y^k(x) = \mathbb{E}_\varphi[\log(m(x, x'))] = \log(\rho)$, where $\mathbb{E}_\varphi$ is the expectation taken over the invariant distribution of x and ρ is the Perron root of the matrix Q. Taking logs of equation (17) gives:

$$\log\left(m(x,x')\right) = \log\left(\beta\right) - \log\left(u_c(c^y(x))\right) + \log\left(u_c(c^o(x'))\right).$$

Furthermore, taking logs in condition (13) and moving it one period forward gives:

$$\log\left(u_c(c^y(x'))\right) - \log\left(u_c(c^o(x'))\right) = \log\left(\frac{\beta}{\delta}\right) + \log\left(1+\nu(\omega')+\eta(x')\right) - \log\left(1+\mu(x')\right).$$

Combining these two equations and using $\nu(\omega') = \mu(x) - \zeta(x)$ gives:

$$\begin{aligned}\log\left(m(x,x')\right) = \log\left(\delta\right) + \log\left(u_c(c^y(x'))\right) - \log\left(u_c(c^y(x))\right) + \log\left(1+\mu(x')\right) - \log\left(1+\mu(x)\right) \\ -\left(\log\left(1+\mu(x)-\zeta(x)+\eta(x')\right) - \log\left(1+\mu(x)\right)\right).\end{aligned}$$

Assume that $\zeta(x) = \eta(x') = 0$ for all x. Then, the last term in the above equation is zero and $\mathbb{E}_\varphi[\log(m(x,x'))] = \log(\delta)$. Hence, comparing to the standard result given in part (*ii*), when $\zeta(x) = \eta(x') = 0$ for all x, $\rho = \delta$. If $\zeta(x) = 0$ for each x, that is, if $\bar{\omega} < \omega_{\max}$, then $\rho \leq \delta$. If $\eta(x) = 0$ for each x, then $\rho \geq \delta$, with strict inequality if $\zeta(x) > 0$ for some x.

(iv) With no aggregate risk, $\omega^f(s)$ is ordered by state. In particular, $\bar{\omega} = \omega^f(1) > \omega^f(S) = \omega_0$. Consider state (ω_0, S). Since ω_0 is a fixed point of the mapping $f(\omega, S)$ and $\mu(\omega_0, S) = 0$, it follows from equations (13) and (17) that $m(\omega_0, S, S) \geq \delta$, with equality if $\eta(\omega_0, S) = 0$. Furthermore, it follows from equation (17) that $m(\omega, s, s')$ is decreasing in s'. Hence, for each $s < S$, $m(\omega_0, S, s) \geq m(\omega_0, S, S) \geq \delta$, with at least one of the inequalities strict. Taking expectations, the bond price $p^1(\omega_0, S) > \delta$. Consequently, the yield $y^1(\omega_0, S) < -\log(\delta)$. Similarly, it can be checked that $m(\bar{\omega}, 1, 1) \leq \delta$, with equality if $\zeta(\bar{\omega}, 1) = 0$, that is, if $\bar{\omega} < \omega_{\max}$. Hence, $m(\bar{\omega}, 1, s) \leq \delta$ for each $s > 1$, with strict inequality for some state, and consequently, $y^1(\bar{\omega}, 1) > -\log(\delta)$. ∎

Proof of Proposition 6.

(i) It follows from parts (*i*) and (*iii*) of Lemma 3 that the functions $c^y(\omega, s)$ are monotonically decreasing in ω and ordered by the state: $c^y(\omega, 1) > c^y(\omega, 2)$. Since $\bar{\omega}$ and ω_0 are fixed points of $f(\omega, 1)$ and $f(\omega, 2)$, it follows from part (*ii*) of Lemma 3 that $c^y(\bar{\omega}, 1) = c^{y*}(1)$ and $c^y(\omega_0, 2) = c^{y*}(2)$. With no aggregate risk, the first-best consumption is independent of the state, $c^{y*}(1) = c^{y*}(2)$, and hence, $c^y(\bar{\omega}, 1) = c^y(\omega_0, 2)$.

(ii) From Proposition 5, the Perron root of the state price matrix is $\rho = \delta$. The properties of the corresponding eigenvector ψ are derived from condition (13), the definition of the stochastic discount factor in equation (17) and the Ross Recovery Theorem that

$$m(x,x') = \rho \frac{\psi(x)}{\psi(x')} = \delta \frac{u_c(c^y(\omega',s'))}{u_c(c^y(\omega,s))} \frac{(1+\mu(\omega',s'))}{(1+\mu(\omega,s))}.$$

Since $\mu(\omega,2)=0$, let $\upsilon(\omega) := 1+\mu(\omega,1)$ and then, using $\rho=\delta$, we can write:

$$\psi(\omega,1) = \frac{1}{u_c(c^y(\omega,1))\upsilon(\omega)} \quad \text{and} \quad \psi(\omega,2) = \frac{1}{u_c(c^y(\omega,2))}. \tag{24}$$

Since $f(\omega,1)$ is increasing in ω, $\upsilon(\omega)$ is also monotonically increasing in ω. It therefore follows that both $\psi(\omega,1)$ and $\psi(\omega,2)$ are decreasing in ω. With CRRA utility, condition (13) can be used to solve for the consumption of the young:

$$c^y(f(\omega,1),1) = \frac{\upsilon(f(\omega,1))^{\frac{1}{\gamma}}}{\upsilon(f(\omega,1))^{\frac{1}{\gamma}} + \left(\frac{\beta}{\delta}\upsilon(\omega)\right)^{\frac{1}{\gamma}}}; \quad c^y(f(\omega,1),2) = \frac{1}{1+\left(\frac{\beta}{\delta}\upsilon(\omega)\right)^{\frac{1}{\gamma}}};$$

$$c^y(\omega_0,1) = \frac{\upsilon(\omega_0)^{\frac{1}{\gamma}}}{\upsilon(\omega_0)^{\frac{1}{\gamma}} + \left(\frac{\beta}{\delta}\right)^{\frac{1}{\gamma}}}; \; c^y(\omega_0,2) = c^y(\bar{\omega},1) = \frac{1}{1+\left(\frac{\beta}{\delta}\right)^{\frac{1}{\gamma}}}; \; c^y(\bar{\omega},2) = \frac{1}{1+\left(\frac{\beta}{\delta}\upsilon(\bar{\omega})\right)^{\frac{1}{\gamma}}}.$$

