

# Internet Appendix to “Long Run Risk and the Persistence of Consumption Shocks”\*

## A. Roadmap

This Internet Appendix contains additional material of technical nature that is relevant to the published article. The material is organized as follows. Section B relates the persistence properties of the series  $g_t^{(j)}$  and  $\pi_t^{(j)}$  to their Fourier spectra. Section C details additional results and robustness checks for our test to detect small but persistent components. Section D provides the derivations of our long-run risk model with persistence heterogeneity. Section E presents additional evidence on the consumption and price-dividend series. Section F compares our decomposition with alternative filtering techniques which have been proposed to analyze long-run comovements between economic and financial variables.

## B. A Frequency Interpretation of the Persistence-based Decomposition

The filtering procedure described in **Section “Decomposing time series along the persistence dimension”** and the persistence properties of the series  $\pi_t^{(j)}$  and  $g_t^{(j)}$  can be usefully visualized in the frequency domain in terms of their Fourier spectra.<sup>1</sup> Recall that given a time series  $\{g_t\}_{t \in \mathbb{Z}}$  we consider sample means over windows of past observations with size  $2^j$ :

$$\pi_t^{(j)} = \frac{1}{2^j} \sum_{p=0}^{2^j-1} g_{t-p} \quad (\text{IA.1})$$

where  $j \geq 1$  and  $\pi_t^{(0)} \equiv g_t$ , and then to define the components  $g_t^{(j)}$  we compute the difference between sample means of sizes  $2^{j-1}$  and  $2^j$ , i.e.

$$g_t^{(j)} = \pi_t^{(j-1)} - \pi_t^{(j)} \quad (\text{IA.2})$$

As an example we extract the components of the consumption growth time series using equations (IA.2) and (IA.1); we then display the spectrum of these components in Figure IA.1. The left and right columns report the case where the decomposition stops at  $J = 1$  and the case where it stops at  $J = 2$ , respectively.

The top subplot of Figure IA.1 shows the Fourier spectrum of the aggregate consumption growth time series. The shadowed region in the bottom left panel identifies the part of the spectrum which survives after the first application of the moving average filter, namely the spectrum of  $\pi_t^{(1)}$ . We clearly see that the effect of the simple 2-period moving average is to halve the spectrum and to keep the lowest part. Nevertheless the high frequency part of the spectrum is recovered by the component  $g_t^{(1)}$ . This can be seen in the mid left panel

---

\*Citation format: Fulvio Ortù, Andrea Tamoni and Claudio Tebaldi, [2013], Internet Appendix to “Long Run Risk and the Persistence of Consumption Shocks”

<sup>1</sup>Indeed in the frequency representation, a  $2^j$ -period moving average operator works as a low band pass filter which removes all those components whose frequency is higher than  $2^j$ . Table IA.2 gives an interpretation of the cycle durations corresponding to the persistence level  $j$  in the case of quarterly time series.

where the shadowed region represent the spectrum of  $g_t^{(1)}$ . In some sense we are reassured that, for  $J = 1$  we recover a simple permanent-transitory decomposition.

[Insert Figure IA.1 about here.]

The right column of Figure IA.1 shows the case where we set  $J = 2$  in our decomposition. We thus filter out the first two components of the aggregate consumption growth,  $g_t^{(1)}$  and  $g_t^{(2)}$ . From (IA.2) we know that  $g_t^{(2)}$  is obtained as the difference of  $\pi_t^{(2)}$  and  $\pi_t^{(1)}$  reported in the bottom right and left panels, respectively. The third left panel displays exactly this operation and confirms our previous intuition that by taking the difference between  $\pi_t^{(2)}$  and  $\pi_t^{(1)}$ , we are able to identify the fluctuations of the original time series  $g_t$  that lies in the well defined frequency range  $[1/4, 1/2)$ .<sup>2</sup> The bottom panels of Figure IA.1 show that the Fourier spectrum of the time series  $\pi_t^{(2)}$  differs from that of  $\pi_t^{(1)}$  because the dyadic averaging operation gets rid of the components at frequencies larger than  $1/2^2$  while the low frequency components are essentially left unaffected. Therefore we conclude that increasing the window of values over which the average is made is equivalent to focus one's attention on lower and lower frequencies.

### C. Gencay - Signori test: the DWT Haar Filter case

This section presents some theoretical results related to our test statistics, and discusses additional Monte Carlo simulations. Recall that our test is based on the variance decomposition induced by the multiresolution approach. In fact if we decompose the (zero mean) time series  $g_t$  using the Haar matrix  $\mathcal{T}^{(J)}$

$$\mathcal{T}^{(J)} X_t^{(J)} = \begin{pmatrix} \pi_T^{(J)} \\ g_T^{(J)} \\ g_T^{(J-1)} \\ g_T^{(J-1)} \\ g_{T/2}^{(J-2)} \\ g_T^{(J-2)} \\ g_{3T/4}^{(J-2)} \\ g_{T/2}^{(J-2)} \\ g_{T/4}^{(J-2)} \\ \vdots \\ g_T^{(1)} \\ g_{T-2}^{(1)} \\ \vdots \\ g_2^{(1)} \end{pmatrix} \quad (\text{IA.3})$$

---

<sup>2</sup> Each component has in fact a corresponding Fourier spectrum localized in the finite interval of frequencies  $(\frac{f_{max}}{2^j}, \frac{f_{max}}{2^{j-1}}]$  where  $f_{max}$  is the maximum frequency of observations, and in our case quarterly.

where

$$X_T^{(J)} = \begin{pmatrix} g_T \\ g_{T-1} \\ g_{T-2} \\ \vdots \\ g_1 \end{pmatrix}$$

Letting now  $\mathbf{g}^{(j)} = \{g_{2^j}^{(j)}, \dots, g_{k \cdot 2^j}^{(j)}, \dots, g_T^{(j)}\}$ , it is easy to see that the sample variance can be computed as

$$\begin{aligned} \frac{(X_t^{(J)})^\top X_t^{(J)}}{T} &= \frac{(\mathcal{T}^{(J)} X_t^{(J)})^\top (\mathcal{T}^{(J)} X_t^{(J)})}{T} \\ &= \frac{\sum_{j=1}^J (\mathbf{g}^{(j)})^\top \mathbf{g}^{(j)}}{T} \\ &= \frac{1}{T} \left( \|\mathbf{g}^{(1)}\|^2 + \dots + \|\mathbf{g}^{(J)}\|^2 \right) \end{aligned}$$

Consider

$$\begin{aligned} W_{t,2} &= \frac{1}{2}(g_{4t} + g_{4t-1} - g_{4t-2} - g_{4t-3}), \quad t = 1, \dots, T/4 \\ V_{t,2} &= \frac{1}{2}(g_{4t} + g_{4t-1} + g_{4t-2} + g_{4t-3}), \quad t = 1, \dots, T/4 \end{aligned}$$

and the test statistic:

$$\hat{G}_{T,2} = \frac{\sum_{t=1}^{T/4} W_{t,2}^2}{\sum_{t=1}^{T/2} W_{t,1}^2 + \sum_{t=1}^{T/4} W_{t,2}^2 + \sum_{t=1}^{T/4} V_{t,2}^2} \quad (\text{IA.4})$$

Heuristically  $\hat{G}_{T,2}$  should be close to 1/4 under  $H_0$ .

Lemma 1: *under  $H_0$ ,  $\hat{G}_{T,2} = 1/4 + o_p(1)$ .*

*Proof.* First note that

$$W_{t,2}^2 = \frac{1}{4}(g_{4t}^2 + g_{4t-1}^2 + g_{4t-2}^2 + g_{4t-3}^2 + 2g_{4t}g_{4t-1} - 2g_{4t}g_{4t-2} - 2g_{4t}g_{4t-3} + 2g_{4t-2}g_{4t-3} - 2g_{4t-1}g_{4t-2} - 2g_{4t-2}g_{4t-3}) \quad (\text{IA.5})$$

Using equation (IA.5), together with (IA.4), we obtain the following under  $H_0$

$$\begin{aligned}
\hat{G}_{T,2} &= \frac{\sum_{t=1}^{T/4} W_{t,2}^2}{\sum_{t=1}^{T/2} W_{t,1}^2 + \sum_{t=1}^{T/4} W_{t,2}^2 + \sum_{t=1}^{T/4} V_{t,2}^2} \\
&= \frac{\frac{1}{4} \sum_{t=1}^{T/4} (g_{4t}^2 + g_{4t-1}^2 + g_{4t-2}^2 + g_{4t-3}^2)}{\sum_{t=1}^{T/4} (g_{4t}^2 + g_{4t-1}^2 + g_{4t-2}^2 + g_{4t-3}^2)} + \\
&\quad \frac{\frac{1}{4} \sum_{t=1}^{T/4} (2g_{4t}g_{4t-1} - 2g_{4t}g_{4t-2} - 2g_{4t}g_{4t-3} + 2g_{4t-2}g_{4t-3} - 2g_{4t-1}g_{4t-2} - 2g_{4t-2}g_{4t-3})}{\sum_{t=1}^{T/4} (g_{4t}^2 + g_{4t-1}^2 + g_{4t-2}^2 + g_{4t-3}^2)} \\
&= \frac{1}{4} + \frac{o_p(T)}{O_p(T)}
\end{aligned}$$

□

Theorem C.1: under  $H_0$ ,  $\sqrt{2^j * (2^{2(j-1)}) * \frac{T}{\text{binom}(2^j, 2)}} \left( \hat{G}_{T,j} - \frac{1}{2^j} \right) \Rightarrow N(0, 1)$ .<sup>3</sup>

