The scale of predictability

F.M. Bandi,*B. Perron,†A. Tamoni,‡C. Tebaldi§

First version: November 2012. This version: March 15, 2013

Abstract

We view economic time series as the result of a cascade of shocks occurring at different times and different frequencies. In agreement with this premise, we suggest that equilibrium (predictive) relations that are found to be elusive when using raw data may hold true for different layers, components, or *details*, in the cascade of shocks affecting economic time series. This observation leads to a notion of a *scale-specific predictability*. We show that running predictive regressions on suitably-aggregated data may work as a simple low-pass filter capable of detecting the extent of scale-specific predictability and, importantly, the frequency at which it operates. We also provide a way to extract low-frequency time-series details directly and, therefore, evaluate their relation or lack thereof. Using aggregation, as well as direct extraction of the details, we offer broad evidence of predictability at frequencies lower than business-cycle frequencies. Among other low-frequency relations, we show that a very slow-moving component of the variance process with decade-long periodicity predicts itself as well as a similarly slow-moving component of future excess market returns, i.e., a long-run risk-return trade-off.

JEL classification: C22, C32, E32, E44, G17
Keywords: long-run, persistence, risk-return trade-off.

*We thank Lars P. Hansen, Bryan Kelly and Alex P. Taylor for helpful comments. We would like to thank seminar participants at LSE.

*Johns Hopkins University, Carey Business School, 100 International Drive, Baltimore MD, 21202, USA. E-mail: fbandi1@jhu.edu

†Université de Montrèal, Department of Economics, 3150 Jean-Brillant Street, Montreal QC H3T 1N8, Canada. E-mail: benoit.perron@umontreal.ca.

‡London School of Economics, Department of Finance, Houghton Street, WC2A 2AE London, UK. E-mail: a.g.tamoni@lse.ac.uk.

§Universita' Bocconi, Via Roentgen 1, Milano, Italy. E-mail: claudio.tebaldi@unibocconi.it.

1 Introduction

Low-frequency economic shocks may not just be long-run averages of high-frequency shocks. Similarly, low-frequency economic relations may not imply analogous relations at higher frequencies.

In agreement with this premise, we argue that every frequency of observation may, in general, be impacted by specific shocks and can, in consequence, carry unique information about the validity of economic relations. While such relations are typically tested - and often rejected - at a specific, high frequency, they may be valid at lower frequencies without requiring the researcher to impose, coherently with data, tight constraints at the highest frequency of observation for internal consistency.

To model these ideas parsimoniously, we introduce the notion of *scale-specific predictability*. Equilibrium relations like the presumed dependence between market risk-premia and expected volatility (risk-return trade-offs) or between nominal rates and expected inflation (Fisher's effects) can take on a different look at different frequencies, levels of resolutions or, in our jargon, *scales*. In essence, we view a time series as a sum of scale-specific components, or *details*. We define the details as elements of the observed time series with a specific level of resolution. Higher scales are associated with lower resolution, lower frequencies, and higher persistence. Higher scales are affected by shocks which are relatively smaller in size but persist in the system relatively longer, as is typical of long-run shocks.

We show that direct extraction of the details from observed regressands and regressors can shed important light on the validity of the assumed economic restrictions. Since economic data may be viewed as aggregates of a cascade of scale-specific shocks with different sizes and different half-lives, focusing on individual elements of a time series with specific levels of resolution provides us with a suitable way to disaggregate information occurring at different frequencies. This, in turn, gives us a methodology to zoom in on to specific layers of the cascade of shocks affecting the system at different frequencies, isolate each layer, and identify those layers over which economic restrictions are more likely to be satisfied.

Often-studied economic relations - such as those between returns and volatility or nominal rates and inflation - may be very hard to detect when using the economic series themselves. They may not be satisfied by high-frequency components of the series. However, they may characterize components, or details, at a lower level of resolution. Often, the pattern of predictability in the

details reaches a peak corresponding to scales associated with business-cycle frequencies or lower. At these scales - but only at these scales - the corresponding slope estimates have signs and magnitudes which are consistent with classical economic logic. We find, for instance, that lagged values of the market return variance's 10-year detail forecast future values of the corresponding variance details as well as future values of the excess market return's 10-year details. In essence, we provide evidence for an extremely slow-moving component in market variance which predicts itself and predicts a similarly slow-moving component in the market's excess returns, i.e., a scale-specific risk-return trade-off. Analogous considerations apply to consumption volatility. The consumption variance's 10-year detail is positively autocorrelated and predicts future values of the excess market return's 10-year detail. Said differently, a higher slow-moving component of past consumption volatility's predicts higher future values of the corresponding slow-moving component in market returns because it predicts a higher future value of itself (and this higher volatility component should, in agreement with classical economic theory, be compensated). This is, again, a risk-return trade-off, but on the details. Both in the case of market volatility and consumption volatility, when evaluating risk-return trade-offs by running predictive regressions on the 10-year details, rather than on the original series, we find R^2 values well above 70%. In both cases, short-run shocks simply hide equilibrium relations which careful signal extraction can bring to light. Additional scale-specific equilibrium relations are illustrated below.

What are the implications of predictability on the details for conducting long-run predictive analysis? We show formally that two-way (forward for the regressand, backward for the regressor) adaptive aggregation of the series, as in Bandi and Perron (2008), leads to increased predictive ability precisely at frequencies corresponding to a scale, or level of resolution, over which the economic relation is more likely to apply. We demonstrate that aggregation works as a simple filter capable of eliminating high-frequency shocks, or short-term noise, while highlighting low frequency details to which economic restrictions may apply. In this sense, finding increasing predictability upon forward/backward aggregation, as in the case of the long-run risk-return trade-offs illustrated by Bandi and Perron (2008) for example, is symptomatic of risk compensations which apply to highly persistent, low-frequency components of returns and volatility. Such components are hidden by noisier, high-frequency details when testing the restrictions on disaggregated, raw data. Their signal, however, dominates short-term noise when two-way aggregation is brought to data.

The separation of a time series in terms of details with different levels of resolution is conducted

by virtue of Multiresolution analysis (MRA) as in Mallat (1989), Dijkerman and Mazumdar (1994), Yazici and Kashyap (1997) and, recently, in the work of Ortu, Tamoni, and Tebaldi (2011) and Tamoni (2011) in the context of structural consumption asset pricing models. This paper provides a mechanism, operating at the level of each frequency or scale, to correlate elements of two, or more, series of interest while preserving internal consistency in the re-construction of the original series. We introduce a novel notion of (scale-specific) predictability intended as the relation between components of two, or more, time series at a specific layer (detail) in the cascade of shocks affecting the economy. We show, using low-pass aggregation filters, that there is a (close to) one-to-one mapping between long-run predictive relations obtained on regressands and regressors aggregated forward and backward and predictability on their details at certain frequencies. We use both direct extraction of the details (and predictive regressions run on the details) and two-way aggregation (with analogous predictive regressions run on the aggregated series) to provide broad evidence of predictability over the long run.

The study of low-frequency contributions to economic and financial time series has a very long history, one which we can not attempt to review here. The approach adopted in this paper shares features with successful existing approaches. As in Beveridge and Nelson (1981), who popularized time-series decompositions into stochastic trends and transitory components, we can view the details as components (more than two) with different levels of persistence operating at different frequencies. As in Hansen and Sheinkman (2009), we employ operators to extract low-frequency information (in our case, the low-frequency information embedded in the details). Finally, similarly to Müller and Watson (2008), the extracted low-frequency information can be summarized by a finite number of averages of the original data or, said differently, by a finite number of "typical" points on the extracted details. These "typical" points are used to formalize a notion of frequency-specific or scale-specific predictability shown to be revealed by simple regression filters based on forward/backward aggregation.

The paper proceeds as follows. Section 2 provides intuition for the analysis of time series with multiple resolutions, MRA. We show how data can be viewed as a collection of time-specific and frequency-specific shocks, as in a generalized Wold decomposition, or as the sum of details operating at different frequencies. Section 3 suggests the notion of scale-specific predictability, the idea that economic relations may hold for individual layers in the cascade of shocks affecting the economy and, hence, for individual details but may be hidden by high-frequency perturbations. This section

discusses the potential of simple low-pass filters based on aggregation for the identification of the level of resolution, or layer, at which predictability plays a role. Section 4 applies aggregation, as well as direct of extraction of the details, to provide evidence of long-run risk-return trade-offs in equities and bonds. In Section 5 we verify the assumption of theory using simulations. Specifically, we impose predictability on the "typical" points of certain details to re-construct the original series before showing effectiveness of aggregation in the identification of low-frequency economic relations. Section 6 turns to the validation of long-run versions of the classical Fisher's effect. Section 7 concludes. Technical details are in the Appendices.

2 Time-series analysis with multiple resolutions

Assume the time series $\{x_t\}$ is wide-sense stationary.[1] While alternative choices are possible, in what follows we use Haar wavelets to decompose $\{x_t\}$ into details. The use of Haar wavelets provides a clear connection between aggregation (and long-run dynamics) and time-series components with low levels of resolution and high persistence, something which will be explored in what follows.[2] In general, however, any MRA can be viewed as an aggregation scheme (e.g., Abry, Veitch, and Flandrin (1998)).

2.1 Intuition

Write

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

with

$$\pi_t^{(0)} = x_t,$$

$$x_t^{(j+1)} = \pi_t^{(j)} - \pi_t^{(j+1)}, \tag{2}$$

and

$$x_t = \sum_{j=1}^{J} x_t^{(j)} + \pi_t^{(J)}. \tag{3}$$

[1] Nonstationary processes in the affine stationary class could also be accommodated (see, e.g., Yazici (2004)).

[2] For interesting discussions of wavelet methods for time series analysis, we refer the reader to Percival and Walden (2000) and Gençay, Selçuk, and Whitcher (2001).

Given a choice of J, Eq. (1), Eq. (2), and Eq. (3) provide a way to decompose the original time series into a sum of *details* $x_t^{(j)}$ associated with time (t) and *scale* (j). The collection of values $\left\{x_t^{(j)}\right\}$ with j fixed, corresponds to the representation of a time-series (viewed as a function of time t) at the j-th scale. Consider the case $J = 1$. We have

$$x_t = \underbrace{\frac{x_t - x_{t-1}}{2}}_{x_t^{(1)}} + \underbrace{\left[\frac{x_t + x_{t-1}}{2}\right]}_{\pi_t^{(1)}},$$

which breaks the series down into a transitory and a persistent component. Set, now, $J = 2$. We obtain

$$x_t = \underbrace{\frac{x_t - x_{t-1}}{2}}_{x_t^{(1)}} + \underbrace{\left[\frac{x_t + x_{t-1} - x_{t-2} - x_{t-3}}{4}\right]}_{x_t^{(2)}} + \underbrace{\left[\frac{x_t + x_{t-1} + x_{t-2} + x_{t-3}}{4}\right]}_{\pi_t^{(2)}},$$

which further separates the persistent component $\pi_t^{(1)}$ into an additional transitory and an additional persistent component. The procedure can, of course, be iterated yielding a general expression for the detail $x_t^{(j)}$, i.e.,

$$\begin{aligned} x_t^{(j)} &= \frac{\sum_{i=0}^{2^j-1} x_{t-i}}{2^j} - \frac{\sum_{i=0}^{2^{(j-1)}-1} x_{t-2^{j-1}-i}}{2^{(j-1)}} \\ &= \pi_t^{(j-1)} - \pi_t^{(j)} \end{aligned}$$

with $\pi_t^{(j)} = \frac{\sum_{i=0}^{2^j-1} x_{t-i}}{2^j}$. In essence, the time series can be written as a collection of details $x_t^{(j)}$ with different degrees of resolution (i.e., persistence) along with a low-resolution approximation $\pi_t^{(J)}$. Similarly, it can be written as a telescopic sum

$$x_t = \sum_{j=1}^{J} \underbrace{\left\{\pi_t^{(j-1)} - \pi_t^{(j)}\right\}}_{x_t^{(j)}} + \pi_t^{(J)} = \pi_t^{(0)}, \tag{4}$$

in which the details are naturally viewed as changes in information occurring at different scales, 2^j and 2^{j-1}. The scales are dyadic and, therefore, enlarge with j. The higher j, the lower the level of resolution, and the larger the scale. In particular, the innovations $x_t^{(j)} = \pi_t^{(j-1)} - \pi_t^{(j)}$ become smoother, and more persistent, as j increases.

Importantly, aggregation of the time series uncovers information at different scales or, more

precisely, for scales that are higher than the one corresponding to the aggregation level. To see this, using Eq. (4), write

$$x_{t-2^s+1,t} = \left(\sum_{i=0}^{2^s-1} x_{t-i} \right) / 2^s = \pi_t^{(s)} = \sum_{j=s+1}^{J} \underbrace{\left\{ \pi_t^{(j-1)} - \pi_t^{(j)} \right\}}_{x_t^{(j)}} + \pi_t^{(J)}. \tag{5}$$

where $s = 0, 1, \ldots, J$. The implication of this statement is that economic relations which emerge from aggregation, and may not appear at higher frequencies, may be viewed as *scale-specific*. We provide a channel through which economic restrictions may be satisfied at particular levels of resolution without having to be satisfied at *all* levels of resolution.

