Electronic Companion

EC.1. Proofs and Additional Numerical Results

EC.1.1. Proofs of Theorem 1 and Propositions 1, 2, 3, 4

We will use the following Proposition to prove Theorem 1 (see Corollary 2.5.2 in Topkis (2011)).

PROPOSITION EC.1. *Suppose that Z is a non-empty complete lattice, E is a partially ordered set, and f is an increasing function from $Z \times E$ into Z. Then, for each $e \in E$, the greatest and least fixed points of f exist and are increasing in e on E.*

Proof of Theorem 1 Let $\theta_1, \theta_2 \in \mathcal{P}(S)$ and assume that $\theta_1 \succeq_D \theta_2$. Let μ_1, μ_2 be two invariant distributions of Q. Assume without loss of generality that $h_2 := H(\mu_2) \geq H(\mu_1) := h_1$ and let $f : S \to \mathbb{R}$ be a function such that $f \in D$. We have

$$
\begin{aligned}
&\int_S f(x) M_{h_2}\theta_2(dx) \\
&= \int_S \int_S f(y) Q(x, h_2, dy)\theta_2(dx) \\
&\leq \int_S \int_S f(y) Q(x, h_1, dy)\theta_2(dx) \\
&\leq \int_S \int_S f(y) Q(x, h_1, dy)\theta_1(dx) \\
&= \int_S f(x) M_{h_1}\theta_1(dx).
\end{aligned}
$$

Thus, $M_{h_1}\theta_1 \succeq_D M_{h_2}\theta_2$. The first inequality follows from the fact that Q is D-decreasing. The second inequality follows from the facts that $\theta_1 \succeq_D \theta_2$ and Q is D-preserving. We conclude that $M^n_{h_1}\theta_1 \succeq_D M^n_{h_2}\theta_2$ for all $n \in \mathbb{N}$.

Assume that condition (i) of the theorem holds. The fact that Q satisfies Property (C) implies that $M^n_{h_i}\theta_i$ converges weakly to the unique fixed point of M_{h_i} which is given by μ_{h_i} for $i = 1, 2$. Because μ_1 and μ_2 are invariant distributions of Q we have $\mu_{h_i} = \mu_i$ for $i = 1, 2$. Because $\succeq_D$ is closed with respect to weak convergence, we have $\mu_1 \succeq_D \mu_2$. Using the fact that H is increasing with respect to $\succeq_D$ implies $h_1 \geq h_2$.

We conclude that if μ_1 and μ_2 are invariant distributions of Q then $H(\mu_1) = H(\mu_2)$. Thus, $Q(x, H(\mu_1), B) = Q(x, H(\mu_2), B)$ for all $x \in S$ and $B \in \mathcal{B}(S)$. Because Q satisfies assumption (U) the operators $M_{H(\mu_1)}$ and $M_{H(\mu_2)}$ have unique fixed points. Thus, $\mu_{H(\mu_1)} = \mu_{H(\mu_2)}$, i.e., $\mu_1 = \mu_2$. We conclude that if an invariant distribution of Q exists, it is unique.

Now assume that condition (ii) of the theorem holds. Define the order $\geq'$ as the reverse of the usual order $\geq$: for x, y, we write $x \geq' y$ if and only if $y \geq x$. Under this assumption, the arguments above imply that the operator M is increasing from $\mathcal{P}(S)\times\mathcal{H}$ to $\mathcal{P}(S)$ on the complete lattice $(\mathcal{P}(S),\succeq_D)$ when $\mathcal{H}$ is endowed with $\geq'$. Then by applying Proposition EC.1 to the increasing operator M we have $\mu_{h_1}\succeq_D \mu_{h_2}$, i.e., $\mu_1 \succeq_D \mu_2$. Now we can use the same arguments as the arguments for the case that condition (i) holds to show that if an invariant distribution of Q exists, it is unique. □

In order to establish the existence of an invariant distribution we will use Schauder-Tychonoff's following fixed-point theorem (see Corollary 17.56 in Aliprantis and Border (2006)).

Proposition EC.2. *(Schauder-Tychonoff) Let K be a non-empty, compact, convex subset of a locally convex Hausdorff space, and let $f: K \to K$ be a continuous function. Then the set of fixed points of f is compact and non-empty.*

Proof of Proposition 1 Because S is a compact Polish space $\mathcal{P}(S)$ is a compact Polish space under the weak topology (see Theorem 15.11 in Aliprantis and Border (2006)). Clearly $\mathcal{P}(S)$ is convex. $\mathcal{P}(S)$ endowed with the weak topology is a locally convex Hausdorff space. Thus, if T is continuous, we can apply Schauder-Tychonoff's fixed point theorem to conclude that T has a fixed point.

To show that T is continuous, take a sequence of measures $\{\mu_n\}$ and assume that it converges weakly to μ.

Let $f: S \to \mathbb{R}$ be a continuous and bounded function. Because Q and H are continuous we have $\lim_{n\to\infty}\int_S f(y)Q(x_n,H(\mu_n),dy) = \int_S f(y)Q(x,H(\mu),dy)$ whenever $x_n \to x$. Define $m_n(x) := \int_S f(y)Q(x,H(\mu_n),dy)$. Then $m_n(x)$ is a uniformly bounded sequence of functions such that $m_n(x_n)\to m(x)$ whenever $x_n \to x$. Thus, by Lebesgue's Convergence Theorem for varying measures (see Theorem 3.5 in Serfozo (1982) and Section 5 in Feinberg et al. (2020)) we have $\lim_{n\to\infty}\int m_n(x)\mu_n(dx) = \int m(x)\mu(dx)$. Hence,

$$\begin{aligned}
&\lim_{n\to\infty}\int_S f(x)T\mu_n(dx)\\
&=\lim_{n\to\infty}\int_S\int_S f(y)Q(x,H(\mu_n),dy)\mu_n(dx)\\
&=\int_S\int_S f(y)Q(x,H(\mu),dy)\mu(dx)\\
&=\int_S f(x)T\mu(dx).
\end{aligned}$$

Thus, $T\mu_n$ converges weakly to $T\mu$. We conclude that T is continuous. Thus, by the Schauder-Tychonoff's fixed point theorem, T has a fixed point. □

Proof of Proposition 2 Consider the function $f(h) = h - H(\mu_h)$ from $[h', h'']$ to $\mathbb{R}$ which is well defined because $\mu_h \in \mathcal{P}(S)$ for all $h \in [h', h'']$.

We first claim that a root of f, say h^*, corresponds to an invariant distribution μ_{h^*} of Q. To see this, let h^* be a root of f, that is, $H(\mu_{h^*}) = h^*$.

From Property (U), μ_{h^*} is the unique probability measure that satisfies

$$\mu_{h^*}(B) = \int Q(x, h^*, B)\mu_{h^*}(dx),$$

so $H(\mu_{h^*}) = h^*$ implies that

$$\mu_{h^*}(B) = \int Q(x, H(\mu_{h^*}), B)\mu_{h^*}(dx),$$

i.e., μ_{h^*} is an invariant distribution of Q.

