Spectral Factor Models

Federico M. Bandi, Shomesh E. Chaudhuri, Andrew W. Lo, Andrea Tamoni

Indiana University, September 3 2019

Introduction

The methodological/econometric contribution

1. Using a suitable (multivariate) Wold representation, we introduce the notion of *spectral factor model*.
2. The spectral factors are (orthogonal) components of an assumed systematic factor (e.g., the market) with cycles of different length.
3. In the model, risk is captured by spectral factor loadings, i.e., *spectral* βs.
4. We show that the traditional β is a linear combination of spectral βs *without* cross-β terms. (Hence, all frequency-specific information is contained in the spectral βs.)
5. Spectral βs can be identified using either *nonparametric* methods or *parametric* methods yielding extraction of the Wold components.
6. Spectral factor models and spectral βs are defined in the *time domain*, rather than in the *frequency domain*, something which should make both applicability and interpretability easier.

Introduction

The economic contribution

1. We provide a modeling framework which captures frequency as a key *dimension of risk*.
2. The framework is suitable to achieve dimensionality reduction in risk assessments and parsimony in the factor structures: classical risk factors may perform better once their signal is extracted properly.

Related literature

- **Frequency-domain econometrics.** Hannan (1963), Engle (1974), Corbae, Ouliaris and Phillips (2002), ...
- **Frequency-domain macro-finance.** Berkowitz (2001), Cogley (2001), Dew-Becker and Giglio (2016), ...
- **Aggregation.** Daniel and Marshall (1997), Parker and Julliard (2005), Jagannathan and Wang (2007), Cohen, Polk and Vuolteenaho (2009), ...
- **Heterogeneity in investment horizon.** Kamara, Korajczyk, Lou and Sadka (2016), Brennan and Zhang (2007), ...
- **Factors, characteristics and dimension reduction.** Harvey, Liu, and Zhu (2016), De Miguel, Martin-Utrera, Nogales, and Uppal (2017), Feng, Giglio, and Xiu (2017), Freyberger, Neuhierl, and Weber (2017), Kozak, Nagel, and Santosh (2017, 2018), Kelly, Pruitt, and Su (2018), ...

Spectral factor models

The intuition

- Assume

$$x_t^1 = \alpha + \beta x_t^2 + \varepsilon_t.$$

- Clearly, $\beta = \frac{\mathbb{C}[x_t^1, x_t^2]}{\mathbb{V}[x_t^2]}$.
- **Now, write** $x_t^1 = x_t^{1,\ <2^{j-1}} + x_t^{1,\ >2^{j-1}}$ **and** $x_t^2 = x_t^{2,\ <2^{j-1}} + x_t^{2,\ >2^{j-1}}$ **with**

$$\mathbb{C}[x_t^{i,\ <2^{j-1}}, x_t^{k,\ >2^{j-1}}] = 0$$

  $\forall i, k$. **The components are orthogonal for each process and across processes.**
- Consider now

$$x_t^1 = \alpha + \beta_1^{<2^{j-1}} x_t^{2,\ <2^{j-1}} + \beta_2^{>2^{j-1}} x_t^{2,\ >2^{j-1}} + \varepsilon_t.$$

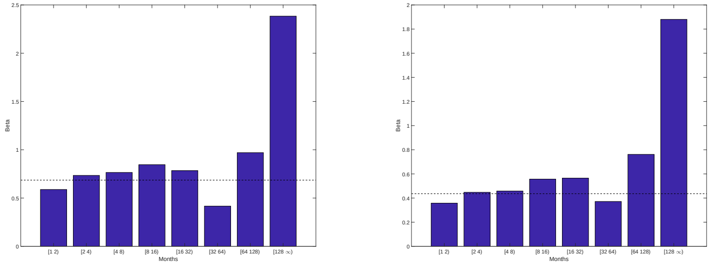
- Given the properties of the decomposition, we have

$$\beta_1^{<2^{j-1}} = \frac{\mathbb{C}[x_t^{1,\ <2^{j-1}}, x_t^{2,\ <2^{j-1}}]}{\mathbb{V}[x_t^{2,\ <2^{j-1}}]}, \quad \beta_2^{>2^{j-1}} = \frac{\mathbb{C}[x_t^{1,\ >2^{j-1}}, x_t^{2,\ >2^{j-1}}]}{\mathbb{V}[x_t^{2,\ >2^{j-1}}]}.$$

- Also,

$$\beta = v^{<2^{j-1}} \beta_1^{<2^{j-1}} + v^{>2^{j-1}} \beta_2^{>2^{j-1}}.$$

A suggestive example: Buffett's spectral βs



(a) Value-Weighted Market Beta. (b) Equal-Weighted Market Beta.

