Online Appendix (not for publication)

A Asymptotic problem

In this appendix we describe how to derive the asymptotic homogeneous problem in an abstract dynamic programming setting. For the notation, we follow Ma and Stachurski (2021). Let

- X be a set called the *state space*;
- A be a set called the *action space*;
- $\Gamma : X \twoheadrightarrow A$ be a nonempty correspondence called the *feasible correspondence*;
- $g : X \times A \to X$ be a function called the *law of motion*;
- $\mathcal{V}$ be a subset of all functions from X to $\mathbb{R} \cup \{-\infty\}$ called the set of *candidate value functions*;
- $Q : X \times A \times \mathcal{V} \to \mathbb{R} \cup \{-\infty\}$ be a map called the *state-action aggregator*.

Then we say that the value function $v \in \mathcal{V}$ satisfies the Bellman equation if

$$v(x) = \max_{a \in \Gamma(x)} Q(x, a, v(g(x, a))) \tag{A.1}$$

for all $x \in X$.

Definition A.1 (Asymptotic homogeneity)**.** We say that the dynamic programming problem is *asymptotically homogeneous* if it has the following properties:

- $X = X_1 \times X_2$, where $\mathbb{R}_+ \subset X_1 \subset \mathbb{R}$;
- $\Gamma(x) = \Gamma_1(x_1, x_2) \times \Gamma_2(x_2)$, where $x = (x_1, x_2) \in X_1 \times X_2$ and $\mathbb{R}^d_+ \subset \Gamma_1(x_1, x_2) \subset \mathbb{R}^d$ for some d;
- $g(x, a) = g_1(x_1, x_2, a_1, a_2) \times g_2(x_2, a_2)$, where $x = (x_1, x_2) \in X_1 \times X_2$ and $(a_1, a_2) \in \Gamma_1(x_1, x_2) \times \Gamma_2(x_2)$;

- $\lim_{\lambda\to\infty} \frac{1}{\lambda}\Gamma_1(\lambda x_1, x_2) = \tilde{\Gamma}_1(x_1, x_2)$ exists for $(x_1, x_2) \in X_1 \times X_2$;
- $\lim_{\lambda\to\infty} \frac{1}{\lambda} g_1(\lambda x_1, x_2, \lambda a_1, a_2) = \tilde{g}_1(x_1, x_2, a_1, a_2)$ exists for $(x_1, x_2) \in X_1 \times X_2$ and $(a_1, a_2) \in \Gamma_1(x_1, x_2) \times \Gamma_2(x_2)$;
- $\lim_{\lambda\to\infty} \frac{1}{\lambda} Q(\lambda x_1, x_2, \lambda a_1, a_2, \lambda v) = \tilde{Q}(x_1, x_2, a_1, a_2, v)$ exists.

Lemma A.2. *Suppose that the dynamic programming problem is asymptotically homogeneous. Then*

(i) $\tilde{\Gamma}_1$ *is homogeneous of degree 1 in* x_1*: for any* $\lambda > 0$ *we have*

$$\tilde{\Gamma}_1(\lambda x_1, x_2) = \lambda \tilde{\Gamma}_1(x_1, x_2).$$

(ii) $\tilde{g}_1$ *is homogeneous of degree 1 in* (x_1, a_1)*: for any* $\lambda > 0$ *we have*

$$\tilde{g}_1(\lambda x_1, x_2, \lambda a_1, a_2) = \lambda \tilde{g}_1(x_1, x_2, a_1, a_2).$$

(iii) $\tilde{Q}$ *is homogeneous of degree 1 in* (x_1, a_1, v)*: for any* $\lambda > 0$ *we have*

$$\tilde{Q}(\lambda x_1, x_2, \lambda a_1, a_2, \lambda v) = \lambda \tilde{Q}(x_1, x_2, a_1, a_2, v).$$

Proof. By the definition of $\tilde{\Gamma}_1$, for any $\lambda > 0$ we have

$$\begin{aligned}\tilde{\Gamma}_1(\lambda x_1, x_2) &= \lim_{\lambda'\to\infty} \frac{1}{\lambda'}\Gamma_1(\lambda'\lambda x_1, x_2) \\ &= \lambda \lim_{\lambda'\to\infty} \frac{1}{\lambda'\lambda}\Gamma_1(\lambda'\lambda x_1, x_2) = \lambda\tilde{\Gamma}_1(x_1, x_2).\end{aligned}$$

The proofs of the other claims are similar. □

When the dynamic programming problem is asymptotically homogeneous, we define the asymptotic problem as follows.

Definition A.3. Suppose that the dynamic programming problem is asymptotically homogeneous. Then the Bellman equation of the asymptotic problem corresponding to (A.1) is defined by

$$v(x_1, x_2) = \max_{(a_1,a_2)\in\tilde{\Gamma}_1(x_1,x_2)\times\Gamma_2(x_2)} \tilde{Q}(x_1, x_2, a_1, a_2, v(\tilde{g}_1(x_1, x_2, a_1, a_2), g_2(x_2, a_2))). \tag{A.2}$$

The following lemma shows that we can reduce the dimension of the asymptotic problem by 1.

