



Timescale betas and the cross section of equity returns: Framework, application, and implications for interpreting the Fama–French factors



Byoung Uk Kang^{a,*}, Francis In^b, Tong Suk Kim^c

^a School of Accounting and Finance, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

^b Department of Accounting, Finance and Economics, Griffith University, QLD 4222, Australia

^c KAIST Business School, 87 Hoegiro, Dongdaemoon-gu, Seoul 130-722, Korea

ARTICLE INFO

JEL classification:

C32

G12

Keywords:

Asset pricing

Timescale betas

Cross section of stock returns

Fama–French factors

Wavelets

ABSTRACT

We show that standard beta pricing models quantify an asset's systematic risk as a weighted combination of a number of different timescale betas. Given this, we develop a wavelet-based framework that examines the cross-sectional pricing implications of isolating these timescale betas. An empirical application to the Fama–French model reveals that the model's well-known empirical success is largely due to the beta components associated with a timescale just short of a business cycle (i.e., wavelet scale 3). This implies that any viable explanation for the success of the Fama–French model that has been applied to the Fama–French factors should apply particularly to the scale 3 components of the factors. We find that a risk-based explanation conforms closely to this implication.

1. Introduction

This paper presents a framework that characterizes an asset's systematic risk in terms of the asset's exposures to a number of different timescale (e.g., short-, intermediate-, long-run) fluctuations in a model's factors. We call measures of such exposures timescale betas. Using two fundamental results in wavelet theory, we show that an asset's standard betas, or factor loadings, can be written as a linear combination of the asset's timescale betas. Empirically, we isolate these betas and explore their relative importance for the cross section of returns. Our exploration is motivated in spirit by the long-run consumption-risk literature, which suggests the possibility that only part of the information in the standard betas is relevant for asset pricing and the relevant part is concentrated in certain timescale betas (see, e.g., Bansal and Yaron, 2004; Bansal et al., 2005; Parker and Julliard, 2005).¹ We demonstrate our framework using the three-factor model of Fama and French (1993; hereafter, FF), among other commonly used models. Our goal and contribution are first to formalize and extend the application of wavelets to asset pricing and also to augment our understanding of what drives the empirical success of the FF three-factor model.

* Corresponding author.

E-mail addresses: byoung.kang@polyu.edu.hk (B.U. Kang), f.in@griffith.edu.au (F. In), tskim@business.kaist.ac.kr (T.S. Kim).

¹ Bansal and Yaron (2004) and Bansal et al. (2005) document that focusing on a long-run component of consumption growth helps account for a range of asset pricing puzzles. Similarly, Parker and Julliard (2005) report that exposure to growth rates in cumulative consumption, rather than contemporaneous consumption, is important for the cross section of equity returns. However, the importance of a long-run component is not universal. When decomposing equity market volatility, for example, Adrian and Rosenberg (2008) find that its short-run component is an important cross-sectional pricing factor.

To conceptualize our use of wavelets, consider a simple case of three timescales (short (*S*), intermediate (*I*), and long (*L*)) and a single-factor model, where an asset's beta is defined as $\beta_i \equiv \text{cov}(R_i, f)/\text{var}(f)$.² With this simplification, the paper's two essential uses of wavelets can be summarized as follows: First, we use wavelets to decompose the factor *f* into its short-, intermediate-, and long-run components, whose sum is equal to the original factor. By replacing the factor *f* in the beta formula with the sum of its three timescale components, it follows that $\beta_i \equiv \omega_S \beta_i^S + \omega_I \beta_i^I + \omega_L \beta_i^L$, where the resulting beta components, or timescale betas, capture the asset's exposure to different timescale fluctuations in the factor. Timescale betas computed in this way have a parallel with the long-run consumption risk literature. For example, [Bansal et al. \(2005\)](#) and [Parker and Julliard \(2005\)](#) estimate an asset's consumption beta as the covariance of the asset's contemporaneous return or cash flow growth with smoothed consumption growth—with the smoothing done by taking moving averages or moving sums of consumption growth. As will be discussed below in [Section 2](#), these commonly used alternatives, however, can only approximate the long-run component but not the short- or intermediate-run component of the factor. More importantly, the long-run component approximated by these approaches contains a great deal of other, unintended timescale components.

Second, we also use wavelets to similarly decompose variance and covariance into the sum of their short-, intermediate-, and long-run components, such that the resulting (co)variance components capture the contribution to the original (co)variance due exclusively to the corresponding timescales. In this use of wavelets, we replace the variance and covariance terms in the beta formula with the sum of their three timescale components, and obtain a similar threefold decomposition of the standard beta as above, where each beta component now represents the effect of conditioning the standard beta on a different timescale. Conceptually, beta components obtained in this way resemble τ -month betas considered, for example, in [Brennan and Zhang \(2013\)](#) and [Kamara et al. \(2016\)](#), who use overlapping monthly observations to estimate τ -month betas as the covariance of τ -month return with τ -month factors, for each $\tau \in \{3, 6, 9, \dots, 60\}$.³ While we focus more on the first use of wavelets in the empirical section of the paper, it follows from either of the uses that the standard beta can be written as a linear combination of three timescale-specific risk measures. Our baseline empirical analysis is then to separate these individual beta components and use them alternately as a regressor in a series of cross-sectional regressions of average returns. Analysis based on the first use of wavelets tells us which timescale component of the factor is important for the cross section of average returns, whereas the second use tells us about the relative importance of different timescales at which to measure the beta, given the cross section of average returns.⁴

Note that the beta components resulting from the second use of wavelets are essentially the *wavelet betas* of [Gençay et al. \(2003, 2005\)](#). With the aim of reevaluating the capital asset pricing model (CAPM) of [Sharpe \(1964\)](#) and [Lintner \(1965\)](#) at different horizons, these authors use daily returns and run a cross-sectional regression, $E(R_i) = \lambda_0 + \lambda_j \beta_i^j$, for each timescale $j \in \{S, I, L\}$ (note that the superscript *j* is on the right-hand side only). Our paper adds to theirs a clear mathematical relationship between the wavelet betas (β_i^j) and the standard beta (β_i), and thus provides a sharper interpretation of what they do: Rather than testing the CAPM at different horizons, their work should be viewed as an attempt to understand the beta components that cause the CAPM to perform relatively well or poorly in matching the cross section of *daily* returns. In our view, the validity of the CAPM at different horizons can be better evaluated by varying the return interval for both sides of the CAPM relationship (as in, e.g., [Handa et al., 1993](#); [Brennan and Zhang, 2013](#)) or by breaking down pricing errors by frequency (as in, e.g., [Cogley, 2001](#)). We discuss and correct this and other interpretations by [Gençay et al. \(2003, 2005\)](#) in [Section 2.4](#).

We demonstrate in detail the empirical implementation of our framework by using the FF three-factor model as an illustrative example. The three factors are the market factor (R_m), the book-to-market factor (*HML*), and the size factor (*SMB*). We use monthly data over the period July 1963 to June 2008, which allows us to entertain up to seven different timescales. The seven timescales are such that the shortest three (scales 1, 2, and 3) correspond to the short run, the next three (scales 4, 5, and 6) to business cycles, and the last (scale 7) to the long run, according to definitions in the business-cycle literature (e.g., [Burns and Mitchell, 1946](#); [Baxter and King, 1999](#)). While [FF \(1993, 1995, 1996\)](#) show that the betas, or factor loadings, on the model's three factors (i.e., a three-factor analog of $\beta_i \equiv \omega_1 \beta_i^1 + \dots + \omega_7 \beta_i^7$) explain many patterns in average stock returns, we show that much of this explanatory power comes from short-run or, specifically, scale 3 components of the betas (i.e., a three-factor analog of β_i^3). Using a standard set of 25 portfolios sorted by size and book-to-market ratio as test assets, we find that the scale 3 betas capture the cross-sectional variation in average returns better than any other timescale betas or overall betas (i.e., the FF model). Additional analyses show that neither portfolio characteristics nor the remaining components of the betas add significant explanatory power beyond that of the scale 3 betas. The relative superiority of the scale 3 betas persists in explaining out-of-sample returns and is not overly affected by the choice of wavelet filter, changes in sample period, the inclusion of other portfolios with little factor structure ([Lewellen et al., 2010](#)), or the handling of boundary conditions.

The scale 3 components of the monthly FF factors capture the fluctuations in the factors that complete a cycle in 8 to 16 months (i.e., the fluctuations with periodicity of 8–16 months). To put this into perspective, the quarterly (annual) FF factors capture the fluctuations in the factors with periodicity greater than 6 (24) months.⁵ Roughly and conceptually, therefore, the scale 3 components can be understood as capturing the fluctuations in quarterly factors that are not contained in annual factors (i.e., the fluctuations

² Economists have long acknowledged the need for more than only two or three timescales to analyze the dynamics of economic time series ([Ramsey and Lampart, 1998a, b](#)). [Section 2](#) presents a general framework with a finer division of timescale and multiple factors.

³ Despite this conceptual similarity, τ -month betas can only approximate long-run betas but not the other timescale betas. As will be clearer below in [Section 2](#), betas estimated using overlapping 48-month return series are essentially long-run betas, but we cannot isolate short- or intermediate-run betas in this manner: For example, betas estimated using 1-month returns are not short-run betas, but a weighted average of short-, intermediate-, and long-run betas. An advantage of using wavelets is that we can cleanly separate out not only long-run betas but also shorter timescale betas so that we can examine their pricing implications in isolation. We discuss this and other advantages of wavelets in [Section 2](#).

⁴ More subtly, our analysis based on the second use of wavelets does not change the pricing kernel of the original factor model and only subset frequencies at which to measure the betas. In contrast, the first use does change the pricing kernel by replacing the original factor with its timescale component.

⁵ This is because the shortest detectable cycle in a time series is one that lasts for two periods ([Baxter and King, 1999](#)).

with periodicity of 6–24 months). Simply using quarterly factors will dilute the information content of the scale 3 component by aggregating it with other, longer timescale components, while simply using annual factors will lose the information content of the scale 3 component completely.⁶

What drives the relative superiority of the scale 3 betas? Consistent with FF (1993), Hahn and Lee (2006) and Petkova, (2006) argue that the FF model works well because the FF factors proxy for risk factors within Merton's (1973) intertemporal CAPM (ICAPM). Specifically, they show that *HML* proxies for innovations in term spread and *SMB* proxies for innovations in default spread. We find that this risk-based explanation for the empirical success of the FF model is also in line with the relative superiority of FFd_3 , given the fact that HML^{d_3} (SMB^{d_3}) proxies for innovations in term (default) spread better than any other timescale component of *HML* (*SMB*) does. Specifically, by decomposing Hahn and Lee's (2006) regressions of *HML* (*SMB*) on innovations in term (default) spread, we find that a substantial fraction of the estimated regression coefficient of *HML* (*SMB*) on the term (default) factor is due to the effect of the scale 3 component of *HML* (*SMB*). Moreover, focusing on the scale 3 components of the FF factors strengthens the FF factors' correlations with the associated state variable innovations, while correlations between the FF factors and the state variable innovations that are otherwise significant become insignificant when we remove the scale 3 components from the FF factors. Overall, these results provide a risk-based explanation for what we find above: Focusing on the scale 3 components of the FF factors sharpens the factors' information content about systematic risk, and hence the model's explanatory power for the cross section of returns.

Recall that FF (1989) observe that the term spread closely tracks business cycles, while the default spread tracks long-run business episodes that span several measured business cycles. Because the scale 3 is shorter than a business cycle, our finding that the scale 3 components of the FF factors covary significantly with the term and default factors may appear at odds with FF's (1989) observation. Note, however, that we do not relate the scale 3 components of the FF factors to the state variable candidates, but rather to their innovations (estimated by changes or residuals). As shown in Baxter, 1994 (and also in Online Appendix Figure A1 of this paper), taking simple changes or residuals from an autoregressive model acts like a filter that removes nearly all of the long-run and many of the business-cycle components, while magnifying the short-run fluctuations in the data. As the ICAPM suggests, priced factors are not state variables per se, but instead are the short-run fluctuations in the state variables captured by their innovations.⁷ What we document here is the strong resonance between such fluctuations in the term and default spreads and the scale 3 components of the FF factors.

2. Beta decompositions via wavelets

This section generalizes and formalizes the beta decompositions described earlier, and develops a framework for examining the pricing implications of isolating different timescale betas. To set the stage, consider a linear K -factor pricing model, where the betas, or factor loadings, of asset i are defined as slope coefficients from the time-series regression

$$R_{i,t} = \alpha_i + \beta_i' \mathbf{f}_t + \varepsilon_{i,t}, \quad E(\varepsilon_{i,t} | \mathbf{f}_t) = 0, \quad \forall i, \quad (1)$$

where $R_{i,t}$ is the excess return on asset i at time t , β_i is a $K \times 1$ vector of betas for asset i , and \mathbf{f}_t is a $K \times 1$ vector of factors at time t . The beta vector, β_i , can be equivalently defined as

$$\beta_i \equiv \text{Var}(\mathbf{f}_t)^{-1} \text{Cov}(\mathbf{f}_t, R_{i,t}), \quad (2)$$

where $\text{Var}(\mathbf{f}_t)$ is the variance–covariance matrix of \mathbf{f}_t , and $\text{Cov}(\mathbf{f}_t, R_{i,t})$ is a $K \times 1$ vector containing the covariances of the factors with $R_{i,t}$. Section 2.1 discusses beta decomposition based on factor decomposition, or formally the *multiresolution analysis* (MRA) of factors; Section 2.2 discusses beta decomposition based on covariance (and variance) decomposition.

One technical note here: Recall that our empirical work involves measuring the covariation of the original time series (e.g., asset returns, state variable innovations) with the decomposed series of the FF factors; hence, it is important that the timing of events in the original time series be perfectly aligned with that in the decomposed series. For this reason, among others, we use the maximal overlap discrete wavelet transform (MODWT), rather than the discrete wavelet transform (DWT), in conducting our empirical analysis (other papers that use the MODWT include Kim and In (2005) and In and Kim (2006)).⁸ Accordingly, the following exposition of our framework is based on the MODWT.

2.1. Beta decomposition via MRA

Wavelets allow us to decompose a time series into the sum of several component series, each capturing variations at a different

⁶ Similarly, using two- or threefold decompositions such as the Hodrick and Prescott (1997) and Baxter and King (1999) filters will dilute the information content of the scale 3 component by aggregating it with its adjacent components into a single noise component. For further details on wavelets and other filtering methods, see Gençay et al. (2002).

⁷ The ICAPM predicts that only innovations earn a risk premium in asset returns (see, e.g., Campbell 1996). Indeed, Petkova (2006) reports that when the levels of the state variables, rather than their innovations, are used as factors, the resulting model does not perform as well (see footnote 17 of her paper).

⁸ There are two main classes of wavelets: The continuous wavelet transform (CWT) and the DWT. While the CWT is designed to work with signals defined over the entire real axis, the DWT deals with signals defined over a discrete set of integers, say, $t = 1, \dots, T$, and hence is directly applicable to time series. The MODWT, a popular variant of the DWT, eliminates certain alignment artifacts in the DWT, while foregoing orthogonality. Unlike the DWT, the MODWT is also associated with zero phase filters, hence making it easy to line up features in an MRA with the original time series (see Section 5.0 in Percival and Walden, 2000] for these and other properties that distinguish the MODWT from the DWT).

timescale. Conceptually, this additive decomposition, or MRA, begins by decomposing a time series x_t into two parts: a moving average (of a particular kind) of the time series and the deviation of the time series from its moving average (i.e., $x_t = s_{1,t} + d_{1,t}$). In wavelet jargon, the outcomes of this first decomposition are called level 1 *smooth* and *detail*, respectively. The MRA can further proceed by subjecting the level 1 smooth to a moving average of the same kind and decomposing it into two parts: the moving average of the level 1 smooth and the deviation of the level 1 smooth from its moving average (i.e., $s_{1,t} = s_{2,t} + d_{2,t}$ or $x_t = s_{2,t} + \sum_{j=1}^2 d_{j,t}$). The outcomes of this second decomposition are called level 2 smooth and detail, respectively. By repeating this procedure J times, a time series can be decomposed into a smooth and J details, each associated with a different timescale (i.e., $x_t = s_{J,t} + \sum_{j=1}^J d_{j,t}$). At higher levels of J , the smooth in effect averages the original series over a longer timescale and becomes smoother in appearance. Because level j detail captures a smooth's fluctuations (relative to a baseline) at level $j - 1$ (i.e., $s_{j-1,t} = s_{j,t} + d_{j,t}$), detail is associated with the same timescale and is about as smooth (or rough) in appearance as the one-level-down smooth is.⁹ Gençay et al. (2003, p. 113) succinctly summarize this recursive procedure by stating: “When decomposing [a time series] using the wavelet transform, we are actually separating layers of information associated with different timescales that increase with the level of the transform.”