Figure: **Buffett's market beta by frequency**. Decomposition of Buffett's market beta into its various frequency components. The $x-$axis displays the frequency cycles, e.g. the first bar captures the beta corresponding to cycles with length between 2^0 and 2^1 months, the second corresponds to cycles with length between 2^1 and 2^2 months, and so on. The dashed line represent the standard market beta (not decomposed across frequencies). We will show below that the standard beta is a weighted average (with weights given by the relative variance, or information, of a specific scale) of the spectral betas. The sample period is 1976/11 to 2017/12.

Scale-wise representations: *extended* Wold

Bandi, Perron, Tamoni, Tebaldi (JoE, 2019), Ortu, Severino, Tamoni, Tebaldi (2017)

- Let $\mathbf{x} = \{x_t\}_{t \in \mathbb{Z}}$ be a covariance-stationary process defined onto the space $L^2(\Omega, \mathcal{F}, \mathbb{P})$. For simplicity, we assume the process is mean zero.
- There exists a unit variance, zero mean white noise process $\boldsymbol{\varepsilon} = \{\varepsilon_t\}_{t \in \mathbb{Z}}$ such that, for any t in $\mathbb{Z}$,

$$x_t = \sum_{k=0}^{+\infty} \alpha_k \varepsilon_{t-k},$$

  where the equality is in the L^2-norm and $\{\alpha_k\}_{k \in \mathbb{N}_0}$ is a square-summable sequence of real coefficients with $\alpha_k = \mathrm{E}(x_t \varepsilon_{t-k})$.
- Let us now define the innovation process at scale j with $j \in \mathbb{N}$. If $j = 1$, the innovation process at scale 1, denoted by $\boldsymbol{\varepsilon}^{(1)} = \left\{\varepsilon_t^{(1)}\right\}_{t \in \mathbb{Z}}$, is the process whose terms are

$$\varepsilon_t^{(1)} = \frac{\varepsilon_t - \varepsilon_{t-1}}{\sqrt{2}}.$$

  We observe that $\varepsilon_t^{(1)}$ has a zero mean and its variance is equal to 1 for all t.
- More generally, we define the innovation process at scale j as the process $\boldsymbol{\varepsilon}^{(j)} = \left\{\varepsilon_t^{(j)}\right\}_{t \in \mathbb{Z}}$ such that

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

- Each sub-series $\left\{\varepsilon_{t-k2^j}^{(j)}\right\}_{k \in \mathbb{Z}}$ is a unit variance, zero mean white noise process on the support $S_t^{(j)} = \{t - k2^j : k \in \mathbb{Z}\}$.

Scale-wise representations: *extended* Wold II

- The definition of the scale-wise shocks induces the following representation of $\mathbf{x}$:

$$x_t = \sum_{j=1}^{+\infty} x_t^{(j)} \quad \text{with} \quad x_t^{(j)} = \sum_{k=0}^{+\infty} \psi_k^{(j)} \varepsilon_{t-k2^j}^{(j)},$$

where the equality is - again - in the L^2-norm, for some square-summable sequence of real coefficients $\left\{\psi_k^{(j)}\right\}_{k\in\mathbb{Z}}$.

- Each coefficient $\psi_k^{(j)}$ is obtained by projecting $\mathbf{x}$ on the linear subspace generated by the (scale-specific) innovations $\varepsilon_{t-k2^j}^{(j)}$:

$$\psi_k^{(j)} = \mathrm{E}\left[x_t \varepsilon_{t-k2^j}^{(j)}\right].$$

- This gives rise to the *extended* Wold representation of x_t, that is

$$x_t = \sum_{j=1}^{+\infty} \sum_{k=0}^{+\infty} \psi_k^{(j)} \varepsilon_{t-k2^j}^{(j)}.$$

- The connection between the coefficients $\psi_k^{(j)}$ of the extended Wold representation and the coefficients α_k of the classical Wold representation of $\mathbf{x}$:

$$\psi_k^{(j)} = \frac{1}{\sqrt{2^j}} \left(\sum_{i=0}^{2^{j-1}-1} \alpha_{k2^j+i} - \sum_{i=0}^{2^{j-1}-1} \alpha_{k2^j+2^{j-1}+i} \right).$$

An example: AR(1)

We formalize the coefficients $\psi_k^{(j)}$ for a weakly stationary AR(1) process $\mathbf{x} = \{x_t\}_{t\in\mathbb{Z}}$, namely

$$x_t = \rho x_{t-1} + \varepsilon_t,$$

where $|\rho| < 1$ and $\boldsymbol{\varepsilon} = \{\varepsilon_t\}_{t\in\mathbb{Z}}$ is a unit variance, zero mean white noise.

