A Theory of the Saving Rate of the Rich*

Qingyin Ma† Alexis Akira Toda‡

July 15, 2020

Abstract

Empirical evidence suggests that the rich have higher propensity to save than do the poor. While this observation may appear to contradict the homotheticity of preferences, we theoretically show that that is not the case. Specifically, we consider an income fluctuation problem with homothetic preferences and general shocks and prove that consumption functions are asymptotically linear, with an exact analytical characterization of asymptotic marginal propensities to consume (MPC). We provide necessary and sufficient conditions for the asymptotic MPCs to be zero. We solve a calibrated model with standard constant relative risk aversion utility and show that asymptotic MPCs can be zero in empirically plausible settings, implying an increasing and large saving rate of the rich.

Keywords: asymptotic linearity, income fluctuation problem, monotone convex map, saving rate.

JEL codes: C65, D15, D52, E21.

1 Introduction

Empirical evidence suggests that the rich have higher propensity to save than do the poor.[1] This fact implies that the rich have lower marginal propensity to consume (MPC), which has important economic consequences. For example, when the rich have lower MPC, the consumption tax, which is a popular tax instrument in many countries, becomes regressive and may not be desirable from equity perspectives. MPC heterogeneity also implies that the wealth distribution matters for determining aggregate demand and hence monetary and fiscal policies (Kaplan, Moll, and Violante, 2018; Mian, Straub, and Sufi, 2020).

*We thank Chris Carroll, Émilien Gouin-Bonenfant, Ben Moll, Johannes Wieland, and seminar participants at CRETA Economic Theory Conference for valuable feedback and suggestions. A previous version of this paper was circulated under the title "Asymptotic Marginal Propensity to Consume".

†International School of Economics and Management, Capital University of Economics and Business. Email: qingyin.ma@cueb.edu.cn.

‡Department of Economics, University of California San Diego. Email: atoda@ucsd.edu.

[1] Quadrini (1999) documents that entrepreneurs (who tend to be rich) have high saving rates. Dynan, Skinner, and Zeldes (2004) document that there is a positive association between saving rates and lifetime income. More recently, using Norwegian administrative data, Fagereng, Holm, Moll, and Natvik (2019) show that among households with positive net worth, saving rates are increasing in wealth.

Why do the rich save so much? Intuition suggests that canonical models of consumption and savings that feature homothetic preferences are unable to explain the high saving rate of the rich: in such models, consumption (hence saving) functions should be asymptotically linear in wealth due to homotheticity, implying an asymptotically constant saving rate. A seemingly obvious explanation for the high saving rate of the rich is that preferences are not homothetic.[2] However, non-homothetic preferences have some undesirable theoretical properties. First, they are inconsistent with balanced growth (whereas many aggregate economic variables such as real per capita GDP are near unit root processes), at least in basic models in which preference parameters are constant. Second, non-homothetic utility functions have more parameters than homothetic ones, which introduces arbitrariness in model specification and calibration.

In this paper we theoretically show that the intuition of "homotheticity implies (asymptotic) linearity" is only partially correct. We consider a standard income fluctuation problem with (homothetic) constant relative risk aversion (CRRA) preferences but with capital and labor income risk in a general Markovian setting. We prove that the consumption functions are asymptotically linear in wealth, or the asymptotic marginal propensities to consume converge to some constants. While this statement is intuitive, there is one surprise: we obtain an exact analytical characterization of the asymptotic MPCs and prove that they can be zero. The asymptotic MPCs depend only on risk aversion and the stochastic processes for the discount factor and return on wealth, and are independent of the income process. Furthermore, we derive necessary and sufficient conditions for zero asymptotic MPCs. When the asymptotic MPCs are zero, the saving rates of the rich converge to one as agents get wealthier. Thus, we provide a potential explanation for why the rich save so much, and we do so with standard homothetic preferences.

To prove that consumption functions are asymptotically linear with particular slopes, we apply policy function iteration as in Li and Stachurski (2014) and Ma, Stachurski, and Toda (2020). Since agents cannot consume more than their financial wealth in the presence of borrowing constraints, a natural upper bound on consumption is asset, which is linear with a slope of 1. Starting from this candidate consumption function, policy function iteration results in increasingly tighter upper bounds. On the other hand, we directly obtain lower bounds by restricting the space of candidate consumption functions such that they have linear lower bounds with specific slopes. We analytically derive these slopes based on the fixed point theory of monotone convex maps developed in Du (1990), which has recently been applied in economics by Toda (2019) and Borovička and Stachurski (2020). Finally, we show that the upper and lower bounds thus obtained have identical slopes, implying the asymptotic linearity of consumption functions with an exact characterization of asymptotic MPCs.

To assess the empirical plausibility of our new mechanism, we numerically solve an income fluctuation problem with CRRA utility and capital income risk

[2] For example, Carroll (2000) considers a 'capitalist spirit' model in which agents directly get utility from holding wealth, where the utility functions for consumption and wealth have different curvatures. De Nardi (2004) considers a model with bequest, which is mathematically similar. Straub (2019) estimates that the elasticity of consumption with respect to permanent income is below 1 (which implies concavity of consumption functions) and uses non-homothetic preferences to explain it. Another possibility is to introduce frictions such as portfolio adjustment costs (Fagereng, Holm, Moll, and Natvik, 2019).

calibrated to the U.S. economy. We find that with moderate risk aversion (above 4–5), the asymptotic MPCs become zero and the saving rates of the rich are increasing and approach 1.

The rest of the paper is organized as follows. After a brief discussion of the related literature, Section 2 introduces a general income fluctuation problem, proves the asymptotic linearity of consumption functions with homothetic preferences, and discusses some examples. Section 3 applies the theory to a calibrated model. Section 4 contains the proofs.

1.1 Related literature

Our paper is related to the theoretical studies of the income fluctuation problem, which is a key building block of heterogeneous-agent models in modern macroeconomics.[3] Chamberlain and Wilson (2000) study the existence of a solution assuming bounded utility and applying the contraction mapping theorem. Li and Stachurski (2014) relax the boundedness assumption and apply policy function iteration. Benhabib, Bisin, and Zhu (2015) consider a special model with CRRA utility, constant discounting, and IID and mutually independent returns and income shocks to study the tail behavior of wealth. Ma, Stachurski, and Toda (2020) allow for stochastic discounting and returns on wealth in a general Markovian setting and discuss the ergodicity, stochastic stability, and tail behavior of wealth. Carroll (2020) examines detailed properties of a special model with CRRA utility, constant discounting and risk-free rate, and IID permanent and transitory income shocks. While the main focus of these papers is the existence, uniqueness, and computation of a solution, we focus on the asymptotic behavior of consumption with general shocks. Carroll and Kimball (1996) show the concavity of consumption functions in a class of income fluctuation problems, which implies asymptotic linearity. However, they do not characterize the asymptotic MPCs as we do.

2 Asymptotic linearity of consumption functions

In this section we introduce a general income fluctuation problem following the setting in Ma, Stachurski, and Toda (2020) and study the asymptotic property of the consumption functions when preferences are homothetic.

2.1 Income fluctuation problem

Time is discrete and denoted by $t = 0, 1, 2, \ldots$. Let a_t be the financial wealth of the agent at the beginning of period t. The agent chooses consumption $c_t \geq 0$ and saves the remaining wealth $a_t - c_t$. The period utility function is u and the discount factor, gross return on wealth, and non-financial income in period

[3] See, for example, Cao (2020) and Açıkgöz (2018) for the existence of equilibrium with and without aggregate shocks, where the theoretical properties of the income fluctuation problem play an important role. Lehrer and Light (2018) and Light (2018) prove comparative statics results regarding savings. Light (2020) proves the uniqueness of stationary equilibrium in an Aiyagari model that exhibits a certain gross substitute property.

t are denoted by β_t, R_t, Y_t, where we normalize $\beta_0 = 1$. Thus the agent solves

$$\begin{aligned} &\text{maximize} && \mathrm{E}_0 \sum_{t=0}^{\infty} \left(\prod_{i=0}^{t} \beta_i \right) u(c_t) \\ &\text{subject to} && a_{t+1} = R_{t+1}(a_t - c_t) + Y_{t+1}, && \text{(2.1a)} \\ & && 0 \le c_t \le a_t, && \text{(2.1b)} \end{aligned}$$

where the initial wealth $a_0 = a > 0$ is given, (2.1a) is the budget constraint, and (2.1b) implies that the agent cannot borrow.[4] The stochastic processes $\{\beta_t, R_t, Y_t\}_{t\ge 1}$ obey

$$\beta_t = \beta(Z_t, \varepsilon_t), \quad R_t = R(Z_t, \zeta_t), \quad Y_t = Y(Z_t, \eta_t), \tag{2.2}$$

where β, R, Y are nonnegative measurable functions, $\{Z_t\}_{t\ge 0}$ is a time-homogeneous finite state Markov chain taking values in $\mathsf{Z} = \{1, \dots, Z\}$ with a transition probability matrix P, and the innovation processes $\{\varepsilon_t\}, \{\zeta_t\}, \{\eta_t\}$ are independent and identically distributed (IID) over time and mutually independent.

We introduce the following notation. For a square matrix A, the scalar $r(A)$ denotes its spectral radius (largest absolute value of all eigenvalues), i.e.,

$$r(A) := \max \{ |\alpha| \mid \alpha \text{ is an eigenvalue of } A \}. \tag{2.3}$$

The spectral radius (2.3) plays an important role in the subsequent discussion. The symbols β, R, Y are shorthand of $\beta(Z, \varepsilon)$, $R(Z, \zeta)$, $Y(Z, \eta)$ and $\hat{\beta}, \hat{R}, \hat{Y}$ are shorthand of $\beta(\hat{Z}, \hat{\varepsilon})$, $R(\hat{Z}, \hat{\zeta})$, $Y(\hat{Z}, \hat{\eta})$. Define the diagonal matrix D_β by

$$D_\beta(z, z) = \mathrm{E}_z \beta = \mathrm{E}[\beta(Z, \varepsilon) \mid Z = z] = \mathrm{E}\, \beta(z, \varepsilon).$$

More generally, for any stochastic process $\{X_t\}$ such that the distribution of X_t conditional on all past information and $Z_t = z$ depends only on z, let D_X be the diagonal matrix such that $D_X(z, z) = \mathrm{E}_z X = \mathrm{E}[X \mid Z = z]$. Consider the following assumptions.

Assumption 1. *The utility function $u : [0, \infty) \to \mathbb{R} \cup \{-\infty\}$ is twice continuously differentiable on $(0, \infty)$ and satisfies $u' > 0$, $u'' < 0$, $u'(0) = \infty$, and $u'(\infty) < 1$.*

Assumption 1 is essentially the usual Inada condition together with monotonicity and concavity.

Assumption 2. *The following conditions hold:*

1. $\mathrm{E}_z \beta < \infty$ *and* $\mathrm{E}_z \beta R < \infty$ *for all* $z \in \mathsf{Z}$,
2. $r(PD_\beta) < 1$ *and* $r(PD_{\beta R}) < 1$,
3. $\mathrm{E}_z Y < \infty$ *and* $\mathrm{E}_z u'(Y) < \infty$ *for all* $z \in \mathsf{Z}$.

The condition $r(PD_\beta) < 1$ generalizes $\beta < 1$ to the case with random discount factors. The condition $r(PD_{\beta R}) < 1$ generalizes the 'impatience' condition $\beta R < 1$ to the stochastic case. Under these two assumptions, the income fluctuation problem (2.1) admits a unique solution.

[4] The no-borrowing condition $a_t - c_t \ge 0$ is without loss of generality as discussed in Chamberlain and Wilson (2000) and Li and Stachurski (2014).

