RFS Advance Access published July 9, 2013

Long-Run Risk and the Persistence of Consumption Shocks

Fulvio Ortu
Bocconi University and IGIER

Andrea Tamoni
London School of Economics

Claudio Tebaldi
Bocconi University, IGIER, and CAREFIN

We propose a decomposition for time series in components classified by levels of persistence. Employing this decomposition, we provide empirical evidence that consumption growth contains predictable components highly correlated with well-known proxies of consumption variability. These components generate a term-structure of sizable risk premia. At low frequencies we identify a component correlated with long-run productivity growth and commanding a yearly premium of approximately 2%. At high frequencies we identify a component with yearly half-life, which contributes to the equity premium for another 2%. Accounting for persistence heterogeneity, we obtain an estimate of the IES strictly above one and robust across subsamples. (*JEL* G12, E21, E32, E44)

1. Introduction

This paper provides a systematic investigation of the persistence properties of consumption growth dynamics and proposes a solution to a major problem in the empirical detection of long-run risk: the lack of statistical power of the conventional tests used to estimate the low volatility, high-persistence component. The existence of such a component of consumption growth has fundamental pricing implications: Agents with Epstein-Zin preferences

---

We thank David Backus, Federico Bandi, Mike Chernov, Francesco Corielli, Itamar Drechsler, Carlo A. Favero, Christian Julliard, Sydney Ludvigson, Robert Ready (EFA discussant), Pedro Santa-Clara, Thomas Sargent, Lukas Schmid, and Stanley Zin for valuable insights, Marina Di Giacinto for editorial assistance. We also thank seminar participants at the Università di Roma "Tor Vergata," EFA 2011 conference participants, and participants at the 2011 4th Financial Forum on Long Term Risks in Paris, at the 2011 Royal Economic Society Annual Conference in London and at the 2010 6th CSEF-IGIER conference on Economics and Institutions in Capri for helpful comments and suggestions. We are especially grateful to Pietro Veronesi, the editor, and to an anonymous referee, for making several useful suggestions that have led to substantial improvements. Any errors or omissions are the solely responsibility of the authors. Funding has been provided solely by Bocconi University. Supplementary data can be found on the *Review of Financial Studies* Web site. Send correspondence to Claudio Tebaldi, Bocconi University, Department of Finance, Via Roentgen 1, Milan, 20136, Italy; telephone: +39-02-58363638. Email: claudio.tebaldi@unibocconi.it

doi:10.1093/rfs/hht038

require a compensation for holding cash flows whose future fluctuations are positively correlated with expected changes in consumption growth, a fact that, in equilibrium, generates a long-run risk premium. To disentangle these components, we develop a tool to decompose a time series into layers with different levels of persistence. These layers do not enlarge the space of shocks generating the time series, rather they classify them by their half-life, so as to capture economic phenomena occurring at different time-scales. To exemplify our methodology, we generate a time series via the aggregation of components, which are all white noise but one. We show that when the persistent component contributes for a very small fraction of the total volatility, standard statistical tests fail to distinguish the time series from a white noise. Our approach, on the contrary, is successful in detecting the persistent layer hidden in the time series.

Motivated by this example, we apply our decomposition to the consumption growth series, and we document that consumption growth has cyclical components defined at different time-scales. We also investigate the existence of reasonable and observable economic proxies for the persistent components filtered out of consumption growth. To search for these proxies, we rely on time series that are economically significant, characterized by a half-life close to the one of the components they are to proxy for, and significantly correlated with such components. In particular, we document a strong correlation between cyclical consumption growth variations captured by the below-business-cycle frequency component and long-run productivity growth. This is similar to Pastor and Veronesi (2006, 2009), Kaltenbrunner and Lochstoer (2010), and Croce (2010), who find that shifts in the long-run rate of productivity growth are a key factor in driving the slow-moving consumption components. The components corresponding to business-cycle frequencies are correlated with well-known economic indicators of economic activity, such as the term and the corporate default spreads. The high-frequency components taken together have a yearly half-life, a value which is close to the half-life of the shocks in Bansal and Yaron (2004). We identify these high-frequency components with the well-documented fourth-quarter effect (see, e.g., Jagannathan and Wang 2007; Moller and Rangvid 2012).

Our persistence-based decomposition implies that the contemporaneous presence of highly persistent components with small volatility and highly volatile components with low persistence can hide the predictability relations that exist at specific levels of persistence. We document this fact by showing that, even if aggregate consumption and dividend growth are unpredictable by the financial ratios (see Beeler and Campbell 2012), there exist specific components of consumption and dividends growth that are predicted by the components of the price-dividend ratio with the same degree of persistence. Consistent with our methodology, moreover, at each given level of persistence the forecast occurs over a time horizon dictated by the half-life of the predicted component. Similarly, we exploit the persistence heterogeneity in consumption growth to obtain a set of regressions that link the components with different degrees of persistence of the risk-free rates to the components with the same level of persistence of consumption growth. Importantly, the slope must be the same across levels of persistence and equal to the inverse of the intertemporal elasticity of substitution (IES). This approach allows us to obtain an estimate of the IES that is both strictly greater than one and robust across subsamples.

Are the predictable components of consumption growth actually priced? We answer this question by analyzing a long-run risk economy in which the exogenous driving variables are the different layers of persistence in the consumption growth series. Our asset pricing model allows us to determine the contribution to the equity premium of these different layers in consumption growth and to construct a term-structure of risk premia. We document that not all components are priced. In particular, the total equity premium implied by our model is determined by the contribution associated with two specific levels of persistence, one on the low-frequency side of the spectrum and the other on the high-frequency side. On the low-frequency side, we identify a thin (in variance) component whose variations occur on time-scales ranging from 8 to 16 years. This component commands a premium of up to 2% per year when the risk aversion takes the reasonable value of 7.5 and the IES is 2.5. On the high-frequency side, we find two short-run predictable components. These components lie close to each other in the spectrum and capture fluctuations with a half-life between one-half and two years. For this reason, we approximate them with a single factor with yearly half-life. When taken together, these two components contribute to the equity premium for up to another sizable 2%. From an empirical point of view, therefore, our persistence-based decomposition generates important implications for the consumption predictability, the intertemporal elasticity of substitution, and the equity premium.

Our contribution places naturally in the fast-growing literature beginning with the seminal paper of Bansal and Yaron (2004), who look at the long-run regime to explain many of the inconsistencies that affect predictions of dynamic asset pricing models. Our model, however, differs from the standard long-run risk economy in terms of what we price. Whereas the long-run risk framework prices the latent persistent conditional mean of consumption growth, in our model we price the different components of consumption growth, each one evolving on a characteristic time-scale. Our focus, therefore, is not just on long-run risk but rather on the entire term structure of consumption risk. Our paper also contributes to the long-run risk literature by reconciling the evidence in Parker and Julliard (2005), Bansal, Dittmar, and Lundblad (2005), and Bansal, Dittmar, and Kiku (2007), who find long-run risk in consumption and cash flows to be an important determinant of asset returns, with the evidence presented in Constantinides and Ghosh (2011) and Beeler and Campbell (2012), who find instead a reversal of the earlier conclusions.

Empirical literature has produced contradictory evidence with regard to the estimation of the intertemporal elasticity of substitution. On the one hand, Hall (1988) and Campbell and Mankiw (1990) estimate an extremely small value of IES on U.S. data and Campbell (2003) summarizes these results and finds similar patterns in international data. On the other hand, Attanasio and Weber (1993) and Beaudry and van Wincoop (1996) find values of the IES higher than one using disaggregated cohort-level and state-level consumption data. Vissing-Jorgensen (2002), moreover, points out that many consumers do not participate actively in asset markets and finds that, in household data, the IES is greater than one for asset market participants. More recently, Avramov and Cederburg (2012) have used a vector autoregressive approach to long-run risk to estimate an IES of 4.5. In line with this second stream of literature we present empirical evidence on the existence of aggregation problems that obscure the relation between consumption growth and the real interest rate. We show that the use of disaggregated consumption data are key to finding a value for the IES greater than one. Our contribution here is to point to persistence heterogeneity in consumption as another possible aggregation mechanism that produces an IES greater than one.

Our work is also close to Calvet and Fisher (2007), who investigate the role of persistence heterogeneity in volatility in a partial equilibrium setup by means of nonlinear regime switching multifractal models. Moreover, our decomposition shares the same insight as the (multiplicative) permanent-transitory decomposition proposed in Hansen and Scheinkman (2009) and used in Hansen, Heaton, and Li (2008). Differently from their decomposition, however, we further decompose the time series to extract all the transitory components with different degrees of persistence present in it. Similar to Fourier analysis, our decomposition breaks a time series into a number of distinct components, each of them lying in a particular frequency range. Our approach, however, retains the simplicity of time-series methods and allows us to model, analyze, and predict the components of the original time series using the conventional tools of time-series analysis.

The remainder of the paper is organized as follows: In Section 2.1 we show how to decompose a time series into components with different levels of persistence. In Section 2.2 we apply this method to consumption growth. Section 3.1 and 3.2 revisit the predictability of consumption growth and the estimation of the intertemporal elasticity of substitution in light of the evidence showing that consumption growth contains persistent components. Section 3.3 quantifies the contribution of the predictable components to risk premia within an asset pricing model that accounts for persistence heterogeneity. In Section 4 we verify the robustness of our results. Section 5 concludes. Appendix A describes the data sources and construction of variables used for the empirical analysis, and Appendix B provides additional details on the implementation of our filtering procedure.

2. Persistence Heterogeneity in Consumption Growth

In this section we document the existence of persistent components in consumption growth. We base our evidence on a decomposition of a time series in multiple layers with different degrees of persistence. Our approach allows us to reject the null "consumption growth is white noise," to extract its persistent components and to show that they correlate well with important macroeconomic and financial variables. These facts suggest an alternative mechanism of propagation of consumption shocks to asset prices, which we detail in Section 3.

2.1 Decomposing time series along the persistence dimension

Given a time series $\{g_t\}_{t\in\mathbb{Z}}$, we begin by constructing moving averages $\pi_t^{(j)}$ of size 2^j:

$$\pi_t^{(j)} = \frac{1}{2^j}\sum_{p=0}^{2^j-1} g_{t-p} \ , \tag{1}$$

where $\pi_t^{(0)} \equiv g_t$. Given the choice of sample size, it is readily observed that these moving averages satisfy the iterative relation:

$$\pi_t^{(j)} = \frac{\pi_t^{(j-1)} + \pi_{t-2^{j-1}}^{(j-1)}}{2} \ . \tag{2}$$

Next, we denote by $g_t^{(j)}$ the difference between moving averages of sizes 2^{j-1} and 2^j, that is,

$$g_t^{(j)} = \pi_t^{(j-1)} - \pi_t^{(j)} \ . \tag{3}$$

Intuitively, $g_t^{(j)}$ captures fluctuations that survive to averaging over 2^{j-1} terms but disappear when the average involves 2^j terms, that is, fluctuations with half-life in the interval $\left[2^{j-1}, 2^j\right)$. Accordingly, the moving average $\pi_t^{(j)}$ includes fluctuations whose half-life exceeds 2^j periods.[1] From now on, we refer to the derived time series $\left\{g_t^{(j)}\right\}_{t\in\mathbb{Z}}$ as to the *component* of the original time series $\{g_t\}_{t\in\mathbb{Z}}$ with level of persistence j. Because $\pi_t^{(0)} \equiv g_t$, by summing up over j, it follows immediately from (3) that

$$g_t = \sum_{j=1}^{J} g_t^{(j)} + \pi_t^{(J)} \tag{4}$$

for any $J \geq 1$. In words, Equation (4) decomposes the time series g_t into a sum of components with half-life belonging to a specific interval, plus a residual term that represents a long-run average.

[1] In the online Technical Appendix, we relate the persistence properties of the series $g_t^{(j)}$ and $\pi_t^{(J)}$ to their Fourier spectra. In particular, we show that each of these components lies in a particular frequency range.

Due to the overlap of the moving averages that define $g_t^{(j)}$, the decomposition (4) can lead to a biased evaluation of the persistence of the time series g_t. The bias emerges clearly, for instance, when g_t is white noise: In that case, the components in (4) would exhibit serial correlation due to the mechanical overlapping of the moving averages, even though there is no persistence in g_t. To address this issue, we select the information in the components $g_t^{(j)}$ and $\pi_t^{(j)}$ in a suitable manner. In particular, because by definition each component $g_t^{(j)}$ is a linear combination of the realizations $g_t,\ g_{t-1},\ldots,g_{t-2^j+1}$, to remove any spurious serial correlation introduced by the overlapping of the moving averages, we restrict our attention to the subseries:

$$\left\{g_t^{(j)}, t=k2^j, k\in\mathbb{Z}\right\}, \tag{5}$$

$$\left\{\pi_t^{(j)}, t=k2^j, k\in\mathbb{Z}\right\}.^{2} \tag{6}$$

We refer to these subseries as the *decimated components* at level of persistence j of the original time series. Clearly, persistence in a decimated component is not an artifact; rather, it represents an actual fluctuation of the original series with an half-life in the interval $\left[2^{j-1}, 2^j\right)$.

The process of decimation controls for spurious persistence by deleting from the components $g_t^{(j)}$ and $\pi_t^{(j)}$ *all and only* the information irrelevant to reconstruct the original time series g_t. Formally, this follows from observing that for any $J\geq 1$ one can define a linear, invertible operator $\mathcal{T}^{(J)}$ that maps the decimated components $\left\{g_t^{(j)}, t=k2^j, k\in\mathbb{Z}\right\}$, $j=1,\ldots,J$, and $\left\{\pi_t^{(j)}, t=k2^j, k\in\mathbb{Z}\right\}$ into the time series $\{g_t\}_{t\in\mathbb{Z}}$. To illustrate how this works for $J=2$, we first observe that in this case (1) yields

$$\pi_t^{(2)} = \frac{g_t+g_{t-1}+g_{t-2}+g_{t-3}}{4}. \tag{7}$$

Next, we substitute (2) into (3) and let $j=1,2$ to obtain

$$\begin{aligned}
g_t^{(2)} &= \frac{\pi_t^{(1)}-\pi_{\cancel{t-1}}^{(1)}}{2} = \frac{1}{2}\left(\frac{g_t+g_{t-1}}{2}-\frac{g_{t-2}+g_{t-3}}{2}\right)\\
g_t^{(1)} &= \frac{\pi_t^{(0)}-\pi_{t-1}^{(0)}}{2} = \left(\frac{g_t-g_{t-1}}{2}\right)\\
g_{t-2}^{(1)} &= \frac{\pi_{t-2}^{(0)}-\pi_{t-3}^{(0)}}{2} = \left(\frac{g_{t-2}-g_{t-3}}{2}\right).
\end{aligned} \tag{8}$$

[2] These subseries are defined up to a translation factor, which results in 2^j degrees of freedom. More precisely for any $h=0,1,\ldots,2^j-1$, we could sample the realizations $\left\{g_{h+k2^j}^{(j)}, k\in\mathbb{Z}\right\}$, $\left\{\pi_{h+k2^J}^{(J)}, k\in\mathbb{Z}\right\}$. The approach described in this section satisfies a translation invariant property, that is, it is independent of the parameter h. This is why in constructing (5) and (6) we let $h=0$ without loss of generality.