- By using the lag operator $\mathbf{L}$ and the identity map I, we can - of course - rewrite the previous equation as

$$x_t = (I - \rho\mathbf{L})^{-1}\varepsilon_t = \varepsilon_t + \sum_{l=1}^{+\infty} \rho^l \varepsilon_{t-l} = \sum_{h=0}^{+\infty} \alpha_h \varepsilon_{t-h},$$

  where $\alpha_h = \rho^h$.
- Let us, now, fix a scale level $j \in \mathbb{N}$. The expression of the coefficients $\psi_k^{(j)}$ can be easily obtained:

$$\psi_k^{(j)} = \frac{\rho^{k2^j}\left(1 - \rho^{2^{j-1}}\right)^2}{\sqrt{2^j}(1-\rho)},$$

  for any $k \in \mathbb{N}_0$.
- The processes $x_t^{(j)}$ are

$$x_t^{(j)} = \frac{\left(1 - \rho^{2^{j-1}}\right)^2}{\sqrt{2^j}(1-\rho)} \sum_{k=0}^{+\infty} \rho^{k2^j} \varepsilon_{t-k2^j}^{(j)}.$$

  We observe that each $x_t^{(j)}$ is proportional to an AR(1) with time steps 2^j and autoregressive coefficient given by ρ^{2^j}. These AR(1) processes are defined on the support $S_t^{(j)} = \{t - k2^j : \quad k \in \mathbb{Z}\}$.
- In essence, then, we can rewrite the original AR(1) as an infinite sum of AR(1)s with time steps 2^j and autoregressive coefficients given by ρ^{2^j}.

Scale-wise representations: multivariate *extended* Wold

- Define the white noise process $\boldsymbol{\varepsilon} = \{(\varepsilon_t^1, \varepsilon_t^2)^\intercal\}_{t\in\mathbb{Z}}$ such that $\mathbb{E}[\boldsymbol{\varepsilon}] = 0$ and $\mathbb{E}[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^\intercal] = \Sigma_\varepsilon$, where Σ_ε is a covariance matrix of dimension 2×2. For any t in $\mathbb{Z}$, $\mathbf{x}$ satisfies the following Wold representation:

$$\begin{pmatrix} x_t^1 \\ x_t^2 \end{pmatrix} = \sum_{k=0}^{\infty} \begin{pmatrix} \alpha_k^1 & \alpha_k^2 \\ \alpha_k^3 & \alpha_k^4 \end{pmatrix} \begin{pmatrix} \varepsilon_{t-k}^1 \\ \varepsilon_{t-k}^2 \end{pmatrix} = \sum_{k=0}^{\infty} \boldsymbol{\alpha}_k \boldsymbol{\varepsilon}_{t-k}, \tag{1}$$

with $\sum_{k=0}^{\infty} \mathrm{tr}^{1/2}(\boldsymbol{\alpha}_k^\intercal \boldsymbol{\alpha}_k) < \infty$ and $\boldsymbol{\alpha}_0 = I_2$, where the equality is in the L^2-norm.

- As before, straightforward aggregation of the system's shocks now leads to the equivalent *extended* multivariate Wold representation:

$$\begin{pmatrix} x_t^1 \\ x_t^2 \end{pmatrix} = \sum_{j=1}^{\infty} \sum_{k=0}^{\infty} \boldsymbol{\Psi}_k^{(j)} \boldsymbol{\varepsilon}_{t-k2^j}^{(j)} = \sum_{j=1}^{\infty} \mathbf{x}_t^{(j)}$$

in which, for any $j \in \mathbb{N}$, the 2×2 matrices $\boldsymbol{\Psi}_k^{(j)}$ are the unique discrete Haar transforms (DHT) of the original Wold coefficients, i.e.,

$$\boldsymbol{\Psi}_k^{(j)} = \frac{1}{\sqrt{2^j}} \left(\sum_{i=0}^{2^{j-1}-1} \boldsymbol{\alpha}_{k2^j+i} - \sum_{i=0}^{2^{j-1}-1} \boldsymbol{\alpha}_{k2^j+2^{j-1}+i} \right),$$

and the 2×1 vectors $\boldsymbol{\varepsilon}_t^{(j)}$ are the DHTs of the original Wold shocks, i.e.,

$$\boldsymbol{\varepsilon}_t^{(j)} = \frac{1}{\sqrt{2^j}} \left(\sum_{i=0}^{2^{j-1}-1} \boldsymbol{\varepsilon}_{t-i} - \sum_{i=0}^{2^{j-1}-1} \boldsymbol{\varepsilon}_{t-2^{j-1}-i} \right).$$

Frequency-specific risk

Theorem 1 (A β representation.)