Applying MRA (of level J) to each element, we can write the vector of factors \mathbf{f}_t as

$$\mathbf{f}_t = \sum_{j=1}^J \mathbf{f}_t^{dj} + \mathbf{f}_t^{sj}, \tag{3}$$

where \mathbf{f}_t^{dj} and \mathbf{f}_t^{sj} are vectors containing level j details and level J smooths for the corresponding elements in \mathbf{f}_t , respectively. For a given decomposition level J , the elements in \mathbf{f}_t^{dj} and \mathbf{f}_t^{sj} represent the j th shortest and the longest timescale components of the corresponding elements in \mathbf{f}_t , respectively. Theoretically, level J can be any arbitrary integer; in practice, it is upper bounded depending on the length of the data available.¹⁰ In an example in Section 3, we consider not only sample size but also economic interpretations to attach to the $J + 1$ timescales when picking J .

It is useful to note from the outset that the notion of timescale can be related to certain bands of frequencies (Percival and Walden, 2000, p. 96–97). Specifically, the equivalent filter that yields the level j smooth is approximately a low-pass filter with a pass-band given by $[0, 1/2^{j+1}]$, and the equivalent filter that yields the level j detail is approximately a band-pass filter with a pass-band given by $[1/2^{j+1}, 1/2^j]$; Figs. 1 and 2 illustrate this point by plotting the gain functions for the equivalent filters associated with smooths and details at different levels (see solid curves).¹¹ Therefore, the longest timescale for a given J relates to frequencies between zero and $1/2^{J+1}$ cycles per period; the j th shortest timescale relates to frequencies between $1/2^{j+1}$ and $1/2^j$ cycles per period. In terms of periodicity (i.e., the inverse of frequency), the longest timescale relates to periodicities greater than 2^{j+1} periods; and the j th shortest timescale relates to periodicities between 2^j and 2^{j+1} periods.

For comparison, Fig. 3 plots the gain functions for simple moving averages; plots for moving sums look the same except for the scale shown on the y-axis. In the long-run consumption risk literature, Bansal et al. (2005) and Parker and Julliard (2005) use essentially simple moving averages and moving sums, respectively. As shown in Fig. 3, these filters reduce the high-frequency components of the data, but are far from completely removing them. See, for example, the amount of remaining high-frequency components when the window size is set equal to 8, as in Bansal et al. (2005), or 11, as in Parker and Julliard (2005). Smooths, in contrast, cleanly remove these unintended components.

Given the factor decomposition in (3), the standard betas in (2) can now be represented as follows:

$$\beta_t \equiv \text{Var}(\mathbf{f}_t)^{-1} \left[\sum_{j=1}^J \text{Cov}(\mathbf{f}_t^{dj}, R_{i,t}) + \text{Cov}(\mathbf{f}_t^{sj}, R_{i,t}) \right] = \sum_{j=1}^J \omega_{dj} \beta_{i,dj} + \omega_{sj} \beta_{i,sj}, \tag{4}$$

where $\beta_{i,dj} \equiv \text{Var}(\mathbf{f}_t^{dj})^{-1} \text{Cov}(\mathbf{f}_t^{dj}, R_{i,t})$ and $\beta_{i,sj} \equiv \text{Var}(\mathbf{f}_t^{sj})^{-1} \text{Cov}(\mathbf{f}_t^{sj}, R_{i,t})$ are $K \times 1$ vectors containing the factor loadings on \mathbf{f}_t^{dj} and \mathbf{f}_t^{sj} , respectively; and $\omega_{dj} \equiv \text{Var}(\mathbf{f}_t)^{-1} \text{Var}(\mathbf{f}_t^{dj})$ and $\omega_{sj} \equiv \text{Var}(\mathbf{f}_t)^{-1} \text{Var}(\mathbf{f}_t^{sj})$ denote the weighting matrices on $\beta_{i,dj}$ and $\beta_{i,sj}$, respectively. Analogous to (1), these beta vectors can be equivalently defined as

$$R_{i,t} = \alpha_i + \beta_{i,dj} \mathbf{f}_t^{dj} + \varepsilon_{i,t} \quad \text{and} \quad R_{i,t} = \alpha_i + \beta_{i,sj} \mathbf{f}_t^{sj} + \varepsilon_{i,t}. \tag{5}$$

Clearly from (4), the betas, or factor loadings, of an asset can be broken into several parts attributable to different timescale fluctuations in the factors, and the overall betas are simply a weighted linear combination of these timescale betas.

As an aside, note that since the MRA in (3) holds for each individual factor (and

⁹ More technically, smooths (details) are obtained by filtering scaling (wavelet) coefficients, which in turn are obtained by filtering the original time series with a scaling (wavelet) filter. The moving average described here refers to a composite filter that yields smooths directly from the original series. Such a composite filter yielding smooths (details) is characterized by zero phase and its gain function is equal to the squared gain function for the scaling (wavelet) filter (see, e.g., Section 5.3 in Percival and Walden, 2000 for details).

¹⁰ A rule of thumb requires the condition $J < \log_2(\frac{T}{L-1} + 1)$, where T is the number of time-series observations, and L is the width of the wavelet filter. Although there are other, less restrictive alternatives, such as $J < \log_2(T)$ or even $J < \log_2(1.5T)$, we are concerned that decomposed series at too high a level would contain too many observations deleteriously affected by the assumption of circularity (Percival and Walden, 2000, p. 197–199). In this light, we use the first, more conservative rule in selecting the upper limit, yet it still permits up to seven different timescales for our empirical exploration in Section 3.

¹¹ The gain (squared gain) at a given frequency indicates the extent to which a filter raises or lowers the standard deviation (variance) of the filtered series relative to that of the original series. Gain and squared gain functions are thus often used to study the consequences of filtering (see, e.g., Baxter and King, 1999).

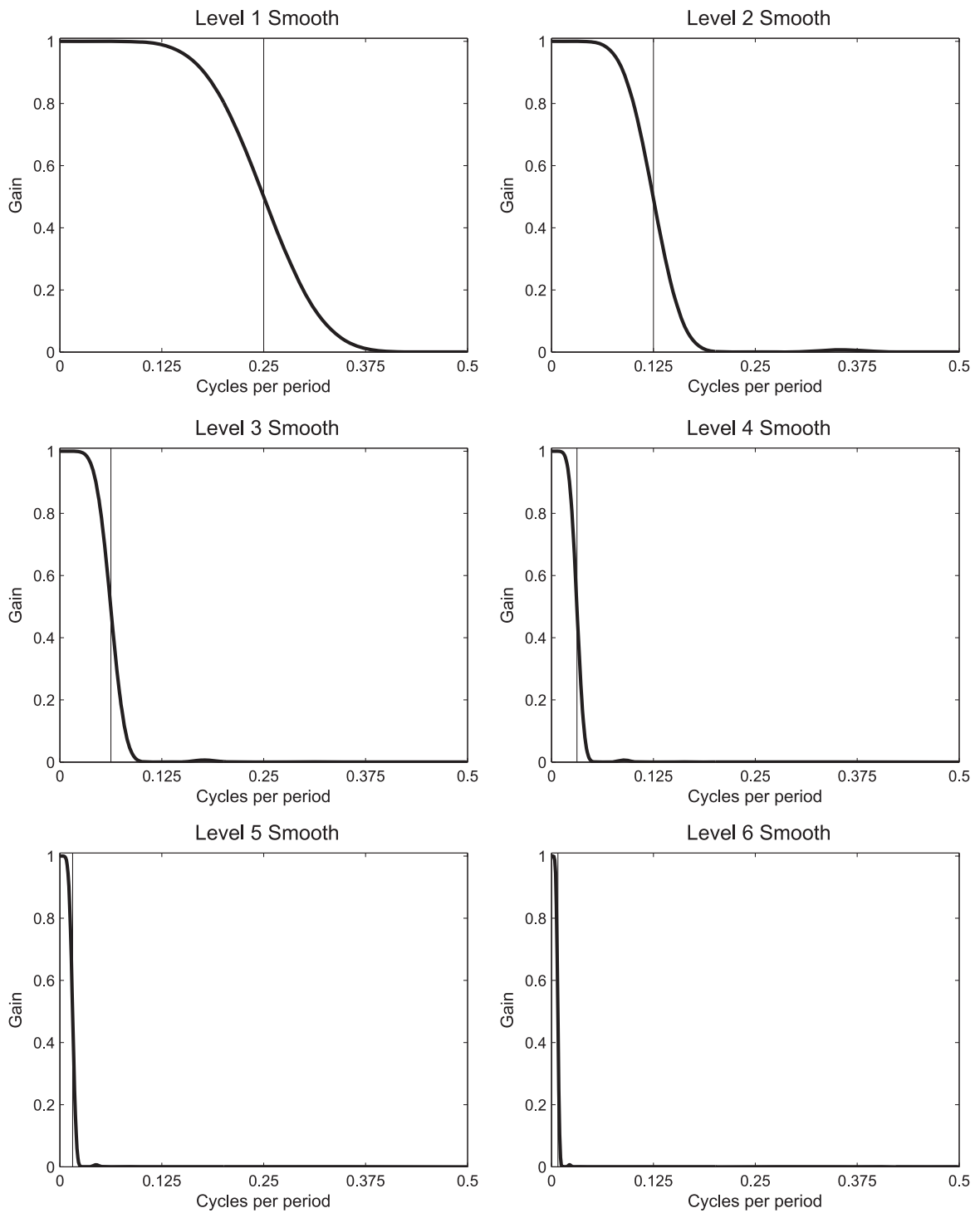


Fig. 1. Gain functions for equivalent filters associated with smooths. This figure shows the gain functions for the ideal low-pass filters (thin vertical lines) and their approximations by the equivalent filters associated with smooths (solid curves). The latter are produced using the LA(8) filter.

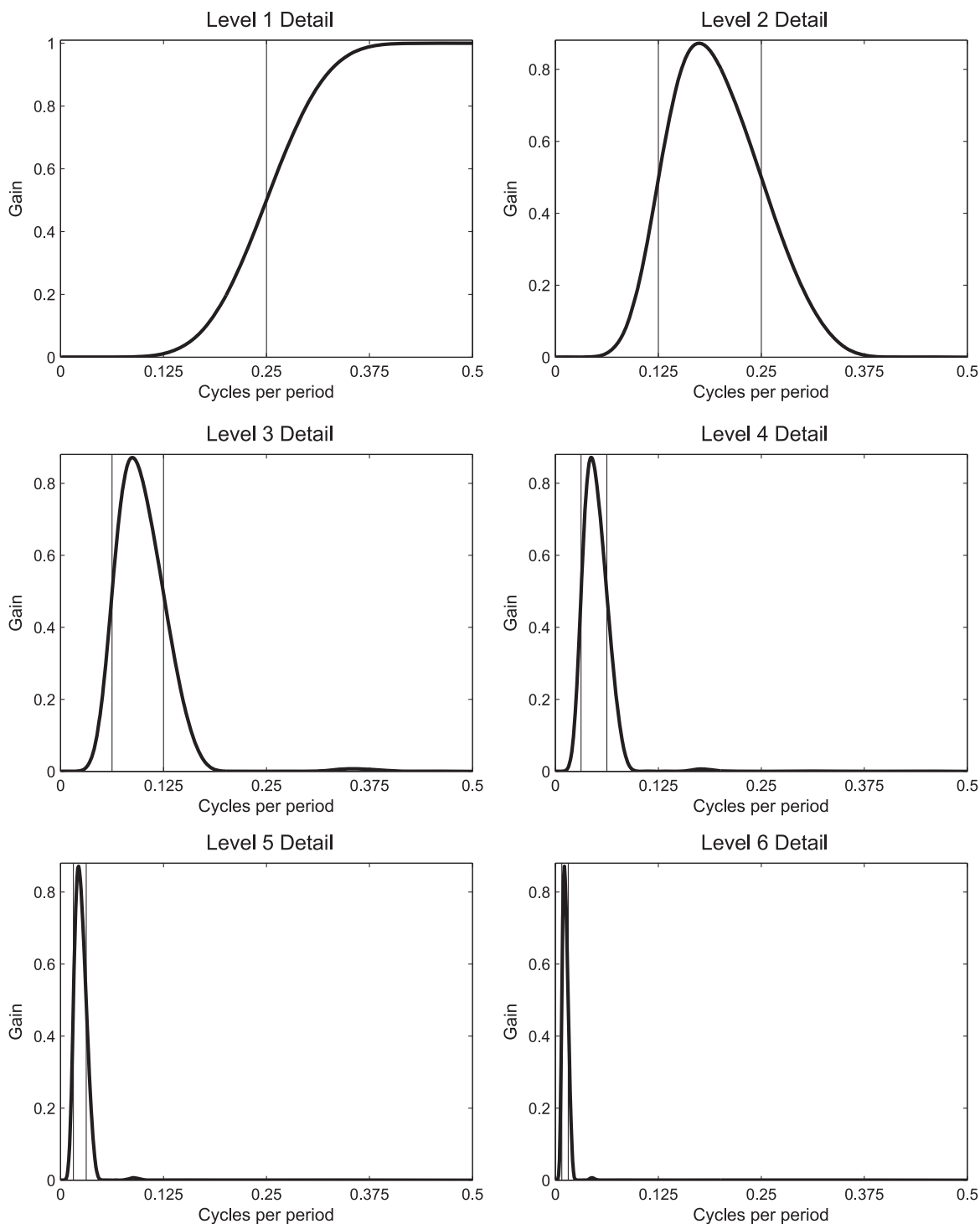


Fig. 2. Gain functions for equivalent filters associated with details. This figure shows the gain functions for the ideal band-pass filters (thin vertical lines) and their approximations by the equivalent filters associated with details (solid curves). The latter are produced using the LA(8) filter.

