

# Necessary and Sufficient Conditions for Existence and Uniqueness of Recursive Utilities: Online Appendix.

Jaroslav Borovička<sup>a</sup> and John Stachurski<sup>b</sup>

<sup>a</sup> New York University, Federal Reserve Bank of Minneapolis and NBER

<sup>b</sup> Research School of Economics, Australian National University

April 22, 2019

## 1. THE BANSAL–YARON MODEL WITH TIME-VARYING VOLATILITY

In the main section of the paper we studied existence and uniqueness of finite valuations of consumption streams under the specification provided in section 1.A of [Bansal and Yaron \(2004\)](#), which features constant volatility. We also studied existence and uniqueness for the specification of [Schorfheide et al. \(2018\)](#). Here we repeat some of these exercises for the consumption specification from section 1.B of [Bansal and Yaron \(2004\)](#), which, like [Schorfheide et al. \(2018\)](#), has time-varying volatility.

The specification of consumption growth in section 1.B of [Bansal and Yaron \(2004\)](#) is

$$\begin{aligned}\ln(C_{t+1}/C_t) &= \mu_c + z_t + \sigma_t \eta_{c,t+1}, \\ z_{t+1} &= \rho z_t + \varphi_z \sigma_t \eta_{z,t+1}, \\ \sigma_{t+1}^2 &= \max \left\{ v \sigma_t^2 + d + \varphi_\sigma \eta_{\sigma,t+1}, 0 \right\}.\end{aligned}$$

Here  $\{\eta_{i,t}\}$  are IID and standard normal for  $i \in \{c, z, \sigma\}$ . The Markov state for consumption growth is taken to be  $X_t = (z_t, \sigma_t)$ . As in [Bansal and Yaron \(2004\)](#), the preference parameters are set to  $\gamma = 10.0$ ,  $\beta = 0.998$  and  $\psi = 1.5$ , while those for in the consumption process are  $\mu_c = 0.0015$ ,  $\rho = 0.979$ ,  $\varphi_z = 0.044$ ,  $v = 0.987$ ,  $d = 7.9092e-7$  and  $\varphi_\sigma = 2.3e-6$ .

Following the same procedure as for the [Schorfheide et al. \(2018\)](#) model, we begin with the Monte Carlo method, so that compactification is implemented implicitly at the largest double precision floating point number. With  $n = m = 1000$ , we obtain a value of

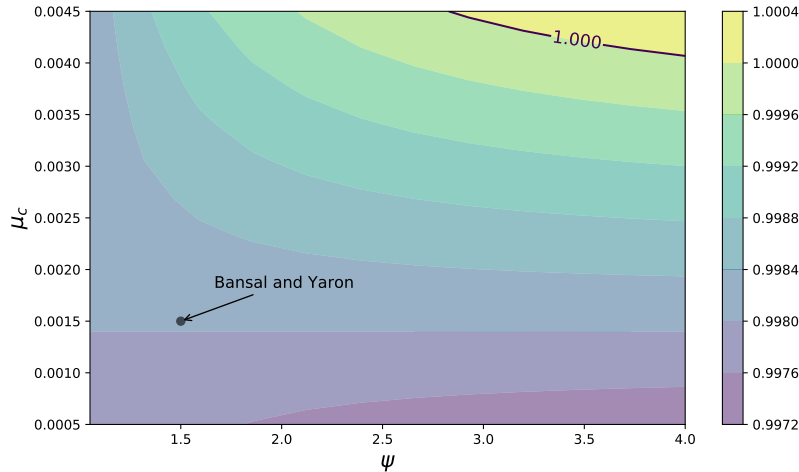


FIGURE 1. Changes in test value,  $\psi$  vs  $\mu_c$

approximately 0.998.<sup>1</sup> Figure 1 reinforces the finding of stability by showing estimated values of  $\Lambda$  at a range of different parameterizations, represented as a contour plot. The figure shows that the conclusion  $\Lambda < 1$  survives large deviations in the intertemporal elasticity of substitution parameter  $\psi$  and the mean consumption growth rate  $\mu_c$ . In particular, starting from the baseline parameterization of [Bansal and Yaron \(2004\)](#), one would have to increase both of these parameters to very high levels in order to obtain a parameterization at which  $\Lambda$  exceeds unity.<sup>2</sup>

## 2. UNBOUNDEDNESS RESULT

In the next part of this online appendix we prove the following result, which was stated in the main part of the paper.

<sup>1</sup>Generating the statistic 1,000 times produced a mean value of 0.998128 with standard deviation 0.000053. We also implemented a discretized version, which led to a similar outcome. Discretization was achieved by multiple iterations of the univariate Rouwenhorst method. We first applied the method to  $\{\sigma_t\}$ , producing a Markov chain with  $I$  states. Then, for each of the  $I$  possible values of  $\sigma_t$ , we again used the Rouwenhorst method to discretize  $\{z_t\}$  across  $J$  possible states. The implementations are identical to those used for the [Schorfheide et al. \(2018\)](#) model, apart from the modified specification for consumption growth. When  $I = J = 4$ , so that the state space has 16 elements, the spectral radius method returned  $\Lambda = 0.998044$ .