Assume $\mathbf{x} = \left\{(x_t^1, x_t^2)^\intercal\right\}_{t\in\mathbb{Z}}$ *satisfies Eq.* (1). *Define the spectral beta associated with frequency j as* $\beta^{(j)} = \frac{\mathbb{E}\left[x_t^{1,(j)} \; x_t^{2,(j)}\right]}{\mathbb{V}\left[x_t^{2,(j)}\right]}$. *The overall beta would, therefore, conform with*

$$\beta = \frac{\mathbb{C}\left[x_t^1, x_t^2\right]}{\mathbb{V}\left[x_t^2\right]} = \sum_{j=1}^{\infty} v^{(j)} \beta^{(j)},$$

where $v^{(j)} = \frac{\mathbb{V}\left[x_t^{2,(j)}\right]}{\mathbb{V}\left[x_t^2\right]}$.

- **Note:** In light of orthogonality of the extended Wold representation, the classical beta can be expressed as a weighted average of spectral betas (without cross-beta terms) with weights directly related to the *relative* informational content of the corresponding frequency. The latter is, of course, defined as $v^{(j)} = \frac{\mathbb{V}\left[x_t^{2,(j)}\right]}{\mathbb{V}\left[x_t^2\right]}$.

▸ Examples

Identification

- **Parametric.** In order to operationalize the extended Wold representation, we first need to compute the classical Wold coefficients, $\boldsymbol{\alpha}_k$.

  To this end, we may assume that the bivariate time series of interest, $\mathbf{x}_t = \left(x_t^1, x_t^2\right)^\intercal$, follows a linear vector autoregressive (VAR) process of order p (VAR(p)) of the form:

  $$\mathbf{x}_t = A_1\mathbf{x}_{t-1} + \ldots + A_p\mathbf{x}_{t-p} + \varepsilon_t,$$

  where the A_is, with $i = 1, \ldots, p$, are 2×2 parameter matrices and the error process, $\varepsilon_t = (\varepsilon_t^1, \varepsilon_t^2)^\intercal$, is a 2-dimensional zero-mean white noise process with covariance matrix $\mathbb{E}(\varepsilon_t\varepsilon_t^\intercal) = \Sigma_\varepsilon$.

- **Nonparametric.** We filter the components directly using a Haar transform Haar which yields:

  $$x_t = \sum_{j=1}^{J} \widehat{x}_t^{(j)} + \pi_t^{(J)},$$

  for any $J \geq 1$.

Identification: continued

Theorem 2 (Disaggregating β into spectral βs nonparametrically.)

Should a Haar transform be applied to the vector $\mathbf{x} = \left\{(x_t^1, x_t^2)^\intercal\right\}_{t\in\mathbb{Z}}$ *to decompose it into* J *decimated components, the resulting beta would conform with*

$$\widehat{\beta} = \frac{\widehat{\mathbb{C}}\left[x_t^1, x_t^2\right]}{\widehat{\mathbb{V}}\left[x_t^2\right]} = \sum_{j=1}^{J} \widehat{v}^{(j)} \widehat{\beta}^{(j)},$$

where $\widehat{\beta}^{(j)} = \frac{\widehat{\mathbb{E}}\left[\widehat{x}^{1,(j)}_{k2^j}\ \widehat{x}^{2,(j)}_{k2^j}\right]}{\widehat{\mathbb{V}}\left[\widehat{x}^{2,(j)}_{k2^j}\right]}$ *and* $\widehat{v}^{(j)} = \frac{\widehat{\mathbb{V}}\left[\widehat{x}^{2,(j)}_{k2^j}\right]}{\widehat{\mathbb{V}}[x_t]}$.