Theorem 2.1. *Suppose Assumptions 1 and 2 hold. Then the income fluctuation problem (2.1) has a unique solution. Furthermore, the consumption function $c(a, z)$ can be computed by policy function iteration.*

Proof. See Ma, Stachurski, and Toda (2020, Theorem 2.2).[5] $\square$

'Policy function iteration' means the following. When the borrowing constraint $c_t \leq a_t$ does not bind, the Euler equation implies

$$u'(c_t) = \mathrm{E}_t\, \beta_{t+1} R_{t+1} u'(c_{t+1}).$$

If $c_t = a_t$, then clearly $u'(c_t) = u'(a_t)$. Therefore combining these two cases, we can compactly express the Euler equation as

$$u'(c_t) = \max\left\{\mathrm{E}_t\, \beta_{t+1} R_{t+1} u'(c_{t+1}), u'(a_t)\right\}.$$

Based on this observation, given a candidate consumption function $c(a, z)$, the policy function iteration updates the consumption function by the value $\xi = Tc(a, z)$ that solves

$$u'(\xi) = \max\left\{\mathrm{E}_z\, \hat{\beta}\hat{R}u'(c(\hat{R}(a-\xi)+\hat{Y}, \hat{Z})), u'(a)\right\}. \tag{2.4}$$

Let $\mathcal{C}$ be the space of candidate consumption functions such that $c : (0, \infty) \times \mathsf{Z} \to \mathbb{R}$ is continuous, is increasing in the first element, $0 < c(a, z) \leq a$ for all $a > 0$ and $z \in \mathsf{Z}$, and

$$\sup_{(a,z)\in(0,\infty)\times\mathsf{Z}} |u'(c(a, z)) - u'(a)| < \infty.$$

For $c, d \in \mathcal{C}$, define

$$\rho(c, d) = \sup_{(a,z)\in(0,\infty)\times\mathsf{Z}} |u'(c(a, z)) - u'(d(a, z))|. \tag{2.5}$$

When Assumptions 1 and 2 hold, Theorem 2.2 of Ma, Stachurski, and Toda (2020) shows that $\mathcal{C}$ is a complete metric space with metric ρ and $T : \mathcal{C} \to \mathcal{C}$ defined as $Tc(a, z) = \xi$ that solves (2.4) is a contraction mapping. We call the operator T the *time iteration operator*.[6]

Exploiting policy function iteration, Ma, Stachurski, and Toda (2020) show several properties such as (i) consumption and savings are increasing in wealth and (ii) consumption is increasing in income.

2.2 Asymptotic linearity of consumption functions

To study the asymptotic behavior of consumption, we strengthen Assumption 1 as follows.

[5] In addition to Assumptions 1 and 2, Ma, Stachurski, and Toda (2020) assume that the transition probability matrix P is irreducible. However, irreducibility is required only for their ergodicity result, not for existence and uniqueness of a solution.

[6] The time iteration operator was introduced by Coleman (1990). Several papers such as Datta, Mirman, and Reffett (2002), Rabault (2002), Morand and Reffett (2003), Kuhn (2013), and Li and Stachurski (2014) use this approach to establish existence of solutions and study theoretical properties.

Assumption 1'. The utility function exhibits constant relative risk aversion $\gamma > 0$: we have

$$u(c) = \begin{cases} \frac{c^{1-\gamma}}{1-\gamma}, & (\gamma \neq 1) \\ \log c. & (\gamma = 1) \end{cases} \tag{2.6}$$

Furthermore, $\mathrm{E}_z \, \beta R^{1-\gamma} < \infty$ for all z.[7]

Theorem 2.2 below, which is our main theoretical result, shows that when the utility function exhibits constant relative risk aversion, the consumption functions are asymptotically linear and characterizes the asymptotic MPCs. To avoid overwhelming the reader with notation and technicalities, we maintain the additional condition $\mathrm{E}_z \, \beta R^{1-\gamma} < \infty$ as in Assumption 1'. Furthermore, Theorem 2.2 only provides a necessary and almost sufficient condition for the asymptotic MPCs to be zero. We provide a complete characterization in Theorem 2.5 below.

Theorem 2.2 (Asymptotic linearity)**.** *Suppose Assumptions 1' and 2 hold. Let $D = D_{\beta R^{1-\gamma}}$ be the diagonal matrix whose (z,z)-th element is $\mathrm{E}_z \, \beta R^{1-\gamma} < \infty$. Then the followings are true:*

1. *If $r(PD) < 1$, then for all $z \in \mathsf{Z}$ we have*

$$\lim_{a \to \infty} \frac{c(a,z)}{a} =: \bar{c}(z) > 0, \tag{2.7}$$

where $\bar{c}(z) = x^*(z)^{-1/\gamma}$ and $x^* = (x^*(z))_{z=1}^Z \in \mathbb{R}_+^Z$ is the unique finite solution to the system of equations

$$x(z) = (Fx)(z) := \left(1 + (PDx)(z)^{1/\gamma}\right)^\gamma, \quad z = 1, \dots, Z. \tag{2.8}$$

2. *If $r(PD) \geq 1$ and PD is irreducible, then for all $z \in \mathsf{Z}$ we have*

$$\lim_{a \to \infty} \frac{c(a,z)}{a} = 0.$$

The proof of Theorem 2.2 is relegated to Section 4. Here we heuristically discuss the intuition for why we would expect the conclusion of Theorem 2.2 to hold. Suppose the limit (2.7) exists. Assuming that the borrowing constraint does not bind, the Euler equation (2.4) implies

$$u'(\xi) = \mathrm{E}_z \, \hat{\beta} \hat{R} u'(c(\hat{R}(a - \xi) + \hat{Y}, \hat{Z})),$$

where $\xi = c(a,z)$. Setting $u'(c) = c^{-\gamma}$ as in Assumption 1', setting $c(a,z) = \bar{c}(z)a$ motivated by (2.7), multiplying both sides by a^γ, letting $a \to \infty$, and interchanging expectations and limits, it must be

$$\bar{c}(z)^{-\gamma} = \mathrm{E}_z \, \hat{\beta} \hat{R}^{1-\gamma} \bar{c}(\hat{Z})^{-\gamma} (1 - \bar{c}(z))^{-\gamma}. \tag{2.9}$$

Dividing both sides of (2.9) by $(1 - \bar{c}(z))^{-\gamma}$ and setting $x(z) = \bar{c}(z)^{-\gamma}$, we obtain

$$x(z) = \left(1 + \left(\mathrm{E}_z \, \hat{\beta} \hat{R}^{1-\gamma} x(\hat{Z})\right)^{1/\gamma}\right)^\gamma, \quad z = 1, \dots, Z. \tag{2.10}$$

[7] We use the convention $\beta R^{1-\gamma} = (\beta R) R^{-\gamma}$ and $0 \cdot \infty = 0$. Then $\mathrm{E}_z \, \beta R^{1-\gamma} \in [0, \infty]$ is well-defined even if $\gamma > 1$ and $(\beta, R) = (0,0)$ with positive probability.

Noting that $\hat{\beta}, \hat{R}$ depend only on $\hat{Z}$ and IID innovations, we have

$$\mathrm{E}_z\, \hat{\beta}\hat{R}^{1-\gamma}x(\hat{Z}) = \sum_{\hat{z}=1}^{Z} P(z,\hat{z})\, \mathrm{E}_{\hat{z}}\, \hat{\beta}\hat{R}^{1-\gamma}x(\hat{z}).$$

Therefore letting P be the transition probability matrix and $D = D_{\beta R^{1-\gamma}}$ be the diagonal matrix whose (z,z)-th element is $\mathrm{E}_z\, \beta R^{1-\gamma} < \infty$, we can rewrite (2.10) as (2.8). This discussion motivates the fixed point equation (2.8).

Next, we discuss the intuition for the spectral condition $r(PD) \gtrless 1$. When the elements of the vector $x \in \mathbb{R}_+^Z$ are large, since PD is a nonnegative matrix, it follows from the definition of F in (2.8) that

$$Fx \approx PDx.$$

Since for large x the function $x \mapsto Fx$ is almost linear, whether iterating $x \mapsto Fx$ converges or not depends on whether the largest eigenvalue of the coefficient matrix PD is less or greater than 1. When $r(PD) < 1$, F in (2.8) behaves like a contraction and we would expect it to have a unique fixed point. When $r(PD) \geq 1$, because F is monotonic, we would expect the iteration of $x \mapsto Fx$ to diverge to infinity, and hence $\bar{c}(z) = x(z)^{-1/\gamma}$ to converge to 0.

Theorem 2.2 roughly says two things: with homothetic preferences, (i) consumption functions are asymptotically linear, and (ii) the asymptotic MPCs can be zero. The first point is not surprising based on the intuition of scale invariance with homothetic preferences. The second point is nontrivial and surprising, and it depends on whether the condition

$$r(PD_{\beta R^{1-\gamma}}) < 1 \tag{2.11}$$

holds or not. A condition of the form $\mathrm{E}_z\, \hat{\beta}\hat{R}^{1-\gamma} < 1$, which Carroll (2009) calls the "finite value condition" and implies (2.11), is often required for the existence of a solution in dynamic programming problems with homothetic preferences.[8] The following proposition explains why this condition has often been assumed in the literature.

Proposition 2.3. *Suppose Assumption 1' holds. Then the optimal consumption-saving problem (2.1) with zero income ($Y \equiv 0$) has a solution (with finite lifetime utility) if and only if the finite value condition (2.11) holds. Under this condition, the optimal consumption function is*

$$c(a,z) = x^*(z)^{-1/\gamma}a, \tag{2.12}$$

where $x^* \in \mathbb{R}_+^Z$ is the unique finite solution to (2.8).

Proof. The case $\gamma \neq 1$ follows from Proposition 1 of Toda (2019). The case $\gamma = 1$ follows from Proposition 10 of Online Appendix C of Toda (2019).[9] $\square$

[8] See, for example, the discussion on p. 244 of Samuelson (1969), Equation (9) of Krebs (2006), Equation (3) of Carroll (2009), Equation (18) of Toda (2014), or Equation (3) of Toda (2019).

[9] In Toda (2019), the discount factor β_t and the return R_t are deterministic functions of the *previous* state Z_{t-1}. In our setting, they are functions of the *current* state Z_t as well as IID shocks as in (2.2). This difference in the timing convention explains the difference in the statements, but the proof is essentially identical and therefore omitted. In fact, we can subsume both settings as follows. Instead of (2.2), assume $\beta_t = \beta(Z_{t-1}, Z_t, \varepsilon_t)$ and similarly for R_t, Y_t. For a random variable X_t, define the matrix M_X by $M_X(z,\hat{z}) = \mathrm{E}[X_t \mid Z_{t-1} = z, Z_t = \hat{z}]$. Then using $P \odot M_X$ instead of PD_X for $X = \beta, \beta R, \beta R^{1-\gamma}$ (where $\odot$ denotes the Hadamard (element-wise) product), we can analyze the problem in a unified way.

Proposition 2.3 implies that for a solution to the income fluctuation problem (2.1) to exist, the finite value condition (2.11) may be violated only if income Y can be positive. In fact, the Inada condition $u'(0) = \infty$ in Assumption 1 and the condition $\mathrm{E}_z\, u'(Y) < \infty$ in Assumption 2 imply that $Y > 0$ almost surely. Contrary to the intuition from the zero-income model, Ma, Stachurski, and Toda (2020, Theorem 2.2) show that Assumptions 1 and 2 are sufficient for the existence of a solution to general income fluctuation problems, and no conditions on risk aversion (including the finite value condition) are necessary.

A natural question is whether the case $r(PD) \geq 1$ (and hence zero asymptotic MPCs) is empirically plausible, or even theoretically possible. We argue in Section 3 that $r(PD) \geq 1$ is empirically plausible. The following proposition shows that $\gamma > 1$ is necessary for zero asymptotic MPCs.

Proposition 2.4. *If Assumption 2 holds and $\gamma \leq 1$, then $r(PD_{\beta R^{1-\gamma}}) < 1$.*

Example 2.3 below (with IID lognormal returns) shows that zero asymptotic MPCs are possible for any $\gamma > 1$.