Lemma A.4. *Suppose that the dynamic programming problem is asymptotically homogeneous. Consider the following "normalized" Bellman equation:*

$$\tilde{v}(x_2) = \max_{(a_1,a_2)\in\tilde{\Gamma}_1(1,x_2)\times\Gamma_2(x_2)} \tilde{Q}(1, x_2, a_1, a_2, \tilde{g}_1(1, x_2, a_1, a_2)\tilde{v}(g_2(x_2, a_2))). \tag{A.3}$$

If (A.3) has a solution $\tilde{v}(x_2)$, then $v(x_1, x_2) = x_1\tilde{v}(x_2)$ is a solution to the asymptotic Bellman equation (A.2). Furthermore, letting $\tilde{a} = (\tilde{a}_1, \tilde{a}_2)$ be the policy function of the normalized Bellman equation (A.3), the policy function $a = (a_1, a_2)$ of the asymptotic Bellman equation (A.2) is given by $a_1(x_1, x_2) = x_1\tilde{a}_1(x_2)$ and $a_2(x_1, x_2) = \tilde{a}_2(x_2)$.

Proof. Immediate by multiplying both sides of (A.3) by $x_1 > 0$ and using the homogeneity of $\tilde{\Gamma}_1, \tilde{g}_1, \tilde{Q}$ established in Lemma A.2. □

The following proposition shows that if a dynamic programming problem is asymptotically homogeneous, then the value function and policy functions are asymptotically linear. (The statement and proof are somewhat heuristic but see Ma and Toda (2021a) for a rigorous treatment for the case of an income fluctuation problem.)

Proposition A.5. *Suppose that the dynamic programming problem is asymptotically homogeneous. Suppose that the Bellman equation (A.1) has a solution $v(x)$, and it can be computed by value function iteration starting from $v(x) \equiv 0$. Then under some regularity conditions, the value function and policy functions are asymptotically linear: we have*

$$\begin{aligned}
v(x_1, x_2) &= x_1\tilde{v}(x_2) + o(x_1),\\
a_1(x_1, x_2) &= x_1\tilde{a}_1(x_2) + o(x_1),\\
a_2(x_1, x_2) &= \tilde{a}_2(x_2) + o(x_1)
\end{aligned}$$

as $x_1 \to \infty$, where $\tilde{v}(x_2)$, $\tilde{a}_1(x_2)$, and $\tilde{a}_2(x_2)$ are defined as in the normalized Bellman equation (A.3).

Proof. Define the operator $T : \mathcal{V} \to \mathcal{V}$ by the right-hand side of (A.1). Let $v^{(0)} \equiv 0$ and $v^{(k)} = Tv^{(k-1)} = T^k 0$. Let us show by induction that

$$\lim_{\lambda\to\infty} \frac{1}{\lambda} v^{(k)}(\lambda x_1, x_2) = \tilde{v}^{(k)}(x_1, x_2)$$

exists. If $k = 0$, the claim is trivial since $v^{(0)} \equiv 0$. Suppose the claim holds for some $k-1$. Then by Lemma A.2, we obtain

$$\begin{aligned} &\frac{1}{\lambda} v^{(k)}(\lambda x_1, x_2) = \frac{1}{\lambda}(Tv^{(k-1)})(\lambda x_1, x_2) \\ &= \max_{\substack{(a_1,a_2)\in \\ \frac{1}{\lambda}\Gamma_1(\lambda x_1,x_2)\times\Gamma_2(x_2)}} Q\left(\lambda x_1, x_2, \lambda a_1, a_2, v^{(k-1)}\left(\lambda \frac{1}{\lambda} g_1(\lambda x_1, x_2, \lambda a_1, a_2), g_2(x_2, a_2)\right)\right). \end{aligned}$$

Using the asymptotic homogeneity of Γ_1, g_1, Q established in Lemma A.2, the asymptotic homogeneity of $v^{(k-1)}$, and assuming that we can interchange the limit and maximization (e.g., assuming enough conditions to apply the Maximum Theorem), it follows that $v^{(k)}$ is asymptotically homogeneous. Since by assumption $v^{(k)} \to v$ as $k \to \infty$ point-wise, assuming that the limit of $k \to \infty$ and $\lambda \to \infty$ can be interchanged (which is the case if $v^{(k)}$ converges to v monotonically, which is often the case in particular applications), then v is asymptotically homogeneous in the sense that $\lim_{\lambda\to\infty} \frac{1}{\lambda} v(\lambda x_1, x_2)$ exists.

