The Income Fluctuation Problem and the Evolution of Wealth[1]

Qingyin Ma[a], John Stachurski[b] and Alexis Akira Toda[c]

[a, b]Research School of Economics, Australian National University

[c]Department of Economics, University of California San Diego

November 17, 2019

Abstract. We analyze the household savings problem in a general setting where returns on assets, non-financial income and impatience are all state dependent and fluctuate over time. All three processes can be serially correlated and mutually dependent. Rewards can be bounded or unbounded and wealth can be arbitrarily large. Extending classic results from an earlier literature, we determine conditions under which (a) solutions exist, are unique and are globally computable, (b) the resulting wealth dynamics are stationary, ergodic and geometrically mixing, and (c) the wealth distribution has a Pareto tail. We show how these results can be used to extend recent studies of the wealth distribution. Our conditions have natural economic interpretations in terms of asymptotic growth rates for discounting and return on savings.

Keywords: Income fluctuation, optimality, stochastic stability, wealth distribution.

1. Introduction

It has been observed that, in the US and several other large economies, the wealth distribution is heavy tailed and wealth inequality has risen sharply over the last few decades.[2] This matters not only for its direct impact on taxation and redistribution

[1]This paper has benefitted from discussion with many colleagues. We particularly thank Fedor Iskhakov, Larry Liu and Chung Tran for valuable feedback and suggestions. The second author gratefully acknowledges financial support from ARC grant FT160100423.

Email addresses: `qingyin.ma@anu.edu.au`, `john.stachurski@anu.edu.au`, `atoda@ucsd.edu`.

[2]For example, in a study based on capital income data, Saez and Zucman (2016) find that, in the case of the US, the share of total household wealth held by the top 0.1% increased from 7 percent to 22 percent between 1978 and 2012. For a discussion of the heavy-tailed property of the wealth

policies, but also for potential flow-on effects for productivity growth, business cycles and fiscal policy, as well as for the political environment that shapes these and other economic outcomes.[3]

At present, our understanding of these phenomena is hampered by the fact that standard tools of analysis—such as those used for heterogeneous agent models—are not well adapted to studying the wealth distribution as it stands. For example, while we have sound understanding of the household problem when returns on savings and rates of time discount are constant (see, e.g., Schechtman (1976), Schechtman and Escudero (1977), Deaton and Laroque (1992), Carroll (1997), or Açıkgöz (2018)), our knowledge is far more limited in settings where these values are stochastic. This is problematic, since injecting such features into the household problem is essential for accurately representing the joint distribution of income and wealth (e.g., Benhabib et al. (2015), Benhabib et al. (2017), Stachurski and Toda (2019)).[4] Moreover, models with time-varying discount rates and returns on assets are at the forefront of recent quantitative analysis of wealth and inequality.[5]

While it might be hoped that the analysis of the income fluctuation problem (or household consumption and savings problem) changes little when we shift from constant to state dependent asset returns and rates of time discount, this turns out not to be the case. Effectively modeling these features and the way they map to the wealth distribution requires significant advances in our understanding of choice and stochastic dynamics in the setting of optimal savings.

distribution, see Pareto (1896), Davies and Shorrocks (2000), Benhabib and Bisin (2018), Vermeulen (2018) or references therein.

[3]One analysis of the two-way interactions between inequality and political decision making can be found in Acemoglu and Robinson (2002). Glaeser et al. (2003) show how inequality can alter economic and social outcomes through subversion of institutions. The same study contains references on linkages between inequality and growth. Regarding fiscal policy, Brinca et al. (2016) find strong correlations between wealth inequality and the magnitude of fiscal multipliers, while Bhandari et al. (2018) study the connection between fiscal-monetary policy, business cycles and inequality. Ahn et al. (2018) discuss the impact of distributional properties on macroeconomic aggregates.

[4]Also related is the recent experimental study of Epper et al. (2018), which finds a strong positive connection between dispersion in subjective rates of time discounting across the population and realized dispersion in the wealth distribution. This in turn is consistent with earlier empirical studies such as Lawrance (1991).

[5]For a recent quantitative study see, for example, Hubmer et al. (2018), where returns on savings and discount rates are both state dependent (as is labor income). Kaymak et al. (2018) find that asset return heterogeneity is required to match the upper tail of the wealth distribution.

One difficulty is that state-dependent discounting takes us beyond the bounds of traditional dynamic programming theory. This matters little if there exists some constant $\bar{\beta} < 1$ such that the discount process $\{\beta_t\}$ satisfies $\beta_t \leq \bar{\beta}$ for all t with probability one, since, in this case, a standard contraction mapping argument can still be applied (see, e.g., Miao (2006) or Cao (2018)). However, recent quantitative studies extend beyond such settings. For example, AR(1) specifications are increasingly common, in which case the support of β_t is unbounded above at every point in time.[6] Even if discretization is employed, the outcome $\beta_t \geq 1$ can occur with positive probability when the approximation is sufficiently fine. Moreover, such outcomes are not inconsistent with empirical and experimental evidence, at least for some households in some states of the world.[7] Do there exist conditions on $\{\beta_t\}$ that allow for $\beta_t \geq 1$ in some states and yet imply existence of optimal polices and practical computational techniques?

Another source of complexity for the income fluctuation problem in the general setting considered here is that the set of possible values for household assets is typically unbounded above. For example, when returns on assets are stochastic, a sufficiently long sequence of favorable returns can compound one another to project a household to arbitrarily high levels of wealth. This model feature is desirable: We wish to analyze these kinds of outcomes rather than rule them out. Indeed, Benhabib et al. (2015) and other related studies argue convincingly that such outcomes are a key causal mechanism behind the heavy tail of the current distribution of wealth.[8] However, if we accept this logic, then stationarity and ergodicity of the wealth process—which are fundamental both for estimation and for simulation-based numerical methods—must now be established in a setting where the wealth distribution has unbounded support. In such a scenario, what conditions on preferences and financial and labor income are necessary for these properties to hold?

A final and related example of the need for deeper analysis is as follows: To understand the upper tail of the wealth distribution, we must avoid unnecessarily truncating the

[6]See, for example, Hills and Nakata (2018), Hubmer et al. (2018) or Schorfheide et al. (2018).

[7]See, for example, Loewenstein and Prelec (1991) and Loewenstein and Sicherman (1991).

[8]One related study is Benhabib et al. (2011), who show that capital income risk is the driving force of the heavy-tail properties of the stationary wealth distribution. In Blanchard-Yaari style economies, Toda (2014), Toda and Walsh (2015) and Benhabib et al. (2016) show that idiosyncratic investment risk generates a double Pareto stationary wealth distribution. Gabaix et al. (2016) point out that a positive correlation of returns with wealth ("scale dependence") in addition to persistent heterogeneity in returns ("type dependence") can well explain the speed of changes in the tail inequality observed in the data.

upper tail of the set of possible asset values in quantitative work. While truncation is convenient because finite or compact state spaces are easier to handle computationally, we can attain greater accuracy in modeling the wealth distribution if truncation at the upper tail can be replaced locally by a parameterized savings function, such as a linear function (Gouin-Bonenfant and Toda, 2018). However, any such approximation must be justified by theory. What conditions can be imposed on primitives to generate such properties while still maintaining realistic assumptions for asset returns and non-financial income?

In this paper we address all of these questions, along with other key properties of the income fluctuation problem, such as continuity and monotonicity of the optimal consumption policy. Our setting admits capital income risk, labor earnings shocks and time-varying discount rates, driven by a combination of IID innovations and an exogenous Markov chain $\{Z_t\}$. The supports of the innovations can be unbounded, so we admit practical innovation sequences such as normal and lognormal. As a whole, this environment allows for a range of realistic features, such as stochastic volatility in returns on asset holdings, or correlation in the shocks impacting asset returns and non-financial income. The utility function can be unbounded both above and below, with no specific structure imposed beyond differentiability, concavity and the usual slope (Inada) conditions.[9]

To begin, when considering optimality in the household problem, we require a condition on the state dependent discount process $\{\beta_t\}$ that generalizes the classical condition $\beta < 1$ from the constant case and, for reasons discussed above, permits $\beta_t > 1$ with positive probability. To this end, we introduce the restriction[10]

$$G_\beta < 1 \quad \text{where} \quad G_\beta := \lim_{n \to \infty} \left(\mathbb{E} \prod_{t=1}^{n} \beta_t \right)^{1/n}. \tag{1}$$

Condition (1) clearly generalizes the classical condition $\beta < 1$ for the constant discount case. In the stochastic case, $\ln G_\beta$ can be understood as the asymptotic growth rate of the probability weighted average discount factor. Indeed, if $B_n := \mathbb{E} \prod_{t=1}^{n} \beta_t$ is the average n-period discount factor, then, from the definition of G_β and some straightforward analysis, we obtain $\ln(B_{n+1}/B_n) \to \ln G_\beta$, so the condition $G_\beta < 1$

[9]While the assumption that the exogenous state process $\{Z_t\}$ is a (finite state) Markov chain might appear restrictive, it fits most practical settings and avoids a host of technical issues that tend to obscure the key ideas. Moreover, the innovation shocks are not restricted to be discrete, and the same is true for assets and consumption.

[10]Here and below we set $\beta_0 \equiv 1$, so $\prod_{t=1}^{n} \beta_t = \prod_{t=0}^{n} \beta_t$.

implies that the asymptotic growth rate of the average n-period discount factor is negative, drifting down from its initial condition $\beta_0 \equiv 1$ at the rate $\ln G_\beta$. This does not, of course, preclude the possibility that $\beta_t > 1$ at any given t.

We show that condition (1) is in fact a necessary condition in those settings where the classical condition is necessary for finite lifetime values. In this sense it cannot be further weakened for the income fluctuation problem apart from special cases. At the same time, it admits the use of convenient specifications such as the discretized AR(1) process from Hubmer et al. (2018). In addition, we prove that G_β can be represented as the spectral radius of a positive matrix, and hence can be computed by numerical linear algebra (as discussed below).

We also generalize the standard condition $\beta R < 1$, where R is the gross interest rate in the constant case, which is used to ensure stability of the asset path and finiteness of lifetime valuations, as well as existence of stationary Markov policies (see, e.g., Deaton and Laroque (1992), Chamberlain and Wilson (2000) or Li and Stachurski (2014)). Analogous to (1), we introduce the generalized condition

$$G_{\beta R} < 1 \quad \text{where} \quad G_{\beta R} := \lim_{n \to \infty} \left(\mathbb{E} \prod_{t=1}^{n} \beta_t R_t \right)^{1/n}. \tag{2}$$

Here $\{R_t\}$ is a stochastic capital income process. Analogous to the case of G_β, the value $\ln G_{\beta R}$ can be understood as the asymptotic growth rate of average gross payoff on assets, discounted to present value.

We show that, when Conditions (1)–(2) hold and non-financial income satisfies two moment conditions, a unique optimal consumption policy exists. We also show that the policy can be computed by successive approximations and analyze its properties, such as monotonicity and asymptotic linearity. This asymptotic linearity can be used to successfully model wealth inequality by accurately representing asset path dynamics for very high wealth households (Gouin-Bonenfant and Toda, 2018).

One important feature of Conditions (1)–(2) is that they take into account the autocorrelation structure of preference shocks and asset returns. For example, if these processes depend only on IID innovations, then (1) reduces to $\mathbb{E}\beta_t < 1$ and (2) reduces to $\mathbb{E}\beta_t R_t < 1$. But returns on assets are typically not IID, since both mean returns and volatility are, in general, time varying, and preference shocks are typically modeled as correlated (see, e.g., Hubmer et al. (2018) or Schorfheide et al. (2018)). This dependence must be and is accounted for in (2), since long upswings in $\{\beta_t\}$ and $\{R_t\}$ can lead to explosive paths for valuations and assets.

Next we study asymptotic stability, stationarity and ergodicity of wealth. Such properties are essential to existence of stationary equilibria in heterogeneous agent models (e.g., Huggett (1993), Aiyagari (1994) or Cao (2018)), as well as standard estimation, calibration and simulation techniques that connect time series averages with cross-sectional moments.[11] These properties require an additional restriction, placed on the asymptotic growth rate of mean returns. Analogous to (1) and (2), this is defined as

$$G_R := \lim_{n\to\infty} \left(\mathbb{E} \prod_{t=1}^{n} R_t \right)^{1/n}. \tag{3}$$

We show that if G_R is sufficiently restricted and a degree of social mobility is present, then there exists a unique stationary distribution for the state process, the distributional path of the state process under the optimal path converges globally to the stationary distribution, and the stationary distribution is ergodic. We also show that, under some mild additional conditions, the rate of convergence of marginal distributions to the stationary distribution is geometric, and that a version of the Central Limit Theorem is valid. Finally, under some mild additional conditions, we prove that the stationary distribution of assets is Pareto tailed, consistent with the data.

Our study is related to Benhabib et al. (2015), who prove the existence of a heavy-tailed wealth distribution in an infinite horizon heterogeneous agent economy with capital income risk. In the process, they show that households facing a stochastic return on savings possess a unique optimal consumption policy characterized by the (boundary constraint-contingent) Euler equation, and that a unique and unbounded stationary distribution exists for wealth under this consumption policy. They assume isoelastic utility, constant discounting, and mutually independent, IID returns and labor income processes, both supported on bounded closed intervals with strictly positive lower bounds. We relax all of these assumptions. Apart from allowing more general utility and state dependent discounting, this permits such realistic features for household income as positive correlations between labor earnings and wealth returns (an extension that was suggested by Benhabib et al. (2015)), or time varying volatility in returns.[12]

[11] A well-known example of a computational technique that uses ergodicity can be found in Krusell and Smith (1998). On the estimation side see, for example, Hansen and West (2002).

[12] Empirical motivation for these kinds of extensions can be found in numerous studies, including Guvenen and Smith (2014) and Fagereng et al. (2016a,b).

Another related paper is Chamberlain and Wilson (2000), which studies an income fluctuation problem with stochastic income and asset returns and obtains many significant results on asymptotic properties of consumption. Their study imposes relatively few restrictions on the wealth return and labor income processes. Our paper extends their work by allowing for random discounting, as well as dropping their boundedness restriction on the utility, which prevents their work from being used in many standard settings such as constant relative risk aversion. We also develop a set of new results on stability and ergodicity, as well as asymptotic normality of the wealth process.

Our optimality theory draws on techniques found in Li and Stachurski (2014), who show that the time iteration operator is a contraction mapping with respect to a metric that evaluates consumption differences in terms of marginal utility, while assuming a constant discount factor and constant rate of return on assets.[13] We show that these ideas extend to a setting where both returns and discount rates are stochastic and time varying. Our results on dynamics under the optimal policy have no counterparts in Li and Stachurski (2014).

In a similar vein, our work is related to several other papers that treat the standard income fluctuation problems with constant rates of return on assets and constant discount rates, such as Rabault (2002), Carroll (2004) and Kuhn (2013). While Carroll (2004) constructs a weighted supremum norm contraction and works with the Bellman operator, the other two papers focus on time iteration. In particular, Rabault (2002) exploits the monotonicity structure, while Kuhn (2013) applies a version of the Tarski fixed point theorem. Our techniques for studying optimality are close to those in Li and Stachurski (2014), as discussed above.

The rest of this paper is structured as follows. Section 2 formulates the problem and establishes optimality results. Sufficient conditions for the existence and uniqueness of optimal policies are discussed. Section 3 focuses on stochastic stability. Section 4 discusses our key conditions and how they can be checked. Section 5 provides a set of applications and Section 6 concludes. All proofs are deferred to the appendix. Code that generates our figures can be found at `https://github.com/jstac/ifp_public`.

[13] Coleman (1990) introduced the time iteration operator as a constructive method for solving stochastic growth models. It has since been used in Datta et al. (2002), Morand and Reffett (2003) and many other studies.

2. The Income Fluctuation Problem and Optimality Results

This section formulates the income fluctuation problem we consider, establishes the existence, uniqueness and computability of a solution, and derives its properties.

2.1. Problem Statement.

We consider a general income fluctuation problem, where a household chooses a consumption-asset path $\{(c_t, a_t)\}$ to solve

$$\begin{aligned} &\max \, \mathbb{E}_0 \left\{ \sum_{t=0}^{\infty} \left(\prod_{i=0}^{t} \beta_i \right) u(c_t) \right\} \\ \text{s.t.} \quad & a_{t+1} = R_{t+1}(a_t - c_t) + Y_{t+1}, \\ & 0 \le \; c_t \le a_t, \quad (a_0, Z_0) = (a, z) \text{ given.} \end{aligned} \tag{4}$$

Here u is the utility function, $\{\beta_t\}_{t \ge 0}$ is discount factor process with $\beta_0 = 1$, $\{R_t\}_{t \ge 1}$ is the gross rate of return on wealth, and $\{Y_t\}_{t \ge 1}$ is non-financial income. These stochastic processes obey

$$\beta_t = \beta\left(Z_t, \varepsilon_t\right), \quad R_t = R\left(Z_t, \zeta_t\right), \quad \text{and} \quad Y_t = Y\left(Z_t, \eta_t\right), \tag{5}$$

where β, R and Y are measurable nonnegative functions and $\{Z_t\}_{t \ge 0}$ is an irreducible time-homogeneous Z-valued Markov chain taking values in finite set Z. Let $P(z, \hat{z})$ be the probability of transitioning from z to $\hat{z}$ in one step. The innovation processes $\{\varepsilon_t\}$, $\{\zeta_t\}$ and $\{\eta_t\}$ are IID independent and their supports can be continuous and vector-valued.

The function u maps $\mathbb{R}_+$ to $\{-\infty\} \cup \mathbb{R}$, is twice differentiable on $(0, \infty)$, satisfies $u' > 0$ and $u'' < 0$ everywhere on $(0, \infty)$, and that $u'(c) \to \infty$ as $c \to 0$ and $u'(c) < 1$ as $c \to \infty$. We define

$$\mathbb{E}_{a,z} := \mathbb{E}\left[\, \cdot \,\middle|\, (a_0, Z_0) = (a, z)\right] \quad \text{and} \quad \mathbb{E}_z := \mathbb{E}\left[\, \cdot \,\middle|\, Z_0 = z\right]. \tag{6}$$

The next period value of a random variable X is typically denoted $\hat{X}$. Expectation without a subscript refers to the stationary process, where Z_0 is drawn from its (necessarily unique) stationary distribution.

Example 2.1. The optimization problem stated above includes the problem faced by households in Cao (2018), which in turn builds on Krusell and Smith (1998). In that model, randomness in the discount factor process is driven by idiosyncratic preference shocks $\{i_t\}$, while returns on assets fluctuate due to an aggregate shock process $\{s_t\}$ impacting productivity. Non-financial income is affected by both the idiosyncratic shocks and the aggregate shocks. His scenario fits our framework when $Z_t = (s_t, i_t)$. The innovation vectors in (5) are not required.

2.2. **Key Conditions.** Our conditions for optimality are listed below. In what follows, G_β is the asymptotic growth rate of the discount process as defined in (1).

Assumption 2.1. The discount factor process satisfies $G_\beta < 1$.

Assumption 2.1 is a natural extension of the standard condition $\beta < 1$ from the constant discount case. If $\beta_t \equiv \beta$ for all t, then $G_\beta = \beta$, as follows immediately from the definition. It is weaker than the obvious sufficient condition $\beta_t \leq \bar{\beta}$ with probability one for some constant $\bar{\beta} < 1$, since in such a setting we have $G_\beta \leq \bar{\beta} < 1$. In fact it cannot be significantly weakened, as the theorem shows.

Proposition 2.1 (Necessity of the discount condition)**.** *Let β_t and $u(Y_t)$ be positive with probability one for all t and all initial states z in Z. If, in this setting, we have $G_\beta \geq 1$, then the objective in (4) is infinite at every initial state (a, z).*

The positivity assumed here may or may not hold in applications, but Proposition 2.1 shows that special conditions will have to be imposed on preferences if Assumption 2.1 fails. Put differently, allowing $G_\beta \geq 1$ is tantamount to allowing $\beta \geq 1$ in the case when the discount rate is constant.

Next, we need to ensure that the present discounted value of wealth does not grow too quickly, which requires a joint restriction on asset returns and discounting. When $\{R_t\}$ and $\{\beta_t\}$ are constant at values R and β, the standard restriction from the existing literature is $\beta R < 1$. A generalization using $G_{\beta R}$ as defined in (2) is

Assumption 2.2. The discount factor and return processes satisfy $G_{\beta R} < 1$.

Finally, we impose routine technical restrictions on non-financial income. The second restriction is needed to exploit first order conditions.

Assumption 2.3. $\mathbb{E}\, Y < \infty$ and $\mathbb{E}\, u'(Y) < \infty$.