If $h'' = H(\mu_{h''})$ or $h' = H(\mu_{h'})$ then f has a root, and hence, Q has an invariant distribution. If $h'' > H(\mu_{h''})$ and $h' < H(\mu_{h'})$, we have $f(h'') > 0 > f(h')$ so if f is continuous we can apply the intermediate value theorem to prove that f has a root, that is, Q has an invariant distribution.

We will now show that f is continuous to conclude the proof.

Consider a sequence $\{h_n\}$, $h_n \in [h', h'']$ such that h_n converges to h and let $\{\mu_{h_k}\}$ be a subsequence of $\{\mu_{h_n}\}$ that converges to λ. From Lebesgue's Convergence Theorem for varying measures (see Theorem 3.5 in Serfozo (1982)) and using the same logic as in the proof of Proposition 1, for every continuous and bounded function $m : S \to \mathbb{R}$, we have

$$\begin{aligned}
&\lim_{k\to\infty} \int_S m(x)\mu_{h_k}(dx) \\
&= \lim_{k\to\infty} \int_S \int_S m(y)Q(x, h_k, dy)\mu_{h_k}(dx) \\
&= \int_S \int_S m(y)Q(x, h, dy)\lambda(dx) \\
&= \int_S m(x)M_h\lambda(dx).
\end{aligned}$$

Because $\{\mu_{h_k}\}$ converges to λ we also have

$$\lim_{k\to\infty} \int_S m(x)\mu_{h_k}(dx) = \int_S m(x)\lambda(dx).$$

Thus, $\lambda = M_h\lambda$. From assumption (U), μ_h is the unique fixed point of M_h, and thus, $\lambda = \mu_h$.

We conclude that any subsequence of $\{\mu_{h_n}\}$ that converges weakly at all converges weakly to μ_h. Furthermore, from assumption, the sequence $\{\mu_{h_n}\}$ is a tight sequence of probability measures. Thus, $\{\mu_{h_n}\}$ converges weakly to μ_h (see the Corollary after Theorem 25.10 in Billingsley (1995)).

Because H is continuous, we conclude that $f(h) = h - H(\mu_h)$ is continuous on $[h', h'']$ which completes the proof. □

Proof of Proposition 3 From Proposition 2 the function f is continuous and has opposite signs at h_1 and h_2. Hence, the sequence h_n defined in the statement of the proposition converges linearly to the root of f (see for example, Theorem 2.1 in Burden and Faires (1985)).

From Proposition 2 if h^* is a root of f, then μ_{h^*} is an invariant distribution of Q which completes the proof. □

Proof of Proposition 4 The proof is similar to the proof of Theorem 1. We provide it here for completeness. Let $\theta_1, \theta_2 \in \mathcal{W}$ such that $\theta_1 \succeq_D \theta_2$. Let $\mu_1, \mu_2 \in \mathcal{W}$ be two invariant distributions of Q.

Assume without loss of generality that $h_2 := H(\mu_2) \geq H(\mu_1) := h_1$ so $h_1, h_2 \in \mathcal{H}_{\mathcal{W}}$ and let $f : S \to \mathbb{R}$ be a function such that $f \in D$. We have

$$\begin{aligned}
&\int_S f(x) M_{h_2}\theta_2(dx) \\
&= \int_S \int_S f(y) Q(x, h_2, dy)\theta_2(dx) \\
&\leq \int_S \int_S f(y) Q(x, h_1, dy)\theta_2(dx) \\
&\leq \int_S \int_S f(y) Q(x, h_1, dy)\theta_1(dx) \\
&= \int_S f(x) M_{h_1}\theta_1(dx).
\end{aligned}$$

Thus, $M_{h_1}\theta_1 \succeq_D M_{h_2}\theta_2$. The first inequality follows from the fact that Q is D-decreasing on $\mathcal{W}$. The second inequality follows from the facts that $\theta_1 \succeq_D \theta_2$ and Q is D-preserving on $\mathcal{W}$. Now because $\theta_1, \theta_2 \in \mathcal{W}$ and $h_1, h_2 \in \mathcal{H}_{\mathcal{W}}$, we have $M_{h_1}\theta_1, M_{h_2}\theta_2 \in \mathcal{W}$. Applying the same argument as above again, we conclude that $M^n_{h_1}\theta_1 \succeq_D M^n_{h_2}\theta_2$ for all $n \in \mathbb{N}$.

Now the proof continues exactly as in the proof of Theorem 1. □

EC.1.2. Proof of Corollaries 1, 3, 4

Proof of Corollary 1 Let $H(\mu) = \int_{\mathbb{R}_+} x\mu(dx)$, $Law(S_t) = Law(S)$ and $Law(T_t(h)) = Law(T(h))$. Let D be the set of increasing functions, so $\succeq_D$ is equivalent to the first order stochastic dominance order $\succeq_{SD}$ and H is increasing with respect to $\succeq_D$. From Theorem 19.3.5 in Meyn and Tweedie (2012), Property (C) is satisfied because $\mathbb{E}(T(h)) \geq \mathbb{E}(T(0)) > \mathbb{E}(S)$ for all $h \geq 0$.

Let $f : \mathbb{R} \to \mathbb{R}$ be increasing. Because T is stochastically increasing in h the function

$$\int f(y) Q(x, h, dy) = \mathbb{E} f(\max\{x + S - T(h), 0\})$$

is increasing in x and decreasing in h where the expectation is taken with respect to the random variables S and $T(h)$. Thus, Q is D-preserving and D-decreasing. Hence, from Theorem 1 we conclude that the nonlinear Markov chain given in Equation (7) has at most one invariant distribution.

For existence, first note that the function $H(\mu_h)$ is bounded from below by 0 so $h' \leq H(\mu_{h'})$ for $h' = 0$. In addition, from the proof of Theorem 1 the function $H(\mu_h)$ is decreasing in h so $H(\mu_h) \leq H(\mu_0) < \infty$ as $\mathbb{E}(T(0)) > \mathbb{E}(S)$. Hence, we can find $h'' \geq H(\mu_{h''})$ for some $h'' > 0$.

We already established that property (C) holds, and hence, property (U) holds too. Further, it is immediate to verify that H and Q are continuous.