2.2 Scale-specific shocks

Assume, without loss of generality and for convenience, that the time series $\{x_t\}$ is mean zero. The details $\left\{x^{(j)}\right\}$ are also mean zero and wide-sense stationary for any fixed j (Wong, 1993). One can write the representation (understood in the mean-squared sense):

$$x_t = \sum_{j=1}^{J} \sum_{k=0}^{\infty} a_{j,k} \varepsilon_{t-k2^j}^{(j)},$$

where $\varepsilon_{j,t} = x_t^{(j)} - \mathcal{P}_{\mathcal{M}_{j,t-2^j}} x_t^{(j)}$ and $\mathcal{P}_{\mathcal{M}_{j,t-2^j}}$ is a projection mapping onto the closed subspace $\mathcal{M}_{j,t-2^j}$ spanned by $\left\{x_{t-k2^j}^{(j)}\right\}_{k \in \mathcal{Z}}$. This is a Wold representation which applies to every scale. Specifically, it is a decomposition which explicitly represents the time series of interest as a linear combination of shocks classified on the basis of their arrival time, as is typical in the analysis of linear stationary time-series, as well as their scale. The decomposition will reduce to a classical Wold representation for linear stationary processes *if*[3]

$$\underbrace{\varepsilon_t^{(j)}}_{x_t^{(j)} - \mathcal{P}_{\mathcal{M}_{j,t-2^j}} x_t^{(j)}} = \underbrace{\left(\sum_{i=0}^{2^{j-1}-1} \varepsilon_{t-i} - \sum_{i=0}^{2^{j-1}-1} \varepsilon_{t-2^{j-1}-i} \right) / \sqrt{2^j}}_{\sum_{i=0}^{2^j-1} \left(x_{t-i} - \mathcal{P}_{\mathcal{M}_{t-i-1}} x_{t-i} \right)}.$$

Intuitively, if the scale-specific innovations are sums of high-frequency innovations, then the information contained by the series at every scale is an aggregate of that contained at higher frequencies.

[3]For an illustration, see Appendix A.

If, on the other hand, this is not the case, then there is a separation between scales - in terms of their informational content - which still preserves their consistency. This would translate into shocks which are specific to individual scales, thereby giving meaning to economic relations which, again, may only be satisfied at certain frequencies. The logic behind multiscale spectral analysis, and its implications for detecting short and long-run dynamics in economic time series, is further described in Appendix B.

2.3 Modeling the details

Following the approach in Dijkerman and Mazumdar (1994) one may represent the details as linear autoregressive processes. A convenient (2-parameter) way to do so is to write

$$x_t^{(j)} = \rho_j x_{t-2^j}^{(j)} + \varepsilon_t^{(j)},$$

where the shocks $\varepsilon_{t+2^j}^{(j)}$ are independent across scales, white noise, and with a scale-specific variance σ_j. The model is autoregressive of order 1 in the dilated time of the scale being considered.

Notice that the decomposition into autoregressive details only resembles the superposition model of Barndorff-Nielsen and Shephard (2001). Here, however, heterogeneity is specific to each scale.

3 Scale-specific predictability and backward/forward aggregation

Consider a predictive variable y_t and a predictor x_t. It is standard in macroeconomics and finance to verify predictability by computing linear, or nonlinear, projections at the highest frequency of observation. It is also common to aggregate the regressand. A recent approach proposed by Bandi and Perron (2008) aggregates both the regressand (forward) and the regressor (backwards). The aggregate regressor is adapted to time t information and is, therefore, non anticipative. The logic for aggregating *both* the regressand and the regressor resides in the intuition according to which equilibrium implications of economic models may impact highly persistent components of the variables $\{y_t, x_t\}$ while being hidden by short-term noise (Bandi and Perron, 2008). Aggregation provides a natural way to filter out noise, thereby yielding a cleaner signal. We now formalize this

intuition. Assume, for simplicity, the following predictive model postulated in terms of details:

$$y_{t+2^j}^{(j)} = \alpha^j + \beta^j x_t^{(j)}$$

In this specification, the assumed relation applies to each j-th scale, but may be confounded by the presence of uncorrelated shocks at alternative scales. Said differently, the (scale) time $t^* + 1$ value of the j-th scale of the variable y is linearly related to the time t^* value of the j-th scale of the variable x. Since time is scale specific and dilated, the time $t^* + 1$ value of the j-th scale is expressed as $t + 2^j - 1$ in calendar time units.

Using Eq. (5), forward/backward aggregation yields

$$\begin{aligned} y_{t+1,t+2^s} &= \left(\sum_{i=1}^{2^s} y_{t+i}\right)/2^s = \left(\sum_{i=1}^{2^s} y_{(t+2^s-1)-i}\right)/2^s = \\ \pi_{t+2^s}^{(s)} &= \sum_{j=s+1}^{J} y_{t+2^s}^{(j)} = \sum_{j=s+1}^{J} \left\{\alpha^j + \beta^j x_t^{(j)}\right\}. \end{aligned} \tag{6}$$

Hence, economic relations which apply to highly persistent components will be revealed by averaging since higher frequency dynamics will not affect inference. Importantly, one can also directly inspect the details to evaluate the extent of the existing relations at different scales. Without decomposing the series into details, simple averaging will highlight relations that exist at scales corresponding to the level of aggregation and, possibly, at lower scales. In what follows, we adopt both approaches. We first show the outcome of two-way aggregation (and predictive regressions run on the aggregated raw series). We then turn to regressions on the extracted details and illustrate the consistency of their findings with those obtained from two-way aggregation. This consistency is further confirmed by simulation as well as in the context of an illustrative example in Appendix C.

Next, we expand on the implications of our framework to understand - and broaden the scope of - classical predictability relations in the literature. We focus on risk-return trade-offs and Fisher's effects.

4 Risk-return trade-offs

4.1 Equity returns on *market* volatility

The basic observation which drives our understanding of the analysis of risk-return trade-offs at different levels of aggregation performed by Bandi and Perron (2008) is the following: the "basis" of independent shocks which drive return time series must be classified along two dimensions: their time of arrival and their scale or level of persistence.

As illustrated above, we propose an adapted linear decomposition which represents a time series as a sum of orthogonal details whose spectrum is concentrated on an interval of characteristic time scales (inverse of frequencies) ranging from 2^{j-1} to 2^j. Each element of the time series x_t is decomposed into a sum of detail components $x_t^{(j)}$ classified by their degree of persistence ρ_j plus a permanent component $\pi_t^{(\infty)}$. By construction each details $x_t^{(j)}$ is stationary and mean zero. The permanent component may be viewed as the sum of deterministic and stochastic trend components with infinite persistence.

We apply the decomposition to monthly excess returns and monthly realized variance series, $r_{t,t+h}$ and v_t^2.[4] The details are shown in Figure 1. We run a VAR(1) for the detail components $\{r_t^{(j)}\}_{j=1}^J$ and $\{v_t^{2(j)}\}_{j=1}^J$ (each detail on lagged values of the other details). The t-statistics are strongly significant in correspondence with the autoregressive components. The hypothesis of independence among detail components with different degrees of persistence is not in contradiction with data.

Next, we consider the forward/backward regressions

$$r_{t,t+h} = \alpha_h + \beta_h v_{t,t-h}^2 + u_{t,t+h}, \tag{7}$$

where $r_{t,t+h}$ and $v_{t,t-h}^2$ are aggregates of excess market returns and return variances over an horizon of length h. Empirical results are reported in Table 9 - Panel A. We confirm the findings in Bandi and Perron (2008), namely *future* excess market returns are correlated with *past* market variance. Dependence increases with the horizon, and is strong in the long run, with R^2-values between 8 and 10 years ranging between 30% and 51%.

A crucial observation in Bandi and Perron (2008) is that "*the long-run results are not compatible*

[4] Appendix D describes the data and construction of variables.

with classical short-term risk return trade-offs." Proposition 3 in that paper discusses the asymptotic properties of the long-run regressions' slope estimates and R^2s and shows that disaggregated asset pricing models which solely imply dependence between excess market returns and (autoregressive) conditional variance at the highest resolution *cannot* deliver the corresponding results upon aggregation.

Here, we argue that the relation may be viewed as being *scale-specific*. In agreement with the implications of Eq. (6), aggregation is helpful to reveal low-frequency risk compensations. To corroborate this logic, we run detail-wise predictive regressions analogous to (7), namely

$$r^{(j)}_{k2^j+2^j} = \beta_j v^{2(j)}_{k2^j} + u^{(j)}_{k2^j+2^j} \tag{8}$$

with $j = 1, ..., 4$ and k an integer. Similarly, we run

$$v^{2(j)}_{k2^j+2^j} = \rho_j v^{2(j)}_{k2^j} + \varepsilon^{(j)}_{k2^j+2^j}. \tag{9}$$

Note that the element at time-scale j is defined on a time-grid whose time unit is 2^j times the unit scale of the original time series of observations. These elements can be obtained by sampling the components constructed as in (2), i.e. $\left\{x^{(j)}_t, t = k2^j, k \in \mathbb{Z}\right\}$. We refer to these sub-series as to the *decimated* elements at time-scale j of the original time series.[5] One can view them as containing, in our framework, similar information content about low-frequency dynamics as the small number of data averages in Müller and Watson (2008). The results are based on yearly data and are reported in Table 9 - Panel B. For a clear interpretation of the corresponding levels of resolution, we refer to Table 1. The strongest predictability is for $j = 4$, which corresponds to a time length of 8 to 16 years. The importance of this scale relates back to the increased significance of backward-aggregated variance as a predictor of forward-aggregated excess market returns at the same low frequencies. For $j = 4$, the R^2 on the detail-wise predictive regression is 72%. The R^2 on the detail-wise autoregression is 26%. In both cases, the slope is positive. In the case of the detail-wise predictive regression, its value is - coherently with theory - similar to that obtained from two-way aggregation. In essence, we find that, at scale $j = 4$, a slow-moving component of the variance process predicts

[5] In general, we can use the components $x^{(j)}_t$, $j = 1, .., J$, and $\pi^{(J)}_t$ in their entirety to reconstruct the time series using (3). This is the *redundant* decomposition of a time series proposed in Renaud, Starck, and Murtagh (2005). Alternatively one can reconstruct the time series signal from the decimated components using the (inverse of the) Haar unitary matrix, see Appendix B.

itself as well as future excess market returns. Higher past values of the variance detail predict higher future values of the same detail and, consequently, higher excess market returns, as required by conventional logic behind risk compensations.

4.2 Equity returns on *consumption* volatility

Replacing market variance with consumption variance, as justified structurally by Tamoni (2011), does not modify the previous results upon two-way aggregation (Table 10 - Panel A). The R^2-values between 8 and 10 years range between 23% and 52%. Running detail-wise predictive regressions leads to maximum predictability (and an R^2 of 83%) associated with low frequency cycles of about a decade on average, i.e., $j = 4$ (Table 10 - Panel B). Similarly, for $j = 4$, detail-wise autoregressions of future consumption volatility on past consumption volatility yield a positive autocorrelation and an R^2 value of 61%. In sum, because it predicts itself, a slow-moving component of consumption volatility has forecasting ability for the corresponding slow-moving component of excess market returns. In other words, again, low-frequency consumption risk appears to be compensated.

4.3 Bond returns on bond return volatility

Excess bond returns may depend on past bond return volatility. We consider 5 maturities: 1 year, 2 years, 5 years, 7 years, and 10 years (in Tables 11, 12, 13 14 and 15, respectively). For the one- and two-year maturity cases, two-way aggregation reveals, once more, a 10-year peak in predictability. For the five-, seven and ten-year maturity, two-way aggregation reveals instead a 5-year peak in predictability. In all cases, two-way aggregation provides substantial evidence of predictability at horizons greater than 24 months.

These findings are confirmed by the detail-wise regressions. In fact, for all maturities, we find support for predictability at time scales greater than or equal to $j = 5$. The evidence hints at fluctuations occurring at time-scale $j = 5$ and $j = 7$ (corresponding to the time interval $16 - 32$ and $64 - 128$ months, respectively, see Table 1) as the main drivers of the predictable relation between bond returns and volatility. The R^2 values for the detail-wise predictive regressions range between about 12% (the 5- and 10-year cases) and 26% (the 1-year case), for the case $j = 5$, and between 50% (the 5-year case) and 75% (the 10-year case), for the time-scale $j = 7$.[6] Importantly consistent

[6]For the 7-year maturity bond we find mixed evidence for predictability at time-scale $j = 6, 7$. The relatively low t-statistics at these scales can be explained upon recalling that the peak in predictability for the 7-maturity is reached

with theory, the predictive slope on the details is numerically close to the predictive slope obtained from two-way regressions.

When we look at the values on the volatility details' autoregressions, we find strong evidence for an autoregressive component at time-scale $j = 5$, the R^2 ranging between about 13% (the 1-year case) and 51% (the 10-year case). The autoregressive coefficient is, again, positive. The autoregessive coefficient on the slow-moving detail at scale $j = 7$ is positive but poorly estimated.

Taken together, these results point to two slow-moving predictable components affecting volatility and future bond returns: one with an half-life of about 2-years and one whose half-life belongs to the interval $64 - 128$ months.