We emphasize that this representation has two key features:

- First, the Haar transform delivers a beta expressed as a linear combination of betas defined with respect to *inner products* rather than with respect to *covariances*, thereby capturing (for all samples) the zero-mean nature of the Wold components in the extended Wold.
- Second, and more importantly, the cross-beta terms do not appear, thereby representing (for all samples) the uncorrelatedness, across frequencies, of the Wold components, once more. ▸ empirical evaluation

An economic metric

Portfolio selection

- **A classical factor model.** Let $R_{i,t}$ denote the return on asset i in a universe of N stocks. Assuming M factors, the vector $\beta_i = (\beta_{i,1}, \ldots, \beta_{i,M})$ represents the asset i's sensitivities to the M factors, namely $f_t = (f_{1,t}, \ldots, f_{M,t})$. A factor decomposition of asset i's returns has the form

$$R_{i,t} = \alpha_i + \beta_i f_t^{\mathsf{T}} + \varepsilon_{i,t}.$$

It is commonly assumed that the asset-specific shocks $\varepsilon_t = (\varepsilon_{1,t}, \ldots, \varepsilon_{N,t})$ are cross-sectionally uncorrelated so that $\mathbb{E}[\varepsilon_t \varepsilon_t^{\mathsf{T}}] = D$, where D is a diagonal matrix. Letting B be an $N \times M$-matrix of factor betas and V be the $M \times M$ covariance matrix of the factors, the covariance matrix of returns Σ_R can expressed as

$$\mathbf{\Sigma_R} = \mathbf{BVB^{\mathsf{T}}} + \mathbf{D}.$$

- **A spectral factor model.** We may write a J-component spectral analogue to the previous model, i.e.,

$$R_{i,t} = \alpha_i + \sum_{j=1}^{J} \beta_i^{(j)} (f_t^{(j)})^{\mathsf{T}} + \sum_{j=1}^{J} \varepsilon_{i,t}^{(j)}.$$

Since the Wold components are *orthogonal* to one another, we have

$$\widetilde{\mathbf{\Sigma}}_{\mathbf{R}} = \sum_{\mathbf{j=1}}^{\mathbf{J}} \widetilde{\mathbf{\Sigma}}_{\mathbf{R}}^{(\mathbf{j})} \quad \text{with} \quad \widetilde{\mathbf{\Sigma}}_{\mathbf{R}}^{(\mathbf{j})} = \mathbf{B^{(j)} V^{(j)} B^{(j)\,\mathsf{T}}} + \mathbf{D^{(j)}}.$$

- **Note: $\widetilde{\mathbf{\Sigma}}_{\mathbf{R}} = \mathbf{\Sigma_R}$ if $\beta_{\mathbf{i}}^{(\mathbf{j})} = \beta_{\mathbf{i}}$ for all j, i. Hence, the classical model can be viewed as a restriction on the spectral model.**

An economic metric

The optimal portfolio problem

The optimization problem is standard:

$$\tilde{\mathbf{w}} = \arg\min_{\mathbf{w}} \mathbf{w}^{\mathsf{T}} \boldsymbol{\Sigma} \mathbf{w},$$

subject to the constraints

$$\mathbf{w}^{\mathsf{T}} \boldsymbol{\mu} = \tilde{\mu} \quad \text{and} \quad \mathbf{w}^{\mathsf{T}} \mathbf{1} = 1,$$

where w_i is the portfolio weight on the i-th security. A related optimization problem minimizes variance in the absence of restrictions on the portfolio expected return (the global minimum variance problem). The weights may be constrained to be positive (long-*only* portfolios) or may be positive and negative (should short sales be allowed).

The data:

1. The dataset and test criteria are similar to those in Ledoit and Wolf (2003). Stock return data are extracted from the University of Chicago's Center for Research in Securities Prices (CRSP) monthly database. Only U.S. common stocks traded on the New York Stock Exchange (NYSE) and the American Stock Exchange (AMEX) are included, which eliminates REIT's, ADR's, and other types of securities.
2. For $t = 1952$ to $t = 2018$, we use an *in-sample* period from August of year $t - 10$ to July of year t to form an estimate of the covariance matrix of stock returns.
3. The estimate is used in the optimizations problem(s) above.
4. The *out-of-sample* period spans the time period from August of year t to the end of July of year $t + 1$.
5. The measure of performance is the portfolio's out-of-sample standard deviation in the period from August 1962 to July 2018.