Finally, for any sequence of non-negative numbers h_n that converges to some h, the assumptions that $\mathbb{E}(T(0)) > \mathbb{E}(S)$ and that $T(h)$ and S have bounded variances, guarantee that the sequence of invariant distributions of the G/G/1 queue μ_{h_n} has bounded first two moments, and hence, it is tight. Thus, we can apply Proposition 2 to conclude that an invariant distribution exists which completes the proof. □

Proof of Corollary 3 Existence follows immediately from Proposition 1. For uniqueness, we need to show that the conditions of Theorem 1 holds. We let $S = \{1, \ldots, n\}$ with the standard order, $H(\mu) = G(\mu(\{1\}), \ldots, \mu(\{n\}))$, and D to be the set of increasing functions, so $\succeq_D$ is equivalent to $\succeq_{SD}$ and $(\mathcal{P}(S), \succeq_D)$ is a complete lattice. Note that H is increasing with respect to $\succeq_{SD}$ because $\mu \succeq_{SD} \mu'$ holds if and only if $(\mu(\{1\}), \ldots, \mu(\{n\})) \succeq_m (\mu'(\{1\}), \ldots, \mu'(\{n\}))$ and from the assumption that G is increasing with respect to $\succeq_m$.

Condition (1) implies that Q is D-preserving, Condition (2) implies that Q is D-decreasing, and Condition (3) implies that Property (U) holds. Thus, we can apply Theorem 1 to prove that Q has at most one invariant distribution.

We can identify Q with the stochastic matrix P by $P_{ij}(\cdot) = Q(i, \cdot, \{j\})$, and hence, using the definition of the invariant distribution, the Corollary follows from Theorem 1. □

Proof of Corollary 4 For existence, continuity of H and Q follows immediately from the assumptions. Now note that the state space is bounded because the random variables R_i, Y, and the policy functions g_i are bounded. In particular, we let the state space be the compact set $S = [0, nM\overline{r} + \overline{y}]$. Hence, we can use Proposition 1 to conclude that Q has an invariant distribution.

For uniqueness, we need to show that the conditions of Theorem 1 hold. We let D to be the set of increasing functions, so $\succeq_D$ is equivalent to $\succeq_{SD}$.

It is immediate that Condition (2) implies that Q is D-preserving and Conditions (2) and (3) imply that Q is D-decreasing. Thus, we can apply Theorem 1 to prove that Q has at most one invariant distribution. □

EC.1.3. Proof of Claims 1,2,3,4,5

Proof of Claim 1 We let D to be the set of all increasing functions. Clearly H is increasing with respect to $\succeq_D$ because m is increasing. Property (C) holds for AR(1) process with $a \in (0, 1)$, (see, for example, Light (2024)). Let $f : \mathbb{R} \to \mathbb{R}$ be increasing. Then

$$\int f(y) Q(x, h, dy) = \int f(ax - h + \epsilon) \phi(d(\epsilon))$$

is increasing in x and decreasing in h so Q is D-preserving and D-decreasing. Hence, we can apply Theorem 1 to conclude that Q has at most one invariant distribution.

For existence, note that the quadratic growth condition and the fact that the variance of ϵ_t is finite imply that $H(\mu_h)$ is finite for every h. Furthermore, if h_n converges to h, then it follows that the sequence $\mu_{h_n}(dx)$ of invariant distributions of the AR(1) process given the parameter h_n has bounded first two moments, and hence, $\{\mu_{h_n}\}$ is a tight sequence of probability measures. In addition it is immediate that Q is continuous and H is continuous on $\{\mu_h : h \in [h', h'']\}$ as μ_h has bounded first two moments and m is continuous with a quadratic bound.

From the proof of Theorem 1 we have that $H(\mu_h)$ is decreasing in h. This implies that we can find $h', h'' \in \mathbb{R}$, $h'' > h'$, such that $h'' \geq H(\mu_{h''})$ and $h' \leq H(\mu_{h'})$ (e.g., by letting $h' = -|c| - 1$ and $h'' = |c| + 1$ where $H(\mu_0) = c$).

Thus, from Proposition 2 existence follows. □

Proof of Claim 2 We let D to be the set of all the functions that are increasing in the first argument. Clearly H is increasing with respect to $\succeq_D$. We need to show that Q is D-preserving and D-decreasing in order to use Theorem 1. Let $f \in \mathbb{R}^{\mathbb{R}^2}$ be increasing in the first argument. Then

$$\int f(y_1, y_2) Q((x_1, x_2), h, dy) = \int f(ax_1 - h + \epsilon_1, k(x_2) + \epsilon_2)\phi(d(\epsilon_1, \epsilon_2))$$

is increasing in the first argument and decreasing in h so Q is D-preserving and D-decreasing. Hence, we can apply Theorem 1 to conclude that Q has at most one invariant distribution. Existence of an invariant distribution follows by the same argument as in Claim 1. □

Proof of Claim 3 Consider the set of functions D such that $f : \mathbb{R}^n \to \mathbb{R}$ is in D if $f(x) = \sum_{i=1}^n y_i x_i + c$ for some $y \in O$ and $c \in \mathbb{R}$. Property (C) holds (see Example 1 in Light (2024)). It is immediate that H is increasing with respect to $\succeq_D$.

We now show that Q is D-preserving and D-decreasing. Let $f \in D$ so $f(x) = \sum_{i=1}^n y_i x_i + b$ for some $y \in O$. We have

$$\begin{aligned} v(x) &:= \int f(x') Q(x, h, dx') \\ &= \int f\big(a_1 x_1 - \beta_1 h + \epsilon_1, \ldots, \\ &\qquad a_n x_n - \beta_n h + \epsilon_n\big)\phi(d\epsilon) \\ &= \int \sum_{i=1}^n y_i (a_i x_i - \beta_i h + \epsilon_i)\phi(d\epsilon) + b \\ &= \sum_{i=1}^n y_i' x_i + b' \end{aligned}$$

with $y_i' = a_i y_i$ and $b' = \int \sum_{i=1}^n y_i(-\beta_i h + \epsilon_i)\phi(d\epsilon) + b$. Note that y' is in O as $y \in O$ and $a_i \geq 0$ for all i. Hence, v is in D which means that Q is D-preserving.

To show that Q is D-decreasing let $h_2 \geq h_1$ and note that

$$\begin{aligned}
&\int f(x')Q(x, h_2, dx') \\
&= \int \sum_{i=1}^{n} y_i(a_i x_i - \beta_i h_2 + \epsilon_i)\phi(d\epsilon) + b \\
&\leq \int \sum_{i=1}^{n} y_i(a_i x_i - \beta_i h_1 + \epsilon_i)\phi(d\epsilon) + b \\
&= \int f(x')Q(x, h_1, dx')
\end{aligned}$$

where the inequality follows from the fact that y and β are in O so $y_i\beta_i \geq 0$ for all i. Thus, Q is D-decreasing.

To prove existence, note that we can find $H(\mu_h)$ directly. A simple calculation shows that $H(\mu_h) = \sum_{i=1}^n \gamma_i(-h + e_i)/(1 - a_i)$ where e_i is the expected value of ϵ_i. Thus, we can find $h', h'' \in \mathbb{R}$, $h'' > h'$, such that $h'' > H(\mu_{h''})$ and $h' < H(\mu_{h'})$. In addition, it is easy to see that the tightness condition of Proposition 2 holds as the sequence $\{\mu_{h_k}\}$ has bounded first two moments whenever h_k converges to some h. □

Proof of Claim 4 Fix the finite lattice $S = \{0, 1\}^n$ and let D be the set of all super-modular functions on S.