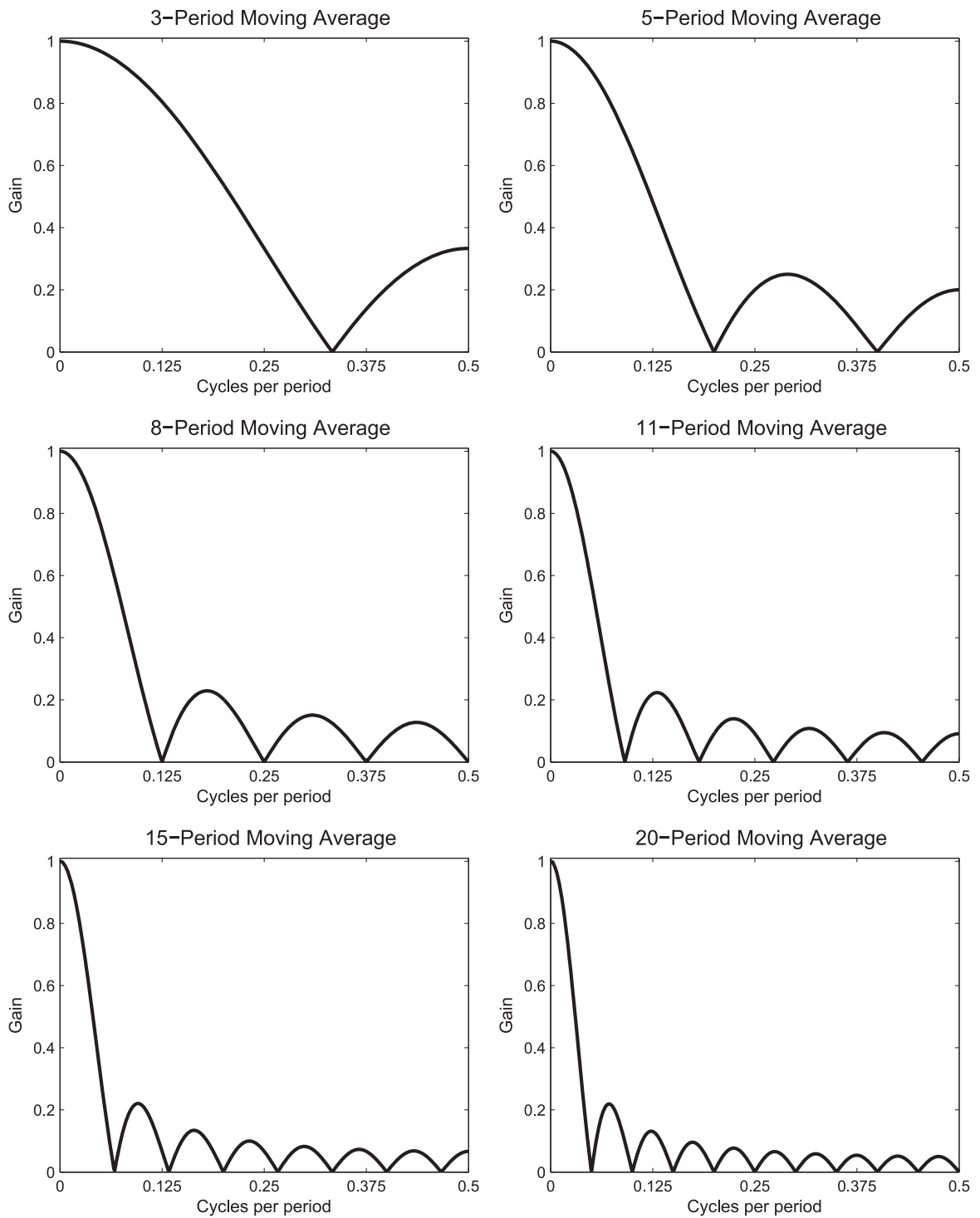


Fig. 3. Gain functions for simple moving averages. This figure shows the gain functions for simple moving averages with different window sizes.

$x_t = d_{1,t} + d_{2,t} + \dots + d_{j,t} + s_{j,t} = s_{j,t} + d_{j,t} + \dots + d_{2,t} + d_{1,t} = \dots$), it is mathematically possible to rewrite (3), and hence (4) and (5), in a total of $[(J + 1)!]^{K-1}$ different ways.¹² However, none of these $[(J + 1)!]^{K-1}$ decompositions, except for the one presented here, account for the effect of other factors when estimating timescale betas in time-series regressions (e.g., (5)), because factors at different timescales are nearly orthogonal (or perfectly so in the case of the DWT; see footnote 8) even between factors. Therefore, without including all factors at the same timescale, the estimated beta on a given factor at a given timescale would be confounded by the betas on other, omitted factors at that timescale (i.e., cross-factor confounding would accompany timescale decomposition).

2.2. Beta decomposition via ANOVA/ANCOVA

In addition to MRA, wavelets also permit a similar additive decomposition for (co)variance, i.e., into the sum of J wavelet (co)variances and one scaling (co)variance, each representing the contribution to the original (co)variance due to a different timescale. As a shorthand, we refer to these decompositions as analysis of variance (ANOVA) and analysis of covariance (ANCOVA), respectively. Applying the former to diagonal elements and the latter to off-diagonal elements, we can write $\text{Var}(\mathbf{f}_t)$ as

$$\text{Var}(\mathbf{f}_t) = \sum_{j=1}^J \text{Var}(\mathbf{f}_t^{wj}) + \text{Var}(\mathbf{f}_t^{vj}), \tag{6}$$

where \mathbf{f}_t^{wj} is a vector whose variance–covariance matrix $\text{Var}(\mathbf{f}_t^{wj})$ contains level j wavelet variances on its diagonal and level j wavelet covariances elsewhere, and \mathbf{f}_t^{vj} is a vector whose variance–covariance matrix $\text{Var}(\mathbf{f}_t^{vj})$ contains level J scaling variances on its diagonal and level J scaling covariances elsewhere. In wavelet jargon, the elements in \mathbf{f}_t^{wj} and \mathbf{f}_t^{vj} are called, respectively, level j wavelet coefficients and level J scaling coefficients for the corresponding elements in \mathbf{f}_t .¹³ Applying ANCOVA for each element, we can also write $\text{Cov}(\mathbf{f}_t, R_{i,t})$ as

$$\text{Cov}(\mathbf{f}_t, R_{i,t}) = \sum_{j=1}^J \text{Cov}(\mathbf{f}_t^{wj}, R_{i,t}^{wj}) + \text{Cov}(\mathbf{f}_t^{vj}, R_{i,t}^{vj}), \tag{7}$$

where $R_{i,t}^{wj}$ and $R_{i,t}^{vj}$ are level j wavelet coefficient and level J scaling coefficient for $R_{i,t}$, respectively. For a given decomposition level J , $\text{Var}(\mathbf{f}_t^{wj})$ and $\text{Cov}(\mathbf{f}_t^{wj}, R_{i,t}^{wj})$ represent the contribution to $\text{Var}(\mathbf{f}_t)$ and $\text{Cov}(\mathbf{f}_t, R_{i,t})$, respectively, due to the j th shortest timescale; $\text{Var}(\mathbf{f}_t^{vj})$ and $\text{Cov}(\mathbf{f}_t^{vj}, R_{i,t}^{vj})$ represent the contribution to $\text{Var}(\mathbf{f}_t)$ and $\text{Cov}(\mathbf{f}_t, R_{i,t})$, respectively, due to the longest timescale.

Although wavelet (scaling) coefficient and detail (smooth) at a given level are associated with the same timescale, and thus the same band of frequencies,¹⁴ the ways in which they are put to use are quite different. Specifically, while wavelet and scaling coefficients can be used to form an ANOVA or ANCOVA, details and smooths cannot be used to form such analyses (e.g., $\text{Var}(\mathbf{f}_t) \neq \sum_{j=1}^J \text{Var}(\mathbf{f}_t^{dj}) + \text{Var}(\mathbf{f}_t^{sj})$). Conversely, while details and smooths can be used to form an MRA, wavelet and scaling coefficients cannot be used to form such an analysis (e.g., $\mathbf{f}_t \neq \sum_{j=1}^J \mathbf{f}_t^{wj} + \mathbf{f}_t^{vj}$) (Percival and Walden, 2000, p. 180–182).

Let us now define $K \times 1$ vectors $\beta_{i,wj}$ and $\beta_{i,vj}$ as

$$\beta_{i,wj} \equiv \text{Var}(\mathbf{f}_t^{wj})^{-1} \text{Cov}(\mathbf{f}_t^{wj}, R_{i,t}^{wj}) \quad \text{and} \quad \beta_{i,vj} \equiv \text{Var}(\mathbf{f}_t^{vj})^{-1} \text{Cov}(\mathbf{f}_t^{vj}, R_{i,t}^{vj}), \tag{8}$$

respectively. As in (1), these vectors can be equivalently defined as slope coefficients

$$R_{i,t}^{wj} = \alpha_i + \beta_{i,wj} \mathbf{f}_t^{wj} + \varepsilon_{i,t} \quad \text{and} \quad R_{i,t}^{vj} = \alpha_i + \beta_{i,vj} \mathbf{f}_t^{vj} + \varepsilon_{i,t}. \tag{9}$$

As the definitions suggest, $\beta_{i,wj}$ and $\beta_{i,vj}$ are betas when the variation in each variable has been restricted to the j th shortest and the longest timescales, respectively. An immediate interpretation is to view them as the quantities of risk measured at the corresponding timescales. Alternatively, one can also view them as the amounts of risk that investors operating at the corresponding timescale bear by holding asset i . Following Gençay et al. (2003, 2005), who coined the term wavelet beta for a single-factor version of $\beta_{i,wj}$, we call

¹² For example, (3), which states

$$\begin{pmatrix} f_{1,t} \\ \vdots \\ f_{K,t} \end{pmatrix} = \begin{pmatrix} f_{1,t}^{d1} \\ \vdots \\ f_{K,t}^{d1} \end{pmatrix} + \begin{pmatrix} f_{1,t}^{d2} \\ \vdots \\ f_{K,t}^{d2} \end{pmatrix} + \dots + \begin{pmatrix} f_{1,t}^{dj} \\ \vdots \\ f_{K,t}^{dj} \end{pmatrix} + \begin{pmatrix} f_{1,t}^{sj} \\ \vdots \\ f_{K,t}^{sj} \end{pmatrix}$$

can be rewritten as

$$\begin{pmatrix} f_{1,t} \\ \vdots \\ f_{K,t} \end{pmatrix} = \begin{pmatrix} f_{1,t}^{sj} \\ \vdots \\ f_{K,t}^{sj} \end{pmatrix} + \begin{pmatrix} f_{1,t}^{dj} \\ \vdots \\ f_{K,t}^{dj} \end{pmatrix} + \dots + \begin{pmatrix} f_{1,t}^{d2} \\ \vdots \\ f_{K,t}^{d2} \end{pmatrix} + \begin{pmatrix} f_{1,t}^{d1} \\ \vdots \\ f_{K,t}^{d1} \end{pmatrix}$$

and so on. We thank the referee for bringing to our attention this possibility.

¹³ As indicated, wavelet (scaling) covariance at a given level is simply the covariance between the wavelet (scaling) coefficients at that level; and wavelet (scaling) variance at a given level is simply the variance of the wavelet (scaling) coefficients at that level.

¹⁴ The gain function associated with detail (smooth) at a given level is equal to the squared gain function associated with the wavelet (scaling) coefficient at that level (see footnote 9). Thus, the ideal brick wall filters that they approximate are associated with the same band of frequencies.

the vector β_{i,w_j} the vector of level j wavelet betas, and similarly the vector β_{i,v_j} the vector of level J scaling betas.

Given the decompositions in (6) and (7) and the definitions in (8), the standard betas in (2) can be represented as follows:

$$\beta_i \equiv \text{Var}(\mathbf{f}_i)^{-1} \left[\sum_{j=1}^J \text{Cov}(\mathbf{f}_i^{w_j}, R_{i,t}^{w_j}) + \text{Cov}(\mathbf{f}_i^{v_j}, R_{i,t}^{v_j}) \right] = \sum_{j=1}^J \omega_{w_j} \beta_{i,w_j} + \omega_{v_j} \beta_{i,v_j}, \tag{10}$$

where $\omega_{w_j} \equiv \text{Var}(\mathbf{f}_i)^{-1} \text{Var}(\mathbf{f}_i^{w_j})$, $\omega_{v_j} \equiv \text{Var}(\mathbf{f}_i)^{-1} \text{Var}(\mathbf{f}_i^{v_j})$, and $\sum_{j=1}^J \omega_{w_j} + \omega_{v_j} = I$.¹⁵ Clearly from (10), the betas, or factor loadings, of an asset can be broken into several parts, each conditioned on a different timescale, and the overall betas are simply a weighted average of these conditional betas.

Some useful analogy can be drawn between wavelet/scaling betas and betas estimated using a longer return interval (as in, e.g., Roll, 1981; Hawawini, 1983; Brennan and Zhang, 2013; Gilbert et al., 2014; Kamara et al., 2016). For example, using *overlapping* monthly observations, Brennan and Zhang (2013) and Kamara et al. (2016) estimate τ -month betas by regressing τ -month return on τ -month factors, for each $\tau \in \{3, 6, 9, \dots, 60\}$. Like (9), these regressions use a different subset of frequencies contained in monthly data; unlike (9), however, these subsets are not mutually exclusive and all contain the zero frequency (i.e., infinite periodicity). To understand this, recall that a monthly time series contains components with periodicities greater than 2 months; a quarterly time series contains a subset consisting of components with periodicities greater than 2 quarters (6 months); and a half-yearly time series contains an even smaller subset, consisting of those with periodicities greater than one year (12 months). Continuing this, one can see that betas estimated using 4-year returns are essentially long-run betas, under Baxter and King's (1999) definition of business cycle (i.e., periodicities of 18–96 months; see Section 3.1 for further discussion). Importantly, however, we cannot isolate short-run or business-cycle betas in this manner. For example, betas estimated using 9-month returns are not business-cycle betas, but a weighted average of business-cycle and long-run betas. Similarly, betas estimated using monthly returns are not short-run betas, but a weighted average of short-run, business-cycle, and long-run betas (see (10)). An advantage of (9) over the abovementioned regressions using overlapping τ -month series is that the former cleanly separates not only long-run betas but also shorter-run betas so that we can examine their pricing implications in isolation.

Wavelet/scaling betas have at least two additional advantages over betas estimated using *nonoverlapping* τ -period data (as in, e.g., Roll, 1981; Hawawini, 1983; Gilbert et al., 2014). First, wavelet/scaling betas do not lose any data point, whereas nonoverlapping τ -period betas suffer from decreasing number of data points as the time interval increases, resulting in loss of information. Second, wavelet/scaling betas are not sensitive to the sampling rule employed to construct a particular time series, whereas nonoverlapping annual betas, for example, may vary nontrivially depending on which quarter of the year is used in computing annual series (e.g., Q1-Q1 versus Q4-Q4; see Jagannathan and Wang, 2007).

2.3. Framework

Under the expected return–beta representation, the K -factor pricing model is given by

$$E(R_{i,t}) = \lambda' \beta_i, \tag{11}$$

where λ is a $K \times 1$ vector of risk premiums. The model imposes that the cross-sectional variation in average returns arises from the cross-sectional variation in the betas. Identities (4) and (10) suggest that the cross-sectional variation in β_i derives from the cross-sectional variation in its components. By combining these, we ask how much of the cross-sectional variation in average returns can be attributed to the cross-sectional variation in each component of β_i . Specifically, we propose to examine a cross-sectional regression

$$E(R_{i,t}) = \lambda' \beta_{i,\bullet} \text{ for each } \bullet \in \{d_1, \dots, d_j, s_j\}, \tag{12}$$

where λ is used (with an abuse of notation) to denote a $K \times 1$ vector of free constants. In addition, we also examine a cross-sectional regression

$$E(R_{i,t}) = \lambda' \beta_{i,\bullet} \text{ for each } \bullet \in \{w_1, \dots, w_j, v_j\}, \tag{13}$$

where λ again denotes a $K \times 1$ vector of constants (without necessarily abusing notation; see Section 3.2 for further discussion). If the priced information in β_i is concentrated in $\beta_{i,\bullet}$ for a particular \bullet , the corresponding regression will produce lower pricing errors (or higher cross-sectional R^2) than the original model (11) as well as the remaining J regressions.

It is important to note that $\beta_{i,\bullet}$ can be cross-sectionally correlated with other timescale components of β_i , meaning that finding significant lambdas in (12) or (13) may simply be due to other, omitted beta components. In this light, we also propose additional specifications with all the remaining J beta components on the right-hand side—either individually, as in $E(R_{i,t}) = \lambda_1' \beta_{i,d_1} + \dots + \lambda_j' \beta_{i,d_j} + \lambda_{j+1}' \beta_{i,s_j}$, or combined, as in

$$E(R_{i,t}) = \lambda_1' \beta_{i,\bullet} + \lambda_2' \beta_{i,\bullet}^\perp, \tag{14}$$

where $\beta_{i,\bullet}^\perp$ denotes a weighted combination of the remaining J beta components other than $\beta_{i,\bullet}$, which can be written as

¹⁵ It follows from (6) that the weighting matrices sum to the identity matrix. However, this is not the case for the weighting matrices in (4) because we cannot in general create a valid ANOVA or ANCOVA using the MODWT details and smooths.