We then consider the system obtained by stacking (7) on top of (8), which in matrix notation becomes

$$\begin{pmatrix} \pi_t^{(2)} \\ g_t^{(2)} \\ g_t^{(1)} \\ g_{t-2}^{(1)} \end{pmatrix} = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ \frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} g_t \\ g_{t-1} \\ g_{t-2} \\ g_{t-3} \end{pmatrix}. \tag{9}$$

Denoting by $\mathcal{T}^{(2)}$ the (4×4) matrix in (9), we notice that $\mathcal{T}^{(2)}$ is orthogonal, that is, $\Lambda^{(2)} \equiv \mathcal{T}^{(2)}\left(\mathcal{T}^{(2)}\right)^{\top}$ is diagonal. Moreover, the diagonal elements of $\Lambda^{(2)}$ are nonvanishing so that $\left(\mathcal{T}^{(2)}\right)^{-1} = \left(\mathcal{T}^{(2)}\right)^{\top}\left(\Lambda^{(2)}\right)^{-1}$ is well defined, and hence

$$\begin{pmatrix} g_t \\ g_{t-1} \\ g_{t-2} \\ g_{t-3} \end{pmatrix} = \left(\mathcal{T}^{(2)}\right)^{-1} \begin{pmatrix} \pi_t^{(2)} \\ g_t^{(2)} \\ g_t^{(1)} \\ g_{t-2}^{(1)} \end{pmatrix}. \tag{10}$$

By letting t vary in the set $\left\{t = k2^2, k \in \mathbb{Z}\right\}$, Equation (10) shows how to reconstruct uniquely the entire time series $\{g_t\}_{t\in\mathbb{Z}}$ from the decimated components $\left\{g_t^{(j)}, t = k2^j, k \in \mathbb{Z}\right\}$, $j = 1,2$, and $\left\{\pi_t^{(2)}, t = k2^2, k \in \mathbb{Z}\right\}$.[3]

Figure 1 visualizes how the decomposition (4) uses redundant information to reconstruct the original time series g_t, while the decimated decomposition (10) reconstructs g_t employing optimally the information contained in the decimated components. Looking, for instance, at any block $g_t, g_{t-1}, g_{t-2}, g_{t-3}$ the figure shows how (4) uses twelve data points, whereas (10) uses only four data points; that is, exactly as many as those in the original block. In this sense, the decimated components contain the minimal information necessary to reconstruct the time series. In particular, the decimated decomposition formalized by the operator $\mathcal{T}^{(2)}$ leaves the number of shocks unaltered, while classifying them in terms of their time of arrival and persistence level, as Panel B of Figure 1 highlights.

2.1.1 Persistence versus white noise: An example. To illustrate the advantage of decomposing time series along the persistence dimension, we produce a time series that would be judged as white noise by standard statistical tests, while it contains a persistent component detected by the decimated decomposition introduced above. To do so, we directly model the dynamics

[3] For an extension of this procedure to any $J \geq 2$ and a recursive algorithm for the construction of the matrix $\mathcal{T}^{(J)}$, associated to an arbitrary level of persistence J, see, for example, Mallat (1989, 1989). The recursive construction of $\mathcal{T}^{(J)}$ shows that the diagonal elements of the matrix $\Lambda^{(J)} \equiv \mathcal{T}^{(J)}\left(\mathcal{T}^{(J)}\right)^{T}$ are $\lambda_1 = \lambda_2 = 1/2^J$, $\lambda_k = 1/2^{J-j+1}$, $k = 2^{j-1}+1,\ldots,2^j$, $j = 2,\ldots,J$, a fact useful in developing our statistical test in Section 2.1.2.

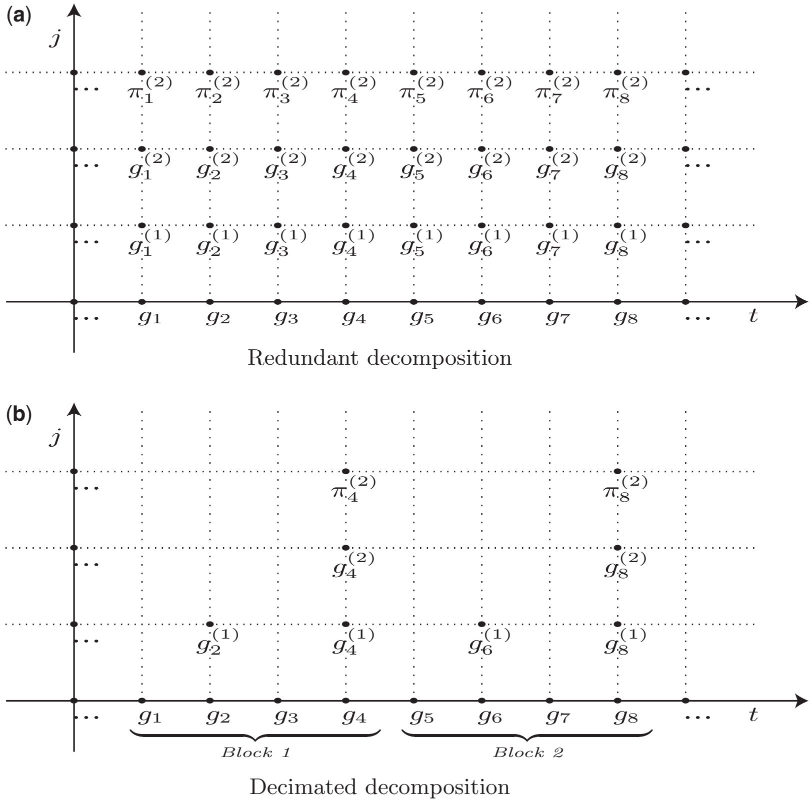

Figure 1
Redundant versus decimated decomposition
This figure displays the components of the time series g_t before (Panel A) and after decimation (Panel B).

of the decimated components defined in (5) and (6) and then use the (inverse of the) operator $\mathcal{T}^{(J)}$ to reconstruct the process g_t. We assume, in particular, that all decimated components are independent normal innovations, except for one which follows an autoregressive process on the time domain defined by decimation. More formally, for $t = k2^j, k \in \mathbb{Z}$, we let

$$
\begin{aligned}
g_t^{(j)} &= \epsilon_t^{(j)}, \quad \forall j < J^* \\
g_{t+2^{J^*}}^{(J^*)} &= \rho_{J^*} g_t^{(J^*)} + \epsilon_{t+2^{J^*}}^{(J^*)}, \\
\pi_t^{(J^*)} &= \eta_t^{(J^*)}
\end{aligned}
\tag{11}
$$

with $\epsilon_t^{(j)} \sim \mathrm{N}(0, 2^{-j})$, $j < J^*$, $\epsilon_t^{(J^*)} \sim \mathrm{N}(0, 2^{-J^*}(1-\rho_{J^*}^2))$, and $\eta_t^{(J^*)} \sim \mathrm{N}(0, 2^{-J^*})$. Finally, we assume that the serially independent innovations $\epsilon_t^{(j)}$, $\epsilon_t^{(j')}$ are uncorrelated for $j \neq j'$ and that $\epsilon_t^{(j)}$ is uncorrelated with $\eta_t^{(J^*)}$ for all j as well, that is, the decimated components are independent across levels of persistence.

Observe that by construction $\mathrm{Var}(g_t)=\sum_{j=1}^{J^*}\mathrm{Var}\left(g_t^{(j)}\right)+\mathrm{Var}(\pi_t^{(J^*)})=\sum_{j=1}^{J^*}2^{-j}+2^{-J^*}=1$ so that the persistent component $g_t^{(J^*)}$ explains a fraction 2^{-J^*} of the total variability of g_t.[4] As J^* increases, therefore, the predictable component accounts for a small percentage of the total variance of g_t.

The point of this example is to show that the persistent component can be obscured by the layers of serially independent innovations unless a suitable filtering procedure is used to disentangle the different degrees of persistence. To this end we let $J^*=4$ so that the persistent component accounts only for 6% of the total variance; we set $\rho_{J^*}=0.5$; and we simulate the components using the dynamics described in (11). We then use the inverse of the matrix $\mathcal{T}^{(J)}$ introduced in (9), here for $J=4$, to reconstruct the series g_t using the procedure described in the previous section. Figure 2 displays the results of this exercise. The top left panel shows the reconstructed time series g_t. Below the top panel we report the components $g_t^{(j)}$, $j=1,\ldots,4$ and $\pi_t^{(4)}$. We report on the right panels the autocorrelation functions associated with these series to assess the presence of autocorrelation at individual lags. The autocorrelation function of the series g_t clearly resembles the one of a white noise process.[5] In fact, the Kolmogorov-Smirnov test to determine if the series g_t comes from a standard normal distribution cannot reject the null hypothesis at a 70% significance level. Hence, this statistical method fails to detect the dependency pattern, which, by construction, is in the data. The precise sense in which the persistence induced in g_t by $g_t^{(4)}$ is obscured is that it is dominated by the nonpredictable, large variance components. It is interesting to observe that this conclusion is robust to a wide range of parameters choices. This is so because when J^* increases, the total variance explained by the persistent component decreases, and the set of possible values of ρ_{J^*} for which the persistent component still goes undetected converges to the open interval $(-1,1)$. For the case $J^*=4$, for instance, the range of ρ_{J^*} for which the test fails to detect the persistent component is approximately $(-0.8,0.8)$.

In sum, standard statistical tests that focus on the aggregate behavior of the time series fail to detect the component localized at a specific level of persistence and too often accept the white noise hypothesis. How to detect empirically the

[4] This is so because our example is tailored in such a way that $\mathrm{Var}\left(\mathcal{T}^{(J^*)}X_t^{(J^*)}\right)=\Lambda^{(J^*)}$, where $X_t^{(J^*)}$ is the vector containing the 2^J observations of our time series up to time t, and $\Lambda^{(J^*)}$ is the matrix defined in footnote 3. This implies that $\mathrm{Var}\left(X_t^{(J^*)}\right)=\mathrm{Var}\left(\left(\mathcal{T}^{(J^*)}\right)^{-1}\mathcal{T}^{(J)}X_t^{(J^*)}\right)=I$, where I is the identity matrix, highlighting the fact that the correlation of the time series manifests itself at a time-scale greater that 2^{J^*} periods.

[5] We use a Ljung-Box Q-test to test for autocorrelation at multiple lags jointly. We cannot reject the null hypothesis that the first $m=1,2,\ldots,6$ autocorrelations are jointly zero. Furthermore, the degree of over rejection magnifies significantly at different lag lengths.

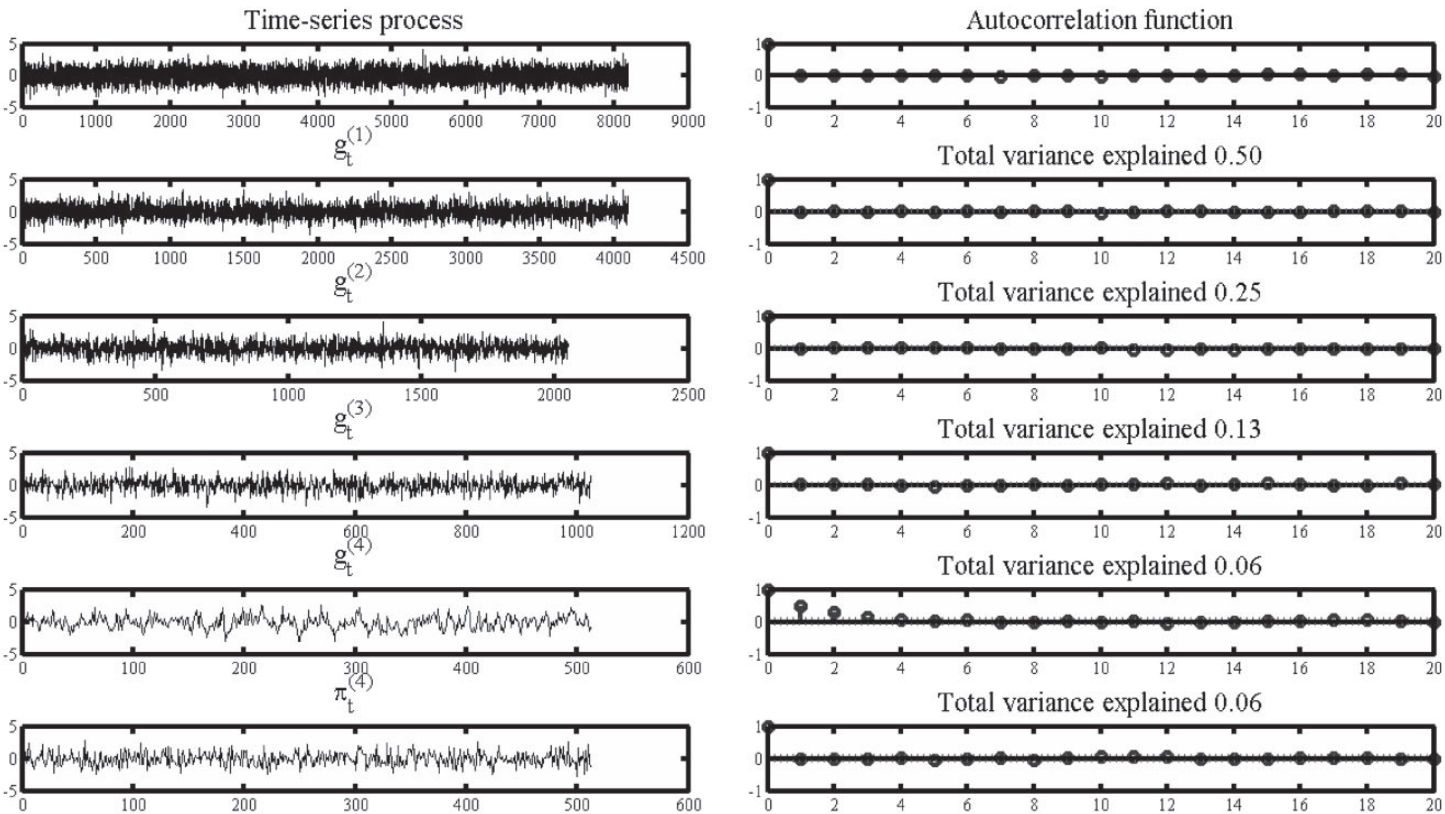

Figure 2
Persistence versus white noise: The role of the components

We simulate a process with one persistent component at time-scale J^*. In particular $g_t^{(j)} = \epsilon_t^{(j)}, \quad \forall j < J^*$, and $g_{t+2^{J^*}}^{(J^*)} = \rho_{J^*} g_t^{(J^*)} + \epsilon_{t+2^{J^*}}^{(J^*)}$ where we set $J^* = 4$ and $\rho_{J^*} = 0.5$. We obtain the stochastic process in the time domain by antitransforming its component. The unconditional variance of the (antitransformed) process in the time domain is set equal to one. The figure displays in the top left panel the antitransformed process. The remaining panels on the left display the time series of the components. The right panels display the autocorrelation functions of the antitransformed process (top right) and of the components.

existence of persistence in an apparently white noise series is covered in the next section.