Since $\upsilon(\omega) > 1$ for all ω, these consumption values can be substituted into (24) to show that $\psi(\omega,1) < \psi(\omega,2)$. Furthermore, with $\psi(\omega,1)$ and $\psi(\omega,2)$ decreasing in ω, $\psi_{\max} = \psi(\omega_0,2) = 1/u_c(c^{y*}(2))$ and $\psi_{\min} = \psi(\bar{\omega},1) = 1/(u_c(c^{y*}(1))\upsilon(\bar{\omega}))$. With no aggregate risk, $c^{y*}(1) = c^{y*}(2)$ and hence, $\Upsilon = \log(\psi_{\max}/\psi_{\min}) = \log(\upsilon(\bar{\omega}))$. To compute $\upsilon(\bar{\omega})$, we use the fact that both the participation constraint of the young in state 1 and the promise-keeping constraint bind at $\omega=\bar{\omega}$, that is,

$$\beta\left(\pi\left(u\left(c^o(\bar{\omega},1)\right) - u\left(e^o(1)\right)\right) + (1-\pi)\left(u\left(c^o(\bar{\omega},2)\right) - u\left(e^o(2)\right)\right)\right) = u\left(e^y(1)\right) - u\left(c^y(\bar{\omega},1)\right).$$

From this, with CRRA utility and the consumption values given above, we get:

$$\Upsilon = \log\left(\upsilon(\bar{\omega})\right) = \log(\delta) - \log(\beta) - \gamma \log\left(\Xi_\gamma^{\frac{1}{\gamma-1}} - 1\right), \quad \text{where}$$

$$\Xi_\gamma = \left(\tfrac{1}{\beta(1-\pi)}\right)(\kappa+\sigma)^{1-\gamma} + \beta\left(\tfrac{\pi}{1-\pi}\right)(1-\kappa-\sigma)^{1-\gamma} + \left(1-\kappa+\sigma\tfrac{\pi}{1-\pi}\right)^{1-\gamma}$$
$$- \left(\tfrac{1}{\beta(1-\pi)}\right)\left(\left(\tfrac{1}{1+\left(\frac{\beta}{\delta}\right)^{\frac{1}{\gamma}}}\right)^{1-\gamma} + \beta\pi\left(\tfrac{\left(\frac{\beta}{\delta}\right)^{\frac{1}{\gamma}}}{1+\left(\frac{\beta}{\delta}\right)^{\frac{1}{\gamma}}}\right)^{1-\gamma}\right).$$

In the limit as $\gamma \to 1$, we have:

$$\Xi_1 = \left(\frac{\delta}{\beta}\right)^{\frac{\pi}{1-\pi}} \left(\frac{\beta+\delta}{\delta}\right)^{\frac{1+\beta\pi}{\beta(1-\pi)}} (\kappa+\sigma)^{\frac{1}{\beta(1-\pi)}} (1-\kappa-\sigma)^{\frac{\pi}{1-\pi}} \left(1-\kappa+\sigma\frac{\pi}{1-\pi}\right).$$

(iii) Conditional entropy is defined as:

$$L(x) = -\sum\nolimits_{x'} \pi(x,x') \log\left(\frac{\varrho(x,x')}{\pi(x,x')}\right),$$

where $\pi(x,x') = \pi(s')$. With $x = (\omega, s)$ and $x' = (f(\omega,s), s')$, it follows that

$$\frac{\varrho(x,x')}{\pi(x,x')} = \frac{m(\omega,s,s')}{\pi m(\omega,s,1) + (1-\pi)m(\omega,s,2)},$$

where $m(\omega,s,s')$ is defined in the proof of Proposition 5. Therefore,

$$L(\omega,s) = \log\left(\pi m(\omega,s,1) + (1-\pi)m(\omega,s,2)\right) - \pi\log(m(\omega,s,1)) - (1-\pi)\log(m(\omega,s,2)). \quad (25)$$

From the Ross Recovery Theorem, $m(\omega,s,s') = \rho\psi(\omega,s)/\psi(f(\omega,s),s')$ and, by assumption, $f(\omega,2) = \omega_0$. Therefore, substituting $m(\omega,s,s')$ into equation (25) gives equations (19) in the statement of the proposition. ■

Proof of Proposition 7.

(i) Since the promised utility ω is reset to ω_0 whenever state 2 occurs, the probability that the promised utility is ω_0 is $1-\pi$, irrespective of the date or history. Therefore, T periods after such a resetting, the distribution of ω is:

$$\phi_T\left(\{\omega_0^{(n)}\}\right) = (1-\pi)\pi^n \quad \text{for} \quad n = 0,1,2,3,\ldots,T-1 \quad \text{and} \quad \phi_T\left(\{\omega_0^{(T)}\}\right) = \pi^T.$$

The distribution ϕ_T satisfies the recursion

$$\phi_{T+1}\left(\{\omega_0)\}\right) = (1-\pi)\sum\nolimits_{n=0}^{T} \phi_T\left(\{(\omega_0^{(n)}\}\right) \quad \text{and}$$
$$\phi_{T+1}\left(\{\omega_0^{(n+1)}\}\right) = \pi\phi_T\left(\{\omega_0^{(n)}\}\right) \quad \text{for} \quad n = 0,1,2,3,\ldots,T.$$

In the limit, ϕ_T converges to the invariant distribution $\phi(\{\omega_0^{(n)}\}) = (1-\pi)\pi^n$ for $n = 0,1,\ldots,\infty$, which is a simple geometric distribution. The invariant distribution $\varphi(x)$ is easily calculated from ϕ because $\varphi(\omega,s) = \pi(s)\phi(\omega)$.

(ii) Mean entropy is computed from the conditional entropy given in Proposition 6 and the invariant distribution derived in part (i).

(iii) From the first-order condition (13), $\log(g(\omega, 2)) = \log(g(\omega, 1)) + \log(\upsilon(\omega))$ where $g(\omega, s) := u_c(c^y(\omega, s))/u_c(e(s) - c^y(\omega, s))$ and $\upsilon(\omega) := 1 + \mu(\omega, 1)$. The two variables $\log(g(\omega, 1))$ and $\log(\upsilon(\omega))$ are co-monotonic increasing in ω. Therefore, it follows by applying Chebyshev's order inequality that their covariance is positive. Computing the variance at the invariant distribution gives:

$$\mathrm{var}(\log(g(\omega, 2))) = \mathrm{var}(\log(g(\omega, 1))) + \mathrm{var}(\log(\upsilon(\omega))) + \mathrm{cov}(\log(g(\omega, 1)), \log(\upsilon(\omega)))$$

and hence, $\mathrm{var}(\log(g(\omega, 1))) > \mathrm{var}(\log(g(\omega, 1)))$.