Note that the nonlinear Markov kernel Q is defined for each $y \in \{0, 1\}^n$ by

$$\begin{aligned}
Q\big(x, h; \{y\}\big) = \rho(x, h)\, \mathbf{1}_{\{y=(z,\ldots,z)\}}\, p^z(1-p)^{1-z} \\
+ (1 - \rho(x, h)) \prod_{i=1}^{n} p^{y_i}(1-p)^{1-y_i}.
\end{aligned}$$

Let $f \in D$ and let $C_0 = \mathbb{E}\big[f(Z_1, \ldots, Z_n)\big]$ and $C_1 = \mathbb{E}\big[f(Z, \ldots, Z)\big]$. It can be shown that $(Z, \ldots, Z)$ dominates $(Z_1, \ldots, Z_n)$ in super-modular order (e.g., Hu et al. (2005)). Hence, $C_1 \geq C_0$.

For each $x \in S$ and $h \in [0, 1]$, we have

$$\sum_{y \in S} f(y)Q(x, h, \{y\}) = C_0 + \big(C_1 - C_0\big)\rho(x, h).$$

which is supermodular in x and decreasing in h because the coefficient $C_1 - C_0$ is non-negative and $\rho(\cdot, h)$ is super-modular in x and decreasing in h by assumption. Hence Q is D-preserving and D-decreasing.

To show that H is increasing let $\mu_2 \geq_D \mu_1$ and define $f^*(y) = \mathbf{1}_{\{y=(1,\dots,1)\}}$, which is super-modular on $\{0,1\}^n$. Then

$$\begin{aligned} H(\mu_2) &= \mu_2\big(\{(1,\dots,1)\}\big) = \int f^*(y)\,\mu_2(dy) \\ &\geq \int f^*\,d\mu_1(dy) = \mu_1\big(\{(1,\dots,1)\}\big) \\ &= H(\mu_1), \end{aligned}$$

and H is increasing with respect to $\geq_D$.

In addition, for each h the linear Markov chain $Q(\cdot,h,\cdot)$ assigns strictly positive probability to every state of S because $0<p<1$ and $1-\rho(x,h)>0$ Hence the chain is irreducible and aperiodic on a finite state space and Property (C) holds.

Thus, from Theorem 1 the nonlinear Markov chain Q has at most one invariant distribution. Existence readily follows from using the continuity of ρ and applying Proposition 1. This proves Claim 4.

Now assume that $D_\uparrow$ is the set of all increasing functions on $S=\{0,1\}^n$ and assume for simplicity $n=2$. Let $k(y) = y_1+y_2-y_1y_2$, which is increasing in each coordinate. We have

$$\begin{aligned} &\sum_y k(y)Q(x,h,\{y\}) \\ &= Q(x,h,\{1,0\}) + Q(x,h,\{1,1\}) + Q(x,h,\{0,1\}) \\ &= 2p-p^2-(p-p^2)\rho(x,h), \end{aligned}$$

which is not necessarily increasing. Hence, Q fails to be $D_\uparrow$-preserving. $\square$

Proof of Claim 5 Let

$$h = \frac{\mathbb{E}(S^2)}{\sqrt{\mathbb{E}(S)^2+2\mathbb{E}(S^2)}-\mathbb{E}(S)} \tag{EC.1}$$

and consider the linear Markov chain $W_{t+1} = \max(0, W_t + S_t - T_t(h))$. Then it has a unique invariant distribution if $\mathbb{E}T_t(h) = h > \mathbb{E}(S)$ (see Theorem 19.3.5 in Meyn and Tweedie (2012)) which holds because

$$\mathbb{E}(S)\sqrt{\mathbb{E}(S)^2+2\mathbb{E}(S^2)} = \sqrt{\mathbb{E}(S)^4+2\mathbb{E}(S^2)\mathbb{E}(S)^2} < \sqrt{\big(\mathbb{E}(S)^2+\mathbb{E}(S^2)\big)^2} = \mathbb{E}(S)^2+\mathbb{E}(S^2)$$

which implies that $h>\mathbb{E}(S)$. Let W_∞ be the random variable with the law μ^* where μ^* is unique invariant distribution of the linear Markov chain $(W_t)_{t\in\mathbb{N}}$.

From the Pollaczek-Khinchin formula (see Equation (8.1) in Chapter 8 in Cooper (1972)) the stationary expected waiting time is given by $\mathbb{E}(W_\infty) = \lambda(h)\mathbb{E}(S^2)/(2(1-\lambda(h)\mathbb{E}(S)))$. Using the fact that $\lambda(h)=1/h$, and algebraic manipulations, we see that $h = \mathbb{E}(W_\infty)$. Hence, μ^* is an invariant distribution of the nonlinear Markov chain given in Equation (7). Uniqueness follows from Corollary 1.

For $M/M/1$ queue S is an exponential random variable with a parameter μ, so $\mathbb{E}(S) = 1/\mu$ and $\mathbb{E}(S^2) = 2/\mu^2$ and we get

$$\mathbb{E}(W_\infty) = \frac{2}{(\sqrt{5}-1)\mu} = \frac{2\mathbb{E}(S)}{\sqrt{5}-1}$$

which completes the proof. □

EC.1.4. Uniqueness via the Convex Stochastic Order

In this section, we expand Section 3.3 by presenting an example where the convex stochastic order is used to establish the uniqueness of an invariant distribution and the standard first order stochastic dominance would not satisfy the required conditions for uniqueness.

Example EC.1 (Convex Stochastic Order). Suppose that the state space is $S = \mathbb{R}$ and $a \in (-1, 1)$. Let $(\epsilon_t)_{t\geq 1}$ be i.i.d. with mean zero, finite moments and law ϕ. Let $\sigma(h)$ be a positive, continuous and decreasing function and consider the nonlinear Markov chain

$$X_{t+1} = a\,X_t + \sigma\big(H(\mu_t)\big)\,\epsilon_{t+1},$$

where $H(\mu) = \int m(x)\mu(dx)$ for some continuous and convex function m on $\mathbb{R}$ such that $|m(x)| \leq C_0 + C_1 x^k$ for some $k \geq 1$ and constants C_0, C_1. Let D be the set of all convex functions on S so $\succeq_D$ is the convex stochastic order.

Clearly H is increasing with respect to $\succeq_D$. Property (C) follows from standard arguments as in Claim 1.

Let $f : \mathbb{R} \to \mathbb{R}$ be convex. Then the function

$$v(x,h) := \int f(y)Q(x,h,dy) = \mathbb{E}f(ax + \sigma(h)\epsilon)$$

where the expectation is with respect to ϵ, is convex in x as a composition of a convex and a linear function. Hence, Q is D-preserving.

Now let $h_2 \geq h_1$ and $Y = \sigma(h_2)\epsilon$, $c = \sigma(h_1)/\sigma(h_2)$. Note that $c \geq 1$ because σ is decreasing.