2.1.2 Detecting small but persistent components. We construct a test that distinguishes a white noise process from a process whose decimated components are serially correlated. Our test, which builds on a new family of tests for serial correlation as introduced by Gencay and Signori (2012), has desirable size and power in small samples.

We assume $\{g_t\}$ to be weakly stationary with $E[g_t]=0$, $Var(g_t)=\sigma^2$ and we denote with $\left(X_T^{(J)}\right)^\intercal=[g_T,g_{T-1},\ldots,g_1]$ the vector collecting the observations of $\{g_t\}$.[6] We then use the operator $\mathcal{T}^{(J)}$ introduced in (9) to obtain a variance decomposition for the series g_t. To do so, we express the sample variance of g_t as the sum of the variances of its decimated components $\mathbf{g^{(j)}}=\left[g_{2^j}^{(j)},\ldots,g_{k\cdot 2^j}^{(j)},\ldots,g_T^{(j)}\right]^\intercal$:

$$\frac{\left(X_T^{(J)}\right)^\intercal X_T^{(J)}}{T}=\frac{\left(\left(\Lambda^{(J)}\right)^{-1/2}\mathcal{T}^{(J)}X_T^{(J)}\right)^\intercal\left(\left(\Lambda^{(J)}\right)^{-1/2}\mathcal{T}^{(J)}X_T^{(J)}\right)}{T}=\frac{\sum_{j=1}^J 2^j\cdot(\mathbf{g^{(j)}})^\intercal\mathbf{g^{(j)}}}{T}$$

The first equality exploits the fact that the matrix $\left(\Lambda^{(J)}\right)^{-1/2}\mathcal{T}^{(J)}$ is orthonormal, whereas the second equality follows upon recalling the expression of the diagonal elements of the matrix $\Lambda^{(J)}$.[7] The presence of the factor 2^j in our variance decomposition is readily understood upon recalling that, due to decimation, the component $\mathbf{g^{(j)}}$ with level of persistence j has in fact $T/2^j$ observations so that its sample variance is exactly $(\mathbf{g^{(j)}})^\intercal\mathbf{g^{(j)}}/\left(T/2^j\right)$.

Our test statistics build on the above variance decomposition and rely on the comparison of the contribution to total variance of the different components of a white noise process on one side and of a process with serially correlated decimated components on the other side. Specifically, we employ as test statistics the ratio of the sample variance of the decimated components to the sample variance of the time series, that is,

$$\hat{\xi}_j=\frac{2^j\cdot(\mathbf{g^{(j)}})^\intercal\mathbf{g^{(j)}}}{\left(X_T^{(J)}\right)^\intercal X_T^{(J)}}\ .$$

Letting $Cov(g_t,g_{t-k})/Var(g_t)\equiv\rho_k$, we test the null hypothesis $\rho_k=0$ for all $k\geq 1$ against $\rho_k\neq 0$ for some $k\geq 1$. It can now be shown that, under the null

[6] The dimensionality of the sample is dictated by the decimated procedure necessary to avoid spurious correlation. In this section we assume to have a sample of $T=2^J$ observations. How to deal with this issue in the empirical applications is discussed in Appendix B.

[7] See footnote 3. Observe also that the variance decomposition of g_t should include the term $(\boldsymbol{\pi}^{(J)})^\intercal\boldsymbol{\pi}^{(J)}$. Given the stationarity assumption and as long as the sample is large enough, it follows from (1) that $\boldsymbol{\pi}^{(J)}$ is an unbiased estimator of the population mean which by assumption is zero. This is why we disregard this term.

of no serial correlation and for a suitable choice of rescaling factors a_j, the properly rescaled variance ratio statistics converge in distribution to a standard normal, that is,[8]

$$\sqrt{\frac{T}{a_j}}\left(\hat{\xi}_j - \frac{1}{2^j}\right) \xrightarrow{d} N(0,1)\ . \tag{12}$$

In other words, because the distribution of the variance of the components of a white noise process is uniform in the degree of persistence, and because the component at level of persistence j captures fluctuations in the range $[2^{j-1}, 2^j)$, then, under the null, the suitably rescaled deviation of the variance ratio $\hat{\xi}_j$ from its large sample mean $1/2^j$ approaches a standard normal.[9] The suitable choice of rescaling factors guarantees that this property holds independently of the degree of persistence of the decimated components.

We now investigate by means of Monte Carlo simulations the finite sample performance of our variance ratio test in the framework of the example introduced in the previous section (see (11)). In this simulation exercise we set $T=256$ so that the simulated sample size matches the postwar quarterly data sample used in the empirical analysis carried on in the rest of the paper. We choose $J^*=6$, and we let ρ_{J*} be either 0.2 or 0.4. We assume that the variance of the persistent component explains alternatively $3\%, 5\%$, and 7% of total variance. We remark that, for any such combination of parameters, the moments of the simulated series are compatible with those of actual consumption growth in the postwar quarterly sample. In particular, the simulated series has volatility of about 1%, first-order autocorrelation less than 0.1, and higher-order autocorrelations that are not significant based on a Ljung-Box test.[10]

For each simulation we compute the rescaled test statistics $\hat{\xi}_j$, and we carry out a two-tailed test. We repeat this experiment $N=5{,}000$ times, and in Table 1 we report for each level of persistence j the probability of rejecting the null.[11] Table 1, Panel A, shows that, when the alternative is a process with a serially correlated decimated component, the variance ratio test $\hat{\xi}_j$ displays strong power exactly at the time-scale at which the persistent component is localized. In particular, the test statistics $\hat{\xi}_6$ has a rejection

[8] To establish this fact, we apply Theorem 10 in Gencay and Signori (2012) to our decimated components. The details of the derivation can be found in the online Technical Appendix. To compute the values of a_j for different resolution scales, we use Corollary 12 in Gencay and Signori (2012).

[9] We are implicitly relying on the relation between the spectral density function $S_g(f)$ and the variance by which the contribution of the frequencies in a small interval $\Delta f = [1/2^{j-1}, 1/2^j]$ is approximately $S_g(f)\Delta f$. In particular, the spectral density function of an uncorrelated process is constant.

[10] In the online Technical Appendix, we also report the results for the case $T=128$, $J^*=4$, $\rho_{J*}=0.2$, or $\rho_{J*}=0.4$, which closely resembles our longer annual data sample.

[11] We adjust the empirical power for variations in the empirical size, and we compute the rejection rate with respect to a 5% empirical size.

Table 1
Rejection rates under the null hypothesis

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

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

Panel B: $T=2048$, $J^*=6$

ρ_{J^*}	$\text{Var}(g_t^{(J^*)})/\text{Var}(g_t)$	Persistence level $j=$ 1	2	3	4	5	6	7
0.2	0.03	0.037	0.047	0.051	0.060	0.046	0.761	0.055
0.2	0.05	0.061	0.096	0.069	0.051	0.052	0.991	0.043
0.2	0.07	0.081	0.091	0.106	0.060	0.047	0.997	0.053
0.4	0.03	0.050	0.054	0.058	0.055	0.038	0.710	0.053
0.4	0.05	0.063	0.093	0.076	0.048	0.046	0.982	0.055
0.4	0.07	0.080	0.095	0.078	0.068	0.054	0.995	0.051

This table reports the rejection probabilities of our variance ratio test with nominal levels of 5% against the multiscale autoregressive process. The null hypotheses is that the process g_t is white noise. We simulate T observation from a multiscale autoregressive process in which the only persistent component is the one at level J^*. All simulations are based on 5,000 replications.

rate between 20% and 60% (depending on the variance of the persistent component) even in a small sample of 256 observations. By looking at the i.i.d components, we see that our variance ratio test statistics $\hat{\xi}_j$, $j \neq 6$ do not overreject the null hypotheses. Overall, our test has desirable size and power in small samples. Finally, Table 1, Panel B, reports results from a large sample simulation with $T=2,048$ observations. From this table it appears clear that the power of the test increases steadily with the sample size, approaching one in population.

This Monte Carlo exercise shows that our variance ratio test has good size and power properties against the alternative of a process with a persistent component localized at a given level j. Instead of looking at the aggregate behavior of the time series, as it is typically done in standard portmanteau tests for serial correlation, this test filters out the components with different degrees of persistence and exploits the fact that for a serially correlated process each component contributes for a different percentage to the variance of the process. This persistence localization feature is what determines the small sample power improvements of our test.

2.1.3 Determining the optimal number of components. From a practical point of view a crucial issue is to determine when to stop the extraction of

Table 2
The variance ratio test for consumption growth

	Persistence level $j=$						
	1	2	3	4	5	6	7
$\sqrt{\frac{T}{a_j}}\left(\hat{\xi}_j - \frac{1}{2^j}\right)$	**−4.38**	**−1.86**	1.22	**3.04**	**3.05**	**3.53**	0.97

This table reports the rescaled variance ratio statistics for the consumption growth series. Bold values reject the null hypothesis of no serial correlation at a 95% confidence level.

the components of a nonwhite noise series. This issue is solved by framing it as a test for the existence of a maximum degree of persistence, based on a sequential analysis of the decimated series $\left\{\pi_t^{(J)}, t = k2^J, k \in \mathbb{Z}\right\}$, $J = 1, 2, \ldots$, defined in (6). Recall that $\pi_t^{(J)}$ by definition incorporates the fluctuations of the time series with persistence greater than 2^J periods, if there are any. Therefore, the existence of a maximum degree of persistence in the original series g_t is equivalent to the existence of J such that $\pi_t^{(J)}$ is white noise.

Given this property, the criterion to determine the optimal number of components to be extracted from a given sample consists of applying sequentially our variance ratio test to the series $\pi_t^{(J)}$, $J = 1, 2, \ldots$ stopping the iteration at the smallest J for which we cannot reject that $\pi_t^{(J)}$ is white noise.[12] The intuition behind our criterion is that the optimal number of components to be extracted is the result of optimizing the trade-off between the volatility of a component and its degree of persistence. Our sequential variance ratio test achieves this aim exactly by relying on these two key dimensions. A standard factor analysis approach, on the contrary, would fail to select the right number of components, because in relying on the sole evidence of total variance explained, it would miss the persistence dimension of the problem.

2.2 Filtering and identifying the consumption components

Our variance ratio test shows that consumption growth is not a white noise process.[13] Table 2 reports the rescaled variance ratio statistics for different levels of persistence. In this table bold values represent rejection at the 95% confidence level. Inspecting the table, we see that a white noise model for consumption growth is rejected at multiple levels of persistence.

This evidence implies that consumption growth contains persistent components. We determine the optimal number of components to be extracted using the criteria developed in the previous section. Recalling (1) and given the length of our postwar quarterly data sample, it is readily seen that $\pi_t^{(8)}$ is the sample mean of g_t so that we apply our test to $\pi_t^{(J)}$, $J = 1, 2, \ldots, 7$. Table 3

12 Clearly, if in a given sample of T observations the null $\pi_t^{(J)}$ is white noise is rejected for all $J = 1, \ldots, \lfloor \log_2 T \rfloor$, then the conclusion is that the maximum degree of persistence, if it exists, exceeds $\lfloor \log_2 T \rfloor$.

13 Appendix A describes the data sources and construction of variables used for the empirical analysis.

Table 3
Determining the optimal number of components

Scale $k=$	1	2	3	4	5	6	7
$\boldsymbol{\pi}^{(1)}$	**−6.29**	−0.38	**2.88**	**2.12**	1.22	−0.77	−0.20
$\boldsymbol{\pi}^{(2)}$	**−4.65**	−0.15	0.26	0.05	−1.52	−0.29	
$\boldsymbol{\pi}^{(3)}$	**−3.28**	**−2.15**	−1.40	**−2.57**	−0.44		
$\boldsymbol{\pi}^{(4)}$	**−3.15**	**−2.59**	**−4.04**	−0.65			
$\boldsymbol{\pi}^{(5)}$	−0.31	**−5.59**	−0.98				
$\boldsymbol{\pi}^{(6)}$	**−3.56**	−1.50					
$\boldsymbol{\pi}^{(7)}$	1						

This table reports the results of applying the variance ratio test sequentially to the series $\boldsymbol{\pi}^{(J)}$, $J=1,2,\ldots,7$. Bold values reject the null hypothesis of no serial correlation at a 95% confidence level. The first row contains the rescaled test statistics for $\pi_t^{(1)}$, the second row contains the test statistics for $\pi_t^{(2)}$ and so on. Observe that the component at level of persistence $k=1,2,\ldots$ extracted from $\pi_t^{(J)}$ coincides with the component at level of persistence $j=J+k$ of the original time series.

reports the statistics $\hat{\xi}_j^{(J)}$ for each of the series $\pi_t^{(J)}$, $J=1,\ldots,7$, with bold values denoting rejection at the 95% confidence level.[14] The first row contains the rescaled test statistics for $\pi_t^{(1)}$, the second row contains the test statistics for $\pi_t^{(2)}$, and so on. To determine the maximum degree of persistence, we search for the first row J that does not contain bold values, that is, the first J for which we cannot reject $\pi_t^{(J)}$ being white noise. The test statistics in Table 3 rejects $\pi_t^{(J)}$ is white noise for $J=1,\ldots,6$.

In particular, our test shows that it is suboptimal to analyze only the first five components $g_t^{(j)}$ of the original series, that is, the ones with half-life, and hence persistence, not exceeding eight years. This is so because $\pi_t^{(5)}$, which incorporates the fluctuations of the time series with persistence greater than $2^5=32$ quarters, is not white noise as highlighted by the test statistics $\hat{\xi}_k^{(5)}$, $k=2$. The test statistics $\hat{\xi}_k^{(6)}$, $k=1$, associated with $\pi_t^{(6)}$ hints also at the presence of even longer fluctuations with half-life between 16 and 32 years. In this case, however, the evidence is not as strong as the one for $\pi^{(5)}$ given the sparsity of data. Finally, the value taken by the test statistics applied to $\pi_t^{(7)}$ denotes that no further information can be extracted because the entire sample has been saturated. In summary, our evidence strongly supports the view that the persistence of consumption growth is not exhausted at business-cycle frequencies.