*Proof.* We show the proof for the case  $j = 2$ . Note that

$$\begin{aligned}
\hat{G}_{T,2} - 1/4 &= \frac{\frac{1}{2} \sum_{t=1}^{T/4} (g_{4t}g_{4t-1} - g_{4t}g_{4t-2} - g_{4t}g_{4t-3} + g_{4t-2}g_{4t-3} - g_{4t-1}g_{4t-2} - g_{4t-2}g_{4t-3})}{\sum_{t=1}^{T/4} (g_{4t}^2 + g_{4t-1}^2 + g_{4t-2}^2 + g_{4t-3}^2)} \\
&= \frac{\frac{1}{2} * \sqrt{T/4} * N(0, \sigma^4) * \sqrt{\binom{4}{2}}}{T * \sigma^2} = \frac{N(0, 1)}{\sqrt{\frac{4 * T * 2^2}{\binom{4}{2}}}}
\end{aligned}$$

since we have that the term

$$\begin{aligned}
&\sqrt{T/4} * \sqrt{T/4} * \frac{\sum_{t=1}^{T/4} (g_{4t}g_{4t-1})}{T/4} \Rightarrow \sqrt{T/4} * N(0, \sigma^4) \\
&T * \frac{\sum_{t=1}^{T/4} (g_{4t}^2 + g_{4t-1}^2 + g_{4t-2}^2 + g_{4t-3}^2)}{T} \Rightarrow T * \sigma^2
\end{aligned}$$

□

### I. Monte Carlo simulations

This subsection reports additional Monte Carlo simulations which complement the one in **Section “Detecting small but persistent components”** in the main text. We run the following two experiments:

- $T = 256$ ,  $J^* = 6$ ,  $\rho_{J^*} = 0.2$  or  $\rho_{J^*} = 0.4$ . We choose the variance of  $\epsilon_t^{(J^*)}$  so that the component explains 3%, 5%, 7% of total variance.
- $T = 128$ ,  $J^* = 4$ ,  $\rho_{J^*} = 0.2$  or  $\rho_{J^*} = 0.4$ . We choose the variance of  $\epsilon_t^{(J^*)}$  so that the component explains 3%, 5%, 7% of total variance.

---

<sup>3</sup> $2^{2(j-1)}$  is obtained from  $\frac{1}{(2 * (\frac{1}{\sqrt{2^j}})^2)^2}$

In every experiment we check that the original time series is indistinguishable from a white noise series (i.e. we check that the first order autocorrelation is not significant and that the K-S test cannot reject the null hypothesis that the values come from a  $N(0,1)$ ). We report the results in Table IA.1.

[Insert Table IA.1 about here.]

#### D. The Long-run Risk Model with Persistence Heterogeneity

In this section we derive formally the results presented in **Section “Persistence Heterogeneity, Predictability and Long-run Valuation”** of the main text. We first show how to solve for the coefficients  $A_{0,j}, A_j, A_{0,j}^m, A_j^m$  of the financial ratios in terms of the parameters of the model, we then compute the pricing kernel and the equity premia on both the consumption claim asset and the market return, and finally we derive the risk-free rate.

We start by recalling that in our endowment economy populated by a representative agent with recursive preferences à-la Epstein and Zin (1989, 1991) the log consumption growth series  $g_t$  satisfies two requirements. First,  $g_t$  is a weakly stationary process with autocovariances decaying to zero as the lag increases. Second, its decimated components  $g_t^{(j)}$  (see Equation (2) in the main text) follow a multiscale autoregressive process on the time domain defined by decimation, i.e.

$$g_{t+2^j}^{(j)} = \rho_j g_t^{(j)} + \varepsilon_{t+2^j}^{(j)} \quad , \quad t = k2^j, k \in \mathbb{Z} \quad (\text{IA.6})$$

with the shocks possibly correlated across levels of persistence (for fixed time  $t$ ) but not across time (for fixed persistence level  $j$ ). We observe in particular that one can rewrite these dynamics as

$$g_{(k+1) \cdot 2^j}^{(j)} = \rho_j g_{k \cdot 2^j}^{(j)} + \varepsilon_{(k+1) \cdot 2^j}^{(j)} \quad , \quad k \in \mathbb{Z}$$

which highlights the fact that the *decimated* components are autoregressive process of order one up to the scale  $2^j$  necessary to avoid spurious correlation. We also remark that the translation invariance property of decimation, i.e. the fact that the matrix  $\mathcal{T}^{(j)}$  is independent from  $h$  (see footnote 6 in the main text), together with the assumption of weak stationarity of the series  $g_t$ , implies that  $\forall h = 0, \dots, 2^{j-1}$

$$g_{h+(k+1) \cdot 2^j}^{(j)} = \rho_j g_{h+k \cdot 2^j}^{(j)} + \varepsilon_{h+(k+1) \cdot 2^j}^{(j)} \quad (\text{IA.7})$$

i.e. the dynamics in (IA.6) are translation invariant. Putting now (IA.6) together with (IA.7) implies that

$$g_{t+2^j}^{(j)} = \rho_j g_t^{(j)} + \varepsilon_{t+2^j}^{(j)} \quad , \quad t \in \mathbb{Z} \quad (\text{IA.8})$$

a fact that proves useful to solve the Euler equation of our model.

To account for the possibility that unhedgeable income risk draws a wedge between consumption and dividends, we allow for the presence of a leverage effect possibly different across levels of persistence. Specifically we assume that the weakly stationary log dividend growth series  $gd_t$  has vanishing autocovariances and decimated components  $\{gd_t^{(j)}\}_{j=1}^J$  related to the consumption growth components through a persistence dependent leverage parameter  $\phi_j$ , i.e.

$$gd_{t+2^j}^{(j)} = \phi_j g_t^{(j)} + \eta_{t+2^j}^{(j)} \quad , \quad t \in \mathbb{Z} \quad (\text{IA.9})$$

with the shocks  $\eta_{t+2j}^{(j)} \sim N(0, \sigma_j^2)$  uncorrelated both across time (for fixed persistence level  $j$ ) and across levels of persistence (for fixed time  $t$ ) and independent from the shocks  $\varepsilon_{t+2j}^{(j)}$  to the consumption growth components.

### I. The Financial Ratios

We solve first for the price-consumption coefficients  $A_j$  and then for the price-dividend coefficients  $A_j^m$ . To this end recall that the Euler equation for the representative agent is:

$$E_t \left[ e^{m_{t+1} + r_{t+1}^i} \right] = 1 \quad (\text{IA.10})$$

where  $m_{t+1}$  is the log stochastic discount factor given by

$$m_{t+1} = \theta \log \beta - \frac{\theta}{\psi} g_{t+1} + (\theta - 1) r_{t+1}^a, \quad (\text{IA.11})$$

$r_{t+1}^a$  is the log return on the claim which distributes a dividend equals to aggregate consumption and  $r_{t+1}^i$  is the log return on any asset  $i$ . The parameter  $\beta$  is the preference discount factor. The preference parameter  $\psi$  measures the intertemporal elasticity of substitution,  $\gamma$  measures risk aversion and  $\theta = (1 - \gamma) / (1 - 1/\psi)$ .

To determine the coefficients  $A_{0,j}, A_j$  we compute (IA.10) for  $r_{i,t+1} = r_{a,t+1}$  to obtain

$$E_t \left[ \exp \left( \theta \log \beta - \frac{\theta}{\psi} g_{t+1} + \theta r_{a,t+1} \right) \right] = 1$$

We then log-linearize returns à-la Campbell and Shiller (1988), i.e.

$$\begin{aligned} r_{a,t+1} &= \kappa_0 + \kappa_1 z_{a,t+1} - z_{a,t} + g_{t+1} \\ r_{m,t+1} &= \kappa_{0,m} + \kappa_{1,m} z_{m,t+1} - z_{m,t} + g_{t+1} \end{aligned} \quad (\text{IA.12})$$

and we substitute the log-linearized returns into the Euler equation to obtain:

$$E_t \left[ \exp \left( \theta \log \beta - \frac{\theta}{\psi} g_{t+1} + \theta \underbrace{(\kappa_0 + \kappa_1 z_{a,t+1} - z_{a,t})}_{r_{a,t+1}} + g_{t+1} \right) \right] = 1 \quad (\text{IA.13})$$