As discussed above, Theorem 2.2 does not cover all possible cases, as $\mathrm{E}_z\, \beta R^{1-\gamma}$ could be infinite or the matrix PD need not be irreducible in particular applications. We can generalize Theorem 2.2 to cover all possible cases at the cost of making the notation slightly more complicated. To this end, let $K = PD$ be as in Theorem 2.2, where the diagonal element $D(z,z) = \mathrm{E}_z\, \beta R^{1-\gamma}$ could be infinite. By relabeling the states $z = 1, \dots, Z$ if necessary, without loss of generality we may assume that K is block upper triangular,

$$K = \begin{bmatrix} K_1 & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & K_J \end{bmatrix}, \tag{2.13}$$

where each diagonal block K_j is irreducible.[10] Partition Z as $\mathsf{Z} = \mathsf{Z}_1 \cup \dots \cup \mathsf{Z}_J$ accordingly. Then we have the following complete characterization.

Theorem 2.5 (Complete characterization)**.** *Suppose Assumption 2 holds and the utility function exhibits constant relative risk aversion $\gamma > 0$. Express $K = PD$ as in (2.13). Define the sequence $\{x_n\}_{n=0}^{\infty} \in [0,\infty]^Z$ by $x_0 = 1$ and $x_n = Fx_{n-1}$, where F is as in (2.8) and we apply the convention $0 \cdot \infty = 0$. Then $\{x_n\}$ monotonically converges to $x^* \in [1,\infty]^Z$, and the limit (2.7) holds with $\bar{c}(z) = x^*(z)^{-1/\gamma} \in [0,1]$.*

Furthermore, $\bar{c}(z) = 0$ if and only if there exist j, $\hat{z} \in \mathsf{Z}_j$, and $m \in \mathbb{N}$ such that $K^m(z,\hat{z}) > 0$ and $r(K_j) \geq 1$, where $r(K_j) = \infty$ if some element of K_j is infinite.

An interesting implication of Theorems 2.2 and 2.5 is that the asymptotic MPCs $\bar{c}(z)$ depend only on the matrix PD, which in turn depends only on relative risk aversion γ as well as "multiplicative shocks" β and R, and not on "additive shocks" Y. The following corollary verifies the intuition in Gouin-Bonenfant and Toda (2018) that only multiplicative shocks matter for characterizing the behavior of wealthy agents.

[10] Recall that a square matrix A is reducible if there exists a permutation matrix P such that $P^\top AP$ is block upper triangular with at least two diagonal blocks. Matrices that are not reducible are called irreducible. Hence by induction a decomposition of the form (2.13) is always possible. By definition scalars (1×1 matrices, including zero) are irreducible, so some K_j in (2.13) can be zero if it is 1×1.

Corollary 2.6 (Irrelevance of additive shocks)**.** *Let everything be as in Theorem 2.5. The asymptotic MPCs $\bar{c}(z)$ depend only on the relative risk aversion γ, transition probability matrix P, the discount factor β, and the return on wealth R, and not on income Y.*

2.3 Examples

The system of fixed point equations (2.8) is in general nonlinear and does not admit a closed-form solution. Below, we discuss several examples with explicit solutions.

Example 2.1. If $\gamma = 1$, then (2.8) becomes

$$x^* = 1 + PDx^* \iff x^* = (I - PD)^{-1}1,$$

where $D = D_\beta = \operatorname{diag}(\dots, \mathrm{E}_z\,\beta, \dots)$.[11] A corollary is that with log utility, we always have $\bar{c}(z) > 0$.

Example 2.2. If $b = b(z) = \mathrm{E}_z\,\beta R^{1-\gamma}$ does not depend on z, then $D = bI$. If $x = k1$ is a multiple of the vector 1, then $PDx = bPk1 = bk1$ because P is a transition probability matrix. Thus if $b < 1$, (2.8) reduces to

$$x^*(z) = (1 + (bx^*(z))^{1/\gamma})^\gamma \iff x^*(z) = (1 - b^{1/\gamma})^{-\gamma} \iff \bar{c}(z) = 1 - b^{1/\gamma}.$$

This example shows that with constant discounting ($\beta(z, \varepsilon) \equiv \beta$) and risk-free saving ($R(z, \zeta) \equiv R$), the asymptotic MPC is constant regardless of the income shocks:

$$\bar{c}(z) = \begin{cases} 1 - (\beta R^{1-\gamma})^{1/\gamma} & \text{if } \beta R^{1-\gamma} < 1, \\ 0 & \text{otherwise.} \end{cases}$$

Example 2.3. Suppose the return on wealth $R_t = R(Z_t, \zeta_t)$ does not depend on Z_t, so $R_t = R(\zeta_t)$. Assume further that $\log R_t$ is normally distributed with standard deviation σ and mean $\mu - \sigma^2/2$, so $\mathrm{E}\,R = \mathrm{e}^\mu$. Let the discount factor $\beta = \mathrm{e}^{-\delta}$ be constant, where $\delta > 0$ is the discount rate. Then using the property of the normal distribution, we obtain

$$\begin{aligned} &1 > \mathrm{E}\,\beta R = \mathrm{e}^{-\delta+\mu} \iff \delta > \mu, \\ &1 > \mathrm{E}\,\beta R^{1-\gamma} = \mathrm{e}^{-\delta+(1-\gamma)(\mu-\gamma\sigma^2/2)} \iff \delta > (1-\gamma)\left(\mu - \frac{1}{2}\gamma\sigma^2\right). \end{aligned}$$

Therefore assuming $\delta > \mu$ for Assumption 2 to hold, it follows from Example 2.2 that

$$\bar{c}(z) = \begin{cases} 1 - \mathrm{e}^{-\psi\delta-(1-\psi)(\mu-\gamma\sigma^2/2)} > 0 & \text{if } \delta > (1-\gamma)\left(\mu - \frac{1}{2}\gamma\sigma^2\right), \\ 0 & \text{otherwise,} \end{cases}$$

where $\psi = 1/\gamma$ is the elasticity of intertemporal substitution. If $\gamma > 1$, then $(1-\gamma)(\mu - \gamma\sigma^2/2) \to \infty$ as $\gamma, \sigma \to \infty$, so the asymptotic MPC is 0 if risk aversion or volatility is sufficiently high.

[11] Note that since $r(PD) = r(PD_\beta) < 1$ by Assumption 2, $(I - PD)^{-1} = \sum_{k=0}^\infty (PD)^k$ exists and is nonnegative.

3 Asymptotic MPCs and saving rates

In this section we apply our theory of asymptotic MPCs to shed light on the saving rate of the rich.

3.1 General theory

We define an agent's saving rate by the change in net worth divided by total income excluding capital loss (to prevent the denominator from becoming negative):

$$s_{t+1} = \frac{a_{t+1} - a_t}{\max\{(R_{t+1} - 1)(a_t - c_t), 0\} + Y_{t+1}}. \tag{3.1}$$

For $x \in \mathbb{R}$, define its positive and negative parts by $x^+ = \max\{x, 0\}$ and $x^- = -\min\{x, 0\}$. Then $x = x^+ - x^-$. Using the budget constraint (2.1a), the saving rate (3.1) can be rewritten as

$$\begin{aligned} s_{t+1} &= \frac{[(R_{t+1} - 1)^+ - (R_{t+1} - 1)^-](a_t - c_t) + Y_{t+1} - c_t}{(R_{t+1} - 1)^+(a_t - c_t) + Y_{t+1}} \\ &= 1 - \frac{(\hat{R} - 1)^-(1 - c/a) + c/a}{(\hat{R} - 1)^+(1 - c/a) + \hat{Y}/a} \in (-\infty, 1). \end{aligned} \tag{3.2}$$

Letting $a \to \infty$, the saving rate of an infinitely wealthy agent becomes

$$\bar{s} := 1 - \frac{(\hat{R} - 1)^-(1 - \bar{c}) + \bar{c}}{(\hat{R} - 1)^+(1 - \bar{c})} \in [-\infty, 1], \tag{3.3}$$

where $\bar{c}$ is the asymptotic MPC. Under what conditions can the saving rate (3.2) be increasing in wealth, and in particular, can the asymptotic saving rate (3.3) become positive? The following proposition provides a negative answer within a class of models.

Proposition 3.1. *Consider a canonical Bewley (1977) model in which agents are infinitely-lived and relative risk aversion γ, discount factor β, and return on wealth $R > 1$ are constant. Then in the stationary equilibrium the asymptotic saving rate (3.3) is negative.*

Proof. Stachurski and Toda (2019) show that it must be $\beta R < 1$ in the stationary equilibrium. Since $R > 1$ by assumption, we obtain $\beta R^{1-\gamma} = (\beta R)R^{-\gamma} < 1$. By Example 2.2, the asymptotic MPC is $\bar{c} = 1 - (\beta R^{1-\gamma})^{1/\gamma} \in (0, 1)$. Therefore using (3.3), we obtain

$$\begin{aligned} \bar{s} = 1 - \frac{\bar{c}}{(R - 1)(1 - \bar{c})} < 0 &\iff (R - 1)(1 - \bar{c}) < \bar{c} \\ &\iff (R - 1)(\beta R^{1-\gamma})^{1/\gamma} < 1 - (\beta R^{1-\gamma})^{1/\gamma} \\ &\iff (\beta R)^{1/\gamma} < 1, \end{aligned}$$

which holds because $\beta R < 1$. $\square$

Proposition 3.1 proves that the negativity of the asymptotic saving rate is inevitable in any canonical Bewley model.[12] Thus, these models are unable to

[12]This result has a similar flavor to Stachurski and Toda (2019), who prove that canonical Bewley models cannot explain the tail behavior of wealth.

explain the observed positive saving rates of the rich. The following proposition shows that just by allowing β or R to be stochastic need not solve the problem when $\bar{c} > 0$.

Proposition 3.2. *Consider a Bewley (1977) model in which agents are infinitely-lived, relative risk aversion γ is constant, and $\{\beta_t, R_t\}_{t\geq 1}$ is* IID *with* $\mathrm{E}\, R > 1$ *and* $\mathrm{E}\, \beta R^{1-\gamma} < 1$*. If the stationary equilibrium wealth distribution has an unbounded support, then the asymptotic saving rate (3.3) evaluated at $\hat{R} = \mathrm{E}\, R$ is nonpositive.*

Proof. Since by assumption $\mathrm{E}\, \beta R^{1-\gamma} < 1$, by Example 2.2 the asymptotic MPC is $\bar{c} = 1 - (\mathrm{E}\, \beta R^{1-\gamma})^{1/\gamma} \in (0, 1)$. Therefore using (3.3), the asymptotic saving rate evaluated at $\mathrm{E}\, R > 1$ is

$$\begin{aligned} \bar{s} = 1 - \frac{\bar{c}}{(\mathrm{E}\, R - 1)(1 - \bar{c})} \leq 0 &\iff (\mathrm{E}\, R - 1)(1 - \bar{c}) \leq \bar{c} \\ &\iff \mathrm{E}\, R(1 - \bar{c}) \leq 1. \end{aligned}$$

Since $\mathrm{E}\, R(1 - \bar{c})$ is the expected growth rate of wealth for infinitely wealthy agents, if the wealth distribution is unbounded and $\mathrm{E}\, R(1 - \bar{c}) > 1$, then wealth will grow at the top, which violates stationarity. Therefore in a stationary equilibrium, it must be $\bar{s} \leq 0$. ☐

One possible explanation for the positive and increasing saving rates is to consider models with discount factor or return heterogeneity. If $r(PD_{\beta R^{1-\gamma}}) \geq 1$, then by Theorem 2.2 we have $\bar{c} = 0$ and hence the asymptotic saving rate becomes $\bar{s} = 1 > 0$ using (3.3).[13]

3.2 Numerical example

To show the theoretical possibility of positive and increasing saving rates, we consider a numerical example calibrated from U.S. data. We present a minimal model to illustrate our theory, and a detailed comparison to the data is beyond the scope of the paper.