Now that asymptotic homogeneity of v is established, from (A.1) we obtain

$$v(\lambda x_1, x_2) = \max_{a\in\Gamma(\lambda x_1,x_2)} Q(\lambda x_1, x_2, a, v(g(\lambda x_1, x_2, a))).$$

Dividing both sides by $\lambda > 0$ and letting $\lambda \to \infty$, using the asymptotic homogeneity of Γ_1, g_1, Q, and v, we obtain the asymptotic Bellman equation (A.2). Thus if in particular (A.3) has a unique solution $\tilde{v}(x_2)$, by Lemma A.4 it must be

$$\lim_{\lambda\to\infty} \frac{1}{\lambda} v(\lambda x_1, x_2) = x_1 \tilde{v}(x_2).$$

Consequently, setting $x_1 = 1$ and $\lambda = x_1$, we obtain $v(x_1, x_2) = x_1\tilde{v}(x_2) + o(x_1)$. The proof for the policy functions is similar. □

B Proofs

Proposition B.1. *The asymptotic Euler equation (3.3) admits a (necessarily unique) positive solution* $\{\bar{c}_s\}_{s=1}^S$ *if and only if the spectral condition (3.4) holds.*

Proof. Setting $x_s = \bar{c}_s^{-\gamma}$, (3.3) can be rewritten as

$$x_s = \left(1 + \left((1-p)R^{1-\gamma}\beta_s \sum_{s'=1}^S p_{ss'}x_{s'}\right)^{1/\gamma}\right)^\gamma.$$

Setting $x = (x_1, \dots, x_S)'$, we can express this equation as

$$x = \left(1 + (Kx)^{1/\gamma}\right)^\gamma, \tag{B.1}$$

where $K = (1-p)R^{1-\gamma}DP$, $D = \operatorname{diag}(\beta_1, \dots, \beta_S)$, and powers are applied entry-wise. By Proposition 14 of Ma and Toda (2021a) (see also discussions in Toda (2019) and Borovička and Stachurski (2020)), (B.1) has a positive solution if and only if $\rho(K) < 1$, in which case the solution is unique. Since $\rho(K) = (1-p)R^{1-\gamma}\rho(DP)$, a necessary and sufficient condition for the existence of a solution is the spectral condition (3.4). □

Proof of Proposition 3.1. If $G_{t+1} \le 1$ always, then by (3.7) we have

$$M_{ss'}(z) = \mathrm{E}\left[G_{t+1}^z \,\middle|\, s_t = s, s_{t+1} = s'\right] \le 1$$

for all $z \ge 0$. Therefore $(1-p)\rho(P \odot M(z)) \le (1-p)\rho(P) = 1-p < 1$, so (3.9) does not have a solution $z > 0$.

Suppose that $M(z)$ is finite for all $z > 0$ and P is irreducible. Define $A(z) = P \odot M(z)$. Define the $S \times S$ matrix $B(z)$ by $B_{ss}(z) = p_{ss}M_{ss}(z) > 0$ for the s satisfying the assumption $p_{ss}\Pr(G_{t+1} > 1 \mid s_t = s_{t+1} = s) > 0$, and 0 for all other entries. Then clearly $A(z) \ge B(z) \ge 0$ entry-wise, so

$$\infty > \rho(P \odot M(z)) = \rho(A(z)) \ge \rho(B(z)) = p_{ss}\,\mathrm{E}\left[G_{t+1}^z \,\middle|\, s_t = s_{t+1} = s\right] \to \infty$$

as $z \to \infty$. Since $(1-p)\rho(P \odot M(0)) = (1-p)\rho(P) = 1-p < 1$, by the intermediate value theorem there exists $z = \zeta > 0$ such that (3.9) holds. Uniqueness is proved in Beare and Toda (2017). □

Proof of Proposition 4.1. Suppose $\{w_t\}$ is a stationary solution to

$$w_t = \begin{cases} G_t w_{t-1} & \text{with probability } 1-p, \\ X & \text{with probability } p, \end{cases}$$

where X is a positive random variable with bounded support.[18] Then for large $w > 0$, we have

$$\begin{aligned} \Pr(s_t = s', w_t > w) &= \Pr(s_t = s') \Pr(w_t > w \mid s_t = s') \\ &= \sum_{s=1}^{S} (1-p) \Pr(s_{t-1} = s) p_{ss'} \operatorname{E}_G[\Pr(w_{t-1} > w/G_t \mid s_{t-1} = s, G_t)], \end{aligned}$$

where the expectation over G is taken conditional on $(s_{t-1}, s_t) = (s, s')$. We know from Beare and Toda (2017) that $\Pr(w_t > w \mid s_t = s) \sim c_s w^{-\zeta}$ for some constant $c_s > 0$ when w is large. Hence setting $\pi_s = \Pr(s_t = s)$, we can write the above equation as

$$\pi_{s'} c_{s'} w^{-\zeta} \sim \sum_{s=1}^{S} (1-p) \pi_s p_{ss'} \operatorname{E}\left[c_s (w/G_t)^{-\zeta} \,\middle|\, s_{t-1} = s, s_t = s'\right].$$

Multiplying both sides by w^{ζ} and letting $w \to \infty$, the approximation becomes exact and we obtain

$$\pi_{s'} c_{s'} = \sum_{s=1}^{S} (1-p) \pi_s c_s p_{ss'} M_{ss'}(\zeta),$$

where $M_{ss'}$ is the conditional moment generating function of $\log G$ in (3.7). Expressing this in matrix form, we obtain

$$y' = y'(1-p) P \odot M(\zeta),$$

where $y' = (\pi_1 c_1, \dots, \pi_S c_S)$. Therefore y is the left eigenvector of $(1-p) P \odot M(\zeta)$ corresponding to the eigenvalue 1. Using the Bayes rule, the

[18] More generally, it suffices to assume that $\operatorname{E}[X^{\zeta+\epsilon}] < \infty$ for some $\epsilon > 0$.

distribution of states conditional on being in the tail is

$$\Pr(s_t = s \mid w_t > w) = \frac{\Pr(s_t = s, w_t > w)}{\Pr(w_t > w)} \sim \frac{\pi_s c_s w^{-\zeta}}{\sum_{s=1}^{N} \pi_s c_s w^{-\zeta}} = \frac{y_s}{\sum_{s=1}^{S} y_s},$$

so $\bar{\pi}$ is the normalized left Perron vector of $P \odot M(\zeta)$.

