Internet Appendix for "Necessary and Sufficient Conditions for Existence and Uniqueness of Recursive Utilities"

Jaroslav Borovička[a] and John Stachurski[b]

[a] New York University, Federal Reserve Bank of Minneapolis and NBER

[b] Research School of Economics, Australian National University

December 30, 2019

1. The Bansal–Yaron Model with Time-Varying Volatility

In the main paper we study 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 study 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

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

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 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

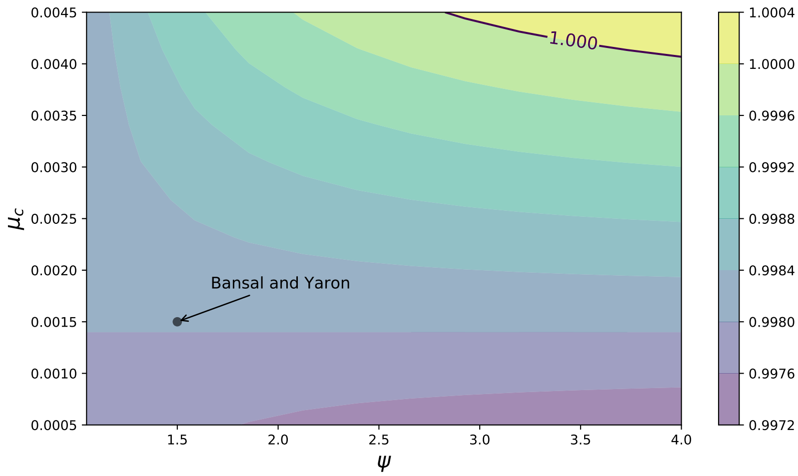

FIGURE 1. **Changes in test value Λ, ψ versus μ_c.**

double-precision floating-point number. With $n = m = 1{,}000$, we obtain a value of approximately 0.998.[1] Figure 1 reinforces the finding of stability by showing estimated values of Λ over 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 ψ and the mean consumption growth rate μ_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 to obtain a parameterization at which Λ exceeds unity.[2]

2. Unboundedness Result

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

Theorem IA.1. *If Assumptions 1 and 3 hold, then the following statements are equivalent:*

[1]Generating the statistic 1,000 times produced a mean value of 0.998128 with standard deviation 0.000053. We also implemented a discretized verion, 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 σ_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 $\Lambda = 0.998044$.

[2]At each parameterization, the value Λ was calculated by the non-discretized Monte Carlo method, with $m = n = 1000$.

(a) $\Lambda < 1$.

(b) A *has a fixed point in* $\mathscr{C}$.

(c) *There exists a* $g \in \mathscr{C}$ *such that* $\{A^n g\}_{n \geqslant 1}$ *converges to an element of* $\mathscr{C}$.

Throughout the following, the state space $\mathbb{X}$ is allowed to be any σ-compact metric space. The collection of Borel-measurable functions g from $\mathbb{X}$ to $\mathbb{R}$ such that $\|g\| := \int |g| \mathrm{d}\pi$ 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 π-almost everywhere. Other definitions are as stated in the main 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 IA.2. *Let* $\{T_n\}$ *and* T *be bounded linear operators on* $L_1(\mathbb{X}, \pi)$ *such that* $0 \leqslant T_n \leqslant T_{n+1} \leqslant T$ *for all* n. *If* $T_n f \to Tf$ *in norm as* $n \to \infty$ *for each* f *in the positive cone* $\mathscr{C}$ *and* T^i *is compact for some* $i \in \mathbb{N}$, *then* $r(T_n) \uparrow r(T)$.

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

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

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

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 by assumption T^i is compact. Further, T^i is 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 IA.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* $T_m u \to Tu$ *for all* $u \in E$. *Let* $\bar{u}_m$ *be a fixed point of* T_m *for each* m *and suppose that* $\bar{u}_m \to \bar{u}$ *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 $T\bar{u} = \lim_{m\to\infty} T_m\bar{u}$, so one only need show that $T_m\bar{u} \to \bar{u}$. To this end, observe that

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

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

$$d(T_m\bar{u}, \bar{u}) \leqslant (L+1)d(\bar{u}_m, \bar{u}) \to 0.$$

Lemma IA.4. *If Assumption 3 holds, then K^i is compact for some $i \in \mathbb{N}$.*