<sup>2</sup>At each parameterization, the value  $\Lambda$  was calculated by the non-discretized Monte Carlo method, with  $m = n = 1000$ .

**Theorem 2.1.** *If assumptions 2.1 and 6.1 hold, then the following statements are equivalent:*

- (a)  $\Lambda < 1$ .
- (b)  $A$  has a fixed point in  $\mathcal{C}$ .
- (c) There exists a  $g \in \mathcal{C}$  such that  $\{A^n g\}_{n \geq 1}$  converges to an element of  $\mathcal{C}$ .

Throughout all of the following, the state space  $\mathbb{X}$  is allowed to be any  $\sigma$ -compact metric space. The collection of Borel measurable functions  $g$  from  $\mathbb{X}$  to  $\mathbb{R}$  such that  $\|g\| := \int |g| d\pi$  is denoted by  $L_1(\mathbb{X}, \pi)$ . Convergence is with respect to  $\|\cdot\|$  unless otherwise stated. For  $g, h \in L_1(\mathbb{X}, \pi)$ , the statement  $g \ll h$  means that  $g < h$  holds  $\pi$ -almost everywhere. Other definitions are as stated in the main section of the paper.

One step of the proof concerns continuity of the spectral radius when a positive operator is approximated from below. The following lemma presents such a result, which builds on a valuable theorem due to [Schep \(1980\)](#).

**Lemma 2.2.** *Let  $\{T_n\}$  and  $T$  be bounded linear operators on  $L_1(\mathbb{X}, \pi)$  such that  $0 \leq T_n \leq T_{n+1} \leq T$  for all  $n$ . If  $T_n f \rightarrow T f$  in norm as  $n \rightarrow \infty$  for each  $f$  in the positive cone  $\mathcal{C}$  and  $T^i$  is compact for some  $i \in \mathbb{N}$ , then  $r(T_n) \uparrow r(T)$ .*

*Proof.* It is clear that  $0 \leq T_n^i \leq T_{n+1}^i \leq T^i$  for all  $n$ . Moreover, given  $f \in \mathcal{C}$ , we have  $T_n^i f \rightarrow T^i f$  as  $n \rightarrow \infty$ , as follows from induction. Indeed,  $T_n f \rightarrow T f$ , and if  $T_n^{i-1} f \rightarrow T^{i-1} f$ , then

$$\begin{aligned} \|T_n^i f - T^i f\| &\leq \|T_n^i f - T_n T^{i-1} f\| + \|T_n T^{i-1} f - T^i f\| \\ &\leq \|T_n\| \cdot \|T_n^{i-1} f - T^{i-1} f\| + \|T_n T^{i-1} f - T^i f\| \\ &\leq \|T\| \cdot \|T_n^{i-1} f - T^{i-1} f\| + \|T_n T^{i-1} f - T T^{i-1} f\| \rightarrow 0. \end{aligned}$$

Here the converges to zero of the first term is by the induction hypothesis, while the second is by the fact that  $T^{i-1} f \in \mathcal{C}$ .

We can now apply the spectral continuity result of [Schep \(1980\)](#), theorem 2.4, to the family of operators  $\{T_n^i, T^i\}$ , noting that  $T^i$  is, by assumption, compact, and also order continuous because it is norm continuous (see [Zaanen \(1997\)](#), p. 147). This gives  $r(T_n^i) \uparrow r(T^i)$ . But then  $r(T_n) \uparrow r(T)$  also holds.  $\square$

The next lemma gives conditions under which the limit of a sequence of fixed points associated with a sequence of approximating maps is itself a fixed point.

**Lemma 2.3.** *Let  $(E, d)$  be a metric space and let  $T$  and  $\{T_m\}_{m \in \mathbb{N}}$  be self-maps on  $E$  with the property  $T_m u \rightarrow Tu$  for all  $u \in E$ . Let  $\bar{u}_m$  be a fixed point of  $T_m$  for each  $m$  and suppose that  $\bar{u}_m \rightarrow \bar{u}$  for some  $\bar{u} \in E$ . If  $T$  is continuous on  $E$  and the maps  $\{T_m\}$  are uniformly Lipschitz continuous, then  $\bar{u}$  is a fixed point of  $T$ .*

*Proof.* We have  $T\bar{u} = \lim_{m \rightarrow \infty} T_m \bar{u}$  so one only need show that  $T_m \bar{u} \rightarrow \bar{u}$ . To this end, observe that

$$d(T_m \bar{u}, \bar{u}) \leq d(T_m \bar{u}, T_m \bar{u}_m) + d(T_m \bar{u}_m, \bar{u}) = d(T_m \bar{u}, T_m \bar{u}_m) + d(\bar{u}_m, \bar{u}).$$