The agent has constant discount factor β and relative risk aversion $\gamma > 0$. We suppose that wealthy agents invest their wealth into stocks, private businesses, and a risk-free asset in constant proportions subject to a capital income tax.[14] Let R^s_t, R^b_t be the gross returns on stock and business between time $t-1$ and t, and let R^f be the gross risk-free rate. The stock return process $\{R^s_t\}$ exhibits constant expected return $\mathrm{E}\, R^s_t = \mathrm{e}^\mu$ with GARCH$(1, 1)$ innovations:

$$\log R^s_t = \mu - \frac{1}{2}\sigma^2_t + \epsilon_t, \tag{3.4a}$$

$$\epsilon_t = \sigma_t \zeta_t, \quad \zeta_t \sim \mathrm{IID}\, N(0, 1) \tag{3.4b}$$

$$\sigma^2_t = \omega + \alpha \epsilon^2_{t-1} + \rho \sigma^2_{t-1}, \tag{3.4c}$$

[13] Another possibility is to consider overlapping generations models. Stachurski and Toda (2019, Theorem 9) present a model with random birth/death and show that it is possible to have $\beta R > 1$ in equilibrium. In this case, by the proof of Proposition 3.1, we have $\bar{s} > 0$.

[14] Since our focus is the individual optimization problem (2.1), the distinction between stocks (which are subject to only aggregate shocks) and private businesses (which are subject to both aggregate and idiosyncratic shocks) is unimportant. We include both assets only to reflect the evidence on individual portfolio cited below.

where $\sigma_t > 0$ is conditional volatility, ϵ_t is a zero mean innovation, and we assume $\omega, \alpha, \rho > 0$ and $\alpha + \rho < 1$ to ensure stationarity. We model business returns parsimoniously and set

$$R_t^b = \begin{cases} \frac{1}{1-p_b} R_t^s & \text{with probability } 1 - p_b, \\ 0 & \text{with probability } p_b, \end{cases}$$

so that private businesses go bankrupt with probability p_b but otherwise business returns are perfectly correlated with the stock return with identical mean.[15] Letting τ be the capital income tax rate, the after-tax gross portfolio return is

$$R_t(\theta) := 1 + (1 - \tau)(\theta^s R_t^s + \theta^b R_t^b + \theta^f R^f - 1), \tag{3.5}$$

where $\theta = (\theta^s, \theta^b, \theta^f)$ is the portfolio with $\theta^s + \theta^b + \theta^f = 1$.

To calibrate the stock return parameters, we use the 1947–2018 monthly data for U.S. stock market returns (volume-weighted index including dividends) and risk-free rates from the updated spreadsheet of Welch and Goyal (2008).[16] Their spreadsheet contains monthly nominal stock and risk-free returns as well as the inflation. From these we construct the real gross stock and risk-free returns R_t^s, R_t^f, define the residual $\widehat{\epsilon}_t$ in (3.4a) by demeaning the log excess returns $\log R_t^s - \log R_t^f$, and estimate the GARCH parameters $\omega = 9.1297 \times 10^{-5}$, $\alpha = 0.8354$, and $\rho = 0.1188$. We estimate the log risk-free rate as $\log R_f = \mathrm{E}[\log R_t^f] = 5.3477 \times 10^{-4}$ (annual rate 0.65%). We estimate the log expected return as $\mu = \log(\mathrm{E}\, R_t^s) = 6.8011 \times 10^{-3}$ (annual rate 8.19%). Because our model requires a finite state Markov chain, we discretize the GARCH(1, 1) process (3.4) using the Farmer and Toda (2017) method as described in Appendix A with $N_v = 3$ points for the volatility state and $N_\epsilon = 15$ for the return state.

To calibrate the portfolio shares $\theta = (\theta^s, \theta^b, \theta^f)$, we use the 1913–2012 wealth share data of the wealthiest households in U.S. estimated by Saez and Zucman (2016). Specifically, in Table B5b of their Online Appendix, they report the composition of wealth of the top 0.01% across asset groups (equities, fixed income claims, housing, business assets, and pensions). We classify equities and pension as "stock", business assets as "business", and fixed income claims and housing as "risk-free asset" to compute the portfolio share θ for all years,[17] take the average across all years, and obtain $(\theta^s, \theta^b, \theta^f) = (0.5546, 0.0827, 0.3627)$.

We calibrate the remaining parameters as follows. The discount factor is $\beta = \mathrm{e}^{-0.05/12}$ so that the annual discounting is 5%. The bankruptcy probability is $p_b = 1 - \mathrm{e}^{-0.025/12}$ so that the annual exit rate is 2.5% as documented in Luttmer (2010) for firms with more than 500 employees. The capital income tax rate is $\tau = 0.25$ based on the estimate in McDaniel (2007) using national account statistics.

To solve the income fluctuation problem (2.1), we need to specify the income process. Because the U.S. economy has been growing, and by Corollary 2.6

[15] Since business returns are modeled as a mean-preserving spread of stock returns, risk averse and unconstrained agents would never hold business assets. Here we suppose that wealthy agents hold business assets for other reasons, for example retaining voting rights in shareholder meetings.

[16] http://www.hec.unil.ch/agoyal/docs/PredictorData2018.xlsx.

[17] These portfolio shares are relatively stable over time. Although the classification of housing and pension may be ambiguous, because these two categories comprise a small fraction (about 10%) of the portfolio, choosing different classifications give quantitatively similar results.

the details on the income process is irrelevant for the asymptotic MPCs, for simplicity we assume that income grows at a constant rate g, so $Y_t = \mathrm{e}^{gt}$. We calibrate the growth rate g from the U.S. real per capita GDP in 1947–2018 and obtain $g = 1.6208 \times 10^{-3}$ at the monthly frequency. Although the theory in Ma, Stachurski, and Toda (2020) requires a stationary process for income, it is straightforward to allow for constant growth in income by detrending the model when the utility function is CRRA. After simple algebra, it suffices to use

$$\begin{aligned}\tilde{R}_t(\theta) &= R_t(\theta)\mathrm{e}^{-g},\\ \tilde{\beta} &= \beta \mathrm{e}^{(1-\gamma)g},\\ \tilde{Y}_t &= Y_t \mathrm{e}^{-gt} = 1.\end{aligned}$$

In the current setting, Assumption 1' and conditions 1 and 3 of Assumption 2 obviously hold. To apply Theorems 2.1 and 2.2, it remains to verify $r(PD_{\beta R}) < 1$ and determined whether $r(PD) \gtrless 1$, where $D = D_{\beta R^{1-\gamma}}$. Figure 1 shows the determination of the asymptotic MPC $\bar{c}(z)$ when we change the relative risk aversion γ and the annual discount rate δ. We see that the asymptotic MPCs can be zero if relative risk aversion is moderately high (above 4–5).

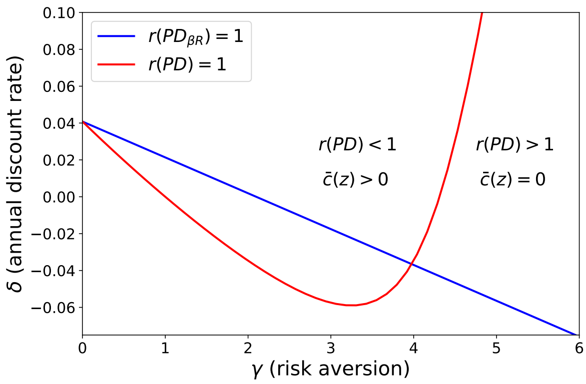

Figure 1: Determination of asymptotic MPCs with GARCH(1, 1) returns.

Is the possibility of zero asymptotic MPC empirically plausible? To address this concern, we do a simple calculation similar to Friend and Blume (1975). Although our paper abstracts from portfolio choice, suppose that wealthy agents choose the portfolio θ^s, θ^f (fixing θ^b) by maximizing the certainty equivalent of return $\mathrm{E}[R(\theta)^{1-\gamma}]^{\frac{1}{1-\gamma}}$, where $R(\theta)$ is the gross portfolio return in (3.5) and the expectation is taken over the ergodic distribution of asset returns. The first-order condition of this optimization problem is

$$\mathrm{E}[R(\theta)^{-\gamma}(R^s - R^f)] = 0. \tag{3.6}$$

Using our discretized asset returns and the portfolio share θ, the relative risk aversion that makes (3.6) hold is $\gamma = 6.38$. According to Figure 1, this level of risk aversion makes the asymptotic MPC $\bar{c}(z)$ equal to zero for any reasonable discount rates.

We next solve the model for $\gamma = 3, 5$ using policy function iteration. According to Figure 1 and Theorem 2.1, a unique solution exists in each case given the calibrated discount factor. Figure 2 shows the optimal consumption rule. Consistent with our theory, for $\gamma = 3$ the consumption functions are approximately linear with positive slopes for high asset level. When $\gamma = 5$, the consumption functions show a more concave pattern.

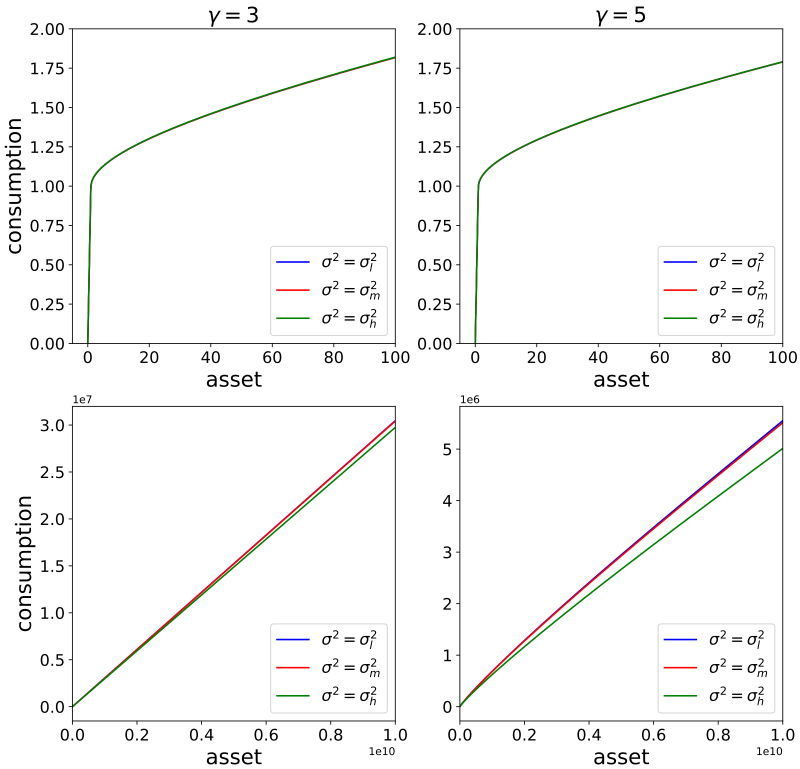

Figure 2: Optimal consumption rule.

Note: The top and bottom panels plot the consumption functions in the range $a \in [0, 100]$ and $a \in [0, 10^{10}]$, respectively. Here and in other figures, the left (right) panels correspond to $\gamma = 3$ ($\gamma = 5$). For visibility, we plot across asset and the three volatility states $\sigma_l^2 < \sigma_m^2 < \sigma_h^2$ holding $\epsilon = 0$ constant.

Figure 3 plots the consumption rates (solid lines) in log-log scale. We see that the consumption rates are decreasing in wealth for each realized volatility. For $\gamma = 3$, as asset level gets large, the asymptotic MPCs approach to positive constants that coincide with the theoretical values calculated based on Theorem 2.2 (dotted lines), indicating that the consumption functions are asymptotically linear, consistent with the theorem. For $\gamma = 5$, the consumption rates exhibit a clear decreasing trend even when asset is extremely large ($a \approx 10^{10}$), which is consistent with zero asymptotic MPC established in Theorem 2.2.