Proof: Under Assumption 3, 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 $i = 2m$. □

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

$$K_m g(x) = \mathbb{1}\{x \in F_m\} \int_{F_m} k(x,y) g(y) \mathrm{d}y \qquad (x \in \mathbb{X}).$$

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

Lemma IA.5. *If $f \in \mathscr{C}$, then $\|K_m f - Kf\| \to 0$ as $m \to \infty$.*

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

$$\|K_m f - Kf\| \leqslant \int \int k(x,y) e_m(x,y) |f(y)| \mathrm{d}y \, \pi(\mathrm{d}x),$$

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 $\{F_m\}$. □

Given $g \colon F_m \to \mathbb{R}$, 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 $g \colon \mathbb{X} \to \mathbb{R}$, 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,

$$\bar{K}_m g(x) = \int_{F_m} k(x,y) g(y) \mathrm{d}y \qquad (x \in F_m).$$

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

$$A_m = e_m \bar{A}_m c_m \text{ on } \mathscr{C}, \tag{1}$$

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

Lemma IA.6. *If g in $\mathscr{C}_m$ is a fixed point of $\bar{A}_m$, then $e_m g$ is a fixed point of A_m.*

Proof: For such $g \in \mathscr{C}_m$, we have $A_m e_m g = e_m \bar{A}_m c_m e_m g = e_m \bar{A}_m g = e_m g$. □

Lemma IA.7. *We have $\|\bar{K}_m\| = \|K_m\|$ for all $m \in \mathbb{N}$.*

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

$$\|\bar{f}\| = \int |\bar{f}| \bar{\pi}(x) \mathrm{d}x \leqslant \|f\| \leqslant 1.$$

We have

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

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

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

$$\|f\| = \int |f| \pi(x) \mathrm{d}x = \int |\bar{f}| \bar{\pi}(x) \mathrm{d}x = \|\bar{f}\| \leqslant 1.$$

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

Lemma IA.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 $m \geqslant M$.*

Proof: In view of Lemma IA.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 $r(K_m) \to r(K)$, which by Lemma IA.2 will hold if (a) K^i is compact for some $i \in \mathbb{N}$, (b) $0 \leqslant K_m \leqslant K_{m+1} \leqslant K$ for all m, and (c) $K_m f \to K f$ in norm for each f in $L_p(\mathbb{X}, \pi)_+$. We already have (a) by Lemma IA.4 and (b) is true by construction. Finally, (c) holds by Lemma IA.5. □

Proposition IA.9. *Under the conditions of Theorem IA.1, Λ is well defined and satisfies $\Lambda = \beta\, r(K)^{1/\theta}$.*

The proof is identical to that of Proposition C.4 in the paper itself.

Lemma IA.10. *If $\theta < 0$ and $\Lambda < 1$, then there exists an $M \in \mathbb{N}$ such that, for all $m \geqslant M$, the operator A_m has a nonzero fixed point g_m in $\mathscr{C}$, and $g_m \leqslant g_{m+1}$ for all such m.*

Proof: If $\theta < 0$ and $\Lambda < 1$, then, by Proposition IA.9, we have $r(K) > 1/\beta^\theta$. Now let M be as in Lemma IA.8 and take $m \geqslant M$. Observe that $\bar{A}_m$ has a unique nonzero fixed point $\bar{g}_m$ in $\mathscr{C}_m$, since F_m is compact (as follows from our results in the main 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 $g_m \leqslant g_{m+1}$ for all $m \geqslant M$. To see this, pick any such m and observe that, since $K_m \leqslant K_{m+1}$ on $\mathscr{C}$ and φ is increasing, we have $A_{m+1} g_m \geqslant A_m g_m = g_m$. Using isotonicity of A_{m+1} and iterating now gives $A_{m+1}^n g_m \geqslant g_m$ for all n. Moreover, since g_m is nonzero on F_m and hence on F_{m+1}, the convergence result in Theorem 2.1 from the main text applied on the compact state space F_{m+1} implies that $A_{m+1}^n g_m \to g_{m+1}$ uniformly. Hence, $g_{m+1} \geqslant g_m$, as was to be shown. $\square$

Lemma IA.11. *If $\theta < 0$, then the family $\{A_n\}$ is uniformly Lipschitz-continuous on $\mathscr{C}$.*

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

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

$$\therefore \quad \|A_m f - A_m g\| \leqslant \|K\| \cdot \|f - g\|.$$

Proof of Theorem IA.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 paper) does not require compactness of $\mathbb{X}$. Hence, below we focus on the case $\theta < 0$. Our proof uses Lemma IA.3, with the metric space $(\mathscr{C}, \|\cdot\|)$ and maps $\{A_m\}$ and A.