With  $L$  as the uniform Lipschitz constant for  $\{T_m\}$ , this yields

$$d(T_m \bar{u}, \bar{u}) \leq (L + 1)d(\bar{u}_m, \bar{u}) \rightarrow 0. \quad \square$$

**Lemma 2.4.** *If assumption 6.1 holds, then  $K^i$  is compact for some  $i \in \mathbb{N}$ .*

*Proof.* Under assumption 6.1, there exists an  $m \in \mathbb{N}$  such that  $K^m$  is weakly compact. But then, by theorem 9.9 of Schaefer (1974), corollary 1,  $K^{2m}$  is compact. Set  $i = 2m$ .  $\square$

Now we turn to the proof of theorem 2.1. The proof of existence of a fixed point in theorem 2.1 uses a limiting argument based on approximating  $\mathbb{X}$  with compact sets. To set this up, let  $\{F_m\}_{m \geq 1}$  be a sequence of compact subsets of  $\mathbb{X}$  with  $F_m \subset F_{m+1}$  for all  $m \in \mathbb{N}$  and  $\cup_{m \geq 1} F_m = \mathbb{X}$ . Let  $K_m$  be the operator on  $L_1(\mathbb{X}, \pi)$  defined by

$$K_m g(x) = \mathbb{1}\{x \in F_m\} \int_{F_m} k(x, y) g(y) dy \quad (x \in \mathbb{X})$$

Note that  $K_m$  is also a positive linear operator and  $0 \leq K_m \leq K_{m+1} \leq K$  for all  $m \in \mathbb{N}$ . It follows that  $K_m$  is a bounded linear operator on  $L_1(\mathbb{X}, \pi)$ .

**Lemma 2.5.** *If  $f \in \mathcal{C}$ , then  $\|K_m f - Kf\| \rightarrow 0$  as  $m \rightarrow \infty$ .*

*Proof.* Fix  $f \in \mathcal{C}$ . For any given  $m \in \mathbb{N}$ , we have

$$\|K_m f - Kf\| \leq \int \int k(x, y) e_m(x, y) |f(y)| dy \pi(dx)$$

where  $e_m(x, y) := 1 - \mathbb{1}_{F_m}(x) \mathbb{1}_{F_m}(y)$ . Since  $K$  is a bounded linear operator, the integral on the right hand side is finite, so we need only show that the integrand converges pointwise to zero. But this is immediate from the definition of  $\{F_m\}$ .  $\square$

Given  $g: F_m \rightarrow \mathbb{R}$ , its *extension*  $e_m g$  to  $\mathbb{X}$  is defined as the function equal to  $g$  on  $F_m$  and zero on  $F_m^c$ . Given  $g: \mathbb{X} \rightarrow \mathbb{R}$ , its *restriction*  $c_m g$  to  $F_m$  is defined as the function

$c_m g$  equal to  $g$  on  $F_m$ . In addition, let  $\bar{K}$  be the restriction of  $K_m$  to real functions on  $F_m$ . That is,

$$\bar{K}_m g(x) = \int_{F_m} k(x, y) g(y) dy \quad (x \in F_m).$$

We regard  $\bar{K}_m$  as a mapping on  $L_1(F_m, \bar{\pi}_m)$ , where  $\bar{\pi}_m := c_m \pi$ . Note that

$$A_m = e_m \bar{A}_m c_m \text{ on } \mathcal{C} \quad (1)$$

where  $A_m := \varphi \circ K_m$  and  $\bar{A}_m := \varphi \circ \bar{K}_m$ . The latter is a self-mapping on  $\mathcal{C}_m$ , the positive cone of  $L_1(F_m, \bar{\pi}_m)$ .

**Lemma 2.6.** *If  $g$  in  $\mathcal{C}_m$  is a fixed point of  $\bar{A}_m$ , then  $e_m g$  is a fixed point of  $A_m$ .*

*Proof.* For such  $g \in \mathcal{C}_m$  we have  $A_m e_m g = e_m \bar{A}_m c_m e_m g = e_m \bar{A}_m g = e_m g$ .  $\square$

**Lemma 2.7.** *We have  $\|\bar{K}_m\| = \|K_m\|$  for all  $m \in \mathbb{N}$ .*

*Proof.* Fix  $f \in L_1(\mathbb{X}, \pi)$  with  $\|f\| \leq 1$ . Let  $\bar{f}$  be the restriction of  $f$  to  $F_m$ . Note that

$$\|\bar{f}\| = \int |\bar{f}| \bar{\pi}(x) dx \leq \|f\| \leq 1.$$

We have

$$\|\bar{K}_m \bar{f}\| = \int_{F_m} \left| \int_{F_m} k(x, y) \bar{f}(y) dy \right| \bar{\pi}(x) dx = \int |K_m f(x)| \pi(x) dx = \|K_m f\|.$$

In particular,  $\|K_m f\| = \|\bar{K}_m \bar{f}\| \leq \|\bar{K}_m\|$  and hence, taking the supremum on the left hand side,  $\|K_m\| \leq \|\bar{K}_m\|$ .