(iv) Using the invariant distribution, the conditional expected promised utility for $t+1$ is $\mathbb{E}[\omega_{t+1} \mid \omega_t = \omega_0^{(n)}] = (1-\pi)\omega_0 + \pi\omega_0^{(n+1)}$. Since $\omega_0^{(n)}$ is monotonically increasing in n, so too is the above conditional expectation. Thus, ω_t and $\mathbb{E}[\omega_{t+1} \mid \omega_t]$) are co-monotonic. Since $\mathrm{cov}(\omega_t, \omega_{t+1}) = \mathrm{cov}(\omega_t, \mathbb{E}[\omega_{t+1} \mid \omega_t])$, it follows that $\mathrm{cov}(\omega_t, \omega_{t+1}) > 0$.

(v) The argument of part (iv) can be applied to the conditional auto-covariance. Consumption $c^y(\omega, s)$ is decreasing in ω from Proposition 6. The expectation of the consumption of the young next period conditional on the current endowment state is:

$$\mathbb{E}[c^y_{t+1} \mid c^y_t = c^y(\omega, 1)] = \pi c^y(f(\omega, 1), 1) + (1-\pi) c^y(f(\omega, 1), 2),$$
$$\mathbb{E}[c^y_{t+1} \mid c^y_t = c^y(\omega, 2)] = \pi c^y(\omega_0, 1) + (1-\pi) c^y(\omega_0, 2).$$

Since the first expectation is decreasing in ω, $\mathrm{cov}(c^y_t, c^y_{t+1} \mid s_t = 1) > 0$. The second expectation is independent of ω and hence, $\mathrm{cov}(c^y_t, c^y_{t+1} \mid s_t = 2) = 0$. ■

Supplementary Appendix

These appendices present supplementary material referenced in the paper. Part A provides evidence from the Luxembourg Income Study Database 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 1 and 3 from Sections 1 and 2 together with the proof of Lemma 1 from Section 3. Part C supplements Section 7 and provides an illustration of the comparative statics for the insurance coefficient and consumption-equivalent welfare change measures. Part D derives results stated in Section 5. Part E presents the shooting algorithm used to derive the optimal allocation in Section 6. Part F describes the pseudo-code for the numerical algorithms used in the paper.

A Change in Relative Income of Young and Old

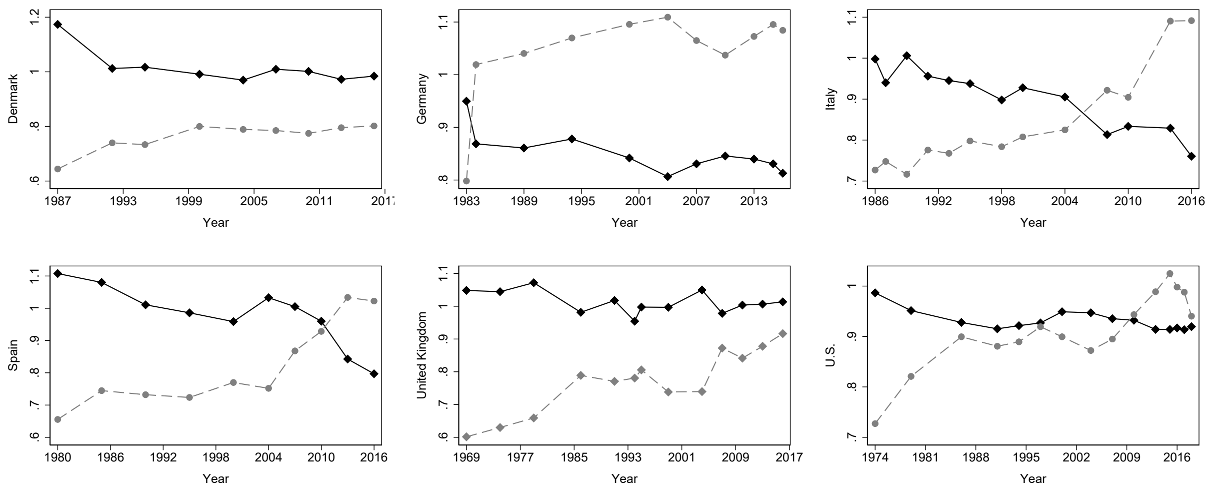


Figure A.1: Relative Income of Young and Old for six OECD Countries

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 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 Proof of Propositions 1 and 3 and Lemma 1

Proof of Proposition 1. The lifetime endowment utility of an agent born in state r is:

$$\hat{v}(r) := u(e^y(r)) + \beta \sum_s \pi(s) u(e^o(s)).$$

Consider a small transfer $\mathrm{d}\tau(r)$ in state r 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

$$-u_c(e^y(r))\mathrm{d}\tau(r) + \beta \sum_s \pi(s) u_c(e^o(s))\mathrm{d}\tau(s) \geq 0. \tag{B.1}$$

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

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

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 $\hat{q}(r,s) = \pi(s)\hat{m}(r,s)$. 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 3. Define the critical transfer τ_1^c by:

$$u(e^y - \tau_1^c) + \beta u(e^o + \tau^*) = \hat{v} := u(e^y) + \beta u(e^o).$$

Define τ_j^c recursively by:

$$u(e^y - \tau_j^c) + \beta u(e^o + \tau_{j-1}^c) = \hat{v} \quad \text{for } j = 1, 2, \ldots, \infty.$$

From the strict concavity of utility function u, $\tau_j^c > \tau_{j-1}^c$ and $\lim_{j\to\infty} \tau_j^c = \bar{\tau} = e^y - e^o$. Correspondingly, define $\omega_j^c := u(e^o + \tau_j^c)$. We have $\omega_0^c = \omega^*$ and $\lim_{j\to\infty} \omega_j^c = \bar{\omega}$. Let $v^* = u(e^y + \tau^*) + (\beta/\delta)\omega^*$. With some abuse of notation, write $V_n(\omega)$ to denote the value function when $\omega \in (\omega_{n-1}^c, \omega_n^c]$. We have:

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

For $\omega \leq \omega^*$, $\tau(\omega) = \tau^*$ and $\omega' = \omega^*$. Therefore, $V(\omega) = v^*/(1-\delta)$ for $\omega \in [u(e^o), \omega^*]$. For $\omega \in (\omega^*, \omega_1^c]$,

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

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

$$\frac{\mathrm{d}V_1(\omega)}{\mathrm{d}\omega} = \frac{\beta}{\delta} - \frac{u_c(e - u^{-1}(\omega))}{u_c(u^{-1}(\omega))}.$$

Let $g(\omega) := u_c(e - u^{-1}(\omega))/u_c(u^{-1}(\omega))$. Since $\omega > \omega^*$, $g(\omega) > \beta/\delta$ and $\mathrm{d}V_1(\omega)/\mathrm{d}\omega < 0$. Note that $g(\omega^*) = \beta/\delta$ 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 $g(\omega)$ is increasing given the strict concavity of u. Having established that $V_1(\omega)$ is decreasing and strictly concave, 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. It is easy to establish continuity by showing that $\lim_{\omega\to\omega_n^c} V_{n+1}(\omega) = V_n(\omega_n^c)$. 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 $\omega \in (\omega_n^c, \omega_{n+1}^c)$:

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

Starting with $n = 1$, we have:

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

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

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

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:

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

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

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

Since $g(\omega_j^c) \in [(\beta/\delta), \lambda^{-1})$, taking the limit as $n \to \infty$, or equivalently, $\omega \to \bar{\omega}$, gives $\lim_{\omega\to\bar{\omega}} \mathrm{d}V(\omega)/\mathrm{d}\omega = -\infty$. ■

Proof of Lemma 1. We establish the domain and concavity and differentiability properties of the value function $V(\omega)$.

Domain: Since $\tau(s) \geq 0$ for all $s \in \mathcal{S}$, $\omega \geq \omega_{\min} := \sum_s \pi(s)u(e^o(s))$. The largest feasible ω, $\omega_{\max}$, can be found by solving the problem of choosing $(\tau(s), \omega'(s))$ to maximize $\sum_s \pi(s)u(e^o(s)+\tau(s))$ subject to $\tau(s) \geq 0$ and constraints (9) and (10). This is a strictly concave programming problem and the objective and constraint functions are continuous. Thus, there exists a unique solution. The constraint set is non-empty by Proposition 1. All constraints in (9) bind at the solution: if one of these constraints did not bind, say in state r, then it would be possible to increase the maximand by increasing $\tau(r)$ without violating the other constraints. Equally, it is desirable to choose $\omega'(s)$ as large as possible because an increase in $\omega'(s)$ allows $\tau(s)$ to be increased without violating constraint (9), increasing the maximand. Thus, the solution involves $\omega'(s) = \omega_{\max}$ for each s. Let $\tau^\sharp(s)$ denote the solution for the transfer and define $\omega^\sharp := \sum_s \pi(s)u(e^o(s) + \tau^\sharp(s))$. Since constraint (9) binds for each s,

$$\tau^\sharp(s) = e^y(s) - u^{-1}(u(e^y(s)) - \beta(\omega_{\max} - \omega_{\min})).$$

By definition $\omega_{\max} = \omega^\sharp$. Thus, $\omega_{\max}$ can be found as the root of

$$\sum\nolimits_s \pi(s) \left(u(e(s) - u^{-1}(u(e^y(s)) - \beta(\omega - \omega_{\min})))\right) - \omega.$$

We next show that the root $\omega_{\max} \in (\omega_{\min}, \sum_s \pi(s)u(e(s)))$. To see this, first note that $\omega_{\max} > \omega_{\min}$ by Proposition 1. For all $\omega > \omega_{\min}$, $\tau^\sharp(s) > 0$. Secondly, suppose that $\tau^\sharp(r) = e^y(r)$ for some state r. Then, $u(e^y(r)) - \beta(\omega - \omega_{\min}) = u(0)$ or

$$u(0) \geq u(e^y(r)) - \beta \sum\nolimits_s \pi(s) \left(u(e(s)) - u(e^o(s))\right),$$

which provides a contradiction since it violates Assumption 1. Hence, $\tau^\sharp(r) < e^y(r)$ for all r and consequently, $\omega_{\max} < \sum_s \pi(s)u(e(s))$.

Concavity: We first show that $V(\omega)$ is concave. Consider the mapping T defined by

$$(TJ)(\omega) = \max_{\{c^y(s),\omega'(s)\}\in\Phi} \left[\sum\nolimits_s \pi(s) \left(\tfrac{\beta}{\delta} u(e(s) - c^y(s)) + u(c^y(s)) + \delta J(\omega'(s))\right)\right].$$

Consider $J = V^*$, the first-best frontier. Proposition 2 established that $V^*(\omega)$ is concave. It follows from the definitions of T and V^* that $TV^*(\omega) \leq V^*(\omega)$ because $V^*(\omega) \leq v^*/(1-\delta)$ and the mapping T adds the participation constraints (9). That is, $T^nV^*(\omega) \leq T^{n-1}V^*(\omega)$ for $n = 1$. Now, make the induction hypothesis that $T^nV^*(\omega) \leq T^{n-1}V^*(\omega)$ for $n \geq 2$ and apply the mapping T to the two functions $T^nV^*(\omega)$ and $T^{n-1}V^*(\omega)$. It is straightforward to show that $T^{n+1}V^*(\omega) \leq T^nV^*(\omega)$, 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 $T^nV^*(\omega)$ is non-increasing and converges. Let $V^\infty(\omega) = \lim_{n\to\infty} T^nV^*(\omega)$, 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 u are concave. The constraint set Φ is convex. Hence, $TV^*(\omega)$ is concave. By induction, $T^nV^*(\omega)$ and the limit function V are also concave.