The above results illustrate one important point of our technique. The multiresolution approach decomposes a time-series in a time-scale space where both the frequency and time dimensions can live together. This in turns allows us to link predictability (associated with the notion of past and future instants of time) to periodic fluctuations in the time series (associated with the frequency dimension). This comes at a cost. In fact a time-scale j identifies a range of frequencies and not just a single one. As for the 7 year maturity bond, if the fluctuation responsible for the predictability is localized at the edge between two time-scales, then the Haar decomposition can be inconclusive on which of the adjacent time-scales is the predictable one. Of course one could then think of using a very selective filter peaked, in our case around the 60-month horizon. In this case, however, we would give up the localization property so useful to study predictability.[7]

One important observation about two-way aggregation is in order. One may argue that aggregation itself could lead to spurious predictability. If this were the case, contemporaneous aggregation should also lead to similar patterns as those found with forward/backward aggregation. In all cases above, one could show that this is not the case.[8] In other words, contemporaneous aggregation does not lead to any of the effects illustrated above (including consistency between findings based on aggregated series and findings based on details). In a similar vein, one could also argue that

at the 60 months horizon see Table 14-Panel A. This horizon is right at the boundary between the fluctuations' intervals captured by the details at scale $j = 6$ and $j = 7$. This in turn explains why it is difficult to discriminate which one between these two details induces the predictable fluctuations in bond returns.

[7] On top of the issue just mentioned, one has to remember that there is leakage across adjacent time-scale. In this case however it is possible to reduce the impact of leakage on the results by replacing the Haar filter with another one in the multiresolution family, for instance the Daubechies one. The investigation of which multiresolution filter is the most suitable for the purpose of time-scale predictability is out of the scope of this paper.

[8] The corresponding tables are not reported for conciseness but can be provided by the authors upon request. Bandi and Perron (2008) report identical results.

spuriousness would prevent a tent-shape pattern to arise in predictive regressions on the aggregated series. Yet, tent-shape patterns may readily arise as the case of Fisher's effects (in Section 6) exemplifies.

Below, we show by simulation that scale-wise predictability translates into predictability upon two-way aggregation. Tent-shaped patterns are possible provided, of course, predictability occurs at the appropriate scale. We show that, if predictability on the details applies, contemporaneous aggregation leads to insignificant outcomes. Similarly, if no predictability on the details applies, two-way aggregation also leads to insignificant outcomes. In sum, the effects reported in this section point to a genuine 10-year cycle in the predictable variation of equity and bond risk premia.

5 Simulating scale-specific predictability

We argue that a predictive relation localized around a specific scale (that is, with a characteristic persistence level) can produce a pattern of R^2 which has a peak when the aggregation interval corresponds to the characteristic persistence level. The postulated process for variance and returns of each detail component is specified by the following relation:

$$\begin{aligned} v^{2(j)}_{k2^j+2^j} &= \rho_j v^{2(j)}_{k2^j} + \varepsilon^{(j)}_{k2^j+2^j} \\ r^{(j)}_{k2^j+2^j} &= v^{2(j)}_{k2^j} \end{aligned} \tag{10}$$

for $j = j^*$ and

$$\begin{aligned} v^{2(j)}_{k2^j+2^j} &= \varepsilon^{(j)}_{k2^j+2^j} \\ r^{(j)}_{k2^j+2^j} &= u^{(j)}_{k2^j,k2^j+2^j}, \end{aligned}$$

for $j \neq j^*$, where k is defined as above and $j = 1, \ldots, J = 9$. The linear processes $\varepsilon^{(j)}_t$ and $u^{(j)}_t$ satisfy $corr(u^{(j)}_t, \varepsilon^{(j)}_t) = 0 \ \forall t, j$. The scales are defined here at the monthly level. The data generating process is formulated for a "deconstructed" or "decimated" process. One then has to recover the time series. To do so, we simulate the process at scale j every 2^j steps and multiply it by the inverse Haar Matrix. Appendix C illustrates within a simple example the simulation procedure in the time-scale domain and the reconstruction steps in the time domain.

5.1 Running the predictive regression

Table 2 shows the results obtained by running the regression in (7) on simulated time series generated from (10). We compare these results with those in Table 5, where no predictability is assumed.

When imposing the relation at scale $j^* = 6$, namely for a time span of 32 to 64 months, we reach a peak in the R^2 of the two-way regressions at 5 years. The 5-year R^2 is more than thirty times larger than the one that we obtain in the case of a spurious regression. Moreover, the slopes increase reaching a maximum at 5 years (and approaching the theoretical value of 1 on the details) before decreasing again. This is a tent-shaped pattern which readily derives from solely imposing scale-wise predictability at a frequency lower than business cycle frequency but not as low as the 10-year frequency.

5.2 Predicting variance

Bandi and Perron (2008) find that, in spite of its predictive ability for long-run excess market returns, long-run past variance does not predict its future values. As emphasized by Bandi and Perron (2008), this appears to contradict classical economic logic behind predictability. Why is realized variance aggregated over a certain time-horizon not predictable? We provide a justification based on measurement error.

In the previous section we simulated a model in which realized variance was characterized by an autoregressive component at scale $j^* = 6$. Table 4 reports simulation results from running linear regressions of h-period realized market variances: $v^2_{t,t+h} = \sum_{i=1}^{h} v^2_{t+i-1,t+i}$ on h-period past realized variance $v^2_{t,t-h} = \sum_{i=1}^{h} v^2_{t-i+1,t-i}$, i.e.,

$$v^2_{t,t+h} = \rho_h v^2_{t,t-h} + \varepsilon_{t,t+h}, \tag{11}$$

when the underlying data generating process is the one with $j^* = 6$. It is readily verified that the level 6 component generates predictability on the interval of scales $h = [2^{j^*-1}, 2^{j^*}) = [32, 64)$. In effect, for the horizon $h = 36$, the t-statistic is very significant and the R^2 value (23%) is non negligible. The sign of the coefficient (negative) at this aggregation level is unsurprising. It is expected that a positive trend in the past 36 periods would revert to the mean in the next 36 periods.

Having made this point, empirical evidence of some variance predictability seems to be an

important condition for the above data generating process to be an acceptable model specification. On the other hand, the findings in Table 4 are at odds with the findings in Table 9 of Bandi and Perron (2008) who also report estimates from a linear regression of realized variance on itself h periods in the past (see their Eq. (11)). The results in their Table 9 show that "the autocorrelations become quickly statistically insignificant with the level of aggregation". This finding led the authors to conclude that "realized variance is virtually uncorrelated in the long run".

Next, we show that this apparent contradiction is easily re-solved by taking into account the measurement error generated by heterogeneity in the levels of persistence. As earlier, we denote by $v^{(j)}$ the j-th component sampled over time intervals 2^j. We evaluate the effect on the dynamics of the data generating process of persistent heterogeneity, i.e., the presence of independent predictable components at different frequencies. We simulate according to the following data generating process:

$$v^{2(j)}_{k2^j+2^j} = \rho_j v^{2(j)}_{k2^j} + \varepsilon^{(j)}_{k2^j+2^j}$$

with $j = j^* \in S = \{6, 9\}$ and

$$v^{2(j)}_{k2^j+2^j} = \varepsilon^{(j)}_{k2^j+2^j}$$

for $j \notin S$, $j = 1, \ldots, J = 9$, where $corr(\varepsilon_{t',j'}, \varepsilon_{t,j}) = 0 \ \forall t, j, t', j'$ with $t' \neq t$ and $j' \neq j$. Hence, we consider a variance process generated by two autoregressive processes, one at scale 6 and one at scale 9.

Table 8 reports results. Comparing these results with the one in Table 4, we note that the effect of adding an additional predictable component close to a permanent one is two-fold. First, the autoregressive coefficient at horizon $h = 36$ is no longer significant and "realized variance is virtually uncorrelated in the long run" even though two autoregressive components are truly present. Second, the t-statistics and the R^2 for very short horizons, namely 3 months and 6 months, are now higher and significant.

Having said this, the conclusions drawn above continue to hold. Earlier, we postulated a predictable relation between returns and realized variance only at scale $j^* = 6$. The variance process had only one predictable component. We now consider the case where there are two components of variance which are not white noise ($j = \{6, 9\}$). However, we impose that the only relevant predictability relation between returns and variance occurs at $j^* = 6$. The $j^* = 9$ detail is in-

distinguishable from a permanent component effect and is not relevant for prediction. Hence, we assume

$$v^{2(j)}_{k2^j+2^j} = \begin{cases} \rho_j v^{2(j)}_{k2^j} + \varepsilon^{(j)}_{k2^j+2^j} & \text{if } j = 6, 9, \\ \varepsilon^{(j)}_{k2^j+2^j} & \text{otherwise} \end{cases}$$

and

$$r^{(j)}_{k2^j,k2^j+2^j} = \begin{cases} v^{2(j)}_{k2^j} & \text{if } j = 6, \\ u^{(j)}_t & \text{otherwise} \end{cases}$$

Results are reported in Table 6. Table 6 confirms the findings of the previous section, namely the relation at scale $j^* = 6$ implies a spike in the β coefficients, t-statistics, and R^2s exactly at horizon $h = 60$. Compared to Table 2, we note that, when the variance process is characterized by two autoregressive components, Table 6 shows that the relation is softened but continues to remain significant.

5.3 The case of contemporaneous aggregation

No predictability is detected at horizons $[2^{j^*-1}, 2^{j^*}]$ with $j^* = 6$ when both the regressor and the regressand are aggregated over the same time interval. Comparing Table 3, where we postulate only one predictable component in variance, to Table 7 where there are two predictable components in the variance process, we see that the effect of multiple predictable components in the regressor is to lower the R^2 at all time scales and, in particular, at the time scale where the predictable relation holds.

6 Fisher's effects

Fisher's hypothesis postulates that nominal interest rates move one-to-one with expected inflation (see, e.g., Boudoukh and Richardson (1993)). To this extent, it is often tested whether the slope of the predictive regression is equal to 1. Mishkin (1992) and Fisher and Seater (1993), for instance, run regressions of k-period continuously-compounded nominal interest rates (and h-period GDP growth) on the contemporaneous h-period expected inflation (and the growth of nominal money supply).

Tamoni and Gozluklu (2012) explore the relationship between nominal stock market returns and inflation using the backward/forward aggregation approach. The authors find that h-period continuously-compounded nominal returns are correlated with *past* h-period realized inflation; moreover this dependence increases with the horizon and is statistically significant across countries and time periods. We show here that the same logic as that employed for long-run risk-return trade-offs can be applied to explain their findings. If low-frequency details of the nominal rates depend on low-frequency details of realized inflation, two-way aggregation will uncover this dependence. Direct extraction of the details will subsequently allow us to zoom in on to individual layers of information.

The results are reported in Tables 16 - Panels A, B, and C. For nominal stock market returns, we find a tent-shaped predictability pattern (as aggregation increases) with a peak around 6, 7 years. Importantly, as predictability increases, the corresponding beta coefficient approaches the value of 1. The 1-year beta is 0.44 with a t-statistic of 0.58 and an R^2 of 0.82%. The 7-year beta, on the contrary, is equal to 1.13 with a t-statistic of 4.76 and an R^2 of 28.29%. Turning to the details, at scale $j = 3$, i.e., for frequencies between 4 and 8 years, the R^2 of the detail-wise regression reaches a peak of 38.86%. Analogous results apply to nominal interest rates. With two-way aggregation, the 1-year beta is 0.37 with a t-statistic of 2.92 and an R^2 of 21.00%. The 7-year beta is double and equal to 0.72 with a t-statistic of 3.07 and an R^2 of 38.51%. The corresponding details ($j = 3$) yield a predictive regression with a significant, positive slope. At the same scale, the detail-wise autoregressive coefficient of future inflation on past inflation is positive albeit rather small (0.10) and inaccurately estimated.

In essence, predictable slow-moving components of the inflation process appear to correlate with slow-moving components of nominal stock returns and interest rates. Higher values of past inflation details predict higher future values of the same details, as well as higher nominal rates, thereby yielding compensation for inflation risk.

7 Conclusions

Shocks to economic time series can be time specific and, importantly for our purposes, frequency specific. We suggest that economic relations may apply to individual layers in the cascade of shocks affecting the economy and be hidden by effects at alternative, higher frequencies. These layers, and the frequency at which they operate, can be identified. In particular, the nature and the magnitude

of the existing, low-frequency, economic relations can be studied. To do so, this paper proposes *direct* extraction of the time-series details by virtue of multiresolution time series analysis - and regressions on the details - as well as *indirect* extraction by means of two-way aggregation of the raw series - and regressions on forward/backward aggregates of the raw series. The mapping between the two methods is established and their close relation is exploited empirically. While the direct method allows one to identify, up to the customary estimation error, the data generating process (i.e., the details and, upon reconstruction, the original series), the indirect method provides one with a rather immediate way to evaluate the frequency at which layers in the information flow are connected across economic variables and employ this information for prediction. By re-writing the way in which one may implement long-run predictability (aggregated regressand on past aggregated regressor, rather than on past regressor over one period), two-way aggregation provides a natural way to exploit scale-specific predictability (in asset allocation for the long run, for example). Using both direct extraction of the details and aggregation, we provide evidence of the long-run validity of economic relations, like risk-return trade-offs and Fisher's effect, typically found to be elusive when working with raw data at the highest frequency of observation.