To see that the reverse inequality holds, fix  $\bar{f} \in L_1(F_m, \bar{\pi})$  with  $\|\bar{f}\| \leq 1$ . Let  $f \in L_1(\mathbb{X}, \pi)$  be defined by  $f = \bar{f}$  on  $F_m$  and  $f = 0$  elsewhere. Note that

$$\|f\| = \int |f| \pi(x) dx = \int |\bar{f}| \bar{\pi}(x) dx = \|\bar{f}\| \leq 1.$$

In addition, by an identical argument to that given just above, we have  $\|\bar{K}_m \bar{f}\| = \|K_m f\|$ . It follows that  $\|\bar{K}_m \bar{f}\| \leq \|K_m\|$ , and taking the supremum on the left hand over all such  $\bar{f}$  yields  $\|\bar{K}_m\| \leq \|K_m\|$ .  $\square$

**Lemma 2.8.** *If  $r(K) > 1/\beta^\theta$ , then there exists an  $M \in \mathbb{N}$  such that  $r(\bar{K}_m) > 1/\beta^\theta$  whenever  $m \geq M$ .*

*Proof.* In view of lemma 2.7 and the definition of the spectral radius, it suffices to prove that  $r(K_m) > 1$  for sufficiently large  $m$ . This will be true if  $r(K_m) \rightarrow r(K)$ , which, by lemma 2.2, will hold if (a)  $K^i$  is compact for some  $i \in \mathbb{N}$ , (b)  $0 \leq K_m \leq K_{m+1} \leq K$  for all  $m$  and (c)  $K_m f \rightarrow K f$  in norm for each  $f$  in  $L_p(\mathbb{X}, \pi)_+$ . We already have (a) by lemma 2.4 and (b) is true by construction. Finally, (c) holds by lemma 2.5.  $\square$

**Proposition 2.9.** *Under the conditions of theorem 2.1,  $\Lambda$  is well defined and satisfies  $\Lambda = \beta r(K)^{1/\theta}$ .*

The proof is identical to that of proposition B.4 in the paper itself.

**Lemma 2.10.** *If  $\theta < 0$  and  $\Lambda < 1$ , then there exists an  $M \in \mathbb{N}$  such that, for all  $m \geq M$ , the operator  $A_m$  has a nonzero fixed point  $g_m$  in  $\mathcal{C}$ , and  $g_m \leq g_{m+1}$  for all such  $m$ .*

*Proof.* If  $\theta < 0$  and  $\Lambda < 1$ , then, by proposition 2.9, we have  $r(K) > 1/\beta^\theta$ . Now let  $M$  be as in lemma 2.8 and take  $m \geq M$ . Observe that  $\bar{A}_m$  has a unique nonzero fixed point  $\bar{g}_m$  in  $\mathcal{C}_m$ , since  $F_m$  is compact (as follows from our results in the main section of the paper). It then follows from (1) that  $A_m e_m \bar{g}_m = e_m \bar{A}_m c_m e_m \bar{g}_m = e_m \bar{g}_m$ , so  $g_m := e_m \bar{g}_m$  is a fixed point of  $A_m$ . Since  $\bar{g}_m$  is nonzero on  $F_m$ , the function  $g_m$  is nonzero on  $\mathbb{X}$ .

It remains to prove that  $g_m \leq g_{m+1}$  for all  $m \geq M$ . To see that this is so, pick any such  $m$  and observe that, since  $K_m \leq K_{m+1}$  on  $\mathcal{C}$  and  $\varphi$  is increasing, we have  $A_{m+1} g_m \geq A_m g_m = g_m$ . Using isotonicity of  $A_{m+1}$  and iterating now gives  $A_{m+1}^n g_m \geq g_m$  for all  $n$ . Moreover, since  $g_m$  is nonzero on  $F_m$  and hence on  $F_{m+1}$ , the convergence result in theorem 3.1 from the main text applied on the compact state space  $F_{m+1}$  implies that  $A_{m+1}^n g_m \rightarrow g_{m+1}$  uniformly. Hence  $g_{m+1} \geq g_m$ , as was to be shown.  $\square$