In addition, from Jensen's inequality we have

$$\mathbb{E}f(ax+Y) \leq \frac{1}{c}\mathbb{E}f(ax+cY) + \left(1 - \frac{1}{c}\right) f(ax).$$

Hence, using Jensen's inequality again we have

$$\mathbb{E}f(ax+Y) = c\mathbb{E}f(ax+Y) - (c-1)\,\mathbb{E}f(ax+Y) \leq c\mathbb{E}f(ax+Y) - (c-1)\,f(ax) \leq \mathbb{E}f(ax+cY).$$

That is,

$$v(x,h_2) = \mathbb{E}f(ax+\sigma(h_2)\epsilon) \leq \mathbb{E}f(ax+\sigma(h_1)\epsilon) = v(x,h_1)$$

so D-decreasing. Hence, we can apply Theorem 1 to conclude that Q has at most one invariant distribution.

The proof of existence follows from similar arguments to the proof in Claim 1 so it is omitted. We conclude that Q has a unique invariant distribution.

EC.2. On non-contraction and numerical examples

In this section we show that the nonlinear operator T is generally not a contraction in our examples and applications. We focus on three tractable cases: Example 1 (a one–dimensional nonlinear autoregression), the strategic M/M/1 queue (Section 4.1), and the nonlinear fixed-point equations studied in Section 4.3. In each instance, we show that the associated dynamics are generally not contractive even in simple parameter configurations. We note that the other two applications (inventory dynamics and wealth distributions) have transitions that depend on policy functions that typically come from a dynamic programming problem and do not have closed-form expressions so establishing contraction there would be much more involved. In Section EC.2.0.3 we derive such a policy function numerically in the wealth distribution application and compute the corresponding invariant wealth distribution using Algorithm 1.

We first recall the definitions of the Wasserstein and total variation metrics between two probability measures μ and ν on a metric space $(\mathcal{S}, d)$ that is a Polish space.

For $p \geq 1$, the *Wasserstein distance of order p* is defined as

$$W_p(\mu, \nu) := \inf_{\pi \in \Pi(\mu,\nu)} \left(\int_{\mathcal{S}\times\mathcal{S}} d(x,y)^p \, d\pi(x,y) \right)^{1/p},$$

where $\Pi(\mu, \nu)$ denotes the set of all couplings of μ and ν. If μ and ν are probability measures on $\mathbb{R}$ with cumulative distribution functions (CDFs) F_1 and F_2, then the Wasserstein distance of order $p \geq 1$ is given by:

$$W_p(\mu, \nu) = \left(\int_0^1 \left| F_1^{-1}(q) - F_2^{-1}(q) \right|^p \, dq \right)^{1/p}, \tag{EC.2}$$

where F_1^{-1} and F_2^{-1} are the quantile functions (inverse CDFs).

The *total variation distance* between μ and ν is defined as

$$\|\mu - \nu\|_{TV} := \sup_{A \in \mathcal{B}(\mathcal{S})} |\mu(A) - \nu(A)|.$$

If μ and ν admit densities f and g with respect to a common reference measure, then the total variation distance simplifies to

$$\|\mu - \nu\|_{TV} = \frac{1}{2} \int_{\mathcal{S}} |f(x) - g(x)| \, dx. \tag{EC.3}$$

EC.2.0.1. Contraction in Example 1 Consider the nonlinear Markov chain in Example 1 given in Equation 3 with $m(x) = \beta x$ for some $\beta > 0$. As a special case of Claim 1, this nonlinear Markov chain has a unique invariant distribution using the monotonicity arguments developed in this paper. We now show it is generally not a contraction in the Wasserstein distance and the total variation distance. We then show that augmenting the nonlinear Markov chain with an additional variable to construct a linear Markov chain would not generally be a useful approach to prove uniqueness of an invariant distribution.

Wasserstein Distance. Let's assume $p \geq 1$ and the pth moment of ϵ exists.

Consider two simple Dirac measures $\mu_1 = \delta_h$ and $\mu_2 = \delta_{h+\delta}$ for some $h \in \mathbb{R}$ and $\delta > 0$. Then $W_p(\mu_1, \mu_2) = \delta$.

For $\mu_1 = \delta_h$, we have $x = h$ and $H(\mu_1) = \beta h$ so $T\mu_1 = \text{Law}(\epsilon + (a-\beta)h)$. Similarly, for $\mu_2 = \delta_{h+\delta}$, we have $x = h+\delta$ and $H(\mu_2) = \beta(h+\delta)$ so $T\mu_2 = \text{Law}(\epsilon + (a-\beta)h + (a-\beta)\delta)$. Thus, $T\mu_2$ is a simple translation of $T\mu_1$ by a constant shift of $(a-\beta)\delta$. The Wasserstein distance between such laws $W_p(T\mu_1, T\mu_2)$ is $|a-\beta|\delta$.

Hence, we conclude that a necessary condition for T to be a contraction is $|a-\beta| < 1$ which is not generally the case (e.g., if $\beta \geq 2$). On the other hand, a sufficient condition for contraction in W_1 is the relatively strong requirement $|a| + |\beta| < 1$.[1] Intuitively, the nonlinear mapping T amplifies differences in the mean when β is large: a small change in the current distribution's mean leads to an even larger shift in the next period's mean. This is in addition to the standard auto regressive feedback captured by a. In more complex applications we consider next where the entire distribution determines the dynamics, contraction is typically far more difficult to satisfy.

Total Variation Distance. Let $\mu_1 = \mathcal{N}(m_1, \sigma^2)$ and $\mu_2 = \mathcal{N}(m_2, \sigma^2)$ where $\mathcal{N}(\mu_i, \sigma^2)$ is the normal random variable with mean μ_i and variance σ^2 and probability density function $\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m_i)^2}{2\sigma^2}\right)$.

A simple calculation using Equation (EC.3) shows that the total variation distance is given by

$$d_{TV}(\mu_1, \mu_2) = 2\Phi\left(\frac{|m_2 - m_1|}{2\sigma}\right) - 1, \tag{EC.4}$$

where Φ is the standard normal CDF.

Consider two normal random variables $\mu = \mathcal{N}(0,1)$ and $\nu = \mathcal{N}(1,1)$ as an example and assume that $\epsilon = \mathcal{N}(0,1)$. Then using Equation (EC.4) we have $d_{TV}(\mu, \nu) = 2\Phi\left(\frac{1}{2}\right) - 1$.

Now note that $T\mu = \text{Law}\big(aX + \varepsilon\big) = \mathcal{N}\Big(0,\ a^2 + 1\Big)$ where we used the fact that the sum of two independent normally distributed random variables is normal and its mean is the sum of the two means, and its variance is the sum of the two variances. Similarly, $T\nu = \mathcal{N}(a-\beta, a^2+1)$.