Next we provide one example where Assumptions 2.1–2.3 are easily verified. More complex examples are deferred to Sections 4 and 5.

Example 2.2. Suppose, as in Benhabib et al. (2015), that there is a constant discount factor $\beta < 1$, utility is CRRA with $\gamma \geq 1$, $\{R_t\}$ and $\{Y_t\}$ are IID, mutually independent, supported on bounded closed intervals of strictly positive real numbers, and, moreover,

$$\beta \mathbb{E} R_t^{1-\gamma} < 1 \quad \text{and} \quad (\beta \mathbb{E} R_t^{1-\gamma})^{1/\gamma} \mathbb{E} R_t < 1. \tag{7}$$

Assumptions 2.1–2.3 are all satisfied in this case. To see this, observe that $G_\beta = \beta < 1$ in the constant discount case, so Assumption 2.1 holds. Since $x \mapsto x^{1-\gamma}$ is convex when $\gamma \geq 1$, Jensen's inequality implies that $\mathbb{E} R_t^{1-\gamma} \geq (\mathbb{E} R_t)^{1-\gamma}$. Multiplying both sides of the last inequality by $\beta(\mathbb{E} R_t)^\gamma$ yields

$$G_{\beta R} = \beta \mathbb{E} R_t = \beta (\mathbb{E} R_t)^{1-\gamma} (\mathbb{E} R_t)^\gamma \leq (\beta \mathbb{E} R_t^{1-\gamma})(\mathbb{E} R_t)^\gamma.$$

By the second condition of (7), Assumption 2.2 holds. Assumption 2.3 also holds because Y_t is restricted to a compact subset of the positive reals.

2.3. **Optimality: Definitions and Fundamental Properties.** To consider optimality, we temporarily assume that $a_0 > 0$ and set the asset space to $(0, \infty)$.[14] The state space for $\{(a_t, Z_t)\}_{t \geq 0}$ is then $\mathsf{S}_0 := (0, \infty) \times \mathsf{Z}$. A *feasible policy* is a Borel measurable function $c \colon \mathsf{S}_0 \to \mathbb{R}$ with $0 \leq c(a, z) \leq a$ for all $(a, z) \in \mathsf{S}_0$. A feasible policy c and initial condition $(a, z) \in \mathsf{S}_0$ generate an asset path $\{a_t\}_{t \geq 0}$ via (4) when $c_t = c(a_t, Z_t)$ and $(a_0, Z_0) = (a, z)$. The lifetime value of policy c is

$$V_c(a, z) = \mathbb{E}_{a,z} \sum_{t=0}^{\infty} \beta_0 \cdots \beta_t u \left[c(a_t, Z_t) \right], \tag{8}$$

where $\{a_t\}$ is the asset path generated by $(c, (a, z))$. In the Appendix we show that V_c is well-defined on S_0. A feasible policy c^* is called *optimal* if $V_c \leq V_{c^*}$ on S_0 for any feasible policy c. A feasible policy is said to satisfy the *first order optimality condition* if

$$(u' \circ c)(a, z) \geq \mathbb{E}_z \, \hat{\beta} \hat{R} \, (u' \circ c) \left(\hat{R} \left[a - c(a, z) \right] + \hat{Y}, \, \hat{Z} \right) \tag{9}$$

for all $(a, z) \in \mathsf{S}_0$, and equality holds when $c(a, z) < a$. Noting that u' is decreasing, the first order optimality condition can be compactly stated as

$$(u' \circ c)(a, z) = \max \left\{ \mathbb{E}_z \, \hat{\beta} \hat{R} \, (u' \circ c) \left(\hat{R} \left[a - c(a, z) \right] + \hat{Y}, \, \hat{Z} \right), u'(a) \right\} \tag{10}$$

for all $(a, z) \in \mathsf{S}_0$. A feasible policy is said to satisfy the *transversality condition* if, for all $(a, z) \in \mathsf{S}_0$,

$$\lim_{t \to \infty} \mathbb{E}_{a,z} \, \beta_0 \cdots \beta_t \, (u' \circ c)(a_t, Z_t) \, a_t = 0. \tag{11}$$

[14]Assumption 2.3 combined with $u'(0) = \infty$ implies that $\mathbb{P}\{Y_t > 0\} = 1$ for all $t \geq 1$. Hence, $\mathbb{P}\{a_t > 0\} = 1$ for all $t \geq 1$ and exclude zero from the asset space makes no difference to optimality.

Theorem 2.1 (Sufficiency of first order and transversality conditions)**.** *If Assumptions 2.1–2.3 hold, then every feasible policy satisfying the first order and transversality conditions is an optimal policy.*

2.4. **Existence and Computability of Optimal Consumption.** Let $\mathscr{C}$ be the space of continuous functions $c\colon \mathsf{S}_0 \to \mathbb{R}$ such that c is increasing in the first argument, $0 < c(a,z) \leq a$ for all $(a,z) \in \mathsf{S}_0$, and

$$\sup_{(a,z)\in\mathsf{S}_0} |(u' \circ c)(a,z) - u'(a)| < \infty. \tag{12}$$

To compare two consumption policies, we pair $\mathscr{C}$ with the distance

$$\rho(c,d) := \|u' \circ c - u' \circ d\| := \sup_{(a,z)\in\mathsf{S}_0} \left|(u' \circ c)(a,z) - (u' \circ d)(a,z)\right|, \tag{13}$$

which evaluates the maximal difference in terms of marginal utility. While elements of $\mathscr{C}$ are not generally bounded, ρ is a valid metric on $\mathscr{C}$. In particular, ρ is finite on $\mathscr{C}$ since $\rho(c,d) \leq \|u' \circ c - u'\| + \|u' \circ d - u'\|$, and the last two terms are finite by (12). In Appendix B, we show that $(\mathscr{C}, \rho)$ is a complete metric space. The following proposition shows that, for any policy in $\mathscr{C}$, the first order optimality condition (10) implies the transversality condition.

Proposition 2.2 (Sufficiency of first order condition)**.** *Let Assumptions 2.1–2.3 hold. If $c \in \mathscr{C}$ and the first order optimality condition (10) holds for all $(a,z) \in \mathsf{S}_0$, then c satisfies the transversality condition. In particular, c is an optimal policy.*

We aim to characterize the optimal policy as the fixed point of the *time iteration operator* T defined as follows: for fixed $c \in \mathscr{C}$ and $(a,z) \in \mathsf{S}_0$, the value of the image Tc at (a,z) is defined as the $\xi \in (0,a]$ that solves

$$u'(\xi) = \psi_c(\xi, a, z), \tag{14}$$

where ψ_c is the function on

$$G := \{(\xi, a, z) \in \mathbb{R}_+ \times (0,\infty) \times \mathsf{Z} \colon 0 < \xi \leq a\} \tag{15}$$

defined by

$$\psi_c(\xi, a, z) := \max\left\{ \mathbb{E}_z \, \hat\beta \hat R (u' \circ c)[\hat R(a - \xi) + \hat Y,\ \hat Z],\ u'(a) \right\}. \tag{16}$$

The following theorem shows that the time iteration operator is an n-step contraction mapping on a complete metric space of candidate policies and its fixed point is the unique optimal policy.

Theorem 2.2 (Existence, uniqueness and computability of optimal policies)**.** *If Assumptions 2.1–2.3 hold, then there exists an n in $\mathbb{N}$ such that T^n is a contraction mapping on $(\mathscr{C}, \rho)$. In particular,*

(1) *T has a unique fixed point $c^* \in \mathscr{C}$.*
(2) *The fixed point c^* is the unique optimal policy in $\mathscr{C}$.*
(3) *For all $c \in \mathscr{C}$ we have $\rho(T^k c, c^*) \to 0$ as $k \to \infty$.*

Part (3) shows that, under our conditions, the familiar time iteration algorithm is globally convergent, provided one starts with some policy in the candidate class $\mathscr{C}$.

2.5. **Properties of Optimal Consumption.** In this section we study the properties of the optimal consumption function obtained in Theorem 2.2. Assumptions 2.1–2.3 are held to be true throughout. The following two propositions show the monotonicity of the consumption function, which is intuitive.

Proposition 2.3 (Monotonicity with respect to wealth)**.** *The optimal consumption and savings functions $c^*(a, z)$ and $i^*(a, z) := a - c^*(a, z)$ are increasing in a.*

Proposition 2.4 (Monotonicity with respect to income)**.** *If $\{Y_{1t}\}$ and $\{Y_{2t}\}$ are two income processes satisfying $Y_{1t} \leq Y_{2t}$ for all t and c_1^* and c_2^* are the corresponding optimal consumption functions, then $c_1^* \leq c_2^*$ pointwise on S_0.*

Under further assumptions we can show that the optimal policy is concave and asymptotically linear with respect to the wealth level.

Proposition 2.5 (Concavity and asymptotic linearity of consumption function)**.** *If for each $z \in \mathsf{Z}$ and $c \in \mathscr{C}$ that is concave in its first argument,*

$$x \mapsto (u')^{-1} \left[\mathbb{E}_z \hat{\beta} \hat{R} \, (u' \circ c) \, (\hat{R} x + \hat{Y}, \, \hat{Z}) \right] \quad \textit{is concave on } \mathbb{R}_+, \tag{17}$$

then

(1) *$a \mapsto c^*(a, z)$ is concave, and*
(2) *there exists $\alpha(z) \in [0, 1]$ such that $\lim_{a \to \infty} [c^*(a, z)/a] = \alpha(z)$.*

Remark 2.1. Condition (17) imposes some concavity structure on utility. It holds for the constant relative risk aversion (CRRA) utility function

$$u(c) = \frac{c^{1-\gamma}}{1-\gamma} \quad \text{if } \gamma > 0 \quad \text{and} \quad u(c) = \log c \quad \text{if } \gamma = 1, \tag{18}$$

as shown in Appendix B.

Proposition 2.5 states that $c^*(a,z) \approx \alpha(z)a + b(z)$ for some function $b(z)$ when a is large. This provides justification for linearly extrapolating the policy functions when computing them at high wealth levels.

Together, parts (1) and (2) of Proposition 2.5 imply the linear lower bound $c^*(a,z) \geq \alpha(z)a$, although they do not provide a concrete number for $\alpha(z)$. The following proposition establishes an explicit linear lower bound.

Proposition 2.6 (Linear lower bound on consumption)**.** *If there exists a nonnegative constant $\bar{s}$ such that*

$$\bar{s} < 1 \qquad \textit{and} \qquad \mathbb{E}_z \hat{\beta} \hat{R}\, u'(\hat{R}\, \bar{s}\, a) \leq u'(a) \textit{ for all } (a,z) \in \mathsf{S}_0, \tag{19}$$

then $c^*(a,z) \geq (1-\bar{s})a$ for all $(a,z) \in \mathsf{S}_0$.[15]

The second inequality in (19) restricts marginal utility derived from transferring wealth to the next period and then consuming versus consuming wealth today. The value $\bar{s}$ can be clarified once primitives are specified, as the next example illustrates.

Example 2.3. Suppose that utility is CRRA, as in (18). If we now take

$$\bar{s} := \left(\max_{z \in \mathsf{Z}} \mathbb{E}_z \hat{\beta} \hat{R}^{1-\gamma} \right)^{1/\gamma} \tag{20}$$

and $\bar{s} < 1$, then the conditions of Proposition 2.6 hold. In particular, the second inequality in (19) holds, as follows directly from the definition of $\bar{s}$ and $u'(x) = x^{-\gamma}$. In the case of Benhabib et al. (2015), where the discount rate is constant and returns are IID, the expression in (20) reduces to $\bar{s} := (\beta \mathbb{E} R_t^{1-\gamma})^{1/\gamma}$. The requirement $\bar{s} < 1$ then reduces to $\beta \mathbb{E} R_t^{1-\gamma} < 1$, which is one of their assumptions (see Example 2.2).

3. Stationarity, Ergodicity, and Tail Behavior

This section focuses on stationarity, ergodicity and tail behavior of wealth under the unique optimal policy c^* obtained in Theorem 2.2. So that this policy exists, Assumptions 2.1–2.3 are always taken to be valid. We extend c^* to S by setting

[15] We adopt the convention $0 \cdot \infty = 0$, so condition (19) does not rule out the case $\mathbb{P}\{R_t = 0 \mid Z_{t-1} = z\} > 0$. Indeed, as shown in the proofs, the conclusions still hold if we replace this condition by the weaker alternative $\mathbb{E}_z \hat{\beta} \hat{R}\, u'[\hat{R}\bar{s}a + (1-\bar{s})\hat{Y}] \leq u'(a)$ for all $(a,z) \in \mathsf{S}_0$.

$c^*(0, z) = 0$ for all $z \in \mathsf{Z}$ and consider dynamics of (a_t, Z_t) on $\mathsf{S} := \mathbb{R}_+ \times \mathsf{Z}$, the law of motion for which is

$$a_{t+1} = R\left(Z_{t+1}, \zeta_{t+1}\right)\left[a_t - c^*\left(a_t, Z_t\right)\right] + Y\left(Z_{t+1}, \eta_{t+1}\right), \tag{21a}$$

$$Z_{t+1} \sim P\left(Z_t, \,\cdot\,\right) \tag{21b}$$

Let Q be the joint stochastic kernel of (a_t, Z_t) on S. See Appendix A for this and related definitions.

3.1. **Stationarity.** To obtain existence of a stationary distribution we need to restrict the asymptotic growth rate for asset returns G_R defined in (3).

Assumption 3.1. There exists a constant $\bar{s}$ such that (19) holds and $\bar{s}\, G_R < 1$.

Below is one straightforward example of a setting where this holds, with more complex applications deferred to Sections 4–5.

Example 3.1. Assumption 3.1 holds in the setting of Benhabib et al. (2015). As shown in Example 2.3, with $\bar{s} := (\beta \mathbb{E} R_t^{1-\gamma})^{1/\gamma}$ and the assumptions of Benhabib et al. (2015) in force, the conditions of (19) hold. Moreover, in their IID setting we have $G_R = \mathbb{E} R_t$, so $\bar{s}\, G_R < 1$ reduces to $(\beta \mathbb{E} R_t^{1-\gamma})^{1/\gamma} \mathbb{E} R_t < 1$. This is one of their conditions, as discussed in Example 2.2.

By Proposition 2.6, the value $\bar{s}$ in Assumption 3.1 is an upper bound on the rate of savings. G_R is an asymptotic growth rate for each unit of savings invested. If the product of these is less than one, then probability mass contained in the wealth distribution will not drift to $+\infty$, which allows us to obtain the following result.[16]

Theorem 3.1 (Existence of a stationary distribution)**.** *If Assumption 3.1 holds, then Q admits at least one stationary distribution on S.*

Stationarity of the form obtained in Theorem 3.1 is required to establish existence of stationary recursive equilibria in heterogeneous agent models with idiosyncratic risk, such as Huggett (1993) or Aiyagari (1994).[17]

[16]Assumption 3.1 is weaker than any restriction implying wealth is bounded from above—a common device for compactifying the state space and thereby obtaining a stationary distribution. Indeed, under many specifications of $\{Y_t\}$ and $\{R_t\}$ that fall within our framework, wealth of a given household can and will, over an infinite horizon, exceed any finite bound with probability one. See, for example, Benhabib et al. (2015), Proposition 6.

[17]For models with aggregate shocks, such as Krusell and Smith (1998), a fully specified recursive equilibrium requires that households take the wealth distribution as one component of the state in

3.2. **Ergodicity.** While Assumption 3.1 implies existence of a stationary distribution, it is not in general sufficient for uniqueness or stability. For these additional properties to hold, we must impose sufficient mixing. In doing so, we consider the following two cases:

(Y1) The support of $\{Y_t\}$ is finite.
(Y2) The process $\{Y_t\}$ admits a density representation.

Condition (Y2) means that there exists a function f from $\mathbb{R}_+ \times \mathsf{Z}$ to $\mathbb{R}_+$ such that

$$\mathbb{P}\{Y_t \in A \mid Z_t = z\} = \int_A f(y \mid z)\, \mathrm{d}y \tag{22}$$

for all Borel sets $A \subset \mathbb{R}_+$ and all z in Z.

Assumption 3.2. There exists a $\bar{z}$ in Z such that $P(\bar{z}, \bar{z}) > 0$. Moreover, with $y_\ell \geq 0$ defined as the greatest lower bound of the support of $\{Y_t\}$, either

- (Y1) holds and $\mathbb{P}\{Y_t = y_\ell \mid Z_t = \bar{z}\} > 0$, or
- (Y2) holds and there exists a $\delta > y_\ell$ such that $f(\cdot \mid \bar{z}) > 0$ on (y_ℓ, δ).

Assumption 3.2 requires that there is a positive probability of receiving low labor income at some relatively persistent state of the world $\bar{z}$. This is a mixing condition that enforces social mobility. The reason is that $\{Z_t\}$ is already assumed to be irreducible, so $\bar{z}$ is eventually visited by each household. For any such household, there is a positive probability of low labor income over a long period. Wealth then declines. In other words, currently rich households or dynasties will not be rich forever. This guarantees sufficient social mobility between rich and poor, generating ergodicity.

To state our uniqueness and stability results, let Q^t be the t-step stochastic kernel, let $\|\cdot\|_{TV}$ be total variation norm and let $V(a, z) := a + m_V$, where m_V is a constant to be defined in the proof. For any integrable real-valued function h on S, let

$$\bar{h}(a, z) := h(a, z) - \mathbb{E} h(a_t, Z_t)$$

and

$$\gamma_h^2 := \mathbb{E}\left[\bar{h}^2(a_0, Z_0)\right] + 2 \sum_{t=1}^{\infty} \mathbb{E}\left[\bar{h}(a_0, Z_0)\bar{h}(a_t, Z_t)\right],$$

their savings problem, and that stationarity holds for the entire joint distribution (defined over a product space encompassing both the wealth distribution and the exogenous state process). These problems fall outside the scope of Theorem 3.1, since $\{Z_t\}$ is finite-valued. For a careful treatment of stationary recursive equilibrium in Krusell–Smith type models, see Cao (2018).

where, here and in the theorem below, $\mathbb{E}$ indicates expectation under stationarity.

Theorem 3.2 (Uniqueness, stability, ergodicity and mixing)**.** *If Assumptions 3.1 and 3.2 hold, then*

(1) *the stationary distribution ψ_∞ of Q is unique and there exist constants $\lambda < 1$ and $M < \infty$ such that,*

$$\left\| Q^t\left((a,z),\cdot\right) - \psi_\infty \right\|_{TV} \leq \lambda^t M V(a,z) \quad \textit{for all } (a,z) \in \mathsf{S}.$$

(2) *For all $(a,z) \in \mathsf{S}$ and real-valued function h on S such that $\mathbb{E}|h(a_t, Z_t)| < \infty$,*

$$\mathbb{P}_{a,z}\left\{ \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} h(a_t, Z_t) = \mathbb{E}h(a_t, Z_t) \right\} = 1.$$

(3) *Q is V-geometrically mixing. Moreover, if $\gamma_h^2 > 0$ and h^2/V is bounded,*

$$\frac{1}{\sqrt{T\gamma_h^2}} \sum_{t=1}^{T} \bar{h}(a_t, Z_t) \xrightarrow{d} N(0,1) \quad \textit{as } T \to \infty.$$

Part 1 of Theorem 3.2 states that the stationary distribution ψ_∞ is unique and asymptotically attracting at a geometric rate. Part 2 states that the state process is ergodic, and hence long-run sample moments for individual households coincide with cross-sectional moments. The notion of mixing discussed in Part 3 is defined in the appendix. It states that social mobility holds asymptotically and mixing occurs at a geometric rate, although the rate may be arbitrarily slow. This mixing is enough to provide a Central Limit Theorem for the state process, which is the second claim in Part 3.

3.3. **Tail Behavior.** Having established the stationarity and ergodicity of wealth, we now study the tail behavior of the wealth distribution. We show that the wealth distribution is either bounded or (unbounded and) heavy-tailed under mild conditions. To prove this result we introduce the following assumption.

Assumption 3.3. The assumptions of Proposition 2.5 are satisfied, so the optimal policy $a \mapsto c^*(a,z)$ is concave and asymptotically linear: $\lim_{a\to\infty} c^*(a,z)/a = \alpha(z) \in [0,1]$. Furthermore, there exists $\bar{z} \in \mathsf{Z}$ such that $P(\bar{z},\bar{z}) > 0$ and

$$\mathbb{P}_{\bar{z}}\{R(\bar{z},\hat{\zeta})(1-\alpha(\bar{z})) > 1\} > 0. \tag{23}$$

Remark 3.1. Condition (23) implies that wealth grows with nonzero probability when it is large. Indeed, using the law of motion (21a) and noting that $Y \geq 0$, if $Z_t = Z_{t+1} = \bar{z}$, then by (23) we have

$$\frac{a_{t+1}}{a_t} \geq R(\bar{z}, \zeta_{t+1})\left[1 - c^*(a_t, \bar{z})/a_t\right] > 1$$

with positive probability if a_t is large enough.