A Multiscale vs. standard Wold decomposition

Start from

$$x_t = \sum_{j=1}^{J} \sum_{k=0}^{\infty} a_{j,k} \varepsilon_{t-k2^j}^{(j)},$$

where we let

$$a_{j,k} = \left\langle x_t, \varepsilon_{t-k2^j}^{(j)} \right\rangle$$

Consider the time window $[t-6, t]$ and $J = 2$:

$$\begin{aligned} x_t =& a_{1,0}\varepsilon_t^{(1)} + a_{1,1}\varepsilon_{t-2}^{(1)} + a_{1,2}\varepsilon_{t-4}^{(1)} + a_{1,3}\varepsilon_{t-6}^{(1)} \\ & a_{2,0}\varepsilon_t^{(2)} + a_{2,1}\varepsilon_{t-4}^{(2)} + a_{J,0}\varepsilon_t^{(J)} + \dots \end{aligned}$$

If

$$\varepsilon_t^{(j)} = \frac{\sum_{i=0}^{2^{j-1}-1} \varepsilon_{t-i} - \sum_{i=0}^{2^{j-1}-1} \varepsilon_{t-2^{j-1}-i}}{\sqrt{2^j}}$$

Then

$$\begin{aligned} a_{1,0} &= \left\langle x_t, \varepsilon_t^{(1)} \right\rangle = \frac{\psi_0}{\sqrt{2}} - \frac{\psi_1}{\sqrt{2}} \\ a_{1,1} &= \left\langle x_t, \varepsilon_{t-2}^{(1)} \right\rangle = \frac{\psi_2}{\sqrt{2}} - \frac{\psi_3}{\sqrt{2}} \\ a_{1,2} &= \left\langle x_t, \varepsilon_{t-4}^{(1)} \right\rangle = \frac{\psi_4}{\sqrt{2}} - \frac{\psi_5}{\sqrt{2}} \\ a_{1,3} &= \left\langle x_t, \varepsilon_{t-6}^{(1)} \right\rangle = \frac{\psi_6}{\sqrt{2}} - \frac{\psi_7}{\sqrt{2}} \\ a_{2,0} &= \left\langle x_t, \varepsilon_t^{(2)} \right\rangle = \frac{\psi_0}{2} + \frac{\psi_1}{2} - \frac{\psi_2}{2} - \frac{\psi_3}{2} \\ a_{2,1} &= \left\langle x_t, \varepsilon_{t-4}^{(2)} \right\rangle = \frac{\psi_4}{2} + \frac{\psi_5}{2} - \frac{\psi_6}{2} - \frac{\psi_7}{2} \\ a_{J,0} &= \left\langle x_t, \varepsilon_t^{(J)} \right\rangle = \frac{\psi_0}{2} + \frac{\psi_1}{2} + \frac{\psi_2}{2} + \frac{\psi_3}{2} + \frac{\psi_4}{2} + \frac{\psi_5}{2} + \frac{\psi_6}{2} + \frac{\psi_7}{2} \end{aligned}$$

where we let

$$\psi_j = \langle x_t, \varepsilon_{t-j} \rangle$$

Finally note that

$$\begin{aligned}
\psi_0\left(\frac{1}{\sqrt{2}}\varepsilon_t^{(1)}+\frac{1}{2}\varepsilon_t^{(2)}+\frac{1}{2}\varepsilon_t^{(J)}\right)&=\psi_0\varepsilon_t\\
\psi_1\left(-\frac{1}{\sqrt{2}}\varepsilon_t^{(1)}+\frac{1}{2}\varepsilon_t^{(2)}+\frac{1}{2}\varepsilon_t^{(J)}\right)&=\psi_1\varepsilon_{t-1}\\
\psi_2\left(\frac{1}{\sqrt{2}}\varepsilon_{t-2}^{(1)}-\frac{1}{2}\varepsilon_t^{(2)}+\frac{1}{2}\varepsilon_t^{(J)}\right)&=\psi_2\varepsilon_{t-2}\\
\psi_3\left(-\frac{1}{\sqrt{2}}\varepsilon_{t-2}^{(1)}-\frac{1}{2}\varepsilon_t^{(2)}+\frac{1}{2}\varepsilon_t^{(J)}\right)&=\psi_3\varepsilon_{t-3}\\
\psi_4\left(\frac{1}{\sqrt{2}}\varepsilon_{t-4}^{(1)}+\frac{1}{2}\varepsilon_{t-4}^{(2)}+\frac{1}{2}\varepsilon_t^{(J)}\right)&=\psi_4\varepsilon_{t-4}\\
\psi_5\left(-\frac{1}{\sqrt{2}}\varepsilon_{t-4}^{(1)}+\frac{1}{2}\varepsilon_{t-4}^{(2)}+\frac{1}{2}\varepsilon_t^{(J)}\right)&=\psi_5\varepsilon_{t-5}\\
\psi_6\left(\frac{1}{\sqrt{2}}\varepsilon_{t-6}^{(1)}-\frac{1}{2}\varepsilon_{t-4}^{(2)}+\frac{1}{2}\varepsilon_t^{(J)}\right)&=\psi_6\varepsilon_{t-6}\\
\psi_7\left(-\frac{1}{\sqrt{2}}\varepsilon_{t-6}^{(1)}-\frac{1}{2}\varepsilon_{t-4}^{(2)}+\frac{1}{2}\varepsilon_t^{(J)}\right)&=\psi_7\varepsilon_{t-7}
\end{aligned}$$

i.e. we obtain the standard Wold Decomposition

$$x_t=\psi_0\varepsilon_t+\psi_1\varepsilon_{t-1}+\psi_2\varepsilon_{t-2}+\dots$$

B A primer on multiresolution analysis

In this section we provide a primer on multiresolution analysis (MRA). The fundamental idea behind MRA is to analyze data according at different scales or resolutions. If we look at a time series with a large window, we would notice coarse features. Similarly, if we look at a time series with a small window, we would notice small features. The goal in wavelet analysis is to see both the forest and the trees, so to speak.

Here we assume for convenience that the original data lies in the space V_0[9]. For each vector space, there is another vector space of higher scale (or lower resolution) V_j and each vector space contains all vector spaces that are of lower resolution. For each vector space V_j, there is an orthogonal complement called $W_{j+1}=V_j-V_{j+1}$ and the basis function for this vector space is the wavelet. Hence we can define the coarser and detail spaces as $(V_1,W_1),(V_2,W_2),\dots$, i.e. increasing the index

[9]For example if we have a deterministic signal of length 8 we can assume V_0 to be $\mathbf{R}^8$

in the V-spaces is equivalent to coarsening the approximation to the data. Now we can represent a data series which lies in V_0 by projecting it onto the detail spaces W_j. In particular define the multiresolution decomposition of a series by:

1. $\pi^{(J)}$ is the coarsest scale, and
2. $\pi^{(J-1)} = \pi^{(J)} + x_t^{(J)}$
3. In general, $\pi^{(j-1)} = \pi^{(j)} + x_t^{(j)}$

where $\pi^{(j)} \in V_j$ and $x_t^{(j)} \in W_j$. We can finally define the multiresolution decomposition of a signal as,

$$\{\pi^{(J)}, x_t^{(J)}, x_t^{(J-1)}, \ldots, x_t^{(j)}, \ldots, x_t^{(1)}\}$$

At high frequencies the wavelet is able to focus in on short-lived phenomena, e.g., a singularity point, whereas at low frequencies the wavelet has a large time support allowing it to identify long periodic behavior. This is a useful property for economic and financial systems where many variables operate on a variety of time-scales simultaneously and where the relationship between variables may be different across time-scales. Given that wavelets allow us to decompose a series by time-scale, it seems then reasonable and appropriate to use this basis for the analysis of economic and financial data. In the next subsection we look at how wavelet transform can be reformulated in terms of matrices and operators.

B.1 Expansion of a random process in a Wavelet-basis

This section explains the structure of wavelet algorithms in the context of linear algebra. Let's work in discrete time and assume the following random process[10] $\omega(t) = \{\omega_0, \omega_1, \omega_2, \omega_3\}$ where for simplicity we can assume:

$$\omega_i = \begin{cases} 1, & \text{with probability } p \\ -1, & \text{with probability } 1-p \end{cases}$$

Our goal is to expand the process into $V^2 \oplus W^2 \oplus W^1$. Let's do first a remark: the coefficients of the original signal are the coefficients of the process $\omega(t)$ expanded in V^0, i.e. $\langle \omega, \phi_0^0 \rangle = \omega_0$,

[10]Note that we assume the length (dimension) of the signal to be 2^j for some positive integer j. This simplifies the analysis. However, the procedure can be generalized to any finite data series.

$\langle \omega, \phi_1^0 \rangle = \omega_1$, $\langle \omega, \phi_2^0 \rangle = \omega_2$, $\langle \omega, \phi_3^0 \rangle = \omega_3$. Recall that the basis vectors of V^0 are translation of the Haar mother scaling function which has been dilated so that each basis function has a support equal 1/4.

The first step is to expand $\omega(t)$ into $V^1 \oplus W^1$. We carry out the matrix multiplication:

$$W_1 = \mathcal{T}_1 \omega(t)$$

where

$$\mathcal{T}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

Notice that the first two rows correspond to the basis vectors ϕ_0^1 and ϕ_1^1 which span V^1. The last two rows correspond to the basis vectors ψ_0^1 and ψ_1^1 which span W^1. The matrix multiplication yields 2 averages

$$\pi_1^{(1)} = \frac{(\omega_0 + \omega_1)}{\sqrt{2}}$$
$$\pi_3^{(1)} = \frac{(\omega_2 + \omega_3)}{\sqrt{2}}$$

and 2 wavelet coefficients

$$\delta_1^{(1)} = \frac{(\omega_0 - \omega_1)}{\sqrt{2}}$$
$$\delta_3^{(1)} = \frac{(\omega_2 - \omega_3)}{\sqrt{2}}$$

Hence our process expanded into $V^1 \oplus W^1$ becomes:

$$\omega(t) = \underbrace{\pi_1^{(1)} \phi_0^1 + \pi_3^{(1)} \phi_1^1}_{V^1} + \underbrace{\delta_1^{(1)} \psi_0^1 + \delta_3^{(1)} \psi_1^1}_{W^1}$$

In the next step we expand the process into $V^2 \oplus W^2 \oplus W^1$. We carry out another matrix multiplication

$$W_2 = \mathcal{T}_2 W_1$$

where

$$\mathcal{T}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Note the identity matrix in the bottom left side of $\mathcal{T}_2$. This is because we already have the random coefficients for the basis vectors of W^1 and therefore only the averages $(\pi_1^{(1)}, \pi_3^{(1)})$ become the input for the next step. or combining with the previous step:

$$W_2 = \mathcal{T}_2 \mathcal{T}_1 \omega(t)$$

where

$$\mathcal{T}_2\mathcal{T}_1 = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

Notice that the first row corresponds to the basis vector ϕ_0^2 which spans V^2, the second row corresponds to the basis vector ψ_0^2 which spans W^2. Eventually the process expanded into $V^2 \oplus W^2 \oplus W^1$ becomes:

$$\omega(t) = \underbrace{\pi_4^{(2)} \phi_0^2}_{V^2} + \underbrace{\delta_4^{(2)} \psi_0^2}_{W^2} + \underbrace{\delta_1^{(1)} \psi_0^1 + \delta_3^{(1)} \psi_1^1}_{W^1}$$

where the average and the wavelet coefficients are given by:

$$\pi_4^{(2)} = \frac{\sum_{i=0}^{3} \omega_i}{2}$$
$$\delta_4^{(2)} = \frac{(\omega_0 + \omega_1)/\sqrt{2} - (\omega_2 + \omega_3)/\sqrt{2}}{\sqrt{2}}$$

More generally if the data set $\mathbf{x}_0 = \{x_{0,k}\}_{k=0}^{T-1}$ contains T elements, then the recursive iterations would continue until a single average is calculated. Indeed the matrix is first applied to the original, full-length data vector. Then the vector is smoothed and the matrix applied again. This process would replace the original data set of N elements with an average $\pi_T^{(J)}$, which we define as the smooth

component, followed by a set of wavelet coefficients $D_J \equiv \{\delta_k^{(J)}\}_k, \delta^{(J-1)} \equiv \{\delta_k^{(J-1)}\}_k, \ldots, \delta^{(1)} \equiv \{\delta_k^{(1)}\}_k$, which make up the details components. All in all the discrete wavelet transform of the sequence $\mathbf{x}$ can then be represented as:

$$\mathbf{x}_0 \to \{\pi^{(J)}, \delta^{(J)}, \ldots, \delta^{(j)}, \ldots, \delta^{(2)}, \delta^{(1)}\} \tag{B.1}$$

Note that denoting by $\boldsymbol{W}$ is an $T \times 1$ vector of coefficients given by

$$\mathbf{W} = \mathcal{T}\mathbf{X}$$

and

$$\mathbf{W} = \begin{pmatrix} \boldsymbol{\pi}^{(J)} \\ \boldsymbol{\delta}^{(J)} \\ \vdots \\ \boldsymbol{\delta}^{(2)} \\ \boldsymbol{\delta}^{(1)} \end{pmatrix}.$$