Using again Equation (EC.4) we have $d_{TV}(T\mu, T\nu) = 2\Phi\left(\frac{|a-\beta|}{2\sqrt{a^2+1}}\right) - 1$.

Hence, $d_{TV}(T\mu, T\nu) \geq d_{TV}(\mu, \nu)$ if $|a-\beta| \geq \sqrt{a^2+1}$ and T is generally not a contraction.

Augmented Markov Chain. Another possible approach is to augment the original nonlinear Markov chain to obtain a linear Markov process, thereby enabling the application of standard tools to establish existence and uniqueness of a stationary distribution (Meyn and Tweedie 2012). In the special case where the aggregator is given by $m(x) = \beta x$, this augmentation can be achieved by introducing a single additional variable representing the mean, rather than augmenting with the entire distribution.

To do that we introduce the mean coordinate as an additional variable: $m_t := \mathbb{E}[X_t]$, and consider the augmented Markov chain $Z_t := (X_t, m_t) \in \mathbb{R}^2$. We start from any distribution of X_0 with finite first two moments and $m_0 = \mathbb{E}[X_0]$.

Taking expectations in the original nonlinear Markov chain given in Equation (3) yields

$$m_{t+1} = a\,m_t - \beta\,m_t + e = (a-\beta)m_t + e.$$

where $e = \mathbb{E}(\epsilon_t)$. Hence with

$$A := \begin{pmatrix} a & -\beta \\ 0 & a-\beta \end{pmatrix}, \quad \xi_{t+1} := \begin{pmatrix} \epsilon_{t+1} \\ e \end{pmatrix},$$

the augmented Markov chain is an AR(1) process

$$Z_{t+1} = AZ_t + \xi_{t+1}, \qquad t \geq 0. \tag{EC.5}$$

on $\mathbb{R}^2$.

Assume that $|a-\beta| > 1$. Now it is easy to see that an invariant distribution for the augmented AR(1) process has to be of the form of the product measure $\mu^* \otimes \delta_{m^\star}$ where $m^\star = \dfrac{e}{1-(a-\beta)}$. In addition, $\mu^* \otimes \delta_{m^\star}$ is an invariant distribution for the augmented AR(1) process if and only if μ^* is an invariant distribution of the original nonlinear Markov chain with $\mathbb{E}(X_\infty) = m^*$.

However, when $|a-\beta| > 1$, the mean component m_t diverges for almost every initial condition m_0. Consequently, classical techniques from the theory of linear Markov chains, such as drift conditions Meyn and Tweedie (2012), that imply global stability do not apply.

Remark EC.1. We showed that the operator T for the nonlinear Markov chain in Example 1, given in Equation 3 with $m(x) = \beta x$ and studied above, is Lipschitz with respect to W_1. One may therefore wonder whether there is any necessary connection between Lipschitz continuity of T with respect to the Wasserstein distance W_p and the existence and uniqueness results established via our monotonicity approach. We now demonstrate that the answer is generally no. We present a cubic aggregator for which the associated nonlinear Markov operator T is not globally Lipschitz with respect to W_p for any $p \geq 1$, yet it still possesses a unique invariant distribution, a result that can be established directly using our uniqueness theorem.

Let $S = \mathbb{R}$ be endowed with the usual order. Fix $a \in (0,1)$ and $\beta > 0$, and let $(\varepsilon_t)_{t\geq 1}$ be i.i.d. with $\mathbb{E}[\varepsilon_1] = 0$, $\mathbb{E}[\varepsilon_1^2] = \sigma^2 \in (0,\infty)$, $\mathbb{E}\big[|\varepsilon_1|^3\big] < \infty$. For $r \geq 1$ set $\mathcal{P}_r(\mathbb{R}) := \Big\{\mu \in \mathcal{P}(\mathbb{R}) : \int_{\mathbb{R}} |x|^r \, \mu(dx) < \infty\Big\}$, and define the cubic aggregator $H_3 : \mathcal{P}_3(\mathbb{R}) \to \mathbb{R}$ by $H_3(\mu) := \int_{\mathbb{R}} x^3 \, \mu(dx)$. Consider the nonlinear Markov chain

$$X_{t+1} = aX_t - \beta H_3(\mu_t) + \varepsilon_{t+1}, \qquad \mu_t := \text{Law}(X_t),$$

and the associated one-step operator T given by $T\mu := \text{Law}\big(aX - \beta H_3(\mu) + \varepsilon\big)$.

Fix $p \geq 1$ such that $\mathbb{E}|\varepsilon_1|^p < \infty$ and write $r := \max\{3, p\}$. Then $T(\mathcal{P}_r(\mathbb{R})) \subseteq \mathcal{P}_r(\mathbb{R})$, but $T : (\mathcal{P}_r(\mathbb{R}), W_p) \to (\mathcal{P}_r(\mathbb{R}), W_p)$ is *not* globally Lipschitz. Indeed, for $x \neq y$ set $\mu = \delta_x$ and $\nu = \delta_y$. Then $W_p(\mu,\nu) = |x-y|$ and $H_3(\delta_z) = z^3$.

Moreover, $T\delta_z = \text{Law}\big(az - \beta z^3 + \varepsilon\big)$, which is a translate of $\text{Law}(\varepsilon)$ by the constant $az - \beta z^3$. Hence,

$$\begin{aligned} W_p(T\delta_x, T\delta_y) &= \big|a(x-y) - \beta(x^3 - y^3)\big| \\ &= \left| a - \beta \frac{x^3 - y^3}{x - y} \right| |x-y|. \end{aligned}$$

Therefore,

$$\frac{W_p(T\delta_x, T\delta_y)}{W_p(\delta_x, \delta_y)} = \left| a - \beta(x^2 + xy + y^2) \right|.$$

Since $x^2 + xy + y^2$ is unbounded on $\mathbb{R}^2$, no finite global W_p-Lipschitz constant can exist.

On the other hand, on the domain $W := \mathcal{P}_3(\mathbb{R})$ and with D the set of increasing functions, the monotonicity conditions for uniqueness follow from the same arguments as in Claim 1. Hence Proposition 4 yields uniqueness, and Proposition 2 yields existence of an invariant distribution in W. That is, T admits a unique invariant distribution in $\mathcal{P}_3(\mathbb{R})$.

EC.2.0.2. Contraction in Strategic Queueing Now consider the strategic queueing system we studied in Section 4.1. Suppose for simplicity that the system behaves as a strategic M/M/1 queue where $S \sim \text{Exp}(\mu)$ and $T(h) \sim \text{Exp}(\lambda(h))$ with $\mu = 1$.

We first define the CDF $F_{h;x}(t) := P\{Y_h(x) \le t\}$ of the random variable $Y_h(x) = \max\{0, x + S - T(h)\}$ which is given by 0 for $t < 0$ and

$$F_{h;x}(t) = \begin{cases} \dfrac{1}{1+\lambda(h)} \exp(\lambda(h)\,(t-x)), \\ \quad 0 \le t < x, \\ 1 - \dfrac{\lambda(h)}{1+\lambda(h)} \exp(-(t-x)), \\ \quad t \ge x. \end{cases} \tag{EC.6}$$

Indeed, this follows from the well known fact that the difference between the two exponential distributions S and $T(h)$ with parameters 1 and $\lambda(h)$, respectively is a Laplace distribution.