To state our result on tail behavior, we introduce the following notation. For any nonnegative function $A(z, \hat{z}, \hat{\zeta})$, define the $\mathsf{Z} \times \mathsf{Z}$ matrix-valued function M_A by

$$(M_A(s))(z, \hat{z}) = \mathbb{E}_{z,\hat{z}} A(z, \hat{z}, \hat{\zeta})^s. \tag{24}$$

Elements of $M_A(s)$ are conditional moment generating functions of $\log A$. In the statement below, $\odot$ denotes the Hadamard (entry-wise) product, and $r(\cdot)$ returns the spectral radius of a matrix. Also a_∞ is a random variable with distribution $\psi_\infty(\cdot, \mathsf{Z})$.

Theorem 3.3 (Tail behavior)**.** *Let Assumptions 3.1–3.3 hold and define*

$$G(z, \hat{z}, \hat{\zeta}) = R(\hat{z}, \hat{\zeta})(1 - \alpha(z)), \tag{25a}$$

$$A(z, \hat{z}, \hat{\zeta}) = G(z, \hat{z}, \hat{\zeta})\mathbb{1}\{G(z, \hat{z}, \hat{\zeta}) > 1\}, \text{ and} \tag{25b}$$

$$\lambda(s) = r(P \odot M_A(s)). \tag{25c}$$

Then λ is convex in $s \geq 0$. Assume that there exists $s > 0$ in the interior of the domain of λ such that $1 < \lambda(s) < \infty$ and let

$$\kappa := \inf\{s > 0 \mid \lambda(s) > 1\}. \tag{26}$$

If a_∞ has unbounded support, then it is heavy-tailed. In particular, for any $\varepsilon > 0$,

$$\liminf_{a \to \infty} a^{\kappa+\varepsilon} \mathbb{P}\{a_\infty \geq a\} > 0. \tag{27}$$

Remark 3.2. The assumption $1 < \lambda(s) < \infty$ for some $s > 0$ is weak. Because the $(\bar{z}, \bar{z})$-th element of $P \odot M_A(s)$ is

$$P(\bar{z}, \bar{z})\mathbb{E}_{\bar{z},\bar{z}} G(\bar{z}, \bar{z}, \hat{\zeta})^s \mathbb{1}\{G(\bar{z}, \bar{z}, \hat{\zeta}) > 1\},$$

by the definition of G in (25a) and condition (23), we always have $\lambda(s) \to \infty$ as $s \to \infty$. Hence there exists $s > 0$ such that $\lambda(s) \in (1, \infty)$ if, for example, $\hat{\zeta}$ has a compact support.

Condition (27) implies that for any $\varepsilon > 0$, there exists a constant $C(\varepsilon) > 0$ such that

$$\mathbb{P}\{a_\infty \geq a\} \geq C(\varepsilon) a^{-\kappa-\varepsilon}$$

for large enough a, so the upper tail of the wealth distribution is at least Pareto.

Remark 3.3. Toda (2019) constructs an example of a Huggett (1993) economy with Pareto-tailed wealth distribution when discount factors are random. Theorem 3.3 is significantly more general as we allow for stochastic returns and income. Stachurski and Toda (2019) prove that with constant discount factor, constant asset return, and light-tailed income, the wealth distribution is always light-tailed. Theorem 3.3 shows that sufficient heterogeneity in discount factor or returns generates heavy tails.

Example 3.2. The CRRA-IID setting of Benhabib et al. (2015) satisfies the assumptions of Theorem 3.3. When utility is CRRA, by Proposition 5 of Benhabib et al. (2015), condition (23) holds if $R(\bar{z}, \hat{\zeta}) > 1/\bar{s}$ with positive probability, where $\bar{s}$ is given in Example 2.3. In the IID case, this condition reduces to $\mathbb{P}\{(\beta \mathbb{E} R_t^{1-\gamma})^{1/\gamma} R_t > 1\} > 0$, which holds under the conditions of Benhabib et al. (2015).[18] Thus, Assumption 3.3 holds. The existence of $s > 0$ with $\lambda(s) \in (1, \infty)$ follows from Remark 3.2 and the assumption that R_t has a compact support.

4. Testing the Growth Conditions

The three key conditions in the paper are the restrictions on the growth rates G_β, $G_{\beta R}$ and G_R, with the first two required for optimality and the last for stationarity (see Assumptions 2.1, 2.2 and 3.1 respectively). In this section we explore the restrictions implied by these conditions. We begin with the following result, which yields a straightforward method for computing these growth rates.

Lemma 4.1 (Long-run growth rates and spectral radii). *Let $\varphi_t = \varphi(Z_t, \xi_t)$, where φ is a nonnegative measurable function and $\{\xi_t\}$ is an* IID *sequence with marginal distribution π. In this setting we have*

$$G_\varphi = r(L_\varphi), \quad \text{where} \quad G_\varphi := \lim_{n \to \infty} \left(\mathbb{E} \prod_{t=1}^{n} \varphi_t \right)^{1/n} \tag{28}$$

and $r(L_\varphi)$ is the spectral radius of the matrix defined by

$$L_\varphi(z, \hat{z}) = P(z, \hat{z}) \int \varphi(\hat{z}, \hat{\xi}) \pi(\mathrm{d}\hat{\xi}). \tag{29}$$

---

[18] Benhabib et al. (2015) assume that $\mathbb{P}\{\beta R_t > 1\} > 0$, so it suffices to show that $(\beta \mathbb{E} R_t^{1-\gamma})^{1/\gamma} \geq \beta$ or, equivalently, $\mathbb{E}(\beta R_t)^{1-\gamma} \geq 1$. By Jensen's inequality and their restriction $\gamma \geq 1$, the last bound is true whenever $(\mathbb{E} \beta R_t)^{1-\gamma} \geq 1$. But this must hold because, under their conditions, we have $\beta \mathbb{E} R_t < 1$, as shown in Example 2.2.

The matrix L_φ is expressed as a function on $\mathsf{Z} \times \mathsf{Z}$ in (29) but can be represented in traditional matrix notation by enumerating Z.[19]

What factors determine the long-run average growth rates embedded in our assumptions, such as G_β or G_R? Lemma 4.1 tells us how to compute these values for a given specification of dynamics, but how should we understand them intuitively and what factors determine their size? To address these questions, let us consider an AR(1) discount factor process, which has been adopted in several recent quantitative studies (see, e.g., Hubmer et al. (2018) or Hills and Nakata (2018)). In particular, suppose that the state process follows a discretized version of

$$Z_{t+1} = \rho Z_t + (1-\rho)\mu + (1-\rho^2)^{1/2}\sigma \upsilon_{t+1}, \quad \{\upsilon_t\} \overset{\text{IID}}{\sim} N(0,1), \tag{30}$$

and $\beta_t = Z_t$. (The discretization implies that β_t is always positive.) To simplify interpretation, the process (30) is structured so that the stationary distribution of $\{Z_t\}$ is $N(\mu, \sigma^2)$. We use Rouwenhorst (1995)'s method to discretize $\{Z_t\}$ and then calculate G_β using Lemma 4.1, studying how G_β is affected by the parameters in (30).

Since $\beta_t = Z_t$ for all t, the structure of (30) implies that μ is the long-run unconditional mean of $\{\beta_t\}$. It can therefore be set to standard calibrated value for the discount factor, such as 0.99 from Krusell and Smith (1998). What we wish to understand is how the remaining parameters ρ and σ affect the value of G_β. While no closed form expression is available in this case, Figure 1 sheds some light by providing a contour plot of G_β over a set of (ρ, σ) pairs. The figure shows that G_β grows with both the persistence term ρ and volatility term σ. In particular, the condition $G_\beta < 1$ fails when the persistence and volatility of the discount factor process are sufficiently high. This is because G_β is the limit of $(\mathbb{E} \prod_{t=1}^n \beta_t)^{1/n}$ and, for positive random variables, sequence of large outcomes have a strong compounding effect on their product. High volatility and high persistence reinforce this effect.

This discussion has focused on G_β but similar intuition applies to both G_R and $G_{\beta R}$. If β_t and R_t are both increasing functions of the state process, then these asymptotic growth rates also increase with greater persistence and volatility in the state process, as well as higher unconditional mean. The next section further illustrates these points.

[19] Specifically, if $\mathsf{Z} := \{z_1, \dots, z_N\}$, then $L_\varphi = P D_\varphi$ where P is, as before, the transition matrix for the exogenous state, and $D_\varphi := \operatorname{diag}(\mathbb{E}_{z_1}\varphi, \dots, \mathbb{E}_{z_N}\varphi)$ when $\mathbb{E}_z\varphi := \mathbb{E}_z \varphi(z, \hat{\xi})$. In what follows, D_β, D_R and $D_{\beta R}$ are defined analogously to D_φ.

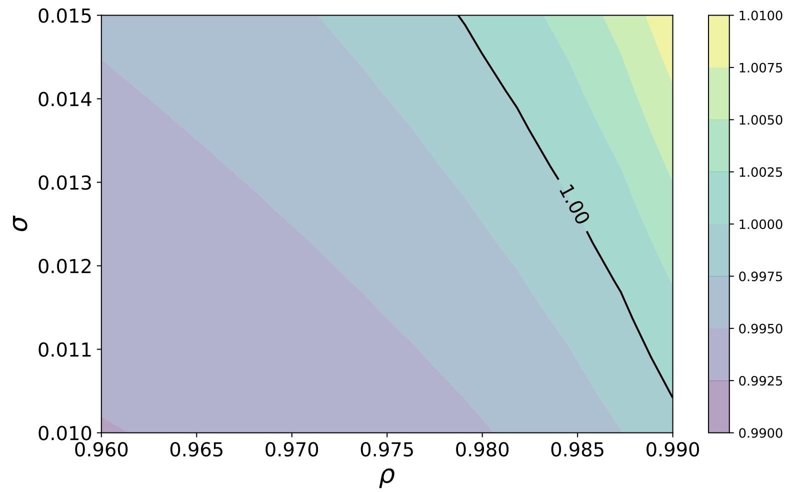


FIGURE 1. Contour plot of G_β under AR(1) discounting

5. Application: Stochastic Volatility and Mean Persistence

We showed in Examples 2.2, 2.3 and 3.1 that, in the setting of Benhabib et al. (2015), where the discount factor is constant and returns and labor income are IID, Assumptions 2.1–2.3 and Assumption 3.1 are all satisfied. Hence, by Theorems 2.2 and 3.1, the household optimization problem has a unique optimal policy and the wealth process under this policy has a stationary solution. If, in addition, the support of Y_t is finite or Y_t has a positive density, say, then the conditions of Theorem 3.2 also hold and the stationary solution is ergodic, geometrically mixing and its time series averages are asymptotically normal.

Let us now bring the model closer to the data by relaxing the IID restrictions on financial and non-financial returns, introducing both mean persistence and time varying volatility in returns on assets.[20] In particular, we set

$$\log R_t = \mu_t + \sigma_t \zeta_t, \tag{31}$$

where $\{\zeta_t\}$ is IID and standard normal and $\{\mu_t\}$ and $\{\sigma_t\}$ are finite-state Markov chains, discretized from

$$\mu_t = (1-\rho_\mu)\bar{\mu} + \rho_\mu \mu_{t-1} + \delta_\mu \upsilon_t^\mu \quad \text{and} \quad \log \sigma_t = (1-\rho_\sigma)\bar{\sigma} + \rho_\sigma \log \sigma_{t-1} + \delta_\sigma \upsilon_t^\sigma.$$

[20]The importance of these features for wealth dynamics was highlighted in Fagereng et al. (2016a).

Innovations are IID and standard normal. Using the data in Fagereng et al. (2016b) on Norwegian financial returns over 1993–2003, we estimate these AR(1) models to obtain $\bar{\mu} = 0.0281$, $\rho_\mu = 0.5722$, $\delta_\mu = 0.0067$, $\bar{\sigma} = -3.2556$, $\rho_\sigma = 0.2895$ and $\delta_\sigma = 0.1896$. Based on this calibration, the stationary mean and standard deviation of $\{R_t\}$ are around 1.03 and 4%, respectively.

To distinguish the effects of stochastic volatility and mean persistence, we consider two subsidiary models. The first reduces $\{\mu_t\}$ to its stationary mean $\bar{\mu}$, while the second reduces $\{\sigma_t\}$ to its stationary mean $\tilde{\sigma} := \mathrm{e}^{\bar{\sigma}+\delta_\sigma^2/2(1-\rho_\sigma^2)}$. In summary,

$$\log R_t = \bar{\mu} + \sigma_t \zeta_t \qquad \text{(Model I)}$$

$$\log R_t = \mu_t + \tilde{\sigma} \zeta_t \qquad \text{(Model II)}$$

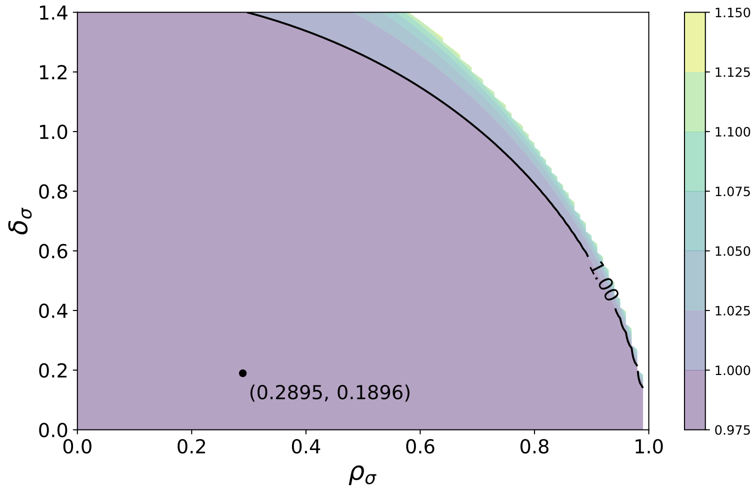
We set $\beta = 0.95$ and $\gamma = 1.5$. To test the stability properties of Model I, we set $K = 5$ and explore a neighborhood of the calibrated $(\rho_\sigma, \delta_\sigma)$ values, while in Model II, we set $M = 5$ and do likewise for (ρ_μ, δ_μ) pairs. In each scenario, other parameters are fixed to the benchmark. The results are shown in Figures 2 and 3.

In part (a) of each figure, we see that $G_{\beta R}$ is increasing in the persistence and volatility parameters of the state process. The intuition behind this feature was explained in Section 4 for the case of G_β and is similar here. (Note that $G_{\beta R} = \beta G_R$ in the present case, since $\beta_t \equiv \beta$ is a constant, so $G_{\beta R}$ has the same shape as G_R in terms of contours.) The dots in the figures show that $G_{\beta R} < 1$ at the estimated parameter values.

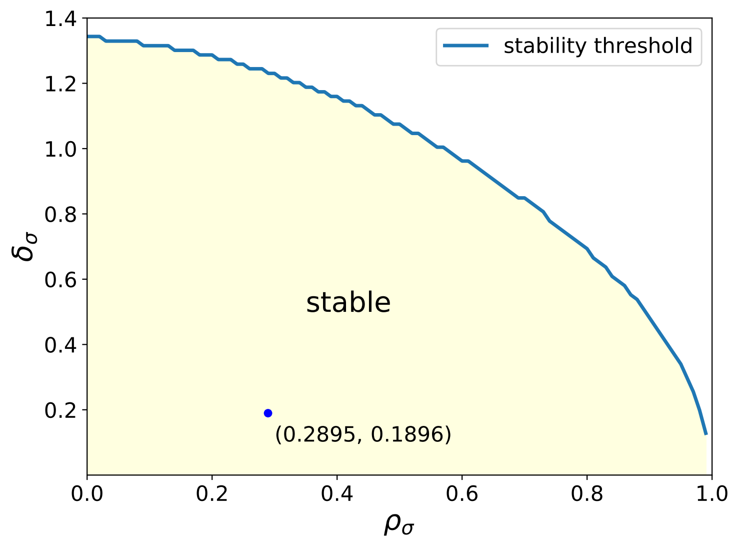
Part (b) of each figure shows the set of parameters under which the model is globally stable and ergodic. The stability threshold is the boundary of the set of parameter pairs that produce $\max\{G_{\beta R}, \bar{s}, \bar{s} G_R\} < 1$, where $\bar{s}$ is given by (20). For such pairs, Assumptions 2.2 and 3.1 both hold, so the conditions of Theorems 3.1–3.2 are satisfied. (We are continuing to suppose that Y_t is finite or has a positive density, so that Assumption 3.2 holds. Assumptions 2.1 and 2.3 are always valid in the current setting). Observe that the estimated parameter values (dot points) lie inside the stable set.

6. Conclusion

We studied an updated version of the income fluctuation problem, the "common ancestor" of modern macroeconomic theory (Ljungqvist and Sargent (2012), p. 3.)



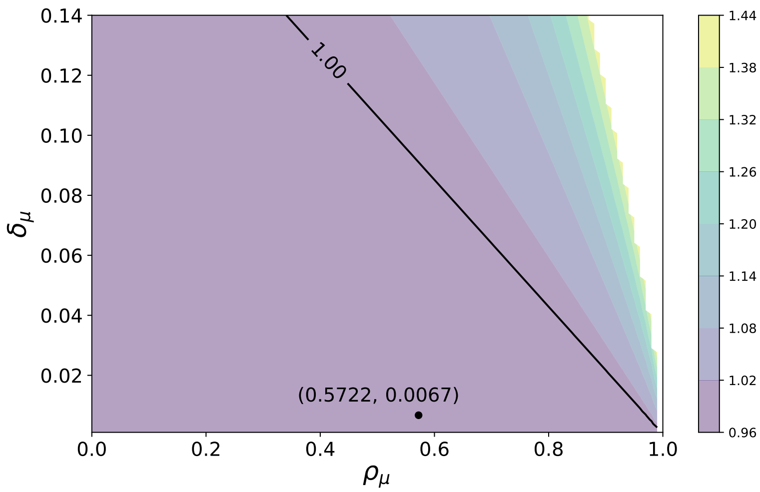
(a) Contour plot of $G_{\beta R}$



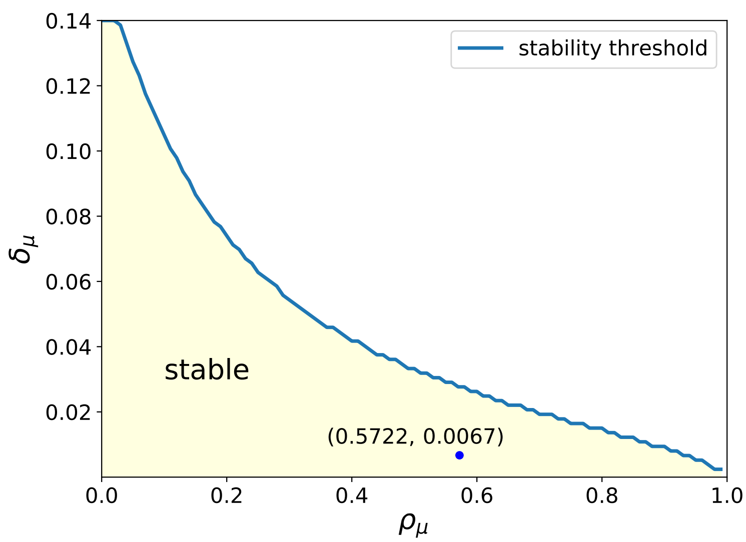
(b) Range and threshold of stability

FIGURE 2. Stability tests for Model I

Working in a setting where returns on financial assets, non-financial income and impatience are all state dependent and fluctuate over time, we obtained conditions under which the household savings problem has a unique solution that can be computed by successive approximations and the wealth process under the optimal savings policy



(a) Contour plot of $G_{\beta R}$



(b) Range and threshold of stability

FIGURE 3. Stability tests for Model II

has a unique stationary distribution with Pareto right tail. We also obtained conditions under which wealth is ergodic and exhibits geometric mixing and asymptotic normality. We investigated the nature of our conditions and provided methods for testing them in applications. While our work was motivated by the desire to better understand the joint distribution of income and wealth, the income fluctuation problem also has applications in asset pricing, life-cycle choice, fiscal policy, monetary policy, optimal taxation, and social security. The ideas contained in this paper should be helpful for those fields after suitable modifications or extensions.