First observe that, by Lemma IA.5, for each given $f \in \mathscr{C}$, we have $K_m f \to K f$ as $m \to \infty$. Since φ is Lipschitz of order one when $\theta < 0$, it follows immediately that $A_m f \to A f$ as $m \to \infty$. By Lemma IA.10, there exists an $M \in \mathbb{N}$ such that, for all $m \geqslant M$, the operator A_m has a nonzero fixed point g_m in $\mathscr{C}$, and $g_m \leqslant g_{m+1}$ for all such m. Since φ is bounded above by $b := (1 - \beta)^\theta$ when $\theta < 0$, it must be the case that $g_m \leqslant g_{m+1} \leqslant b$ 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 IA.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 φ, while uniform Lipschitz continuity of $\{A_m\}$ follows from Lemma IA.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 C.9 in the main text. There, compactness of $\mathbb{X}$ was used only 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 IA.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 $\mathscr{C}$, since the limit $g^* := \lim_{n\to\infty} A^n g$ is then a fixed point of A. $\square$

3. Valuation Under IID Growth

In this section we investigate an interpretation of Λ related to the 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 in which $\ln(C_{t+1}/C_t) = \mu + \sigma W_{t+1}$ for some IID sequence $\{W_t\}$. The risk-adjusted growth rate of consumption is then given by

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

and hence

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

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

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

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

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

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

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

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 3.3 of the paper. The agent starts with a prior $P(X_0 = 1) = \overline{X}_0$ and observes data $Z_t = \ln(C_t/C_{t-1})$, $t = 1, 2, \ldots$. Denote the density of the data conditional on the unobserved state by

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

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

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

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

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

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

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 $z \leqslant z^*$ and $h_x^2(z) = h(x,z)$ for $z \geqslant z^*$, and then write the distribution function as $\overline{Q}(x,y) = 1$ for $y \geqslant h(x,z^*)$ and

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

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

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

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 $\overline{X}_t = x$ 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 $\overline{Q}(x,y)$ of the transition density.

The transition density $\overline{q}(x,y)$ in the benchmark specification of the model does not satisfy the assumptions imposed in Section 2 of the main text. First, the transition density