The strategy now is to express log consumption growth  $g_{t+1}$  and the price-consumption ratios  $z_{a,t}, z_{a,t+1}$  in terms of the time  $t$  consumption components  $g_t^{(j)}$  and of the innovations  $\varepsilon_{t+2j}^{(j)}$ . To this end, we first apply the redundant decomposition (4) in the main text to  $z_{a,t}$  to obtain

$$z_{a,t} = \sum_{j=1}^J z_{a,t}^{(j)} + \pi_{a,t}^{(J)}$$

Observe now that (2) in the main text implies  $\pi_t^{(j-1)} - \pi_t^{(j)} = \pi_t^{(j)} - \pi_{t-2j-1}^{(j-1)}$ , which together with (3) in the main text one has  $g_t^{(j)} = \pi_t^{(j)} - \pi_{t-2j-1}^{(j-1)}$ , from which after some tedious but

otherwise straightforward algebra one obtains

$$g_{t+1} = - \sum_{j=1}^J g_{t+2j}^{(j)} + \pi_{t+2J}^{(J)} \quad (\text{IA.14})$$

and, applying the same logic to the price-consumption ratio

$$z_{a,t+1} = - \sum_{j=1}^J z_{a,t+2j}^{(j)} + \pi_{a,t+2J}^{(J)}$$

We denote now by  $\pi_g, \pi_a$  the mean consumption growth and price-consumption ratio (with  $\pi_a$  to be determined in equilibrium) respectively, and recalling that by definition  $\pi_t^{(J)}, \pi_{a,t}^{(J)}$  are sample means of past realizations of consumption growth and of the price-consumption ratio, we assume that there exists  $J$  large enough so that  $\pi_t^{(J)} \simeq \pi_g$  and  $\pi_{a,t}^{(J)} \simeq \pi_a$  for all  $t$ .<sup>4</sup> Under this assumption the above decompositions become

$$z_{a,t} \simeq \sum_{j=1}^J z_{a,t}^{(j)} + \pi_a \quad (\text{IA.15})$$

$$z_{a,t+1} \simeq - \sum_{j=1}^J z_{a,t+2j}^{(j)} + \pi_a \quad (\text{IA.16})$$

$$g_{t+1} \simeq - \sum_{j=1}^J g_{t+2j}^{(j)} + \pi_g \quad (\text{IA.17})$$

Plugging these expressions into (IA.13) yields:

$$E_t \left[ \exp \left( \theta \log \beta - \frac{\theta}{\psi} \left( \sum_{j=1}^J (-g_{t+2j}^{(j)}) + \pi_g \right) + \theta \left( \kappa_0 + \kappa_1 \left( \sum_{j=1}^J (-z_{a,t+2j}^{(j)}) + \pi_a \right) - \left( \sum_{j=1}^J z_{a,t}^{(j)} + \pi_a \right) + \sum_{j=1}^J (g_{t+2j}^{(j)}) + \pi_g \right) \right) \right] = 1$$

Exploiting now our postulated relation between the components of the price-consumption ratio and those of consumption growth, i.e.

$$\begin{aligned} z_{a,t}^{(j)} &= A_{0,j} + A_j g_t^{(j)} \\ z_{m,t}^{(j)} &= A_{0,j}^m + A_j^m g_t^{(j)} \end{aligned} \quad (\text{IA.18})$$

---

<sup>4</sup>Alternatively one can assume that there exist a maximum level of persistence  $J$  such that  $\pi_t^{(J)} \sim WN(0, \sigma)$  and impose  $\pi_{a,t}^{(J)} = A_{0,J} + A_J \pi_t^{(J)}$ , and then solve for  $A_J$  and  $\pi_{a,t}^{(J)}$  in equilibrium.

plugging it into the Euler equation and rearranging terms we obtain:

$$E_t \left[ \exp \left( \theta \log \beta + \theta \left( 1 - \frac{1}{\psi} \right) \left( \sum_{j=1}^J (-g_{t+2j}^{(j)}) + \pi_g \right) + \theta \left( \kappa_0 + (\kappa_1 - 1)\pi_a - \kappa_1 \left( \sum_{j=1}^J A_{0,j} + \sum_{j=1}^J A_j g_{t+2j}^{(j)} \right) - \left( \sum_{j=1}^J A_{0,j} + \sum_{j=1}^J A_j g_t^{(j)} \right) \right) \right) \right] = 1$$

Sum up now the first equation in (IA.18) over  $j$  to obtain  $\sum_{j=1}^J z_{a,t}^{(j)} = \sum_{j=1}^J A_{0,j} + \sum_{j=1}^J A_j g_t^{(j)}$ , take (unconditional) expectations on both sides and exploit the fact that  $\sum_{j=1}^J z_{a,t}^{(j)}$  has zero mean due to (IA.15), and that  $g_t^{(j)}$  has zero mean by the dynamics assumed in (IA.8), to see that as a consequence  $\sum_{j=1}^J A_{0,j} = 0$ . Using this fact and the dynamics in (IA.8) the first order condition can be rewritten as follows:

$$\begin{aligned} & E_t \left[ \exp \left( \theta (\log \beta + \kappa_0 + (\kappa_1 - 1)\pi_a) + \theta \left( 1 - \frac{1}{\psi} \right) \pi_g + \theta \left( 1 - \frac{1}{\psi} \right) \left( \sum_{j=1}^J (-g_{t+2j}^{(j)}) \right) + \theta \left( \kappa_1 \sum_{j=1}^J A_j (-g_{t+2j}^{(j)}) - \sum_{j=1}^J A_j g_t^{(j)} \right) \right) \right] \\ &= E_t \left[ \exp \left( \theta (\log \beta + \kappa_0 + (\kappa_1 - 1)\pi_a) + \theta \left( 1 - \frac{1}{\psi} \right) \pi_g + \theta \left( 1 - \frac{1}{\psi} \right) \left( \sum_{j=1}^J \underbrace{(e_j M \tilde{g}_t + e_j \varepsilon_{t+1})}_{-g_{t+2j}^{(j)}} \right) + \theta \left( \kappa_1 \sum_{j=1}^J \underbrace{A_j (e_j M \tilde{g}_t + e_j \varepsilon_{t+1})}_{-g_{t+2j}^{(j)}} - \sum_{j=1}^J A_j g_t^{(j)} \right) \right) \right] = 1 \end{aligned}$$

where  $e_j$  is the vector of the canonical basis,  $M$  is the  $J$ -dimensional diagonal matrix with the opposite of the persistence parameters  $\rho_j$  on the diagonal, and

$$\begin{aligned} \tilde{g}_t &\equiv [g_t^{(1)}, \dots, g_t^{(J)}]^\top \\ \varepsilon_{t+1} &\equiv [\varepsilon_{t+21}^{(1)}, \dots, \varepsilon_{t+2J}^{(J)}]^\top \end{aligned}$$

Collecting terms in  $\tilde{g}_t$  yields eventually a system of equations

$$e_j \left( \left( 1 - \frac{1}{\psi} \right) M + A_j (\kappa_1 M - \mathbb{I}_J) \right) = 0$$

for all  $j = 1, \dots, J$ . If we introduce the following column vectors

$$\underline{A} \equiv [A_1, \dots, A_J]^\top$$

the solution to these equations is given by the following vectors of sensitivities:

$$\underline{A} = \left( 1 - \frac{1}{\psi} \right) (\mathbb{I}_J - \kappa_1 M)^{-1} \underbrace{M \underline{1}}_{-(\rho_1, \dots, \rho_J)^\top}$$



To complete our derivation we need to determine the equilibrium mean consumption-price ratio  $\pi_a$ . To this end we first recall that the parameters  $\kappa_0, \kappa_1$  that result from log-linearizing the return  $r_{a,t+1}$  on the aggregate consumption claim, see (IA.12), are indeed functions of the mean consumption-price ratio  $\pi_a$ , i.e.  $\kappa_0 = \kappa_0(\pi_a)$  and  $\kappa_1 = \kappa_1(\pi_a)$ . Plugging then the solution for  $\underline{A}$  back into the Euler equation, using the properties of the Log-Normal distributions to compute the expectation of the innovations  $\varepsilon_{t+1}$  and making explicit the dependence  $\kappa_0 = \kappa_0(\pi_a)$  and  $\kappa_1 = \kappa_1(\pi_a)$  results in a (non-linear) equation in  $\pi_a$  whose solution is indeed the equilibrium mean consumption-price ratio.

To derive the expression for  $A_j^m$  we exploit once again the Euler condition

$$E_t \left[ \exp \left( \theta \log \beta - \frac{\theta}{\psi} g_{t+1} + (\theta - 1) r_{a,t+1} + r_{m,t+1} \right) \right] = 1$$

where now the asset being priced is the market return  $r_{m,t+1}$ . Following the same steps as above, and additionally using the Campbell and Shiller (1988) log-linear approximation for  $r_{m,t+1}$ , see again (IA.12) we rewrite the Euler equation as:

$$E_t \left[ \exp \left( \theta \log \beta + \theta \left( 1 - \frac{1}{\psi} \right) \left( \sum_{j=1}^J (-g_{t+2j}^{(j)}) + \pi_g \right) + \right. \right. \\ \left. \left. (\theta - 1) \left( \kappa_0 + \kappa_1 \left( \sum_{j=1}^J (-z_{a,t+2j}^{(j)}) + \pi_a \right) - \left( \sum_{j=1}^J z_{a,t}^{(j)} + \pi_a \right) \right) - \left( \sum_{j=1}^J (-g_{t+2j}^{(j)}) + \pi_g \right) + \right. \right. \\ \left. \left. \underbrace{\kappa_{0,m} + \kappa_{1,m} \left( \sum_{j=1}^J (-z_{m,t+2j}^{(j)}) + \pi_m \right) - \left( \sum_{j=1}^J z_{m,t}^{(j)} + \pi_m \right) + \left( \sum_{j=1}^J (-gd_{t+2j}^{(j)}) + \pi_{gd} \right)}_{r_{m,t+1}} \right) \right] = 1$$

where we have exploited once again the assumption that for  $J$  large enough  $\pi_{m,t}^{(J)} \simeq \pi_m$  and  $\pi_{gd,t}^{(J)} \simeq \pi_{gd}$ , with  $\pi_m$ , respectively  $\pi_{gd}$ , the mean price-dividend ratio (to be determined in equilibrium) and mean dividend growth. Let's focus now on the term

$$\theta \left( 1 - \frac{1}{\psi} \right) \left( \sum_{j=1}^J (-g_{t+2j}^{(j)}) \right) + (\theta - 1) \left( \kappa_0 + \kappa_1 \left( \sum_{j=1}^J (-z_{a,t+2j}^{(j)}) \right) - \left( \sum_{j=1}^J z_{a,t}^{(j)} \right) \right)$$