$\beta_{i,\bullet}^\perp \equiv \text{Var}(\mathbf{f}_t - \mathbf{f}_t^*)^{-1} \text{Cov}(\mathbf{f}_t - \mathbf{f}_t^*, R_{i,t})$ for $\bullet \in \{d_1, \dots, d_j, s_j\}$ or $\beta_{i,\bullet}^\perp \equiv [\text{Var}(\mathbf{f}_t) - \text{Var}(\mathbf{f}_t^*)]^{-1} [\text{Cov}(\mathbf{f}_t, R_{i,t}) - \text{Cov}(\mathbf{f}_t^*, R_{i,t})]$ for $\bullet \in \{w_1, \dots, w_j, v_j\}$.¹⁶ Under these specifications, the significance of the λ terms, which can be determined by t -tests if $K=1$ or F -tests if $K > 1$, will suggest the timescale location of the priced information. Of course, depending on the number of regressors (i.e., $(J + 1) \times K$) and test assets, the former specification (based on $(J + 1)$ -fold decomposition) may not always be feasible.¹⁷

Our framework provides a new diagnostic tool for econometricians. Rather than simply accepting or rejecting the empirical fit of a particular model, researchers can constructively examine the timescale betas that cause the model to perform relatively well or poorly: The MRA-based analysis will tell us which timescale components of the factors contain valuable (or little) information regarding the cross section of returns, whereas the ANOVA/ANCOVA-based analysis will tell us about the information content of different timescales at which to measure the betas. As discussed, the ANOVA/ANCOVA-based analysis is analogous to estimating τ -period betas, without changing the pricing kernel of the original model. In contrast, MRA-based analysis does change the pricing kernel by replacing the original factors with their details or smooths. In this regard, our framework (in particular the MRA-based one) may also serve as an initiative toward formulating a new pricing model. As will be demonstrated below, focusing on a timescale component of betas, rather than the overall betas, can improve pricing performance substantially.¹⁸

2.4. Discussion

In parallel studies, Gençay et al. (2003, 2005) examine what is essentially a single-factor version of (13) in the context of the CAPM. Using daily returns, these authors measure wavelet betas (but not scaling betas) at six different timescales and show that the (freely estimated) lambda on wavelet beta increases as the timescale at which the beta is measured increases. Based on this, Gençay et al. (2005, p. 68) conclude that “predictions of the CAPM are more relevant at medium- to long-run horizons as compared to short time horizons.” Here, we seek to iron out a couple of points.

First, the authors use terms such as “short-” and “long-run” to denote different timescales. Precisely speaking, however, even the longest timescale they consider is much shorter than a business cycle and hence is short run: Given that they use daily data, the level 6 wavelet coefficient is associated with periodicities between 64 and 128 days (i.e., roughly 3–6 months). Second, and more importantly, in each of the six cross-sectional regressions examined, the authors use average daily returns as the dependence variable; that is, they condition the independent variable on a different timescale, but not the dependent variable. This does not constitute testing the CAPM at different horizons. The predictions of the CAPM at different horizons can be better evaluated by conditioning both sides of the CAPM relationship as in Handa et al. (1993) and Brennan and Zhang (2013) or by breaking down pricing errors by frequency as in Cogley (2001). In light of our discussion above, their work can be viewed rather as a constructive attempt to understand the beta components that cause the CAPM to perform relatively well or poorly in matching the cross section of daily returns.

3. Empirical implementation

This section demonstrates an empirical implementation of our framework using FF's (1993) three-factor model. As an illustrative example, this model is a straightforward choice given its ubiquity in empirical finance.¹⁹ Under FF's specification, the vector of factors \mathbf{f}_t is given by $\mathbf{f}_t = (R_{m,t}, HML_t, SMB_t)'$, and the vector of factor loadings β_t is given by $\beta_t = (\beta_t^m, \beta_t^{HML}, \beta_t^{SMB})'$, where the elements are the slopes in the time-series regression corresponding to (1). The FF model is then written under the expected return–beta representation as

$$E(R_{i,t}) = \lambda_m \beta_i^m + \lambda_h \beta_i^{HML} + \lambda_s \beta_i^{SMB}. \tag{15}$$

We organize Sections 3.1 and 3.2 around (12) and (13), respectively. Other specifications such as (14) will be entertained in Section 3.3. For brevity, we base our discussion of the results mainly on those obtained from the MRA-based approach. Unless otherwise stated, we use monthly data from July 1963 to June 2008 and 25 size and book-to-market sorted portfolios as test assets.²⁰ Results based on subperiods or when the test assets are augmented with other portfolios, as suggested by Lewellen et al (2010), are summarized in Section 3.3. As a technical note, we employ the *least asymmetric* wavelet filter of width $L=8$ (denoted by LA(8)) throughout the paper, but our findings below are not specific to the particular wavelet filter used.

¹⁶ This specification is based on a twofold decomposition of the betas, given by $\beta_t \equiv \omega_\bullet \beta_{t,\bullet} + \omega_\bullet^\perp \beta_{t,\bullet}^\perp$, where $\omega_\bullet^\perp \equiv \text{Var}(\mathbf{f}_t)^{-1} \text{Var}(\mathbf{f}_t - \mathbf{f}_t^*)$ for $\bullet \in \{d_1, \dots, d_j, s_j\}$ or $\omega_\bullet^\perp \equiv \text{Var}(\mathbf{f}_t)^{-1} [\text{Var}(\mathbf{f}_t) - \text{Var}(\mathbf{f}_t^*)]$ for $\bullet \in \{w_1, \dots, w_j, v_j\}$. Again, $\omega_\bullet + \omega_\bullet^\perp = I$ only for $\bullet \in \{w_1, \dots, w_j, v_j\}$.

¹⁷ For example, in our application in Section 3, where $K=3, J=6$, and the number of test assets is 25, such a specification would reduce the power of the tests too much.

¹⁸ This by no means suggests that only one timescale should matter in general. Depending on factors, it is quite possible that the priced information may be more widely dispersed, so more than one timescale may matter. We thank the referee for pointing this out.

¹⁹ Moreover, the FF model allows us to demonstrate the framework in a more general setting than do single-factor models such as the CAPM or the consumption CAPM (CCAPM). Nonetheless, results obtained using the CCAPM are also provided in Online Appendix Table A4. Although beyond the scope of this paper, it might be potentially interesting to link the CCAPM results (e.g., a large and significant lambda on $\beta_{i,s4}^c$) to the work of DewBecker and Giglio (2016), which shows that periodicities greater than 32 quarters are priced in the cross section of stock returns, or to the work of Hansen et al. (2008), which shows that the cash flows of value portfolios have more exposure to consumption variations at low frequencies than do the cash flows of growth portfolios. We thank the referee for suggesting these possible links.

²⁰ We thank Kenneth French for making the portfolio and factor data available on his website.

3.1. Empirical implementation of (12)

The implementation of (12) begins with that of (3), which consists of an element-wise MRA for \mathbf{f}_t that can be written under the FF's specification as

$$\begin{pmatrix} R_{m,t} \\ HML_t \\ SMB_t \end{pmatrix} = \sum_{j=1}^J \begin{pmatrix} R_{m,t}^{d_j} \\ HML_t^{d_j} \\ SMB_t^{d_j} \end{pmatrix} + \begin{pmatrix} R_{m,t}^{s_J} \\ HML_t^{s_J} \\ SMB_t^{s_J} \end{pmatrix}, \tag{16}$$

where the superscripts d_j and s_J denote level j details and level J smooths for the corresponding factors, respectively. An immediate practical consideration is the choice of decomposition level J . In our empirical demonstration, we set $J=6$, that is, the maximum level possible for the size of the data at hand (see footnote 10). Given the exploratory nature of our procedure, it is natural to examine as many timescales as possible, but setting $J=6$ also facilitates mapping the resulting (seven) timescales into three notionally or pedagogically convenient groups: The shortest three timescales correspond to the short run, the next three correspond to intermediate run or business cycles, and the last one corresponds to the long run.

To explain this interpretation, let us recall the following. First, in terms of periodicity, level j detail for monthly data relates to periodicities between 2^j and 2^{j+1} months, and level 6 smooth relates to periodicities greater than 128 months. Second, the business-cycle literature defines the business cycle in terms of a specified range of periodicities. For example, [Baxter and King \(1999\)](#) associate the business cycle with periodicities between 1.5 and 8 years (18–96 months); [Burns and Mitchell \(1946\)](#) define it as ranging up to 10 or 12 (not 8) years (18–120 or 18–144 months). Given that [Baxter and King \(1999\)](#) refer to Burns and Mitchell's (1946) definition of business cycle in support of their choice, and that cycles exceeding 8 years in duration are no longer very rare in the recent era ([Zarnowitz and Ozyildirim, 2006](#)), it is reasonable to regard level 4, 5, and 6 details (with periodicities of 16–32, 32–64, and 64–128 months, respectively) as business-cycle components. As in the business-cycle literature, we can also regard the remaining components with shorter periodicities (i.e., level 1, 2, and 3 details) as irregular or short-run components, and those with longer periodicities (i.e., level 6 smooth) as slow-moving, long-run trend.²¹

Let us now take a look at the dynamics of the FF factors at these seven timescales. [Fig. 4](#) illustrates the level $J=6$ MRA for the *HML* factor; plots for the other factors are omitted for brevity, but available from the authors. The uppermost series is level 6 smooth, and the six series below it are details at levels 6 to 1. As explained, these seven series sum to the original time series, plotted at the bottom of the figure. The figure reveals the movements in the value premium that are transient in nature and limited to certain timescales. For example, the increased volatility following the 1973 oil crisis is localized primarily in the level 1 and 2 details, and the impact of the 1987 stock market crash is limited to the level 1 detail. In contrast, the greater changes in volatility surrounding the burst of the technology bubble in early 2000 seem to be collocated in details at levels 1–5, while there is also a hint of increased variability in the level 3 detail prior to the main portion of the event. It is also interesting to observe that the relatively broad spikes in the value premium at around the start of some recessions seem to arise primarily from the level 3 detail.

We now turn to the cross-sectional pricing implications of isolating different timescale fluctuations in the factors. As outlined, we run a series of cross-sectional regressions in (12), which can be rewritten in the current context as

$$E(R_{i,t}) = \lambda_m \beta_{i,\bullet}^m + \lambda_n \beta_{i,\bullet}^{HML} + \lambda_s \beta_{i,\bullet}^{SMB}, \tag{17}$$

where $\bullet \in \{d_1, \dots, d_6, s_6\}$. Note that the regressors above are estimated for each asset from the time-series regressions corresponding to (5). For convenience, we denote (17) by either FFd_1, \dots, FFd_6 , or FFs_6 , depending on the timescale components of the original factors the model uses as pricing factors.

[Tables 1 and 2](#) report the results from the [Fama and MacBeth's \(1973\)](#) cross-sectional regressions. In each table, Panel A contains seven subpanels, each corresponding to a different variant model, and Panel B corresponds to the original FF model to facilitate comparisons. Reported below the Fama–MacBeth estimates and t -statistics is a second set of t -statistics based on the stationary bootstrap of [Politis and Romano \(1994\)](#). In addition to allowing for time-series dependencies such as heteroskedasticity and autocorrelation, our bootstrap procedure also incorporates error in the estimation of details or smooths, as well as of factor loadings. The last two columns provide the cross-sectional R^2 and weighted sum of squared pricing errors (WSSPE) as summary statistics for the fit of each cross-sectional regression. The WSSPE uses a weighting matrix that gives less weight on noisy observations yet is independent of the specific model, so that we can compare models by their WSSPEs.²² Also reported (in brackets) are the 95% confidence intervals for the true cross-sectional R^2 s, reported following the suggestion of [Lewellen et al. \(2010\)](#),²³ and the p -values associated with the WSSPEs for the null hypothesis that the pricing errors in each model are jointly equal to zero. As are the second set of t -statistics, the confidence intervals and p -values are robust to time-series dependence and the two-layered estimation error. A

²¹ Similarly, [Yogo, 2008](#) draws on [Baxter and King \(1999\)](#) and regards level 1 detail for quarterly data (with periodicities of 2–4 quarters) as high-frequency noise, level 2, 3, and 4 details (with periodicities of 4–8, 8–16, and 16–32 quarters, respectively) as business-cycle components, and level 4 smooth (with periodicities greater than 32 quarters) as long-run trend.

²² Specifically, we follow [Campbell and Vuolteenaho \(2004\)](#) and define the WSSPE as $\hat{\alpha}' \hat{\Omega}^{-1} \hat{\alpha}$, where $\hat{\alpha}$ is the vector of estimated pricing errors and $\hat{\Omega}$ is a diagonal matrix with sample return variance on the main diagonal. An advantage of using $\hat{\Omega}$ over a freely estimated variance–covariance matrix of test-asset returns is that the former behaves better when inverted, especially with a large number of test assets.

²³ If a confidence interval for the true R^2 includes 1, this means we cannot reject the null that the corresponding model works perfectly. If a confidence interval for the true R^2 includes 0, this means we cannot reject the null that the corresponding model has no explanatory power.

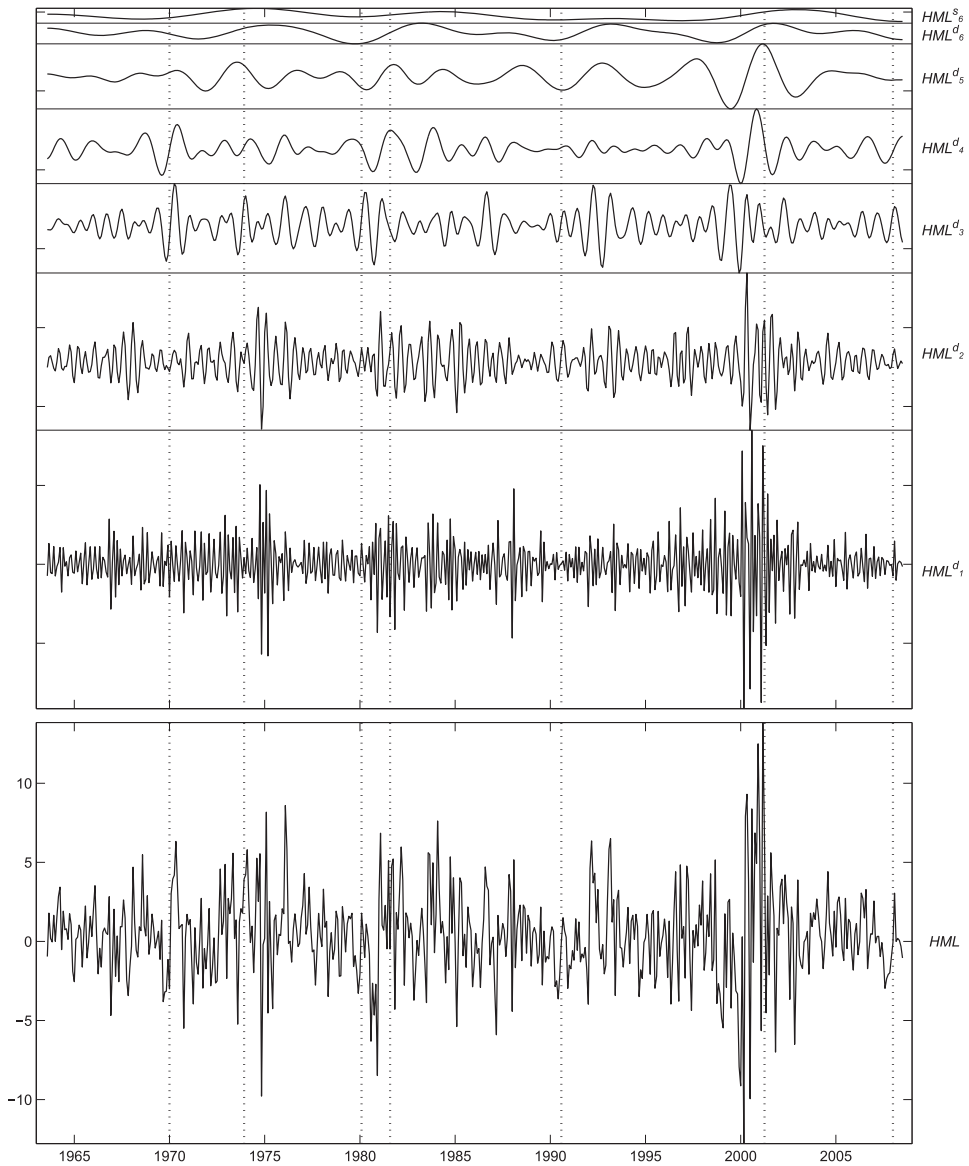


Fig. 4. Dynamics of value premium at seven timescales. MODWT MRA of level $J=6$ for the HML factor (in percentage) using the LA(8) filter. The dotted vertical lines indicate the start of NBER recession periods.

detailed description of our bootstrap procedure for the t -statistics, confidence intervals, and p -values is given in the Online Appendix.