**Lemma 2.11.** *If  $\theta < 0$ , then the family  $\{A_n\}$  is uniformly Lipschitz continuous on  $\mathcal{C}$ .*

*Proof.* When  $\theta < 0$ , the scalar map  $\varphi$  is Lipschitz with Lipschitz constant 1. Hence, for arbitrary  $m \in \mathbb{N}$  and  $f, g \in \mathcal{C}$ , we have

$$|A_m f - A_m g| \leq |K_m f - K_m g| = |K_m(f - g)| \leq K_m |f - g| \leq K |f - g|.$$

$$\therefore \|A_m f - A_m g\| \leq \|K\| \cdot \|f - g\|.$$

$\square$

*Proof of theorem 2.1.* First we show that (a) and (b) are equivalent under the stated assumptions, starting with the implication (a)  $\implies$  (b). For the case  $\theta > 0$  there is nothing to show, since the existence component of the proof for the compact case (from the main section of the paper) does not require compactness of  $\mathbb{X}$ . Hence we focus below on the case  $\theta < 0$ . Our proof uses lemma 2.3, with the metric space  $(\mathcal{C}, \|\cdot\|)$  and maps  $\{A_m\}$  and  $A$ .

First observe that, by lemma 2.5, for each given  $f \in \mathcal{C}$ , we have  $K_m f \rightarrow K f$  as  $m \rightarrow \infty$ . Since  $\varphi$  is Lipschitz of order 1 when  $\theta < 0$ , it follows immediately that  $A_m f \rightarrow A f$  as  $m \rightarrow \infty$ . By lemma 2.10, there exists an  $M \in \mathbb{N}$  such that, for all  $m \geq M$ , the operator  $A_m$  has a nonzero fixed point  $g_m$  in  $\mathcal{C}$ , and  $g_m \leq g_{m+1}$  for all such  $m$ . Since  $\varphi$  is bounded

above by  $b := (1 - \beta)^\theta$  when  $\theta < 0$ , it must be that  $g_m \leq g_{m+1} \leq b$  for all  $m$ . Any order bounded monotone sequence in  $L_1(\mathbb{X}, \pi)$  converges to an element of that set. This limit we denote by  $g$ . In view of lemma 2.3, this  $g$  will be a fixed point of  $A$  whenever  $A$  is continuous and  $\{A_m\}$  is uniformly Lipschitz continuous. Continuity of  $A$  is immediate from the properties of  $K$  and  $\varphi$ , while uniform Lipschitz continuity of  $\{A_m\}$  follows from lemma 2.11. The proof of (a)  $\implies$  (b) is now done.

Now we turn to (b)  $\implies$  (a). Here we can use exactly the same proof that we used when  $\mathbb{X}$  was compact, in proposition B.9 in the main text. There, compactness of  $\mathbb{X}$  was only used to ensure that  $K^i$  is a compact operator in  $L_1(\mathbb{X}, \pi)$  for some  $i \in \mathbb{N}$ . The latter property still holds in our setting, by lemma 2.4.

Finally, (b) and (c) are also equivalent. That (b) implies (c) is obvious. That (c) implies (b) follows from continuity of  $A$  as a self-map on  $\mathcal{C}$ , since then the limit  $g^* := \lim_{n \rightarrow \infty} A^n g$  is a fixed point of  $A$ .  $\square$

### 3. VALUATION UNDER IID GROWTH

In this section we investigate an interpretation of  $\Lambda$  related to valuation of consumption strips. In particular, we show that the condition  $\Lambda < 1$ , which guarantees finite valuation in the recursive utility model, implies an upper bound on the risk-adjusted growth rate of consumption relative to the risk-free rate. To preserve tractability, we study the case where  $\ln(C_{t+1}/C_t) = \mu + \sigma W_{t+1}$  for some IID sequence  $\{W_t\}$ . The risk-adjusted growth rate of consumption is then given by

$$\mathcal{M}_C = \left\{ E \left[ \left( \frac{C_{t+1}}{C_t} \right)^{1-\gamma} \right] \right\}^{\frac{1}{1-\gamma}} = \exp \left( \mu + \frac{1}{2} (1-\gamma) \sigma^2 \right),$$

and hence

$$\Lambda = \beta (\mathcal{M}_C)^{1-\frac{1}{\psi}} = \beta \exp \left( \left( 1 - \frac{1}{\psi} \right) \left( \mu + \frac{1}{2} (1-\gamma) \sigma^2 \right) \right).$$

In this environment, the scaled continuation value  $V_t/C_t \doteq v$  is constant. The Epstein and Zin (1989) stochastic discount factor  $S_t$  can then be written as

$$\frac{S_{t+1}}{S_t} = \beta \left( \frac{C_{t+1}}{C_t} \right)^{-\frac{1}{\psi}} \left[ \frac{V_{t+1}}{E_t [V_{t+1}^{1-\gamma}]^{1/(1-\gamma)}} \right]^{\frac{1}{\psi}-\gamma} = \beta (\mathcal{M}_C)^{\gamma-\frac{1}{\psi}} \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma}.$$

From this, we have the growth rate in the value of consumption strips as their maturity increases

$$E_t \left[ \frac{S_{t+1} C_{t+1}}{S_t C_t} \right] = E_t \left[ \beta (\mathcal{M}_C)^{\gamma - \frac{1}{\psi}} \left( \frac{C_{t+1}}{C_t} \right)^{1-\gamma} \right] = \beta (\mathcal{M}_C)^{1 - \frac{1}{\psi}} = \Lambda.$$

The present value of a consumption strip maturing at  $t + j$  is therefore  $C_t \Lambda^j$ , and the condition  $\Lambda < 1$  assures that the wealth-consumption ratio, given by the present discounted value of these strips  $(1 - \Lambda)^{-1}$ , is finite.