Neglecting innovations that are going to be captured by the constant using the law of log normal distribution, and neglecting constant parameters as well, this can be written as

follows

$$\begin{aligned}
&= \theta \left(1 - \frac{1}{\psi}\right) \left( \sum_{j=1}^J \underbrace{(e_j M \tilde{g}_t + e_j \varepsilon_{t+1})}_{-g_{t+2j}^{(j)}} \right) + (\theta - 1) \left( \kappa_1 \left( \sum_{j=1}^J A_j (-g_{t+2j}^{(j)}) \right) - \left( \sum_{j=1}^J A_j g_t^{(j)} \right) \right) \\
&= \theta \left(1 - \frac{1}{\psi}\right) \left( \sum_{j=1}^J e_j M \tilde{g}_t \right) + (\theta - 1) (\underline{A}^\top (\kappa_1 M - \mathbb{I}_J) \tilde{g}_t)
\end{aligned}$$

Plugging the solution for  $\underline{A}$  obtained above and after some algebra we eventually obtain

$$\begin{aligned}
&= \theta \left(1 - \frac{1}{\psi}\right) (M \underline{1})^\top \tilde{g}_t - (\theta - 1) \left(1 - \frac{1}{\psi}\right) (M \underline{1})^\top \tilde{g}_t \\
&= \left(1 - \frac{1}{\psi}\right) (M \underline{1})^\top \tilde{g}_t
\end{aligned}$$

Plugging this expression into the Euler equation, using the dynamics for the components of the log consumption growth given in formula (IA.6) and rearranging terms we have:

$$\begin{aligned}
&E_t \left[ \exp \left( \theta \log \beta + \left( \theta - \frac{\theta}{\psi} - 1 \right) \pi_g + (\theta - 1) ((\kappa_1 - 1) \pi_a + \kappa_1) + \left(1 - \frac{1}{\psi}\right) (M \underline{1})^\top \tilde{g}_t - \left( \sum_{j=1}^J (-g_{t+2j}^{(j)}) \right) \right. \right. \\
&\quad \left. \left. + \kappa_{0,m} + \kappa_{1,m} \left( \sum_{j=1}^J (-z_{m,t+2j}^{(j)}) + \pi_m \right) - \left( \sum_{j=1}^J z_{m,t}^{(j)} + \pi_m \right) + \left( \sum_{j=1}^J (-g d_{t+2j}^{(j)}) + \pi_{gd} \right) \right) \right] \\
&= E_t \left[ \exp \left( \theta \log \beta + \left( \theta - \frac{\theta}{\psi} - 1 \right) \pi_g + (\theta - 1) ((\kappa_1 - 1) \pi_a + \kappa_1) + \left(1 - \frac{1}{\psi}\right) (M \underline{1})^\top \tilde{g}_t - \sum_{j=1}^J \underbrace{e_j M \tilde{g}_t + e_j \varepsilon_{t+1}}_{-g_{t+2j}^{(j)}} \right. \right. \\
&\quad \left. \left. + \kappa_{0,m} + \kappa_{1,m} \left( \sum_{j=1}^J (-z_{m,t+2j}^{(j)}) + \pi_m \right) - \left( \sum_{j=1}^J z_{m,t}^{(j)} + \pi_m \right) + \left( \sum_{j=1}^J (-g d_{t+2j}^{(j)}) + \pi_{gd} \right) \right) \right] \\
&= E_t \left[ \exp \left( \theta \log \beta + \left( \theta - \frac{\theta}{\psi} - 1 \right) \pi_g + (\theta - 1) ((\kappa_1 - 1) \pi_a + \kappa_1) - \frac{1}{\psi} (M \underline{1})^\top \tilde{g}_t - \sum_{j=1}^J e_j \varepsilon_{t+1} \right. \right. \\
&\quad \left. \left. + \kappa_{0,m} + \kappa_{1,m} \left( \sum_{j=1}^J (-z_{m,t+2j}^{(j)}) + \pi_m \right) - \left( \sum_{j=1}^J z_{m,t}^{(j)} + \pi_m \right) + \left( \sum_{j=1}^J (-g d_{t+2j}^{(j)}) + \pi_{gd} \right) \right) \right] = 1
\end{aligned}$$

Finally, analogously to what we have done for the return on the consumption claim, using the dynamics for the components of the log dividend growth given in equation (IA.9) together with our guess for the components of log price-dividend ratio given in equation (IA.18),

rearranging terms and using the log normal properties of the shocks we obtain:

$$\begin{aligned}
& E_t \left[ \exp \left( \theta \log \beta + \left( \theta - \frac{\theta}{\psi} - 1 \right) \pi_g + (\theta - 1)(\kappa_1 - 1)\pi_a + \kappa_{0,m} + (\kappa_{1,m} - 1)\pi_m - \frac{1}{\psi} (M\mathbb{1})^\top \tilde{g}_t - \sum_{j=1}^J e_j \varepsilon_{t+1} + \right. \right. \\
& \quad \left. \left( \kappa_{1,m} \left( \sum_{j=1}^J A_j^m (-g_{t+2j}^{(j)}) \right) - \left( \sum_{j=1}^J A_j^m g_t^{(j)} \right) \right) + \left( \sum_{j=1}^J \phi_j g_t^{(j)} + \eta_{t+2j}^{(j)} \right) \right) \Big] \\
&= E_t \left[ \exp \left( \theta \log \beta + \left( \theta - \frac{\theta}{\psi} - 1 \right) \pi_g + (\theta - 1)(\kappa_1 - 1)\pi_a + \kappa_{0,m} + (\kappa_{1,m} - 1)\pi_m - \frac{1}{\psi} (M\mathbb{1})^\top \tilde{g}_t - \sum_{j=1}^J e_j \varepsilon_{t+1} + \right. \right. \\
& \quad \left. \left( \kappa_{1,m} \left( \sum_{j=1}^J \underbrace{A_j^m}_{-g_{t+2j}^{(j)}} (e_j M \tilde{g}_t + e_j \varepsilon_{t+1}) \right) \right) - \left( \sum_{j=1}^J A_j^m g_t^{(j)} \right) + \left( \sum_{j=1}^J \phi_j (e_j M \tilde{g}_t) + \eta_{t+2j}^{(j)} \right) \right) \Big] = 1
\end{aligned}$$

where we exploit once again the fact that  $\sum_{j=1}^J z_{m,t}^{(j)}$  has zero mean since  $z_{m,t}^{(j)}$  are the components of the *demeaned* price-dividend ratios, and that  $g_t^{(j)}$  has zero mean by the dynamics assumed in (IA.8), to see that as a consequence  $\sum_{j=1}^J A_{0,j}^m = 0$ .

Define

$$\begin{aligned}
\underline{A}_m &\equiv [A_1^m, \dots, A_J^m]^\top \\
\underline{\phi} &\equiv [\phi_1, \dots, \phi_J]^\top
\end{aligned} \tag{IA.19}$$

In vector notation we have

$$\begin{aligned}
\underline{A}_m (\kappa_{1,m} M - \mathbb{I}_J) &= \frac{1}{\psi} M \mathbb{1} - M \underline{\phi} \\
\underline{A}_m &= (\mathbb{I}_J - \kappa_{1,m} M)^{-1} M \left( \underline{\phi} - \frac{1}{\psi} \mathbb{1} \right)
\end{aligned}$$

To conclude, we need to compute the equilibrium mean price-dividend ratio  $\pi_m$ , given the exogenous mean consumption growth  $\pi_g$  and mean dividend growth  $\pi_{gd}$ , and given the equilibrium mean price-consumption ratio  $\pi_a$  computed above. To this end we recall now that the parameters  $\kappa_{0,m}$ ,  $\kappa_{1,m}$  that result from log-linearizing the return  $r_{m,t+1}$  on the market, see (IA.12), are indeed functions of the mean price-dividend ratio  $\pi_m$ , i.e.  $\kappa_{0,m} = \kappa_{0,m}(\pi_m)$  and  $\kappa_{1,m} = \kappa_{1,m}(\pi_m)$ . Plugging then the solution for  $\pi_a$  and  $\underline{A}_m$  back into the Euler equation, using the properties of the Log-Normal distributions to compute the expectation of the innovations  $\varepsilon_{t+1}$  and making explicit the dependence  $\kappa_{0,m} = \kappa_{0,m}(\pi_m)$  and  $\kappa_{1,m} = \kappa_{1,m}(\pi_m)$  results in a (non-linear) equation in  $\pi_m$  whose solution is indeed the equilibrium mean price-dividend ratio.

## II. The Risk Premia

The risk premium for any asset is determined by the conditional covariance between the return and the SDF, i.e.

$$E_t[r_{i,t+1} - r_{f,t}] + 0.5\sigma_{r_{i,t}}^2 = -cov_t(m_{t+1}, r_{i,t+1})$$

where we exploit the fact that asset returns and the pricing kernel in our economy are conditionally log-normal. Therefore to obtain expressions (25) and (26) in the main text, we need to compute first the innovations in the stochastic discount factor and in the returns.