From Panel A of Table 1, we see considerable differences across the seven variant models in their ability to explain the cross section of the 25 portfolio returns (see the R^2 s and WSSPEs). We find, however, that both the highest R^2 and lowest WSSPE come from the same model, the FFd_3 model. Note that the FFd_3 model also compares favorably with the original FF model, reported in Panel B. The FF model, which quantifies factor loadings as a weighted combination of those used in (17) via (4), indeed produces only “average looking” performance that falls within the range of model fits indicated in Panel A. The results suggest that focusing on only the level 3 details for the FF factors sharpens their explanatory power, to the extent that the pricing errors of the FF model are reduced to insignificance (see the upper bound of the confidence interval for the true R^2 and the p -value associated with the WSSPE). Given the width of the confidence interval for the FFd_3 model compared to that for the original FF model, failure to reject the null hypothesis that the FFd_3 model works perfectly does not seem to be driven much by sampling errors, unlike the case of the FFd_4 , FFd_5 , FFd_6 , and FFs_6 models.

Table 2 provides similar implications. From Panel A, we see that the intercept of the FFd_3 model is much smaller than that of the other variant models and is also statistically insignificant after accounting for sampling error (see the bootstrapped t -statistic). Only the FFs_6 model produces a comparable insignificant intercept, but its cross-sectional fit falls far short of the FFd_3 model's; all other variant models have a larger, statistically significant intercept, even after taking sampling error into account. Panel B shows that the original FF model also leaves a

Table 1
MRA-based cross-sectional regressions.

Panel A: $E(R_{i,t}) = \lambda_m \beta_{i,\bullet}^m + \lambda_h \beta_{i,\bullet}^{HML} + \lambda_s \beta_{i,\bullet}^{SMB}$					
	λ_m	λ_h	λ_s	Adj. R^2	WSSPE
• = d_1					
Estimate	0.33	0.41	0.19	0.58	0.016
t-value	2.02	3.74	1.51	[0.12, 0.93]	[0.007]
Bootstrap-t	2.23	3.60	1.27		
• = d_2					
Estimate	0.30	0.30	0.17	0.63	0.014
t-value	2.43	3.77	1.83	[0.19, 0.95]	[0.010]
Bootstrap-t	2.79	3.61	1.52		
• = d_3					
Estimate	0.30	0.31	0.16	0.72	0.010
t-value	2.53	3.96	1.87	[0.35, 1.00]	[0.053]
Bootstrap-t	3.09	3.95	1.50		
• = d_4					
Estimate	0.23	0.25	0.13	0.65	0.013
t-value	2.03	3.31	1.62	[0.24, 1.00]	[0.057]
Bootstrap-t	2.23	3.10	1.30		
• = d_5					
Estimate	0.27	0.31	0.15	0.63	0.013
t-value	2.03	3.71	1.83	[0.20, 1.00]	[0.143]
Bootstrap-t	2.33	3.56	1.54		
• = d_6					
Estimate	0.34	0.25	0.14	0.58	0.017
t-value	2.43	4.44	1.78	[0.15, 1.00]	[0.310]
Bootstrap-t	2.69	2.64	1.23		
• = s_6					
Estimate	0.40	0.14	0.17	0.60	0.020
t-value	2.41	2.02	1.60	[0.22, 1.00]	[0.501]
Bootstrap-t	2.45	1.09	1.11		
Panel B: $E(R_{i,t}) = \lambda_m \beta_{i,\bullet}^m + \lambda_h \beta_{i,\bullet}^{HML} + \lambda_s \beta_{i,\bullet}^{SMB}$					
	λ_m	λ_h	λ_s	Adj. R^2	WSSPE
Estimate	0.41	0.47	0.22	0.65	0.013
t-value	2.16	3.74	1.56	[0.20, 0.95]	[0.005]
Bootstrap-t	2.52	3.68	1.29		

This table reports the cross-sectional regression results using the excess returns on 25 portfolios sorted by size and book-to-market ratio. Panel A contains results for seven variants of the FF model; Panel B contains results for the original FF model. The slope estimates are expressed as percentage per month. The first set of *t*-statistics stands for the Fama–MacBeth estimate. The second set, indicated by Bootstrap-*t*, is obtained from a bootstrap procedure designed to account for time-series dependence, as well as estimation error in details, smooths, and factor loadings. The fourth column reports the 95% confidence intervals for the true R^2 s (in brackets), given the sample adjusted R^2 s reported above them. The last column reports the weighted sum of squared pricing errors (WSSPE) employed by Campbell and Vuolteenaho (2004) and the corresponding *p*-values (in brackets) for the null hypothesis that the pricing errors are jointly zero.

relatively large and statistically significant intercept. The results suggest that focusing on the level 3 details, again, yields improvement, to the extent that the intercept of the FF model is reduced (by more than one half) to insignificance without sacrificing at all the cross-sectional fit.²⁴

In Fig. 5, we use fitted versus actual returns plots to illustrate the models’ pricing errors for each of the 25 size and book-to-market portfolios. The plot for the original FF model (Panel H of Fig. 5) shows that although the points are generally scattered quite closely around the 45-degree line, several portfolios stand out as problematic for the FF model. The improvement of the FFd_3 model over the original FF model seems to derive mainly from some of these portfolios: the growth portfolios within the largest size quintile (41 and 51) and the growth and value portfolios within the smallest size quintile (11 and 15). As shown in Panel C of Fig. 5, the FFd_3

²⁴ Lewellen et al. (2010) emphasize that theoretical restrictions, such as zero cross-sectional intercept, should be taken seriously, whether imposed *ex ante* (as in Table 1) or tested *ex post* (as in Table 2). They argue that asset pricing tests can be misleading if one cares only about a model’s high cross-sectional R^2 and small pricing errors. Indeed, apparently strong fit of the FFd_6 model in Table 2, for example, is achieved by violating the theoretical restriction of zero cross-sectional intercept.

Table 2
MRA-based cross-sectional regressions with an intercept.

Panel A: $E(R_{i,t}) = \lambda_0 + \lambda_m \beta_{i,\bullet}^m + \lambda_h \beta_{i,\bullet}^{HML} + \lambda_s \beta_{i,\bullet}^{SMB}$						
	λ_0	λ_m	λ_h	λ_s	Adj. R^2	WSSPE
$\bullet = d_1$						
Estimate	1.29	-0.72	0.38	0.19	0.72	0.011
t -value	5.02	-2.75	3.49	1.49	[0.33, 0.98]	[0.010]
Bootstrap- t	2.75	-2.12	3.45	1.22		
$\bullet = d_2$						
Estimate	0.77	-0.17	0.30	0.13	0.67	0.012
t -value	2.96	-0.86	3.71	1.42	[0.25, 0.94]	[0.005]
Bootstrap- t	1.96	-0.77	3.51	1.17		
$\bullet = d_3$						
Estimate	0.59	-0.06	0.30	0.12	0.74	0.010
t -value	2.28	-0.31	3.82	1.48	[0.36, 0.98]	[0.020]
Bootstrap- t	1.46	-0.25	4.09	1.17		
$\bullet = d_4$						
Estimate	1.09	-0.38	0.29	0.10	0.72	0.011
t -value	4.14	-2.26	3.89	1.22	[0.36, 1.00]	[0.028]
Bootstrap- t	2.56	-1.59	3.71	0.97		
$\bullet = d_5$						
Estimate	1.15	-0.48	0.29	0.13	0.79	0.008
t -value	3.59	-2.61	3.52	1.53	[0.56, 1.00]	[0.323]
Bootstrap- t	2.84	-1.91	3.58	1.28		
$\bullet = d_6$						
Estimate	0.80	-0.25	0.25	0.15	0.81	0.008
t -value	3.81	-2.71	4.33	1.89	[0.59, 1.00]	[0.286]
Bootstrap- t	2.29	-1.21	3.06	1.40		
$\bullet = s_6$						
Estimate	0.53	-0.05	0.22	0.13	0.66	0.014
t -value	2.03	-0.36	3.38	1.22	[0.30, 1.00]	[0.119]
Bootstrap- t	1.56	-0.16	2.04	0.93		
Panel B: $E(R_{i,t}) = \lambda_0 + \lambda_m \beta_i^m + \lambda_h \beta_i^{HML} + \lambda_s \beta_i^{SMB}$						
	λ_0	λ_m	λ_h	λ_s	Adj. R^2	WSSPE
Estimate	1.21	-0.74	0.46	0.19	0.73	0.010
t -value	4.14	-2.11	3.60	1.32	[0.37, 0.97]	[0.007]
Bootstrap- t	2.56	-1.78	3.64	1.07		

This table reports the cross-sectional regression results using the excess returns on 25 portfolios sorted by size and book-to-market ratio. Panel A contains results for seven variants of the FF model; Panel B contains results for the original FF model. The intercept and slope estimates are expressed as percentage per month. The first set of t -statistics stands for the Fama–MacBeth estimate. The second set, indicated by Bootstrap- t , is obtained from a bootstrap procedure designed to account for time-series dependence, as well as estimation error in details, smooths, and factor loadings. The fifth column reports the 95% confidence intervals for the true R^2 's (in brackets), given the sample adjusted R^2 's reported above them. The last column reports the weighted sum of squared pricing errors (WSSPE) employed by Campbell and Vuolteenaho (2004) and the corresponding p -values (in brackets) for the null hypothesis that the pricing errors are jointly zero.

model brings these corner portfolios closer to the 45-degree line, graphically illustrating the model's higher R^2 .

Finally, to check how much of our pricing results are just from looking at the market at different timescales versus HML and SMB at different timescales, we run the cross-sectional regressions like those in Tables 1 and 2 but without $\beta_{i,\bullet}^{HML}$ and $\beta_{i,\bullet}^{SMB}$.²⁵ The results are presented in Tables 3 and 4. As expected, looking at the market alone generally lowers the R^2 's and increases the WSSPEs. More importantly, it also renders a difference in the relative performance of the variant models. In particular, we no longer find that the highest R^2 and lowest WSSPE come about at $\bullet = d_3$ (see Table 3). β_{i,d_3}^m also leaves a large, statistically significant intercept, even after accounting for sampling error (see Table 4). These results suggest that the relative superiority of the FF $_{d_3}$ model is not from the effect of β_{i,d_3}^m but is due largely to β_{i,d_3}^{HML} and β_{i,d_3}^{SMB} .²⁶

²⁵ We thank the referee for suggesting this additional analysis.

²⁶ As an aside, note that the market beta works best at $\bullet = s_6$. When considered alone, β_{i,s_6}^m explains 27% of the cross-sectional variation in average returns (Table 3), and adding an intercept does not make the slope turn negative (Table 4).

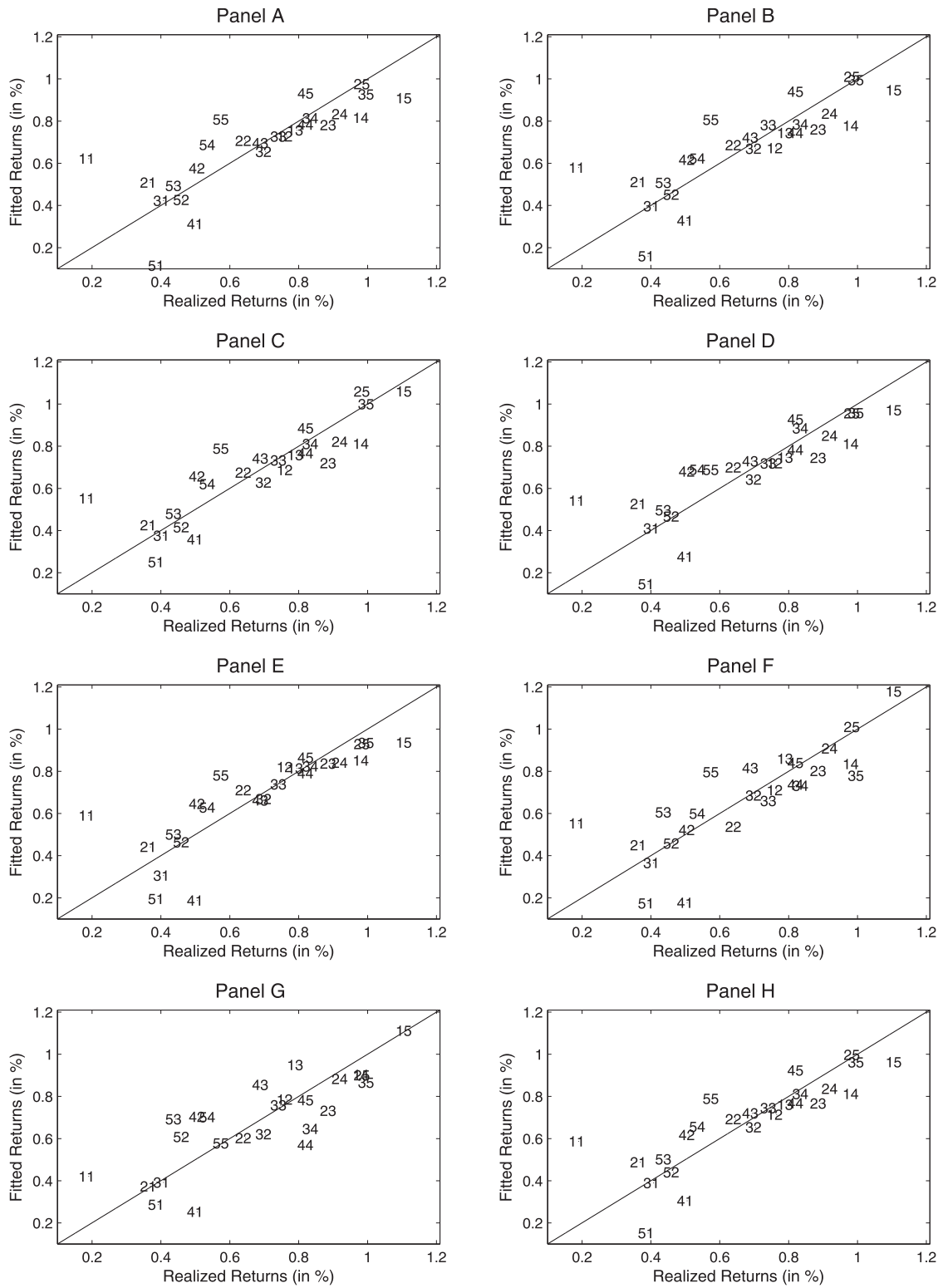


Fig. 5. Fitted versus realized returns. This figure shows the pricing errors for each of the 25 size and book-to-market sorted portfolios of the FF model and its seven variants. Each two-digit number represents one portfolio. The first digit refers to the size quintile (1 being the smallest and 5 the largest), while the second digit refers to the book-to-market quintile (1 being the lowest and 5 the highest). The pricing errors are from the Fama–MacBeth regressions in Table 1. Panels A through G correspond to the seven variant models FFd_1 through FFs_6 ; Panel H corresponds to the original FF model.

Table 3
MRA-based cross-sectional regressions: CAPM.