Wasserstein Distance. To test contraction we let $\mu = \delta_a$ for $a > 1$ and perturb it to $\nu_\varepsilon = (1-\varepsilon)\,\delta_a + \varepsilon\,\delta_c$ for some $c > a$ and small ε. We will focus on W_1 for this example but similar computations to show non-contraction of T can be done to other $p \ge 1$. The 1-Wasserstein distance between the measures is $W_1(\mu, \nu_\varepsilon) = \varepsilon|c - a|$.

We now compute the law of $T\mu$ and $T\nu_\varepsilon$. For $\mu = \delta_a$ the mean is a so the cumulative distribution function is given by $F_{T\mu}(t) = F_{a;a}(t)$ (see Equation EC.6). For $\nu_\varepsilon = (1-\varepsilon)\,\delta_a + \varepsilon\,\delta_c$ the mean is $\bar{h} = (1-\varepsilon)a + \varepsilon c$ so

$$F_{T\nu_\varepsilon}(t) = (1-\varepsilon)\,F_{\bar{h};a}(t) + \varepsilon\,F_{\bar{h};c}(t)$$
$$\bar{h} = (1-\varepsilon)a + \varepsilon c.$$

The 1-Wasserstein distance between these probability measures is

$$W_1(T\mu, T\nu_\varepsilon) = \int_0^\infty \left| F_{T\mu}(t) - F_{T\nu_\varepsilon}(t) \right| dt,$$

which can be evaluated analytically and numerically. In Figure EC.1a we evaluate for small values of ε and plot the ratio

$$K(\varepsilon) = \frac{W_1(T\mu, T\nu_\varepsilon)}{W_1(\mu, \nu_\varepsilon)}.$$

and show that it is above 1, i.e., T is not a global contraction in the W_1 metric.

Total Variation. As in the Wasserstein metric example, to test contraction we let $\mu = \delta_a$ for $a > 1$ and perturb it to $\nu_\varepsilon = (1-\varepsilon)\,\delta_a + \varepsilon\,\delta_c$ for some $c > a$ and small ε. The TV distance is $\|\mu - \nu_\varepsilon\|_{TV} = \varepsilon$. We will consider the measurable set $A = \{0\}$. From Equation (EC.6) and the analysis of the Wasserstein metric we have

$$T\mu(\{0\}) = F_{a;a}(0)$$
$$T\nu_\varepsilon(\{0\}) = (1-\varepsilon)F_{\bar{h}(\varepsilon);a}(0) + \varepsilon F_{\bar{h}(\varepsilon);c}(0).$$

where $\bar{h}(\varepsilon) := (1-\varepsilon)a + \varepsilon c$.

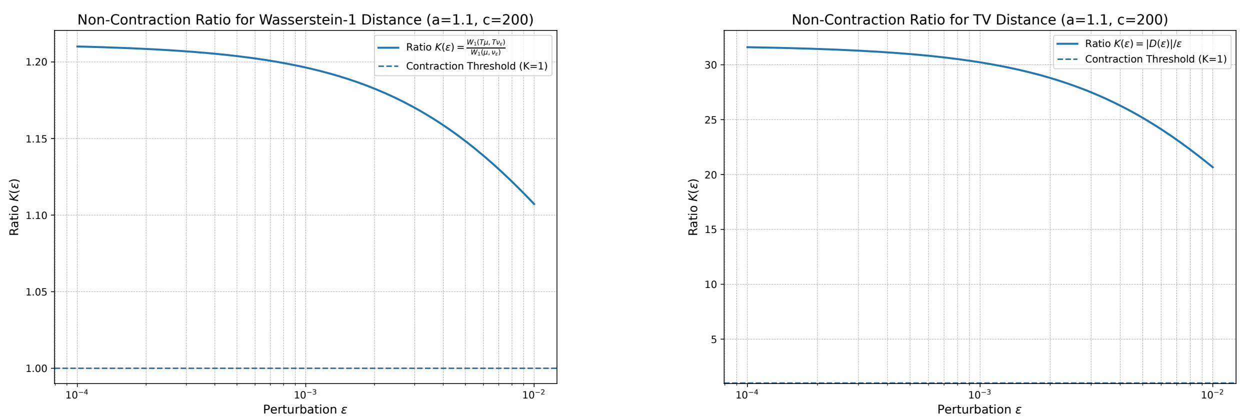
We define the difference $D(\varepsilon) := T\mu(\{0\}) - T\nu_\varepsilon(\{0\})$ and note that

$$\frac{\|T\mu - T\nu_\varepsilon\|_{TV}}{\|\mu - \nu_\varepsilon\|_{TV}} \geq \frac{|D(\varepsilon)|}{\varepsilon}$$

from the definition of the TV metric.

In Figure EC.1b we evaluate for small values of ε and compute numerically $|D(\epsilon)|/\epsilon$ and show that it is above 1, i.e., T is not a global contraction in the TV metric.

Figure EC.1 The Figures show the non-contraction in the Wasserstein and TV metric for the M/M/1 strategic queueing systems with parameters $\lambda(h) = 1/(h+1)$, $a = 1.1$, $c = 200$.

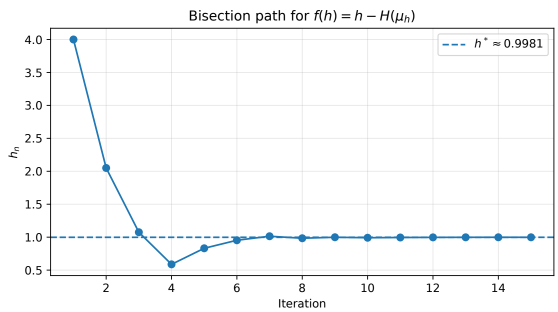


(a) Non-contraction in the Wasserstein metric

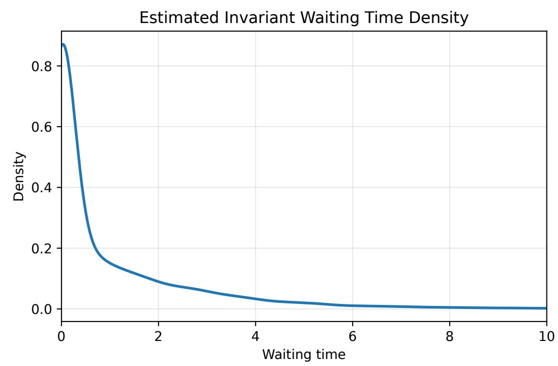
(b) Non-contraction in the TV metric

Despite this non-contraction, Algorithm 1 computes the equilibrium very fast and with only 15 iterations for the bisection method as we see in Figure EC.2a.