The equilibrium return innovations can be found by plugging the expressions (IA.15), (IA.16) and (IA.17) into the Campbell and Shiller (1988) approximation for log returns, see equation (IA.12) to obtain

$$\begin{aligned}
r_{a,t+1} - E_t[r_{a,t+1}] &= \left( \sum_{j=1}^J g_{t+2j}^{(j)} + \pi_g \right) + \kappa_0 + \kappa_1 \left( \sum_{j=1}^J z_{a,t+2j}^{(j)} + \pi_a \right) - \left( \sum_{j=1}^J z_{a,t}^{(j)} + \pi_a \right) - E_t[r_{a,t+1}] \\
&= \sum_{j=1}^J \varepsilon_{t+2j}^{(j)} + \kappa_1 \left( \sum_{j=1}^J A_j \varepsilon_{t+2j}^{(j)} \right) \\
&= \underline{\mathbf{1}} \cdot \boldsymbol{\varepsilon}_{t+1} + \kappa_1 \underline{\mathbf{A}} \cdot \boldsymbol{\varepsilon}_{t+1}
\end{aligned} \tag{IA.20}$$

where in the second line we use our solution for the price-consumption ratio and in the third line we use the definition

$$\begin{aligned}
\boldsymbol{\varepsilon}_{t+1}^\top &\equiv \left[ \varepsilon_{t+2^1}^{(1)}, \dots, \varepsilon_{t+2^J}^{(J)} \right] \\
\boldsymbol{\varepsilon}_{t+1} &\sim N(0, \mathbf{Q})
\end{aligned}$$

Recall that the innovations  $\varepsilon_{t+2^j}^{(j)}$  driving the components of consumption growth at persistence level  $j$  reflect the impact of current uncertainty many years into the future. In analogy with the interest rate literature, they are similar to forward rates and we therefore refer to  $\boldsymbol{\varepsilon}_{t+1}$  as the term structure of risks.

Analogous steps yield the following expression for the market return innovations

$$r_{m,t+1} - E_t[r_{m,t+1}] = \eta_{t+1} + \kappa_{1,m} \underline{\mathbf{A}}_m \cdot \boldsymbol{\varepsilon}_{t+1} \tag{IA.21}$$

with  $\eta_{t+1} \equiv \sum_j \eta_{t+2^j}^{(j)}$ .

To find the innovations in the stochastic discount factor, we plug the expressions (IA.15), (IA.16) and (IA.17), together with the dynamics for the components of log consumption growth given in equation (IA.6) and our guess for the components of price-consumption

ratio solution given in equation (IA.18) into equation (IA.11) to obtain:

$$\begin{aligned}
m_{t+1} &= \theta \log \beta - \frac{\theta}{\psi} g_{t+1} + (\theta - 1) r_{a,t+1} \\
&= \theta \log \beta - \frac{\theta}{\psi} g_{t+1} + (\theta - 1) (\kappa_0 + \kappa_1 z_{a,t+1} - z_{a,t} + g_{t+1}) \\
&= \theta \log \beta + \left( \theta - \frac{\theta}{\psi} - 1 \right) \pi_g + (\theta - 1) (\kappa_1 - 1) \pi_a \\
&\quad - \frac{\theta}{\psi} \sum_{j=1}^J g_{t+2j}^{(j)} + (\theta - 1) \left( \kappa_0 + \kappa_1 \sum_{j=1}^J z_{a,t+2j}^{(j)} - \sum_{j=1}^J z_{a,t}^{(j)} + \sum_{j=1}^J g_{t+2j}^{(j)} \right) \\
&= \theta \log \beta + \left( \theta - \frac{\theta}{\psi} - 1 \right) \pi_g + (\theta - 1) (\kappa_1 - 1) \pi_a \\
&\quad - \left( 1 - \theta + \frac{\theta}{\psi} \right) \sum_{j=1}^J g_{t+2j}^{(j)} + (\theta - 1) \left( \kappa_0 + \kappa_1 \sum_{j=1}^J z_{a,t+2j}^{(j)} - \sum_{j=1}^J z_{a,t}^{(j)} \right) \\
&= \theta \log \beta + \left( \theta - \frac{\theta}{\psi} - 1 \right) \pi_g + (\theta - 1) (\kappa_1 - 1) \pi_a \\
&\quad - \left( 1 - \theta + \frac{\theta}{\psi} \right) \sum_{j=1}^J g_{t+2j}^{(j)} + (\theta - 1) \left( \kappa_0 + \kappa_1 \sum_{j=1}^J (A_{0,j} + A_j g_{t+2j}^{(j)}) - \sum_{j=1}^J (A_{0,j} + A_j g_t^{(j)}) \right)
\end{aligned}$$

Finally using the dynamics for the components of consumption growth, see (IA.6), we obtain

$$\begin{aligned}
m_{t+1} &= \theta \log(\beta) + \left( \theta - \frac{\theta}{\psi} - 1 \right) \pi_g + (\theta - 1) ((\kappa_1 - 1) \pi_a + \kappa_0) \\
&\quad - \left( 1 - \theta + \frac{\theta}{\psi} \right) \sum_{j=1}^J \rho_j g_t^{(j)} + (\theta - 1) \left( \kappa_1 \sum_{j=1}^J (A_{0,j} - A_j \rho_j g_t^{(j)}) - \sum_{j=1}^J (A_{0,j} + A_j g_t^{(j)}) \right) \\
&\quad - \left( 1 - \theta + \frac{\theta}{\psi} \right) \sum_{j=1}^J \varepsilon_{t+2j}^{(j)} + (\theta - 1) \kappa_1 \left( \sum_{j=1}^J A_j \varepsilon_{t+2j}^{(j)} \right)
\end{aligned}$$

which implies

$$\begin{aligned}
m_{t+1} - E_t[m_{t+1}] &= - \left( 1 - \theta + \frac{\theta}{\psi} \right) \sum_{j=1}^J \varepsilon_{t+2j}^{(j)} + (\theta - 1) \kappa_1 \left( \sum_{j=1}^J A_j \varepsilon_{t+2j}^{(j)} \right) \\
&= -\gamma \underline{\mathbf{1}} \cdot \boldsymbol{\varepsilon}_{t+1} - \underline{\lambda}_\varepsilon \cdot \boldsymbol{\varepsilon}_{t+1}
\end{aligned} \tag{IA.22}$$

where we use the fact that  $\left( \frac{\theta}{\psi} - \theta + 1 \right) = \gamma$  and where  $\underline{\lambda}_\varepsilon \equiv \kappa_1 (1 - \theta) \underline{\mathbf{A}}$ . Using the formula for the return on aggregate wealth (IA.20) and the innovation in the SDF (IA.22) we obtain the risk premium for the consumption claim asset,

$$E_t[r_{a,t+1} - r_{f,t}] + 0.5 \sigma_{r_{a,t}}^2 = \gamma \underline{\mathbf{1}}^\top \mathbf{Q} \underline{\mathbf{1}} + \kappa_1 \underline{\lambda}_\varepsilon^\top \mathbf{Q} \underline{\mathbf{A}} \tag{IA.23}$$

where

$$\mathbf{Q} = \mathbf{E}_t [\boldsymbol{\varepsilon}_{t+1} \boldsymbol{\varepsilon}'_{t+1}]$$

Similarly to what we have just done, using the formula for the innovations in the market return (IA.21) and in the SDF (IA.22) the premium to the market return becomes:

$$E_t[r_{m,t+1} - r_{f,t}] + 0.5\sigma_{r_{m,t}}^2 = \kappa_{1,m} \lambda_{\varepsilon}^T \mathbf{Q} \underline{A}_m \quad (\text{IA.24})$$

### III. The Risk-Free Rate and The Intertemporal Elasticity of Substitution

To obtain our expression for the risk-free rate, see eq. (15) in the main text, we start by plugging the log short-term real interest rate  $r_{f,t+1}$  for  $r_{t+1}^i$  into the Euler equation (IA.10). Then by applying the forward decomposition (IA.14) to the consumption growth and to the log returns processes at time  $t+1$  we observe that the risk-free rate between  $t$  and  $t+1$ ,  $r_{f,t+1}$  satisfies the following condition:

$$E_t \left[ \exp \left( \theta \log \delta - \left( \frac{\theta}{\psi} \right) \left( \sum_{j=1}^J (-g_{t+2j}^{(j)}) + \pi_g \right) + (\theta - 1) \sum_{j=1}^J (-r_{a,t+2j}^{(j)}) \right) \right] = \exp(-r_{f,t+1})$$

where once again  $r_{a,t+1}$  is the return on the asset that pays consumption as dividend. Taking logs on both sides and using the log normal properties of the shocks we can rewrite it as follows

$$\begin{aligned} r_{f,t+1} &= -\theta \log \beta + \frac{\theta}{\psi} E_t \left[ \sum_{j=1}^h (-g_{t+2j}^{(j)}) \right] + (1 - \theta) E_t \left[ \sum_{j=1}^J (-r_{a,t+2j}^{(j)}) \right] \\ &\quad - \frac{1}{2} \text{var}_t \left[ \frac{\theta}{\psi} \sum_{j=1}^J (-g_{t+2j}^{(j)}) + (1 - \theta) \sum_{j=1}^J (-r_{a,t+2j}^{(j)}) \right] \\ &= -\log \beta + \frac{1}{\psi} E_t \left[ \sum_{j=1}^J (-g_{t+2j}^{(j)}) \right] + \frac{(1 - \theta)}{\theta} E_t \left[ \sum_{j=1}^h (-r_{a,t+2j}^{(j)} - r_{f,t+2j}^{(j)}) \right] \\ &\quad - \frac{1}{2\theta} \text{Var}_t \left[ \frac{\theta}{\psi} \sum_{j=1}^h (-g_{t+2j}^{(j)}) + (1 - \theta) \sum_{j=1}^h (-r_{a,t+2j}^{(j)}) \right] \end{aligned} \quad (\text{IA.25})$$

where in the last line we subtract  $(1 - \theta)r_{f,t+1}$  from both sides and divide by  $\theta$ , where it is assumed that  $\theta \neq 0$ . Further to solve the above expression, note that

$$\text{Var}_t \left[ \frac{\theta}{\psi} \sum_{j=1}^J (-g_{t+2j}^{(j)}) + (1 - \theta) \sum_{j=1}^J (-r_{a,t+2j}^{(j)}) \right] = \text{Var}_t(m_{t+1})$$