Appendix A. Preliminaries

Given a topological space T, let $\mathscr{B}(\mathsf{T})$ be the Borel σ-algebra and $\mathscr{P}(\mathsf{T})$ be the probability measures on $\mathscr{B}(\mathsf{T})$. A *stochastic kernel* Π on T is a map $\Pi \colon \mathsf{T} \times \mathscr{B}(\mathsf{T}) \to [0,1]$ such that $x \mapsto \Pi(x, B)$ is $\mathscr{B}(\mathsf{T})$-measurable for each $B \in \mathscr{B}(\mathsf{T})$ and $B \mapsto \Pi(x, B)$ is a probability measure on $\mathscr{B}(\mathsf{T})$ for each $x \in \mathsf{T}$. For all $t \in \mathbb{N}$, $x, y \in \mathsf{T}$ and $B \in \mathscr{B}(\mathsf{T})$, we define $\Pi^1 := \Pi$ and $\Pi^t(x, B) := \int \Pi^{t-1}(y, B)\Pi(x, \mathrm{d}y)$. Furthermore, for all $\mu \in \mathscr{P}(\mathsf{T})$, let $(\mu\Pi^t)(B) := \int \Pi^t(x, B)\mu(\mathrm{d}x)$. Π is called *Feller* if $x \mapsto \int h(y)\Pi(x, \mathrm{d}y)$ is continuous on T whenever h is bounded and continuous on T. We call $\psi \in \mathscr{P}(\mathsf{T})$ *stationary* for Π if $\psi\Pi = \psi$.

A sequence $\{\mu_n\} \subset \mathscr{P}(\mathsf{T})$ is called *tight*, if, for all $\varepsilon > 0$, there exists a compact $K \subset \mathsf{T}$ such that $\mu_n(\mathsf{T} \backslash K) \leq \varepsilon$ for all n. A stochastic kernel Π is called *bounded in probability* if the sequence $\{Q^t(x, \cdot)\}_{t \geq 0}$ is tight for all $x \in \mathsf{T}$. Given $\mu \in \mathscr{P}(\mathsf{T})$, we define the total variation norm $\|\mu\|_{TV} := \sup_{g : |g| \leq 1} \left| \int g \, \mathrm{d}\mu \right|$. Given any measurable map $V \colon \mathsf{T} \to [1, \infty)$, we say that Π is *V-geometrically mixing* if there exist constants $M < \infty$ and $\lambda < 1$ such that, for all $x \in \mathsf{T}$ and $t \geq 0$, the corresponding Markov process $\{X_t\}$ satisfies $\sup_{k \geq 0;\, h^2,\, g^2 \leq V} |\mathbb{E}_x g(X_t) h(X_{t+k}) - [\mathbb{E}_x g(X_t)] [\mathbb{E}_x h(X_{t+k})]| \leq \lambda^t M V(x)$.

Below we use $(\Omega, \mathscr{F}, \mathbb{P})$ to denote a fixed probability space on which all random variables are defined. $\mathbb{E}$ is expectations with respect to $\mathbb{P}$. The state process $\{Z_t\}$ and the innovation processes $\{\varepsilon_t\}$, $\{\zeta_t\}$ and $\{\eta_t\}$ introduced in (5) live on this space. In what follows, $\{Z_t\}$ is a stationary version of the chain, where Z_0 is drawn from its unique stationary distribution—henceforth denoted π_Z. The marginal distributions of the innovations are denoted by π_ε, π_ζ and π_η respectively. We let $\{\mathscr{F}_t\}$ be the natural filtration generated by $\{Z_t\}$ and the three innovation processes. $\mathbb{P}_z$ conditions on $Z_0 = z$ and $\mathbb{E}_z$ is expectation under $\mathbb{P}_z$.

We first prove Lemma 4.1, since its implications will be used immediately below. In the proof, we consider the matrix L_φ as a linear operator on $\mathbb{R}^{\mathsf{Z}}$ and identify vectors in $\mathbb{R}^{\mathsf{Z}}$ with real-valued functions on Z.

Proof of Lemma 4.1. A proof by induction confirms that, for any function $h \in \mathbb{R}^{\mathsf{Z}}$,

$$L_\varphi^n h(z) = \mathbb{E}_z \prod_{t=1}^{n} \varphi_t h(Z_t), \tag{32}$$

where L_φ^n is the n-th composition of the operator L_φ with itself (or, equivalently, the n-th power of the matrix L_φ). The positivity of L_φ and Theorem 9.1 of Krasnosel'skii et al. (2012) imply that $r(L_\varphi) = \lim_{n\to\infty} \|L_\varphi^n h\|^{1/n}$ when $\|\cdot\|$ is any norm on $\mathbb{R}^{\mathsf{Z}}$ and h is everywhere positive on Z. With $h \equiv 1$ and $\|f\| = \mathbb{E}|f(Z_0)|$, this becomes

$$r(L_\varphi) = \lim_{n\to\infty} \left(\mathbb{E}\, L_\varphi^n \mathbb{1}(Z_0)\right)^{1/n} = \lim_{n\to\infty} \left(\mathbb{E}\,\mathbb{E}_{Z_0} \prod_{t=1}^{n} \varphi_t\right)^{1/n} = \lim_{n\to\infty} \left(\mathbb{E} \prod_{t=1}^{n} \varphi_t\right)^{1/n} \tag{33}$$

where the second equality is due to (32) and $h = \mathbb{1}$ and the third is by the law of iterated expectations. □

Lemma A.1. *Let $\{\varphi_t\}$ and G_φ be as defined in Lemma 4.1. If $G_\varphi < 1$, then there exists an N in $\mathbb{N}$ and a $\delta < 1$ such that $\max_{z\in\mathsf{Z}} \mathbb{E}_z \prod_{t=1}^{n} \varphi_t < \delta^n$ whenever $n \geq N$.*

Proof. Recalling from the proof of Lemma 4.1 that $r(L_\varphi) = \lim_{n\to\infty} \|L_\varphi^n h\|^{1/n}$ when $\|\cdot\|$ is any norm on $\mathbb{R}^{\mathsf{Z}}$ and h is everywhere positive on Z, we can again take $h \equiv 1$ but now switch to $\|f\| = \max_{z\in\mathsf{Z}} |f(z)|$, so that (33) becomes

$$r(L_\varphi) = \lim_{n\to\infty} \left(\max_{z\in\mathsf{Z}} L_\varphi^n \mathbb{1}(z)\right)^{1/n} = \lim_{n\to\infty} \left(\max_{z\in\mathsf{Z}} \mathbb{E}_z \prod_{t=1}^{n} \varphi_t\right)^{1/n}. \tag{34}$$

Since $r(L_\varphi) = G_\varphi$ and $G_\varphi < 1$, the claim in Lemma A.1 now follows. □

Appendix B. Proof of Section 2 Results

Proof of Proposition 2.1. Pick any $a \geq 0$ and $z \in \mathsf{Z}$. Since $c_t = Y_t$ for all t is dominated by a feasible consumption path, monotonicity of u and the law of iterated expectations give

$$\max\, \mathbb{E}_{a,z} \sum_{t=0}^{\infty} \prod_{i=0}^{t} \beta_i u(c_t) \geq \mathbb{E}_z \sum_{t=0}^{\infty} \prod_{i=0}^{t} \beta_i u(Y_t) = \sum_{t=0}^{\infty} \mathbb{E}_z \prod_{i=0}^{t} \beta_i h(Z_t),$$

where $h(Z_t) := \mathbb{E}_{Z_t} u(Y)$ and the monotone convergence theorem has been employed to pass the expectation through the sum. In view of (32) and $\beta_0 = 1$, we then have

$$\max\, \mathbb{E}_{a,z} \sum_{t=0}^{\infty} \prod_{i=0}^{t} \beta_i u(c_t) \geq \sum_{t=0}^{\infty} L_\beta^t h(z). \tag{35}$$

By the assumed almost sure positivity of β_t and the irreducibility of P, the matrix L_β is irreducible. Hence, by the Perron–Frobenius theorem, we can choose an everywhere positive eigenfunction e such that $L_\beta e = r(L_\beta)e$. By the everywhere positivity of $u(Y_t)$, the function h is everywhere positive on Z, and hence we can choose $\alpha > 0$ such that $e_\alpha := \alpha e$ is less than h pointwise on Z. We then have

$$\sum_{t=0}^{\infty} L_\beta^t \, h(z) \geq \sum_{t=0}^{\infty} L_\beta^t \, e_\alpha(z) = \alpha \sum_{t=0}^{\infty} r(L_\beta)^t \, e(z).$$

By lemma 4.1 we know that $r(L_\beta) \geq 1$, and since α and e are positive, this expression is infinite. Returning to (35), we see that the value function is infinite at our arbitrarily chosen pair (a, z). $\square$

For the rest of this section we suppose that Assumptions 2.1–2.3 hold.

Lemma B.1. *$M_1 := \sum_{t=0}^{\infty} \max_{z \in \mathsf{Z}} \mathbb{E}_z \prod_{i=1}^t \beta_i$ and $M_2 := \sum_{t=0}^{\infty} \max_{z \in \mathsf{Z}} \mathbb{E}_z \prod_{i=1}^t \beta_i R_i$, are finite, as are the constants $M_3 = \max_{z \in \mathsf{Z}} \mathbb{E}_z Y$ and $M_4 = \max_{z \in \mathsf{Z}} \mathbb{E}_z u'(Y)$.*

Proof. That M_1 and M_2 are finite follows directly from Lemma A.1, with $\varphi_t = \beta_t$ and $\varphi_t = \beta_t R_t$ respectively. Regarding M_3, Assumption 2.3 states that $\mathbb{E} Y < \infty$. By the Law of Iterated Expectations, we can write this as $\sum_{z \in \mathsf{Z}} \mathbb{E}_z Y \pi_Z(z) < \infty$. As $\{Z_t\}$ is irreducible, we know that π_Z is positive everywhere on Z. Hence, $M_3 < \infty$ must hold. The proof of $M_4 < \infty$ is similar. $\square$

Lemma B.2. *For the maximal asset path $\{\tilde{a}_t\}$ defined by*

$$\tilde{a}_{t+1} = R_{t+1} \, \tilde{a}_t + Y_{t+1} \quad \textit{and} \quad (\tilde{a}_0, \tilde{z}_0) = (a, z) \textit{ given.} \tag{36}$$

we have, for each $(a, z) \in \mathsf{S}_0$, that $M(a, z) := \sum_{t=0}^{\infty} \mathbb{E}_{a,z} \prod_{i=0}^t \beta_i \, \tilde{a}_t < \infty$.

Proof. Iterating backward on (36), we can show that $\tilde{a}_t = \prod_{i=1}^t R_i \, a + \sum_{j=1}^t Y_j \prod_{i=j+1}^t R_i$. Taking expectation yields

$$\mathbb{E}_{a,z} \prod_{i=0}^t \beta_i \, \tilde{a}_t = \mathbb{E}_z \prod_{i=1}^t \beta_i R_i \, a + \sum_{j=1}^t \mathbb{E}_z \prod_{i=j+1}^t \beta_i R_i \prod_{k=0}^j \beta_k \, Y_j.$$

Then the Monotone Convergence Theorem and the Markov property imply that

$$\begin{aligned} M(a,z) &= \sum_{t=0}^{\infty} \mathbb{E}_z \prod_{i=1}^{t} \beta_i R_i \, a + \sum_{t=0}^{\infty} \sum_{j=1}^{t} \mathbb{E}_z \prod_{i=j+1}^{t} \beta_i R_i \prod_{k=0}^{j} \beta_k \, Y_j \\ &= \mathbb{E}_z \sum_{t=0}^{\infty} \prod_{i=1}^{t} \beta_i R_i \, a + \sum_{j=1}^{\infty} \sum_{i=0}^{\infty} \mathbb{E}_z \prod_{k=0}^{j} \beta_k \, Y_j \prod_{\ell=1}^{i} \beta_{j+\ell} R_{j+\ell} \\ &= \sum_{t=0}^{\infty} \mathbb{E}_z \prod_{i=1}^{t} \beta_i R_i \, a + \sum_{j=1}^{\infty} \mathbb{E}_z \prod_{k=0}^{j} \beta_k \, Y_j \, \mathbb{E}_{Z_j} \sum_{i=0}^{\infty} \prod_{\ell=1}^{i} \beta_\ell R_\ell . \end{aligned}$$

By Lemma B.1, we now have, for all $(a, z) \in \mathsf{S}_0$,

$$M(a,z) \leq M_2 \, a + M_2 \sum_{t=1}^{\infty} \mathbb{E}_z \prod_{i=0}^{t} \beta_i Y_t = M_2 \, a + M_2 \sum_{t=1}^{\infty} \mathbb{E}_z \prod_{i=0}^{t} \beta_i \, \mathbb{E}_{Z_t} Y.$$

Applying Lemma B.1 again gives $M(a, z) < \infty$, as was to be shown. □

Proposition B.1. *The value $V_c(a, z)$ in (8) is well-defined in $\{-\infty\} \cup \mathbb{R}$.*

Proof. By the assumptions on the utility function, there exists a constant $B \in \mathbb{R}_+$ such that $u(c) \leq c + B$, and hence $V_c(a, z) \leq \mathbb{E}_{a,z} \sum_{t=0}^{\infty} \prod_{i=0}^{t} \beta_i \, u(\tilde{a}_t) \leq M(a, z) + B \sum_{t=0}^{\infty} \mathbb{E}_z \prod_{i=0}^{t} \beta_i$. The last term is finite by Lemma A.1. □

Proof of Thoerem 2.1. The proof is a long but relatively straightforward extension of Theorem 1 of Benhabib et al. (2015) and thus omitted. A full proof is available from the authors upon request. □

Proposition B.2. *$(\mathscr{C}, \rho)$ is a complete metric space.*

Proof. The proof is a straightforward extension of Proposition 4.1 of Li and Stachurski (2014) and thus omitted. A full proof is available from the authors upon request. □

Proof of Proposition 2.2. Let c be a policy in $\mathscr{C}$ satisfying (10). To show that any asset path generated by c satisfies the transversality condition (11), observe that, by condition (12), we have

$$c \in \mathscr{C} \implies \exists M \in \mathbb{R}_+ \text{ s.t. } u'(a) \leq (u' \circ c)(a, z) \leq u'(a) + M, \; \forall (a, z) \in \mathsf{S}_0. \tag{37}$$

$$\therefore \quad \mathbb{E}_{a,z} \prod_{i=0}^{t} \beta_i \, (u' \circ c)(a_t, Z_t) a_t \leq \mathbb{E}_{a,z} \prod_{i=0}^{t} \beta_i \, u'(a_t) a_t + M \, \mathbb{E}_{a,z} \prod_{i=0}^{t} \beta_i \, a_t. \tag{38}$$

Regarding the first term on the right hand side of (38), fix $A > 0$ and observe that

$$
\begin{aligned}
u'(a_t)a_t &= u'(a_t)a_t \mathbb{1}\{a_t \leq A\} + u'(a_t)a_t \mathbb{1}\{a_t > A\} \\
&\leq A u'(a_t) + u'(A) a_t \leq A u'(Y_t) + u'(A) \tilde{a}_t
\end{aligned}
$$

with probability one, where $\tilde{a}_t$ is the maximal path defined in (36). We then have

$$
\mathbb{E}_{a,z} \prod_{i=0}^{t} \beta_i \, u'(a_t) a_t \leq A \mathbb{E}_z \prod_{i=0}^{t} \beta_i \, u'(Y_t) + u'(A) \mathbb{E}_{a,z} \prod_{i=0}^{t} \beta_i \, \tilde{a}_t. \tag{39}
$$

By Lemma B.1, we have

$$
A \, \mathbb{E}_z \prod_{i=0}^{t} \beta_i \, u'(Y_t) = A \, \mathbb{E}_z \prod_{i=0}^{t} \beta_i \, \mathbb{E}_{Z_t} u'(Y) \leq M_4 A \, \mathbb{E}_z \prod_{i=0}^{t} \beta_i,
$$

and the last expression converges to zero as $t \to \infty$ by Lemma A.1. The second term in (39) also converges to zero by Lemma B.2. Hence $\mathbb{E}_{a,z} \prod_{i=0}^{t} \beta_i \, u'(a_t) a_t \to 0$ as $t \to \infty$, which, combined with (38) and another application of Lemma B.2, gives our desired result. □

Proposition B.3. *For all $c \in \mathscr{C}$ and $(a, z) \in \mathsf{S}_0$, there exists a unique $\xi \in (0, a]$ that solves (14).*

Proof. Fix $c \in \mathscr{C}$ and $(a, z) \in \mathsf{S}_0$. Because $c \in \mathscr{C}$, the map $\xi \mapsto \psi_c(\xi, a, z)$ is increasing. Since $\xi \mapsto u'(\xi)$ is strictly decreasing, the equation (14) can have at most one solution. Hence uniqueness holds.

Existence follows from the intermediate value theorem provided we can show that

(a) $\xi \mapsto \psi_c(\xi, a, z)$ is a continuous function,
(b) $\exists \xi \in (0, a]$ such that $u'(\xi) \geq \psi_c(\xi, a, z)$, and
(c) $\exists \xi \in (0, a]$ such that $u'(\xi) \leq \psi_c(\xi, a, z)$.

For part (a), it suffices to show that

$$
g(\xi) := \mathbb{E}_z \hat{\beta} \hat{R} \, (u' \circ c) \, [\hat{R}(a - \xi) + \hat{Y}, \hat{Z}]
$$

is continuous on $(0, a]$. To this end, fix $\xi \in (0, a]$ and $\xi_n \to \xi$. By (37) we have

$$
\hat{\beta} \hat{R} \, (u' \circ c) \, [\hat{R} \, (a - \xi) + \hat{Y}, \hat{Z}] \leq \hat{\beta} \hat{R} \, (u' \circ c) \, (\hat{Y}, \hat{Z}) \leq \hat{\beta} \hat{R} u'(\hat{Y}) + \hat{\beta} \hat{R} M. \tag{40}
$$

The last term is integrable, as follows easily from Lemma B.1. Hence the dominated convergence theorem applies. From this fact and the continuity of c, we obtain $g(\xi_n) \to g(\xi)$. Hence, $\xi \mapsto \psi_c(\xi, a, z)$ is continuous.

Part (b) clearly holds, since $u'(\xi) \to \infty$ as $\xi \to 0$ and $\xi \mapsto \psi_c(\xi, a, z)$ is increasing and always finite (since it is continuous as shown in the previous paragraph). Part (c) is also trivial (just set $\xi = a$). □

Proposition B.4. *We have $Tc \in \mathscr{C}$ for all $c \in \mathscr{C}$.*

Proof. Fix $c \in \mathscr{C}$ and let $g(\xi, a, z) := \mathbb{E}_z \hat{\beta} \hat{R} \, (u' \circ c) \, [\hat{R} \, (a - \xi) + \hat{Y}, \, \hat{Z}]$.

Step 1. We show that Tc is continuous. To apply a standard fixed point parametric continuity result such as Theorem B.1.4 of Stachurski (2009), we first show that ψ_c is jointly continuous on the set G defined in (15). This will be true if g is jointly continuous on G. For any $\{(\xi_n, a_n, z_n)\}$ and (ξ, a, z) in G with $(\xi_n, a_n, z_n) \to (\xi, a, z)$, we need to show that $g(\xi_n, a_n, z_n) \to g(\xi, a, z)$. To that end, we define

$$h_1(\xi, a, \hat{Z}, \hat{\varepsilon}, \hat{\zeta}, \hat{\eta}), \, h_2(\xi, a, \hat{Z}, \hat{\varepsilon}, \hat{\zeta}, \hat{\eta}) := \hat{\beta} \hat{R} [u'(\hat{Y}) + M] \pm \hat{\beta} \hat{R} \, (u' \circ c) \, [\hat{R} \, (a - \xi) + \hat{Y}, \hat{Z}],$$

where $\hat{\beta} := \beta(\hat{Z}, \hat{\varepsilon})$, $\hat{R} := R(\hat{Z}, \hat{\zeta})$ and $\hat{Y} := Y(\hat{Z}, \hat{\eta})$ as defined in (5). Then h_1 and h_2 are continuous in $(\xi, a, \hat{Z})$ by the continuity of c and nonnegative by (40).