Finally, Figure 4 shows the saving rates assuming $\sigma_t^2 = \sigma_{t+1}^2 \in \left\{\sigma_l^2, \sigma_m^2, \sigma_h^2\right\}$ and $\epsilon = 0$. When wealth is low, the borrowing constraint binds and labor in-

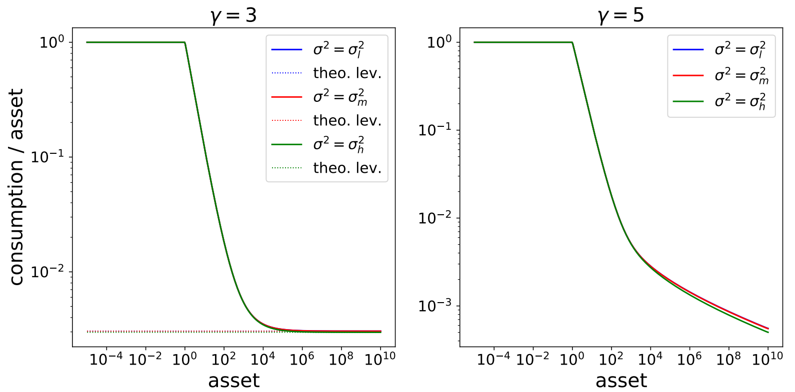

Figure 3: Consumption rate.

come is the only source of income and net worth accumulation, i.e., $s_{t+1} = (Y_{t+1} - a)/Y_{t+1} = 1 - e^{-g}a$, which is decreasing in asset. A moderately greater wealth implies lower saving rates because capital income is used to finance disproportionately large consumption. The saving rate starts to increase when wealth is relatively high (≈ 100). Importantly, when $\gamma = 5$ and $\sigma^2 \in \{\sigma_l^2, \sigma_m^2\}$, the saving rate is increasing in wealth among agents with large asset and the asymptotic saving rate equals 1,[18] as opposed to the increasing but either negative or small positive saving rate when $\gamma = 3$. This example illustrates that the empirically observed large positive and increasing saving rate could potentially be explained by models with capital income risk, particularly those with zero asymptotic MPCs.

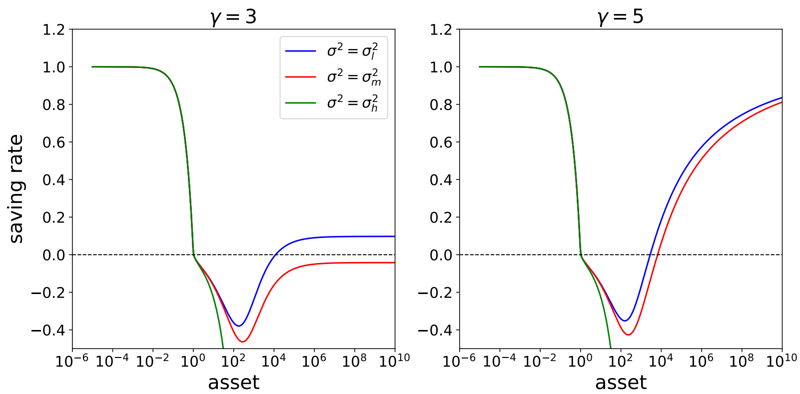

Figure 4: Saving rate.

4 Proofs

The proof of Theorem 2.2 is technical and consists of the following steps:

[18] When $\sigma^2 = \sigma_h^2$, the saving rate becomes negative because $\hat{R} < 1$ when $\epsilon = 0$; see (3.2).

1. show that policy function iteration leads to increasingly tighter upper bounds on consumption functions that are asymptotically linear with explicit slopes,

2. show that the slopes of the upper bounds converge using the fixed point theory of monotone convex maps, and

3. show that the consumption functions have linear lower bounds with identical slopes to the limit of upper bounds, implying asymptotic linearity.

Let $\mathcal{C}$ be the space of candidate consumption functions and $T : \mathcal{C} \to \mathcal{C}$ be the time iteration operator as defined in Section 2. The following proposition allows us to asymptotically bound the consumption rate $c(a,z)/a$ from above.

Proposition 4.1. *Let everything be as in Theorem 2.2. If $c \in \mathcal{C}$ and*

$$\limsup_{a\to\infty} \frac{c(a,z)}{a} \le x(z)^{-1/\gamma}$$

for some $x(z) \ge 1$ for all $z \in \mathsf{Z}$, then

$$\limsup_{a\to\infty} \frac{Tc(a,z)}{a} \le (Fx)(z)^{-1/\gamma}. \tag{4.1}$$

Proof. Let $\alpha = \limsup_{a\to\infty} Tc(a,z)/a$. By definition, we can take an increasing sequence $\{a_n\}$ such that $\alpha = \lim_{n\to\infty} Tc(a_n,z)/a_n$. Define $\alpha_n = Tc(a_n,z)/a_n \in (0,1]$ and

$$\lambda_n = \frac{c(\hat{R}(1-\alpha_n)a_n + \hat{Y}, \hat{Z})}{a_n} > 0. \tag{4.2}$$

Let us show that

$$\limsup_{n\to\infty} \lambda_n \le x(\hat{Z})^{-1/\gamma}\hat{R}(1-\alpha). \tag{4.3}$$

To see this, if $\alpha < 1$ and $\hat{R} > 0$, then since $\hat{R}(1-\alpha_n)a_n \to \hat{R}(1-\alpha)\cdot\infty = \infty$, by assumption we have

$$\begin{aligned}
\limsup_{n\to\infty} \lambda_n &= \limsup_{n\to\infty} \frac{c(\hat{R}(1-\alpha_n)a_n + \hat{Y}, \hat{Z})}{\hat{R}(1-\alpha_n)a_n + \hat{Y}} \left(\hat{R}(1-\alpha_n) + \frac{\hat{Y}}{a_n}\right) \\
&\le \limsup_{a\to\infty} \frac{c(a,\hat{Z})}{a} \times \hat{R}(1-\alpha) \\
&\le x(\hat{Z})^{-1/\gamma}\hat{R}(1-\alpha),
\end{aligned}$$

which is (4.3). If $\alpha = 1$ or $\hat{R} = 0$, then since $c(a,z) \le a$, we have

$$\begin{aligned}
\lambda_n &= \frac{c(\hat{R}(1-\alpha_n)a_n + \hat{Y}, \hat{Z})}{\hat{R}(1-\alpha_n)a_n + \hat{Y}} \left(\hat{R}(1-\alpha_n) + \frac{\hat{Y}}{a_n}\right) \\
&\le \hat{R}(1-\alpha_n) + \frac{\hat{Y}}{a_n} \to \hat{R}(1-\alpha) = 0,
\end{aligned}$$

so again (4.3) holds.

Since $\xi_n \coloneqq Tc(a_n, z) = \alpha_n a_n$ solves the Euler equation, using $u'(c) = c^{-\gamma}$ and the definition of λ_n in (4.2), we have

$$\begin{aligned}
0 &= \frac{u'(\alpha_n a_n)}{u'(a_n)} - \max\left\{\mathrm{E}_z\,\hat{\beta}\hat{R}\frac{u'(c(\hat{R}(1-\alpha_n)a_n + \hat{Y}, \hat{Z}))}{u'(a_n)}, 1\right\} \\
&= \alpha_n^{-\gamma} - \max\left\{\mathrm{E}_z\,\hat{\beta}\hat{R}(c(\hat{R}(1-\alpha_n)a_n + \hat{Y}, \hat{Z})/a_n)^{-\gamma}, 1\right\} \\
&= \alpha_n^{-\gamma} - \max\left\{\mathrm{E}_z\,\hat{\beta}\hat{R}\lambda_n^{-\gamma}, 1\right\} \\
\implies \alpha_n^{-\gamma} &= \max\left\{\mathrm{E}_z\,\hat{\beta}\hat{R}\lambda_n^{-\gamma}, 1\right\} \geq \mathrm{E}_z\,\hat{\beta}\hat{R}\lambda_n^{-\gamma}. \qquad (4.4)
\end{aligned}$$

Now letting $n \to \infty$ in (4.4) and applying Fatou's lemma, we obtain

$$\begin{aligned}
\alpha^{-\gamma} = \lim_{n\to\infty} \alpha_n^{-\gamma} &\geq \liminf_{n\to\infty} \mathrm{E}_z\,\hat{\beta}\hat{R}\lambda_n^{-\gamma} \\
&\geq \mathrm{E}_z \liminf_{n\to\infty} \hat{\beta}\hat{R}\lambda_n^{-\gamma} \\
&= \mathrm{E}_z\,\hat{\beta}\hat{R}\left[\limsup_{n\to\infty} \lambda_n\right]^{-\gamma} \\
&\geq \mathrm{E}_z\,\hat{\beta}\hat{R}\left[x(\hat{Z})^{-1/\gamma}\hat{R}(1-\alpha)\right]^{-\gamma}
\end{aligned}$$

by (4.3). Solving the inequality for α and using the convention $\beta R^{1-\gamma} = (\beta R)R^{-\gamma}$ and $0 \cdot \infty = 0$ (see Footnote 7), we obtain

$$\limsup_{a\to\infty} \frac{Tc(a,z)}{a} = \alpha \leq \frac{1}{1 + \left(\mathrm{E}_z\,\hat{\beta}\hat{R}^{1-\gamma}x(\hat{Z})\right)^{1/\gamma}} = (Fx)(z)^{-1/\gamma}. \qquad \square$$

Starting from the trivial upper bound $c(a,z) \leq a$ and applying Proposition 4.1 repeatedly we obtain increasingly tighter upper bounds of $c(a,z)$. The following proposition characterizes the limits of the slopes of the upper bounds.

Proposition 4.2. *Let everything be as in Theorem 2.2. Then F in (2.8) has a fixed point $x^* \in \mathbb{R}_+^Z$ if and only if $r(PD) < 1$, in which case the fixed point is unique. Take any $x_0 \in \mathbb{R}_+^Z$ and define the sequence $\{x_n\}_{n=1}^\infty \subset \mathbb{R}_+^Z$ by*

$$x_n = Fx_{n-1} \qquad (4.5)$$

for all $n \in \mathbb{N}$. Then the followings are true.

1. *If $r(PD) < 1$, then $\{x_n\}_{n=1}^\infty$ converges to x^*.*
2. *If $r(PD) \geq 1$ and PD is irreducible, then $x_n(z) \to x^*(z) = \infty$ as $n \to \infty$ for all $z \in \mathsf{Z}$.*

Proof. Immediate from Lemmas 4.3 and 4.4 below. $\square$

Lemma 4.3. *Let $\gamma > 0$ and define $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ by $\phi(t) = (1 + t^{1/\gamma})^\gamma$. Then there exist $a \geq 1$ and $b \geq 0$ such that $\phi(t) \leq at + b$. Furthermore, we can take $a \geq 1$ arbitrarily close to 1. (The choice of b may depend on a.)*

Proof. The proof depends on $\gamma \gtrless 1$.

Case 1: $\gamma \le 1$. Let us show that we can take $a = b = 1$. Let $f(t) = 1+t-\phi(t)$. Then $f(0) = 0$ and

$$f'(t) = 1 - \phi'(t) = 1 - \gamma(1 + t^{1/\gamma})^{\gamma-1} \frac{1}{\gamma} t^{1/\gamma-1} = 1 - (t^{-1/\gamma} + 1)^{\gamma-1} \ge 0,$$

so $f(t) \ge 0$ for all $t \ge 0$. Therefore $\phi(t) \le 1 + t$.