Now we show that in our homoskedastic version of the long-run risk with persistence heterogeneity the variance of the stochastic discount factor  $\text{Var}_t(m_{t+1})$  is constant (not function

of time). Indeed recall from (IA.22) that we have

$$\begin{aligned} m_{t+1} - E_t[m_{t+1}] &= -\gamma \underline{\mathbf{1}} \cdot \boldsymbol{\varepsilon}_{t+1} - \kappa_1(1 - \theta) \underline{\mathbf{A}} \cdot \boldsymbol{\varepsilon}_{t+1} \\ &= -\gamma \underline{\mathbf{1}} \cdot \boldsymbol{\varepsilon}_{t+1} - \underline{\boldsymbol{\lambda}}_\varepsilon \cdot \boldsymbol{\varepsilon}_{t+1} \end{aligned}$$

from which we can compute the variance as follows

$$\text{Var}_t(m_{t+1}) = \gamma^2 \underline{\mathbf{1}}^\top \mathbf{Q} \underline{\mathbf{1}} + \kappa_1^2 (1 - \theta)^2 \underline{\mathbf{A}}^\top \mathbf{Q} \underline{\mathbf{A}}$$

Using the dynamics for log consumption growth (IA.6) and taking conditional expectation we can rewrite (IA.25) as follows:

$$r_{f,t+1} = \alpha_f - \frac{1}{\psi} \sum_{j=1}^J \rho_j g_t^{(j)} \quad (\text{IA.26})$$

where  $\alpha_f$  is meant to capture the unconditional mean of the risk-free rate and  $\psi$  is the IES. If we define the variable  $x_t = E_t[g_{t+1}] = \sum_{j=1}^J (-\rho_j) g_t^{(j)}$ , i.e. the expectation of the aggregate consumption, then equation (IA.26) can be rewritten as

$$\begin{aligned} r_{f,t+1} &= \alpha_f + \frac{1}{\psi} x_t \\ &= \alpha_f + \frac{1}{\psi} g_{t+1} + \sigma_f \eta_{t+1} \end{aligned}$$

where in the second equation we introduce measurement errors  $\eta_{t+1}$  to link the latent conditional expectation  $x_t$  to the observable consumption growth. This is the standard regression approach<sup>5</sup> originally suggested by Hansen and Singleton (1983) to estimate the elasticity of intertemporal substitution and followed by Hall (1988) and Campbell and Mankiw (1990), among many others.

#### IV. Understanding the Mechanism

To gain some further perspective on the consequences of our model for the risk premium we consider two alternative simple examples. The first is the case in which consumption follows a pure random walk. The second is instead the case where consumption growth, although it looks like a white noise process, it contains a small but persistent component (see **Section “Persistence Vs. White noise: An example”**). In the first case, since considering decimated components avoids the emergence of spurious correlation,  $\rho_j = 0$  for all  $j$  which implies  $\underline{\mathbf{A}} = \underline{\mathbf{A}}_m = \mathbf{0}$ . In this case moreover  $\mathbf{Q}$  is a diagonal matrix such that  $\underline{\mathbf{1}}^\top \mathbf{Q} \underline{\mathbf{1}} = \text{Var}(g_t)$ . Therefore the right hand side of (IA.23) boils down to

$$E_t[r_{a,t+1} - r_{f,t}] + 0.5\sigma_{r_{a,t}}^2 = \gamma \text{Var}(g_t)$$

i.e. we recover the classical result where the only contribution to the equity premium comes from the independent consumption shocks and the risk compensation for such shocks equals the risk aversion parameter  $\gamma$ . Consider now the second case in which the apparently-white-noise consumption growth hides instead a small in variance but predictable component. If one was to assume mistakenly that consumption growth is independent over time, then one

---

<sup>5</sup>More in details instrumental variables (IV) are used.

would price assets as in the previous case. However since our model is based on the filtering technique that disentangles the different degrees of persistence, the small in variance but predictable component is detected and priced and it shows up in (IA.23). Under the parametrization of our example in **Section “Detecting small but persistent components: An example”** we obtain in fact

$$E_t[r_{a,t+1} - r_{f,t}] + 0.5\sigma_{r_{a,t}}^2 = \gamma\text{Var}(g_t) + \kappa_1^2(1 - \theta) \left(1 - \frac{1}{\psi}\right)^2 \frac{\rho_J^2}{(1 + \kappa_1\rho_J)^2} 2^{-J}(1 - \rho_J^2)$$

where the second addendum is the risk compensation for the small but persistent component. Recalling that  $\kappa_1 \approx 1$  the extra term in the above expression is well approximated by

$$(1 - \theta) \left(1 - \frac{1}{\psi}\right)^2 \frac{\rho_J^2}{(1 + \rho_J)} 2^{-J}(1 - \rho_J)$$

This expression shows that the contribution to the risk premium from the small but persistent component is positive and non monotonic in the interval  $0 \leq \rho_J \leq 1$  where it reaches a local maximum. When  $-1 < \rho_J \leq 0$  instead this contribution diverges positively as  $\rho_J$  approaches minus one.<sup>6</sup> We remark that our model is different from the classical Bansal and Yaron (2004) economy even in this simple case with a single persistent component. This is not the case for two reasons. First, although the expression of the extra term in the equity premium is formally similar to the one obtained in the standard long-run risk economy, the two terms have a different economic interpretation. In the standard long-run risk economy one would in fact price the total expected consumption growth whereas here the agent prices only a specific component of the future consumption growth. Second, while the standard long-run risk framework requires the expected consumption growth to be persistent without further specifying its half-life, our model links the persistence level  $j$  of the priced component with the duration of the fluctuations captured by this component.<sup>7</sup> This implies that a low  $\rho_j$  does not necessarily mean low persistence, since it is the scale index  $j$  that controls the half-life of the process. Importantly the knowledge of the half-life of the priced component turns out to be fundamental to identify the priced components with economic drivers.

To fully appreciate the difference between our approach and the standard long-run risk framework consider the general expression for the risk premia in (IA.23) and (IA.24) when the  $\rho_j$ 's are not restricted as in the examples analyzed above, but they differ in general across different levels of persistence.

### *E. Some additional evidence on consumption and price-dividend*

In this section we validate our model assumption that there exists  $J$  large enough so that  $\pi_t^{(J)} \simeq \pi_g$  for all  $t$ . Recall now that for any given  $J$  the knowledge of the dynamics of  $g_t^{(j)}$ ,  $j = 1, \dots, J$  is in general not enough to fully characterize the time series  $g_t$  since

---

<sup>6</sup>Observe that in this simple example the additional contribution to the equity premium is equal to zero also when  $\rho_J = 1$ . This is not surprising since, under the parametrization of our example in **Section “Detecting small but persistent components: An example”**, the  $J$ -th component has zero variance when  $\rho_J = 1$  and as such bears no risk so that it commands no premium.

<sup>7</sup>In fact, recall that the component at time-scale  $j$  is defined on a time-grid whose time unit is  $2^j$  times the unit scale of the original time series of observations.



one would also need to make assumptions on the dynamics of the component  $\pi_t^{(J)}$  defined in (IA.1). Under the assumption of weak stationarity with vanishing autocovariances and recalling (IA.1), however, it is readily seen that  $\pi_t^{(J)}$  converges to the unconditional mean of the process  $g_t$  when  $J$  diverges. From a practical standpoint, therefore, as long as  $\pi_t^{(J)}$  converges fast enough compared to the sample size we do not need to model its dynamics. Figure IA.2 provides graphical evidence supporting that this is indeed the case for the series of consumption growth. The top panel plots the demeaned consumption growth together with the sum of its components, excluding  $\pi_t^{(J)}$ . The two series are close to each other with a correlation of 0.97. The bottom panel shows instead the difference between the unconditional mean of consumption growth and  $\pi_t^{(J)}$ . This difference vanishes quickly as the sample length increases.

[Insert Figure IA.2 about here.]

Next we report in Table IA.3 the contribution of the components  $g_t^{(j)}$  and  $z_{m,t}^{(j)}$  to the total variance of consumption growth and price-dividend respectively.

[Insert Table IA.3 about here.]