Differentiability: There are $2S$ choice variables and $3S + 1$ constraints, including the non-negativity constraints on transfers. Without differentiability of the value function

V, the first-order condition (14) is replaced by

$$\partial V(f(\omega, s)) \ni -\tfrac{\beta}{\delta}\left(\mu(\omega, s) - \zeta(\omega, s)\right),$$

where $\partial V(\omega)$ denotes the set of superdifferentials of V at ω. 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. We first note that the participation constraints of the young and the old cannot bind simultaneously in a given endowment state. If $\eta(\omega, s) > 0$, that is, the participation constraint of the old binds, then both the young and the old consume their endowments. Since the sustainable intergenerational insurance is non-autarkic, the current young receive a transfer in some endowment state when they are old and hence, they cannot be constrained in the current period, so that $\mu(\omega, s) = 0$. Likewise, if $\mu(\omega, s) > 0$, that is, the participation constraint of the young binds, then the current young are making a transfer and hence, the current old are unconstrained, so that $\eta(\omega, s) = 0$. Similarly, for $\omega < \omega_{\max}$, it is easily checked that not all upper bound constraints can bind for all states. Thus, for $\omega < \omega_{\max}$, there can be at most $2S$ binding constraints. Since utility is strictly increasing, $\beta > 0$, and $\pi(s) > 0$ for each s, it can be checked that the matrix of binding constraints has full rank. Hence, the multipliers are unique and $V(\omega)$ is differentiable on the interior of Ω. Since $V(\omega)$ is concave and differentiable, it is also continuously differentiable. It follows from the envelope condition (15) that $V_\omega(\omega_0) = 0$. Since the promise-keeping constraint (11) is an inequality, it is easily checked that $V(\omega)$ is non-increasing. The multiplier $\nu(\omega) = 0$ for $\omega < \omega_0$ and is increasing in ω for $\omega > \omega_0$. Let $\bar{\nu} := \lim_{\omega \to \omega_{\max}} \nu(\omega)$, then $\lim_{\omega \to \omega_{\max}} V_\omega(\omega) = -(\beta/\delta)\bar{\nu}$, where $\bar{\nu} \in \mathbb{R}_+ \cup \{\infty\}$.

Interiority $\omega_0 \in (\omega_{\min}, \omega^*)$: Any non-autarkic sustainable intergenerational insurance involves some transfer from the young to the old. Thus, by Proposition 1, $\omega_0 > \omega_{\min}$. Since $V(\omega)$ is a concave Pareto frontier (that is, $V(\omega)$ is weakly decreasing and concave), it follows that $V(\omega)$ is constant for $\omega < \omega_0$ and that, by differentiability, $\nu(\omega_0) = 0$. Therefore, from the first-order condition (13), $\tau(\omega_0, s) \leq \tau^*(s)$, with strict equality if the participation constraint of the young is non-binding (that is, if $\mu(\omega_0, s) = 0$). Thus, the utility promised to the old is no greater than ω^*. By Assumption 4, the first best cannot be sustained. Hence, we conclude $\omega_0 < \omega^*$. ■

C Alternative Measures of Risk Sharing

In this Appendix, we compute the insurance coefficient and the consumption equivalent welfare change at the invariant distribution in the two-state example of Section 6. We do this for the range of parameter values considered in the comparative static exercise of Section 7.

The insurance coefficient $\iota(x)$ 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, it is defined conditional on state $x = (\omega, s)$ as follows:

$$\iota(x) = 1 - \frac{\operatorname{cov}\left(\log\left(c^y(f(x), s')\right), \log\left(e^y(s')\right)\right)}{\operatorname{var}\left(\log\left(e^y(s')\right)\right)}.$$

At the first best, and provided that the non-negativity constraint does not bind, consumption is independent of the endowment and the insurance coefficient is one. The insurance coefficient increases as more risk is shared. With two states, the insurance coefficient can be rewritten as:

$$\iota(x) = 1 - \frac{\log\left(c^y(f(x), 1)\right) - \log\left(c^y(f(x), 2)\right)}{\log\left(e^y(1)\right) - \log\left(e^y(2)\right)}.$$

The top row of Figure C.1 plots the average insurance coefficient evaluated at the invariant distribution of x for the three parameters κ, σ and δ when $\beta = \delta$.

We measure the consumption equivalent welfare change relative to the first best for a given ω by solving the following equation in terms of ε:

$$\frac{1}{1-\delta}\left(\mathbb{E}_s[u(c^{y*}(s)(1-\varepsilon))] + \frac{\beta}{\delta}\mathbb{E}_s[u((e(s) - c^{y*}(s))(1-\varepsilon))]\right) = V(\omega).$$

The solution $\varepsilon(\omega)$ measures the proportion by which the first-best consumption needs to be reduced to match the optimal solution for each ω. The consumption equivalent welfare change is smaller when more risk is shared. The long-run welfare loss measure is the average of $\varepsilon(\omega)$ at the invariant distribution of ω. The bottom row of Figure C.1 plots the consumption equivalent welfare change for the three parameters κ, σ and δ when $\beta = \delta$.

Comparing Figure C.1 with the mean entropy measure illustrated in top row of Figure 7, it can be seen that broadly similar conclusions are obtained using mean entropy, the average insurance coefficient or the average consumption equivalent welfare change. The amount of risk shared at the optimal sustainable intergenerational insurances increases with κ and δ but falls with σ.