C Two-way aggregation: additional details

C.1 Dynamics of time-scale components

Assume the following components (or details) dynamics for $j = j*$ where $j* \in \{1, \ldots, J\}$

$$y_{t+2^j}^{(j)} = \beta_j x_t^{(j)} \tag{C.1}$$

$$x_{t+2^j}^{(j)} = \rho_j x_t^{(j)} + \sigma_j \epsilon_{t+2^j} \tag{C.2}$$

and for $j = 1, \ldots, J$ and $j \neq j*$

$$y_t^{(j)} = 0 \tag{C.3}$$

$$x_t^{(j)} = 0 \tag{C.4}$$

Assume for simplicity that $T = 16$, $j* = 2$ and $J = 3$. Arrange the details of x as follows:

$$\begin{pmatrix} \pi_8^{(3)} & \pi_{16}^{(3)} \\ x_8^{(3)} & x_{16}^{(3)} \\ x_8^{(2)} & x_{16}^{(2)} \\ x_4^{(2)} & x_{12}^{(2)} \\ x_8^{(1)} & x_{16}^{(1)} \\ x_6^{(1)} & x_{14}^{(1)} \\ x_4^{(1)} & x_{12}^{(1)} \\ x_2^{(1)} & x_{10}^{(1)} \end{pmatrix} \tag{C.5}$$

and analogously for the details of y. Consider the isometric transformation matrix:

$$\mathcal{T}^{(3)} = \begin{pmatrix} \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} \\ \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & -\frac{1}{\sqrt{8}} & -\frac{1}{\sqrt{8}} & -\frac{1}{\sqrt{8}} & -\frac{1}{\sqrt{8}} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \tag{C.6}$$

To reconstruct the time series of x, we run through each column of the matrix (C.5) and for each column we perform the following operation:

$$X_8^{(3)} = \begin{pmatrix} x_8 \\ x_7 \\ x_6 \\ x_5 \\ x_4 \\ x_3 \\ x_2 \\ x_1 \end{pmatrix} = \left(\mathcal{T}^{(3)}\right)^{-1} \begin{pmatrix} \pi_8^{(3)} \\ x_8^{(3)} \\ x_8^{(2)} \\ x_4^{(2)} \\ x_8^{(1)} \\ x_6^{(1)} \\ x_4^{(1)} \\ x_2^{(1)} \end{pmatrix} \tag{C.7}$$

and

$$X_{16}^{(3)} = \begin{pmatrix} x_{16} \\ x_{15} \\ x_{14} \\ x_{13} \\ x_{12} \\ x_{11} \\ x_{10} \\ x_9 \end{pmatrix} = \left(\mathcal{T}^{(3)}\right)^{-1} \begin{pmatrix} \pi_{16}^{(3)} \\ x_{16}^{(3)} \\ x_{16}^{(2)} \\ x_{12}^{(2)} \\ x_{16}^{(1)} \\ x_{14}^{(1)} \\ x_{12}^{(1)} \\ x_{10}^{(1)} \end{pmatrix} \tag{C.8}$$

and analogously for the details of y. The matrix $\left(\mathcal{T}^{(3)}\right)^{-1}$ takes the following expression

$$\left(\mathcal{T}^{(3)}\right)^{-1} = \begin{pmatrix} \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{2} & 0 & \frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{2} & 0 & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & -\frac{1}{2} & 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & -\frac{1}{2} & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{\sqrt{8}} & -\frac{1}{\sqrt{8}} & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{8}} & -\frac{1}{\sqrt{8}} & 0 & \frac{1}{2} & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{8}} & -\frac{1}{\sqrt{8}} & 0 & -\frac{1}{2} & 0 & 0 & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{8}} & -\frac{1}{\sqrt{8}} & 0 & -\frac{1}{2} & 0 & 0 & 0 & -\frac{1}{\sqrt{2}} \end{pmatrix} \tag{C.9}$$

Using the dynamics for the state (C.2), and (C.7) and (C.8) we obtain

$$X_{16}^{(3)} = \begin{pmatrix} x_{16} = & x_{16}^{(2)}/2 \\ x_{15} = & x_{16}^{(2)}/2 \\ x_{14} = & -x_{16}^{(2)}/2 \\ x_{13} = & -x_{16}^{(2)}/2 \\ x_{12} = & x_{12}^{(2)}/2 \\ x_{11} = & x_{12}^{(2)}/2 \\ x_{10} = & -x_{12}^{(2)}/2 \\ x_{9} = & -x_{12}^{(2)}/2 \end{pmatrix} \tag{C.10}$$

and

$$X_{8}^{(3)} = \begin{pmatrix} x_{8} = & x_{8}^{(2)}/2 \\ x_{7} = & x_{8}^{(2)}/2 \\ x_{6} = & -x_{8}^{(2)}/2 \\ x_{5} = & -x_{8}^{(2)}/2 \\ x_{4} = & x_{4}^{(2)}/2 \\ x_{3} = & x_{4}^{(2)}/2 \\ x_{2} = & -x_{4}^{(2)}/2 \\ x_{1} = & -x_{4}^{(2)}/2 \end{pmatrix} \tag{C.11}$$

C.2 Aggregation

C.2.1 Fitting an AR(1) process to the regressor

We fit an AR(1) process to x_t

$$x_{t+1} = \rho x_{,t} + \epsilon_{t+1}$$

From (C.10) and (C.11) it is easy to see that

$$\rho \to \frac{1 - \rho_j}{4}$$

More generally if the process for x_t is given by (C.2) and (C.4) then

$$\rho \to \frac{\underbrace{1+1+\ldots}_{2^{j^*-1}-1}-1+\underbrace{1+1+\ldots}_{2^{j^*-1}-1}-\rho_j}{2^{j^*}}$$

C.2.2 Regressor and regressand are aggregated over non-overlapping periods

Let us construct the (temporally) aggregated series

$$y_{t+1,t+h} = \sum_{i=1}^{h} y_{t+i}$$

and run the regression of

$$y_{t+1,t+h} = \tilde{\beta} x_{t-h,t} + \epsilon_{t,t+h}$$

For $h = 4$, and using (C.1) and (C.3) together with (C.10) and (C.11), we have

$$\begin{aligned}
y_{13,16} &= 0 & x_{13,16} &= 0 \\
y_{12,15} &= -y_{16}^{(2)} + y_{12}^{(2)} = \beta\left(-x_{12}^{(2)} + x_{8}^{(2)}\right) & x_{12,15} &= -x_{16}^{(2)} + x_{12}^{(2)} \\
y_{11,14} &= -2y_{16}^{(2)} + 2y_{12}^{(2)} = \beta\left(-2x_{12}^{(2)} + 2x_{8}^{(2)}\right) & x_{11,14} &= -2x_{16}^{(2)} + 2x_{12}^{(2)} \\
y_{10,13} &= -y_{16}^{(2)} + y_{12}^{(2)} = \beta\left(-x_{12}^{(2)} + x_{8}^{(2)}\right) & x_{10,13} &= -x_{16}^{(2)} + x_{12}^{(2)} \\
y_{9,12} &= 0 & x_{9,12} &= 0 \\
y_{8,11} &= -y_{12}^{(2)} + y_{8}^{(2)} = \beta\left(-x_{8}^{(2)} + x_{4}^{(2)}\right) & x_{8,11} &= -x_{12}^{(2)} + x_{8}^{(2)} \\
y_{7,10} &= -2y_{12}^{(2)} + 2y_{8}^{(2)} = \beta\left(-2x_{8}^{(2)} + 2x_{4}^{(2)}\right) & x_{7,10} &= -2x_{12}^{(2)} + 2x_{8}^{(2)} \\
y_{6,9} &= -y_{12}^{(2)} + y_{8}^{(2)} = \beta\left(-x_{8}^{(2)} + x_{4}^{(2)}\right) & x_{6,9} &= -x_{12}^{(2)} + x_{8}^{(2)} \\
y_{5,8} &= 0 & x_{5,8} &= 0 \\
y_{4,7} &= -y_{8}^{(2)} + y_{4}^{(2)} = \beta\left(-x_{4}^{(2)} + x_{0}^{(2)}\right) & x_{4,7} &= -x_{8}^{(2)} + x_{4}^{(2)} \\
y_{3,6} &= -2y_{8}^{(2)} + 2y_{4}^{(2)} = \beta\left(-2x_{4}^{(2)} + 2x_{0}^{(2)}\right) & x_{3,6} &= -2x_{8}^{(2)} + 2x_{4}^{(2)} \\
y_{2,5} &= -y_{8}^{(2)} + y_{4}^{(2)} = \beta\left(-x_{4}^{(2)} + x_{0}^{(2)}\right) & x_{2,5} &= -x_{8}^{(2)} + x_{4}^{(2)} \\
y_{1,4} &= 0 & x_{1,4} &= 0
\end{aligned}$$

where we use the implied dynamics for x, see equations (C.10) and (C.11), and the equivalent ones for y together with (C.1) and (C.2). The regression of $y_{t+1,t+4}$ on $x_{t-4,t}$ yields $\tilde{\beta} = \beta$ with $R^2 = 100\%$.

For $h = 2$ we have

$$
\begin{aligned}
y_{15,16} &= y_{16}^{(2)} = \beta x_{12}^{(2)} & x_{15,16} &= x_{16}^{(2)} \\
y_{14,15} &= 0 & x_{14,15} &= 0 \\
y_{13,14} &= -y_{16}^{(2)} = -\beta x_{12}^{(2)} & x_{13,14} &= -x_{16}^{(2)} \\
y_{12,13} &= -y_{16}^{(2)} + y_{12}^{(2)} = \beta(-x_{12}^{(2)} + x_{8}^{(2)}) & x_{12,13} &= -x_{16}^{(2)} + x_{12}^{(2)} \\
y_{11,12} &= y_{12}^{(2)} = \beta x_{8}^{(2)} & x_{11,12} &= x_{12}^{(2)} \\
y_{10,11} &= 0 & x_{10,11} &= 0 \\
y_{9,10} &= -y_{12}^{(2)} = -\beta x_{8}^{(2)} & x_{9,10} &= -x_{12}^{(2)} \\
y_{8,9} &= -y_{12}^{(2)} + y_{8}^{(2)} = \beta(-x_{8}^{(2)} + x_{4}^{(2)}) & x_{8,9} &= -x_{12}^{(2)} + x_{8}^{(2)} \\
y_{7,8} &= y_{8}^{(2)} = \beta x_{4}^{(2)} & x_{7,8} &= x_{8}^{(2)} \\
y_{6,7} &= 0 & x_{6,7} &= 0 \\
y_{5,6} &= -y_{8}^{(2)} = -\beta x_{4}^{(2)} & x_{5,6} &= -x_{8}^{(2)} \\
y_{4,5} &= -y_{8}^{(2)} + y_{4}^{(2)} = \beta(-x_{4}^{(2)} + x_{0}^{(2)}) & x_{4,5} &= -x_{8}^{(2)} + x_{4}^{(2)} \\
y_{3,4} &= y_{4}^{(2)} = \beta x_{0}^{(2)} & x_{3,4} &= x_{4}^{(2)} \\
y_{2,3} &= 0 & x_{2,3} &= 0 \\
y_{1,2} &= -y_{4}^{(2)} = \beta x_{0}^{(2)} & x_{1,2} &= -x_{4}^{(2)}
\end{aligned}
$$

where we use the implied dynamics for x, see equations (C.10) and (C.11), and the equivalent ones for y together with (C.1) and (C.2).

The regression of $y_{t+1,t+2}$ on $x_{t-2,t}$ yields $\tilde{\beta} = -\beta\frac{(1+\rho)}{3}$ coming from the regression of $y_{15,16} = \beta x_{12}^{(2)}$ onto $x_{13,14} = -x_{16}^{(2)}$ and from the regression of $y_{13,14} = -\beta x_{12}^{(2)}$ onto $x_{11,12} = x_{12}^{(2)}$.

C.2.3 Contemporaneous aggregation

Let us construct the aggregated series

$$y_{t+1,t+h} = \sum_{i=1}^{h} y_{t+i}$$

and run the regression of

$$y_{t+1,t+h} = \tilde{\beta} x_{t+1,t+h} + \epsilon_{t+1,t+h}$$

The generic term for $h = 4$ takes the following form:

$$\begin{aligned} y_{t+1,t+h} &= \beta(-x_{12}^{(2)} + x_{8}^{(2)}) \\ x_{t+1,t+h} &= (-x_{16}^{(2)} + x_{12}^{(2)}) \end{aligned}$$

and by taking the covariances we estimate

$$\begin{aligned} \tilde{\beta} &= \beta \frac{\left(-\operatorname{Var}(x_{12}^{(2)}) + \rho_j \operatorname{Var}(x_{12}^{(2)}) - \rho_j^2 \operatorname{Var}(x_{12}^{(2)}) + \rho_j \operatorname{Var}(x_{12}^{(2)}) \right)}{\operatorname{Var}(x_{8}^{(2)}) + \operatorname{Var}(x_{12}^{(2)}) - 2cov(x_{16}^{(2)}, x_{12}^{(2)})} \\ &= \beta \frac{\left(-1 + 2\rho_j - \rho_j^2 \right)}{2(1-\rho_j)} \\ &= -\beta \frac{(1-\rho_j)}{2} \end{aligned}$$

The R^2 is equal to

$$R^2 = \frac{\operatorname{Var} \widehat{y}_{t+1,t+h}}{\operatorname{Var} y_{t+1,t+h}} = \frac{\tilde{\beta}^2 \operatorname{Var}\left(-x_{12}^{(2)} + x_{8}^{(2)} \right)}{\beta^2 \operatorname{Var}\left(-x_{16}^{(2)} + x_{12}^{(2)} \right)} = \left(\frac{1-\rho_j}{2} \right)^2.$$

D Data

The empirical analysis in Sections 4.1, 4.2 and 6 is conducted using annual data on consumption, inflation, stock returns and short-term interest rate for the time period from 1930 till 2011 - the longest available sample. We take the view that this sample better represents the overall variation in asset and macro economic data. Importantly, this long sample also helps in detecting low frequency

variations in the time series of interest.