14 Specifically, our test statistics $\hat{\xi}_j^{(J)}$ follows from (12) upon replacing $\hat{\xi}_j$ with with the ratio of the sample variance of the decimated component $\mathbf{g}^{(\mathbf{j})}$ to the sample variance of the series $\pi_t^{(J)}$, that is,

$$\hat{\xi}_j^{(J)} = \frac{2^{j-J}\cdot(\mathbf{g}^{(\mathbf{j})})^{\top}\mathbf{g}^{(\mathbf{j})}}{\left(\boldsymbol{\pi}^{(J)}\right)^{\top}\boldsymbol{\pi}^{(J)}}$$

for $j>J$. This is so because the recursive definition of components in (3) implies that the component at level of persistence $k=1,2\ldots$ extracted from $\pi_t^{(J)}$ coincides with the component at a level of persistence $j=J+k$ of the original time series.

Based on our sequential criterion, we extract and analyze the components of consumption growth $g_t^{(j)}$, $j=1,\ldots,7$. Because the first component $g_t^{(1)}$ resembles clearly a statistical (random) noise, we identify it with a contemporaneous i.i.d consumption shock and therefore we concentrate our analysis on the components with level of persistence $j=2,\ldots,7$. To investigate the existence of reasonable economic proxies for these consumption components, we search for time series that are economically significant, characterized by a half-life close to the one of the components they are to proxy for, and significantly correlated with such components.

The second and third components $g_t^{(2)}$ and $g_t^{(3)}$ capture fluctuations with a half-life between one-half and two years. One way to identify these two components with observable economic factors is to follow the lead of Jagannathan and Wang (2007) and Moller and Rangvid (2012), who analyze the ability of the fourth-quarter consumption growth rate to predict expected excess returns on stocks. Importantly, this variable aims at capturing economic and financial choices happening with yearly frequency. Figure 3 reports the series $g_t^{(2)}$ and the one used by Moller and Rangvid (2012). Quite remarkably the correlation between these two series is 0.60 in our sample period. Moreover, if the component $g_t^{(3)}$ is added to the component $g_t^{(2)}$, the correlation rises to 0.72, thus hinting at the presence of a common high-frequency factor with a half-life of about one year.[15] Interestingly, this half-life is right at the boundary between the fluctuations' intervals captured by the second and third component.

The intuition behind Jagannathan and Wang (2007) and Moller and Rangvid (2012) is that due to cultural and institutional features (such as Christmas, end-of-year bonuses and the tax consequences of capital gains and losses) consumption and investment decisions are aligned in the fourth quarter. Our second and third components taken together seem to be good candidates to capture this alignment.

As for the fourth and fifth components $g_t^{(4)}$ and $g_t^{(5)}$, taken together they capture fluctuations in consumption growth lasting from two to eight years. Since Burns and Mitchell (1946), it is widely accepted that the frequencies of business-cycle fluctuations in economic activity belong to this interval. To identify $g_t^{(4)}$ and $g_t^{(5)}$, we look therefore at indicators of the business cycle, such as the term spread, that is, the slope of the Treasury yield curve defined as the difference between the ten-year constant-maturity yield and the three-month constant-maturity Treasury yield, and the Baa-Aaa credit spread. It is well-documented that movements in these variables are related to the dynamics of the business cycle. In particular, the term spread and the default spread are low around business-cycle peaks and high near troughs (see Fama and French 1989; Estrella and Hardouvelis 1991). Consistent with this idea, we find a strong

15 We also use the fourth-quarter year-over-year (Q4-Q4) consumption growth measure of Jagannathan and Wang (2007). In this case the correlation with the second component $g_t^{(2)}$ is 42% and raises to 50% when $g_t^{(3)}$ is added.

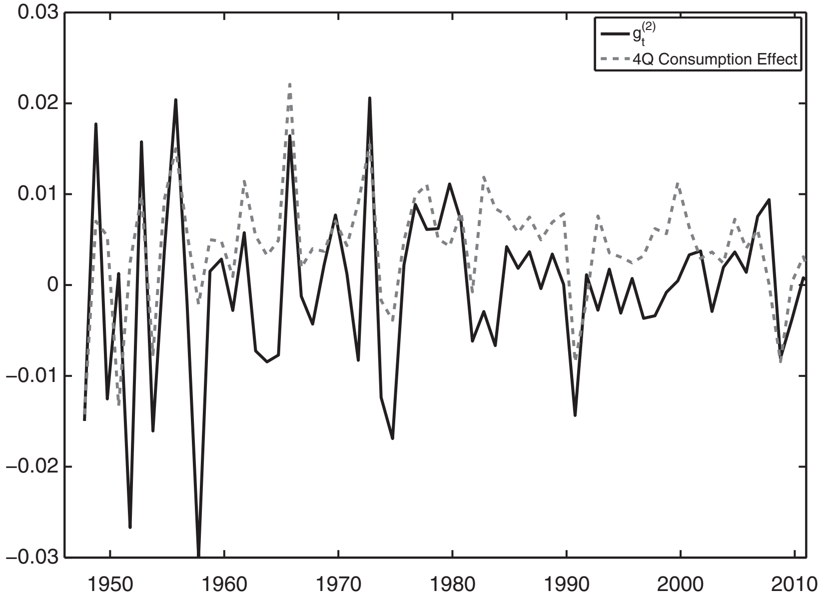

Figure 3
Consumption persistence and the fourth quarter effect
This figure displays the component with level of persistence $j=2$ of consumption growth, $g_t^{(2)}$, along with the real consumption from the third quarter of a calendar year to the fourth quarter as suggested in Moller and Rangvid (2012).

negative correlation between the sum of the fourth and fifth components of consumption growth and these indicators. In particular, this correlation ranges from a negative 30% in the case of the term spread to 45% in the case of the default spread. This fact is documented in Figure 4.

We now turn to the sixth component of consumption growth, $g_t^{(6)}$, which is a slow-moving series with a half-life of about eight years. To search for a valid proxy for this component, we follow the literature on long-run risk with production (see, e.g., Kaltenbrunner and Lochstoer 2010; Croce 2010) and investigate whether shocks to productivity growth can explain these persistent fluctuations in consumption. Figure 5 plots the sixth component of consumption growth together with the sixth component filtered out of the multifactor productivity growth index (TFP). The correlation between the two series is a sound 0.61, giving further support to the idea that highly persistent time-variation in (components of) consumption growth reflects very persistent technological shocks.

Figure 5 shows that the sixth component filtered out of TFP mimics the behavior of U.S. output growth over the postwar period. In particular, the sustained high-output growth during the early to late 1960s was followed by output growth low on average (from the early 1970s to the early 1980s), which in

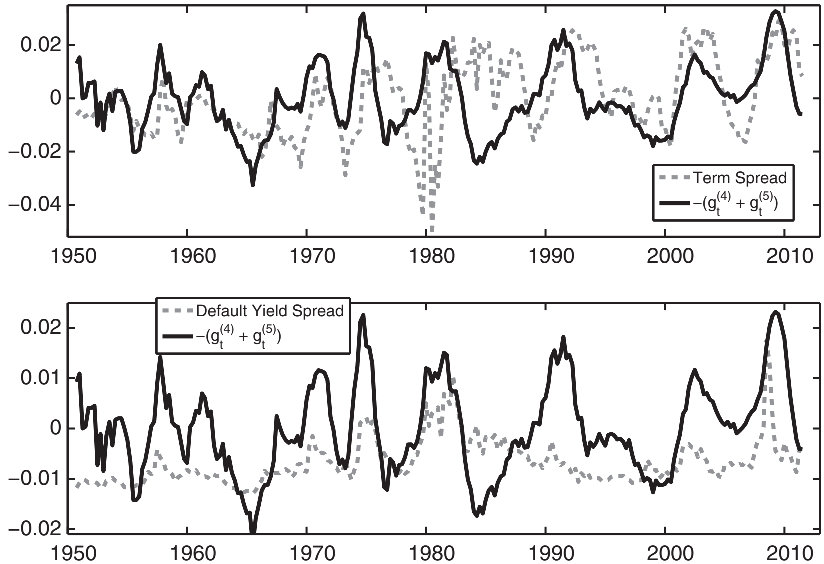

Figure 4
Consumption persistence and business-cycle indicators
This figure displays the sum of the components with level of persistence $j=4,5$ of consumption growth, $g_t^{(4)}$ and $g_t^{(5)}$, along with the the term spread (top panel) and the Baa-Aaa credit spread (bottom panel). We use $2^4/4=4$ years of data at the beginning of the sample to initialize the filtering procedure.

turn was followed by a return to strong growth since the mid-1990s. Therefore, the fluctuations in the sixth component of TFP growth, which have an half-life in between eight and sixteen years, represent a good proxy for the transitions of the U.S. economy from periods of robust growth to periods of relative stagnation. This evidence is strongly in line with Comin and Gertler (2006), who find support for significant medium-frequency oscillations in the U.S. economy corresponding to frequencies between 32 and 200 quarters, and with Garleanu, Panageas, and Yu (2012), who show that predictable components of consumption that occur at cycles between ten and fifteen years are due to the presence of large infrequent embodied technology shocks.

As for the identification of the seventh component, we look at demographic trends as a possible economic proxy. Importantly, live births in the United States have featured alternating twenty-year periods of boom and busts, and therefore are consistent with the half-life of our seventh component (see, e.g., Geanakoplos, Magill, and Quinzii 2004). In the full sample, the correlation between the seventh component of consumption growth and the ratio of middle-aged population to population of young adults (the so-called MY ratio) is equal to 0.3. The demographic variable still leaves place for unexplained variability in the slow-moving component of consumption growth. Also, the sparsity of

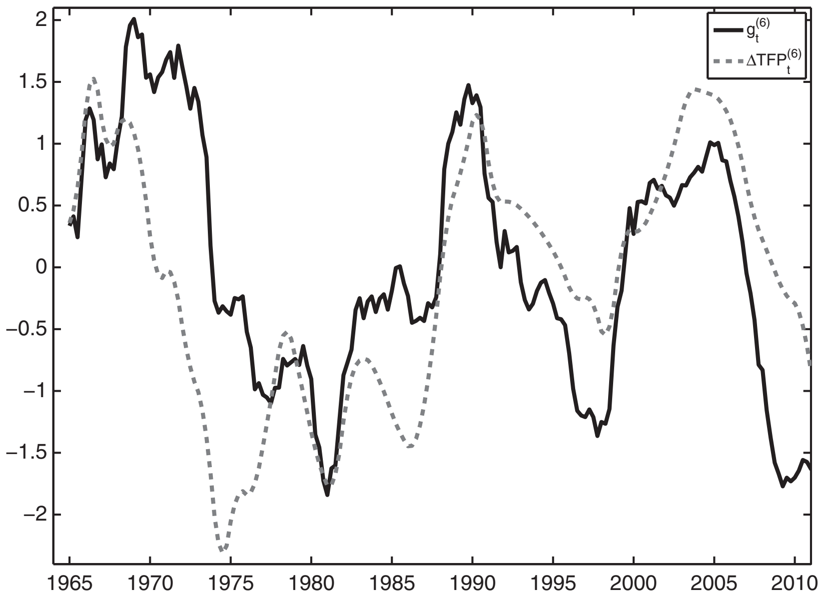

Figure 5
Consumption persistence and total factor productivity
This figure displays the component with level of persistence $j=6$ of consumption growth, $g_t^{(6)}$, and the component of total factor productivity, $\Delta TFP_t^{(6)}$. We use $2^6/4=16$ years of data at the beginning of the sample to initialize the filtering procedure.

the data in this case makes the identification of this component less strong than the previous ones.

In summary, three main economic drivers of the persistence in consumption growth emerge from our empirical analysis. As expected, we find support for one high-frequency component with a yearly half-life and for one that correlates with business-cycle indicators. Along with these two drivers, however, we find strong evidence for a persistent component in consumption with fluctuations occurring over a longer time frame than the one typically considered in conventional business-cycle analysis. Whereas conventional business-cycle detrending methods would remove this component from the analysis by sweeping it into the trend, our persistence-based decomposition brings it to life. In the next section, we show that this persistent component plays a crucial role in financial valuation.

3. Persistence Heterogeneity, Predictability, and Long-Run Valuation

The ability of the price-dividend ratio to predict consumption growth is a key requirement of the long-run risk literature (see, e.g., Bansal and Yaron 2004; Beeler and Campbell 2012). This requirement, however, is empirically rejected.

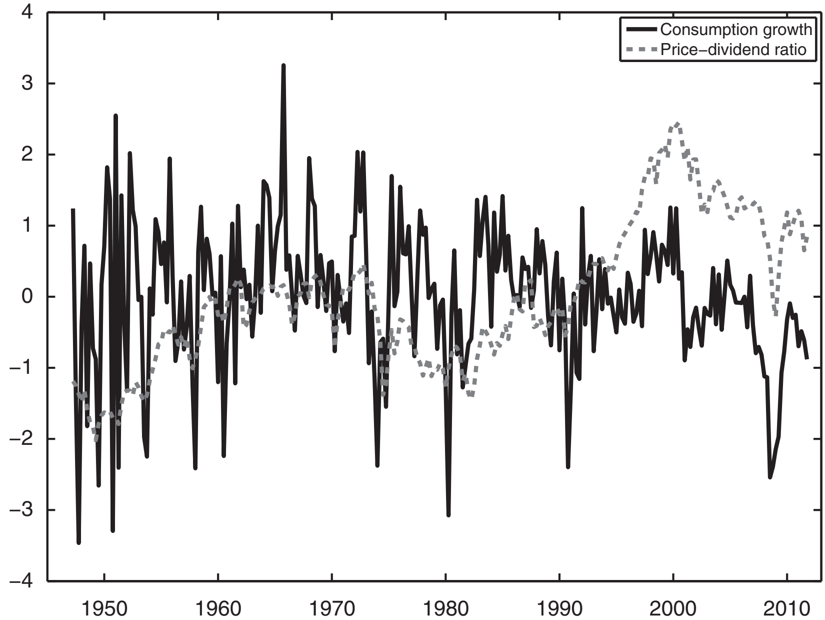

Figure 6
Consumption growth and the price-dividend ratio
This figure displays the series of U.S. consumption growth (nondurables and services) from the Bureau of Economic Analysis and the log price-dividend ratio (dashed line).

On the one hand, in fact, the price-dividend is (close to) a unit root process, as it is well-documented in the literature (see, e.g., Torous, Valkanov, and Yan 2004; Campbell and Yogo 2006; Lettau and Nieuwerburgh 2008). On the other hand, consumption growth closely resembles a white noise and therefore it does not share the high persistence of the price-dividend ratio. This point is synthesized in Figure 6, which displays the very different statistical behavior of the demeaned price-dividend and consumption growth series. The figure shows that over the sample 1947Q2–2011Q4 the price-dividend ratio has crossed its mean value much less often than consumption growth. The intervals between crossings for the price-dividend ratio range from one year to twenty years, the twenty-year interval being the one falling between 1950 and 1970.[16] The persistence of consumption growth, however, is only moderate; the half-life of consumption growth shocks is one year.[17]

[16] On a different sample Campbell and Shiller (2001) report that the price-dividend ratio has crossed its mean value only 29 times since 1872.