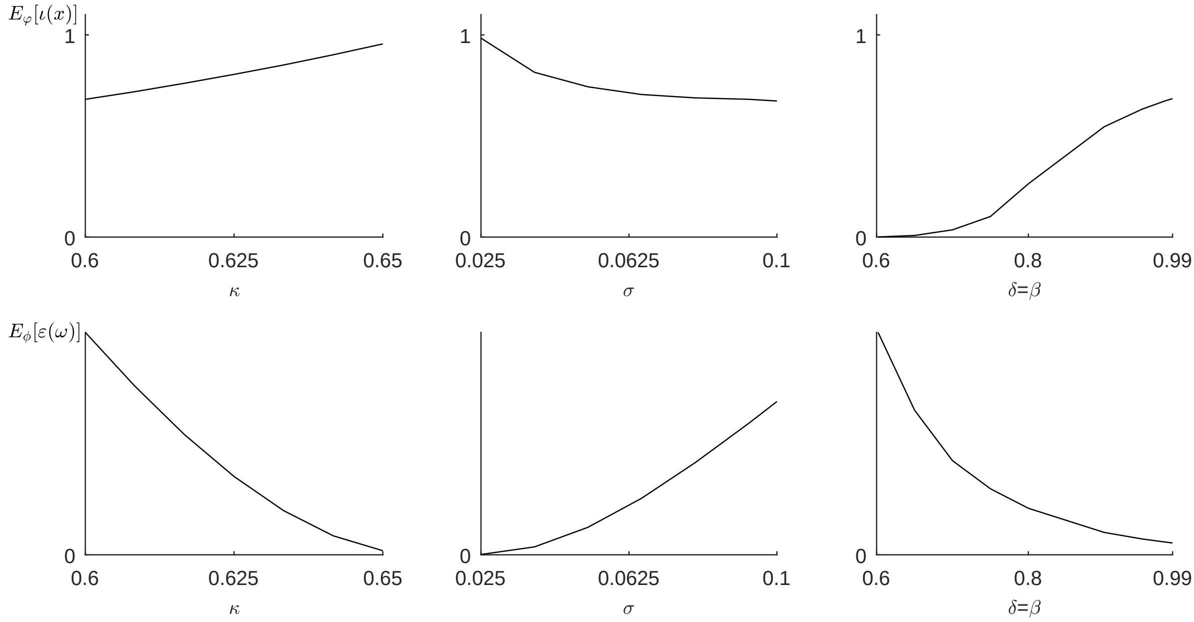


Figure C.1: Insurance Coefficient and the Consumption Equivalent Welfare Change
Note: The top row illustrates the average insurance coefficient $\mathbb{E}_{\varphi}[\iota(x)]$. The bottom row illustrates the average consumption equivalent welfare change $\mathbb{E}_{\phi}[\varepsilon(\omega)]$.

D Derivation of Results of Section 5

Kullback-Leibler divergence The Kullback-Leibler divergence (hereafter, KL) measures the divergence between the corresponding rows of a stochastic matrix Π and a non-negative irreducible matrix Q. The two matrices are compatible if the element $\pi(x, x') = 0$ whenever $q(x, x') = 0$. Let $\Pi(x)$ and $Q(x)$ denote the rows of our transition matrix and state price matrix that correspond to state x. Then, the Kullback-Leibler divergence is:

$$\mathrm{KL}\left(\Pi(x)\|Q(x)\right) = -\sum\nolimits_{x'} \pi(x, x') \log\left(\frac{q(x, x')}{\pi(x, x')}\right).$$

This divergence is zero if and only if the rows are identical. By the log sum inequality, $\mathrm{KL}(\Pi(x)\|Q(x)) \geq y^1(x)$, but $y^1(x)$ could be negative if the row sum of Q corresponding to state x is greater than one. However, there is a lower bound on that depends on the Perron root ρ of Q and the left eigenvector φ of Π. Define the average divergence as $\sum_x \varphi(x)\mathrm{KL}(\Pi(x)\|Q(x))$. Then,

$$\sum\nolimits_x \varphi(x)\mathrm{KL}\left(\Pi(x)\|Q(x)\right) \geq -\log(\rho).$$

Moreover, the bound is attained when

$$m(x,x') := \frac{q(x,x')}{\pi(x,x')} = \rho\frac{\psi(x)}{\psi(x')},$$

where ψ is the right eigenvector of Q corresponding to the eigenvalue ρ, that is when the Ross Recovery Theorem holds. To see this lower bound note that the average divergence can be rewritten as:

$$\sum\nolimits_x \varphi(x)\mathrm{KL}\left(\Pi(x)\|Q(x)\right) = -\sum\nolimits_x \varphi(x)\sum\nolimits_{x'} \pi(x,x')\log\left(\frac{q(x,x')}{\pi(x,x')}\right).$$

Moreover, note that for any probability vector $\tilde{\psi}(x)$:

$$\begin{aligned}\sum\nolimits_x \varphi(x)\sum\nolimits_{x'} \pi(x,x')\log\left(\frac{\tilde{\psi}(x')}{\tilde{\psi}(x)}\right) &= \sum\nolimits_x \psi(x)\sum\nolimits_{x'} \pi(x,x')\log\left(\tilde{\psi}(x')\right)\\ &\quad -\sum\nolimits_x \varphi(x)\sum\nolimits_{x'} \pi(x,x')\log\left(\tilde{\psi}(x)\right)\\ &= \sum\nolimits_{x'} \log\left(\tilde{\psi}(x')\right)\sum\nolimits_x \varphi(x)\pi(x,x')\\ &\quad -\sum\nolimits_x \varphi(x)\log\left(\tilde{\psi}(x)\right)\sum\nolimits_{x'} \pi(x,x') = 0,\end{aligned} \tag{D.1}$$

where the last line follows because $\sum_{x'}\pi(x,x') = 1$ and $\sum_x \varphi(x)\pi(x,x') = \varphi(x')$. Hence, the left hand side of equation (D.1) is independent of $\tilde{\psi}$. Let D_ψ denote the diagonal matrix with the eigenvector ψ on the diagonal. It therefore follows from equation (D.1) that the average divergence can be rewritten as:

$$\begin{aligned}\sum\nolimits_x \varphi(x)\mathrm{KL}\left(\Pi(x)\|Q(x)\right) &= \sum\nolimits_x \varphi(x)\mathrm{KL}\left(\Pi(x)\|IQI^{-1}(x)\right)\\ &= \sum\nolimits_x \varphi(x)\mathrm{KL}\left(\Pi(x)\|\rho^{-1}D_\psi Q D_\psi^{-1}(x)\right) - \log(\rho).\end{aligned}$$

Since the matrix $\rho^{-1}D_\psi Q D_\psi^{-1}$ is stochastic and $\varphi > 0$,

$$\sum\nolimits_x \varphi(x)\mathrm{KL}\left(\Pi(x)\|\rho^{-1}D_\psi Q D_\psi^{-1}(x)\right) \geq 0,$$

with equality if $\rho^{-1}D_\psi Q D_\psi^{-1} = \Pi$.