The risk-free rate in this IID setting is constant as well:

$$r^f = -\log E_t \left[ \frac{S_{t+1}}{S_t} \right] = -\log \beta + \frac{1}{\psi} \left( \mu + \frac{1}{2} (1 - \gamma) \sigma^2 \right) - \frac{1}{2} \gamma \sigma^2,$$

and we can therefore rewrite the condition  $\Lambda < 1$  as

$$\log \Lambda = \log \mathcal{M}_C - r^f - \frac{1}{2} \gamma \sigma^2 < 0$$

implying an upper bound on the risk-adjusted growth rate of consumption relative to the risk-free rate.

#### 4. DETAILS OF THE LEARNING MODEL

In this section, we outline the details of the unobserved state model from section 4.3 of the paper. The agent starts with a prior  $P(X_0 = 1) = \bar{X}_0$  and observes data  $Z_t = \ln(C_t/C_{t-1})$ ,  $t = 1, 2, \dots$ . Denote the density of the data conditional on the unobserved state

$$\varphi_i(Z_t) \doteq \varphi(Z_t | X_t = i) = \frac{1}{\sqrt{2\pi\sigma(i)^2}} \exp\left(-\frac{1}{2\sigma(i)^2} (Z_t - \mu(i))^2\right).$$

and  $p(\bar{X}_t) \doteq [\bar{X}_t q(1, 1) + (1 - \bar{X}_t) q(2, 1)]$  the conditional expectation of  $\bar{X}_{t+1}$  given  $\bar{X}_t$ . The posterior probability  $\bar{X}_{t+1} = P(X_{t+1} = 1 | Z^{t+1})$  can be written recursively as

$$\begin{aligned} \bar{X}_{t+1} &= \frac{P(X_{t+1} = 1, Z_{t+1} | Z^t)}{P(Z_{t+1} | Z^t)} = \frac{\sum_i P(X_{t+1} = 1, X_t = i, Z_{t+1} | Z^t)}{\sum_{i,j} P(X_{t+1} = j, X_t = i, Z_{t+1} | Z^t)} \\ &= \frac{\sum_{i=1}^2 P(Z_{t+1} | X_{t+1} = 1) P(X_{t+1} = 1 | X_t = i) P(X_t = i | Z^t)}{\sum_{i,j=1}^2 P(Z_{t+1} | X_{t+1} = j) P(X_{t+1} = j | X_t = i) P(X_t = i | Z^t)} \\ &= \frac{\varphi_1(Z_{t+1}) [\bar{X}_t q(1, 1) + (1 - \bar{X}_t) q(2, 1)]}{\sum_{j=1}^2 \varphi_j(Z_{t+1}) [\bar{X}_t q(1, j) + (1 - \bar{X}_t) q(2, j)]} \\ &= \frac{p(\bar{X}_t) \varphi_1(Z_{t+1})}{p(\bar{X}_t) (\varphi_1(Z_{t+1}) - \varphi_2(Z_{t+1})) + \varphi_2(Z_{t+1})} \doteq h(\bar{X}_t, Z_{t+1}) \end{aligned}$$



The conditional distribution function  $\bar{Q}(x, y) = P(\bar{X}_{t+1} \leq y | \bar{X}_t = x)$  can then be written as

$$\begin{aligned}\bar{Q}(x, y) &= \int_{z: h(x, z) \leq y} \left( \sum_{j=1}^n \varphi_j(z) P(X_{t+1} = j | \bar{X}_t = x) \right) dz \\ &= \int_{z: h(x, z) \leq y} [p(x) (\varphi_1(z) - \varphi_2(z)) + \varphi_2(z)] dz\end{aligned}$$

The distribution is constructed by integrating over the likelihood of the data  $\ln(C_{t+1}/C_t) = z$  conditional on  $\bar{X}_t$  over all consumption growth realizations  $z$  that lead to a particular updated value  $\bar{X}_{t+1} = h(\bar{X}_t, z) \leq y$ .