[17] Similarly, Paseka and Theocharides (2010) find that the persistence of the latent mean consumption growth corresponds to a half-life of about 1.3 years for the 1934–2005 period.

To reconcile this empirical evidence with the long-run risk framework, we exploit the fact, documented in the previous section, that consumption growth contains layers with different levels of persistence. The highly persistent components indeed contribute for a very small fraction of the total volatility of aggregate consumption growth, and yet they are predictable by the highly persistent components of the financial ratios. Therefore, the contemporaneous presence in the same time series of components with different level of persistence and volatility hides the predictability relation that holds for specific components.

3.1 Predictability of consumption and dividend growth under persistence heterogeneity

To test the ability of the components of financial ratios to predict the components of consumption growth with the same level of persistence, we run the following regressions:

$$g^{(j)}_{t+2^j} = \beta_{0,j} + \beta^{g}_{1,j} z^{(j)}_{m,t} + \varepsilon^{(j)}_{t+2^j} \tag{13}$$

$$g^{(j)}_{t+2^j} = \tilde{\beta}_{0,j} + \tilde{\beta}^{g}_{1,j} z^{(j)}_{a,t} + w^{(j)}_{t+2^j}$$

where $z^{(j)}_{a,t}$ and $z^{(j)}_{m,t}$ denote the components with level of persistence j of the price-consumption and price-dividend series, respectively.[18,19] We remark that, differently from Beeler and Campbell (2012), who focus on the aggregate time series, we analyze instead predictability at different levels of persistence. Results for the quarterly sample are reported in Tables 4 and 5. Table 4 shows that for the consumption components with levels of persistence $j=2,6$, the coefficients on the price-dividend ratio are statistically significant at the 5% level. The sixth component accounts for a great part of the variation in the future consumption growth at the corresponding scale, with the R^2 being 16%. In summary, the components of the price-dividend ratio that actually lead the corresponding consumption growth components have cycles whose length belongs to the intervals $[1/2, 1]$ and $[8, 16]$, measured in years.

[18] The component-wise regressions (13) and (14) can be related to the forward-backward regressions first suggested in Bandi and Perron (2008). Bandi et al. (2013) formally explore the link between these two approaches.

[19] Regressions (13) and (14) are estimated using the full sample instead of the decimated components. The translation invariance property of the dynamics (see footnote 2) guarantees that the OLS coefficients are unbiased. The residuals, however, can be correlated over time. To address this fact we compute the standard errors using heteroscedasticity and autocorrelation consistent (HAC) estimators. Because the choice of bandwidth for HAC estimators depends on the assumed correlation and heteroscedasticity structure, if one knows the pattern of correlation and heteroscedasticity, formulas that impose this knowledge can work better in small samples. Given our decomposition of a time series, we know that spurious autocorrelation at level of persistence j emerges as a result of the $2^j - 1$ overlapping data. Analogously to Cochrane and Piazzesi (2005) we therefore ignore conditional heteroscedasticity, and we impose the idea that error correlation is due only to overlapping observations of homoscedastic forecast errors. In particular, we use a bandwidth equal to the overlap $2^j - 1$ at level of persistence j. Alternatively, we also use the Hansen and Hodrick (1980) estimator that corrects “nonparametrically” for arbitrary error correlation and conditional heteroscedasticity. Conclusions are almost identical to the ones reported in the main text.

Table 4
Predictability of consumption components by the price-dividend ratio

	Persistence level j						
Variable	1	2	3	4	5	6	7
$pd_t^{(j)}$	0.34	**0.69**	0.49	−0.20	0.17	**0.34**	0.06
	(0.67)	(2.01)	(1.27)	(−0.46)	(0.45)	(2.46)	(1.13)
	[0.00]	[0.03]	[0.03]	[0.01]	[0.02]	[0.16]	[0.11]

This table reports the results of predictive regressions of the components of consumption growth $g_{t+2^j}^{(j)}$ on the components of the (log) price-dividend ratio $pd_t^{(j)}$. For each regression, the table reports OLS estimates of the regressors, Hansen and Hodrick corrected t-statistics in parentheses, and adjusted R^2 statistics in square brackets. Significant coefficients at the 5% level are highlighted in bold. The estimated slope coefficients are multiplied by 100. The effective sample is quarterly and spans the period 1947Q2–2011Q4.

Table 5
Predictability of consumption components by the price-consumption ratio

	Persistence level j						
Variable	1	2	3	4	5	6	7
$pc_t^{(j)}$	0.38	0.41	**0.41**	−0.05	0.27	**0.21**	0.05
	(1.09)	(1.40)	(2.11)	(−0.19)	(1.00)	(1.98)	(0.85)
	[0.00]	[0.02]	[0.04]	[0.00]	[0.08]	[0.08]	[0.07]

This table reports the results of predictive regressions of the components of consumption growth $g_{t+2^j}^{(j)}$ on the components of (log) price-consumption ratio $pc_t^{(j)}$. For each regression, the table reports OLS estimates of the regressors, Hansen and Hodrick corrected t-statistics in parentheses, and adjusted R^2 statistics in square brackets. Significant coefficients at the 5% level are highlighted in bold. The estimated slope coefficients are multiplied by 100. The effective sample is quarterly and spans the period 1947Q2–2011Q4.

Table 5 shows that the component of consumption growth with level of persistence $j=6$ is predicted also by the corresponding component of the price-consumption ratio. When using the price-consumption ratio on the right-hand side, moreover, the component with level of persistence $j=3$ becomes significant at the 5% level.

In summary, on the high-frequency part of the spectrum, the second component is found to be predictable when using the price-dividend ratio, whereas the third component is significant when using the price-consumption ratio. On the low-frequency part, instead, the sixth component is the only one that is significant, and it is so both for the price-dividend and the price-consumption ratios. As for the components with degree of persistence $j=1,4,5$, they are statistically insignificant in the regressions (13).

A second set of predictive regressions is obtained by examining at the ability of the financial ratios to forecast cash flows. In particular, in extending the standard long-run risk model to our persistence heterogeneity framework, one would expect that those components of financial ratios that are useful in forecasting the components of consumption growth should also forecast variations at the same levels of persistence in cash flows. We test these

Table 6
Predictability of dividend growth components by the price-dividend ratio

	Persistence level j						
Variable	1	2	3	4	5	6	7
$pd_t^{(j)}$	3.11 (1.66) [0.01]	2.81 (1.85) [0.03]	**4.64** (2.82) [0.11]	2.67 (1.39) [0.07]	0.27 (0.13) [0.00]	**1.30** (2.28) [0.14]	0.77 (1.17) [0.14]

This table reports the results of predictive regressions of the components of dividend growth $gd_{t+2^j}^{(j)}$ on the components of (log) price-dividend ratio $pd_t^{(j)}$. For each regression, the table reports OLS estimates of the regressors, Hansen and Hodrick corrected t-statistics in parentheses, and adjusted R^2 statistics in square brackets. Significant coefficients at the 5% level are highlighted in bold. The estimated slope coefficients are multiplied by 100. The sample is quarterly and spans the period 1947Q2–2011Q4.

implications by running the following set of regressions:

$$gd_{t+2^j}^{(j)} = \beta_{0,j}^{gd} + \beta_{1,j}^{gd} z_{m,t}^{(j)} + \nu_{t+2^j}^{(j)} \tag{14}$$
$$gd_{t+2^j}^{(j)} = \tilde{\beta}_{0,j}^{gd} + \tilde{\beta}_{1,j}^{gd} z_{a,t}^{(j)} + \eta_{t+2^j}^{(j)} \ ,$$

where for the left-hand side we use cash dividends following much of the earlier literature (see, for example, Cochrane 1992). The results reported in Table 6 show that the components of dividend growth that are predictable by the financial ratios are those with levels of persistence $j=2,3,6$, in line with the results found for the components of consumption growth.[20]

The evidence presented in this section shows that consumption growth does contain cyclical components that are predictable by financial ratios. Coupling this componentwise predictability with the evidence in Section 2.2, we can conclude that the components of consumption growth that are predictable are also highly correlated with well-known structural drivers of consumption variability. On the longest side, long-run productivity growth is correlated with the component describing consumption growth variations that occur on time-scales that range between eight and sixteen years. On the shortest side, we identify a high-frequency predictable component with a yearly half-life with the well-documented fourth-quarter effect. The ultimate relevance of the predictability effects in the components of consumption and dividends is related to the ability of these "thin persistent effects" to generate sizable risk premia. To quantify the contribution of the predictable components to risk premia, we first estimate the intertemporal elasticity of substitution (IES), and we then extend the long-run risk valuation framework of Bansal and Yaron (2004) to account for the persistence heterogeneity in consumption growth.

[20] We report results only for the case in which the regressor is the price-dividend ratio. Conclusions do not change when we use the price-consumption series as the regressor.

3.2 Estimating the IES under persistence heterogeneity

The standard approach to estimate the IES (see Hansen and Singleton 1983) is to run the following regression:[21]

$$r_{f,t} = \alpha_f + \frac{1}{\psi} g_t + \sigma_f u_t \ . \tag{15}$$

Empirical tests of (15) typically find an estimate of ψ lower than one (see, e.g., Hall 1988; Campbell and Mankiw 1990, among many others). This evidence contradicts a basic paradigm in the long-run risk literature that requires an IES greater than one for the small-in-volatility but persistent component to contribute to the equity premium. We resolve this puzzle by employing our decomposition of a time series into layers with different levels of persistence. Coherently with our approach, instead of running the regression in (15), we filter both sides of (15) using the transformation matrix $\mathcal{T}^{(J)}$ introduced in Section 2.1 to obtain the following set of J testable restrictions:[22]

$$r_{f,t}^{(j)} = \frac{1}{\psi} g_t^{(j)} + \sigma_{f,j} u_t^{(j)} \ . \tag{16}$$

In words, this set of relations is constrained by the condition that the coefficient linking the information content of the components of the risk-free rate to those of consumption growth must be the same at all levels of persistence.[23] Compared with (15), the system of regressions (16) mandates to properly account for the heterogeneity in consumption growth generated by the mixture of highly volatile and slowly evolving components. This is exactly what we do when we apply the persistence-based decomposition before running the regressions.

Table 7 displays the results. The first row shows the estimate obtained using the full sample 1947Q2–2011Q4. Remarkably, when we decompose the risk-free rate and consumption growth across the different levels of persistence, the estimate of the IES is significant at standard levels and equal to 5.4. Next, we split the sample into two subsamples of equal sizes (128 data points), 1947Q2–1979Q1 and 1979Q1–2010Q4. The second and third rows show that the above finding is robust in these two subsamples. In particular, these subsample estimates are strongly significant, all above one and close to the value obtained in the full sample.

A potential explanation of these results is that (particularly in postwar quarterly data) the real interest rate is highly volatile relative to predictable variation in consumption growth (see, e.g., Beeler and Campbell (2012)

[21] Alternatively, one can reverse the regression and estimate $g_{t+1} = \beta_0 + \psi r_{f,t+1} + \eta_{t+1}$. However, if ψ is large as it will turn out to be the case in our empirical exercise, then it is better to estimate the equation reported in the text.

[22] In the online Technical Appendix, we derive (15) as an equilibrium condition in a long-run risk framework with heterogeneous persistence.

[23] To estimate the system of equations in (16), we adopt the technique suggested in Fadili and Bullmore (2002), who show that in this case the GLS is theoretically close to the best linear unbiased (BLU) estimator.

Table 7
Estimates of the intertemporal elasticity of substitution (IES)

	$r^{(j)}_{f,t+1} = \alpha_f + \frac{1}{\psi} g^{(j)}_{t+1}$	
Asset	Sample	$\hat{\psi}$
$r_{f,t+1}$	1948Q1–2011Q4	5.54 (3.20)
$r_{f,t+1}$	1948Q1–1979Q4	6.15 (1.90)
$r_{f,t+1}$	1980Q1–2011Q4	4.02 (3.30)

This table displays the EIS estimates using the real risk-free rate. The first row reports the estimate obtained from the full sample (256 data points, and hence the maximum number of components is $J = 8$). In the second and third rows, the sample is split into two subsamples of equal sizes, and the estimate is based on 128 data points, hence the maximum number of components is $J = 7$. All estimates are computed using the decimated persistence-based decomposition as suggested in Fadili and Bullmore (2002).

for a discussion). It would be difficult to estimate properly the IES unless one disentangles the highly volatile and noisy components from the more informative but less volatile ones. We obtain a robust estimate larger than one exactly by disentangling these components. Our estimate, moreover, yields support to a key hypothesis of the long-run risk valuation approach. Its magnitude is in line with the recent findings of Avramov and Cederburg (2012) who, in a similar postwar sample, estimate an IES of 4.5. Our empirical findings are also in agreement with previous studies, for example, Attanasio and Weber (1993), Beaudry and van Wincoop (1996), and Vissing-Jorgensen (2002), who find values for ψ higher than one. Similarly to those papers, we present evidence of the fact that the relation between the real interest rate and consumption growth can be obscured by aggregation problems, and we show that using disaggregated consumption data is key to finding a value for the IES greater than one. Whereas the literature cited above focuses on consumption data disaggregated at cohort level, state level or household level, we suggest instead persistence heterogeneity as a key dimension along which to disaggregate consumption.

3.3 Persistence heterogeneity and the term structure of equity premia

The layers of consumption growth with heterogeneous persistence generate a term structure of equity risk premia. To see this, we consider a Bansal and Yaron (2004) economy in which a representative agent with recursive preferences à la Epstein-Zin (see Kreps and Porteus 1978; Epstein and Zin 1989, 1991) faces a consumption stream g_t whose decimated components $g^{(j)}_t$ follow multiscale autoregressive processes, that is,[24]

$$g^{(j)}_{t+2^j} = \rho_j g^{(j)}_t + \varepsilon^{(j)}_{t+2^j} \ , \tag{17}$$

with the shocks possibly correlated across levels of persistence (for fixed time t), but not across time (for fixed persistence level j). The lag of 2^j units of time between the regressand and the regressor means that cyclical fluctuations at time-scale j forecast the next cycle of length 2^j periods.[25] Consistently, we allow the leverage effect to be different across levels of persistence, that is, the decimated components $gd^{(j)}_t$ of the log dividend growth series gd_t satisfy

$$gd^{(j)}_{t+2^j} = \phi_j g^{(j)}_t + \eta^{(j)}_{t+2^j} \tag{18}$$

with the shocks $\eta^{(j)}_{t+2^j} \sim N\left(0, \sigma^2_j\right)$ uncorrelated both across time (for fixed persistence level j) and across levels of persistence (for fixed time t), and independent from the shocks $\varepsilon^{(j)}_{t+2^j}$ to the consumption growth components.