Conditional entropy Let

$$\varrho(x,x') = \frac{q(x,x')}{\sum_{x'} q(x,x')} \quad \text{and} \quad p(x) = \sum\nolimits_{x'} q(x,x'),$$

denote the risk-neutral probability of state x' when the current state is x and the price of the one period risk-free bond in state x. Let Γ denote the matrix of risk-neutral probabilities. Conditional entropy is defined by:

$$L(x) := \mathrm{KL}(\Pi(x)\|\Gamma(x)) = -\sum\nolimits_{x'} \pi(x,x')\log\left(\frac{\varrho(x,x')}{\pi(x,x')}\right).$$

Let $y(x) = -\log(p(x))$ denote the yield on the one period bond. Then,

$$L(x) = \mathrm{KL}(\Pi(x)\|\Gamma(x)) = \mathrm{KL}\left(\Pi(x)\|Q(x)\right) - y(x).$$

Since $q(x,x') = \pi(x,x')m(x,x')$, we can also write:

$$L(x) = \log\left(\sum\nolimits_{x'} \pi(x,x')m(x,x')\right) - \sum\nolimits_{x'} \pi(x,x')\log\left(m(x,x')\right).$$

Mean entropy Let $\varphi(x)$ denote the stationary probability of state i. That is φ is the left eigenvector of Π. The mean entropy is:

$$\bar{L} = \sum\nolimits_x \varphi(x)L(x).$$

Using the Ross Recovery Theorem,

$$q(x,x') = \pi(x,x')\rho\frac{\psi(x)}{\psi(x')},$$

where ρ is Perron root of Q and ψ is the corresponding right eigenvector. That is, $Q = \rho D_\psi \Pi D_\psi^{-1}$ or $\Pi = \rho^{-1}D_\psi^{-1}QD_\psi$. Then, the bound described above is attained, and:

$$\begin{aligned}\sum\nolimits_x \varphi(x)\mathrm{KL}\left(\Pi(x)\|Q(x)\right) &= -\sum\nolimits_x \varphi(x)\sum\nolimits_{x'} \pi(x,x')\log\left(\frac{q(x,x')}{\pi(x,x')}\right)\\ &= -\log(\rho) - \sum\nolimits_{x'}\left(\varphi(x') - \sum\nolimits_x \varphi(x)\pi(x,x')\right)\log\left(\psi(x')\right)\\ &= -\log(\rho),\end{aligned}$$

where the last line follows because $\sum_x \varphi(x)\pi(x,x') = \varphi(x')$. Hence,

$$\bar{L} = \sum\nolimits_x \varphi(x)L(x) = -\log(\rho) - \sum\nolimits_x \varphi(x)y(x).$$

Since $y^\infty = -\log(\rho)$ and letting $\bar{y} = \sum_x \varphi(x)y(x)$ denote the average yield, we have:

$$\bar{L} = \sum\nolimits_x \varphi(x)L(x) = y^\infty - \bar{y}.$$

Note that we can also write:

$$\bar{L} = \sum\nolimits_x \varphi(x)\mathrm{KL}\left(\Pi(x)\|\Gamma(x)\right) = \sum\nolimits_x \varphi(x)\left(\mathrm{KL}\left(\Pi(x)\|Q(x)\right) - y(x)\right),$$

which can be rewritten using $\bar{L} = y^{\infty} - \bar{y}$ as:

$$-\log(\rho) = \sum\nolimits_x \varphi(x)\left(y(x) + \mathrm{KL}\left(\Pi(x)\|\Gamma(x)\right)\right) = \bar{y} + \bar{L}.$$

Martin-Ross measure Let $\psi_{\max} = \max_x \psi(x)$ and $\psi_{\min} = \min_x \psi(x)$. Following Martin and Ross (2019) define the Martin-Ross measure:

$$\Upsilon := \log\left(\frac{\psi_{\max}}{\psi_{\min}}\right).$$

It follows from the Ross Recovery Theorem that for each pair (x, x'):

$$\log\left(m(x, x')\right) - \log(\rho) = \log\left(\frac{\psi(x)}{\psi(x')}\right),$$

and hence, using the definitions of $\psi_{\max}$ and $\psi_{\min}$,

$$-\Upsilon \le \log\left(m(x, x')\right) - \log(\rho) \le \Upsilon.$$

Since ψ is the corresponding eigenvector, we have the following two sets of inequalities:

$$\rho\psi_{\min} \le \rho\psi(x) = \sum\nolimits_{x'} q(x, x')\psi(x') \le \sum\nolimits_{x'} q(x, x')\psi_{\max} = p(x)\psi_{\max},$$
$$\rho\psi_{\max} \ge \rho\psi(x) = \sum\nolimits_{x'} q(x, x')\psi(x') \ge \sum\nolimits_{x'} q(x, x')\psi_{\min} = p(x)\psi_{\min}.$$

Taking logs and using $\log(\rho) = -y^{\infty}$, $|y(x) - y^{\infty}| \le \Upsilon$. Since $\bar{L} = y^{\infty} - \bar{y}$,

$$\bar{L} - \Upsilon \le y(x) - \bar{y} \le \bar{L} + \Upsilon.$$

E Shooting Algorithm

In the two-state economy in Section 6, the multiplier on the participation constraint in state 2 satisfies $\mu(\omega, 2) = 0$ for all $\omega \in [\omega_0, \bar{\omega}]$. Therefore, write $\upsilon(\omega) := 1 + \mu(\omega, 1)$. At the invariant distribution, write $\upsilon^{(n)} = \upsilon(\omega^{(n)})$. Using the updating property of equation (16), $\nu(\omega^{(n+1)}) = \mu(\omega^{(n)})$ and equation (13) can be written as:

$$\begin{aligned}\frac{e(1) - c^y(\omega^{(n)}, 1)}{c^y(\omega^{(n)}, 1)} &= \frac{\beta}{\delta}\left(\frac{\upsilon^{(n-1)}}{\upsilon^{(n)}}\right), \\ \frac{e(2) - c^y(\omega^{(n)}, 2)}{c^y(\omega^{(n)}, 2)} &= \frac{\beta}{\delta}\upsilon^{(n-1)}.\end{aligned}$$