The function  $h(x, z)$ , for a given  $x$ , achieves its maximum at

$$z^* = \frac{\sigma_1^{-2} \mu_1 - \sigma_2^{-2} \mu_2}{\sigma_1^{-2} - \sigma_2^{-2}},$$

irrespective of  $x$ , which is the value of  $z$  that maximizes the ratio of the likelihoods  $\varphi_1(z) / \varphi_2(z)$ . Given the shape of the function  $h(x, z)$ , we can define two strictly monotone functions  $h_x^1(z) = h(x, z)$  for  $z \leq z^*$  and  $h_x^2(z) = h(x, z)$  for  $z \geq z^*$ , and then write the distribution function as  $\bar{Q}(x, y) = 1$  for  $y \geq h(x, z^*)$  and

$$\begin{aligned}\bar{Q}(x, y) &= \int_{-\infty}^{(h_x^1)^{-1}(y)} [p(x) (\varphi_1(z) - \varphi_2(z)) + \varphi_2(z)] dz + \\ &\quad + \int_{(h_x^2)^{-1}(y)}^{\infty} [p(x) (\varphi_1(z) - \varphi_2(z)) + \varphi_2(z)] dz.\end{aligned}$$

We then obtain the density  $\bar{q}(x, y)$  by differentiating  $\bar{Q}(x, y)$  with respect to its second argument, which yields  $\bar{q}(x, y) = 0$  for  $y \geq h(x, z^*)$  and

$$\begin{aligned}\bar{q}(x, y) &= \frac{1}{(h_x^1)'(z)} [p(x) (\varphi_1(z) - \varphi_2(z)) + \varphi_2(z)] \Big|_{z=(h_x^1)^{-1}(y)} \\ &\quad - \frac{1}{(h_x^2)'(z)} [p(x) (\varphi_1(z) - \varphi_2(z)) + \varphi_2(z)] \Big|_{z=(h_x^2)^{-1}(y)}\end{aligned}$$

for  $y < h(x, z^*)$ .

The dashed lines in the top row of figure 2 depict the two conditional densities  $\varphi_i(z)$  in the left panel, and the function  $h(x, z)$  for three values of the current state  $\bar{X}_t = x$  in the right panel, as a function of the realization of the data  $Z_{t+1} = z$ . The bottom row, again depicted in dashed lines, shows the implied cdf  $\bar{Q}(x, y)$  of the transition density.

The transition density  $\bar{q}(x, y)$  in the benchmark specification of the model does not satisfy the assumptions imposed in section 3 of the main text. First, the transition density  $\bar{q}(x, y)$