To determine the term structure of risk premia, we first recall that the Euler equation for the representative agent is

$$E_t\left[e^{\theta \log \beta - \frac{\theta}{\psi} g_{t+1} + (\theta-1) r_{a,t+1} + r_{i,t+1}}\right] = 1 \ , \tag{19}$$

where $r_{a,t+1}$ is the log return on the claim that distributes a dividend equal to aggregate consumption, $r_{i,t+1}$ is the log return on any asset i, the parameters β, ψ, γ measure the subjective discount factor, the intertemporal elasticity of substitution and risk aversion, and $\theta \equiv (1-\gamma)/(1-1/\psi)$.

To solve the model, we log-linearize the return on the consumption claim, $r_{a,t+1}$, and the return on the market portfolio, $r_{m,t+1}$, à la Campbell and Shiller (1988) to express them in terms of the log price-consumption and log price-dividend ratio $z_{a,t}$, $z_{m,t}$ as follows:

$$\begin{aligned} r_{a,t+1} &= \kappa_0 + \kappa_1 z_{a,t+1} - z_{a,t} + g_{t+1} \\ r_{m,t+1} &= \kappa_{0,m} + \kappa_{1,m} z_{m,t+1} - z_{m,t} + gd_{t+1} \ , \end{aligned} \tag{20}$$

---

[24] We focus our attention on the case in which second moments are constant. By assuming constant volatilities of log consumption growth and log dividend growth, we are able to better concentrate on our primary research question. That is, whether fluctuations in the conditional mean of consumption and dividend growth are indeed priced. Tamoni (2011) uses the persistence-based decomposition to study the implications of persistence heterogeneity in consumption volatility for the risk premium dynamics.

[25] In fact the $g^{(j)}_t$'s are autoregressive processes of order one up to the scale 2^j necessarily to avoid spurious correlation, a fact highlighted by rewriting (17) as

$$. \qquad g^{(j)}_{(k+1)\cdot 2^j} = \rho_j g^{(j)}_{k\cdot 2^j} + \varepsilon^{(j)}_{(k+1)\cdot 2^j} \ .$$

and we conjecture a linear relation between the persistent components $z_{a,t}^{(j)}$, $z_{m,t}^{(j)}$ of the financial ratios and the consumption components $g_t^{(j)}$:[26]

$$z_{a,t}^{(j)} = A_{0,j} + A_j g_t^{(j)} \tag{21}$$

$$z_{m,t}^{(j)} = A_{0,j}^m + A_j^m g_t^{(j)} \ .$$

To determine the coefficients A_j, A_j^m in terms of the parameters of the model, one first uses (20), (21) and the law of motions (17) and (18) to express the returns $r_{a,t+1}$, $r_{m,t+1}$ in terms of the consumption and dividend components $g_t^{(j)}, gd_t^{(j)}$ and the innovations $\varepsilon_{t+2^j}^{(j)}$ and $\eta_{t+2^j}^{(j)}$. Plugging the expressions obtained in this way into the Euler equation and denoting by $\underline{A}$, respectively $\underline{A_m}$ the column vectors with entries A_j, A_j^m, $j=1,...,J$, the method of undetermined coefficients yields two systems of equations in the unknowns $\underline{A}$, respectively $\underline{A_m}$, whose solutions are:

$$\underline{A} = \left(1 - \frac{1}{\psi}\right)(\mathbb{I}_J - \kappa_1 M)^{-1} M \underline{1} \tag{22}$$

$$\underline{A_m} = (\mathbb{I}_J - \kappa_{1,m} M)^{-1} M \left(\underline{\phi} - \frac{1}{\psi}\underline{1}\right) , \tag{23}$$

where $\mathbb{I}_J$ is the $(J \times J)$ identity matrix, M is a J-dimensional diagonal matrix with the opposite of the persistence parameters ρ_j on the diagonal, and $\underline{\phi}$ is a column vector whose entries are the leverage parameters ϕ_j.[27]

Aggregating over persistence levels the relation imposed by (21) between financial ratio components and consumption components and using (17) and (18) yields the following linear decomposition of the aggregate financial ratios in terms of cash flows components:

$$z_{a,t} - \pi_a = \sum_{j=1}^{J} z_{a,t}^{(j)} = \sum_{j=1}^{J} c_j E_t\left[g_{t+2^j}^{(j)}\right] \tag{24}$$

$$z_{m,t} - \pi_m = \sum_{j=1}^{J} z_{m,t}^{(j)} = \sum_{j=1}^{J} c_j^m E_t\left[gd_{t+2^j}^{(j)}\right] ,$$

where $c_j \equiv A_j/\rho_j$, $c_j^m \equiv A_j^m/\phi_j$ and the constants π_a, π_m capture the unconditional mean of the financial ratios. Equation (24) decomposes the financial ratios in a weighted sum of cash flow expectations over forecasting

[26] The fact that $z_{a,t}^{(j)}, z_{m,t}^{(j)}$ are the components of the *demeaned* price-consumption and price-dividend ratios implies that $\sum A_{0,j} = 0$ and $\sum A_{0,j}^m = 0$. Moreover as in Bansal, Yaron, and Kiku (2012), in our model the log-linearization parameters κ_1 and $\kappa_{1,m}$, are determined endogenously because the long-run means π_a, π_m depend on κ_1, $\kappa_{1,m}$ which in turns are functions of the price-consumption and price-dividend ratio mean.

[27] All the details behind these computations are available in the online Technical Appendix.

horizons determined by the degrees of persistence present in the consumption and dividends series. This decomposition hints at an alternative approach to studying the contribution of cash flows to the variation over time of financial ratios. In particular, whereas most of the literature has focused on the importance of the timing of dividend payments to determine the value of financial ratios (see, for instance, Van Binsbergen, Brandt, and Koijen (2012); Van Binsbergen et al. (2012), who look at dividend strips of different maturities), our decomposition suggests instead studying the contribution of the components with different levels of persistence disentangled from the full dividend stream.

In our model with persistence heterogeneity, the equity premia for the consumption claim asset, $r_{a,t+1}$, and for the market portfolio, $r_{m,t+1}$, take the form:[28]

$$E_t[r_{a,t+1}-r_{f,t}]+0.5\sigma^2_{r_{a,t}}=\gamma\,\underline{1}^{\intercal}\mathbf{Q}\underline{1}+\kappa_1\underline{\lambda}_{\varepsilon}^{\intercal}\mathbf{Q}\underline{A}\ , \tag{25}$$

$$E_t[r_{m,t+1}-r_{f,t}]+0.5\sigma^2_{r_{m,t}}=\kappa_{1,m}\underline{\lambda}_{\varepsilon}^{\intercal}\mathbf{Q}\underline{A_m}\ , \tag{26}$$

where $\underline{\lambda}_{\varepsilon}\equiv\kappa_1(1-\theta)\underline{A}$ and $\mathbf{Q}$ is the variance-covariance matrix of the innovations $\varepsilon^{(j)}_{t+2^j}$ to the components of consumption growth. The expression for the equity premium shows that, whereas the standard long-run risk framework prices only a single, persistent shock to *aggregate* consumption growth, the presence of persistence heterogeneity permits an analysis of the pricing impact of the shocks driving the *components* of consumption growth, each characterized by its own level of persistence. The risk compensations for these innovations are collected in the vector $\underline{\lambda}_{\varepsilon}$, whereas the market return's exposures to these shocks are represented by the vector $\mathbf{Q}\underline{A_m}$. Importantly, the exposure of the market return is determined simultaneously by the size of the shocks as measured by their instantaneous volatility, captured by $\mathbf{Q}$, and by their persistence. To obtain the entire term structure of risk-return trade-offs it is therefore key to decompose the aggregate shocks that impinge an economy along the two competing dimensions of volatility and persistence.[29]

To quantify the effect and the contribution of the predictable components in consumption and dividends, we compute the equity premium for two different levels of the IES, $\psi=2.5$ and $\psi=5$. The first is the value estimated in Bansal, Yaron, and Kiku (2012). The second value is coherent with the estimate obtained in the previous section. For the risk aversion parameter we consider the alternative values of $\gamma=5$, close to the postwar estimate in Ghysels, Santa-Clara, and Valkanov (2004), and $\gamma=7.5$, that has been used in the calibration of Bansal and Yaron (2004).[30]

[28] See again the online Technical Appendix.

[29] This is similar in spirit to Hansen and Scheinkman (2009), Borovicka et al. (2011), and Lettau and Wachter (2011), who look at the entire term structure of risk prices.

[30] We set $\kappa_{1,m}$ and κ_1 to 0.988, a value consistent with the magnitude of the mean of the financial ratios in our sample and with magnitudes used in Campbell and Shiller (1988).

To estimate the entries A_j, A_j^m of the vectors $\underline{A}$ and $\underline{A_m}$, we first observe that by solving (21) for the consumption growth components and plugging into the right-hand side of the consumption dynamics (17) we obtain

$$g_{t+2^j}^{(j)} = -\frac{\rho_j A_{0,j}}{A_j} + \frac{\rho_j}{A_j} z_{a,t}^{(j)} + \varepsilon_{t+2^j}^{(j)}$$

$$g_{t+2^j}^{(j)} = -\frac{\rho_j A_{0,j}^m}{A_j^m} + \frac{\rho_j}{A_j^m} z_{m,t}^{(j)} + \varepsilon_{t+2^j}^{(j)} ,$$

which is exactly the set of predictive regressions discussed in Section 3.1. Denoting therefore by $\widehat{\beta_{1,j}^g} = \widehat{\left(\frac{\rho_j}{A_j}\right)}$ and $\widehat{\tilde{\beta}_{1,j}^g} = \widehat{\left(\frac{\rho_j}{A_j^m}\right)}$ the slopes estimates of (13), and given an estimate of the persistence parameter $\hat{\rho}_j$, we estimate the sensitivities A_j, A_j^m via

$$\widehat{A_j} = \hat{\rho}_j / \widehat{\left(\frac{\rho_j}{A_j}\right)}, \quad \widehat{A_j^m} = \hat{\rho}_j / \widehat{\left(\frac{\rho_j}{A_j^m}\right)} .$$

Estimates $\hat{\rho}_j$ of the autoregressive coefficients, together with the R^2, are shown in Table 8. The off-diagonal terms of the estimated variance-covariance matrix $\widehat{\mathbf{Q}}$ of the innovations $\varepsilon_{t+2^j}^{(j)}$ are small and insignificant, that is, $\widehat{\mathbf{Q}}$ is approximately a diagonal.[31] Therefore, we find no evidence of interaction across components with different degrees of persistence even though they are allowed to be correlated. This allows us to approximate the equity premium in (26) as follows

$$E_t[r_{m,t+1} - r_{f,t}] + 0.5\sigma_{r_{m,t}}^2 \approx \kappa_{1,m}\kappa_1(1-\theta)\sum_{j=1}^{J} \hat{A}_j \hat{Q}_{jj} \hat{A}_j^m .$$

With these parameter values at hand, we compute the equity premium disaggregated across different levels of persistence. Results are reported in Tables 9 and 10. We first observe that the equity premium is close to zero at level of persistence $j=1$. This is consistent with the fact that, as Table 8 shows, the first component is not persistent and as such bears no risk so that it must command no premium. The predictive regressions discussed in Section 3.1 (see, in particular, Tables 4 and 6) show that the only components of consumption and dividends that are predictable by financial ratios are those at level of persistence $j=2,3,6$, that is, those with an half-life in the interval, measured in years, [0.5, 1), [1, 2), and [8, 16), respectively. In other words, the risk exposures and the prices of risk of the other components are statistically insignificant. For this reason, we concentrate our discussion only on the premium for the predictable components. We start by the low-frequency persistent component $j=6$ because its contribution to the equity premium is

[31] The Pearson's correlation test indicates that almost all of the consumption growth components residuals are pairwise uncorrelated. The correlation is significant at standard levels only between the sixth and seventh components, yielding a Pearson's p-values of about 0.07. Similar results are obtained using the Spearman rank-order correlation coefficients test.

Table 8
Multiscale autoregressive process estimates

Variables	ρ_j	HL (years)	$\bar{R}^2$
$g^{(1)}_{t+2}$	−0.00 (−0.08)	–	[0.00]
$g^{(2)}_{t+2^2}$	**−0.14** (−2.37)	0.4	[0.02]
$g^{(3)}_{t+2^3}$	**−0.13** (−2.13)	0.7	[0.02]
$g^{(4)}_{t+2^4}$	−0.11 (−1.75)	–	[0.01]
$g^{(5)}_{t+2^5}$	**−0.16** (−2.66)	3.1	[0.03]
$g^{(6)}_{t+2^6}$	**−0.21** (−3.19)	7.1	[0.05]
$g^{(7)}_{t+2^7}$	**0.14** (4.40)	10.8	[0.13]

This table reports the estimation results of the multiscale autoregressive system. For each level of persistence $j \in \{1,\ldots,7\}$, we run a regression of the consumption growth component $g^{(j)}_{t+2^j}$ on its own lagged component $g^{(j)}_t$. For each regression, the table reports OLS estimates of the regressors, Hansen and Hodrick corrected t-statistics in parentheses, and adjusted R^2 statistics in square brackets. Bold coefficients denote statistical significance at the 5% level. Half-lives (in annual units) are obtained by $HL = \cancel{4/2^j} \times \log(2)/\log(|\rho_j|)$. The sample is quarterly and spans the period 1947Q2–2010Q4.

the largest. This component commands a premium that ranges from 1% for the case in which $\psi=5$ and $\gamma=5$ to a maximum of almost 2% for the case in which $\psi=2.5$ and $\gamma=7.5$. Considering now the high-frequency components with level of persistence $j=2,3$, we note that they command a smaller, but by no means trivial, premium. In particular, the premium to each of this component is about 0.5% for the case in which $\psi=5$ and $\gamma=5$ and raises to 1% for the case in which $\psi=2.5$ and $\gamma=7.5$. In sum, the contributions of the different components deliver an equity premium in the range of 2%–4% per annum, for a moderate amount of the risk aversion and an IES greater than one.[32]

Recalling the identification of the predictable layers of consumption in Section 2.2, our calibration exercise shows that two components, one on the low-frequency side and correlated to long-run productivity growth and one on the high-frequency side and correlated with the fourth-quarter effect, deliver a sizable contribution to the equity premium for a moderate amount of risk

32 To check the accuracy of the equity premium term structure obtained via predictive regressions, we compute the vectors $\underline{A}$ and $\underline{A_m}$ directly from expressions (22) and (23). In particular, we insert in (22) and (23) the estimates of ρ_j, see Table 2, $\phi_j = (\phi_j/A^m_j)/(\rho_j/A^m_j)\cdot\rho_j$ and the value of the IES estimated in Section 3.2 By doing so, we obtain alternative estimates for $\underline{A}$ and $\underline{A_m}$ that are slightly larger, but in reasonable agreement with those obtained via OLS regressions. The term structure of risk premia reported in Tables 9 and 10 makes use of the most conservative OLS estimates.