Case 2: $\gamma > 1$. By simple algebra we obtain

$$\phi''(t) = (\gamma - 1)(t^{-1/\gamma} + 1)^{\gamma-2} \left(-\frac{1}{\gamma} t^{-1/\gamma-1} \right) < 0, \tag{4.6}$$

so ϕ is increasing and concave. Therefore $\phi(t) \le \phi(u) + \phi'(u)(t - u)$ for all t, u. Letting $a = \phi'(u)$ and $b = \max\{0, \phi(u) - \phi'(u)u\}$, we obtain $\phi(t) \le at + b$. Furthermore, since $\phi'(t) = (t^{-1/\gamma}+1)^{\gamma-1} \to 1$ as $t \to \infty$, we can take $a = \phi'(u)$ arbitrarily close to 1 by taking u large enough. □

Lemma 4.4. *Let $\gamma > 0$ and K be a $Z \times Z$ nonnegative matrix. Define $F : \mathbb{R}_+^Z \to \mathbb{R}_+^Z$ by $Fx = \phi(Kx)$, where ϕ is as in Lemma 4.3 and is applied element-wise. Then F has a fixed point $x^* \in \mathbb{R}_+^Z$ if and only if $r(K) < 1$, in which case x^* is unique.*

Take any $x_0 \in \mathbb{R}_+^Z$ and define the sequence $\{x_n\}_{n=1}^\infty \subset \mathbb{R}_+^Z$ by $x_n = Fx_{n-1}$ for all $n \in \mathbb{N}$. Then the followings are true.

1. *If $r(K) < 1$, then $\{x_n\}_{n=1}^\infty$ converges to x^*.*
2. *If $r(K) \ge 1$ and K is irreducible, then $x_n(z) \to x^*(z) = \infty$ as $n \to \infty$ for all $z \in \mathsf{Z}$.*

Proof. We divide the proof into several steps.

Step 1. If $r(K) \ge 1$, then F does not have a fixed point. If in addition K is irreducible, then $x_n(z) \to \infty$ for all $z \in \mathbb{Z}$.

We prove the contrapositive. Suppose that F has a fixed point $x^* \in \mathbb{R}_+^Z$. Since $\phi > 0$, we have $x^* \gg 0$. Since clearly $\phi(t) > t$ for all $t \ge 0$, we have $x^* = \phi(Kx^*) \gg Kx^*$. Since K is a nonnegative matrix, by the Perron-Frobenius theorem, we can take a right eigenvector $y > 0$ such that $y'K = r(K)y'$. Since $x^* \gg Kx^*$ and $y > 0$, we obtain

$$0 < y'(x^* - Kx^*) \implies r(K)y'x^* < y'x^*.$$

Dividing both sides by $y'x^* > 0$, we obtain $r(K) < 1$.

Suppose that $r(K) \ge 1$ and K is irreducible. Since K is nonnegative and ϕ is strictly increasing, $F = \phi \circ K$ is a monotone map. Therefore to show $x_n(z) \to \infty$, it suffices to show this when $x_0 = 0$. Since $x_1 = Fx_0 = F0 = 1 \ge 0$, applying F^{n-1} we obtain $x_n \ge x_{n-1}$ for all n. Since $\{x_n\}_{n=0}^\infty$ is an increasing sequence in $\mathbb{R}_+^Z$, if it is bounded, then it converges to some $x^* \in \mathbb{R}_+^Z$. By continuity, x^* is a fixed point of F, which is a contradiction. Therefore $\{x_n\}_{n=0}^\infty$ is unbounded, so $x_n(\hat{z}) \to \infty$ for at least one $\hat{z} \in \mathsf{Z}$. Since by assumption K is irreducible, for each $(z, \hat{z}) \in \mathsf{Z}^2$, there exists $m \in \mathbb{N}$ such that $K^m(z, \hat{z}) > 0$. Therefore

$$x_{m+n}(z) \ge K^m(z, \hat{z})x_n(\hat{z}) \to \infty$$

as $n \to \infty$, so $x_n(z) \to \infty$ for all $z \in \mathsf{Z}$.

Step 2. If $r(K) < 1$, then F has a unique fixed point x^* in $\mathbb{R}^{\mathsf{Z}}_+$. If we take $a \in [1, 1/r(K))$ and $b > 0$ as in Lemma 4.3, then

$$1 \leq x^* \ll (I - aK)^{-1} b 1. \tag{4.7}$$

Take any fixed point $x^* \in \mathbb{R}^{\mathsf{Z}}_+$ of F. Since $\phi(t) \geq 1$ for all $t \geq 0$, clearly $x^* \geq 1$. Since K is nonnegative and $ar(K) < 1$, the inverse $(I - aK)^{-1} = \sum_{k=0}^{\infty} (aK)^k$ exists and is nonnegative. Therefore

$$x^* = Fx^* \ll aKx^* + b1 \implies x^* \ll (I - aK)^{-1} b1,$$

which is (4.7).

The proof of existence and uniqueness uses a similar strategy to Borovička and Stachurski (2020). Clearly F is a monotone map. Using (4.6), it follows that F is convex if $\gamma \leq 1$ and concave if $\gamma \geq 1$. Define $u_0 = 0$ and $v_0 = (I - aK)^{-1} b1 \gg 0$. Then $Fu_0 = 1 \gg 0 = u_0$ and $Fv_0 = \phi(Kv_0) \ll aKv_0 + b1 = v_0$. Hence by Theorem 2.1.2 of Zhang (2013), which is based on Theorem 3.1 of Du (1990), F has a unique fixed point in $[u_0, v_0] = [0, v_0]$. Since by (4.7) any fixed point x^* must lie in this interval, it follows that F has a unique fixed point in $\mathbb{R}^{\mathsf{Z}}_+$.

Step 3. If $r(K) < 1$, then $\{x_n\}_{n=1}^{\infty}$ converges to x^*.

Let $a \in [1, 1/r(K))$, $b > 0$, and $v_0 \gg 0$ be as in the previous step. Since $Fx = \phi(Kx)$, we obtain

$$x_n = Fx_{n-1} = \phi(Kx_{n-1}) \ll aKx_{n-1} + b1.$$

Iterating, we obtain

$$\begin{aligned} x_n &\ll (aK)^n x_0 + \sum_{k=0}^{n-1} (aK)^k (b1) \\ &= (aK)^n x_0 + \sum_{k=0}^{\infty} (aK)^k (b1) - \sum_{k=n}^{\infty} (aK)^k (b1) \\ &= (aK)^n (x_0 - v_0) + v_0. \end{aligned}$$

Since $r(aK) = ar(K) < 1$, we have $(aK)^n (x_0 - v_0) \to 0$ as $n \to \infty$. Therefore $0 = u_0 \ll x_n \ll v_0$ for large enough n. Again by Theorem 2.1.2 of Zhang (2013), we have $x_n \to x^*$ as $n \to \infty$. □

The following proposition allows us to obtain explicit linear lower bounds on consumption functions.

Proposition 4.5. *Let everything be as in Theorem 2.2. Suppose $r(PD) < 1$ and let $x^* \in \mathbb{R}^{\mathsf{Z}}_{++}$ be the unique fixed point of F in (2.8). Restrict the candidate space to*

$$\mathcal{C}_0 = \{c \in \mathcal{C} \mid c(a, z) \geq \epsilon(z) a \quad \text{for all } a > 0 \text{ and } z \in \mathsf{Z}\}, \tag{4.8}$$

where $\epsilon(z) = x^*(z)^{-1/\gamma} \in (0, 1]$. Then $T\mathcal{C}_0 \subset \mathcal{C}_0$.

Proof. Suppose to the contrary that $T\mathcal{C}_0 \not\subset \mathcal{C}_0$. Then there exists $c \in \mathcal{C}_0$ such that for some $a > 0$ and $z \in \mathsf{Z}$, we have $\xi := Tc(a,z) < \epsilon(z)a$.

Since u' is strictly decreasing and $\epsilon(z) \in (0,1]$, it follows from (2.4) that

$$u'(a) \leq u'(\epsilon(z)a) < u'(\xi) = \max\left\{\mathrm{E}_z\,\hat{\beta}\hat{R}u'(c(\hat{R}(a-\xi)+\hat{Y},\hat{Z})), u'(a)\right\}.$$

Therefore it must be $u'(a) < \mathrm{E}_z\,\hat{\beta}\hat{R}u'(c(\hat{R}(a-\xi)+\hat{Y},\hat{Z}))$. Since u' is strictly decreasing and $c \in \mathcal{C}_0$, we obtain

$$\begin{aligned}u'(\epsilon(z)a) < u'(\xi) &= \mathrm{E}_z\,\hat{\beta}\hat{R}u'(c(\hat{R}(a-\xi)+\hat{Y},\hat{Z}))\\ &\leq \mathrm{E}_z\,\hat{\beta}\hat{R}u'(\epsilon(\hat{Z})(\hat{R}(a-\xi)+\hat{Y}))\\ &\leq \mathrm{E}_z\,\hat{\beta}\hat{R}u'(\epsilon(\hat{Z})\hat{R}[1-\epsilon(z)]a).\end{aligned}$$

Using $u'(c) = c^{-\gamma}$ and $\epsilon(z) = x^*(z)^{-1/\gamma}$, we obtain

$$\begin{aligned}&x^*(z) < \mathrm{E}_z\,\hat{\beta}\hat{R}^{1-\gamma}x^*(\hat{Z})[1-x^*(z)^{-1/\gamma}]^{-\gamma}\\ \iff\, &x^*(z) < \left(1+(\mathrm{E}_z\,\hat{\beta}\hat{R}^{1-\gamma}x^*(\hat{Z}))^{1/\gamma}\right)^\gamma = \left(1+(PDx^*)(z)^{1/\gamma}\right)^\gamma,\end{aligned}$$

which is a contradiction because x^* is a fixed point of F in (2.8). $\square$

With all the above preparations, we can prove Theorem 2.2.

Proof of Theoreom 2.2. Define the sequence $\{c_n\} \subset \mathcal{C}$ by $c_0(a,z) = a$ and $c_n = Tc_{n-1}$ for all $n \geq 1$. Since $Tc(a,z) \leq a$ for any $c \in \mathcal{C}$, in particular $c_1(a,z) = Tc_0(a,z) \leq a = c_0(a,z)$. Since $T : \mathcal{C} \to \mathcal{C}$ is order preserving by Lemma B.4 of Ma, Stachurski, and Toda (2020), by induction $0 \leq c_n \leq c_{n-1}$ for all n and $c(a,z) = \lim_{n\to\infty} c_n(a,z)$ exists. Then by Theorem 2.2 of Ma, Stachurski, and Toda (2020), this c is the unique fixed point of T and also the unique solution to the income fluctuation problem (2.1).

Define the sequence $\{x_n\} \subset \mathbb{R}^Z_{++}$ by $x_0 = 1$ and $x_n = Fx_{n-1}$, where F is as in (2.8). By definition, we have $c_0(a,z)/a = 1 = x_0(z)^{-1/\gamma}$, so in particular $\limsup_{a\to\infty} c_0(a,z)/a \leq x_0(z)^{-1/\gamma}$ for all $z \in \mathsf{Z}$. Since $c_n \downarrow c \geq 0$ point-wise, a repeated application of Proposition 4.1 implies that

$$0 \leq \limsup_{a\to\infty} \frac{c(a,z)}{a} \leq \limsup_{a\to\infty} \frac{c_n(a,z)}{a} \leq x_n(z)^{-1/\gamma}. \tag{4.9}$$

Case 1: $r(PD) \geq 1$ and PD is irreducible. By Proposition 4.2 we have $x_n(z) \to \infty$ for all $z \in \mathsf{Z}$. Letting $n \to \infty$ in (4.9), we obtain

$$\lim_{a\to\infty} \frac{c(a,z)}{a} = 0.$$

Case 2: $r(PD) < 1$. By Proposition 4.2 we have $x_n(z) \to x^*(z)$, where x^* is the unique fixed point of F in (2.8). Letting $n \to \infty$ in (4.9), we obtain

$$\limsup_{a\to\infty} \frac{c(a,z)}{a} \leq x^*(z)^{-1/\gamma}. \tag{4.10}$$

On the other hand, a repeated application of Proposition 4.5 implies that $c_n(a,z) \geq x^*(z)^{-1/\gamma}a$ for all $a > 0$ and $z \in \mathsf{Z}$. Since $c_n \to c$ point-wise, letting $n \to \infty$, dividing both sides by $a > 0$, and letting $a \to \infty$, we obtain

$$\liminf_{a\to\infty} \frac{c(a,z)}{a} \geq x^*(z)^{-1/\gamma}. \tag{4.11}$$

Combining (4.10) and (4.11), we obtain $\lim_{a\to\infty} c(a,z)/a = \bar{c}(z) = x^*(z)^{-1/\gamma}$. □