By Fatou's lemma and Theorem 1.1 of Feinberg et al. (2014),

$$\begin{aligned}
&\iiint \sum_{\hat{z} \in \mathsf{Z}} h_i(\xi, a, \hat{z}, \hat{\varepsilon}, \hat{\zeta}, \hat{\eta}) P(z, \hat{z}) \pi_\varepsilon(\mathrm{d}\hat{\varepsilon}) \pi_\zeta(\mathrm{d}\hat{\zeta}) \pi_\eta(\mathrm{d}\hat{\eta}) \\
&\leq \iiint \liminf_{n \to \infty} \sum_{\hat{z} \in \mathsf{Z}} h_i(\xi_n, a_n, \hat{z}, \hat{\varepsilon}, \hat{\zeta}, \hat{\eta}) P(z_n, \hat{z}) \pi_\varepsilon(\mathrm{d}\hat{\varepsilon}) \pi_\zeta(\mathrm{d}\hat{\zeta}) \pi_\eta(\mathrm{d}\hat{\eta}) \\
&\leq \liminf_{n \to \infty} \iiint \sum_{\hat{z} \in \mathsf{Z}} h_i(\xi_n, a_n, \hat{z}, \hat{\varepsilon}, \hat{\zeta}, \hat{\eta}) P(z_n, \hat{z}) \pi_\varepsilon(\mathrm{d}\hat{\varepsilon}) \pi_\zeta(\mathrm{d}\hat{\zeta}) \pi_\eta(\mathrm{d}\hat{\eta}).
\end{aligned}$$

This implies that

$$\liminf_{n \to \infty} \left(\pm \mathbb{E}_{z_n} \hat{\beta} \hat{R} \, (u' \circ c) \, [\hat{R} \, (a_n - \xi_n) + \hat{Y}, \hat{Z}] \right) \geq \left(\pm \mathbb{E}_z \hat{\beta} \hat{R} \, (u' \circ c) \, [\hat{R} \, (a - \xi) + \hat{Y}, \hat{Z}] \right).$$

The function g is then continuous, since the above inequality is equivalent to the statement $\liminf_{n \to \infty} g(\xi_n, a_n, z_n) \geq g(\xi, a, z) \geq \limsup_{n \to \infty} g(\xi_n, a_n, z_n)$. Hence, ψ_c is continuous on G, as was to be shown. Moreover, since $\xi \mapsto \psi_c(\xi, a, z)$ takes values in the closed interval $I(a, z) := [u'(a), u'(a) + \mathbb{E}_z \hat{\beta} \hat{R} (u'(\hat{Y}) + M)]$, and the correspondence $(a, z) \mapsto I(a, z)$ is nonempty, compact-valued and continuous, Theorem B.1.4 of Stachurski (2009) then implies that Tc is continuous on S_0.

Step 2. We show that Tc is increasing in a. Suppose that for some $z \in \mathsf{Z}$ and $a_1, a_2 \in (0, \infty)$ with $a_1 < a_2$, we have $\xi_1 := Tc(a_1, z) > Tc(a_2, z) =: \xi_2$. Since c

is increasing in a by assumption, ψ_c is increasing in ξ and decreasing in a. Then $u'(\xi_1) < u'(\xi_2) = \psi_c(\xi_2, a_2, z) \leq \psi_c(\xi_1, a_1, z) = u'(\xi_1)$. This is a contradiction.

Step 3. We have shown in Proposition B.3 that $Tc(a, z) \in (0, a]$ for all $(a, z) \in \mathsf{S}_0$.

Step 4. We show that $\|u' \circ (Tc) - u'\| < \infty$. Since $u'[Tc(a, z)] \geq u'(a)$, we have

$$\begin{aligned}&|u'[Tc(a, z)] - u'(a)| = u'[Tc(a, z)] - u'(a)\\&\leq \mathbb{E}_z \hat{\beta}\hat{R}\,(u' \circ c)\,(\hat{R}[a - Tc(a, z)] + \hat{Y},\, \hat{Z}) \leq \mathbb{E}_z \hat{\beta}\hat{R}[u'(\hat{Y}) + M]\end{aligned}$$

for all $(a, z) \in \mathsf{S}_0$. The right hand side is easily shown to be finite via Lemma B.1. ☐

To prove Theorem 2.2, let $\mathscr{H}$ be all continuous functions $h : \mathsf{S}_0 \to \mathbb{R}$ that is decreasing in its first argument and $(a, z) \mapsto h(a, z) - u'(a)$ is bounded and nonnegative. Given $h \in \mathscr{H}$, let $\tilde{T}h$ be the function mapping $(a, z) \in \mathsf{S}_0$ into the κ that solves

$$\kappa = \max\{\mathbb{E}_z\, \hat{\beta}\hat{R}\, h(\hat{R}\,[a - (u')^{-1}(\kappa)] + \hat{Y}, \hat{Z}),\, u'(a)\}. \tag{41}$$

Moreover, consider the bijection $H : \mathscr{C} \to \mathscr{H}$ defined by $Hc := u' \circ c$.

Lemma B.3. *The operator $\tilde{T} \colon \mathscr{H} \to \mathscr{H}$ and satisfies $\tilde{T}H = HT$ on $\mathscr{C}$.*

Proof. Pick any $c \in \mathscr{C}$ and $(a, z) \in \mathsf{S}_0$. Let $\xi := Tc(a, z)$, then ξ solves

$$u'(\xi) = \max\{\mathbb{E}_z\, \hat{\beta}\hat{R}\,(u' \circ c)\,[\hat{R}\,(a - \xi) + \hat{Y}, \hat{Z}],\, u'(a)\}. \tag{42}$$

We need to show that HTc and $\tilde{T}Hc$ evaluate to the same number at (a, z). In other words, we need to show that $u'(\xi)$ is the solution to

$$\kappa = \max\{\mathbb{E}_z\, \hat{\beta}\hat{R}\,(u' \circ c)\,(\hat{R}\,[a - (u')^{-1}(\kappa)] + \hat{Y}, \hat{Z}),\, u'(a)\}.$$

But this is immediate from (42). Hence, we have shown that $\tilde{T}H = HT$ on $\mathscr{C}$. Since $H \colon \mathscr{C} \to \mathscr{H}$ is a bijection, we have $\tilde{T} = HTH^{-1}$. Since in addition $T \colon \mathscr{C} \to \mathscr{C}$ by Proposition B.4, we have $\tilde{T} \colon \mathscr{H} \to \mathscr{H}$. This concludes the proof. ☐

Lemma B.4. *$\tilde{T}$ is order preserving on $\mathscr{H}$. That is, $\tilde{T}h_1 \leq \tilde{T}h_2$ for all $h_1, h_2 \in \mathscr{H}$ with $h_1 \leq h_2$.*

Proof. Let h_1, h_2 be functions in $\mathscr{H}$ with $h_1 \leq h_2$. Suppose to the contrary that there exists $(a, z) \in \mathsf{S}_0$ such that $\kappa_1 := \tilde{T}h_1(a, z) > \tilde{T}h_2(a, z) =: \kappa_2$. Since functions in $\mathscr{H}$

are decreasing in the first argument, we have

$$\begin{aligned}\kappa_1 &= \max\{\mathbb{E}_z\, \hat{\beta}\hat{R}\, h_1(\hat{R}\,[a-(u')^{-1}(\kappa_1)]+\hat{Y},\hat{Z}),\, u'(a)\}\\ &\leq \max\{\mathbb{E}_z\, \hat{\beta}\hat{R}\, h_2(\hat{R}\,[a-(u')^{-1}(\kappa_1)]+\hat{Y},\hat{Z}),\, u'(a)\}\\ &\leq \max\{\mathbb{E}_z\, \hat{\beta}\hat{R}\, h_2(\hat{R}\,[a-(u')^{-1}(\kappa_2)]+\hat{Y},\hat{Z}),\, u'(a)\} = \kappa_2.\end{aligned}$$

This is a contradiction. Hence, $\tilde{T}$ is order preserving. □

Lemma B.5. *There exists an $n \in \mathbb{N}$ and $\theta < 1$ such that $\tilde{T}^n$ is a contraction mapping of modulus θ on $(\mathscr{H}, d_\infty)$.*

Proof. Since $\tilde{T}$ is order preserving and $\mathscr{H}$ is closed under the addition of nonnegative constants, based on Blackwell (1965), it remains to verify the existence of $n \in \mathbb{N}$ and $\theta < 1$ such that $\tilde{T}^n(h+\gamma) \leq \tilde{T}^n h + \theta\gamma$ for all $h \in \mathscr{H}$ and $\gamma \geq 0$. By Lemma A.1 and Assumption 2.2, it suffices to show that for all $k \in \mathbb{N}$ and $(a,z) \in \mathsf{S}_0$, we have

$$\tilde{T}^k(h+\gamma)(a,z) \leq \tilde{T}^k h(a,z) + \gamma\, \mathbb{E}_z \prod_{i=1}^k \beta_i R_i. \tag{43}$$

Fix $h \in \mathscr{H}$, $\gamma \geq 0$, and let $h_\gamma(a,z) := h(a,z)+\gamma$. By the definition of $\tilde{T}$, we have

$$\begin{aligned}\tilde{T}h_\gamma(a,z) &= \max\{\mathbb{E}_z\, \hat{\beta}\hat{R}\, h_\gamma(\hat{R}\,[a-(u')^{-1}(\tilde{T}h_\gamma)(a,z)]+\hat{Y},\hat{Z}), u'(a)\}\\ &\leq \max\{\mathbb{E}_z\, \hat{\beta}\hat{R}\, h(\hat{R}\,[a-(u')^{-1}(\tilde{T}h_\gamma)(a,z)]+\hat{Y},\hat{Z}), u'(a)\} + \gamma\mathbb{E}_z\beta_1 R_1\\ &\leq \max\{\mathbb{E}_z\, \hat{\beta}\hat{R}\, h(\hat{R}\,[a-(u')^{-1}(\tilde{T}h)(a,z)]+\hat{Y},\hat{Z}), u'(a)\} + \gamma\mathbb{E}_z\beta_1 R_1.\end{aligned}$$

Here, the first inequality is elementary and the second is due to the fact that $h \leq h_\gamma$ and $\tilde{T}$ is order preserving. Hence, $\tilde{T}(h+\gamma)(a,z) \leq \tilde{T}h(a,z) + \gamma\mathbb{E}_z\beta_1 R_1$ and (43) holds for $k=1$. Suppose (43) holds for arbitrary k. It remains to show that it holds for $k+1$. For $z \in \mathsf{Z}$, define $f(z) := \gamma\mathbb{E}_z\beta_1 R_1\cdots\beta_k R_k$. By the induction hypothesis, the monotonicity of $\tilde{T}$ and the Markov property,

$$\begin{aligned}\tilde{T}^{k+1}h_\gamma(a,z) &= \max\{\mathbb{E}_z\, \hat{\beta}\hat{R}\, (\tilde{T}^k h_\gamma)(\hat{R}\,[a-(u')^{-1}(\tilde{T}^{k+1}h_\gamma)(a,z)]+\hat{Y},\hat{Z}), u'(a)\}\\ &\leq \max\{\mathbb{E}_z\, \hat{\beta}\hat{R}\, (\tilde{T}^k h + f)(\hat{R}\,[a-(u')^{-1}(\tilde{T}^{k+1}h_\gamma)(a,z)]+\hat{Y},\hat{Z}), u'(a)\}\\ &\leq \max\{\mathbb{E}_z\, \hat{\beta}\hat{R}\, (\tilde{T}^k h)(\hat{R}\,[a-(u')^{-1}(\tilde{T}^{k+1}h_\gamma)(a,z)]+\hat{Y},\hat{Z}), u'(a)\}\\ &\quad + \mathbb{E}_z\, \beta_1 R_1 f(Z_1)\\ &\leq \max\{\mathbb{E}_z\, \hat{\beta}\hat{R}\, (\tilde{T}^k h)(\hat{R}\,[a-(u')^{-1}(\tilde{T}^{k+1}h)(a,z)]+\hat{Y},\hat{Z}), u'(a)\}\\ &\quad + \gamma\mathbb{E}_z\, \beta_1 R_1\, \mathbb{E}_{Z_1}\beta_1 R_1\cdots\beta_k R_k\\ &= \tilde{T}^{k+1}h(a,z) + \gamma\mathbb{E}_z\, \beta_1 R_1\cdots\beta_{k+1}R_{k+1}.\end{aligned}$$

Hence, (43) is verified by induction. This concludes the proof. □

Proof of Theorem 2.2. Let n and θ be as in Lemma B.5. In view of Propositions 2.2, B.2 and B.4, to show that T^n is a contraction and verify claims (1)–(3) of Theorem 2.2, based on the Banach contraction mapping theorem, it suffices to show that $\rho(T^n c, T^n d) \leq \theta \rho(c, d)$ for all $c, d \in \mathscr{C}$. To this end, pick any $c, d \in \mathscr{C}$. Note that the topological conjugacy result established in Lemma B.3 implies that $\tilde{T} = HTH^{-1}$. Hence, $\tilde{T}^n = (HTH^{-1}) \cdots (HTH^{-1}) = HT^nH^{-1}$ and $\tilde{T}^n H = HT^n$. By the definition of ρ and the contraction property established in Lemma B.5,

$$\rho(T^n c, T^n d) = d_\infty(HT^n c, HT^n d) = d_\infty(\tilde{T}^n Hc, \tilde{T}^n Hd) \leq \theta d_\infty(Hc, Hd) = \theta \rho(c, d).$$

Hence, T^n is a contraction and claims (1)–(3) are verified. □

Our next goal is to prove Proposition 2.3. To begin with, we define

$$\mathscr{C}_0 = \{c \in \mathscr{C} \colon a \mapsto a - c(a, z) \text{ is increasing for all } z \in \mathsf{Z}\}.$$

Lemma B.6. *$\mathscr{C}_0$ is a closed subset of $\mathscr{C}$, and $Tc \in \mathscr{C}_0$ for all $c \in \mathscr{C}_0$.*

Proof. To see that $\mathscr{C}_0$ is closed, for a given sequence $\{c_n\}$ in $\mathscr{C}_0$ and $c \in \mathscr{C}$ with $\rho(c_n, c) \to 0$, we need to show that $c \in \mathscr{C}_0$. This obviously holds since $a \mapsto a - c_n(a, z)$ is increasing for all n, and, in addition, $\rho(c_n, c) \to 0$ implies that $c_n(a, z) \to c(a, z)$ for all $(a, z) \in \mathsf{S}_0$.

Fix $c \in \mathscr{C}_0$. We now show that $\xi := Tc \in \mathscr{C}_0$. Since $\xi \in \mathscr{C}$ by Proposition B.4, it remains to show that $a \mapsto a - \xi(a, z)$ is increasing. Suppose the claim is false, then there exist $z \in \mathsf{Z}$ and $a_1, a_2 \in (0, \infty)$ such that $a_1 < a_2$ and $a_1 - \xi(a_1, z) > a_2 - \xi(a_2, z)$. Since $a_1 - \xi(a_1, z) \geq 0$, $a_2 - \xi(a_2, z) \geq 0$ and $\xi(a_1, z) \leq \xi(a_2, z)$ by Proposition B.4, we have $\xi(a_1, z) < a_1$ and $\xi(a_1, z) < \xi(a_2, z)$. However, based on the property of the time iteration operator, we then have

$$\begin{aligned}
(u' \circ \xi)(a_1, z) &= \mathbb{E}_z \hat{\beta} \hat{R} (u' \circ c)(\hat{R}\,[a_1 - \xi(a_1, z)] + \hat{Y}, \hat{Z}) \\
&\leq \mathbb{E}_z \hat{\beta} \hat{R} (u' \circ c)(\hat{R}\,[a_2 - \xi(a_2, z)] + \hat{Y}, \hat{Z}) \leq (u' \circ \xi)(a_2, z),
\end{aligned}$$

which implies that $\xi(a_1, z) \geq \xi(a_2, z)$. This is a contradiction. Hence, $a \mapsto a - \xi(a, z)$ is increasing, and T is a self-map on $\mathscr{C}_0$. □

Proof of Proposition 2.3. Since T maps elements of the closed subset $\mathscr{C}_0$ into itself by Lemma B.6, Theorem 2.2 implies that $c^* \in \mathscr{C}_0$. Hence, the stated claims hold. □

Proof of Proposition 2.4. Let T_j be the time iteration operator for the income process j established in Proposition B.4. It suffices to show $T_1 c \leq T_2 c$ for all $c \in \mathscr{C}$. To see this, note that by Lemma B.4, we have $T_j c_1 \leq T_j c_2$ whenever $c_1 \leq c_2$. Therefore if $T_1 c \leq T_2 c$ for all $c \in \mathscr{C}$, we obtain $T_1 c_1 \leq T_1 c_2 \leq T_2 c_2$. Iterating this starting from any $c \in \mathscr{C}$, by Theorem 2.2, it follows that $c_1^* = \lim_{n\to\infty}(T_1)^n c \leq \lim_{n\to\infty}(T_2)^n c = c_2^*$, completing the proof.

To show that $T_1 c \leq T_2 c$ for any $c \in \mathscr{C}$, take any $(a,z) \in \mathsf{S}_0$ and define $\xi_j = (T_j c)(a,z)$. To show $\xi_1 \leq \xi_2$, suppose on the contrary that $\xi_1 > \xi_2$. Since c is increasing in a and $u'' < 0$ (hence u' is decreasing), it follows from the definition of the time iteration operator in (14)–(16), $Y_1 \leq Y_2$, $u'' < 0$ and the monotonicity of $c \in \mathscr{C}$ that

$$\begin{aligned} u'(\xi_2) > u'(\xi_1) &= \max\{\mathbb{E}_z\, \hat{\beta}\hat{R}\,(u' \circ c)[\hat{R}(a - \xi_1) + \hat{Y}_1, \hat{Z}], u'(a)\} \\ &\geq \max\{\mathbb{E}_z\, \hat{\beta}\hat{R}\,(u' \circ c)[\hat{R}(a - \xi_2) + \hat{Y}_2, \hat{Z}], u'(a)\} = u'(\xi_2), \end{aligned}$$

which is a contradiction. □

To prove Proposition 2.5, we need several lemmas.

Lemma B.7. *For all $c \in \mathscr{C}_0$, there exists a threshold $\bar{a}_c(z)$ such that $Tc(a,z) = a$ if and only if $a \leq \bar{a}_c(z)$. In particular, there exists a threshold $\bar{a}(z)$ such that $c^*(a,z) = a$ if and only if $a \leq \bar{a}(z)$.*

Proof. Recall that, for all $c \in \mathscr{C}_0$, $\xi(a,z) := Tc(a,z)$ solves

$$(u' \circ \xi)\,(a,z) = \max\{\mathbb{E}_z\, \hat{\beta}\hat{R}\,(u' \circ c)\,(\hat{R}\,[a - \xi(a,z)] + \hat{Y}, \hat{Z}), u'(a)\}. \tag{44}$$

For each $z \in \mathsf{Z}$ and $c \in \mathscr{C}_0$, define

$$\bar{a}_c(z) := (u')^{-1}\,[\mathbb{E}_z\, \hat{\beta}\hat{R}\,(u' \circ c)\,(\hat{Y}, \hat{Z})] \quad \text{and} \quad \bar{a}(z) := \bar{a}_{c^*}(z). \tag{45}$$

To prove the first claim, by Lemma B.6, it suffices to show that $\xi(a,z) < a$ implies $a > \bar{a}_c(z)$. This obviously holds since in view of (44), the former implies that

$$u'(a) < \mathbb{E}_z\, \hat{\beta}\hat{R}\,(u' \circ c)\,(\hat{R}\,[a - \xi(a,z)] + \hat{Y}, \hat{Z}) \leq \mathbb{E}_z\, \hat{\beta}\hat{R}\,(u' \circ c)\,(\hat{Y}, \hat{Z}) = u'[\bar{a}_c(z)],$$

which then yields $a > \bar{a}_c(z)$. The second claim follows immediately from the first claim and the fact that $c^* \in \mathscr{C}_0$ is the unique fixed point of T in $\mathscr{C}$. □

Consider a subset $\mathscr{C}_1$ defined by $\mathscr{C}_1 := \{c \in \mathscr{C}_0 \colon a \mapsto c(a,z) \text{ is concave for all } z \in \mathsf{Z}\}$.

Lemma B.8. *$\mathscr{C}_1$ is a closed subset of $\mathscr{C}_0$ and $\mathscr{C}$, and, $Tc \in \mathscr{C}_1$ for all $c \in \mathscr{C}_1$.*

Proof. The first claim is immediate because limits of concave functions are concave. To prove the second claim, fix $c \in \mathscr{C}_1$. We have $Tc \in \mathscr{C}_0$ by Lemma B.6. It remains to show that $a \mapsto \xi(a, z) := Tc(a, z)$ is concave for all $z \in \mathsf{Z}$. Given $z \in \mathsf{Z}$, Lemma B.7 implies that $\xi(a, z) = a$ for $a \leq \bar{a}_c(z)$ and that $\xi(a, z) < a$ for $a > \bar{a}_c(z)$. Since in addition $a \mapsto \xi(a, z)$ is continuous and increasing, to show the concavity of ξ with respect to a, it suffices to show that $a \mapsto \xi(a, z)$ is concave on $(\bar{a}_c(z), \infty)$.