Table 9
The term structure of risk premia, IES = 5

Panel A: Risk aversion and IES: $\gamma=5$, $\psi=5$

Scale $j=$	Half-life (years)	Q_{jj} (1×10^{-5})	Risk exposure (1×10^{-6})	Risk price	Risk premium (%)
1	0.06	1.06	15.77	1.83	0.00
2	0.4	0.54	117.21	57.57	0.67
3	0.7	0.45	125.45	42.20	0.53
4	1.2	0.32	145.27	302.13	4.38
5	3.1	0.17	195.48	101.51	1.98
6	7.1	0.09	60.21	159.81	0.96
7	10.8	0.02	36.07	456.14	1.64

Panel B: Risk aversion and IES: $\gamma=7.5$, $\psi=5$

Scale $j=$	Half-life (years)	Q_{jj} (1×10^{-5})	Risk exposure (1×10^{-6})	Risk price	Risk premium (%)
1	0.06	1.06	15.77	2.78	0.00
2	0.4	0.54	117.21	87.55	1.03
3	0.7	0.45	125.45	64.18	0.81
4	1.2	0.32	145.27	459.49	6.67
5	3.1	0.17	195.48	154.38	3.02
6	7.1	0.09	60.21	243.04	1.46
7	10.8	0.02	36.07	693.71	2.50

This table reports equity premium (in %) $E_t[r_{m,t+1}-r_{f,t}]$ decomposed by level of persistence. We set $\psi=5$, $\gamma=5$ (Panel A), and $\gamma=7.5$ (Panel B). Risk exposure and risk price are annualized. The annual percentage risk premium is given by the risk exposure times the risk price (multiplied by 100). Highlighted rows denote the predictable components of consumption growth.

Table 10
The term structure of risk premia, IES = 2.5

Panel A: Risk aversion and IES: $\gamma=5$, $\psi=2.5$

Scale $j=$	Half-life (years)	Q_{jj} (1×10^{-5})	Risk exposure (1×10^{-6})	Risk price	Risk premium (%)
1	0.06	1.06	15.77	2.34	0.00
2	0.4	0.54	117.21	73.56	0.86
3	0.7	0.45	125.45	53.92	0.68
4	1.2	0.32	145.27	386.06	5.60
5	3.1	0.17	195.48	129.71	2.53
6	7.1	0.09	60.21	204.20	1.23
7	10.8	0.02	36.07	582.85	2.10

Panel B: Risk aversion and IES: $\gamma=7.5$, $\psi=2.5$

Scale $j=$	Half-life (years)	Q_{jj} (1×10^{-5})	Risk exposure (1×10^{-6})	Risk price	Risk premium (%)
1	0.06	1.06	15.77	2.78	0.01
2	0.4	0.54	117.21	87.55	1.33
3	0.7	0.45	125.45	64.18	1.04
4	1.2	0.32	145.27	459.49	8.65
5	3.1	0.17	195.48	154.38	3.91
6	7.1	0.09	60.21	243.04	1.90
7	10.8	0.02	36.07	693.71	3.24

This table reports equity premium (in %) $E_t[r_{m,t+1}-r_{f,t}]$ decomposed by level of persistence. We set $\psi=2.5$, $\gamma=5$ (Panel A), and $\gamma=7.5$ (Panel B). Risk exposure and risk price are annualized. The annual percentage risk premium is given by the risk exposure times the risk price (multiplied by 100). Highlighted rows denote the predictable components of consumption growth.

aversion. This fact further highlights the importance of our heterogeneous persistence approach to understanding the behavior of the equity premium across different time horizons.

4. Robustness

In this section we conduct some checks to analyze the robustness of our predictive regressions and of our estimate of the IES to the sampling frequency and to the choice of financial ratio. First, to understand whether measurement errors in quarterly consumption data change our conclusions, we use annual rather than quarterly data on consumption and asset prices. In fact, annual data are less susceptible to measurement errors (see, e.g., Bansal, Kiku, and Yaron 2012). We also extend our sample to span the time period from 1930 to 2011. As noted by Bansal, Kiku, and Yaron (2012), this sample covers a wide range of macroeconomic events and various episodes of high turbulence in asset markets that potentially contain additional important information regarding variation in expected consumption growth. Finally, we test the robustness of our results to the case in which the price-earnings ratio is used instead of the price-dividend ratio.

We start by rerunning the regression (13) in Section 3.1 when both consumption growth and the price-dividend are sampled at an annual frequency over the period 1948–2011. The results, reported in Table 11, Panel A, show that with annual data the regressions yield significant slopes only for the levels of persistence $j=1,4$. This is consistent with the results in Section 3.1 using quarterly data. To see this, observe that the component at level of persistence $j=1$ extracted using annual data has a half-life falling between 1 and 2 years, that is, four and eight quarters, which with quarterly data corresponds to the component at level $j=3$. Likewise, the component at level of persistence $j=4$ extracted using annual data captures cycles of length in the interval $[8,16)$ years, that is, $[32,64)$ quarters, which with quarterly data corresponds to the component at level $j=6$. This allows us to conclude that, aside from a change in time unit, going from quarterly to annual data leaves our results unaltered.[33] Also, we note that both the coefficient and the R^2 obtained using annual data are largely in agreement with those using quarterly data, and even more so for components with a high level of persistence. For example, by annualizing the coefficient on the sixth component obtained in the quarterly sample, that is, $0.34\times 4=1.36$, one obtains a result close to the one in the annual data. The rationale behind these results rests on the fact that the component at level of persistence j in the annual sample corresponds to the component with level of persistence $j+2$ in the quarterly sample and, for any given j, formal

[33] Clearly, the second component that was found to be significant using quarterly data cannot be captured by the annual sample because it reflects cyclicality between half-year and a year, which is well below the maximum observation frequency achievable using annual data.

Table 11
Predictability of consumption components by the price-dividend ratio: Annual data

Panel A: 1948–2011

	Persistence level j				
Variable	1	2	3	4	5
	3.74	−0.34	0.50	**1.25**	0.14
$pd_t^{(j)}$	(4.72)	(−0.25)	(0.33)	(2.01)	(0.40)
	[0.23]	[0.00]	[0.01]	[0.14]	[0.04]

Panel B: 1930–2011

	Persistence level j				
Variable	1	2	3	4	5
	4.98	−0.43	0.21	**1.68**	0.03
$pd_t^{(j)}$	(3.53)	(−0.20)	(0.10)	(2.09)	(0.06)
	[0.15]	[0.00]	[0.00]	[0.15]	[0.00]

This table reports the results of predictive regressions of the components of consumption growth on the components of the (log) price-dividend ratio. For each regression, the table reports OLS estimates of the regressors, Hansen and Hodrick corrected t-statistics in parentheses, and adjusted R^2 statistics in square brackets. Significant coefficients at the 5% level are highlighted in bold. The estimated slope coefficients are multiplied by 100. The sample is annual and spans the period 1948-2011 (Panel A) and the period 1930–2011 (Panel B).

correlation tests (not reported) show that the correlation is high and approaches one as the level of persistence increases. In conclusion, data with different frequency of observation (e.g., quarterly vs. annual) do not drive the inference, and any measurement error, if present, vanishes as the level of persistence increases.

We also run the regression (13) at the annual frequency but using the longest available data span, that is, 1930–2011 and report the results in Table 11, Panel B. The results are consistent with those obtained for the quarterly series: The components with level of persistence $j = 1, 4$ are the only statistically significant ones.

We then turn our attention to the robustness of our results to the choice of the financial ratio. Following Bansal, Khatchatrian, and Yaron (2005), we use the price-earnings as an alternative to the price-dividend ratio and then run predictive regression of the components of consumption growth on the respective components of the price-earnings ratio on the annual sample 1948–2011. We report the results in Table 12. Our results show that using the price-earnings ratio instead of the price-dividend ratio does not impact our findings obtained using both quarterly and annual data (see Tables 4 and 11). The sign of the components of consumption growth with level of persistence $j = 1, 4$ is positive, as predicted by our economic model, and estimates have significant robust t-statistics.

In summary, the robustness checks presented above show that the two components with an approximate half-life of one and eight years of both the price-dividend and the price-earnings ratio are useful in predicting the respective components of future consumption growth. Moreover, the results are

Table 12
Predictability of consumption components by the price-earnings ratio: Annual data

	Persistence level j				
Variable	1	2	3	4	5
$pe_t^{(j)}$	**3.95**	0.63	−0.03	**0.72**	0.09
	(4.14)	(0.43)	(−0.02)	(1.91)	(0.18)
	[0.21]	[0.01]	[0.00]	[0.10]	[0.01]

This table reports the results of predictive regressions of the components of consumption growth on the components of (log) price-earnings ratio. Trailing ten-year earnings are used. For each regression, the table reports OLS estimates of the regressors, Hansen and Hodrick corrected t-statistics in parentheses, and adjusted R^2 statistics in square brackets. Significant coefficients at the 5% level are highlighted in bold. The estimated slope coefficients are multiplied by 100. The sample is annual and spans the period 1948–2011.

robust to the inclusion of the prewar period and to the frequency of observations (quarterly versus annual data).

A natural question that arises at this point is that of whether we can understand the long-run fluctuations in the data rather than simply the fluctuations represented by a component with a specific level of persistence j. To do so, we use Equations (1) and (3) to show that when we sum the components with level of persistence greater than $j=6$ to the constant term $\pi_t^{(J)}$ we obtain $\sum_{j=6}^{J} g_t^{(j)} + \pi_t^{(J)} = g_{t-2^6-1,t}$, where $g_{t-2^6-1,t}$ represents the log of consumption growth over 32 quarters. In other words, the sum of all components with level of persistence greater than six produces a smoothed consumption growth series that corresponds to frequencies of 32 quarters and above and is in strict analogy with the medium-frequency component of Comin and Gertler (2006). We plot in Figure 7 this truncated sum of components for the consumption growth and TFP series. The striking comovement patterns between long-run consumption growth and TFP is confirmed by a correlation of about 47%. From a closer inspection of Figure 7, moreover, another interesting point emerges: Even though our truncated sum of components allows for all possible fluctuations whose cycle length exceeds eight years, the cycles are in fact on the order of a decade, and therefore most of the variations in the consumption and total factor productivity growth rates over 32 quarters are indeed captured by the sixth component of such a series.

In essence, the vast majority of the action in terms of long-run fluctuations in consumption and TFP growth rates comes from the sixth component, that is, cycles with a half-life between eight and sixteen years, whereas components with higher half-life seem to be of lesser importance. To further support this statement, Figure 8 displays in the top panel the component of consumption growth with level of persistence $j=6$, $g_t^{(6)}$, along with the long-term growth rate of consumption growth $g_{t-2^6-1,t}$ and in the bottom panel the analogous for the total factor productivity. The sixth component explains 85% and 54% of the long-term consumption and TFP growth rates, respectively. This highlights the predominance of the specific component $j=6$ in explaining the long-term fluctuations in the data.

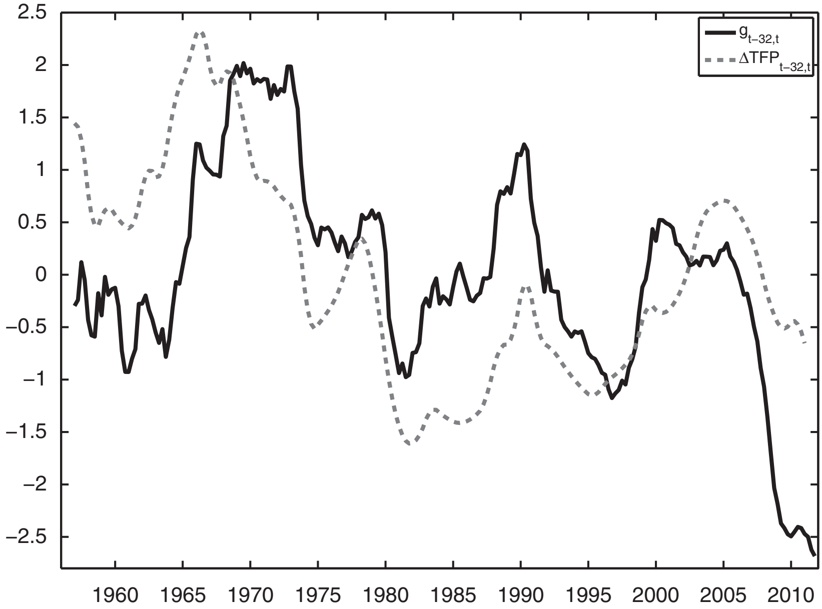

Figure 7
Consumption and total factor productivity at medium-low frequencies
This figure displays the components of consumption growth and total factor productivity with frequency greater than 32 quarters. These components are obtained by summing the components with level of persistence $j \geq 6$ of consumption growth, $\sum_{j=6}^{J} g_t^{(j)} = g_{t-32,t}$ and of total factor productivity, $\sum_{j=6}^{J} \Delta TFP_t^{(j)} = \Delta TFP_{t-32,t}$. These components are related to the Comin and Gertler (2006) medium-frequency components, that is, those components related to a frequency range of 32 to 200 quarters.

Table 13
Estimates of the intertemporal elasticity of substitution (IES): Annual data

	$r_{f,j,t+1} = \alpha_f + \frac{1}{\psi} g_{t+1,j}$	
Asset	Sample	$\hat{\psi}$
$r_{f,t+1}$	1930–2011	2.09 (3.22)

This table displays the EIS estimates using the risk-free rate. The estimate is obtained from the full sample and is computed using the decimated decomposition as suggested in Fadili and Bullmore (2002).

We conclude this robustness section by using the set of equations (16) to estimate the intertemporal elasticity of substitution employing the long annual sample from 1930 to 2011. The point estimate for this long sample, reported in Table 13, is strictly greater than one and statistically significant. This allows us to highlight once again the importance of considering cyclical fluctuations in the time series of interest.

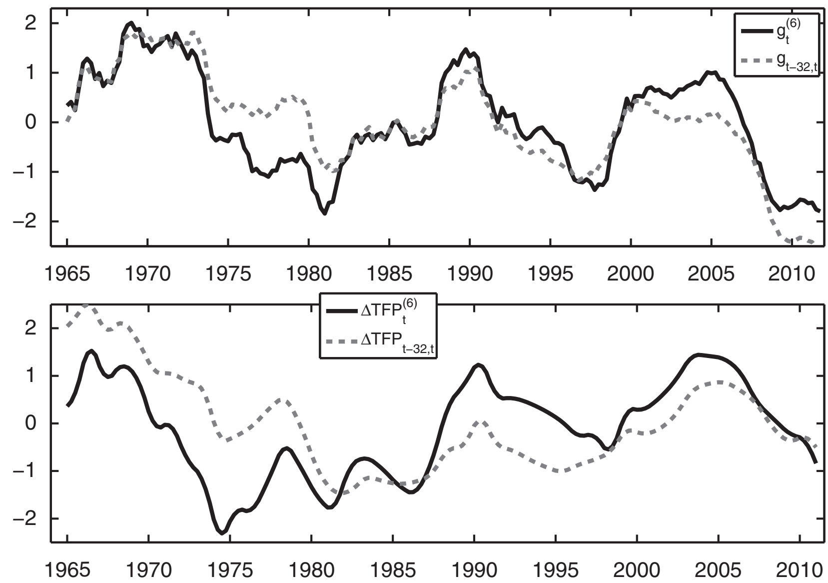

Figure 8
Comparing medium to low frequencies
This figure displays the component with level of persistence $j=6$ and the components with a frequency greater than 32 quarters filtered out of consumption growth (top panel) and total factor productivity (bottom panel). We use $2^6/4=16$ years of data at the beginning of the sample to initialize the filtering procedure.