For the consumption growth series the contribution of the components with different degrees of persistence to the total variance decreases with the persistence level. In particular the sixth component yields about 5% of total variance. The opposite happens for the price-dividend series: the contribution to the total variance increases with the persistence and in fact the components at levels 6 and 7 account for more than half of the total variance. This evidence contributes to explain why the aggregate time series of consumption and price-dividend have a very different persistence behavior. In fact, since the great part of the variability in consumption is explained by high frequency (i.e. low  $j$ ) components and vice versa for the price-dividend ratio, then the aggregate time series of consumption growth and price-dividend resemble a white noise and a (close to a) unit root process, respectively. The fact that the highly persistent components contribute for a very small fraction to the total volatility of aggregate consumption growth explains also why the predictability that in the next section we find for certain level of persistence  $j$  disappears at the aggregate level. Indeed since at the aggregate level the long-run component of consumption at level of persistence  $j = 6$  is clearly entangled with the high frequency noise produced by low persistent components then the comovements highlighted in Figure IA.3 between consumption growth and price-dividend ratio at high levels of persistence do not emerge unless we suitably disentangle the informative low frequency components from the noisy high frequency ones.

## *F. Further Comparisons with the Literature: Persistence and Long-run Comovements*

It is interesting to compare the approach used in this paper with other techniques which have been proposed to analyze long-run comovements between economic and financial variables. Among these we will focus in particular on band-pass filtering analysis, cointegration and long-horizon regressions.

To draw a comparison with traditional band-pass filtering, it is important to recall that the conventional decompositions of the data with the band pass filter associate the business cycle with frequencies between 2 and 32 quarters (e.g. Baxter and King, 1999, and Christiano and Fitzgerald, 2003). However the persistent predictable components common to consumption and cash flows, identified by our persistence based decomposition of time series coupled with componentwise predictive regressions, occur over cycles between 32 – 64 quarters and therefore over a longer time frame than is typically considered in conventional business cycle analysis. To put it differently, conventional business cycle detrending methods tend to sweep these oscillations into the trend, thereby removing them from the analysis. To our knowledge the only exception is Comin and Gertler (2006) who avoid this problem by constructing an alternative trend/cycle decomposition which includes in the analysis all frequencies between 2 and 200 quarters. Our levels of persistence  $j = 1, \dots, 7$  span altogether a range comprised between 1 and 128 quarters and hence it is comparable to that used in Comin and Gertler (2006). However our decomposition refines theirs on multiple dimensions. In fact whereas Comin and Gertler (2006) refer to the frequencies between 2 and 32 quarters as the high-frequency component of the medium-term cycle, and frequencies between 32 and 200 quarters as the medium-frequency component, our decomposition technique further allows us to disentangle multiple driving forces with different persistence levels inside these two intervals. Moreover Comin and Gertler (2006) use a band pass filter, which is basically a two-sided moving average filter and therefore less useful to study predictability. In contrast, the filtering technique introduced in **Section “Decomposing time series along the persistence dimension”** is one-sided into the past and can be implemented in real time.

Consider now the cointegrated approach which has been used in recent empirical work, e.g. BDK (2007) and Ferson (2010), to model the persistent component in consumption and dividend processes. In particular BDK (2007) argue that the cointegrating relation between dividends and consumption is a good measure of long-run consumption risk. Our long-run risk model with heterogeneity in persistence does not rule out this possibility. In fact the presence in consumption and dividends series of seasonal patterns with a very long half-life suggests that the component at level  $j = 6$  can be interpreted as the common trend driving the cointegrating relation between consumption and dividends. However our interpretation points to a kind of cointegration different from the traditional one used in BDK (2007) and Ferson (2010), since our components are persistent but not permanent. Therefore our suggested cointegration should rely upon the notion of seasonal cointegration relationships (see Osborn (1993)), which is a generalization of the classic concept of cointegration proposed in the seminal work of Engle and Granger (1987) where we may consider cointegrating relations not only at zero frequency but also at other frequencies connected with long-run cycles (Engle et al., 1993).<sup>8</sup> In this framework the parallel common movements in the component at persistence levels 6 in the dividends and consumption variables may be therefore interpreted as a cointegration at seasonal frequencies and long-run risk may well be captured by (seasonal) cointegration relations.

With regard to long-horizon regressions, we note that our approach, which uses filtered

---

<sup>8</sup>According to the definition proposed primarily by Hylleberg et al. (1990) seasonal cointegration means the cointegration only at seasonal frequencies. However in many theoretical and practical works (Johansen and Schaumburg, 1999), the term seasonal cointegration analysis corresponds to the cointegration analysis conducted at not only seasonal but nonseasonal frequencies as well. We use this term in this more general meaning.

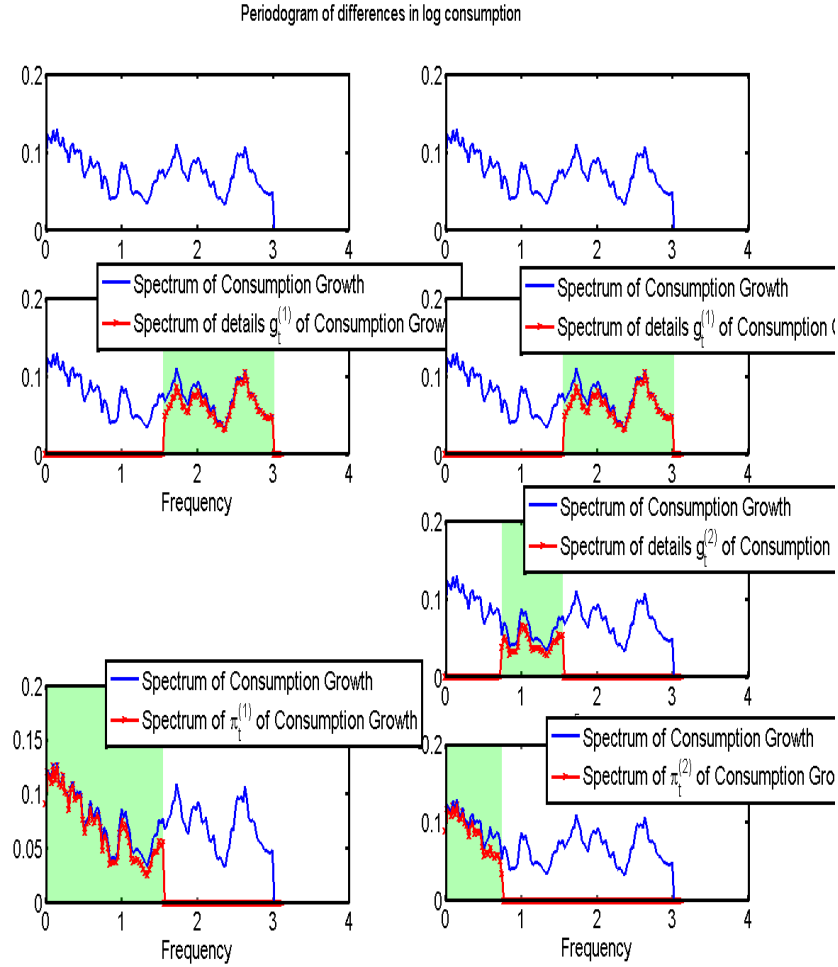
regressand and regressors in ordinary least squares, relates to Valkanov (2003)<sup>9</sup> and Bandi and Perron (2008) who study predictability relations aggregating both the regressor and regressand over non-overlapping periods with the same length. This is very similar to what we do in our componentwise regressions where we regress the consumption components obtained using aggregate consumption data in the time window  $t + 1$  to  $t + 2^j$  onto the price-dividend components in the time window  $t - 2^j$  to  $t$ .

---

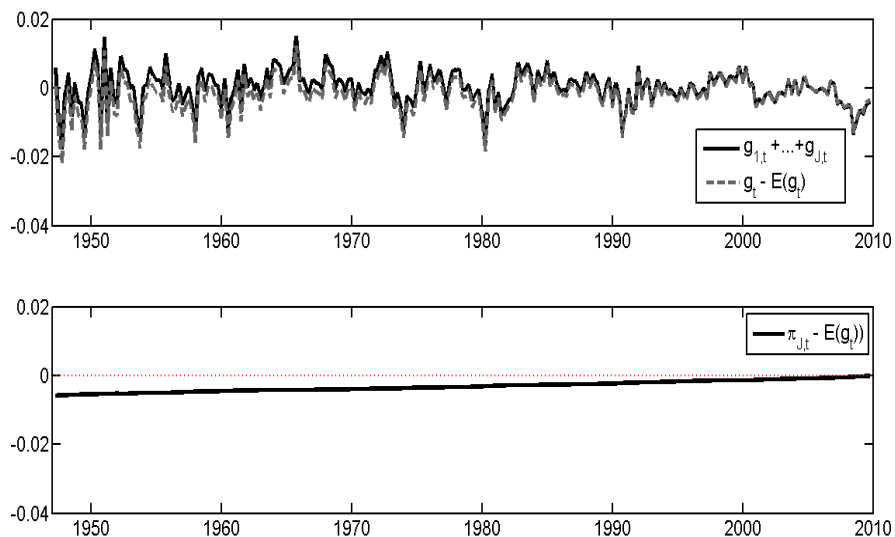
<sup>9</sup>In fact Valkanov (2003) shows that the OLS estimator is consistent when both the regressor and the regressand are aggregated over non-overlapping periods (cases 2 and 4 in their paper), i.e. regressing a long-horizon variable against the other.