By a similar argument, the probability that a type s agent remains in the top tail is

$$\begin{aligned}
&\Pr(w_{t+1} > w \mid w_t > w, s_t = s) \\
&= (1-p)\Pr(w_t > w/G_{t+1} \mid w_t > w, s_t = s) \\
&= (1-p)\frac{\Pr(w_t > w/G_{t+1}, w_t > w, s_t = s)}{\Pr(w_t > w, s_t = s)} \\
&= (1-p)\frac{\Pr\left(w_t > \max\left\{1, G_{t+1}^{-1}\right\} w, s_t = s\right)}{\Pr(w_t > w, s_t = s)} \\
&\sim (1-p)\frac{\mathrm{E}\left[\pi_s c_s \left(\max\left\{1, G_{t+1}^{-1}\right\} w\right)^{-\zeta} \,\middle|\, s_t = s\right]}{\pi_s c_s w^{-\zeta}} \\
&= (1-p)\,\mathrm{E}\left[\min\left\{1, G_{t+1}^{\zeta}\right\} \,\middle|\, s_t = s\right].
\end{aligned}$$

Therefore (4.5) holds. □

Proof of Proposition 4.2. 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 $a < c < \frac{a+b}{2}$. Therefore, for any such c, there exists an exponential grid with median point c. □

C Theoretical properties of analytical model

In this appendix we analytically characterize the equilibrium of the model in Section 5. We consider the firm's problem, the single agent problem, and the existence of a stationary equilibrium.

C.1 Firm's problem

The firm's problem (5.3) is static. Suppressing the time subscript and noting that the production function is Cobb-Douglass $AF(K, L) = AK^\alpha L^{1-\alpha}$, the first-order conditions of the profit maximization problem are

$$\omega = AF_L(K, L) = A(1-\alpha)(K/L)^\alpha,$$
$$R - 1 + \delta = AF_K(K, L) = A\alpha(K/L)^{\alpha-1}.$$

Solving for the capital-labor ratio, we obtain

$$\left(\frac{R-1+\delta}{A\alpha}\right)^{\frac{1}{\alpha-1}} = \frac{K}{L} = \left(\frac{\omega}{A(1-\alpha)}\right)^{\frac{1}{\alpha}}. \tag{C.1}$$

In particular, the wage ω and gross risk-free rate R are related as

$$\omega = A^{\frac{1}{1-\alpha}}(1-\alpha)\left(\frac{R-1+\delta}{\alpha}\right)^{\frac{\alpha}{\alpha-1}}. \tag{C.2}$$

C.2 Single agent problem

For simplicity we first consider a single agent problem without mortality risk ($p = 0$) and general minimum consumption, income, and gross risk-free rate $\{\bar{c}_t, y_t, R_t\}_{t=0}^\infty$. Thus the optimization problem becomes

$$\begin{aligned} &\text{maximize} && \mathrm{E}_0 \sum_{t=0}^{\infty} \left(\prod_{i=0}^{t-1} \beta_{s_i}\right) \log(c_t - \underline{c}_t) \\ &\text{subject to} && a_{t+1} = R_{t+1}(a_t - c_t + y_t), \end{aligned}$$

where a_t is financial wealth at the beginning of time t excluding labor income, R_{t+1} is risk-free rate from time t to $t+1$, and c_t is consumption. The agent takes $\{\underline{c}_t, y_t, R_t\}_{t=0}^{\infty}$ as given (perfect foresight). For survival we assume $0 \le \underline{c}_t < y_t$ for all t.

We can solve this problem as follows. First, define the "surplus" consumption $x_t = c_t - \underline{c}_t$. Then the objective function is the standard log utility and the budget constraint becomes

$$a_{t+1} = R_{t+1}(a_t - x_t - \underline{c}_t + y_t) = R_{t+1}(a_t - x_t + \tilde{y}_t),$$

where $\tilde{y}_t \coloneqq y_t - \underline{c}_t > 0$ is "disposable income" after financing the minimum consumption $\underline{c}_t$. To convert the problem to a homogeneous problem, we define the "human wealth" h_t to satisfy

$$h_{t+1} = R_{t+1}(h_t - \tilde{y}_t),$$

so

$$h_t = \sum_{i=0}^{\infty} \left(\prod_{\ell=1}^{i} R_{t+\ell} \right)^{-1} \tilde{y}_{t+i}, \tag{C.3}$$

where we use the convention $\prod_{\ell=1}^{0} R_{t+\ell} = 1$. Using the budget constraint and the definition of human wealth, we obtain

$$w_{t+1} = R_{t+1}(w_t - x_t),$$

where $w_t = a_t + h_t$ is total (financial plus human) wealth. Now the problem reduces to

$$\begin{aligned} &\text{maximize} && \mathrm{E}_0 \sum_{t=0}^{\infty} \left(\prod_{i=0}^{t-1} \beta_{s_i} \right) \log x_t \\ &\text{subject to} && w_{t+1} = R_{t+1}(w_t - x_t) \ge 0, \end{aligned}$$

which is a homogeneous problem. Because $w_t = a_t + h_t \ge 0$, the natural borrowing constraint is then

$$a_t \ge -h_t, \tag{C.4}$$

where the human wealth h_t is defined by (C.3).