Aggregate consumption is from the Bureau of Economic Analysis (BEA), series 7.1 and is defined as consumer expenditures on non-durables and services. Growth rates are constructed by taking the first difference of the corresponding log series. Our measure of consumption volatility is based on modeling consumption growth as following an AR(1)-GARCH(1,1), as in Bansal, Khatchatrian, and Yaron (2005).

For yearly inflation, we use the seasonally unadjusted CPI from the Bureau of Labor Statistics. Yearly inflation is the log December over December growth rate in the CPI.

We use the NYSE/Amex value-weighted index with dividends as our market proxy, R_{t+1}. Return data on the value-market index are obtained from the Chicago Center for Research in Security Prices (CRSP). The nominal short-term rate ($R_{f,t+1}$) is the annualized yield on the 3-month Treasury bill taken from the CRSP treasury files. The h-horizon continuously compounded excess return is calculated as $r_{t,t+h} = r^e_{t+1} + \ldots + r^e_{t+h}$ where $r^e_{t+j} = \ln(R_{t+j}) - \ln(R_{f,t+j})$ is the 1-year excess log stock return between dates $t+j-1$ and $t+j$; R_{t+j} is the simple gross return; and $R_{f,t+j}$ is the gross risk-free rate (3-month Treasury bill) at the beginning of period $t+j$.

The empirical analysis in Section 4.3 is conducted using returns on bonds extracted from the Fixed Term Indices File of the Chicago Center of Research in Security Prices (CRSP). The CRSP Fixed Term Indices File provides daily total returns and yields on constant maturity coupon bonds, with maturities ranging from 1 year to 30 years, starting on June 14, 1961.

The realized variance of stock and bond returns between the end of period t and the end of period $t+n$, a measure of the integrated instantaneous volatility, is measured as

$$v^2_{t,t+n} = \sum_{d=t_1}^{t_D} r^2_{i,d}$$

where $[t_1, t_D]$ denotes the sample of the available daily returns between the end of period t and the end of period $t+n$, $i =$ S and B for stocks and bonds respectively and $r_{S,d}$ and $r_{B,d}$ are the stock and bond log returns on day d. We build time series of realized second moments of bond and stock returns using daily returns on stocks and bonds extracted from the Fixed Term Indices File and the Daily Stock Indices File of the Center of Research in Security Prices (CRSP) at the University of Chicago.

References

Abry, P., D. Veitch, and P. Flandrin (1998): "Long-Range Dependence: revisiting Aggregation with Wavelets.," *Journal of Time Series Analysis*, 19, 253–266.

Bandi, F. M., and B. Perron (2008): "Long-run risk-return trade-offs," *Journal of Econometrics*, 143(2), 349–374.

Bansal, R., V. Khatchatrian, and A. Yaron (2005): "Interpretable asset markets?," *European Economic Review*, 49(3), 531 – 560.

Barndorff-Nielsen, O. E., and N. Shephard (2001): "Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics," *Journal Of The Royal Statistical Society Series B*, 63(2), 167–241.

Beveridge, S., and C. R. Nelson (1981): "A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle'," *Journal of Monetary Economics*, 7(2), 151–174.

Boudoukh, J., and M. Richardson (1993): "Stock Returns and Inflation: A Long-Horizon Perspective," *American Economic Review*, 83(5), 1346–55.

Dijkerman, R., and R. Mazumdar (1994): "Wavelet representations of stochastic processes and multiresolution stochastic models," *Signal Processing, IEEE Transactions on*, 42(7), 1640 –1652.

Fisher, M. E., and J. J. Seater (1993): "Long-Run Neutrality and Superneutrality in an ARIMA Framework," *American Economic Review*, 83(3), 402–15.

Gençay, R., F. Selçuk, and B. Whitcher (2001): *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*. Academic Press, New York, first edn.

Hansen, L. P., and J. A. Sheinkman (2009): "Long-term Risk: An Operator Approach," *Econometrica*, 77(1), 177–234.

Mallat, S. G. (1989): "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, 11, 674–693.

Mishkin, F. S. (1992): "Is the Fisher effect for real? : A reexamination of the relationship between inflation and interest rates," *Journal of Monetary Economics*, 30(2), 195–215.

Müller, U. K., and M. W. Watson (2008): "Testing Models of Low-Frequency Variability," *Econometrica*, 76(5), 979–1016.

Ortu, F., A. Tamoni, and C. Tebaldi (2011): "Long-Run Risk and the Persistence of Consumption Shocks," Working Papers 14260.

Percival, D. B., and A. T. Walden (2000): *Wavelet Methods for Time Series Analysis (Cambridge Series in Statistical and Probabilistic Mathematics)*. Cambridge University Press.

Renaud, O., J.-L. Starck, and F. Murtagh (2005): "Wavelet-Based Combined Signal Filtering and Prediction," *IEEE Transactions SMC, Part B*, 35, 1241 – 1251.

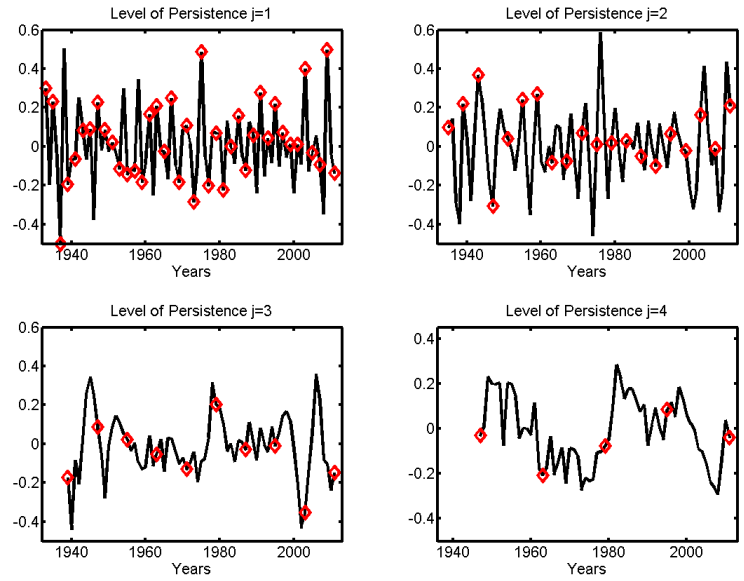
Tamoni, A. (2011): "The multi-horizon dynamics of risk and returns," *SSRN eLibrary*.

Tamoni, A., and A. Gozluklu (2012): "Nominal inertia and stock returns over the long-run," *Working Paper*.

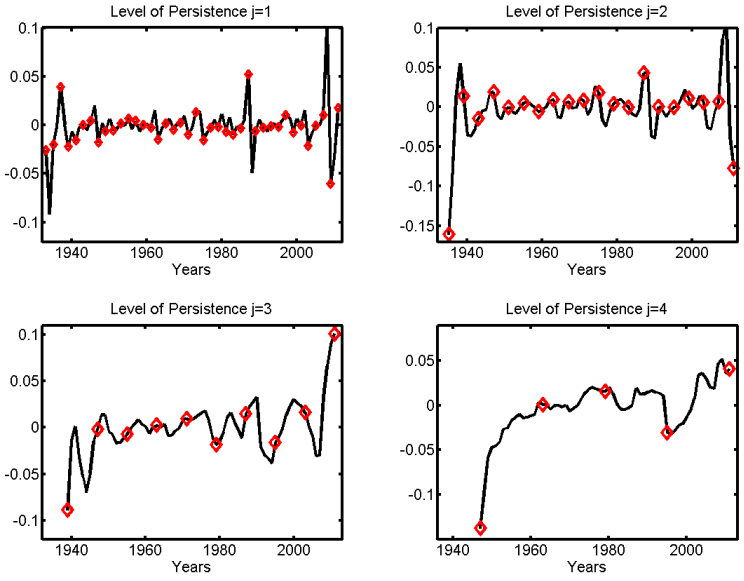
Wong, P. W. (1993): "Wavelet decomposition of harmonizable random processes," *IEEE Transactions on Information Theory*, 39(1), 7–18.

Yazici, B. (2004): "Stochastic deconvolution over groups," *IEEE Transactions on Information Theory*, 20(3), 494–510.

Yazici, B., and R. Kashyap (1997): "A class of second order self-similar processes for 1/f phenomena," *IEEE Transactions on Signal Processing*, 45(2), 396–410.



(a) Time-scale decomposition for market returns



(b) Time-scale decomposition for market realized variance

Figure 1: Panel A displays the time-scale decomposition for the excess stock market returns. Panel B displays the time-scale decomposition for market realized variance.

Panel A: Annual calendar time

Time-scale	Frequency resolution
$j = 1$	$1 - 2$ years
$j = 2$	$2 - 4$ years
$j = 3$	$4 - 8$ years
$j = 4$	$8 - 16$ years
$j = 5$	$16 - 32$ years
$\pi_t^{(5)}$	> 32 years

Panel B: Quarterly calendar time

Time-scale	Frequency resolution
$j = 1$	$1 - 2$ quarters
$j = 2$	$2 - 4$ quarters
$j = 3$	$1 - 2$ years
$j = 4$	$2 - 4$ years
$j = 5$	$4 - 8$ years
$j = 6$	$8 - 16$ years
$j = 7$	$16 - 32$ years
$\pi_t^{(7)}$	> 32 years

Panel C: Monthly calendar time

Time-scale	Frequency resolution
$j = 1$	$1 - 2$ months
$j = 2$	$2 - 4$ months
$j = 3$	$4 - 8$ months
$j = 4$	$8 - 16$ months
$j = 5$	$16 - 32$ months
$j = 6$	$32 - 64$ months
$j = 7$	$64 - 128$ months
$\pi_t^{(7)}$	> 128 years

Table 1: Interpretation of the time-scale (or persistence level) j in terms of time spans in the case of annual (Panel A), quarterly (Panel B) and monthly (Panel C) time series.

	Horizon h (in months)											
	3	6	12	24	36	48	60	72	84	96	108	120
$v^2_{t-h,t}$	0.04	0.04	-0.03	-0.34	-0.21	0.36	0.62	0.58	0.33	0.07	0.04	0.09
	(0.64)	(0.60)	(-0.39)	(-3.52)	(-1.96)	(3.92)	(7.36)	(6.92)	(3.52)	(0.71)	(0.45)	(0.83)
$Adj.R^2$	[0.12]	[0.35]	[0.70]	[10.56]	[5.35]	[12.67]	[36.56]	[32.61]	[12.05]	[3.12]	[1.66]	[3.05]

Table 2: Simulation. We simulate excess market returns and market variance under the assumption of predictability at scale $j^* = 6$. We simulate $v^2_{t,j} = \rho_j v^2_{t-2^j,j} + \epsilon_{t,j}$ for $j = 6$ and $v^2_{t,j} = \epsilon_{t,j}$ otherwise. We implement 500 replications. We set $T = 1024$. We run linear regressions (with an intercept) of h-period continuously compounded excess market returns on h-period past realized market variances. For each regression, the table reports OLS estimates of the regressors, Newey-West t-statistics with 2*(horizon-1) lags in parentheses and adjusted R^2 statistics in square brackets.

	Horizon h (in months)											
	3	6	12	24	36	48	60	72	84	96	108	120
$v^2_{t,t+h}$	0.04	0.08	0.12	0.13	0.07	-0.07	-0.20	-0.17	-0.01	0.10	0.13	0.10
	(0.72)	(1.27)	(1.53)	(1.36)	(0.61)	(-0.51)	(-1.48)	(-1.26)	(0.08)	(0.98)	(1.20)	(0.87)
$Adj.R^2$	[0.18]	[0.75]	[2.24]	[4.11]	[3.73]	[3.64]	[6.40]	[5.52]	[3.76]	[4.69]	[4.68]	[3.96]

Table 3: Simulation - contemporaneous aggregation. We simulate continuously compounded excess market returns and market variances under the assumption of predictability at scale $j^* = 6$. We implement 500 replications. We set $T = 1024$. We run linear regressions (with an intercept) of h-period continuously compounded excess market returns on h-period realized market variances. For each regression, the table reports OLS estimates of the regressors, Newey-West t-statistics with 2*(horizon-1) lags in parentheses and adjusted R^2 statistics in square brackets.