## REFERENCES

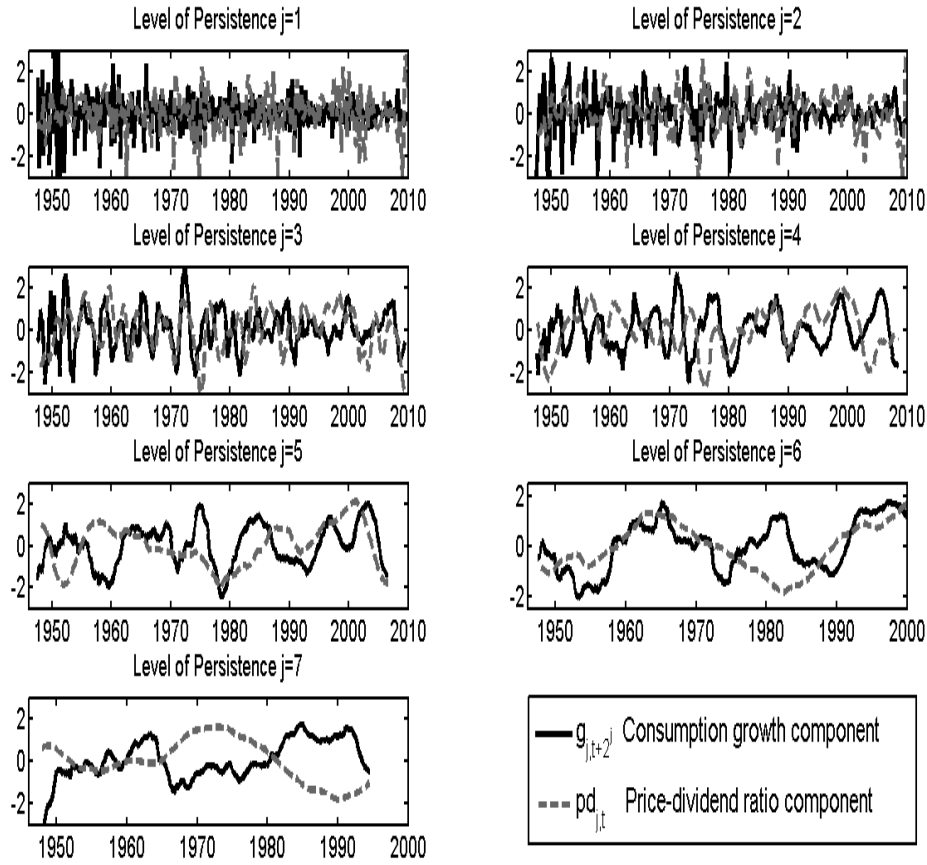
- Bandi, Federico M. and Benoit Perron, 2008, Long-run risk-return trade-offs, *Journal of Econometrics* 143, 349-374.
- Bansal, Ravi, and Amir Yaron, 2004, Risks for the long run: A potential resolution of asset pricing puzzles, *Journal of Finance* 59, 1481-1509.
- Bansal, Ravi, Dana Kiku and Amir Yaron, 2010, Risks For the Long Run: Estimation and Inference, *Working Papers* No-15504.
- Bansal, Ravi, Dana Kiku and Amir Yaron, 2009, An Empirical Evaluation of the Long-Run Risks Model for Asset Prices, *NBER Working Papers* No-15504.
- Bansal, Ravi, Robert Dittmar and Dana Kiku, 2007, Cointegration and Consumption Risks in Asset Returns, *Review of Financial Studies* 22, 1343-1375.
- Campbell, John Y. and N. Gregory Mankiw, 1990, Consumption, Income, and Interest Rates: Reinterpreting the Time Series Evidence, *NBER Working Papers* 2924.
- Comin, Diego and Mark Gertler, 2006, Medium-Term Business Cycles, *American Economic Review* 96, 523-551.
- Epstein, Larry G., and Stanley E. Zin, 1989, Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework, *Econometrica* 57, 937-69.
- Epstein, Larry G., and Stanley E. Zin, 1990, First-order' risk aversion and the equity premium puzzle, *Journal of Monetary Economics* 26, 387-407.
- Epstein, Larry G., and Stanley E. Zin, 1991, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis *Journal of Political Economy* 99, 263-286.
- Ferson, Wayne E. and Suresh Nallareddy and Biqin Xie, 2010, The Out-of-sample Performance of Long-Run Risk Models, *NBER Working Papers*.
- Hall, Robert E, 1988, Intertemporal Substitution in Consumption, *Journal of Political Economy* 96, 339-57.
- Gencay, Ramazan and Nikola Gradojevic, 2011, Errors in Variables Estimation with No Instruments, *Journal of Statistical Computation and Simulation* 81, 15451564.
- Valkanov, Rossen 2003, Long Horizon Regressions: Theoretical Results and Applications, *Journal of Financial Economics* 68, 201232.



**Figure IA.1.** The figure displays the effects of the persistence based decomposition of the consumption time series applied up to level  $J = 1$  (left panels) and  $J = 2$  (right panels). In particular the top panels displays the smoothed periodogram the consumption process for the data. An equally weighted “nearest neighbor” kernel was used to perform the smoothing, equally weighting the 4 nearest frequencies. The bottom right panel displays the Fourier spectrum of the time series  $\pi_t^{(2)}$  whereas the bottom left panel displays the Fourier spectrum of the time series  $\pi_t^{(1)}$ .



**Figure IA.2.** The figure displays the effect of removing the permanent component  $\pi_{J,t}$ . The top panel reports the *demeaned* series of US consumption growth (nondurables and services) from the Bureau of Economic Analysis consumption growth (dashed line) together with the sum of its components, excluding the permanent one (solid line). The correlation between the two series is 0.97. The bottom panel shows the difference between the permanent component  $\pi_{J,t}$  and the sample mean of the consumption growth. This difference vanishes as the sample length increases. Both panels are obtained for  $J = 7$ .



**Figure IA.3.** Time-scale decomposition for the log price-dividend  $pd_t$  and log consumption growth based upon quarterly data. The figure displays the future components of consumption growth  $g_{t+2^j}^{(j)}$ , along with the current components of the price-dividend ratio  $z_{m,t}^{(j)}$ . The components are obtained using the decomposition described in Section II.B.1. The sample spans the period 1947Q2-2010Q4.

**Panel A:**  $T = 256, J^* = 6$ 

Scale $j =$		1	2	3	4	5	6	7	8
$\rho$	$\text{Var}(g_t^{(j)})/\text{TotVar}$								
0.2	0.03	0.041	0.051	0.049	0.049	0.043	0.212	0.055	0.058
0.2	0.05	0.042	0.048	0.057	0.041	0.052	0.457	0.043	0.059
0.2	0.07	0.043	0.060	0.051	0.039	0.054	0.604	0.045	0.053
0.4	0.03	0.039	0.050	0.048	0.044	0.047	0.200	0.042	0.048
0.4	0.05	0.036	0.064	0.049	0.041	0.045	0.416	0.051	0.055
0.4	0.07	0.051	0.062	0.046	0.054	0.052	0.554	0.046	0.039

**Panel B:**  $T = 128, J^* = 4$ 

Scale $j =$		1	2	3	4	5	6	7
$\rho$	$\text{Var}(g_t^{(j)})/\text{TotVar}$							
0.2	0.12	0.055	0.057	0.043	0.266	0.048	0.051	0.057
0.2	0.20	0.061	0.070	0.061	0.431	0.044	0.045	0.051
0.2	0.07	0.085	0.090	0.066	0.639	0.046	0.041	0.045
0.4	0.12	0.051	0.057	0.056	0.256	0.041	0.048	0.046
0.4	0.20	0.068	0.071	0.048	0.405	0.054	0.044	0.041
0.4	0.26	0.099	0.091	0.067	0.614	0.051	0.039	0.043

**Table IA.1** This table reports the probability of rejecting the null hypotheses that the component at level  $j$  is white noise. We simulate  $T$  observation from a multiscale autoregressive process where the only persistent component is the one at level  $J^*$ . We repeat the experiment  $N = 2500$  times.



Component	Quarterly-frequency resolution
$g_t^{(1)}$	1 – 2 quarters
$g_t^{(2)}$	2 – 4 quarters
$g_t^{(3)}$	1 – 2 years
$g_t^{(4)}$	2 – 4 years
$g_t^{(5)}$	4 – 8 years
$g_t^{(6)}$	8 – 16 years
$g_t^{(7)}$	16 – 32 years
$\pi_t^{(7)}$	> 32 years

**Table IA.2** Frequency interpretation of the component  $g_t^{(j)}$  at level of persistence  $j$ . We assume the original time series  $g_t$  to be observed at quarterly intervals.

$j =$	Component at persistence level $j$						
	1	2	3	4	5	6	7
$\frac{Var(g_t^{(j)})}{Var(\sum g_t^{(j)})}$	0.369 [0.338, 0.427]	0.191 [0.177, 0.216]	0.158 [0.157, 0.159]	0.126 [0.108, 0.136]	0.068 [0.046, 0.080]	0.043 [0.026, 0.048]	0.045 [0.020, 0.062]
$\frac{Var(z_{m,t}^{(j)})}{Var(\sum z_{m,t}^{(j)})}$	0.016 [0.011, 0.067]	0.028 [0.019, 0.104]	0.048 [0.035, 0.163]	0.069 [0.056, 0.178]	0.126 [0.117, 0.208]	0.233 [0.101, 0.247]	0.480 [0.180, 0.515]

**Table IA.3** Contribution to total unconditional variance of the different components  $g_t^{(j)}$  of the log consumption growth and  $z_{j,t}^m$  of the log price-dividend ratio. Note that  $Var(\sum g_t^{(j)}) = Var(g_t)$  and  $Var(\sum z_{m,t}^{(j)}) = Var(z_t^m)$ . Approximate confidence intervals for the variance of the components are reported in square brackets.