Let $V_{t,s}(w)$ be the value function in period t given state s and wealth w. The Bellman equation is then

$$V_{t,s}(w) = \max_x \left\{\log x + \beta_s \operatorname{E}\left[V_{t+1,s'}(R_{t+1}(w-x)) \mid s\right]\right\}.$$

Guess that the value function takes the form $V_{t,s}(w) = A_{t,s} + B_s \log w$ for some constants $A_{t,s}, B_s$, with $B_s > 0$ depending only on s. The first-order condition is

$$0 = \frac{1}{x} - \beta_s \operatorname{E}\left[B_{s'} \mid s\right] \frac{1}{w-x} \iff x = \frac{w}{1 + \beta_s \operatorname{E}\left[B_{s'} \mid s\right]}.$$

Substituting this x into the Bellman equation and comparing the coefficients of $\log w$, we obtain

$$B_s = 1 + \beta_s \operatorname{E}\left[B_{s'} \mid s\right].$$

Setting $B = (B_1, \dots, B_S)'$ and expressing using a matrix, we obtain

$$B = 1 + DPB \iff B = (I - DP)^{-1} 1,$$

where $D = \operatorname{diag}(\dots, \beta_s, \dots)$ is the diagonal matrix with s-th diagonal element β_s. The optimal consumption rule is then $x = w/B_s$ in state s. The case with constant β corresponds to $D = \beta$ and $P = 1$, so $B = 1/(1-\beta)$ and the optimal consumption rule is $x = (1-\beta)w$, as is well known. See the Online Appendix in Toda (2019) for a more complete analysis.

Let $m_s = 1/B_s < 1$ be the asymptotic marginal propensity to consume in state s. Substituting the optimal consumption rule $x = m_s w$ into the budget constraint, we get $w_{t+1} = R_{t+1}(1 - m_{s_t}) w_t$, or

$$\begin{aligned} a_{t+1} &= R_{t+1}(1 - m_{s_t})(a_t + h_t) - h_{t+1} \\ &= R_{t+1}((1 - m_{s_t}) a_t - m_{s_t} h_t + \tilde{y}_t) \end{aligned}$$

after using the definition of human wealth in (C.3). Note that the savings (capital holdings) in period t is given by

$$\begin{aligned}k_{t+1} = a_t - x_t + \tilde{y}_t &= \frac{a_{t+1}}{R_{t+1}} \\ &= (1 - m_{s_t})a_t - m_{s_t}h_t + \tilde{y}_t.\end{aligned}$$

If the agent is subject to mortality risk, then all the above derivations are valid by replacing β_s, R with the effective discount factor $\tilde{\beta}_s = \beta_s(1 - p)$ and effective gross risk-free rate $\tilde{R}_t = R_t/(1 - p)$. We collect these results into the following proposition.

Proposition C.1. *Consider an agent with log utility, labor endowment 1, discount factor* (β_s) *with transition probability matrix* $P = (p_{ss'})$*, birth/death probability* p*, and minimum consumption ratio* ϕ*. Suppose that* $(1-p)\rho(DP) < 1$*, where* $D = \operatorname{diag}(\dots, \beta_s, \dots)$*, and let* $B = (I - (1 - p)DP)^{-1}1 \gg 1$ *and* $m_s = 1/B_s < 1$*.*

Let the sequence of wage and gross risk-free rate $\{\omega_t, R_t\}_{t=0}^{\infty}$ *be given,* $\tilde{R}_t = R_t/(1 - p)$*, and assume* $\limsup_{t\to\infty} \omega_t < \infty$ *and* $\liminf_{t\to\infty} \tilde{R}_t > 1$*. Then the optimal consumption of the agent is given by*

$$c_t = m_{s_t}(a_t + h_t) + \phi\omega_t, \tag{C.5}$$

where

$$h_t = (1 - \phi) \sum_{i=0}^{\infty} \left(\prod_{\ell=1}^{i} \tilde{R}_{t+\ell} \right)^{-1} \omega_{t+i}. \tag{C.6}$$

The law of motion for financial wealth is

$$a_{t+1} = \tilde{R}_{t+1}((1 - m_{s_t})a_t - m_{s_t}h_t + (1 - \phi)\omega_t). \tag{C.7}$$

C.3 Aggregation

We aggregate the individual behavior and characterize the aggregate quantities. Since the minimum consumption is proportional to labor income, the disposable income of an agent is

$$\tilde{y}_t = (1 - \phi)\omega_t. \tag{C.8}$$

We denote aggregate quantities by capital letters. By (C.6), the human wealth satisfies

$$h_t = \tilde{y}_t + \frac{1}{\tilde{R}_{t+1}} h_{t+1} = \tilde{y}_t + \frac{1-p}{R_{t+1}} h_{t+1}. \tag{C.9}$$