	Horizon h (in months)											
	3	6	12	24	36	48	60	72	84	96	108	120
$v^2_{t-h,t}$	0.07	0.09	0.04	-0.32	-0.47	-0.39	-0.36	-0.37	-0.41	-0.46	-0.45	-0.44
	(1.64)	(1.62)	(0.48)	(-3.44)	(-4.55)	(-3.61)	(-2.79)	(-2.82)	(-3.48)	(-3.69)	(-3.75)	(-3.34)
$Adj.R^2$	[0.74]	[1.40]	[0.80]	[11.04]	[23.06]	[16.09]	[15.27]	[15.34]	[18.43]	[22.93]	[21.44]	[21.57]

Table 4: We simulate $v^2_{t,j} = \rho_j v^2_{t-2^j,j} + \epsilon_{t,j}$ for $j = 6$ and $v^2_{t,j} = \epsilon_{t,j}$ otherwise. We implement 500 replications. We set $T = 1024$. We run linear regressions of h-period realized market variances $x_{t,t+h}$ on h-period past realized market variances $x_{t-h,t}$. For each regression, the table reports OLS estimates of the regressors, Newey-West t-statistics with 2*(horizon-1) lags in parentheses and adjusted R^2 statistics in square brackets.

	Horizon h (in months)											
	3	6	12	24	36	48	60	72	84	96	108	120
$v^2_{t-h,t}$	-0.00	-0.00	0.01	0.00	-0.01	-0.01	-0.00	-0.00	-0.00	-0.00	-0.01	-0.00
	(-0.04)	(-0.01)	(0.19)	(0.05)	(-0.13)	(-0.10)	(-0.04)	(-0.07)	(-0.05)	(-0.08)	(-0.15)	(-0.07)
$Adj.R^2$	[0.02]	[0.10]	[0.36]	[0.75]	[0.86]	[1.03]	[0.99]	[1.10]	[1.24]	[1.14]	[1.42]	[1.26]

Table 5: Simulation. We simulate excess market returns and market variance under the assumption of no predictability. We simulate $v^2_{t,j} = \rho_j v^2_{t-2^j,j} + \epsilon_{t,j}$ for $j = 6$ and $v^2_{t,j} = \epsilon_{t,j}$ otherwise. We implement 500 replications. We set $T = 1024$. We then run linear regressions (with an intercept) of h-period continuously compounded excess market returns on h-period past realized market variances. For each regression, the table reports OLS estimates of the regressors, Newey-West t-statistics with 2*(horizon-1) lags in parentheses and adjusted R^2 statistics in square brackets.

	Horizon h (in months)											
	3	6	12	24	36	48	60	72	84	96	108	120
$v^2_{t-h,t}$	0.04	0.05	-0.01	-0.29	-0.18	0.24	0.37	0.31	0.16	0.01	0.00	0.03
	(0.66)	(0.72)	(-0.19)	(-3.19)	(-1.83)	(2.76)	(4.33)	(3.84)	(1.88)	(0.17)	(0.13)	(0.44)
$Adj.R^2$	[0.15]	[0.37]	[0.58]	[9.36]	[4.91]	[8.34]	[22.00]	[18.37]	[6.86]	[3.30]	[2.80]	[3.94]

Table 6: Simulation. We simulate excess market returns and market variance under the assumption of predictability at scale $j^* = 6$. We simulate $v^2_{t,j} = \rho_j v^2_{t-2^j,j} + \epsilon_{t,j}$ for $j = 6, 9$ and $v^2_{t,j} = \epsilon_{t,j}$ otherwise. We implement 500 replications. We set $T = 1024$. We then run linear regressions (with an intercept) of h-period continuously compounded excess market returns on h-period past realized market variances. For each regression, the table reports OLS estimates of the regressors, Newey-West t-statistics with 2*(horizon-1) lags in parentheses and adjusted R^2 statistics in square brackets.

	Horizon h (in months)											
	3	6	12	24	36	48	60	72	84	96	108	120
$v^2_{t,t+h}$	0.05	0.09	0.12	0.13	0.07	-0.03	-0.12	-0.09	-0.00	0.06	0.06	0.04
	(0.87)	(1.38)	(1.65)	(1.48)	(0.75)	(-0.29)	(-1.13)	(-0.86)	(-0.01)	(0.62)	(0.68)	(0.45)
$Adj.R^2$	[0.19]	[0.82]	[2.37]	[4.16]	[3.34]	[2.66]	[4.64]	[4.27]	[3.46]	[4.19]	[4.26]	[4.02]

Table 7: Simulation - contemporaneous aggregation. We simulate continuously compounded excess market returns and market variances under the assumption of predictability at scale $j^* = 6$. We simulate $v^2_{t,j} = \rho_j v^2_{t-2^j,j} + \epsilon_{t,j}$ for $j = 6, 9$ and $v^2_{t,j} = \epsilon_{t,j}$ otherwise. We implement 500 replications. We set $T = 1024$. We run linear regressions (with an intercept) of h-period continuously compounded excess market returns on h-period realized market variances. For each regression, the table reports OLS estimates of the regressors, Newey-West t-statistics with 2*(horizon-1) lags in parentheses and adjusted R^2 statistics in square brackets.

	Horizon h (in months)											
	3	6	12	24	36	48	60	72	84	96	108	120
$v^2_{t-h,t}$	0.11	0.15	0.13	-0.13	-0.18	-0.02	0.06	0.08	0.04	-0.01	-0.03	-0.07
	(2.53)	(2.63)	(1.65)	(-1.39)	(-1.65)	(-0.28)	(0.36)	(0.38)	(0.02)	(-0.29)	(-0.40)	(-0.57)
$Adj.R^2$	[1.56]	[3.05]	[2.92]	[5.24]	[9.48]	[7.15]	[8.83]	[8.92]	[7.89]	[7.21]	[5.86]	[5.15]

Table 8: We simulate $v^2_{t,j} = \rho_j v^2_{t-2^j,j} + \epsilon_{t,j}$ for $j = 6, 9$ and $v^2_{t,j} = \epsilon_{t,j}$ otherwise. We implement 500 replications. We set $T = 1024$. We run linear regressions of h-period realized market variances $x_{t,t+h}$ on h-period past realized market variances $x_{t-h,t}$. For each regression, the table reports OLS estimates of the regressors, Newey-West t-statistics with 2*(horizon-1) lags in parentheses and adjusted R^2 statistics in square brackets.

Panel A: $(r_{t,t+h} - rf_t) = \alpha_h + \beta_h v_{t-h,t} + \epsilon_{t+h}$

	Horizon h (in years)									
	1	2	3	4	5	6	7	8	9	10
$v_{t-h,t}$	0.19	0.12	-0.21	-0.02	0.36	0.58	0.74	1.06	1.41	1.53
	(0.90)	(0.39)	(-0.73)	(-0.07)	(1.16)	(1.71)	(2.25)	(3.57)	(5.71)	(7.34)
$R^2(\%)$	[0.53]	[0.29]	[0.95]	[0.01]	[3.78]	[10.41]	[16.51]	[29.19]	[45.09]	[50.71]

Panel B: $r^{(j)}_{t+2^j} - rf^{(j)}_{t+2^j} = \beta_j v^{(j)}_t + \epsilon_{t+2^j}$

	Time-scale j			
	1	2	3	4
$v^{(j)}_t$	0.81	0.21	-0.58	1.33
	(0.37)	(0.14)	(-0.35)	(3.01)
$R^2(\%)$	[1.26]	[2.83]	[2.24]	[72.70]

Panel C: $v^{(j)}_{t+2^j} = \rho_j v^{(j)}_t + \epsilon_{t+2^j}$

	Time-scale j			
	1	2	3	4
$v^{(j)}_t$	-0.17	-0.06	-0.08	0.44
	(-0.95)	(-0.04)	(-0.58)	(1.38)
$R^2(\%)$	[3.32]	[5.01]	[17.42]	[25.76]

Table 9: **Market Volatility. Panel A:** We run linear regressions (with an intercept) of h-period continuously compounded market returns on the CRSP value-weighted index in excess of a 1-year Treasury bill rate on h-period past market variance. For each regression, the table reports OLS estimates of the regressors, Hansen and Hodrick corrected t-statistics in parentheses. **Panel B:** results of componentwise predictive regressions of the components of excess stock market returns on the components of market variance. For each regression, the table reports OLS estimates of the regressors, OLS t-statistics in parentheses and R^2 statistics in square brackets. **Panel C:** estimation results of the multiscale autoregressive system. For each level of persistence $j \in \{1, \ldots, 4\}$, we run a regression of the market variance component $v^{(j)}_{t+2^j}$ on its own lagged component $v^{(j)}_t$. For each regression, the table reports OLS estimates of the regressors, t-statistics in parentheses and adjusted R^2 statistics in square brackets. The sample is annual and spans the period 1930-2011. For the translation of time-scales into appropriate range of time horizons refer to Table 1.

Panel A: $(r_{t,t+h} - rf_t) = \alpha_h + \beta_h v_{t-h,t} + \epsilon_{t+h}$

	Horizon h (in years)									
	1	2	3	4	5	6	7	8	9	10
$v_{t-h,t}$	1.69	1.75	0.60	0.31	0.28	1.47	2.78	3.14	4.16	5.13
	(1.12)	(1.69)	(0.62)	(0.30)	(0.21)	(1.17)	(2.53)	(2.58)	(3.62)	(4.82)
$R^2(\%)$	[1.73]	[3.56]	[0.68]	[0.22]	[0.19]	[5.65]	[20.39]	[23.85]	[37.76]	[52.63]

Panel B: $r^{(j)}_{t+2^j} - rf^{(j)}_{t+2^j} = \beta_j v^{(j)}_t + \epsilon_{t+2^j}$

	Time-scale j			
	1	2	3	4
$v^{(j)}_t$	3.31	-10.66	-2.43	3.85
	(1.01)	(-2.30)	(-0.59)	(3.98)
$R^2(\%)$	[8.70]	[22.73]	[4.68]	[83.69]

Panel C: $v^{(j)}_{t+2^j} = \rho_j v^{(j)}_t + \epsilon_{t+2^j}$

	Time-scale j			
	1	2	3	4
$v^{(j)}_t$	-0.18	0.09	0.02	0.19
	(-1.34)	(0.32)	(0.16)	(2.27)
$R^2(\%)$	[6.28]	[5.66]	[25.58]	[61.28]

Table 10: **Consumption Volatility. Panel A:** We run linear regressions (with an intercept) of h-period continuously compounded market returns on the CRSP value-weighted index in excess of a 1-year Treasury bill rate on h-period past consumption variance $v_{t-h,t}$. For each regression, the table reports OLS estimates of the regressors, Hansen and Hodrick corrected t-statistics in parentheses. **Panel B:** results of componentwise predictive regressions of the components of excess stock market returns on the components of consumption variance $v^2_{j,t}$. For each regression, the table reports OLS estimates of the regressors, OLS t-statistics in parentheses and R^2 statistics in square brackets. **Panel C:** estimation results of the multiscale autoregressive system. For each level of persistence $j \in \{1, \ldots, 4\}$, we run a regression of the consumption variance component $v^{(j)}_{t+2^j}$ on its own lagged component $v^{(j)}_t$. For each regression, the table reports OLS estimates of the regressors, t-statistics in parentheses and adjusted R^2 statistics in square brackets. The sample is annual and spans the period 1930-2011. For the translation of time-scales into appropriate range of time horizons refer to Table 1.