Suppose there exist some $z \in \mathsf{Z}$, $\alpha \in [0, 1]$, and $a_1, a_2 \in (\bar{a}_c(z), \infty)$ such that

$$\xi\left((1-\alpha)a_1 + \alpha a_2,\ z\right) < (1-\alpha)\xi(a_1, z) + \alpha\xi(a_2, z). \tag{46}$$

Let $h(a, z, \hat{\omega}) := \hat{R}\left[a - \xi(a, z)\right] + \hat{Y}$, where $\hat{\omega} := (\hat{R}, \hat{Y})$. Then by Lemma B.7 and noting that consumption is interior, we have

$$\begin{aligned}(u' \circ \xi)\left((1-\alpha)a_1 + \alpha a_2,\ z\right) &= \mathbb{E}_z\, \hat{\beta}\hat{R}\,(u' \circ c)\,\{h[(1-\alpha)a_1 + \alpha a_2,\ z,\ \hat{\omega}], \hat{Z}\} \\ &\leq \mathbb{E}_z\, \hat{\beta}\hat{R}\,(u' \circ c)\,[(1-\alpha)h(a_1, z, \hat{\omega}) + \alpha h(a_2, z, \hat{\omega}), \hat{Z}].\end{aligned}$$

Using condition (17) then yields

$$\begin{aligned}&\xi((1-\alpha)a_1 + \alpha a_2, z) \geq (u')^{-1}\{\mathbb{E}_z\, \hat{\beta}\hat{R}(u' \circ c)[(1-\alpha)h(a_1, z, \hat{\omega}) + \alpha h(a_2, z, \hat{\omega}), \hat{Z}]\} \\ &\geq (1-\alpha)(u')^{-1}\{\mathbb{E}_z\, \hat{\beta}\hat{R}(u' \circ c)[h(a_1, z, \hat{\omega}), \hat{Z}]\} + \alpha(u')^{-1}\{\mathbb{E}_z\, \hat{\beta}\hat{R}(u' \circ c)[h(a_2, z, \hat{\omega}), \hat{Z}]\} \\ &= (1-\alpha)(u')^{-1}\{(u' \circ \xi)(a_1, z)\} + \alpha(u')^{-1}\{(u' \circ \xi)(a_2, z)\} = (1-\alpha)\xi(a_1, z) + \alpha\xi(a_2, z),\end{aligned}$$

which contradicts (46). Hence, $a \mapsto \xi(a, z)$ is concave for all $z \in \mathsf{Z}$. $\square$

Proof of Proposition 2.5. By Theorem 2.2, $T \colon \mathscr{C} \to \mathscr{C}$ is a contraction mapping with unique fixed point c^*. Since $\mathscr{C}_1$ is a closed subset of $\mathscr{C}$ and $T\mathscr{C}_1 \subset \mathscr{C}_1$ by Lemma B.8, we know that $c^* \in \mathscr{C}_1$. The first claim is verified. Regarding the second claim, note that $c^* \in \mathscr{C}_1$ implies that $a \mapsto c^*(a, z)$ is increasing and concave for all $z \in \mathsf{Z}$. Hence, $a \mapsto c^*(a, z)/a$ is a decreasing function for all $z \in \mathsf{Z}$. Since $0 \leq c^*(a, z) \leq a$ for all $(a, z) \in \mathsf{S}_0$, $\alpha(z) := \lim_{a \to \infty} c^*(a, z)/a$ is well-defined and $\alpha(z) \in [0, 1]$. $\square$

Proof of Remark 2.1. For each c in $\mathscr{C}$ concave in its first argument, let $h_c(x, \hat{\omega}) := c(\hat{R}x + \hat{Y}, \hat{z})$, where $\hat{\omega} := (\hat{R}, \hat{Y}, \hat{z})$. Then $x \mapsto h_c(x, \hat{\omega})$ is concave. Based on the generalized Minkowski's inequality (see, e.g., Hardy et al. (1952), page 146, theorem

198), we have

$$
\begin{aligned}
&[\mathbb{E}_z\, \hat{\beta}\hat{R}\, h_c(\alpha x_1 + (1-\alpha)x_2, \hat{\omega})^{-\gamma}]^{-\frac{1}{\gamma}} \geq \{\mathbb{E}_z\, \hat{\beta}\hat{R}\, [\alpha h_c(x_1, \hat{\omega}) + (1-\alpha) h_c(x_2, \hat{\omega})]^{-\gamma}\}^{-\frac{1}{\gamma}} \\
&= \{\mathbb{E}_z[\, \alpha(\hat{\beta}\hat{R})^{-\frac{1}{\gamma}} h_c(x_1, \hat{\omega}) + (1-\alpha)(\hat{\beta}\hat{R})^{-\frac{1}{\gamma}} h_c(x_2, \hat{\omega})\,]^{-\gamma}\}^{-\frac{1}{\gamma}} \\
&\geq (\mathbb{E}_z[\alpha(\hat{\beta}\hat{R})^{-\frac{1}{\gamma}} h_c(x_1, \hat{\omega})]^{-\gamma})^{-\frac{1}{\gamma}} + (\mathbb{E}_z[(1-\alpha)(\hat{\beta}\hat{R})^{-\frac{1}{\gamma}} h_c(x_2, \hat{\omega})]^{-\gamma})^{-\frac{1}{\gamma}} \\
&= \alpha[\mathbb{E}_z\, \hat{\beta}\hat{R}\, h_c(x_1, \hat{\omega})^{-\gamma}]^{-\frac{1}{\gamma}} + (1-\alpha)[\mathbb{E}_z\, \hat{\beta}\hat{R}\, h_c(x_2, \hat{\omega})^{-\gamma}]^{-\frac{1}{\gamma}},
\end{aligned}
$$

Since $u'(c) = c^{-\gamma}$, the above inequality implies that condition (17) holds. □

To prove Proposition 2.6, let $\bar{s}$ be as in (19) and define

$$
\mathscr{C}_2 := \{c \in \mathscr{C} : c(a,z) \geq (1-\bar{s})a \;\text{ for all } (a,z) \in \mathsf{S}_0\}\,. \tag{47}
$$

Lemma B.9. *$\mathscr{C}_2$ is a closed subset of $\mathscr{C}$, and $Tc \in \mathscr{C}_2$ for all $c \in \mathscr{C}_2$.*

Proof. To see that $\mathscr{C}_2$ is closed, for a given sequence $\{c_n\}$ in $\mathscr{C}_2$ and $c \in \mathscr{C}$ with $\rho(c_n, c) \to 0$, we need to verify that $c \in \mathscr{C}_2$. This obviously holds since $c_n(a,z)/a \geq 1 - \bar{s}$ for all n and $(a,z) \in \mathsf{S}_0$, and, on the other hand, $\rho(c_n, c) \to 0$ implies that $c_n(a,z) \to c(a,z)$ for all $(a,z) \in \mathsf{S}_0$.

We next show that T is a self-map on $\mathscr{C}_2$. Fix $c \in \mathscr{C}_2$. We have $Tc \in \mathscr{C}$ since T is a self-map on $\mathscr{C}$. It remains to show that $\xi := Tc$ satisfies $\xi(a,z) \geq (1-\bar{s})a$ for all $(a,z) \in \mathsf{S}_0$. Suppose $\xi(a,z) < (1-\bar{s})a$ for some $(a,z) \in \mathsf{S}_0$. Then

$$
u'((1-\bar{s})a) < (u' \circ \xi)(a,z) = \max\{\mathbb{E}_z\, \hat{\beta}\hat{R}\, (u' \circ c)\, (\hat{R}\, [a - \xi(a,z)] + \hat{Y}, \hat{Z}),\, u'(a)\}.
$$

Since $u'((1-\bar{s})a) > u'(a)$ and $c \in \mathscr{C}_2$, this implies that

$$
\begin{aligned}
u'((1-\bar{s})a) &< \mathbb{E}_z\, \hat{\beta}\hat{R}\, (u' \circ c)\, (\hat{R}\, [a - \xi(a,z)] + \hat{Y}, \hat{Z}) \\
&\leq \mathbb{E}_z\, \hat{\beta}\hat{R}\, u'\, \{(1-\bar{s})\hat{R}\, [a - \xi(a,z)] + (1-\bar{s})\hat{Y}\} \\
&\leq \mathbb{E}_z\, \hat{\beta}\hat{R}\, u'\, [(1-\bar{s})\hat{R}\bar{s}a + (1-\bar{s})\hat{Y}] \leq \mathbb{E}_z\, \hat{\beta}\hat{R}\, u'\, [\hat{R}\bar{s}(1-\bar{s})a],
\end{aligned}
$$

which contradicts (19) since $((1-\bar{s})a, z) \in \mathsf{S}_0$. As a result, $\xi(a,z) \geq (1-\bar{s})a$ for all $(a,z) \in \mathsf{S}_0$ and we conclude that $Tc \in \mathscr{C}_2$. □

Proof of Proposition 2.6. We have shown in Theorem 2.2 that T is a contraction mapping on the complete metric space $(\mathscr{C}, \rho)$, with unique fixed point c^*. Since in addition $\mathscr{C}_2$ is a closed subset of $\mathscr{C}$ and $T\mathscr{C}_2 \subset \mathscr{C}_2$ by Lemma B.9, we know that $c^* \in \mathscr{C}_2$. The stated claim is verified. □

Appendix C. Proof of Section 3 Results

As before, Assumptions 2.1–2.3 are in force. Notice that Assumption 2.2, Assumption 3.1 and Lemma A.1 imply existence of an n in $\mathbb{N}$ such that

$$\theta := \max_{z \in \mathsf{Z}} \mathbb{E}_z \prod_{t=1}^{n} \beta_t R_t < 1 \quad \text{and} \quad \gamma := \bar{s}^n \max_{z \in \mathsf{Z}} \mathbb{E}_z \prod_{t=1}^{n} R_t < 1. \tag{48}$$

Lemma C.1. *For all* $(a, z) \in \mathsf{S}$*, we have* $\sup_{t \geq 0} \mathbb{E}_{a,z}\, a_t < \infty$.

Proof. Since $c^*(0, z) = 0$, Proposition 2.6 implies that $c^*(a, z) \geq (1 - \bar{s})a$ for all $(a, z) \in \mathsf{S}$. For all $t \geq 1$, we have $t = kn + j$ in general, where the integers $k \geq 0$ and $j \in \{0, 1, \ldots, n-1\}$. Using these facts and (4), we have:

$$\begin{aligned} a_t &\leq \bar{s}^t R_t \cdots R_1 a + \bar{s}^{t-1} R_t \cdots R_2 Y_1 + \cdots + \bar{s} R_t Y_{t-1} + Y_t \\ &= \bar{s}^{kn+j} R_{kn+j} \cdots R_1 a + \sum_{\ell=1}^{j} \bar{s}^{kn+j-\ell} R_{kn+j} \cdots R_{\ell+1} Y_\ell \\ &\quad + \sum_{m=1}^{k} \sum_{\ell=1}^{n} \bar{s}^{mn-\ell} R_{kn+j} \cdots R_{(k-m)n+j+\ell+1} Y_{(k-m)n+j+\ell} \end{aligned}$$

with probability one. Taking expectations of the above while noting that $M_0 := \max_{1 \leq \ell \leq n,\, z \in \mathsf{Z}} \mathbb{E}_z \prod_{t=1}^{\ell} R_t < \infty$ by Assumption 3.1 and Lemma A.1, we have

$$\begin{aligned} \mathbb{E}_{a,z} a_t &\leq \gamma^k \bar{s}^j \mathbb{E}_z R_j \cdots R_1 a + \gamma^k \sum_{\ell=1}^{j} \bar{s}^{j-\ell} \mathbb{E}_z R_j \cdots R_{\ell+1} Y_\ell \\ &\quad + \sum_{m=0}^{k-1} \gamma^m \sum_{\ell=1}^{n} \bar{s}^{n-\ell} \mathbb{E}_z R_{(k-m)n+j} \cdots R_{(k-m-1)n+j+\ell+1} Y_{(k-m)n+j+\ell} \\ &\leq \gamma^k M_0 a + \gamma^k M_0 \sum_{\ell=1}^{j} \mathbb{E}_z Y_\ell + \sum_{m=0}^{k-1} \gamma^m M_0 \sum_{\ell=1}^{n} \mathbb{E}_z Y_{(k-m-1)n+j+\ell} \\ &\leq M_0 a + M_0 M_3 n + \sum_{m=0}^{\infty} \gamma^m M_0 M_3 n < \infty. \end{aligned}$$

or all $(a, z) \in \mathsf{S}$ and $t \geq 0$. Here we have used M_3 in Lemma B.1 and the Markov property. Hence, $\sup_{t \geq 0} \mathbb{E}_{a,z}\, a_t < \infty$ for all $(a, z) \in \mathsf{S}$, as was claimed. $\square$

A function $w^* \colon \mathsf{S} \to \mathbb{R}_+$ is called *norm-like* if all its sublevel sets (i.e., sets of the form $\{x \in \mathsf{S} \colon w(x) \leq b\}, b \in \mathbb{R}_+$) are precompact in S (i.e., any sequence in a given sublevel set has a subsequence that converges to a point of S).

Proof of Theorem 3.1. Based on Lemma D.5.3 of Meyn and Tweedie (2009), a stochastic kernel Q is bounded in probability if and only if for all $x \in \mathsf{S}$, there exists a norm-like function $w_x^* \colon \mathsf{S} \to \mathbb{R}_+$ such that the (Q, x)-Markov process $\{X_t\}_{t \geq 0}$ satisfies $\limsup_{t \to \infty} \mathbb{E}_x \left[w_x^*(X_t)\right] < \infty$. Fix $(a, z) \in \mathsf{S}$. Since Z is finite, P is bounded in probability. Hence, there exists a norm-like function $w \colon \mathsf{Z} \to \mathbb{R}_+$ such that $\limsup_{t \to \infty} \mathbb{E}_z w(Z_t) < \infty$. Then $w^* \colon \mathsf{S} \to \mathbb{R}_+$ defined by $w^*(a_0, Z_0) := a_0 + w(Z_0)$ is a norm-like function on S. The stochastic kernel Q is then bounded in probability since Lemma C.1 implies that $\limsup_{t \to \infty} \mathbb{E}_{a,z} \, w^*(a_t, Z_t) \leq \sup_{t \geq 0} \mathbb{E}_{a,z} \, a_t + \limsup_{t \to \infty} \mathbb{E}_z \, w(Z_t) < \infty$. Regarding existence of stationary distribution, since P is Feller (due to the finiteness of Z), whenever $z_n \to z$, the product measure satisfies

$$P(z_n, \cdot) \otimes \pi_\zeta \otimes \pi_\eta \xrightarrow{w} P(z, \cdot) \otimes \pi_\zeta \otimes \pi_\eta.$$

Since in addition c^* is continuous, a simple application of the generalized Fatou's lemma of Feinberg et al. (2014) (Theorem 1.1) shows that the stochastic kernel Q is Feller. Moreover, since Q is bounded in probability, based on the Krylov-Bogolubov theorem (see, e.g., Meyn and Tweedie (2009), Proposition 12.1.3 and Lemma D.5.3), Q admits at least one stationary distribution. $\square$

Lemma C.2. *The borrowing constraint binds in finite time with positive probability. That is, for all $(a, z) \in \mathsf{S}$, we have $\mathbb{P}_{a,z} \left(\cup_{t \geq 0} \{c_t = a_t\}\right) > 0$.*

Proof. The claim holds trivially when $a = 0$. Suppose the claim does not hold on S_0 (recall that $\mathsf{S}_0 = \mathsf{S} \backslash \{0\}$), then $\mathbb{P}_{a,z} \left(\cap_{t \geq 0} \{c_t < a_t\}\right) = 1$ for some $(a, z) \in \mathsf{S}_0$, i.e., the borrowing constraint never binds with probability one. Hence,

$$\mathbb{P}_{a,z} \left\{(u' \circ c^*)(a_t, Z_t) = \mathbb{E} \left[\beta_{t+1} R_{t+1} (u' \circ c^*)(a_{t+1}, Z_{t+1}) \middle| \mathscr{F}_t\right]\right\} = 1$$

for all $t \geq 0$. Then we have

$$\begin{aligned}(u' \circ c^*)(a, z) &= \mathbb{E}_{a,z} \, \beta_1 R_1 \cdots \beta_t R_t \, (u' \circ c^*)(a_t, Z_t) \\ &\leq \mathbb{E}_{a,z} \, \beta_1 R_1 \cdots \beta_t R_t \left[u'(a_t) + M\right] \leq \mathbb{E}_z \, \beta_1 R_1 \cdots \beta_t R_t \left[u'(Y_t) + M\right] \quad (49)\end{aligned}$$

for all $t \geq 1$. Let n and θ be defined by (48). Let $t = kn + 1$. Based on the Markov property and Lemma B.1, as $k \to \infty$,

$$\begin{aligned}&\mathbb{E}_z \beta_1 R_1 \cdots \beta_t R_t = \mathbb{E}_z \beta_1 R_1 \cdots \beta_{t-1} R_{t-1} \mathbb{E}_{Z_{t-1}} \beta_1 R_1 \\ &\leq \left(\max_{z \in \mathsf{Z}} \mathbb{E}_z \beta_1 R_1\right) \left(\mathbb{E}_z \beta_1 R_1 \cdots \beta_{nk} R_{nk}\right) \leq \left(\max_{z \in \mathsf{Z}} \mathbb{E}_z \beta_1 R_1\right) \theta^k \to 0.\end{aligned}$$

Similarly, as $k \to \infty$,

$$\mathbb{E}_z\, \beta_1 R_1 \cdots \beta_t R_t u'(Y_t) = \mathbb{E}_z\, \beta_1 R_1 \cdots \beta_{t-1} R_{t-1} \mathbb{E}_{Z_{t-1}} \left[\beta_1 R_1 u'(Y_1)\right]$$
$$\leq \left[\max_{z \in \mathsf{Z}} \mathbb{E}_z \hat{\beta} \hat{R} u'(\hat{Y})\right] \mathbb{E}_z \beta_1 R_1 \cdots \beta_{nk} R_{nk} \leq \left[\max_{z \in \mathsf{Z}} \mathbb{E}_z \hat{\beta} \hat{R} u'(\hat{Y})\right] \theta^k \to 0.$$

Letting $k \to \infty$. (49) then implies that $(u' \circ c^*)(a, z) \leq 0$, contradicted with the fact that $u' > 0$. Thus, we must have $\mathbb{P}_{a,z}\left(\cup_{t \geq 0}\{c_t = a_t\}\right) > 0$ for all $(a, z) \in \mathsf{S}$. $\square$

Our next goal is to prove Theorem 3.2. In proofs we apply the theory of Meyn and Tweedie (2009). Important definitions (their information in the textbook) include: ψ-irreducibility (Section 4.2), small set (page 102), strong aperiodicity (page 114), petite set (page 117), Harris chain (page 199), and positivity (page 230).

Recall that $\mathbb{R}^m$ paired with its Euclidean topology is a second countable topological space (i.e., its topology has a countable base). Since $\mathbb{R}_+$ and Z are respectively Borel subsets of $\mathbb{R}$ and $\mathbb{R}^m$ paired with the relative topologies, they are also second countable. Hence, $\mathsf{S} := \mathbb{R}_+ \times \mathsf{Z}$ satisfies $\mathscr{B}(\mathsf{S}) = \mathscr{B}(\mathbb{R}_+) \otimes \mathscr{B}(\mathsf{Z})$ (see, e.g., page 149, Theorem 4.44 of Aliprantis and Border (2006)). Recall (22). With slight abuse of notation, in proofs, we use f to denote the density of $\{Y_t\}$ in both cases (Y1) and (Y2) and write $\mathrm{d}y = \nu(\mathrm{d}y)$, where ν is the related measure. Specifically, ν is the Lebesgue measure when (Y2) holds. Moreover, Let ϑ be the counting measure.

Recall $\bar{z} \in \mathsf{Z}$ and the greatest lower bound $y_\ell \geq 0$ of the support of $\{Y_t\}$ given by Assumption 3.2. Let $\bar{p} := P(\bar{z}, \bar{z})$. Then $\bar{p} > 0$ by Assumption 3.2.