Given Assumptions 3 and 4, the participation constraint of the young in state 1 and the promise-keeping constraint are binding. That is,

$$\pi \log\left(\tfrac{\beta\upsilon^{(n-1)}e(1)}{\beta\upsilon^{(n-1)}+\delta\upsilon^{(n)}}\right) + (1-\pi)\log\left(\tfrac{\beta\upsilon^{(n-1)}e(2)}{\beta\upsilon^{(n-1)}+\delta}\right) = \omega^{(n)}, \tag{E.1}$$

$$\log\left(\tfrac{\delta\upsilon^{(n)}e(1)}{\beta\upsilon^{(n-1)}+\delta\upsilon^{(n)}}\right) + \beta\omega^{(n+1)} = \log(e^y(1)) + \beta\omega_{\min}, \tag{E.2}$$

for $n \geq 0$ where $\upsilon^{(-1)} = 1$. For $n = 0$,

$$\pi \log\left(\tfrac{\beta e(1)}{\beta+\delta\upsilon^{(0)}}\right) + (1-\pi)\log\left(\tfrac{\beta e(2)}{\beta+\delta}\right) = \omega_0,$$

while for n that tends to infinity,

$$\pi \log\left(\tfrac{\beta e(1)}{\beta+\delta}\right) + (1-\pi)\log\left(\tfrac{\beta\upsilon^{(\infty)}e(2)}{\beta\upsilon^{(\infty)}+\delta}\right) = \bar{\omega}, \tag{E.3}$$

where $\upsilon^{(\infty)} = \lim_{n\to\infty}\upsilon^{(n)}$ and

$$\bar{\omega} = \frac{1}{\beta}\left(\log(e^y\,(1)) - \log\left(\tfrac{\beta e(1)}{\beta+\delta}\right)\right) + \pi\log(e^o\,(1)) + (1-\pi)\log(e^o\,(2)). \tag{E.4}$$

Substituting equation (E.4) into (E.3), we have:

$$\upsilon^{(\infty)} = \frac{\delta}{\beta}\left(-1+\left(\left(\tfrac{\delta}{\beta}\right)^{\frac{\pi}{1-\pi}}\left(\tfrac{\beta+\delta}{\delta}\right)^{\frac{1+\beta\pi}{\beta(1-\pi)}}\left(\tfrac{e^y(1)}{e(1)}\right)^{\frac{1}{\beta(1-\pi)}}\left(\tfrac{e^o(1)}{e(1)}\right)^{\frac{\pi}{1-\pi}}\tfrac{e^o(2)}{e(2)}\right)^{-1}\right)^{-1}. \tag{E.5}$$

Using the equations (E.1) and (E.2), we can derive a second-order difference equation for $\upsilon^{(n)}$ where

$$\upsilon^{(n+1)} = \tfrac{\beta}{\delta}\upsilon^{(n)}\left(-1+\left(\tfrac{\beta\upsilon^{(n)}}{\beta\upsilon^{(n)}+\delta}\right)^{\frac{1-\pi}{\pi}}\left(\tfrac{\beta\upsilon^{(\infty)}+\delta}{\beta\upsilon^{(\infty)}}\right)^{\frac{1-\pi}{\pi}}\left(\tfrac{\beta+\delta}{\delta}\right)^{\frac{1}{\beta\pi}}\left(\tfrac{\beta+\delta}{\beta}\right)\left(1+\tfrac{\beta}{\delta}\tfrac{\upsilon^{(n-1)}}{\upsilon^{(n)}}\right)^{-\frac{1}{\beta\pi}}\right). \tag{E.6}$$

It can be shown that the second-order difference equation in (E.6) has a unique saddle path solution. Recalling that $\upsilon^{(-1)} = 1$, the solution can be found by a *forward shooting* algorithm to search for an $\upsilon^{(0)}$ such that the absolute difference between $\upsilon^{(\infty)}$ (given in (E.5)) and $\upsilon^{(N+1)}$ (given in (E.6)) is sufficiently close to zero for N sufficiently large.

F Pseudo-code for Numerical Algorithms

Algorithm 1: Shooting Algorithm

procedure ▷ Find $\upsilon^{(0)} = 1 + \mu^{(0)}$ in two state economy (Section 6)
  target ← $\upsilon^{(\infty)}$ ▷ Use equation (E.5) in Appendix E
  tolerance ← $\epsilon > 0$ ▷ $\epsilon = 10^{-10}$
  repeat
    initialization ← $\upsilon_0^{(0)} > 0$
    Compute $\upsilon_0^{(N)}$ for $N = 20$ ▷ Use equation (E.6) in Appendix E
    $d \leftarrow d(\upsilon_0^{(N)}, \upsilon^{(\infty)})$ ▷ $d(\upsilon_0^{(N)}, \upsilon^{(\infty)}) = |\upsilon_0^{(N)} - \upsilon^{(\infty)}|$
  until $d < \epsilon$
  $\upsilon^{(0)} \leftarrow \upsilon_0^{(0)}$
end procedure

Algorithm 2: Find Value and Policy Functions

procedure ▷ Find solution to functional equation (12)
  $\Omega \leftarrow [\omega_{\min}, \bar{\omega}]$ ▷ $\omega_{\min}$ and $\bar{\omega}$ computed
  gridpoints ← gp ▷ Discretize Ω: $gp = 200$ Chebyshev interpolation points
  tolerance ← $\epsilon > 0$ ▷ $\epsilon = 10^{-6}$
  $J \leftarrow V^*$ ▷ V^* is first best
  repeat
    Compute TJ from J ▷ Use equation (12) and interpolate
    $d \leftarrow d(TJ, J)$ ▷ $d(TJ, J) = \max_\omega |TJ(\omega) - J(\omega)|$
    $J \leftarrow TJ$
  until $d < \epsilon$
  $V \leftarrow J$
  Compute $f(\omega, s)$ and $c^y(\omega, s)$ ▷ Using the function V just computed.
end procedure

Algorithm 3: Computing the Invariant Distribution

procedure ▷ Find invariant distribution for $x = (\omega, s) \in X \subset \mathbb{R}^{nS\times 1}$
  initialization ← $a_0 = \mathbf{e}(1/nS)$ ▷ $\mathbf{e} = (1, 1, ..., 1) \in \mathbb{R}^{nS\times 1}$
  Compute $a = \Pi a_0$ ▷ Use the transition probability $\Pi \subset \mathbb{R}^{nS\times nS}$
  tolerance ← $\epsilon > 0$ ▷ $\epsilon = 10^{-8}$
  repeat
    Compute $a = \Pi a$ ▷ a is eigenvector associated with 1
    $d \leftarrow d(\Pi a, a)$ ▷ $d(\Pi a, a) = \max_x |\Pi a(x) - a(x)|$
    $a \leftarrow \Pi a$
  until $d < \epsilon$
  $\varphi \leftarrow a / \sum_x a(x)$ ▷ φ is normalized invariant distribution
end procedure

Algorithms are implemented in MATLAB®. At each iteration, the optimization uses the nonlinear programming solver command `fsolve` in Algorithm 1 and command `fmincon` in Algorithm 2. Interpolation of the value function uses the spline method embedded in the `interp1` command. In a typical example, the value function converges within 300 iterations. Each iteration takes about 1 second on a standard pc.