By (C.7), the law of motion at time t for financial wealth and type is

$$(a, s) \mapsto \begin{cases} (\tilde{R}_{t+1}((1-m_s)a - m_s h_t + \tilde{y}_t), s') & \text{with probability } (1-p)p_{ss'}), \\ (0, s') & \text{with probability } p\pi_{s'}). \end{cases}$$

Letting $A_{t,s}$ be the aggregate financial wealth held by type s agents at time t and aggregating the law of motion for financial wealth, we obtain

$$A_{t+1,s'} = \sum_{s=1}^{S} (1-p)p_{ss'} \tilde{R}_{t+1}((1-m_s)A_{t,s} - m_s \pi_s h_t + \pi_s \tilde{y}_t).$$

Letting $\mathbf{A}_t = (A_{t,1}, \dots, A_{t,S})'$ and $M = \operatorname{diag}(\dots, m_s, \dots)$, we can express the above equation as

$$\begin{aligned} \mathbf{A}_{t+1} &= (1-p)\tilde{R}_{t+1} P'((I-M)\mathbf{A}_t - h_t M\pi + \tilde{y}_t \pi) \\ &= R_{t+1} P'((I-M)\mathbf{A}_t - h_t M\pi + \tilde{y}_t \pi). \end{aligned} \tag{C.10}$$

The vector of aggregate capital holdings of type s agents at the end of time t is then

$$\mathbf{K}_{t+1} = \mathbf{A}_{t+1}/R_{t+1} = P'((I-M)\mathbf{A}_t - h_t M\pi + \tilde{y}_t \pi). \tag{C.11}$$

Multiplying the vector 1 from the left as inner product and noting that $P1 = 1$ and $1'\pi = 1$, aggregate capital is

$$K_{t+1} = 1'((I-M)\mathbf{A}_t - h_t M\pi) + \tilde{y}_t.$$

C.4 Stationary equilibrium

The following theorem establishes the existence of a stationary equilibrium.

Theorem C.2 (Existence)**.** *Let everything be as in Proposition C.1 and* *define the diagonal matrix* $M = \operatorname{diag}(\dots, m_s, \dots)$. *If*

$$P'(I-M)\pi \geq (1-p)\rho(P'(I-M))\pi, \tag{C.12}$$

then a stationary equilibrium exists.

Proof. If a stationary equilibrium with gross risk-free rate R exists, by (C.1) it must be $R > 1-\delta$. Setting $h_t = h$ in (C.9), we obtain

$$h = \frac{\tilde{y}}{1-\frac{1-p}{R}},$$

where $R > 1-p$ is necessary for convergence. By (C.10) we obtain

$$\mathbf{A} = RP'((I-M)\mathbf{A} - hM\pi + \tilde{y}\pi).$$

For convergence we need $R\rho(P'(I-M)) < 1 \iff R < 1/\rho(P'(I-M))$ is necessary, in which case

$$\begin{aligned}\mathbf{A} &= (I-RP'(I-M))^{-1}RP'(I-M)(\tilde{y}I-hM)\pi \\ &= \frac{(1-\phi)\omega}{1-\frac{1-p}{R}}(I-RP'(I-M))^{-1}R\left(P'(I-M)\pi - \frac{1-p}{R}\pi\right).\end{aligned}$$

The vector of aggregate capital supply $\mathbf{K} = (K_1, \dots, K_S)'$ is therefore

$$\mathbf{K} = \frac{(1-\phi)\omega}{R-1+p}(I-RP'(I-M))^{-1}\left(RP'(I-M)\pi - (1-p)\pi\right). \tag{C.13}$$

Define the lower and upper bounds on R by

$$\begin{aligned}\underline{R} &= \max\{1-p, 1-\delta\} \leq 1, \\ \bar{R} &= 1/\rho(P'(I-M)) > 1,\end{aligned}$$

where we have used $m_s \in (0,1)$ and $\rho(P) = 1$. Define the excess demand function of aggregate capital by

$$f(R) = \left(\frac{R-1+\delta}{A\alpha}\right)^{\frac{1}{\alpha-1}} - 1'\mathbf{K},$$

where $\mathbf{K}$ is as in (C.13) and ω is as in (C.2). Clearly f is continuous on $(\underline{R}, \bar{R})$. If we can show $f(\underline{R}+) > 0$ and $f(\bar{R}-) < 0$, then by the intermediate theorem there exists $R \in (\underline{R}, \bar{R})$ and we obtain a stationary equilibrium.

First, consider the behavior of f as $R \downarrow \underline{R}$. If $1 - \delta > 1 - p$, then $\underline{R} = 1 - \delta$ and

$$\left(\frac{R - 1 + \delta}{A\alpha}\right)^{\frac{1}{\alpha-1}} \to \infty,$$
$$1'\mathbf{K} \to \text{finite value}$$

as $R \downarrow \underline{R}$, so $f(R) \to \infty$. If $1 - \delta \leq 1 - p$, then $\underline{R} = 1 - p$. Noting that $m_s \in (0, 1)$ and π is the left Perron vector of P, it follows that

$$\begin{aligned}\underline{R}P'(I - M)\pi - (1 - p)\pi &= (1 - p)(P'(I - M)\pi - \pi) \\ &= -(1 - p)P'M\pi \ll 0.\end{aligned}$$

Therefore by (C.13)

$$\left(\frac{R - 1 + \delta}{A\alpha}\right)^{\frac{1}{\alpha-1}} \geq 0,$$
$$1'\mathbf{K} \to -\infty$$

as $R \downarrow \underline{R}$, so $f(R) \to \infty$.