Lemma C.3. $\mathbb{P}_{(a,\bar{z})}\left\{\cup_{t \geq 0}\left[\{c_t = a_t\} \cap \left(\cap_{i=0}^t \{Z_i = \bar{z}\}\right)\right]\right\} > 0$ *for all* $a \in (0, \infty)$.

Proof. Fix $a \in (0, \infty)$. If $a \leq \bar{a}(\bar{z})$, the claim holds trivially by Lemma B.7. Now consider the case $a > \bar{a}(\bar{z})$. Suppose $\mathbb{P}_{(a,\bar{z})}\left\{\cup_{t \geq 0}\left[\{c_t = a_t\} \cap \left(\cap_{i=0}^t \{Z_i = \bar{z}\}\right)\right]\right\} = 0$. Then, based on the De Morgan's law, we have

$$\mathbb{P}_{(a,\bar{z})}\left\{\cap_{t \geq 0}\left[\{c_t < a_t\} \cup \left(\cup_{i=0}^t \{Z_i \neq \bar{z}\}\right)\right]\right\} = 1.$$
$$\therefore \quad \mathbb{P}_{(a,\bar{z})}\left\{\{c_t < a_t\} \cup \left(\cup_{i=0}^t \{Z_i \neq \bar{z}\}\right)\right\} = 1 \text{ for all } t \in \mathbb{N}.$$
$$\therefore \quad \mathbb{P}_{(a,\bar{z})}\left\{\{c_k < a_k\} \cup \left(\cup_{i=0}^t \{Z_i \neq \bar{z}\}\right)\right\} = 1 \text{ for all } k, t \in \mathbb{N} \text{ with } k \leq t.$$
$$\therefore \quad \mathbb{P}_{(a,\bar{z})}\left\{\left(\cap_{i=0}^t \{c_i < a_i\}\right) \cup \left(\cup_{i=0}^t \{Z_i \neq \bar{z}\}\right)\right\} = 1 \text{ for all } t \in \mathbb{N}.$$

Note that the set $\triangle(t) := (\cap_{i=0}^t \{c_i < a_i\}) \cup (\cup_{i=0}^t \{Z_i \neq \bar{z}\})$ can be written as

$$\triangle(t) = \triangle_1(t) \cup \triangle_2(t), \quad \text{ where } \triangle_1(t) \cap \triangle_2(t) = \emptyset,$$
$$\triangle_1(t) := \left(\cap_{i=0}^t \{c_i < a_i\}\right) \cap \left(\cap_{i=0}^t \{Z_i = \bar{z}\}\right) \text{ and } \triangle_2(t) := \cup_{i=0}^t \{Z_i \neq \bar{z}\}.$$

Assumption 3.2 then implies that, for all $t \geq 0$,

$$\mathbb{P}_{(a,\bar{z})}\{\triangle_1(t)\} = 1 - \mathbb{P}_{\bar{z}}\{\triangle_2(t)\} = \mathbb{P}_{\bar{z}}\left\{\cap_{i=0}^{t}\{Z_i = \bar{z}\}\right\} = \bar{p}^t > 0.$$

Let n and θ be defined by (48) and let $t = kn+1$. Similar to the proof of Lemma B.7, we can show that, with probability $\bar{p}^t > 0$,

$$(u' \circ c^*)(a, \bar{z}) \leq \theta^k \left[\max_{z \in \mathsf{Z}} \mathbb{E}_z \hat{\beta}\hat{R}u'(\hat{Y}) + M \max_{z \in \mathsf{Z}} \mathbb{E}_z \hat{\beta}\hat{R}\right]$$

for some constant $M \in \mathbb{R}_+$. Since $\theta \in (0,1)$ and $(u' \circ c^*)(a, \bar{z}) > 0$, Lemma B.1 implies that there exists $N \in \mathbb{N}$ such that

$$\theta^N \left[\max_{z \in \mathsf{Z}} \mathbb{E}_z \hat{\beta}\hat{R}u'(\hat{Y}) + M \max_{z \in \mathsf{Z}} \mathbb{E}_z \hat{\beta}\hat{R}\right] < (u' \circ c^*)(a, \bar{z}).$$

As a result, we have $(u' \circ c^*)(a, \bar{z}) < (u' \circ c^*)(a, \bar{z})$ with probability $\bar{p}^{Nn+1} > 0$. This is a contradiction. Hence the stated claim is verified. $\square$

Let $F(\mathrm{d}a_{t+1} \mid a_t, Z_t, Z_{t+1})$ be defined such that $\mathbb{P}\{a_{t+1} \in A \mid (a_t, Z_t, Z_{t+1}) = (a, z, z')\} = \int \mathbb{1}\{a' \in A\} F(\mathrm{d}a' \mid a, z, z')$ at $A \in \mathscr{B}(\mathbb{R}_+)$.

Lemma C.4. *Let $h : \mathsf{S} \to \mathbb{R}_+$ be an integrable map such that $a \mapsto h(a, z)$ is decreasing for all $z \in \mathsf{Z}$. Then, for all $t \in \mathbb{N}$ and $z \in \mathsf{Z}$, the map $a \mapsto \ell(a, z, t) := \int h(a', z') Q^t((a, z), \mathrm{d}(a', z'))$ is decreasing.*

Proof. Fix $z \in \mathsf{Z}$. When $t = 1$, (21a) implies that

$$\ell(a, z, 1) = \int \left[\int h(a', z') F(\mathrm{d}a' \mid a, z, z')\right] P(z, z') \vartheta(\mathrm{d}z').$$

Since $a \mapsto h(a, z)$ is decreasing, and by Proposition 2.3 and (21a), the optimal asset accumulation path a_{t+1} is increasing in a_t with probability one, we know that $a \mapsto \int h(a', z') F(\mathrm{d}a' \mid a, z, z')$ is decreasing for all $z' \in \mathsf{Z}$. Thus, $a \mapsto \ell(a, z, 1)$ is decreasing. The claim holds for $t = 1$. Suppose this claim holds for arbitrary t, it remains to show that it holds for $t + 1$. Note that

$$\begin{aligned}
\ell(a, z, t+1) &= \iint h(a'', z'') Q^t((a', z'), \mathrm{d}(a'', z'')) Q((a, z), \mathrm{d}(a', z')) \\
&= \int \ell(a', z', t) Q((a, z), \mathrm{d}(a', z')).
\end{aligned}$$

Since $a' \mapsto \ell(a', z', t)$ is decreasing for all $z' \in \mathsf{Z}$, based on the induction argument, $a \mapsto \ell(a, z, t+1)$ is decreasing. The stated claim then follows. $\square$

Lemma C.5. *The Markov process $\{(a_t, Z_t)\}_{t \geq 0}$ is ψ-irreducible.*

Proof. Recall $\delta > y_\ell$ given by Assumption 3.2. Let $\mathsf{D} \in \mathscr{B}(\mathsf{S})$ be defined by $\mathsf{D} := \{y_\ell\} \times \{\bar{z}\}$ if (Y1) holds and $\mathsf{D} := (y_\ell, \delta) \times \{\bar{z}\}$ if (Y2) holds. We define the measure φ on $\mathscr{B}(\mathsf{S})$ by $\varphi(A) := (\nu \times \vartheta)(A \cap \mathsf{D})$ for $A \in \mathscr{B}(\mathsf{S})$. Clearly φ is a nontrivial measure. In particular, $\vartheta(\{\bar{z}\}) = 1$ as ϑ is the counting measure. Moreover, since y_ℓ is the greatest lower bound of the support of $\{Y_t\}$, it must be the case that $\nu(\{y_\ell\}) > 0$ if (Y1) holds and that $\nu((y_\ell, \delta)) > 0$ if (Y2) holds. As a result, $\varphi(\mathsf{S}) = \nu(\{y_\ell\}) \times \vartheta(\{\bar{z}\}) > 0$ when (Y1) holds and $\varphi(\mathsf{S}) = \nu((y_\ell, \delta)) \times \vartheta(\{\bar{z}\}) > 0$ when (Y2) holds.

We first show that $\{(a_t, Z_t)\}$ is φ-irreducible. Let A be an element of $\mathscr{B}(\mathsf{S})$ such that $\varphi(A) > 0$. Fix $(a, z) \in \mathsf{S}$. We need to show that $\{(a_t, Z_t)\}$ visits set A in finite time with positive probability.

Since $\{z_t\}$ is irreducible, $\mathbb{P}_z\{Z_{N_0} = \bar{z}\} > 0$ for some integer $N_0 \geq 0$. By Lemma C.1, there exists $\tilde{a} < \infty$ such that $\mathbb{P}_{(a,z)}\{a_{N_0} < \tilde{a}, z_{N_0} = \bar{z}\} > 0$. By Lemma C.3, there exists $T \in \mathbb{N}$ such that $\mathbb{P}_{(\tilde{a},\bar{z})}\left\{c_T = a_T,\, Z_T = \bar{z}\right\} \geq \mathbb{P}_{(\tilde{a},\bar{z})}\left\{c_T = a_T,\, \cap_{i=0}^{T}\{Z_i = \bar{z}\}\right\} > 0$. Lemma B.7 and Lemma C.4 then imply that $\mathbb{P}_{(a',\bar{z})}\left\{c_T = a_T,\, Z_T = \bar{z}\right\} > 0$ for all $a' \in (0, \tilde{a})$. Hence, for $N := N_0 + T$ and $E := \{c_N = a_N,\, Z_N = \bar{z}\}$, we have

$$\mathbb{P}_{(a,z)}(E) \geq \int_{\{a' \leq \tilde{a},\, z' = \bar{z}\}} \mathbb{P}_{(a',\bar{z})}\{c_T = a_T, Z_T = \bar{z}\} Q^{N_0}((a,z), \mathrm{d}(a', z')) > 0 \tag{50}$$

based on the Markov property. By (21a), we have

$$\begin{aligned}
\mathbb{P}_{(a,z)}\{(a_{N+1}, Z_{N+1}) \in A\} &\geq \mathbb{P}_{(a,z)}\left\{(a_{N+1}, Z_{N+1}) \in A,\, a_N = c_N,\, Z_N = \bar{z}\right\} \\
&= \mathbb{P}_{(a,z)}\left\{(a_{N+1}, Z_{N+1}) \in A \mid a_N = c_N,\, Z_N = \bar{z}\right\} \mathbb{P}_{(a,z)}(E) \\
&= \mathbb{P}_{(a,z)}\left\{(Y_{N+1}, Z_{N+1}) \in A,\, a_N = c_N,\, Z_N = \bar{z}\right\}. \qquad (51)
\end{aligned}$$

Note that, by Assumption 3.2, $f(y'' \mid z'')P(\bar{z}, z'') > 0$ whenever $(y'', z'') \in \mathsf{D}$. Since in addition $\varphi(A) = (\nu \times \vartheta)(A \cap \mathsf{D}) > 0$, we have

$$\int_A f(y'' \mid z'')P(\bar{z}, z'')(\nu \times \vartheta)[\mathrm{d}(y'', z'')] > 0.$$

Let $\triangle := \mathbb{P}_{(a,z)}\{(a_{N+1}, Z_{N+1}) \in A\}$. Then (50) and (51) imply that

$$\triangle \geq \int_E \left\{\int_A f(y'' \mid z'')P(z', z'')(\nu \times \vartheta)[\mathrm{d}(y'', z'')]\right\} Q^N\left((a,z), \mathrm{d}(a', z')\right) > 0.$$

Therefore, we have shown that any measurable subset with positive φ measure can be reached in finite time with positive probability, i.e., $\{(a_t, Z_t)\}$ is φ-irreducible. Based on Proposition 4.2.2 of Meyn and Tweedie (2009), there exists a maximal probability measure ψ on $\mathscr{B}(\mathsf{S})$ such that $\{(a_t, Z_t)\}$ is ψ-irreducible. $\square$

Lemma C.6. *Let the function $\bar{a}$ be defined as in (45). Then $\bar{a}(\bar{z}) \geq y_\ell$ if (Y1) holds, while $\bar{a}(\bar{z}) > y_\ell$ if (Y2) holds.*

Proof. Suppose (Y1) holds and $\bar{a}(\bar{z}) < y_\ell$. Then, by Lemma B.7, for all $t \in \mathbb{N}$,

$$\begin{aligned}\left[\{c_t = a_t\} \cap \left(\cap_{i=0}^t \{Z_i = \bar{z}\}\right)\right] &= \left[\{a_t \leq \bar{a}(Z_t)\} \cap \left(\cap_{i=0}^t \{Z_i = \bar{z}\}\right)\right] \\ &\subset \left[\{a_t < y_\ell\} \cap \left(\cap_{i=0}^t \{Z_i = \bar{z}\}\right)\right] \subset \{a_t < y_\ell\}. \end{aligned} \tag{52}$$

Hence, for all $a \in (0, \infty)$ and $t \in \mathbb{N}$,

$$\mathbb{P}_{(a,\bar{z})}\left[\{c_t = a_t\} \cap \left(\cap_{i=0}^t \{Z_i = \bar{z}\}\right)\right] \leq \mathbb{P}_{(a,\bar{z})}\{a_t < y_\ell\} = 0,$$

where the last equality follows from (21a), which implies that $a_t \geq Y_t \geq y_\ell$ with probability one. This is contradicted with Lemma C.3.

Suppose (Y2) holds and $\bar{a}(\bar{z}) \leq y_\ell$. By definition, $\mathbb{P}_z\{Y_t \leq y_\ell\} = 0$ for all $z \in \mathsf{Z}$ and $t \in \mathbb{N}$. Since $a_t \geq Y_t$ with probability one, we have $\mathbb{P}_{(a,z)}\{a_t \leq y_\ell\} = 0$ for all $(a, z) \in \mathsf{S}$ and $t \in \mathbb{N}$. Via similar analysis to (52), Lemma B.7 implies that $[\{c_t = a_t\} \cap (\cap_{i=0}^t \{Z_i = \bar{z}\})] \subset \{a_t \leq y_\ell\}$ for all $t \in \mathbb{N}$. Hence, for all $a \in (0, 1)$ and $t \in \mathbb{N}$, we have $\mathbb{P}_{(a,\bar{z})}[\{c_t = a_t\} \cap (\cap_{i=0}^t \{Z_i = \bar{z}\})] \leq \mathbb{P}_{(a,\bar{z})}\{a_t \leq y_\ell\} = 0$. Again, this contradicts Lemma C.3. □

Lemma C.7. *The Markov process $\{(a_t, Z_t)\}_{t \geq 0}$ is strongly aperiodic.*

Proof. By the definition of strong aperiodicity, we need to show that there exists a v_1-small set D_1 with $v_1(\mathsf{D}_1) > 0$, i.e., there exists a nontrivial measure v_1 on $\mathscr{B}(\mathsf{S})$ and a subset $\mathsf{D}_1 \in \mathscr{B}(\mathsf{S})$ such that $v_1(\mathsf{D}_1) > 0$ and

$$\inf_{(a,z) \in \mathsf{D}_1} Q\left((a, z), A\right) \geq v_1\left(A\right) \quad \text{for all } A \in \mathscr{B}(\mathsf{S}). \tag{53}$$

For $\delta > 0$ given by Assumption 3.2, let $\mathsf{C} := (y_\ell, \min\{\delta, \, \bar{a}(\bar{z})\})$ and let $\mathsf{D}_1 := \{y_\ell\} \times \{\bar{z}\}$ if (Y1) holds and $\mathsf{D}_1 := \mathsf{C} \times \{\bar{z}\}$ if (Y2) holds. We now show that D_1 satisfies the above conditions. Define $r(a', z') := f(a' \,|\, z') P(\bar{z}, z')$ and note that $r(a', z') > 0$ on D_1. Define the measure v_1 on $\mathscr{B}(\mathsf{S})$ by $v_1(A) := \int_A r(a', z')(\nu \times \vartheta)[\mathrm{d}(a', z')]$. If (Y1) holds, then $\nu(\{y_\ell\}) > 0$ as shown above, and, if (Y2) holds, Lemma C.6 implies that $\nu(\mathsf{C}) > 0$. Since in addition $\vartheta(\{\bar{z}\}) > 0$, it always holds that $(\nu \times \vartheta)(\mathsf{D}_1) > 0$. Moreover, since $r(a', z') > 0$ on D_1, we have $v_1(\mathsf{D}_1) > 0$ and v_1 is a nontrivial measure.

For all $(a, z) \in \mathsf{D}_1$ and $A \in \mathscr{B}(S)$, Lemma B.7 implies that

$$Q\left((a, z), A\right) = \int_A r(a', z')(\nu \times \vartheta)[\mathrm{d}(a', z')] = v_1(A).$$

Hence, D_1 satisfies (53) and $\{(a_t, Z_t)\}_{t \geq 0}$ is strongly aperiodic. □

Lemma C.8. *The set $[0, d] \times \mathsf{Z}$ is a petite set for all $d \in \mathbb{R}_+$.*

Proof. Fix $d \in (0, \infty)$ and $z \in \mathsf{Z}$. Let $B := [0, d] \times \{z\}$. By Lemma C.3,

$$\mathbb{P}_{(d,z)}\{c_{N-1} = a_{N-1}, Z_{N-1} = \bar{z}\} > 0 \quad \text{for some } N \in \mathbb{N}. \tag{54}$$

We start by showing that there exists a nontrivial measure v_N on $\mathscr{B}(\mathsf{S})$ such that

$$\inf_{(a,z)\in B} Q^N((a, z), A) \geq v_N(A) \quad \text{for all } A \in \mathscr{B}(\mathsf{S}). \tag{55}$$

In other words, B is a v_N-small set. Fix $A \in \mathscr{B}(\mathsf{S})$. For all $z' \in \mathsf{Z}$, define

$$m(z') := \int \left[\int \mathbb{1}\{(y'', z'') \in A\} f(y'' \mid z'') \, \mathrm{d}y'' \right] P(z', z'') \vartheta(\mathrm{d}z'').$$

Note that for all $(a, z) \in B$, Lemma B.7 implies that

$$\begin{aligned} Q^N((a, z), A) &\geq \mathbb{P}_{a,z} \{(Y_N, Z_N) \in A, \, a_{N-1} \leq \bar{a}(Z_{N-1}), \, Z_{N-1} = \bar{z}\} \\ &= \int m(z') \mathbb{1}\{a' \leq \bar{a}(z'), z' = \bar{z}\} Q^{N-1}((a, z), \mathrm{d}(a', z')). \end{aligned}$$

Since $a' \mapsto m(z')\mathbb{1}\{a' \leq \bar{a}(z'), z' = \bar{z}\}$ is decreasing for all $z' \in \mathsf{Z}$, by Lemma C.4,

$$\begin{aligned} Q^N((a, z), A) &\geq \int m(z') \mathbb{1}\{a' \leq \bar{a}(z'), z' = \bar{z}\} Q^{N-1}((d, z), \mathrm{d}(a', z')) \\ &= \mathbb{P}_{d,z} \{(Y_N, Z_N) \in A, \, c_{N-1} = a_{N-1}, \, Z_{N-1} = \bar{z}\} =: v_N(A). \end{aligned}$$

Note that v_N is a nontrivial measure on $\mathscr{B}(\mathsf{S})$ since (54) implies that $v_N(\mathsf{S}) > 0$. Furthermore, since (a, z) is chosen arbitrarily, the above inequality implies that (55) holds. We have shown that B is a v_N-small set, and hence a petite set. Since finite union of petite sets is petite for ψ-irreducible chains (see, e.g., Proposition 5.5.5 of Meyn and Tweedie (2009)), the set $[0, d] \times \mathsf{Z}$ must also be petite. $\square$

Recall $\bar{s} \in [0, 1)$ in Assumption 3.1, $n \in \mathbb{N}$ and $\gamma \in (0, 1)$ in (48). Let $B := [0, d] \times \mathsf{Z}$.