Proof of Proposition 2.4. If $\gamma = 1$, then $r(PD_{\beta R^{1-\gamma}}) = r(PD_\beta) < 1$ by condition 1 of Assumption 2. Suppose $\gamma \in (0,1)$. For a nonnegative matrix A and $\theta > 0$, let $A^{(\theta)} = (A(z,\hat{z})^\theta)$ be the matrix of θ-th power. Also, let $\odot$ denote the Hadamard (element-wise) product. Applying Hölder's inequality, we obtain

$$\mathrm{E}_z\, \beta R^{1-\gamma} = \mathrm{E}_z\, \beta^\gamma (\beta R)^{1-\gamma} \leq (\mathrm{E}_z\, \beta)^\gamma (\mathrm{E}_z\, \beta R)^{1-\gamma}.$$

Constructing diagonal matrices, we obtain

$$D_{\beta R^{1-\gamma}} \leq D_\beta^{(\gamma)} \odot D_{\beta R}^{(1-\gamma)}.$$

Multiplying P from left and noting that D_X is diagonal, it follows that

$$PD_{\beta R^{1-\gamma}} \leq P(D_\beta^{(\gamma)} \odot D_{\beta R}^{(1-\gamma)}) = (PD_\beta)^{(\gamma)} \odot (PD_{\beta R})^{(1-\gamma)}.$$

Applying Theorem 1 of Elsner, Johnson, and Dias da Silva (1988), we obtain

$$r(PD_{\beta R^{1-\gamma}}) \leq r(PD_\beta)^\gamma r(PD_{\beta R})^{1-\gamma} < 1$$

by conditions 1 and 2 of Assumption 2. □

The proof of Theorem 2.5 follows from the same idea as Theorem 2.2 by considering each diagonal block separately.

Proof of Theorem 2.5. Since K is a nonnegative matrix (with elements that are potentially infinite), the map F in (2.8) is monotone and therefore $\{x_n\}_{n=0}^\infty$ monotonically converges to some $x^* \in [1,\infty]^Z$. To characterize $x^*(z)$ and $\bar{c}(z)$, we consider two cases.

Case 1: There exist j, $\hat{z} \in \mathsf{Z}_j$, and $m \in \mathbb{N}$ such that $K^m(z,\hat{z}) > 0$ and $r(K_j) \geq 1$. Define the block diagonal matrix $\tilde{K} = \operatorname{diag}(K_1,\dots,K_J)$ and the sequence $\{\tilde{x}_n\}_{n=0}^\infty \subset [0,\infty]^Z$ by $\tilde{x}_0 = 1$ and iterating (2.8), where K is replaced by $\tilde{K}$. Since $K \geq \tilde{K} \geq 0$, clearly $x_n \geq \tilde{x}_n \geq 1$ for all n. Since by definition $\tilde{K}$ is block diagonal with each diagonal block irreducible, by Lemma 4.4 we have $\tilde{x}_n(z) \to \infty$ as $n \to \infty$ if and only if there exists j such that $z \in \mathsf{Z}_j$ and $r(K_j) \geq 1$. (Although Lemma 4.4 assumes the elements of K are finite, the infinite case is similar.) Replacing the vector 1 in (2.8) by 0 and iterating, we obtain

$$x_{m+n} \geq K^m x_n \geq K^m \tilde{x}_n.$$

Therefore if there exist j, $\hat{z} \in \mathsf{Z}_j$ and $m \in \mathbb{N}$ such that $K^m(z,\hat{z}) > 0$ and $r(K_j) \geq 1$, then

$$x_{m+n}(z) \geq K^m(z,\hat{z})\tilde{x}_n(\hat{z}) \to \infty$$

as $n \to \infty$, so $x^*(z) = \infty$. In this case we obtain $\bar{c}(z) = 0$ by the same argument as in the proof of Proposition 4.1.

Case 2: For all j, either $r(K_j) < 1$ or $K^m(z, \hat{z}) = 0$ for all $\hat{z} \in \mathsf{Z}_j$ and $m \in \mathbb{N}$. For any $\hat{z}$ such that $K^m(z, \hat{z}) = 0$ for all m, by (2.8) the value of $x_n(z)$ is unaffected by all previous $x_k(\hat{z})$ for $k < n$. Therefore for the purpose of computing $x_n(z)$, we may drop all rows and columns of K corresponding to such $\hat{z}$. The resulting matrix has block diagonal elements K_j with $r(K_j) < 1$ only, so this matrix has spectral radius less than 1. Therefore by Lemma 4.4, we have $x_n(z) \to x^*(z) < \infty$ as $n \to \infty$. In this case we obtain $\bar{c}(z) = x^*(z)^{-1/\gamma}$ by the same argument as in the proof of Theorem 2.2. □

References

Ömer T. Açıkgöz. On the existence and uniqueness of stationary equilibrium in Bewley economies with production. *Journal of Economic Theory*, 173:18–55, January 2018. doi:10.1016/j.jet.2017.10.006.

Jess Benhabib, Alberto Bisin, and Shenghao Zhu. The wealth distribution in Bewley economies with capital income risk. *Journal of Economic Theory*, 159 (A):489–515, September 2015. doi:10.1016/j.jet.2015.07.013.

Truman F. Bewley. The permanent income hypothesis: A theoretical formulation. *Journal of Economic Theory*, 16(2):252–292, December 1977. doi:10.1016/0022-0531(77)90009-6.

Jaroslav Borovička and John Stachurski. Necessary and sufficient conditions for existence and uniqueness of recursive utilities. *Journal of Finance*, 75(3): 1457–1493, June 2020. doi:10.1111/jofi.12877.

Dan Cao. Recursive equilibrium in Krusell and Smith (1998). *Journal of Economic Theory*, 186:104978, March 2020. doi:10.1016/j.jet.2019.104978.

Christopher D. Carroll. Why do the rich save so much? In Joel B. Slemrod, editor, *Does Atlas Shrug? The Economic Consequences of Taxing the Rich*, chapter 14, pages 465–484. Harvard University Press, Cambridge, MA, 2000.

Christopher D. Carroll. Precautionary saving and the marginal propensity to consume out of permanent income. *Journal of Monetary Economics*, 56(6): 780–790, September 2009. doi:10.1016/j.jmoneco.2009.06.016.

Christopher D. Carroll. Theoretical foundations of buffer stock saving. *Quantitative Economics*, 2020. Forthcoming.

Christopher D. Carroll and Miles S. Kimball. On the concavity of the consumption function. *Econometrica*, 64(4):981–992, July 1996. doi:10.2307/2171853.

Gary Chamberlain and Charles A. Wilson. Optimal intertemporal consumption under uncertainty. *Review of Economic Dynamics*, 3(3):365–395, July 2000. doi:10.1006/redy.2000.0098.

Wilbur John Coleman, II. Solving the stochastic growth model by policy-function iteration. *Journal of Business and Economic Statistics*, 8(1):27–29, January 1990. doi:10.1080/07350015.1990.10509769.

Manjira Datta, Leonard J. Mirman, and Kevin L. Reffett. Existence and uniqueness of equilibrium in distorted dynamic economies with capital and labor. *Journal of Economic Theory*, 103(2):377–410, April 2002. doi:10.1006/jeth.2000.2789.

Mariacristina De Nardi. Wealth inequality and intergenerational links. *Review of Economic Studies*, 71(3):743–768, July 2004. doi:10.1111/j.1467-937X.2004.00302.x.

Yihong Du. Fixed points of increasing operators in ordered Banach spaces and applications. *Applicable Analysis*, 38(1-2):1–20, 1990. doi:10.1080/00036819008839957.

Karen E. Dynan, Jonathan Skinner, and Stephen P. Zeldes. Do the rich save more? *Journal of Political Economy*, 112(2):397–444, April 2004. doi:10.1086/381475.

Ludwig Elsner, Charles R. Johnson, and José António Dias da Silva. The Perron root of a weighted geometric mean of nonneagative matrices. *Linear and Multilinear Algebra*, 24(1):1–13, November 1988. doi:10.1080/03081088808817892.

Andreas Fagereng, Martin Blomhoff Holm, Benjamin Moll, and Gisle Natvik. Saving behavior across the wealth distribution: The importance of capital gains. 2019.

Leland E. Farmer and Alexis Akira Toda. Discretizing nonlinear, non-Gaussian Markov processes with exact conditional moments. *Quantitative Economics*, 8(2):651–683, July 2017. doi:10.3982/QE737.

Irwin Friend and Marshall E. Blume. The demand for risky assets. *American Economic Review*, 65(5):900–922, December 1975.

Émilien Gouin-Bonenfant and Alexis Akira Toda. Pareto extrapolation: An analytical framework for studying tail inequality. 2018. URL `https://ssrn.com/abstract=3260899`.

Greg Kaplan, Benjamin Moll, and Giovanni L. Violante. Monetary policy according to HANK. *American Economic Review*, 108(3):697–743, March 2018. doi:10.1257/aer.20160042.

Tom Krebs. Recursive equilibrium in endogenous growth models with incomplete markets. *Economic Theory*, 29(3):505–523, 2006. doi:10.1016/S0165-1889(03)00062-9.

Moritz Kuhn. Recursive equilibria in an Aiyagari-style economy with permanent income shocks. *International Economic Review*, 54(3):807–835, August 2013. doi:10.1111/iere.12018.

Ehud Lehrer and Bar Light. The effect of interest rates on consumption in an income fluctuation problem. *Journal of Economic Dynamics and Control*, 94: 63–71, September 2018. doi:10.1016/j.jedc.2018.07.004.

Huiyu Li and John Stachurski. Solving the income fluctuation problem with unbounded rewards. *Journal of Economic Dynamics and Control*, 45:353–365, August 2014. doi:10.1016/j.jedc.2014.06.003.

Bar Light. Precautionary saving in a Markovian earnings environment. *Review of Economic Dynamics*, 29:138–147, July 2018. doi:10.1016/j.red.2017.12.004.

Bar Light. Uniqueness of equilibrium in a Bewley-Aiyagari economy. *Economic Theory*, 69:435–450, 2020. doi:10.1007/s00199-018-1167-z.

Erzo G. J. Luttmer. Models of growth and firm heterogeneity. *Annual Review of Economics*, 2:547–576, 2010. doi:10.1146/annurev.economics.102308.124410.

Qingyin Ma, John Stachurski, and Alexis Akira Toda. The income fluctuation problem and the evolution of wealth. *Journal of Economic Theory*, 187: 105003, May 2020. doi:10.1016/j.jet.2020.105003.

Cara McDaniel. Average tax rates on consumption, investment, labor and capital in the OECD 1950-2003. 2007.

Atif R. Mian, Ludwig Straub, and Amir Sufi. Indebted demand. NBER Working Paper 26940, 2020. URL `https://www.nber.org/papers/w26940`.

Olivier F. Morand and Kevin L. Reffett. Existence and uniqueness of equilibrium in nonoptimal unbounded infinite horizon economies. *Journal of Monetary Economics*, 50(6):1351–1373, September 2003. doi:10.1016/S0304-3932(03)00082-5.

Vincenzo Quadrini. The importance of entrepreneurship for wealth concentration and mobility. *Review of Income and Wealth*, 45(1):1–19, March 1999. doi:10.1111/j.1475-4991.1999.tb00309.x.

Guillaume Rabault. When do borrowing constraints bind? Some new results on the income fluctuation problem. *Journal of Economic Dynamics and Control*, 26(2):217–245, February 2002. doi:10.1016/S0165-1889(00)00042-7.

Emmanuel Saez and Gabriel Zucman. Wealth inequality in the United States since 1913: Evidence from capitalized income tax data. *Quarterly Journal of Economics*, 131(2):519–578, May 2016. doi:10.1093/qje/qjw004.