Next, consider the behavior of f as $R \uparrow \bar{R}$. Let us show that the inequality in (C.12) is strict for at least one entry. Suppose not. Then

$$P'(I - M)\pi = (1 - p)\rho(P'(I - M))\pi.$$

Multiplying the left Perron vector v of the irreducible nonnegative matrix $P'(I - M)$ and dividing by $(v'\pi)\rho(P'(I - M)) > 0$, we obtain the contradiction $1 = 1 - p$. Now letting $R \uparrow \bar{R}$, the entries of $(I - RP'(I - M))^{-1}$ diverge to ∞ and

$$RP'(I - M)\pi - (1 - p)\pi \to \frac{1}{\rho(P'(I - M))}P'(I - M)\pi - (1 - p)\pi > 0$$

as we have just shown, so $1'\mathbf{K} \to \infty$. Therefore $f(R) \to -\infty$. □

The following theorem shows that the stationary wealth distribution has a Pareto upper tail.

Theorem C.3 (Pareto exponent)**.** *Suppose* $p_{ss} > 0$ *for all* s*. Then the stationary equilibrium wealth distribution has a Pareto upper tail with exponent* $\zeta > 1$ *that solves*

$$(1-p)\tilde{R}^z \rho(P(I-M)^{(z)}) = 1. \tag{C.14}$$

Proof. Let $R \in (\underline{R}, \bar{R})$ be the equilibrium gross risk-free rate. By (C.7), the asymptotic gross growth rate of financial wealth is $\tilde{R}(1-m_s)$ in state s. Let us show that $\tilde{R}(1-m_s) > 1$ for some s. Suppose not. Then $\tilde{R}(I-M) \leq I$, so

$$\begin{aligned} RP'(I-M)\pi - (1-p)\pi &= (1-p)\tilde{R}P'(I-M)\pi - (1-p)\pi \\ &\leq (1-p)P'I\pi - (1-p)\pi = 0, \end{aligned}$$

where we have used the fact that $P'\pi = \pi$. Then $\mathbf{K} \leq 0$ by (C.13) and hence $f(R) > 0$, which contradicts the equilibrium condition $f(R) = 0$.

Now let

$$\lambda(z) = (1-p)\tilde{R}^z \rho(P(I-M)^{(z)}).$$

Then

$$\lambda(1) = (1-p)\tilde{R}\rho(P(I-M)) = R\rho(P'(I-M)) = R/\bar{R} < 1.$$

Since $p_{ss} > 0$ for all s and $\tilde{R}(1-m_s) > 1$ for some s, by Proposition 3.1 $\lambda(z) = 1$ has a unique solution $\zeta \in (1, \infty)$, which is the Pareto exponent of the wealth distribution. □

D Simulation

One may argue that the numerical issues discussed throughout the paper are specific to the particular algorithm that involves truncation, and other solution methods such as simulation (Aiyagari, 1994; Krusell and Smith, 1998) may not be subject to those issues. As we see below, however, the situation is equally problematic. Simulation-based methods essentially use the law of large numbers to evaluate the market clearing condition. Suppose we simulate I agents and compute the sample mean of wealth $\frac{1}{I}\sum_{i=1}^{I} w_i$. The question is how fast the sample mean converges to the population mean. If the Pareto exponent ζ exceeds 2, then wealth has finite variance and we can apply the Central Limit Theorem. In this case the sample mean converges at rate $I^{1/2}$. If $\zeta < 2$ on the other hand, it is well known that the rate of convergence to the stable law is only $I^{1-1/\zeta}$. Therefore solving a model accurately may require an impractically large number of agents.

As an illustration, Table 4 shows the order of error $I^{\max\{-1/2,1/\zeta-1\}}$ in the sample mean for various sample size I and Pareto exponent ζ. If $\zeta \geq 2$ and we use 10,000 agents (the number used in Aiyagari (1994)), then the order of the error in the sample mean is $10000^{-1/2} = 1/100 = 1\%$. However, the error order is much larger if the Pareto exponent is smaller. With $\zeta = 1.5$ (a typical number for the wealth distribution according to Vermeulen (2018)), the error order with 10,000 agents is 4.6%, which is substantial. If the Pareto exponent is 1.1 (a typical number for the firm size distribution, which obeys Zipf's law (Axtell, 2001)), then even with ten billion agents ($I = 10^{10}$), which is about the same order of magnitude as the world population, the error order is still 12.3%. To drive the error down to 1%, quite a modest number, the required sample size for $\zeta = 1.1$ is $I = 100^{\frac{\zeta}{\zeta-1}} = 10^{22}$ (ten sextillion), which is about the same order of magnitude as the number of stars in the universe or sand grains on earth.[19] Therefore we cannot expect to solve such models accurately using simulation.