Lemma C.9. *There exist constants $b \in \mathbb{R}_+$, $\rho \in (0, 1)$ and a measurable map $V : \mathsf{S} \to [n/\rho, \infty)$ that is bounded on B, such that, for sufficiently large $d \in \mathbb{R}_+$ and all $(a, z) \in \mathsf{S}$, we have $\mathbb{E}_{a,z} V(a_n, Z_n) - V(a, z) \leq -\rho V(a, z) + b\mathbb{1}\{(a, z) \in B\}$.*

Proof. Since $c^*(a,z) \geq (1-\bar{s})a$ by Proposition 2.6 and $M_0 := \max_{z\in \mathsf{Z}} \mathbb{E}_z \hat{R} < \infty$ by Assumption 3.1 and Lemma A.1, by Lemma B.1 and the Markov property,

$$\begin{aligned}\mathbb{E}_{a,z} a_n &\leq \bar{s}^n \mathbb{E}_z R_n \cdots R_1 a + \sum_{t=1}^n \bar{s}^{n-t} \mathbb{E}_z R_n \cdots R_{t+1} Y_t \\ &\leq \gamma a + \sum_{t=1}^n \bar{s}^{n-t} \mathbb{E}_z Y_t \, \mathbb{E}_{Z_t} R_{t+1} \cdots R_n \leq \gamma a + \sum_{t=1}^n \bar{s}^{n-t} M_0^{n-t} M_3.\end{aligned}$$

Define $b_0 := \sum_{t=1}^n \bar{s}^{n-t} M_0^{n-t} M_3$. Note that $b_0 < \infty$. Choose $\rho \in (0, 1-\gamma)$, $m_V \geq n/\rho$ and $d \in \mathbb{R}_+$ such that $(1-\gamma-\rho)d \geq b_0 + \rho m_V$. Then, for $V(a,z) := a + m_V$,

$$\begin{aligned}\mathbb{E}_{a,z} V(a_n, Z_n) - V(a,z) &\leq -(1-\gamma)a + b_0 = -\rho a - (1-\gamma-\rho)a + b_0 \\ &= -\rho V(a,z) - (1-\gamma-\rho)a + b_0 + \rho m_V. \end{aligned} \tag{56}$$

In particular, if $(a,z) \notin B$, then $a > d$ and (56) implies that

$$\mathbb{E}_{a,z} V(a_n, Z_n) - V(a,z) \leq -\rho V(a,z) - (1-\gamma-\rho)d + b_0 + \rho m_V \leq -\rho V(a,z). \tag{57}$$

Let $b := b_0 + \rho m_V$. Then the stated claim follows from (56)–(57) and the fact that V is bounded on B. $\square$

Proof of Theorem 3.2. Claim (1) can be proved by applying Theorem 19.1.3 (or a combination of Proposition 5.4.5 and Theorem 15.0.1) of Meyn and Tweedie (2009). The required conditions in those theorems have been established by Lemmas C.5, C.7, C.8 and C.9 above. Regarding claim (2), Lemmas C.8 and C.9 imply that $\mathbb{E}_{a,z} V(a_n, Z_n) - V(a,z) \leq -n + b \mathbb{1}\{(a,z) \in B\}$ for all $(a,z) \in \mathsf{S}$, where $B := [0,d] \times \mathsf{Z}$ is petite. Since in addition $\{(a_t, Z_t)\}$ is ψ-irreducible by Lemma C.5, Theorem 19.1.2 of Meyn and Tweedie (2009) implies that $\{(a_t, Z_t)\}$ is a positive Harris chain. Claim (2) then follows from Theorem 17.1.7 of Meyn and Tweedie (2009).

To verify claim (3), since we have shown that $\Phi := \{(a_t, Z_t)\}$ is positive Harris with stationary distribution ψ_∞, based on Theorem 16.1.5 and Theorem 17.5.4 of Meyn and Tweedie (2009), it suffices to show that Q is V-uniformly ergodic. Let Φ^n be the n-skeleton of Φ (see page 62 of Meyn and Tweedie (2009)). Then Φ^n is ψ-irreducible and aperiodic by Proposition 5.4.5 of Meyn and Tweedie (2009). Theorem 16.0.1 of Meyn and Tweedie (2009) and Lemmas C.8 and C.9 then imply that Φ^n is V-uniformly ergodic, and, there exists $N \in \mathbb{N}$ such that $|||Q^{nN} - 1 \otimes \psi_\infty|||_V < 1$, where $\|\mu\|_V := \sup_{g : |g| \leq V} |\int g \, \mathrm{d}\mu|$ for $\mu \in \mathscr{P}(\mathsf{S})$ and, for all $t \in \mathbb{N}$,

$$|||Q^t - 1 \otimes \psi_\infty|||_V := \sup_{(a,z) \in \mathsf{S}} \frac{\|Q^t((a,z), \cdot) - \psi_\infty\|_V}{V(a,z)}.$$

To show that Q is V-uniformly ergodic, by Theorem 16.0.1 of Meyn and Tweedie (2009), it remains to verify: $|||Q^t - 1 \otimes \psi_\infty|||_V < \infty$ for $t \leq nN$. This obviously holds since, by the proof of Lemma C.9, there exist $L_0, L_1 \in \mathbb{R}$ such that, for all $t \in \mathbb{N}$,

$$\begin{aligned} |||Q^t - 1 \otimes \psi_\infty|||_V &\leq \sup_{(a,z)\in \mathsf{S}} \sup_{\|f\|\leq V} \frac{\int |f(a',z')| Q^t((a,z), \mathrm{d}(a',z'))}{V(a,z)} + L_0 \\ &\leq \sup_{(a,z)\in \mathsf{S}} \frac{\int V(a',z') Q^t((a,z), \mathrm{d}(a',z'))}{V(a,z)} + L_0 \leq L_0 + L_1 < \infty. \end{aligned}$$

Hence, Q is V-uniformly ergodic and claim (3) follows. The proof is now complete. $\square$

Proof of Theorem 3.3. Take an arbitrarily large constant $k < 1$ such that

$$P(\bar{z}, \bar{z}) > 0 \quad \text{and} \quad \mathbb{P}_{\bar{z}}\{kG(\bar{z}, \bar{z}, \hat{\zeta}) > 1\} > 0,$$

which is possible by Assumption 3.3 and the definition of G in (25a). For this k, since $\lim_{a\to\infty} c^*(a,z)/a = \alpha(z)$ and Z is a finite set, we can take $\bar{a} > 0$ such that

$$1 - \frac{c^*(a,z)}{a} \geq k(1 - \alpha(z))$$

for all $z \in \mathsf{Z}$ and $a \geq \bar{a}$. Multiplying both sides by $R(\hat{z}, \hat{\zeta}) \geq 0$, it follows from the law of motion (21a), $Y(\hat{z}, \hat{\eta}) \geq 0$, and the definition of G in (25a) that for $a \geq \bar{a}$,

$$\begin{aligned} \hat{a} &= R(\hat{z}, \hat{\zeta})(a - c^*(a,z)) + Y(\hat{z}, \hat{\eta}) \\ &\geq R(\hat{z}, \hat{\zeta})(a - c^*(a,z)) = R(\hat{z}, \hat{\zeta}) \left(1 - \frac{c^*(a,z)}{a}\right) a \\ &\geq R(\hat{z}, \hat{\zeta}) k (1 - \alpha(z)) a = kG(z, \hat{z}, \hat{\zeta}) a. \end{aligned}$$

Let $\tilde{A}(z, \hat{z}, \hat{\zeta}) := kG(z, \hat{z}, \hat{\zeta}) \mathbb{1}\{kG(z, \hat{z}, \hat{\zeta}) > 1\}$. Then for all $z, \hat{z}, \hat{\zeta}, \hat{\eta}$ and all $a \geq \bar{a}$,

$$\hat{a} \geq \tilde{A}(z, \hat{z}, \hat{\zeta}) a. \tag{58}$$

Start the wealth accumulation process a_t from $a_0 \geq \bar{a}$. Consider the following process:

$$S_{t+1} = \tilde{A}(Z_t, Z_{t+1}, \zeta_{t+1}) S_t,$$

where $S_0 = a_0$. We now show that $a_t \geq S_t$ with probability one for all t by induction. Since $S_0 = a_0$, the case $t = 0$ is trivial. Suppose the claim holds up to t. Because $a_t \geq 0$ and S_t remains 0 once it becomes 0, without loss of generality we may assume $S_0, \dots, S_t$ are all positive. Hence $\tilde{A}_1, \dots, \tilde{A}_t > 0$. By the definition of $\tilde{A}$, we have $\tilde{A} > 1$ whenever $\tilde{A} > 0$. Therefore

$$S_t = \tilde{A}_t \cdots \tilde{A}_1 S_0 \geq S_0 = a_0 \geq \bar{a}.$$

Hence applying (58), we get

$$a_{t+1} \geq \tilde{A}(Z_t, Z_{t+1}, \zeta_{t+1}) a_t \geq \tilde{A}(Z_t, Z_{t+1}, \zeta_{t+1}) S_t = S_{t+1}.$$

Now take any $p \in (0,1)$ and let T be a geometric random variable with mean $1/p$ that is independent of everything. Define

$$\tilde{\lambda}(s) = (1-p) r(P \odot M_{\tilde{A}}(s)),$$

where $M_{\tilde{A}}(s)$ is as in (24). Since clearly $A \geq \tilde{A}$ and $p > 0$, we have $\lambda > \tilde{\lambda}$. By Lemma 3.1 of Beare and Toda (2017), $\lambda, \tilde{\lambda}$ are convex, and hence continuous in the interior of their domains. Therefore $\lambda(\kappa) = 1$ and $\lambda(s) > 1$ for small enough $s > \kappa$. Hence, for any $\varepsilon > 0$, we can take small enough $p \in (0,1)$ and large enough $k < 1$ such that $\tilde{\lambda}(\kappa) < 1 < \tilde{\lambda}(\kappa + \varepsilon) < \infty$. By Lemma 3.1 of Beare and Toda (2017), there exists a unique $\tilde{\kappa} \in (\kappa, \kappa + \varepsilon)$ such that $\tilde{\lambda}(\tilde{\kappa}) = 1$. Theorem 3.4 of Beare and Toda (2017) then implies that

$$\liminf_{a \to \infty} a^{\tilde{\kappa}} \mathbb{P}_{a_0, z_0}\{S_T > a\} > 0$$

for all $(a_0, z_0) \in \mathsf{S}$. In particular, for any initial $(a_0, z_0) \in \mathsf{S}$ with $a_0 \geq \bar{a}$,

$$\liminf_{a \to \infty} a^{\kappa + \varepsilon} \mathbb{P}_{a_0, z_0}\{S_T > a\} > 0. \tag{59}$$

Now suppose that we draw a_0 from the ergodic distribution. Then a_t has the same distribution as a_∞, and so does a_T. Therefore

$$\begin{aligned} \mathbb{P}\{a_\infty > a\} &= \mathbb{P}\{a_T > a\} \\ &= \mathbb{P}\{a_0 < \bar{a}\}\mathbb{P}\{a_T > a \mid a_0 < \bar{a}\} + \mathbb{P}\{a_0 \geq \bar{a}\}\mathbb{P}\{a_T > a \mid a_0 \geq \bar{a}\}. \end{aligned} \tag{60}$$

If the ergodic distribution of $\{a_t\}$ has unbounded support, then $\mathbb{P}\{a_0 \geq \bar{a}\} > 0$. As we have seen above, conditional on $a_0 \geq \bar{a}$, we have $a_t \geq S_t$ for all t. Therefore

$$\liminf_{a \to \infty} a^{\kappa + \varepsilon} \mathbb{P}\{a_T > a \mid a_0 \geq \bar{a}\} \geq \liminf_{a \to \infty} a^{\kappa + \varepsilon} \mathbb{P}\{S_T > a \mid a_0 \geq \bar{a}\} > 0 \tag{61}$$

by (59), and so (27) follows from (60) and (61). $\square$

References

Acemoglu, D. and J. A. Robinson (2002): "The Political Economy of the Kuznets curve," *Review of Development Economics*, 6, 183–203.

Açikgöz, Ö. T. (2018): "On the Existence and Uniqueness of Stationary Equilibrium in Bewley Economies with Production," *Journal of Economic Theory*, 173, 18–55.

Ahn, S., G. Kaplan, B. Moll, T. Winberry, and C. Wolf (2018): "When Inequality Matters for Macro and Macro Matters for Inequality," *NBER Macroeconomics Annual*, 32, 1–75.

Aiyagari, S. R. (1994): "Uninsured Idiosyncratic Risk and Aggregate Saving," *Quarterly Journal of Economics*, 109, 659–684.

Aliprantis, C. D. and K. C. Border (2006): *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Springer.

Beare, B. K. and A. A. Toda (2017): "Geometrically Stopped Markovian Random Growth Processes and Pareto Tails," Tech. rep., UC San Diego.

Benhabib, J. and A. Bisin (2018): "Skewed Wealth Distributions: Theory and Empirics," *Journal of Economic Literature*, 56, 1261–1291.

Benhabib, J., A. Bisin, and M. Luo (2017): "Earnings Inequality and Other Determinants of Wealth Inequality," *American Economic Review: Papers and Proceedings*, 107, 593–597.

Benhabib, J., A. Bisin, and S. Zhu (2011): "The Distribution of Wealth and Fiscal Policy in Economies with Finitely Lived Agents," *Econometrica*, 79, 123–157.

——— (2015): "The Wealth Distribution in Bewley Economies with Capital Income Risk," *Journal of Economic Theory*, 159, 489–515.

——— (2016): "The Distribution of Wealth in the Blanchard–Yaari Model," *Macroeconomic Dynamics*, 20, 466–481.

Bhandari, A., D. Evans, M. Golosov, and T. J. Sargent (2018): "Inequality, Business Cycles, and Monetary-Fiscal Policy," Tech. rep., National Bureau of Economic Research.

Blackwell, D. (1965): "Discounted Dynamic Programming," *Annals of Mathematical Statistics*, 36, 226–235.

Brinca, P., H. A. Holter, P. Krusell, and L. Malafry (2016): "Fiscal Multipliers in the 21st Century," *Journal of Monetary Economics*, 77, 53–69.

Cao, D. (2018): "Recursive Equilibrium in Krusell and Smith (1998)," Tech. rep., SSRN 2863349.

Carroll, C. (2004): "Theoretical Foundations of Buffer Stock Saving," Tech. rep., National Bureau of Economic Research.

Carroll, C. D. (1997): "Buffer-stock Saving and the Life Cycle/Permanent Income Hypothesis," *Quarterly Journal of Economics*, 112, 1–55.

Chamberlain, G. and C. A. Wilson (2000): "Optimal Intertemporal Consumption under Uncertainty," *Review of Economic Dynamics*, 3, 365–395.

Coleman, II, W. J. (1990): "Solving the Stochastic Growth Model by Policy-Function Iteration," *Journal of Business and Economic Statistics*, 8, 27–29.

Datta, M., L. J. Mirman, and K. L. Reffett (2002): "Existence and Uniqueness of Equilibrium in Distorted Dynamic Economies with Capital and Labor," *Journal of Economic Theory*, 103, 377–410.

Davies, J. B. and A. F. Shorrocks (2000): "The Distribution of Wealth," in *Handbook of Income Distribution*, Elsevier, vol. 1, 605–675.

Deaton, A. and G. Laroque (1992): "On the Behaviour of Commodity Prices," *Review of Economic Studies*, 59, 1–23.

Epper, T., E. Fehr, H. Fehr-Duda, C. Kreiner, D. Lassen, S. Leth-Petersen, and G. Rasmussen (2018): "Time Discounting and Wealth Inequality," Tech. rep., Working paper.

Fagereng, A., L. Guiso, D. Malacrino, and L. Pistaferri (2016a): "Heterogeneity and Persistence in Returns to Wealth," Tech. rep., National Bureau of Economic Research.

——— (2016b): "Heterogeneity in Returns to Wealth and the Measurement of Wealth Inequality," *American Economic Review: Papers and Proceedings*, 106, 651–655.

Feinberg, E. A., P. O. Kasyanov, and N. V. Zadoianchuk (2014): "Fatou's Lemma for Weakly Converging Probabilities," *Theory of Probability & Its Applications*, 58, 683–689.

Gabaix, X., J.-M. Lasry, P.-L. Lions, and B. Moll (2016): "The Dynamics of Inequality," *Econometrica*, 84, 2071–2111.

Glaeser, E., J. Scheinkman, and A. Shleifer (2003): "The Injustice of Inequality," *Journal of Monetary Economics*, 50, 199–222.

Gouin-Bonenfant, E. and A. A. Toda (2018): "Pareto Extrapolation: Bridging Theoretical and Quantitative Models of Wealth Inequality," Tech. rep., SSRN 3260899.

Guvenen, F. and A. A. Smith (2014): "Inferring Labor Income Risk and Partial Insurance from Economic Choices," *Econometrica*, 82, 2085–2129.

Hansen, B. E. and K. D. West (2002): "Generalized Method of Moments and Macroeconomics," *Journal of Business & Economic Statistics*, 20, 460–469.

Hardy, G. H., J. E. Littlewood, and G. Pólya (1952): *Inequalities*, Cambridge University Press.

Hills, T. S. and T. Nakata (2018): "Fiscal Multipliers at the Zero Lower Bound: The Role of Policy Inertia," *Journal of Money, Credit and Banking*, 50, 155–172.

Hubmer, J., P. Krusell, and A. A. Smith, Jr. (2018): "A Comprehensive Quantitative Theory of the US Wealth Distribution," Tech. rep., Yale.

Huggett, M. (1993): "The Risk-free Rate in Heterogeneous-agent Incomplete-insurance Economies," *Journal of Economic Dynamics and Control*, 17, 953–969.

Kaymak, B., C. S. Leung, and M. Poschke (2018): "The Determinants of Wealth Inequality and Their Implications for Economic Policy," Tech. rep., Society for Economic Dynamics.

Krasnosel'skii, M. A., G. M. Vainikko, R. Zabreyko, Y. B. Ruticki, and V. V. Stet'senko (2012): *Approximate Solution of Operator Equations*, Springer Netherlands.

Krusell, P. and A. A. Smith, Jr. (1998): "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106, 867–896.

Kuhn, M. (2013): "Recursive Equilibria in an Aiyagari-style Economy with Permanent Income Shocks," *International Economic Review*, 54, 807–835.

Lawrance, E. C. (1991): "Poverty and the Rate of Time Preference: Evidence from Panel Data," *Journal of Political Economy*, 99, 54–77.

Li, H. and J. Stachurski (2014): "Solving the Income Fluctuation Problem with Unbounded Rewards," *Journal of Economic Dynamics and Control*, 45, 353–365.

Ljungqvist, L. and T. J. Sargent (2012): *Recursive Macroeconomic Theory*, MIT Press, 4 ed.

Loewenstein, G. and D. Prelec (1991): "Negative Time Preference," *American Economic Review*, 81, 347–352.

Loewenstein, G. and N. Sicherman (1991): "Do Workers Prefer Increasing Wage Profiles?" *Journal of Labor Economics*, 9, 67–84.

Meyn, S. P. and R. L. Tweedie (2009): *Markov Chains and Stochastic Stability*, Springer Science & Business Media.

Miao, J. (2006): "Competitive Equilibria of Economies with a Continuum of Consumers and Aggregate Shocks," *Journal of Economic Theory*, 128, 274–298.

Morand, O. F. and K. L. Reffett (2003): "Existence and Uniqueness of Equilibrium in Nonoptimal Unbounded Infinite Horizon Economies," *Journal of Monetary Economics*, 50, 1351–1373.

Pareto, V. (1896): *La Courbe de la Répartition de la Richesse*, Lausanne: Imprimerie Ch. Viret-Genton.

Rabault, G. (2002): "When Do Borrowing Constraints Bind? Some New Results on the Income Fluctuation Problem," *Journal of Economic Dynamics and Control*, 26, 217–245.

Rouwenhorst, K. G. (1995): "Asset Pricing Implications of Equilibrium Business Cycle Models," in *Frontiers of Business Cycle Research*, ed. by T. F. Cooley, Princeton University Press, chap. 10, 294–330.

Saez, E. and G. Zucman (2016): "Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data," *Quarterly Journal of Economics*, 131, 519–578.

Schechtman, J. (1976): "An Income Fluctuation Problem," *Journal of Economic Theory*, 12, 218–241.

Schechtman, J. and V. L. S. Escudero (1977): "Some Results on "An Income Fluctuation Problem"," *Journal of Economic Theory*, 16, 151–166.

Schorfheide, F., D. Song, and A. Yaron (2018): "Identifying Long-Run Risks: A Bayesian Mixed-Frequency Approach," *Econometrica*, 86, 617–654.

Stachurski, J. (2009): *Economic Dynamics: Theory and Computation*, MIT Press.

Stachurski, J. and A. A. Toda (2019): "An Impossibility Theorem for Wealth in Heterogeneous-agent Models with Limited Heterogeneity," *Journal of Economic Theory*, 182, 1–24.

Toda, A. A. (2014): "Incomplete Market Dynamics and Cross-Sectional Distributions," *Journal of Economic Theory*, 154, 310–348.

——— (2019): "Wealth Distribution with Random Discount Factors," *Journal of Monetary Economics*, 104, 101–113.

Toda, A. A. and K. Walsh (2015): "The Double Power Law in Consumption and Implications for Testing Euler Equations," *Journal of Political Economy*, 123, 1177–1200.

Vermeulen, P. (2018): "How Fat Is the Top Tail of the Wealth Distribution?" *Review of Income and Wealth*, 64, 357–387.