Figure EC.2 Figure EC.2a shows the sequence of midpoints $\{h_n\}$ generated by (Algorithm 1) when applied to the strategic M/M/1 queue with $S \sim \text{Exp}(1)$ and arrival rate $\lambda(h) = 1/(1+h)$. At each node the stationary mean $H(\mu_h) = \mathbb{E}[X_h]$ is estimated by a Monte Carlo simulation of the linear $M/M/1$ queue. Algorithm 1 finds the unique fixed point $h^* \approx 1.00$ (which can be calculated explicitly in this simple setting as in Claim 5) in fewer than 15 iterations with tolerance level 10^{-4}. Figure EC.2b shows the estimated invariant density of waiting times, obtained by applying a Gaussian kernel density estimator.



(a) Algorithm 1 convergence



(b) Equilibrium waiting time distribution

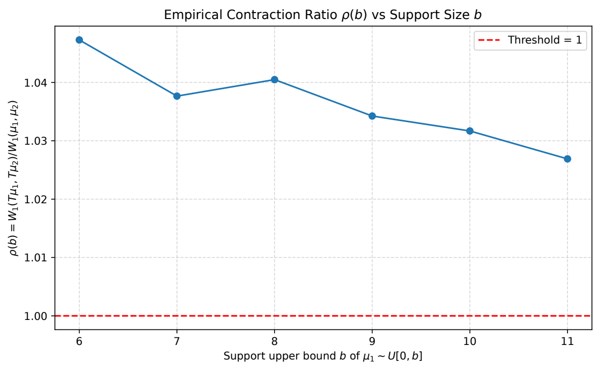
EC.2.0.3. Contraction in the Wealth–Distribution Model In this section we consider the wealth distribution application presented in Section 4.4. In particular, we consider a single asset case with a fixed interest rate that is determined by the aggregate agents' behavior in the economy.[2] In this model the optimal savings policy does not have closed-form solution. We compute it by using the value function iteration algorithm. In Figure EC.3 we first test numerically if T is a global contraction in the W_1 metric.[3] We then show that Algorithm 1 converges fast to the unique invariant distribution of this model and plot the corresponding wealth distribution.

While the monotonicity conditions of Theorem 1 guarantee uniqueness of the invariant distribution, the numerical simulations suggest that standard contraction arguments are inapplicable, as contraction is again not satisfied in this application.

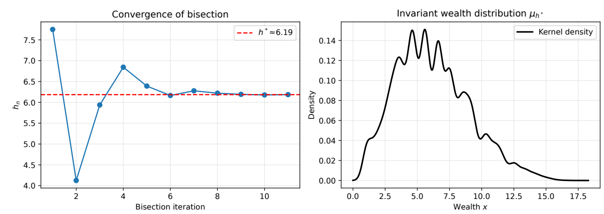
Notes for the Electronic Companion

[1]Indeed, let π be an optimal coupling of μ and ν, so $(X, Y) \sim \pi$ satisfies $W_1(\mu, \nu) = \mathbb{E}_\pi\left[|X - Y|\right]$. Let $\epsilon \sim \phi$ be an independent noise term. Define the valid coupling of $T\mu$ and $T\nu$, $(X', Y') = (aX - \beta\mathbb{E}[X] +$

Figure EC.3 In Figure EC.3a we compare two initial laws: $\mu_1 \sim U[0,b]$ with $b \in \{6,7,\dots,11\}$, and $\mu_2 \sim U[0,5]$. With $N = 4 \times 10^5$ draws from these distributions we compute $W_1(\mu_1,\mu_2)$ and $W_1(T\mu_1,T\mu_2)$, then plot the ratio $\rho(b) = \frac{W_1(T\mu_1,T\mu_2)}{W_1(\mu_1,\mu_2)}$ against b. We see that $\rho(b) > 1$; so T is not a global contraction. In Figure EC.3b we apply Algorithm 1 to compute the invariant distribution. The left panel displays the midpoint sequence $\{h_n\}$ generated by the bisection and convergence to the fixed point h^* is achieved in only 11 iterations. The right panel plots the density of the invariant wealth distribution μ_{h^*} where we apply a Gaussian kernel density estimator to plot it.



(a) Non-contraction in the wealth distribution model



(b) Computation of the unique invariant wealth distribution

$\epsilon, aY - \beta\mathbb{E}[Y] + \epsilon)$. Therefore by using the triangle inequality and Jensen's inequality we have

$$\begin{aligned} W_1(T\mu, T\nu) &\le \mathbb{E}_\pi\left[|a(X-Y) - \beta\mathbb{E}_\pi[X-Y]|\right] \\ &\le |a|\,\mathbb{E}_\pi\left[|X-Y|\right] + |\beta|\,|\mathbb{E}_\pi[X-Y]| \\ &\le \left(|a| + |\beta|\right)\mathbb{E}_\pi\left[|X-Y|\right]. \end{aligned}$$

That is,

$$W_1(T\mu, T\nu) \le \left(|a| + |\beta|\right) W_1(\mu, \nu).$$

[2]Specifically, we consider the Aiyagari model (Aiyagari 1994) as presented in Light (2020) (see the full details there). The production function is given by $f(k) = k^\alpha$ with $\alpha = 0.5$ and the interest rate is therefore given by $R(H(\mu)) = \alpha H(\mu)^{\alpha-1} - \delta + 1$ with $\delta = 0.1$ (as in Light (2020)) where $H(\mu) = \int x\mu(dx)$. The nonlinear Markov chain is then given by $X_{t+1} = R(H(\mu_t))\, g\left(X_t, R(H(\mu_t))\right) + Y_{t+1}$, where $\{Y_t\}$ are i.i.d. labor income shocks and $g(\cdot\,; R)$ is the optimal saving policy that is determined by an income–fluctuation problem with logarithmic utility. In this setting it can be shown that the savings policy function is increasing in the interest rate and current wealth and Property (C) holds (Light 2020) so we can use Corollary 4 to show that the nonlinear Markov chain has indeed a unique invariant distribution.

[3]In particular, given a fixed rate R, we solve the Bellman equation for the income fluctuation problem

$$V(x) = \max_{a \in [0,x]} \left\{\log(x-a) + \beta\mathbb{E}[V(Ra + Y')]\right\},$$

on the cash–on–hand grid $x \in [10^{-3}, 15]$ using value–function iteration with 800 iterations and a 25-point grid for the savings between 0 and x and interpolate the optimal savings to derive the optimal policy function $g(x, R)$. Then, to implement T, for any empirical wealth distribution represented by a sample $\{x_i\}_{i=1}^N$ we set $H(\mu) = \frac{1}{N} \sum_i x_i$, compute $R = R(H(\mu))$ and the corresponding policy $g(\cdot, R)$, draw i.i.d. income $Y_i \in \{1, 3\}$ with equal probabilities for each agent, and return $x_i' = R\, g_R(x_i) + Y_i$.