Table 4: Order of error $I^{\max\{-1/2,1/\zeta-1\}}$ in sample mean.

Sample size I	Pareto exponent ζ			
	≥ 2	1.5	1.3	1.1
$10^0 = 1$	1.00000	1.00000	1.00000	1.00000
10^2	0.10000	0.21544	0.34551	0.65793
10^4	0.01000	0.04642	0.11938	0.43288
10^6	0.00100	0.01000	0.04125	0.28480
10^8	0.00010	0.00215	0.01425	0.18738
10^{10}	0.00001	0.00046	0.00492	0.12328

[19] http://www.abc.net.au/science/articles/2015/08/19/4293562.htm

E Matlab files

The MATLAB files for implementing the Pareto extrapolation algorithm can be downloaded from `https://github.com/alexisakira/Pareto-extrapolation`. There are three main functionalities for Pareto extrapolation:

- `getZeta.m`
- `getQ.m`
- `getTopShares.m`

In addition, `example.m` contains a simple example.

E.1 Computing Pareto exponent

`getZeta.m` computes the Pareto exponent using the Beare and Toda (2017) formula. The usage is

```
zeta = getZeta(PS,PJ,p,G,zetaBound)
```

where

- `PS` is the $S \times S$ transition probability matrix of exogenous states indexed by $s = 1, \dots, S$,
- `PJ` is the $S^2 \times J$ matrix of conditional probabilities of transitory states indexed by $j = 1, \dots, J$,
- `p` is the birth/death probability $p \in [0, 1)$,
- `G` is the $S^2 \times J$ matrix of gross growth rates, and
- `zetaBound` is a vector $(\underline{\zeta}, \bar{\zeta})$ that specifies the lower and upper bounds to search for the Pareto exponent (optional).

The S^2 rows in `PJ` and `G` should be ordered such that

$$(s, s') = (1, 1), \dots, (1, S); \dots; (s, 1), \dots, (s, S); \dots; (S, 1), \dots, (S, S).$$

If `PS` $= P = (p_{ss'})$, `PJ` $= (\pi_{ss'j})$, and `G` $= (G_{ss'j})$, then the Pareto exponent $z = \zeta$ is the solution to

$$(1-p)\rho(P \odot M(z)) = 1,$$

where ρ is the spectral radius and $M(z) = (M_{ss'}(z))$,

$$M_{ss'}(z) = \sum_{j=1}^{J} \pi_{ss'j} G_{ss'j}^{z},$$

and $\odot$ is the Hadamard (entry-wise) product.

`PJ` can be either $1 \times J$, $S \times J$, or $S^2 \times J$. If it is $1 \times J$, it assumes $\pi_{ss'j} = \pi_j$ depends only on j. If it is $S \times J$, it assumes $\pi_{ss'j} = \pi_{sj}$ depends only on (s, j).

`G` can be either $S \times J$ or $S^2 \times J$. If it is $S \times J$, it assumes $G_{ss'j} = G_{sj}$ depends only on (s, j).

E.2 Computing joint transition probability matrix

`getQ.m` computes the $SN \times SN$ joint transition probability matrix $Q = (q_{sn,s'n'})$ and the stationary distribution $\pi = (\pi_{sn})$ for the exogenous state s and wealth. The usage is

```
[Q,pi] = getQ(PS,PJ,p,x0,xGrid,gstjn,Gstj,zeta)
```

where

- `PS`, `PJ`, `p` are as above,
- `x0` is the initial wealth of newborn agents,
- `xGrid` is the $1 \times N$ grid of wealth (size variable) w_n,
- `gstjn` is the $S^2 \times JN$ matrix of law of motion for wealth $g_{ss'j}(w_n)$,
- `Gstj` is the $S^2 \times J$ matrix of asymptotic slopes of law of motion $G_{ss'j}$ (optional), and
- `zeta` is the Pareto exponent (optional).

The JN columns of `gstjn` must be ordered such that the first N columns correspond to $j = 1$, the next N columns correspond to $j = 2$, and so on. `Gstj` is the same as `G` in `getZeta.m`. If unspecified, it uses the slope of the law of motion between the two largest grid points. If `zeta` is unspecified, it calls `getZeta.m` to compute.

`gstjn` can be either $S \times JN$ or $S^2 \times JN$. If it is $S \times JN$, it assumes $g_{ss'j}(w_n) = g_{sj}(w_n)$ depends only on (s, j, n).

E.3 Computing top wealth shares

`getTopShares.m` computes the top wealth shares. The usage is

```
topShare = getTopShares(topProb,wGrid,wDist,zeta)
```

where

- `topProb` is the vector of top probabilities to evaluate top shares,
- `wGrid` is the $1 \times N$ vector of wealth grid,
- `wDist` is the $1 \times N$ vector of wealth distribution, and
- `zeta` is the Pareto exponent (optional).

Given the stationary distribution π computed using `getQ.m`, one can compute the wealth distribution as $\pi_n = \sum_{s=1}^{S} \pi_{sn}$. If `zeta` is unspecified, `getTopShares.m` uses spline interpolation to compute top wealth shares.