Panel A: $E(R_{i,t}) = \lambda_m \beta_{i,\bullet}^m$			
	λ_m	Adj. R^2	WSSPE
• = d_1			
Estimate	0.55	-0.15	0.046
t-value	3.19		[0.009]
Bootstrap-t	3.91		
• = d_2			
Estimate	0.44	-0.14	0.047
t-value	3.21		[0.012]
Bootstrap-t	3.99		
• = d_3			
Estimate	0.43	-0.18	0.050
t-value	3.21		[0.009]
Bootstrap-t	4.07		
• = d_4			
Estimate	0.38	0.01	0.043
t-value	3.25		[0.026]
Bootstrap-t	3.74		
• = d_5			
Estimate	0.45	-0.20	0.050
t-value	3.19		[0.016]
Bootstrap-t	4.17		
• = d_6			
Estimate	0.50	-0.23	0.050
t-value	3.19		[0.066]
Bootstrap-t	3.95		
• = s_6			
Estimate	0.57	0.27	0.042
t-value	3.41		[0.345]
Bootstrap-t	3.56		
Panel B: $E(R_{i,t}) = \lambda_m \beta_i^m$			
	λ_m	Adj. R^2	WSSPE
Estimate	0.66	-0.13	0.046
t-value	3.20		[0.011]
Bootstrap-t	4.02		

This table re-estimates the cross-sectional regressions in Table 1 but without $\beta_{i,\bullet}^{HML}$ and $\beta_{i,\bullet}^{SMB}$ (in Panel A) or β_i^{HML} and β_i^{SMB} (in Panel B). The betas are estimated from the same first-stage time-series regressions as in Tables 1 or 2, in order to avoid the cross-factor confounding discussed in Section 2.1, as well as to facilitate comparison. The slope estimates are expressed as percentage per month. The first set of t-statistics stands for the Fama–MacBeth estimate. The second set, indicated by Bootstrap-t, is obtained from a bootstrap procedure designed to account for time-series dependence, as well as estimation error in details, smooths, and factor loadings. The second column reports the adjusted cross-sectional R^2 s. The last column reports the weighted sum of squared pricing errors (WSSPE) employed by Campbell and Vuolteenaho (2004) and the corresponding p-values (in brackets) for the null hypothesis that the pricing errors are jointly zero.

In summary, our results in this subsection suggest that isolating different timescale fluctuations in the factors can be important in pricing risky assets, and that in the case of the FF factors much of the priced information seems contained in their level 3 details. Interestingly, the FF d_3 model improves over the FF model in the same portfolios as those in which Hahn and Lee's (2006) and Petkova's (2006) models improve over the FF model (see, e.g., Fig. 1 in Petkova, 2006).²⁷ In Section 4, we will show that the FF factors' ability to proxy for the state variable innovations considered in Hahn and Lee (2006) and Petkova (2006) comes from the level 3 details for the FF factors.

3.2. Empirical implementation of (13)

When evaluating (13), there is another theoretical restriction to take into account (in addition to zero cross-sectional intercept) if the factors in the original model are portfolio returns, as in the FF model. That is, the cross-sectional slopes should be equal to the

²⁷ Online Appendix Figure A2 also replicates and summarizes the pricing results of Hahn and Lee (2006) and Petkova (2006).

Table 4
MRA-based cross-sectional regressions with an intercept: CAPM.

Panel A: $E(R_{i,t}) = \lambda_0 + \lambda_m \beta_{i,\bullet}^m$				
	λ_0	λ_m	Adj. R^2	WSSPE
• = d_1				
Estimate	1.68	-0.81	0.04	0.044
t -value	6.03	-2.77		[0.015]
Bootstrap- t	3.71	-2.24		
• = d_2				
Estimate	1.18	-0.32	-0.01	0.042
t -value	2.55	-1.24		[0.013]
Bootstrap- t	2.19	-1.08		
• = d_3				
Estimate	1.42	-0.46	0.02	0.040
t -value	2.80	-1.74		[0.015]
Bootstrap- t	2.34	-1.29		
• = d_4				
Estimate	0.26	0.24	-0.02	0.042
t -value	0.38	0.75		[0.013]
Bootstrap- t	0.34	0.54		
• = d_5				
Estimate	1.52	-0.56	0.05	0.040
t -value	5.02	-3.37		[0.021]
Bootstrap- t	2.03	-1.22		
• = d_6				
Estimate	0.84	-0.12	-0.03	0.043
t -value	4.00	-0.82		[0.011]
Bootstrap- t	1.42	-0.34		
• = s_6				
Estimate	0.06	0.52	0.24	0.041
t -value	0.16	2.54		[0.051]
Bootstrap- t	0.15	1.53		
Panel B: $E(R_{i,t}) = \lambda_0 + \lambda_m \beta_i^m$				
	λ_0	λ_m	Adj. R^2	WSSPE
Estimate	1.74	-1.02	0.03	0.043
t -value	4.58	-2.77		[0.015]
Bootstrap- t	3.70	-2.42		

This table re-estimates the cross-sectional regressions in Table 2 but without $\beta_{i,\bullet}^{HML}$ and $\beta_{i,\bullet}^{SMB}$ (in Panel A) or β_i^{HML} and β_i^{SMB} (in Panel B). The betas are estimated from the same first-stage time-series regressions as in Tables 1 or 2, in order to avoid the cross-factor confounding discussed in Section 2.1, as well as to facilitate comparison. The intercept and slope estimates are expressed as percentage per month. The first set of t -statistics stands for the Fama–MacBeth estimate. The second set, indicated by Bootstrap- t , is obtained from a bootstrap procedure designed to account for time-series dependence, as well as estimation error in details, smooths, and factor loadings. The third column reports the adjusted cross-sectional R^2 s. The last column reports the weighted sum of squared pricing errors (WSSPE) employed by Campbell and Vuolteenaho (2004) and the corresponding p -values (in brackets) for the null hypothesis that the pricing errors are jointly zero.

factors’ expected returns (i.e., $\lambda = E(\mathbf{f}_t)$). This restriction arises because each factor has a wavelet (or scaling) beta of one on itself and zero on all the other factors from (9), and returns on any asset, including those on the factor portfolios, should satisfy (13). Under this restriction, pricing errors of (13) are computed as $\hat{\alpha}_i = \bar{R}_i - \bar{\mathbf{f}}' \hat{\beta}_{i,\bullet}$, where bars denote time-series averages. Of course, the same restriction also applies to the original model, in which case the pricing errors, $\hat{\alpha}_i$, are equal to time-series regression intercepts from (1) and the WSSPE becomes analogous to the F -statistic of Gibbons et al. (1989). The difference is that, unlike the F -statistic, the WSSPE facilitates model comparison because its weighting matrix does not vary across models. In Online Appendix Table A1, we perform this comparison and show that the performance metrics across models exhibit a similar pattern to that in Table 1. Other results obtained from the ANOVA/ANCOVA-based approach are available from the authors upon request.

3.3. Additional MRA-based analyses

Since Jagannathan and Wang (1998), some authors evaluate models by including portfolio characteristics in the cross-sectional

Table 5
Cross-sectional regressions including portfolio characteristics.

Panel A: $E(R_{i,t}) = \lambda_m \beta_{i,\bullet}^m + \lambda_h \beta_{i,\bullet}^{HML} + \lambda_s \beta_{i,\bullet}^{SMB} + \lambda_{bm} BM_{i,t-1} + \lambda_{size} SIZE_{i,t-1}$						
	λ_m	λ_h	λ_s	λ_{bm}	λ_{size}	Adj. R^2
• = d_1						
Estimate	-0.22	0.01	0.32	0.36	0.08	0.76
t-value	-0.90	0.07	2.39	3.46	4.12	
• = d_2						
Estimate	0.24	0.00	0.16	0.35	0.03	0.71
t-value	1.29	0.02	1.67	2.94	1.59	
• = d_3						
Estimate	0.21	0.20	0.16	0.14	0.02	0.76
t-value	1.19	1.40	1.95	1.03	0.84	
• = d_4						
Estimate	-0.05	0.16	0.16	0.17	0.05	0.72
t-value	-0.27	1.23	1.94	1.37	2.51	
• = d_5						
Estimate	-0.48	0.08	0.27	0.30	0.11	0.82
t-value	-2.99	0.72	3.36	3.18	5.75	
• = d_6						
Estimate	-0.24	0.21	0.23	0.11	0.07	0.76
t-value	-2.89	2.83	2.84	1.06	4.31	
• = s_6						
Estimate	0.20	0.04	0.22	0.24	0.04	0.65
t-value	1.69	0.56	1.93	2.34	1.79	
Panel B: $E(R_{i,t}) = \lambda_m \beta_{i,\bullet}^m + \lambda_h \beta_{i,\bullet}^{HML} + \lambda_s \beta_{i,\bullet}^{SMB} + \lambda_{bm} BM_{i,t-1} + \lambda_{size} SIZE_{i,t-1}$						
	λ_m	λ_h	λ_s	λ_{bm}	λ_{size}	Adj. R^2
Estimate	-0.09	0.18	0.31	0.25	0.06	0.74
t-value	-0.29	0.77	2.12	2.02	2.70	

This table reports the cross-sectional regression results using the excess returns on 25 portfolios sorted by size and book-to-market ratio. Panel A examines whether the log of the book-to-market ratio (BM) and the log of size (SIZE) have incremental explanatory power in seven variants of the FF model. Panel B examines whether BM and SIZE have incremental explanatory power in the original FF model. As in FF (1992), BM measured in December of year $t - 1$ and SIZE measured in June of year t are matched with returns from July of year t to June of year $t + 1$. The slope estimates are expressed as percentage per month, and the t -statistics are the Fama–MacBeth estimates. The last column reports adjusted R^2 s.

regression. The idea is that, if a model is well specified, portfolio characteristics should capture no residual variation in the size and book-to-market sorted portfolios. In Table 5, we present results when the log of the book-to-market ratio and log of size are added to each of the seven variant models and to the original FF model. Consistent with Daniel and Titman (1997), when the two characteristics are added to the FF model, their slope coefficients are strongly significant. After controlling for the betas prescribed by the FFd_3 model, however, the magnitudes of these coefficients are much reduced, both to insignificance. Likewise, the cross-sectional adjusted R^2 s also indicate that the two characteristics add considerable explanatory power to the FF model and to its other variants, but relatively little to the FFd_3 model. Overall, these results provide further support for the earlier finding that focusing on the level 3 details for the FF factors refines the factors’ information content about cross-sectional returns.

We also introduce to each variant model the remaining part of the original betas, that is, the part of the betas that is not already in use by each model. Note that this amounts to an empirical implementation of (14). Although (14) can be used as a model of its own, our purpose here is to examine the robustness of the FFd_3 model in the presence of the remaining part of the betas. If the priced information in the FF factors is indeed concentrated in their level 3 details, the rest should not add much to the explanatory power of the FFd_3 model. The first column of Table 6 presents the results of joint tests on the significance of the betas prescribed by each variant model (i.e., the null of $\lambda_1 = 0$); the second column presents corresponding results for the significance of the original betas deprived of those prescribed by each variant model (i.e., the null of $\lambda_2 = 0$). The results confirm the “information tilt” toward the level 3 details: In the case of • = d_3 , the null of $\lambda_1 = 0$ is soundly rejected by the data, whereas the null of $\lambda_2 = 0$ cannot be rejected at any significance level (and this failure to reject the null is not due to the sampling error, as in the case of • = s_6). None of the other beta components remains significant in the presence of the remaining part, while at the same time completely subsuming the explanatory power of the latter.

In Online Appendix Table A2, we conduct an extensive set of robustness checks. First, we consider alternative wavelet filters.

Table 6
Cross-sectional regressions including remaining components of betas.

	Unrestricted: $E(R_{i,t}) = \lambda_1' \beta_{i,\bullet} + \lambda_2' \beta_{i,\bullet}^\perp$	
	Restricted: $E(R_{i,t}) = \lambda_2' \beta_{i,\bullet}^\perp$	Restricted: $E(R_{i,t}) = \lambda_1' \beta_{i,\bullet}$
	• = d_1	
F-statistic	3.228	6.320
p-value	0.046	0.004
Bootstrap-p	0.080	0.013
	• = d_2	
F-statistic	1.941	2.439
p-value	0.157	0.096
Bootstrap-p	0.239	0.092
	• = d_3	
F-statistic	2.904	0.699
p-value	0.062	0.564
Bootstrap-p	0.039	0.538
	• = d_4	
F-statistic	0.411	0.233
p-value	0.747	0.872
Bootstrap-p	0.815	0.911
	• = d_5	
F-statistic	1.100	1.377
p-value	0.373	0.280
Bootstrap-p	0.477	0.426
	• = d_6	
F-statistic	6.851	8.927
p-value	0.003	0.001
Bootstrap-p	0.018	0.068
	• = s_6	
F-statistic	5.663	7.158
p-value	0.006	0.002
Bootstrap-p	0.081	0.306

This table examines an MRA-based implementation of (14) in the context of the FF model using the excess returns on 25 portfolios sorted by size and book-to-market ratio. The first column reports results of joint tests on the significance of the betas prescribed by each variant model (i.e., the null of $\lambda_1 = 0$); the second column reports corresponding results for the significance of the original betas deprived of those prescribed by each variant model (i.e., the null of $\lambda_2 = 0$). The first set of p-values is based on the *F* distribution. The second set, indicated by Bootstrap-p, is based on empirical distributions obtained from a bootstrap procedure designed to account for time-series dependence, as well as estimation error in details, smooths, and factor loadings.

Second, we use rolling windows and extending windows in the first-stage time-series regressions, to compare models' out-of-sample performance.²⁸ Third, in response to a comment that our results may arise simply because the periodicities associated with the level 3 detail (i.e., 8-16 months) coincide with the rebalancing frequency of the FF data (i.e., 12 months), we rerun the analysis using data constructed in the same way as described in FF (1993), except that they are rebalanced every five years. Fourth, given that the performance of the FF model as well as the size and book-to-market effects have evolved since the publication of FF (1992, 1993) (see, e.g., Schwert, 2003; Zhang, 2008), we examine separately the periods before and after January 1992, to see if the relative superiority of the FFd₃ model is sensitive to this change.²⁹ Fifth, following the suggestion of Lewellen et al. (2010), we augment the test assets with other portfolios with little factor structure, such as industry, CAPM beta, and the cluster portfolios of Ahn et al. (2009). Finally, we analyze whether our results change when we discard observations heavily influenced by boundary conditions.³⁰ Overall, the results are consistent with the main analysis: The FFd₃ model continues to stand out from its siblings and also compares

²⁸ The 120-month rolling windows and extending windows start in July 1953, so the second-stage cross-sectional regressions start in July 1963. To ensure out-of-sample setting, we perform MRA for windowed time series and use the resulting details and smooth rather than performing MRA for the full time series and windowing the resulting details and smooth.

²⁹ We also tried two and three subperiods of equal length.

³⁰ Since the MODWT (and DWT) makes use of circular filtering, which assumes that the beginning and end of a time series are joined in a loop, a filtering operation near the beginning and end of the series can introduce undesirable effects, especially when there is a large mismatch between the first and last observation. Throughout the paper, we employ reflection boundary conditions, which append a reversed version of a time series onto the end of the original series, in an effort to reduce the impact of circularity. A more conservative approach would be to drop observations heavily affected by the circularity assumption. For this robustness test, we use data from January 1947 to conduct the MRA and discard the first and last 198 observations in the resulting series. Consequently, the details and smooth cover the same period as in our other analyses and yet are virtually free of influence from circularity.

favorably with the original FF model.

Finally, to see how likely (or unlikely) it is that the FFd_3 model outperforms others by chance, we conduct the same bootstrap procedure as in the main analysis, except that we impose the null of no difference in performance across seven models: Under this null, any realized difference in performance is due to chance. To impose this null, we resample the original factors and the test assets separately, so that all seven models, as well as the FF model, become “useless” (hence, no difference in performance). Overall, it seems unlikely that our results are due to chance: Out of 5000 bootstrap simulations, we find 402 times (i.e., 8.04%) that the FFd_3 model outperforms others in the way it does across Tables 1 and 2. In particular, the “robustness” that we see above across alternative wavelet filters, alternative rebalancing frequencies, and augmented sets of test assets (i.e., across Panels A, B, E, and H of Online Appendix Table A2) occurs only 82 times in 5000 bootstrap simulations (i.e., 1.64%). These results suggest that unless the FFd_3 model is indeed better than the other models, there is at best a 8.04% (1.64%) chance that we would see the results that we see in Tables 1 and 2 (in Tables 1 and 2, as well as in Panels A, B, E, and H of Online Appendix Table A2).