Paul A. Samuelson. Lifetime portfolio selection by dynamic stochastic programming. *Review of Economics and Statistics*, 51(3):239–246, August 1969. doi:10.2307/1926559.

John Stachurski and Alexis Akira Toda. An impossibility theorem for wealth in heterogeneous-agent models with limited heterogeneity. *Journal of Economic Theory*, 182:1–24, July 2019. doi:10.1016/j.jet.2019.04.001.

Ludwig Straub. Consumption, savings, and the distribution of permanent income. 2019.

Ken'ichiro Tanaka and Alexis Akira Toda. Discrete approximations of continuous distributions by maximum entropy. *Economics Letters*, 118(3):445–450, March 2013. doi:10.1016/j.econlet.2012.12.020.

Ken'ichiro Tanaka and Alexis Akira Toda. Discretizing distributions with exact moments: Error estimate and convergence analysis. *SIAM Journal on Numerical Analysis*, 53(5):2158–2177, 2015. doi:10.1137/140971269.

Alexis Akira Toda. Incomplete market dynamics and cross-sectional distributions. *Journal of Economic Theory*, 154:310–348, November 2014. doi:10.1016/j.jet.2014.09.015.

Alexis Akira Toda. Wealth distribution with random discount factors. *Journal of Monetary Economics*, 104:101–113, June 2019. doi:10.1016/j.jmoneco.2018.09.006.

Ivo Welch and Amit Goyal. A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, 21(4):1455–1508, July 2008. doi:10.1093/rfs/hhm014.

Zhitao Zhang. *Variational, Topological, and Partial Order Methods with Their Applications*, volume 29 of *Developments in Mathematics*. Springer, 2013. doi:10.1007/978-3-642-30709-6.

A Discretizing the GARCH$(1,1)$ process

In this appendix we explain how to discretize the GARCH$(1,1)$ process (3.4).

A.1 Constructing the grid

Let $v_t = \sigma_t^2$. Using the properties of the GARCH process, it is known that the expected conditional variance is

$$\mathrm{E}[v_t] = \frac{\omega}{1-\alpha-\rho}.$$

Therefore it is natural to take an evenly-spaced grid $\{\bar{\epsilon}_n\}_{n=1}^{N_\epsilon}$, where N_ϵ is an odd number and the largest grid point $\bar{\epsilon} := \bar{\epsilon}_{N_\epsilon}$ is some multiple of $\sqrt{\frac{\omega}{1-\alpha-\rho}}$. Because the conditional variance of the GARCH process can be quite large, it is also natural to choose an exponential grid (as discussed in Appendix A.3) $\{\bar{v}_n\}_{n=1}^{N_v}$ such that the median point of the grid is $\frac{\omega}{1-\alpha-\rho}$.

To determine the end points, let $\underline{v} = v_1$ and $\bar{v} = v_{N_v}$. In principle $v_t = \sigma_t^2$ can be arbitrarily close to ω, so we set $\underline{v} = \omega$. For $v_t = \sigma_t^2$ to remain in the grid when ϵ_{t-1} and σ_{t-1}^2 are at their maximum value, we need

$$\bar{v} \ge \omega + \alpha\bar{\epsilon}^2 + \rho\bar{v} \iff \bar{v} \ge \frac{\omega + \alpha\bar{\epsilon}^2}{1-\rho}.$$

Setting $\bar{\epsilon} = k\sqrt{\frac{\omega}{1-\alpha-\rho}}$ for some $k > 0$, we obtain

$$\bar{v} \ge \frac{1-\rho+(k^2-1)\alpha}{1-\rho}\frac{\omega}{1-\alpha-\rho}. \tag{A.1}$$

In order to be able to match up to the second moments of ϵ_t when $v_t = \bar{v}$, it is necessary and sufficient that

$$\bar{\epsilon} \geq \sqrt{\bar{v}} \iff \bar{v} \leq \bar{\epsilon}^2 = \frac{k^2\omega}{1-\alpha-\rho}. \tag{A.2}$$

We have the following result.

Proposition A.1. *Consider the GARCH$(1,1)$ process (3.4) with $\alpha+\rho<1$ and set $v_t = \sigma_t^2$. Let $N_\epsilon \geq 3$ be an odd number and $N_v \geq 2$. Then there exists a discretization such that*

1. *$\{\bar{\epsilon}_n\}_{n=1}^{N_\epsilon}$ is evenly spaced and centered around 0,*
2. *$\{\bar{v}_n\}_{n=1}^{N_v}$ is exponentially spaced with minimum point ω and median point $\frac{\omega}{1-\alpha-\rho}$, and*
3. *the conditional mean of v_t and the conditional mean and variance of ϵ_t are exact.*

Proof. Set $\bar{\epsilon} = \bar{\epsilon}_{N_\epsilon} = k\sqrt{\frac{\omega}{1-\alpha-\rho}}$ for some $k>0$. Combining (A.1) and (A.2), we obtain

$$\begin{aligned}
&\frac{1-\rho+(k^2-1)\alpha}{1-\rho}\frac{\omega}{1-\alpha-\rho} \leq \frac{k^2\omega}{1-\alpha-\rho} \\
\iff &1-\rho+(k^2-1)\alpha \leq (1-\rho)k^2 \\
\iff &(k^2-1)(\alpha+\rho-1) \leq 0,
\end{aligned}$$

which always holds if $k \geq 1$ because $\alpha+\rho<1$. Set

$$(\underline{v}, \bar{v}) = \left(\omega, \frac{1-\rho+(k^2-1)\alpha}{1-\rho}\frac{\omega}{1-\alpha-\rho}\right)$$

for some $k \geq 1$ and construct an exponential grid $\{v_n\}_{n=1}^{N_v}$ on $[\underline{v}, \bar{v}]$ with median point $\frac{\omega}{1-\alpha-\rho}$ as explained in Appendix A.3. For the exponential grid to be well-defined, we need

$$\begin{aligned}
&\frac{\omega}{1-\alpha-\rho} < \frac{\underline{v}+\bar{v}}{2} = \frac{1}{2}\left(\omega + \frac{1-\rho+(k^2-1)\alpha}{1-\rho}\frac{\omega}{1-\alpha-\rho}\right) \\
\iff &2 < 1-\alpha-\rho + \frac{1-\rho+(k^2-1)\alpha}{1-\rho} \\
\iff &(1-\rho)(1+\alpha+\rho) < 1-\rho+(k^2-1)\alpha \\
\iff &(1-\rho)(\alpha+\rho) < (k^2-1)\alpha,
\end{aligned} \tag{A.3}$$

which holds for large enough $k \geq 1$ because $\alpha > 0$. To make (A.3) true, for example we can set

$$k^2 - 1 = N_\epsilon(1-\rho)(1+\rho/\alpha) \iff k = \sqrt{1+N_\epsilon(1-\rho)(1+\rho/\alpha)},$$

which satisfies $k \geq 1$.[19] In this case we have

$$\bar{v} = (1+N_\epsilon(\alpha+\rho))\frac{\omega}{1-\alpha-\rho}. \quad \square$$

[19]Setting $k \sim \sqrt{N_\epsilon}$ is advocated in Farmer and Toda (2017) based on the trapezoidal rule for quadrature.

Applying Proposition A.1 and its proof, we can construct the grid $\{\bar{v}_n\}_{n=1}^{N_v}$ and $\{\bar{\epsilon}_n\}_{n=1}^{N_\epsilon}$ as follows.

Constructing grid for GARCH.

1. Select the number of grid points $N_\epsilon \geq 3$ and $N_v \geq 2$ to discretize the return and variance states.
2. Set $\bar{\epsilon} = \sqrt{(1 + N_\epsilon(1-\rho)(1+\rho/\alpha))\frac{\omega}{1-\alpha-\rho}}$ and construct the evenly-spaced grid $\{\bar{\epsilon}_n\}_{n=1}^{N_\epsilon}$ on $[-\bar{\epsilon}, \bar{\epsilon}]$.
3. Set
$$(a, b, c) = \left(\omega, (1 + N_\epsilon(\alpha+\rho))\frac{\omega}{1-\alpha-\rho}, \frac{\omega}{1-\alpha-\rho}\right)$$
and construct the exponentially-spaced grid $\{\bar{v}_n\}_{n=1}^{N_v}$ on $[a, b]$ with median point c as in Appendix A.3.

A.2 Constructing transition probabilities

Having constructed the grid, it remains to construct transition probabilities. Let $\mathsf{Z} = \{1, \dots, N_v\} \times \{1, \dots, N_\epsilon\}$ be the state space. If $z = (m, n) \in \mathsf{Z}$, then the current conditional variance and return are $(v, \epsilon) = (\bar{v}_m, \bar{\epsilon}_n)$. The next period's conditional variance is then

$$\hat{v} = \omega + \alpha\bar{\epsilon}_n^2 + \rho\bar{v}_m.$$

This $\hat{v}$ will in general not be a grid point. However, we can approximate the transition to $\hat{v}$ by assigning probabilities $1-\theta, \theta$ to the two points $\bar{v}_{m'}, \bar{v}_{m'+1}$ such that

$$\hat{v} = (1-\theta)\bar{v}_{m'} + \theta\bar{v}_{m'+1},$$

where m' is uniquely determined such that $\bar{v}_{m'} < \hat{v} \leq \hat{v}_{m'+1}$.

Because the distribution of $\hat{\epsilon}$ is $N(0, \hat{v})$, we can assign probabilities on the grid points $\{\bar{\epsilon}_n\}_{n=1}^{N_\epsilon}$ such that the mean and variance are exact. For this purpose, we can use the maximum entropy method of Tanaka and Toda (2013, 2015) and Farmer and Toda (2017). If $N_\epsilon = 3$, we can avoid optimizing because there is a closed-form solution as follows. Assign probabilities $p, 1-2p, p$ to points $-\bar{\epsilon}, 0, \bar{\epsilon}$ so that $\mathrm{E}[\epsilon] = 0$ and $\mathrm{Var}[\epsilon] = \hat{v}$. For this purpose, we can set

$$\hat{v} = 2p\bar{\epsilon}^2 \iff p = \frac{\hat{v}}{2\bar{\epsilon}^2},$$

which is always in $(0, 1/2)$ because $\hat{v} \leq \bar{v} < \bar{\epsilon}^2$.

A.3 Exponential grid

In many models, the state variable may become negative (e.g., asset holdings), which causes a problem for constructing an exponentially-spaced grid because we cannot take the logarithm of a negative number. Suppose we would like to construct an N-point exponential grid on a given interval (a, b). A natural idea to deal with such a case is as follows.

Constructing exponential grid.

1. Choose a shift parameter $s > -a$.
2. Construct an N-point evenly-spaced grid on $(\log(a+s), \log(b+s))$.
3. Take the exponential.
4. Subtract s.

The remaining question is how to choose the shift parameter s. Suppose we would like to specify the median grid point as $c \in (a, b)$. Since the median of the evenly-spaced grid on $(\log(a+s), \log(b+s))$ is $\frac{1}{2}(\log(a+s) + \log(b+s))$, we need to take $s > -a$ such that

$$
\begin{aligned}
& c = \exp\left(\frac{1}{2}(\log(a+s) + \log(b+s))\right) - s \\
\iff & c + s = \sqrt{(a+s)(b+s)} \\
\iff & (c+s)^2 = (a+s)(b+s) \\
\iff & c^2 + 2cs + s^2 = ab + (a+b)s + s^2 \\
\iff & s = \frac{c^2 - ab}{a + b - 2c}.
\end{aligned}
$$

Note that in this case

$$
s + a = \frac{c^2 - ab}{a + b - 2c} + a = \frac{(c-a)^2}{a + b - 2c},
$$

so $s + a$ is positive if and only if $c < \frac{a+b}{2}$. Therefore, for any $c \in \left(a, \frac{a+b}{2}\right)$, it is possible to construct an exponentially-spaced grid with end points (a, b) and median point c.