5. Conclusion

This paper shows that a long-run risk model, with the effects of persistence heterogeneity properly taken into account, offers a credible explanation to many empirical results which seemed to contradict the long-run valuation picture. Our results clearly indicate that any systematic empirical test of a long-run risk model must classify shocks across two competing dimensions: their size as measured by volatility and their persistence as measured by their half-life. In fact, a classification of shocks based on the persistence-based decomposition provides a significant improvement in the ability to detect pricing implications within a long-run risk framework.

Our proposal, the use of a persistence-based decomposition, offers interesting developments. Previously, pros and cons of filtering procedures have been discussed in the macroeconomic literature (see, for instance, Canova 1998; Christiano and Fitzgerald 2003) and it has been observed that sometimes results are not robust to different choices of the filtering criterion. In fact, the decomposition procedure introduces an additional source of model risk, and hence an uncertainty-averse agent should take it into account while forming expectations. In our analysis we assumed that the representative agent is uncertainty-indifferent leaving for future research the analysis of the effects of ambiguity aversion on valuation.

Our contribution raises a number of questions both on the methodological and on the empirical side. On the methodological side, a promising direction of research is to explore the relation between our decomposition based on multiresolution and the spectral approach discussed in Hansen, Heaton, and Li (2008) and formalized in Hansen and Scheinkman (2009). On the empirical side, an interesting next step is to apply our persistence-based decomposition to the analysis of bond prices and state dependent volatility. We leave these topics for future research.

Appendix

A. Data

This appendix describes the data used in the paper. To evaluate the implications of the long-run risks model with persistence heterogeneity we concentrate on five variables: the changes in log consumption and dividends, the log price-dividend ratio, the log price-consumption ratio, and the real interest rate. Following Bansal and Yaron (2004) and Beeler and Campbell (2012),[34] we use data on U.S. nondurables and services consumption from the Bureau of Economic Analysis. We make the standard "end-of-period" timing assumption that consumption during period t takes place at the end of the period. Growth rates are constructed by taking the first difference of the corresponding log series. The price-dividend ratio and dividend growth rates are obtained from the CRSP files. All nominal quantities are converted to real using the personal consumption deflator. To proxy for the price-consumption, we follow Duffee (2005) and use the ratio of the market capitalization of publicly traded stocks to total consumption on nondurables and services. Stock market wealth is measured by the month-end market capitalization of the CRSP value-weighted index, expressed in real per capita terms for comparability to the consumption data. Following Bansal, Yaron, and Kiku (2012) the ex ante real risk-free rate is constructed as a fitted value from a projection of the ex post real rate on the current nominal yield and inflation over the previous year. To run the predictive regression, we use monthly observations on the three-month nominal yield from the CRSP Fama Risk-Free Rate tapes and CPI series. We consider a postwar quarterly U.S. series over the period 1947:Q2–2011:Q4, and for robustness we use a long-run annual series over the period 1930–2010.[35] To initialize the filtering procedure for the quarterly and annual sample we use CRSP data from 1927Q1–1946Q4 and Shiller's annual dataset[36] from 1900–1929, respectively.[37] Finally, the time series for annual productivity is from the Bureau of Labor Statistics; the sample spans 1948–2011. We use a multifactor productivity index that takes into account capital accumulation. In particular, the index adopted measures the value-added output per combined unit of labor and capital input in private business and private nonfarm business, available at ftp://ftp.bls.gov/pub/special.requests/opt/mp/prod3.mfptablehis.zip (as Kaltenbrunner and Lochstoer 2010).

[34] We thank Jason Beeler for kindly providing us with the data.

[35] We do not consider monthly consumption data because they are plagued with measurement errors (Wilcox 1992).

[36] The data have been obtained from Robert Shiller's home page (www.econ.yale.edu/ shiller/).

[37] See Appendix B for further details on the initialization procedure.

B. The initialization of the filtering procedure

In this section we tackle the problem of initializing the filtering procedure in the case where the time-series x_t has only a finite number of observations, $x_1,\dots,x_T$. In fact, formally the persistence-based decomposition described in Section 2.1 is defined for infinite-length signals and finite-length processes must be extended before their components can be extracted. Common extension methods include periodic or mirror-image replication, zero padding, and linear extrapolation (see Mallat 1989 for a detailed description of the various methods). An important drawback of all these methods is that they influence the first 2^j-1 elements of each of the components j and they can therefore generate spurious correlations if such components are used as regressand/regressors in the estimation procedure. In this paper, therefore, we adopt a different approach to deal with the initialization. Specifically, for a given maximum level J for the decomposition, we initialize the components using the first 2^J observations which are then discarded. This of course implies a reduction of our effective sample.

In particular, for the quarterly sample, we use CRSP data from 1927Q1 to 1946Q4 to initialize the components of the price-dividend ratio. These are twenty years of data that suffice to initialize the components with level of persistence $j=1,\dots,6$. In order to extract the seventh component, which requires thirty-two years of data, we either zero-pad the time series or we extend it circularly. The main text presents the results obtained with zero padding. Conclusions do not change when we use circular extension. For the annual sample we rely on the yearly dataset of Shiller beginning in 1900. Our regressions start in 1931 and therefore use thirty-one years of data to initialize all of the components.

To estimate the system of equations in (16) we adopt the technique suggested in Fadili and Bullmore (2002). This requires that we apply the transformation matrix $\mathcal{T}^{(J)}$ to a time series with $T=2^J$ elements to obtain J components each with $T/2^j$ elements. We use the sample period 1948Q1–2011Q4 which has length equal to $T=256$. It is possible, however, to handle sample periods with length less than 2^J by using either zero padding or reflecting boundaries to reach the critical dimension. Fadili and Bullmore (2002) study in detail the effect of artifactual intercoefficient correlations introduced by boundary correction and show that the generalized least squares estimator (GLS) is unbiased over a wide range of data conditions and that its efficiency closely approximates theoretically derived limits.

References

Attanasio, O. P., and G. Weber. 1993. Consumption growth, the interest rate and aggregation. *Review of Economic Studies* 60:631–49.

Avramov, D., and S. Cederburg. 2012. Implications of long-run risk for asset allocation decisions. Working Paper.

Bandi, F. M., and B. Perron. 2008. Long-run risk-return trade-offs. *Journal of Econometrics* 143:349–74.

Bandi, F. M., B. Perron, A. Tamoni, and C. Tebaldi. 2013. The scale of predictability. Working Paper.

Bansal, R., R. F. Dittmar, and C. T. Lundblad. 2005. Consumption, dividends, and the cross section of equity returns. *Journal of Finance* 60:1639–72.

Bansal, R., R. Dittmar, and D. Kiku. 2007. Cointegration and consumption risks in asset returns. *Review of Financial Studies* 22:1343–75.

Bansal, R., V. Khatchatrian, and A. Yaron. 2005. Interpretable asset markets? *European Economic Review* 49:531–60.

Bansal, R., D. Kiku, and A. Yaron. 2012. An empirical evaluation of the long-run risks model for asset prices. *Critical Finance Review* 1:183–221.

Bansal, R., and A. Yaron. 2004. Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance* 59:1481–509.

Bansal, R., A. Yaron, and D. Kiku. 2012. Risks for the long run: Estimation with time aggregation. Working Paper.

Beaudry, P., and E. van Wincoop. 1996. The intertemporal elasticity of substitution: An exploration using a US panel of state data. *Economica* 63:495–512.

Beeler, J., and J. Y. Campbell. 2012. The long-run risks model and aggregate asset prices: An empirical assessment. *Critical Finance Review* 1:141–82.

Borovicka, J., L. P. Hansen, M. Hendricks, and J. A. Scheinkman. 2011. Risk-price dynamics. *Journal of Financial Econometrics* 9:3–65.

Burns, A. F., and W. C. Mitchell. 1946. *Measuring business cycles*. New York, National Bureau of Economic Research.

Calvet, L. E., and A. J. Fisher. 2007. Multifrequency news and stock returns. *Journal of Financial Economics* 86:178–212.

Campbell, J. Y. 2003. Consumption-based asset pricing. In *Handbook of the economics of finance*, vol. 1, chap. 13, pp. 803–887. Eds. G. Constantinides, M. Harris, and R. M. Stulz. Amsterdam: Elsevier.

Campbell, J. Y., and N. G. Mankiw. 1990. Consumption, income, and interest rates: Reinterpreting the time series evidence. NBER Working Papers 2924.

Campbell, J. Y., and R. J. Shiller. 1988. The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies* 1:195–228.

——. 2001. Valuation ratios and the long-run stock market outlook: An update. Cowles Foundation Discussion Papers 1295, Yale University.

Campbell, J. Y., and M. Yogo. 2006. Efficient tests of stock return predictability. *Journal of Financial Economics* 81:27–60.

Canova, F. 1998. Detrending and business cycle facts. *Journal of Monetary Economics* 41:475–512.

Christiano, L. J., and T. J. Fitzgerald. 2003. The band pass filter. *International Economic Review* 44:435–65.

Cochrane, J. H. 1992. Explaining the variance of price-dividend ratios. *Review of Financial Studies* 5:243–80.

Cochrane, J. H., and M. Piazzesi. 2005. Bond risk premia. *American Economic Review* 95:138–60.

Comin, D., and M. Gertler. 2006. Medium-term business cycles. *American Economic Review* 96:523–51.

Constantinides, G. M., and A. Ghosh. 2011. Asset pricing tests with long-run risks in consumption growth. *Review of Asset Pricing Studies* 1:96–136.

Croce, M. M. 2010. Long-run productivity risk: A new hope for production-based asset pricing? Working Paper.

Duffee, G. R. 2005. Time variation in the covariance between stock returns and consumption growth. *Journal of Finance* 60:1673–1712.

Epstein, L. G., and S. E. Zin. 1989. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica* 57:937–69.

——. 1991. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis. *Journal of Political Economy* 99:263–86.

Estrella, A., and G. A. Hardouvelis. 1991. The term structure as a predictor of real economic activity. *Journal of Finance* 46:555–76.

Fadili, M., and E. Bullmore. 2002. Wavelet-generalized least squares: A new blu estimator of linear regression models with 1/f errors. *NeuroImage* 15:217–32.

Fama, E. F., and K. R. French. 1989. Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics* 25:23–49.

Garleanu, N. B., S. Panageas, and J. Yu. 2012. Technological growth and asset pricing. *Journal of Finance* 67:1265–92.

Geanakoplos, J., M. Magill, and M. Quinzii. 2004. Demography and the long-run predictability of the stock market. *Brookings Papers on Economic Activity* 35:241–326.

Gencay, R., and D. Signori. 2012. Multi-scale tests for serial correlation. Working Paper.

Ghysels, E., P. Santa-Clara, and R. Valkanov. 2004. The MIDAS touch: Mixed data sampling regression models. Cirano Working Papers, CIRANO.

Hall, R. E. 1988. Intertemporal substitution in consumption. *Journal of Political Economy* 96:339–57.

Hansen, L. P., J. C. Heaton, and N. Li. 2008. Consumption strikes back? Measuring long-run risk. *Journal of Political Economy* 116:260–302.

Hansen, L. P., and R. J. Hodrick. 1980. Forward exchange rates as optimal predictors of future spot rates: An econometric analysis. *Journal of Political Economy* 88:829–53.

Hansen, L. P., and J. A. Scheinkman. 2009. Long-term risk: An operator approach. *Econometrica* 77:177–234.

Hansen, L. P., and K. J. Singleton. 1983. Stochastic consumption, risk aversion, and the temporal behavior of asset returns. *Journal of Political Economy* 91:249–65.

Jagannathan, R., and Y. Wang. 2007. Lazy investors, discretionary consumption, and the cross-section of stock returns. *Journal of Finance* 62:1623–61.

Kaltenbrunner, G., and L. A. Lochstoer. 2010. Long-run risk through consumption smoothing. *Review of Financial Studies* 23:3190–224.

Kreps, D. M., and E. L. Porteus. 1978. Temporal resolution of uncertainty and dynamic choice theory. *Econometrica* 46:185–200.

Lettau, M., and S. V. Nieuwerburgh. 2008. Reconciling the return predictability evidence. *Review of Financial Studies* 21:1607–52.

Lettau, M., and J. A. Wachter. 2011. The term structures of equity and interest rates. *Journal of Financial Economics* 101:90–113.

Mallat, S. G. 1989. Multiresolution approximations and wavelet orthonormal bases of L2(R). *Transactions of the American Mathematical Society* 315:69–87.

——. 1989. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11:674–93.

Moller, S. V., and J. Rangvid. 2012. End-of-the-year economic growth and time-varying expected returns. Working Paper.

Parker, J. A., and C. Julliard. 2005. Consumption risk and the cross section of expected returns. *Journal of Political Economy* 113:185–222.

Paseka, A., and G. Theocharides. 2010. Predictability in consumption growth and equity returns: A Bayesian investigation. *Financial Review* 45:167–203.

Pastor, L., and P. Veronesi. 2006. Was there a Nasdaq bubble in the late 1990s? *Journal of Financial Economics* 81:61–100.

——. 2009. Technological revolutions and stock prices. *American Economic Review* 99:1451-–83.

Tamoni, A. 2011. The multi-horizon dynamics of risk and returns. *SSRN eLibrary*.

Torous, W., R. Valkanov, and S. Yan. 2004. On predicting stock returns with nearly integrated explanatory variables. *Journal of Business* 77:937–66.

Van Binsbergen, J. H., M. W. Brandt, and R. S. Koijen. 2012. On the timing and pricing of dividends. *American Economic Review* 102:1596–618.

Van Binsbergen, J. H., W. Hueskes, R. S. Koijen, and E. B. Vrugt. 2012. Equity yields. Working Paper.

Vissing-Jorgensen, A. 2002. Limited asset market participation and the elasticity of intertemporal substitution. *Journal of Political Economy* 110:825–53.

Wilcox, D. W. 1992. The construction of U.S. consumption data: Some facts and their implications for empirical work. *American Economic Review* 82:922–41.