4. What drives our results?

Much of the controversy surrounding the empirical success of the FF model comes from the lack of clear economic links between the FF factors and systematic risk. This is because the model's two empirically motivated factors, *HML* and *SMB*, are not designed to capture or proxy for any particular economic factor. Recently, however, progress has been made in establishing an empirical linkage between the FF factors and more theoretically structured risk factors. For example, Hahn and Lee (2006) and Petkova (2006) examine several candidate state variables within Merton's (1973) ICAPM (i.e., dividend yield, term spread, default spread, and short-term rate) and find that *HML* is related to innovations (estimated by changes or residuals) in term spread, and *SMB* is related to innovations in default spread.³¹ Consistent with FF's (1993) conjecture that the FF factors proxy for risk factors within Merton's (1973) ICAPM or Ross's (1976) arbitrage pricing theory (APT), these authors suggest that innovations in term spread and default spread are the two underlying risk factors proxied by *HML* and *SMB*, respectively.³²

This section builds on the recent development in the empirical literature and seeks to provide a risk-based explanation for what we document in Section 3. Specifically, given the finding of Hahn and Lee (2006) and Petkova (2006) that the significant relationship between the FF factors and state variable innovations gives rise to the significant explanatory power of the FF model, we hypothesize and show that the FFd_3 model does a better job of pricing cross-sectional returns than the other variants of the FF model or the FF model itself because the level 3 details for the FF factors relate more closely to innovations in term and default spreads than the other timescale components of the FF factors and the FF factors themselves. To the extent that the priced information in the FF factors is concentrated in their level 3 details, a rational risk-based story would predict the observed empirical linkage between the FF factors and state variable innovations to arise from the association of the latter with the level 3 details for the former.

As a first step toward testing this prediction, we develop an approach, inspired by Hecht and Vuolteenaho (2006), that decomposes the following regressions used in Hahn and Lee (2006) (regressions (1) and (2) in their paper):

$$HML_t \text{ (or } SMB_t) = a + b R_{m,t} + c \Delta term_t + d \Delta def_t + e_t, \quad (18)$$

where $\Delta term$ denotes changes in term spread and Δdef denotes changes in default spread. Specifically, we split the regression into seven component regressions as follows:

$$\begin{aligned} HML_t^{d_1} \text{ (or } SMB_t^{d_1}) &= a_1 + b_1 R_{m,t} + c_1 \Delta term_t + d_1 \Delta def_t + e_{1,t}, \quad ; HML_t^{d_6} \text{ (or } SMB_t^{d_6}) \\ &= a_6 + b_6 R_{m,t} + c_6 \Delta term_t + d_6 \Delta def_t + e_{6,t}, \quad HML_t^{s_6} \text{ (or } SMB_t^{s_6}) = a_7 + b_7 R_{m,t} + c_7 \Delta term_t + d_7 \Delta def_t + e_{7,t}. \end{aligned} \quad (19)$$

Recall that Hahn and Lee (2006) find, when regressing *HML*, a significant coefficient on $\Delta term$ (i.e., c) and, when regressing *SMB*, a significant coefficient on Δdef (i.e., d). An advantage of (19) is that it quantifies the relative importance of the seven components of the FF factors in establishing these empirical linkages: Because the independent variables in each component regression are the same as those in the original regression (18), it follows from summing the seven component regressions that $a = a_1 + \dots + a_7$, $b = b_1 + \dots + b_7$, $c = c_1 + \dots + c_7$, and $d = d_1 + \dots + d_7$. Consequently, the coefficients in each component regression measure the contribution from a different timescale component of *HML* (or *SMB*) in driving the regression coefficients in the original specification.

Table 7 presents the estimation results for these regressions. Panel A refers to the case where the regression of *HML* is decomposed, and Panel B refers to the case where the regression of *SMB* is decomposed. The results reveal a substantial contribution from HML^{d_3} and SMB^{d_3} to the relationships between *HML* and the term factor and between *SMB* and the default factor, respectively. In the top two rows of each panel, we first replicate and confirm Hahn and Lee's (2006) results that *HML* relates to the term factor (i.e., $\hat{c} = 1.06$ with a t -statistics of 2.30) and *SMB* relates to the default factor (i.e., $\hat{d} = -2.99$ with a t -statistics of -4.33) after

³¹ In addition to time-series evidence, Hahn and Lee (2006) and Petkova (2006) find that the loadings of the 25 portfolios on *HML* and the term factor share the same systematic pattern along the book-to-market dimension, while the loadings on *SMB* and the default factor show the same systematic pattern along the size dimension. In contrast, Petkova (2006) finds that the loadings on dividend yield and short-term rate innovations do not exhibit any systematic patterns related to size or book-to-market ratio.

³² There is also evidence linking the FF factors to future GDP growth (Liew and Vassalou, 2000; Vassalou, 2003). According to Campbell (1996), however, variables shown to forecast stock market returns, such as those considered in Hahn and Lee (2006) and Petkova (2006), conform more closely to the ICAPM than important macroeconomic variables such as GDP growth. Moreover, Petkova (2006) shows that the mimicking portfolio of Vassalou (2003) that captures news about future GDP growth loses its explanatory power for the cross section in the presence of the state variable innovations factors.

Table 7
Decomposing Hahn and Lee's (2006) regressions.

Dep. var.	Independent variables				
	Const.	R_m	$\Delta term$	Δdef	Adj. R^2
Panel A: Decomposing the regression of <i>HML</i> on the independent variables					
<i>HML</i>	0.55 (5.00)	-0.28 (-4.72)	1.06 (2.30)	0.46 (0.98)	0.17
<i>HML</i> ^{d1}	0.05 (2.07)	-0.13 (-4.08)	0.42 (2.02)	0.29 (0.95)	0.09
<i>HML</i> ^{d2}	0.03 (2.14)	-0.07 (-4.53)	-0.05 (-0.25)	0.19 (0.63)	0.06
<i>HML</i> ^{d3}	0.02 (1.99)	-0.04 (-4.46)	0.55 (4.03)	-0.02 (-0.13)	0.06
<i>HML</i> ^{d4}	0.01 (1.85)	-0.02 (-2.76)	0.07 (0.97)	0.13 (1.25)	0.02
<i>HML</i> ^{d5}	0.01 (1.92)	-0.01 (-2.28)	-0.03 (-0.41)	0.07 (0.95)	0.01
<i>HML</i> ^{d6}	0.00 (1.30)	-0.00 (-1.13)	0.08 (1.75)	-0.13 (-2.70)	0.01
<i>HML</i> ³⁶	0.43 (3.65)	-0.00 (-0.83)	0.02 (1.15)	-0.06 (-2.52)	0.00
Panel B: Decomposing the regression of <i>SMB</i> on the independent variables					
<i>SMB</i>	0.15 (0.93)	0.22 (5.27)	0.48 (1.43)	-2.99 (-4.33)	0.11
<i>SMB</i> ^{d1}	-0.02 (-1.77)	0.04 (1.92)	-0.26 (-1.21)	-0.18 (-0.68)	0.00
<i>SMB</i> ^{d2}	-0.03 (-2.60)	0.08 (5.93)	0.48 (2.46)	-1.19 (-3.11)	0.09
<i>SMB</i> ^{d3}	-0.02 (-2.35)	0.06 (6.71)	0.18 (1.45)	-0.96 (-5.24)	0.11
<i>SMB</i> ^{d4}	-0.01 (-2.08)	0.03 (4.42)	0.05 (0.59)	-0.53 (-4.28)	0.06
<i>SMB</i> ^{d5}	-0.00 (-1.81)	0.01 (2.12)	0.01 (0.22)	-0.06 (-0.87)	0.01
<i>SMB</i> ^{d6}	0.00 (0.68)	-0.00 (-0.58)	0.03 (0.75)	-0.07 (-1.36)	-0.00
<i>SMB</i> ³⁶	0.24 (1.49)	-0.00 (-0.29)	-0.01 (-0.28)	0.00 (0.02)	-0.01

Panel A decomposes the time-series regressions of *HML* on the term factor and the default factor, controlling for market excess returns. Panel B decomposes the time-series regressions of *SMB* on the same set of independent variables. The numbers reported are the coefficient estimates of the regressions with the associated *t*-statistics in parentheses. The *t*-statistics are obtained from a bootstrap procedure designed to account for time-series dependence, as well as estimation error in the dependent variables and independent variables. The last column reports adjusted R^2 's. The term factor and default factor are simple changes in term and default spreads, respectively. Following Hahn and Lee (2006), we define the term spread as the spread between 10-year and 1-year Treasury constant maturity rates, and the default spread as the spread between yield to maturity on a Baa corporate bond and 10-year Treasury constant maturity rate. These data are obtained from the Federal Reserve Bank of St. Louis.

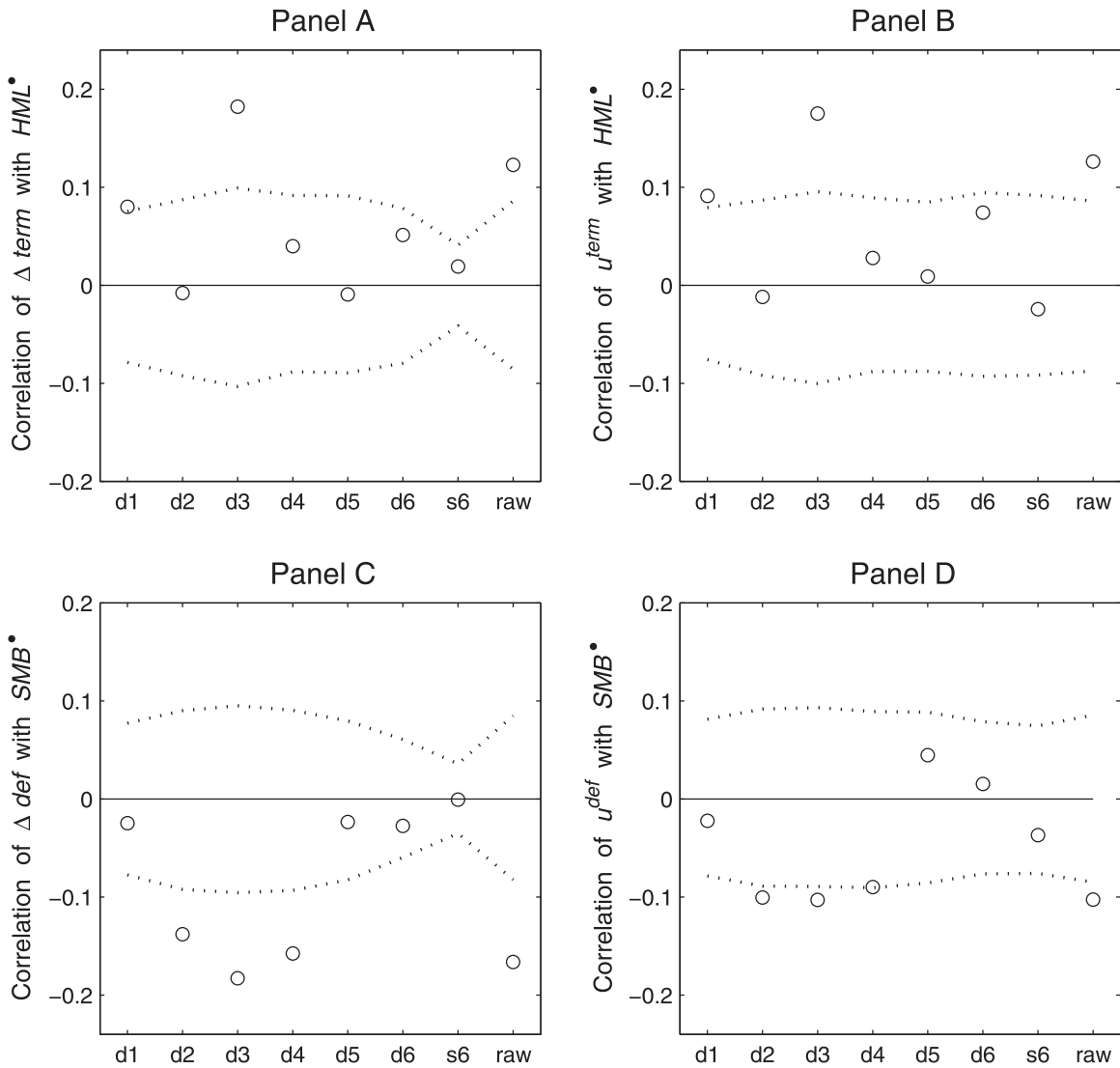


Fig. 6. Correlations between the Fama–French factors and state variable innovations when focusing on a timescale component of the former. In Panels A and B, the first seven points indicate the partial correlation of HML^* , for $\bullet \in \{d_1, \dots, d_6, s_6\}$, with the term factor, while the last point indicates that of HML . Similarly, in Panels C and D, the first seven points indicate the partial correlation of SMB^* , for $\bullet \in \{d_1, \dots, d_6, s_6\}$, with the default factor, while the last point indicates that of SMB . The partial correlations control for excess market returns. Indicated in dotted lines are the 5% critical values, under the null of zero correlation, obtained from a bootstrap procedure designed to account for time-series dependence and estimation error.

controlling for market excess returns. The third column of Panel A then reveals that more than half of the estimated regression coefficient of HML on $\Delta term$ comes from the effect of HML^{d_3} (i.e., $\hat{c}_3 = 0.55$ with a t -statistic of 4.03), while the fourth column of Panel B shows that about one third of the estimated regression coefficient of SMB on Δdef comes from the effect of SMB^{d_3} (i.e., $\hat{d}_3 = -0.96$ with a t -statistic of -5.24). As indicated by the t -statistics, the contribution from the level 3 details is also statistically significant.

The component regression results are insensitive to the definitions of the term and default factors and to the inclusion of other state variable risk proxies considered in Petkova (2006). In an unreported table, we follow Hahn and Lee (2006) and use residuals, rather than simple changes, estimated from a simple autoregressive specification. The results are almost identical.³³ In Online Appendix Table A1, we also follow Petkova (2006) and use residuals in dividend yield (u^{div}), term spread (u^{term}), default spread (u^{def}), and short-term rate (u^{r^f}), estimated from a vector autoregressive (VAR) model.³⁴ The contribution from HML^{d_3} and SMB^{d_3} to

³³ If anything, the results are cleaner. For example, the significant negative coefficients of HML^{d_6} and HML^{s_6} on the default factor are no longer observed using residuals.

³⁴ Following Petkova (2006), we specify a VAR for the vector including the FF three factors and the four state variables. We also specify a VAR for the vector including the market factor and the four state variables. The results are very similar. Thus, we report only the latter results. In any case, we follow Petkova (2006) and orthogonalize the innovations in the four state variables with respect to the market factor, and scale all innovations to have the same variance as the innovation in the market factor.