Panel A: $n = 1$ $\left(r^{(n)}_{t+h} - rf_t\right) = \alpha_h + \beta_h v^{(n)}_{t-h,t} + \epsilon_{t+h}$

	Horizon h (in months)						
	1	3	6	12	24	60	120
$v^{(1)}_{t-h,t}$	0.29	0.13	0.26	0.28	0.31	0.28	0.33
	(2.10)	(0.86)	(1.11)	(1.96)	(4.00)	(5.19)	(6.48)
$R^2(\%)$	[2.70]	[0.76]	[4.45]	[9.03]	[15.56]	[22.01]	[48.82]

Panel B: $n = 1$ $\left(r^{(n)}_{t+2^j,j} - rf_{t+2^j,j}\right) = \beta_j v^{(n)}_{t,j} + \epsilon_{t+2^j}$

	Time-scale j						
	1	2	3	4	5	6	7
$v^{(1)}_{t,j}$	0.08	-0.65	-0.26	-0.31	1.77	0.16	0.21
	(0.62)	(-3.63)	(-1.09)	(-0.96)	(2.35)	(0.45)	(2.43)
$R^2(\%)$	[0.13]	[8.27]	[1.63]	[2.55]	[25.61]	[2.85]	[74.70]

Panel C: $n = 1$ $v^{(n)}_{t+2^j,j} = \rho_j v^{(n)}_{t,j} + \epsilon_{t+2^j}$

	Time-scale j						
	1	2	3	4	5	6	7
$v^{(1)}_{t,j}$	-0.10	-0.32	0.08	-0.12	0.37	0.10	0.22
	(-1.74)	(-4.06)	(0.67)	(-0.72)	(1.58)	(0.27)	(0.32)
$R^2(\%)$	[1.02]	[10.16]	[0.62]	[1.46]	[13.45]	[1.01]	[4.91]

Table 11:

Panel A: $n = 2$ $\left(r^{(n)}_{t+h} - rf_t\right) = \alpha_h + \beta_h v^{(n)}_{t-h,t} + \epsilon_{t+h}$

	Horizon h (in months)						
	1	3	6	12	24	60	120
$v^{(2)}_{t-h,t}$	0.07	0.02	0.05	0.08	0.13	0.13	0.12
	(0.67)	(0.27)	(0.40)	(0.95)	(4.73)	(4.07)	(2.62)
$R^2(\%)$	[0.44]	[0.08]	[0.57]	[3.22]	[11.81]	[21.28]	[39.20]

Panel B: $n = 2$ $\left(r^{(n)}_{t+2^j,j} - rf_{t+2^j,j}\right) = \beta_j v^{(n)}_{t,j} + \epsilon_{t+2^j}$

	Time-scale j						
	1	2	3	4	5	6	7
$v^{(2)}_{t,j}$	0.17	-0.35	-0.09	0.06	0.47	0.09	0.1
	(1.90)	(-2.42)	(-0.59)	(0.32)	(1.48)	(0.61)	(2.12)
$R^2(\%)$	[1.21]	[3.86]	[0.48]	[0.28]	[12.09]	[5.09]	[69.25]

Panel C: $n = 2$ $v^{(n)}_{t+2^j,j} = \rho_j v^{(n)}_{t,j} + \epsilon_{t+2^j}$

	Time-scale j						
	1	2	3	4	5	6	7
$v^{(2)}_{t,j}$	0.15	-0.47	-0.17	0.18	0.55	0.07	0.05
	(2.11)	(0.89)	(-1.49)	(1.06)	(2.60)	(0.18)	(0.07)
$R^2(\%)$	[2.32]	[21.91]	[3.01]	[3.12]	[29.76]	[0.45]	[0.23]

Table 12:

Panel A: $n = 5$ $\left(r^{(n)}_{t+h} - rf_t\right) = \alpha_h + \beta_h v^{(n)}_{t-h,t} + \epsilon_{t+h}$

	Horizon h (in months)						
	1	3	6	12	24	60	120
$v^{(5)}_{t-h,t}$	0.02	-0.00	-0.02	0.01	0.06	0.08	0.05
	(1.12)	(-0.11)	(-0.85)	(0.44)	(3.49)	(2.80)	(1.60)
$R^2(\%)$	[0.26]	[0.01]	[0.96]	[0.43]	[10.58]	[23.52]	[13.53]

Panel B: $n = 5$ $\left(r^{(n)}_{t+2^j,j} - rf_{t+2^j,j}\right) = \beta_j v^{(n)}_{t,j} + \epsilon_{t+2^j}$

	Time-scale j						
	1	2	3	4	5	6	7
$v^{(5)}_{t,j}$	0.02	0.00	-0.03	0.08	0.10	0.01	0.08
	(1.21)	(0.05)	(-0.93)	(1.74)	(1.46)	(0.19)	(1.42)
$R^2(\%)$	[0.49]	[0.00]	[1.18]	[7.94]	[11.76]	[0.49]	[50.30]

Panel C: $n = 5$ $v^{(n)}_{t+2^j,j} = \rho_j v^{(n)}_{t,j} + \epsilon_{t+2^j}$

	Time-scale j						
	1	2	3	4	5	6	7
$v^{(5)}_{t,j}$	-0.13	-0.27	-0.16	0.35	0.66	-0.11	0.04
	(-2.18)	(-3.40)	(-1.39)	(2.09)	(1.58)	(-0.26)	(0.05)
$R^2(\%)$	[1.59]	[7.38]	[2.69]	[11.71]	[15.11]	[1.35]	[0.12]

Table 13:

Panel A: $n = 7$ $\left(r^{(n)}_{t+h} - rf_t\right) = \alpha_h + \beta_h v^{(n)}_{t-h,t} + \epsilon_{t+h}$

	Horizon h (in months)						
	1	3	6	12	24	60	120
$v^{(7)}_{t-h,t}$	0.02	-0.00	-0.00	0.02	0.05	0.08	0.05
	(0.86)	(-0.02)	(-0.13)	(0.62)	(2.67)	(3.11)	(1.93)
$R^2(\%)$	[0.26]	[0.00]	[0.04]	[0.92]	[8.62]	[35.09]	[23.54]

Panel B: $n = 7$ $\left(r^{(n)}_{t+2^j,j} - rf_{t+2^j,j}\right) = \beta_j v^{(n)}_{t,j} + \epsilon_{t+2^j}$

	Time-scale j						
	1	2	3	4	5	6	7
$v^{(7)}_{t,j}$	0.03	-0.04	-0.00	0.06	0.08	0.05	0.08
	(1.22)	(-1.01)	(-0.09)	(1.15)	(0.92)	(0.97)	(1.35)
$R^2(\%)$	[0.50]	[0.70]	[0.01]	[3.63]	[5.07]	[11.95]	[47.70]

Panel C: $n = 7$ $v^{(n)}_{t+2^j,j} = \rho_j v^{(n)}_{t,j} + \epsilon_{t+2^j}$

	Time-scale j						
	1	2	3	4	5	6	7
$v^{(7)}_{t,j}$	0.01	-0.11	0.02	0.08	0.13	0.15	-0.02
	(0.18)	(-1.33)	(0.15)	(0.46)	(0.55)	(1.39)	(-0.02)
$R^2(\%)$	[0.01]	[1.19]	[0.03]	[0.60]	[1.88]	[2.13]	[0.02]

Table 14:

Panel A: $n = 10 \quad \left(r^{(n)}_{t+h} - rf_t\right) = \alpha_h + \beta_h v^{(n)}_{t-h,t} + \epsilon_{t+h}$

	Horizon h (in months)						
	1	3	6	12	24	60	120
$v^{(10)}_{t-h,t}$	0.02	0.00	-0.01	0.01	0.05	0.06	0.04
	(1.43)	(0.24)	(-0.35)	(0.59)	(3.40)	(3.55)	(2.82)
$R^2(\%)$	[0.55]	[0.02]	[0.12]	[0.61]	[13.54]	[42.11]	[26.10]

Panel B: $n = 10 \quad \left(r^{(n)}_{t+2^j,j} - rf_{t+2^j,j}\right) = \beta_j v^{(n)}_{t,j} + \epsilon_{t+2^j}$

	Time-scale j						
	1	2	3	4	5	6	7
$v^{(10)}_{t,j}$	0.01	-0.01	-0.02	0.02	0.11	0.05	0.09
	(0.40)	(-0.42)	(-0.84)	(0.55)	(1.41)	(1.35)	(2.42)
$R^2(\%)$	[0.06]	[0.12]	[0.97]	[0.87]	[10.99]	[13.66]	[74.53]

Panel C: $n = 10 \quad v^{(n)}_{t+2^j,j} = \rho_j v^{(n)}_{t,j} + \epsilon_{t+2^j}$

	Time-scale j						
	1	2	3	4	5	6	7
$v^{(10)}_{t,j}$	0.13	0.22	-0.18	0.26	0.52	0.12	0.01
	(2.26)	(2.77)	(-1.52)	(1.54)	(3.79)	(0.25)	(1.49)
$R^2(\%)$	[1.71]	[5.03]	[3.21]	[6.67]	[50.69]	[1.45]	[1.01]

Table 15: **Bond Return Volatility. Panel A:** We run linear regressions (with an intercept) of h-period continuously compounded bond returns with maturity n in excess of the short rate on h-period past bond variance. The short rate is the log yield on the 30-day Treasury Bill from CRSP. For each regression, the table reports OLS estimates of the regressors, Hansen and Hodrick corrected t-statistics in parentheses. **Panel B:** results of componentwise predictive regressions of the components of bond returns with maturity n in excess of the short rate on the components of bond variance. **Panel C:** estimation results of the multiscale autoregressive system. For each level of persistence $j \in \{1, \ldots, 7\}$, we run a regression of the bond variance $v^{(j)}_{t+2^j}$ on its own lagged component $v^{(j)}_t$. For each regression, the table reports OLS estimates of the regressors, OLS t-statistics in parentheses and R^2 statistics in square brackets. The sample is monthly and spans the period 1962M7-2011M12. For the translation of time-scales into appropriate range of time horizons refer to Table 2.

Panel A: $r_{t+h} = \alpha_h + \beta_h \pi_{t-h,t} + \epsilon_{t+h}$

	Horizon h (in years)									
	1	2	3	4	5	6	7	8	9	10
π_t	0.44	0.02	0.11	0.66	0.98	1.22	1.13	1.06	1.14	1.10
	(0.58)	(0.05)	(0.22)	(3.39)	(5.40)	(5.81)	(4.76)	(3.28)	(3.56)	(2.63)
$R^2(\%)$	[0.82]	[0.00]	[0.15]	[6.72]	[15.50]	[27.41]	[28.29]	[27.23]	[30.98]	[27.93]

Panel B: $r^{(j)}_{t+2^j} = \beta_j \pi^{(j)}_t + \epsilon_{t+2^j}$

	Time-scale j			
	1	2	3	4
$\pi^{(j)}_t$	0.72	0.03	2.46	-0.10
	(0.54)	(0.10)	(2.12)	(-0.44)
$R^2(\%)$	[1.28]	[10.84]	[35.86]	[46.04]

Panel C: $\pi^{(j)}_{t+2^j} = \rho_j \pi^{(j)}_t + \epsilon_{t+2^j}$

	Time-scale j			
	1	2	3	4
$\pi^{(j)}_t$	-0.17	0.11	0.10	-0.35
	(-1.11)	(0.51)	(0.35)	(-1.61)
$R^2(\%)$	[5.14]	[1.82]	[2.09]	[37.51]

Table 16: **Stock market return. Panel A:** We run linear regressions (with an intercept) of h-period continuously compounded nominal stock market returns $r_{t,t+h}$ on h-period past realized inflation. We consider values of h equal to $1-10$ years. 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. **Panel B:** results of componentwise predictive regressions of the components of nominal excess stock market returns on the components of inflation $\pi_{j,t}$. For each regression, the table reports OLS estimates of the regressors, OLS t-statistics in parentheses and R^2 statistics in square brackets. **Panel C:** estimation results of the multiscale autoregressive system. For each level of persistence $j \in \{1, \ldots, 4\}$, we run a regression of the inflation rate $\pi^{(j)}_{t+2^j}$ on its own lagged component $\pi^{(j)}_t$. For each regression, the table reports OLS coefficient estimates of the regressors, t-statistics in parentheses and adjusted R^2 statistics in square brackets. The sample is annual and spans the period 1930-2011. For the translation of time-scales into appropriate range of time horizons refer to Table D Panel A.

Panel A: $rf_{t,t+h} = \alpha_h + \beta_h \pi_{t-h,t} + \epsilon_{t+h}$

	Horizon h (in years)									
	1	2	3	4	5	6	7	8	9	10
π_t	0.37	0.43	0.50	0.61	0.69	0.73	0.72	0.70	0.65	0.60
	(2.92)	(2.49)	(2.56)	(2.82)	(2.99)	(3.04)	(3.07)	(3.14)	(3.20)	(3.10)
$R^2(\%)$	[21.00]	[24.05]	[26.36]	[32.45]	[37.66]	[39.60]	[38.51]	[34.95]	[30.53]	[25.06]

Panel B: $rf^{(j)}_{t+2^j} = \beta_j \pi^{(j)}_t + \epsilon_{t+2^j}$

	Time-scale j			
	1	2	3	4
$\pi^{(j)}_t$	-0.01	-0.01	0.22	-0.02
	(-0.12)	(-0.04)	(1.90)	(-0.06)
$R^2(\%)$	[0.06]	[5.61]	[13.24]	[0.30]

Panel C: $\pi^{(j)}_{t+2^j} = \rho_j \pi^{(j)}_t + \epsilon_{t+2^j}$

	Time-scale j			
	1	2	3	4
$\pi^{(j)}_t$	-0.17	0.11	0.10	-0.35
	(-1.11)	(0.51)	(0.35)	(-1.61)
$R^2(\%)$	[5.14]	[1.82]	[2.09]	[37.51]

Table 17: **Risk-free rate. Panel A:** We run linear regressions (with an intercept) of h-period continuously compounded nominal risk-free rate $r_{t,t+h}$ on h-period past realized inflation. We consider values of h equal to $1-10$ years. 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. **Panel B:** results of componentwise predictive regressions of the components of nominal risk-free rate on the components of inflation $\pi_{j,t}$. For each regression, the table reports OLS estimates of the regressors, OLS t-statistics in parentheses and R^2 statistics in square brackets. **Panel C:** estimation results of the multiscale autoregressive system. For each level of persistence $j \in \{1, \ldots, 4\}$, we run a regression of the inflation rate $\pi^{(j)}_{t+2^j}$ on its own lagged component $\pi^{(j)}_t$. For each regression, the table reports OLS coefficient estimates of the regressors, t-statistics in parentheses and adjusted R^2 statistics in square brackets. The sample is annual and spans the period 1930-2011. For the translation of time-scales into appropriate range of time horizons refer to Table D Panel A.