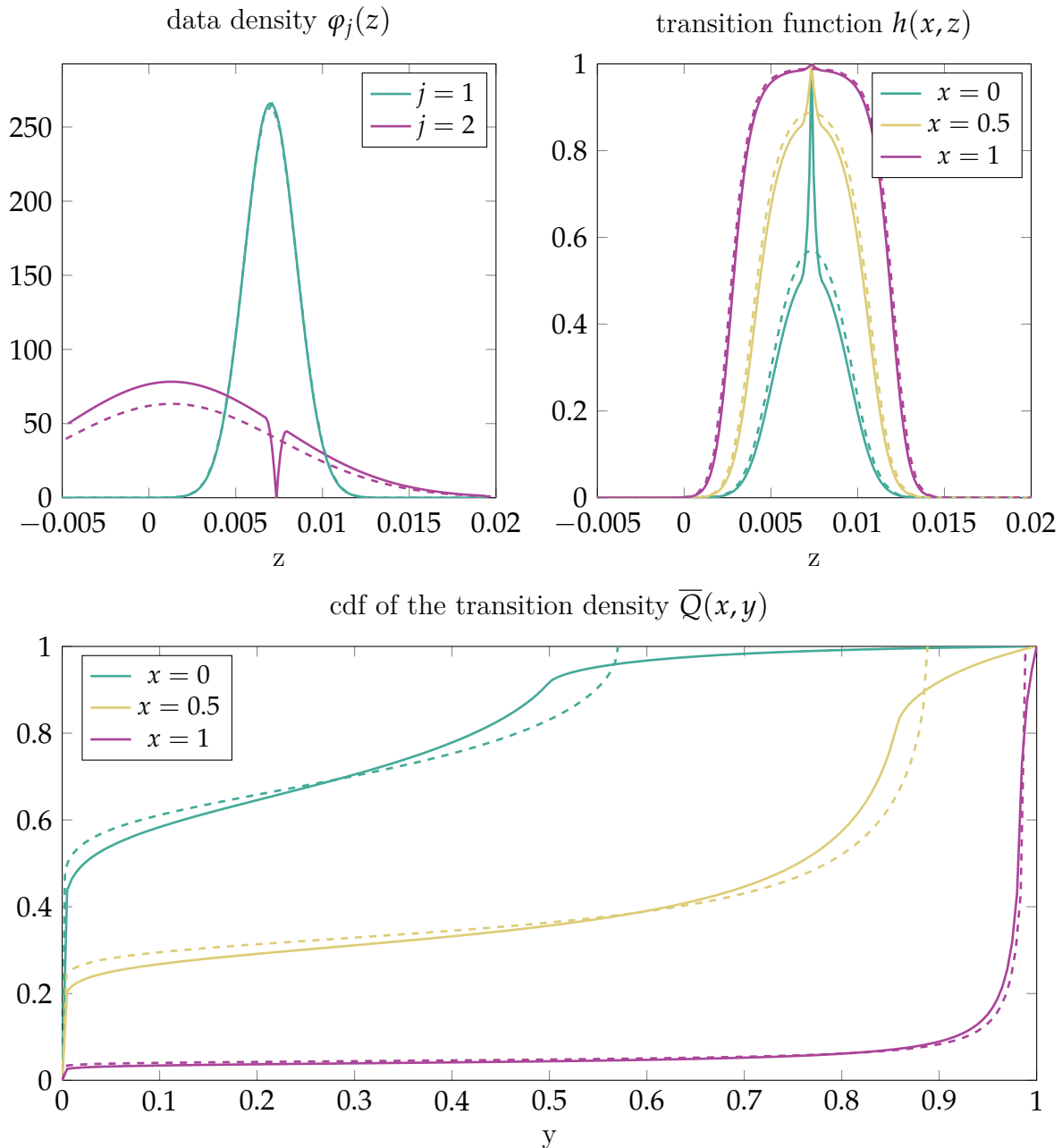


FIGURE 2. Original (dashed lines) and perturbed (solid lines) densities and transition functions in the learning model.

is zero for any  $y > h(x, z^*)$ , and hence the  $l$ -step transition density  $\bar{q}^l$  is not everywhere positive for any finite  $l$ . This is apparent from the top right plot of figure 2 where the function  $h(x, z)$  (in dashed lines) does not reach one, and is caused by the fact that the likelihood ratio  $\varphi_1(z) / \varphi_2(z)$  is bounded from above. In other words, given the two densities  $\varphi_j(z)$ , there is no realization of consumption growth  $Z_{t+1} = z$  that would

allow the agent to conclude with an arbitrary degree of certainty that the current state is  $X_{t+1} = 1$ . Second,  $\bar{q}(x, \mathbf{y})$  diverges to infinity as  $\mathbf{y} \searrow 0$ . This indicates that the tails of  $\varphi_2(z)$  are ‘too fat’ relative to  $\varphi_1(z)$  and therefore extreme realizations of  $Z_{t+1} = z$  that allow the agent to determine with a high degree of certainty that the state is  $X_{t+1} = 2$  are too likely to keep  $\bar{q}(x, \mathbf{y})$  bounded as  $\mathbf{y} \searrow 0$ .

Nevertheless, we can, construct a perturbation  $\tilde{\varphi}_j(z)$  of the data densities  $\varphi_j(z)$  such that

- (a) the original full information problem introduced in section 4.2 of the paper is unchanged;
- (b) the perturbed data densities  $\tilde{\varphi}_j(z)$  and the implied perturbed transition distribution  $\tilde{Q}(x, \cdot)$  are arbitrarily close to the original objects  $\varphi_j(z)$  and  $\bar{Q}(x, \mathbf{y})$ , in the sense that for any  $\varepsilon > 0$ , we can construct a perturbation  $\tilde{\varphi}_j(z)$  such that

$$\sup_{A \subseteq \mathbb{R}} \int_A |\varphi_j(z) - \tilde{\varphi}_j(z)| dz < \varepsilon$$

$$\sup_{x \in [0,1], B \subseteq [0,1]} \left| \bar{Q}(x, B) - \tilde{Q}(x, B) \right| < \varepsilon$$

The first statement holds when the densities  $\tilde{\varphi}_j(z)$  imply the same moments  $\zeta(j)$  in equation (32) of the main text as in the original specification. The second statement implies that the operator  $\bar{K}$  in (32) is approximated to an arbitrary degree of accuracy for the class of bounded functions. Details of the construction of  $\tilde{\varphi}_j(z)$  can be obtained from the authors on request.

## REFERENCES

- Bansal, R. and Yaron, A. (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *The Journal of Finance*, 59(4):1481–1509.
- Epstein, L. G. and Zin, S. E. (1989). Risk aversion and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica*, 57(4):937–969.
- Schaefer, H. (1974). *Banach Lattices and Positive Operators*. Berlin-Heidelberg, New York.
- Schep, A. (1980). Positive diagonal and triangular operators. *Journal of Operator Theory*, pages 165–178.
- Schorfheide, F., Song, D., and Yaron, A. (2018). Identifying long-run risks: A bayesian mixed-frequency approach. *Econometrica*, 86(2):617–654.
- Zaanen, A. C. (1997). *Introduction to Operator Theory in Riesz Spaces*. Springer.