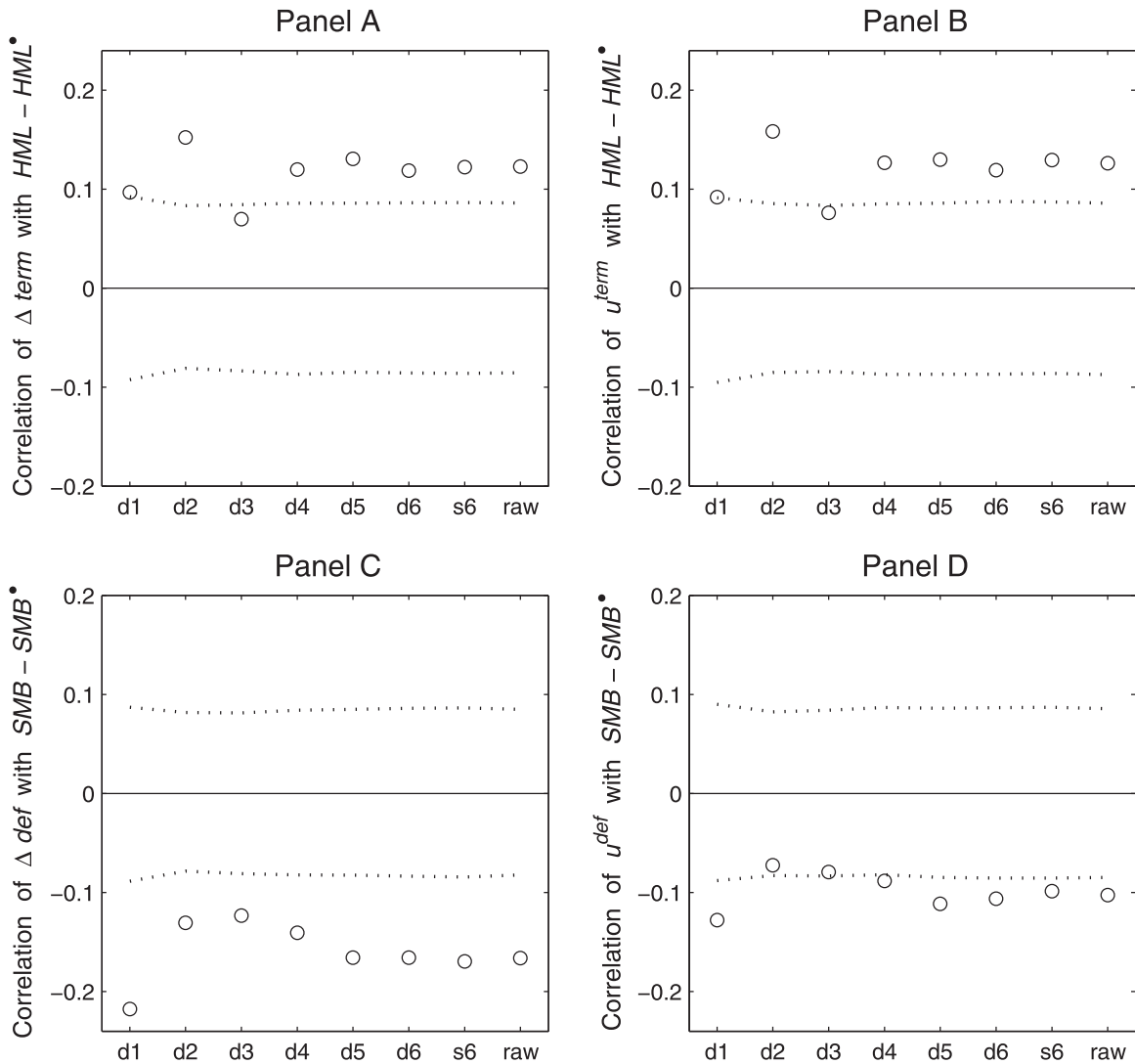


Fig. 7. Correlations between the Fama–French factors and state variable innovations when excluding a timescale component of the former. In Panels A and B, the first seven points indicate the partial correlation of $HML - HML^*$, for $\bullet \in \{d_1, \dots, d_6, s_6\}$, with the term factor, while the last point indicates that of HML . Similarly, in Panels C and D, the first seven points indicate the partial correlation of $SMB - SMB^*$, for $\bullet \in \{d_1, \dots, d_6, s_6\}$, with the default factor, while the last point indicates that of SMB . The partial correlations control for excess market returns. Indicated in dotted lines are the 5% critical values, under the null of zero correlation, obtained from a bootstrap procedure designed to account for time-series dependence and estimation error.

the relationships between HML and the term factor and between SMB and the default factor is similar to, if not greater than, that indicated in Table 7: As shown in Online Appendix Table A1, HML^{d_3} (SMB^{d_3}) accounts for about 60% (32%) of the estimated regression coefficient of HML (SMB) on the term (default) factor, with a t -statistic of 3.16 (-2.84). In contrast, the contribution from the other timescale components is mostly smaller and statistically insignificant.

It is also useful to examine how the correlations between the FF factors and the state variable innovations change when we focus on, or omit, a different timescale component of the former. To do so, we display in Fig. 6 the correlations of HML^* and SMB^* , for each $\bullet \in \{d_1, \dots, d_6, s_6\}$, with the term and default factors, respectively, and in Fig. 7 the correlations of $HML - HML^*$ and $SMB - SMB^*$, for each $\bullet \in \{d_1, \dots, d_6, s_6\}$, with the term and default factors, respectively. In each panel of Figs. 6 and 7, the last (right-most) point indicates the correlation of the original HML or SMB for comparison, while the dotted lines indicate the 5% critical values, under the null hypothesis of zero correlation, obtained from a bootstrap procedure designed to account for time-series dependence as well as estimation error (see the Online Appendix). Fig. 6 shows that focusing on the level 3 details for the FF factors generally strengthens the correlations with the associated state variable innovations, while focusing on the other timescale components reduces them, often to insignificance. Conversely, Fig. 7 shows that removing the level 3 details from the FF factors generally reduces the correlations with the state variable innovations, often to insignificance, whereas removing the other timescale components does not reduce the correlations as much, if at all.

Overall, the results are consistent with a risk-based story for what we find in Section 3: The superior performance of the FFd_3 model relative to the other variants of the FF model and the FF model itself comes from the fact that the level 3 details for the FF

factors are better proxies for risk than the other timescale components of the FF factors and the FF factors themselves are. To the extent that the innovations in term and default spreads represent the two underlying risk factors proxied by *HML* and *SMB*, focusing on the level 3 details for the FF factors refines the factors' information content about systematic risk, and thereby pricing performance.

5. Concluding remarks

Standard beta pricing models quantify an asset's systematic risk as a weighted combination of different timescale betas. This paper presents a wavelet-based framework that disentangles the explanatory power of these timescale betas for the cross section of returns and thereby locates the priced information in the standard betas. Our framework extends the earlier work of Gençay et al. (2003, 2005) by adding MRA-based regressions while rectifying their interpretation of ANOVA/ANCOVA-based regressions. The framework provides a diagnostic tool to understand what makes a certain model perform relatively well or poorly by examining the pricing implications of each timescale beta in isolation. One may thus use our framework to work backward from “diagnostic” results to arrive at a new model, especially when the basic premise of a modeler is similar to that underlying, for example, Bansal and Yaron (2004), Bansal et al. (2005), and Parker and Julliard (2005), that is, only part of the information in the standard betas is relevant for pricing risky assets and the relevant part is concentrated in certain timescale betas.

In an application to the FF three-factor model, we not only demonstrate in detail an empirical implementation of our framework but also obtain results that pose a new challenge to existing explanations for the model's well-known empirical success. Specifically, we find that the FFd_3 model explains the cross-sectional patterns in average returns better than any other variants of the FF model or the FF model itself do. The FFd_3 model also drives out portfolio characteristics and the remaining components of the betas in the cross-sectional regression. Given the robust evidence of the information tilt in the standard betas toward their scale 3 components, we expect that any viable explanation for the success of the FF model that has been applied to the FF factors should apply particularly to the level 3 details for the FF factors, but not to the other timescale components.

In this light, FF's (1993) risk-based explanation, for example, should show that the ability of the FF factors to proxy for underlying risk factors comes from the level 3 details for the FF factors, but not from the other timescale components. Drawing on the work of Hahn and Lee (2006) and Petkova (2006) for the identity of the underlying risk factors that *HML* and *SMB* proxy for, we find evidence validating the risk-based explanation. Specifically, by revisiting Hahn and Lee's (2006) regressions of *HML* or *SMB* on innovations in term and default spreads, we find that a substantial fraction of the estimated regression coefficient of *HML* (*SMB*) on the term (default) factor comes from the effect of the level 3 details for *HML* (*SMB*). In addition, while focusing on the level 3 details for the FF factors strengthens the FF factors' correlations with the associated state variable innovations, correlations between the FF factors and the state variable innovations that are otherwise significant become insignificant when we remove the level 3 details from the FF factors. While it would be interesting to evaluate in this manner other explanations for the success of the FF model, we leave this task for future work.

Acknowledgement

The authors acknowledge useful comments on earlier versions of this paper from Louis Cheng, Timothy Chue, Andy Chui, Philip Gharghori, Kewei Hou, Steven Jordan, Jinyong Kim, Myung-Jig Kim, Eric Lam, Inmoo Lee, Tse-Chun Lin, Srikant Marakani, Mujtaba Mian, Peter Pham, Wilson Tong, Cameron Truong, Steven Wei, and seminar and conference participants at the University of Sydney, Monash University, Hong Kong Polytechnic University, City University of Hong Kong, KAIST Business School, University of Hong Kong, the 2009 Financial Management Association Annual Meeting, and the 4th Annual Conference on Asia-Pacific Financial Markets. The authors are also most grateful to an anonymous referee for particularly helpful suggestions. Any errors are our own. Kang acknowledges funding support from Hong Kong Polytechnic University (grant no. A-PJ95).

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.jempfin.2017.01.004>.

References

- Adrian, T., Rosenberg, J., 2008. Stock returns and volatility: pricing the short-run and long-run components of market risk. *J. Financ.* 63, 2997–3030.
- Ahn, D.-H., Conrad, J., Dittmar, R.F., 2009. Basis assets. *Rev. Financ. Stud.* 22, 5133–5174.
- Bansal, R., Dittmar, R.F., Lundblad, C.T., 2005. Consumption, dividends, and the cross section of equity returns. *J. Financ.* 60, 1639–1672.
- Bansal, R., Yaron, A., 2004. Risks for the long run: a potential resolution of asset pricing puzzles. *J. Financ.* 59, 1481–1509.
- Baxter, M., 1994. Real exchange rates and real interest differentials: have we missed the business-cycle relationship? *J. Monet. Econ.* 33, 5–37.
- Baxter, M., King, R.G., 1999. Measuring business cycles: approximate band-pass filters for economic time series. *Rev. Econ. Stat.* 81, 575–593.
- Brennan, M.J., Zhang, Y., 2013. Capital asset pricing with a stochastic horizon. *Work. Pap.*
- Burns, A.F., Mitchell, W.C., 1946. *Measuring Business Cycles*. National Bureau of Economic Research, New York, NY.
- Campbell, J.Y., 1996. Understanding risk and return. *J. Political Econ.* 104, 298–345.
- Campbell, J.Y., Vuolteenaho, T., 2004. Bad beta, good beta. *Am. Econ. Rev.* 94, 1249–1275.
- Cogley, T., 2001. A frequency decomposition of approximation errors in stochastic discount factor models. *Int. Econ. Rev.* 42, 473–503.
- Daniel, K., Titman, S., 1997. Evidence on the characteristics of cross sectional variation in stock returns. *J. Financ.* 52, 1–33.

- Dew-Becker, I., Giglio, S., 2016. Asset pricing in the frequency domain: theory and empirics. *Rev. Financ. Stud.* 29, 2029–2068.
- Fama, E.F., French, K.R., 1989. Business conditions and expected returns on stocks and bonds. *J. Financ. Econ.* 25, 23–49.
- Fama, E.F., French, K.R., 1992. The cross-section of expected stock returns. *J. Financ.* 47, 427–465.
- Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. *J. Financ. Econ.* 33, 3–56.
- Fama, E.F., French, K.R., 1995. Size and book-to-market factors in earnings and returns. *J. Financ.* 50, 131–155.
- Fama, E.F., French, K.R., 1996. Multifactor explanations of asset pricing anomalies. *J. Financ.* 51, 55–84.
- Fama, E.F., MacBeth, J.D., 1973. Risk, return, and equilibrium: empirical tests. *J. Political Econ.* 81, 607–636.
- Gençay, R., Selçuk, F., Whitcher, B., 2002. *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*. Academic Press, San Diego, CA.
- Gençay, R., Selçuk, F., Whitcher, B., 2003. Systematic risk and timescales. *Quant. Financ.* 3, 108–116.
- Gençay, R., Selçuk, F., Whitcher, B., 2005. Multiscale systematic risk. *J. Int. Money Financ.* 24, 55–70.
- Gibbons, M.R., Ross, S.A., Shanken, J., 1989. A test of the efficiency of a given portfolio. *Econometrica* 57, 1121–1152.
- Gilbert, T., Hrdlicka, C., Kalodimos, J., Siegel, S., 2014. Daily data is bad for beta: opacity and frequency-dependent betas. *Rev. Asset Pricing Stud.* 4, 78–117.
- Hahn, J., Lee, H., 2006. Yield spreads as alternative risk factors for size and book-to-market. *J. Financ. Quant. Anal.* 41, 245–269.
- Handa, P., Kothari, S.P., Wasley, C., 1993. Sensitivity of multivariate tests of the capital asset-pricing model to the return measurement interval. *J. Financ.* 48, 1543–1551.
- Hansen, L.P., Heaton, J., Li, N., 2008. Consumption strikes back? Measuring long-run risk. *J. Political Econ.* 116, 260–302.
- Hawawini, G., 1983. Why beta shifts as the return interval changes. *Financ. Anal. J.*, 73–77.
- Hecht, P., Vuolteenaho, T., 2006. Explaining returns with cash-flow proxies. *Rev. Financ. Stud.* 19, 159–194.
- Hodrick, R.J., Prescott, E.C., 1997. Postwar U.S. business cycles: an empirical investigation. *J. Money, Credit, Bank.* 29, 1–16.
- In, F., Kim, S., 2006. The hedge ratio and the empirical relationship between the stock and futures markets: a new approach using wavelet analysis. *J. Bus.* 79, 799–820.
- Jagannathan, R., Wang, Z., 1998. An asymptotic theory for estimating beta-pricing models using cross-sectional regression. *J. Financ.* 53, 1285–1309.
- Jagannathan, R., Wang, Y., 2007. Lazy investors, discretionary consumption, and the cross-section of stock returns. *J. Financ.* 62, 1623–1661.
- Kamara, A., Korajczyk, R.A., Lou, X., Sadka, R., 2016. Horizon pricing. *J. Financ. Quant. Anal.* 51, 1769–1793.
- Kim, S., In, F., 2005. The relationship between stock returns and inflation: new evidence from wavelet analysis. *J. Empir. Financ.* 12, 435–444.
- Lewellen, J., Nagel, S., Shanken, J., 2010. A skeptical appraisal of asset pricing tests. *J. Financ. Econ.* 96, 175–194.
- Liew, J., Vassalou, M., 2000. Can book-to-market, size and momentum be risk factors that predict economic growth? *J. Financ. Econ.* 57, 221–245.
- Lintner, J., 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Rev. Econ. Stat.* 47, 13–37.
- Merton, R.C., 1973. An intertemporal capital asset pricing model. *Econometrica* 41, 867–887.
- Parker, J.A., Julliard, C., 2005. Consumption risk and the cross section of expected returns. *J. Political Econ.* 113, 185–222.
- Percival, D.B., Walden, A.T., 2000. *Wavelet Methods for Time Series Analysis*. Cambridge University Press, New York, NY.
- Petkova, R., 2006. Do the Fama-French factors proxy for innovations in predictive variables? *J. Financ.* 61, 581–612.
- Politis, D.N., Romano, J.P., 1994. The stationary bootstrap. *J. Am. Stat. Assoc.* 89, 1303–1313.
- Ramsey, J.B., Lampart, C., 1998a. Decomposition of economic relationships by timescale using wavelets: money and income. *Macroecon. Dyn.* 2, 49–71.
- Ramsey, J.B., Lampart, C., 1998b. Decomposition of economic relationships by timescale using wavelets: expenditure and income. *Stud. Nonlinear Dyn. Econ.* 3, 23–42.
- Roll, R., 1981. A possible explanation of the small firm effect. *J. Financ.* 36, 879–888.
- Ross, S.A., 1976. The arbitrage theory of capital asset pricing. *J. Econ. Theory* 13, 341–360.
- Schwert, G.W., 2003. Anomalies and market efficiency. In: Constantinides, G.M., Harris, M., Stulz, R. (Eds.), *Handbook of the Economics of Finance*. Amsterdam, The Netherlands, North-Holland.
- Sharpe, W.F., 1964. Capital asset prices: a theory of market equilibrium under conditions of risk. *J. Financ.* 19, 425–442.
- Vassalou, M., 2003. News related to future GDP growth as a risk factor in equity returns. *J. Financ. Econ.* 68, 47–73.
- Yogo, M., 2008. Measuring business cycles: a wavelet analysis of economic time series. *Econ. Lett.* 100, 208–212.
- Zarnowitz, V., Ozyildirim, A., 2006. Time series decomposition and measurement of business cycles, trends and growth cycles. *J. Monet. Econ.* 53, 1717–1739.
- Zhang, C., 2008. Decomposed Fama-French factors for the size and book-to-market effects. *Work. Pap.*