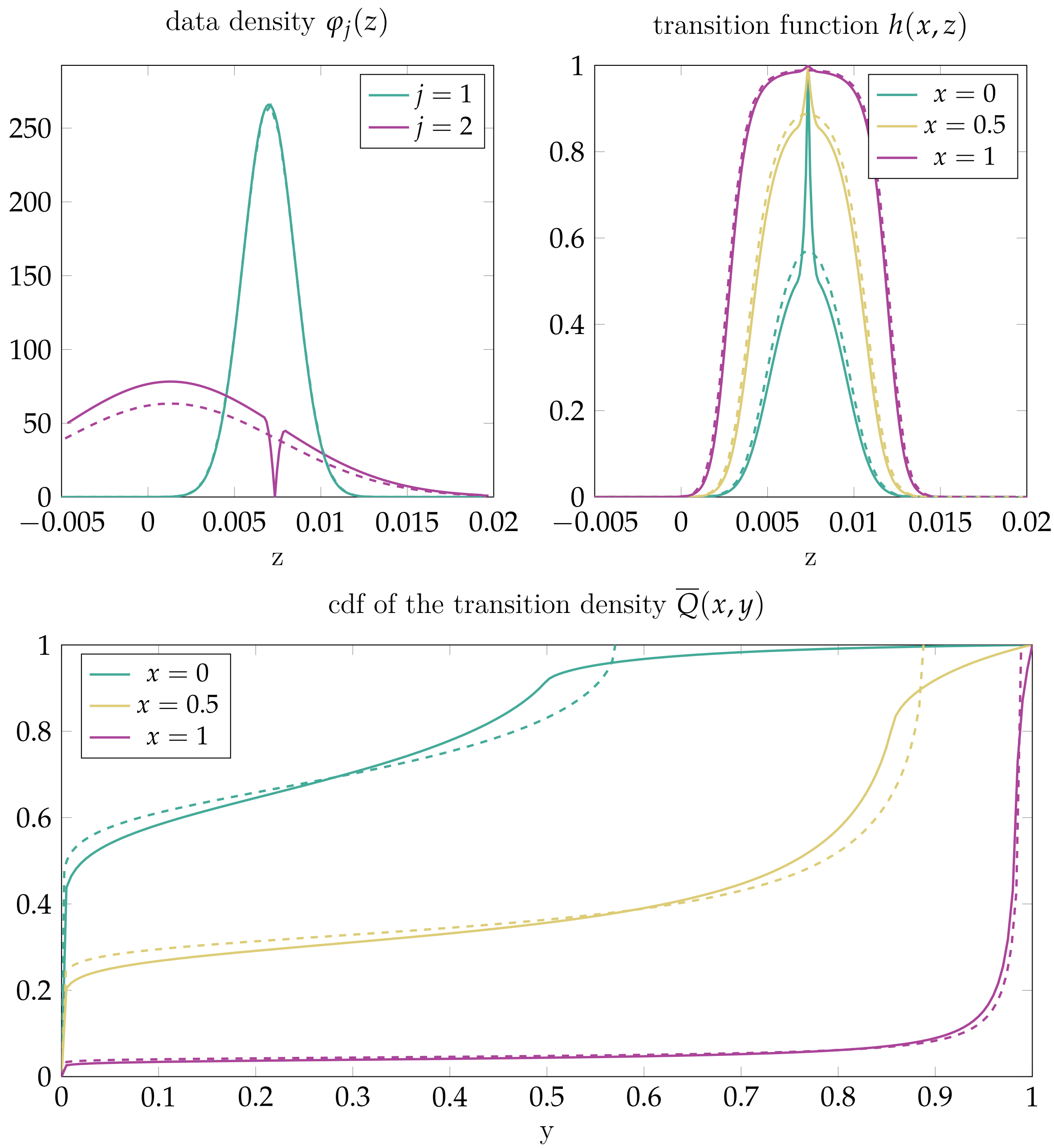

FIGURE 2. **Impact of the perturbation of the learning model.** Dashed lines represent the original densities and transition functions, solid lines represent their perturbed counterparts.

$\overline{q}\,(x,y)$ is zero for any $y > h\,(x,z^*)$, and hence the l-step transition density $\overline{q}^l$ 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, $\overline{q}(x, y)$ diverges to infinity as $y \searrow 0$. 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 $\overline{q}(x, y)$ bounded as $y \searrow 0$.

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

(a) the original full-information problem introduced in Section 3.2 of the paper is unchanged;

(b) the perturbed data densities $\widetilde{\varphi}_j(z)$ and the implied perturbed transition distribution $\widetilde{Q}(x, \cdot)$ are arbitrarily close to the original objects $\varphi_j(z)$ and $\overline{Q}(x, y)$, in the sense that for any $\varepsilon > 0$, we can construct a perturbation $\widetilde{\varphi}_j(z)$ such that

$$\begin{aligned} \sup_{A \subseteq R} \int_A \left| \varphi_j(z) - \widetilde{\varphi}_j(z) \right| dz &< \varepsilon \\ \sup_{x \in [0,1], B \subseteq [0,1]} \left| \overline{Q}(x, B) - \widetilde{Q}(x, B) \right| &< \varepsilon. \end{aligned}$$

The first statement holds when the densities $\widetilde{\varphi}_j(z)$ imply the same moments $\xi(j)$ in equation (27) of the main text as in the original specification. The second statement implies that the operator $\overline{K}$ in (27) is approximated to an arbitrary degree of accuracy for the class of bounded functions. Details on the construction of $\widetilde{\varphi}_j(z)$ 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. *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. H. (1974). *Banach Lattices and Positive Operators*. Springer-Verlag Berlin Heidelberg.

Schep, A. R. (1980). Positive diagonal and triangular operators. *Journal of Operator Theory*, 3(2):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.