Panel A: August 1962 to July 1989

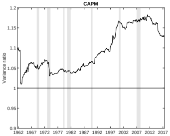
	SD Global Min	SD Min $\mid E[R] = 10\%$	SD Min $\mid E[R] = 20\%$
CAPM	1.058 (0.016)	1.060 (0.008)	1.120 (0.000)
Fama-French model	1.083 (0.309)	1.053 (0.049)	1.076 (0.023)
Fama-French plus Momentum	1.082 (0.099)	1.059 (0.044)	1.095 (0.007)
Five Principal Components	1.044 (0.092)	1.045 (0.021)	1.074 (0.001)

Panel B: August 1990 to July 2018

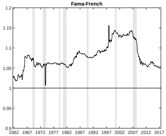
	SD Global Min	SD Min $\mid E[R] = 10\%$	SD Min $\mid E[R] = 20\%$
CAPM	1.143 (0.001)	1.138 (0.003)	1.137 (0.002)
Fama-French model	1.049 (0.082)	1.048 (0.085)	1.063 (0.123)
Fama-French plus Momentum	1.071 (0.023)	1.071 (0.019)	1.103 (0.018)
Five Principal Components	1.064 (0.015)	1.063 (0.026)	1.086 (0.039)

Panel C: August 1962 to July 2018

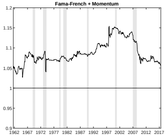
	SD Global Min	SD Min $\mid E[R] = 10\%$	SD Min $\mid E[R] = 20\%$
CAPM	1.097 (0.000)	1.095 (0.000)	1.128 (0.000)
Fama-French model	1.067 (0.065)	1.050 (0.008)	1.070 (0.003)
Fama-French plus Momentum	1.076 (0.009)	1.063 (0.004)	1.098 (0.001)
Five Principal Components	1.052 (0.004)	1.053 (0.001)	1.079 (0.000)



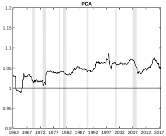
(a) GMV



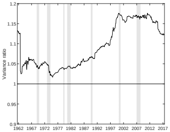
(b) GMV



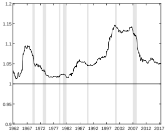
(c) GMV



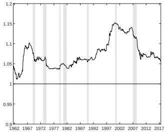
(d) GMV



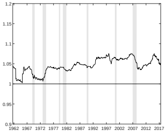
(e) Target 10%



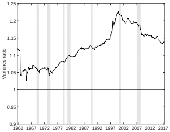
(f) Target 10%



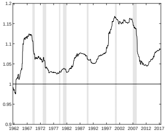
(g) Target 10%



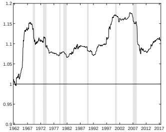
(h) Target 10%



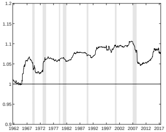
(i) Target 20%



(j) Target 20%



(k) Target 20%



(l) Target 20%

A re-evaluation of the Consumption CAPM

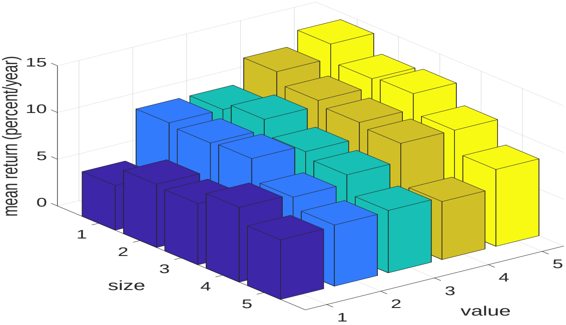
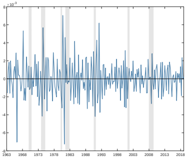
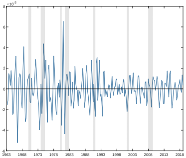


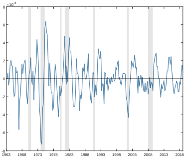
Figure: Average realized returns of the 25 Fama-French portfolios sorted on size and book-to-market.



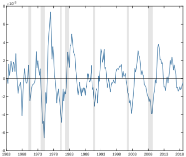
(a) $j = 1$



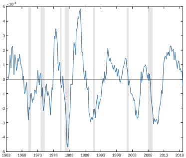
(b) $j = 2$



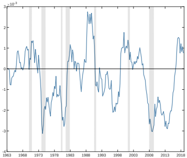
(c) $j = 3$



(d) $j = 4$

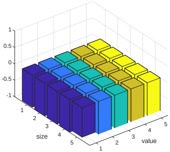


(e) $j = 5$

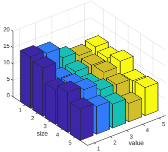


(f) $j = 6$

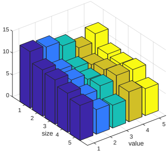
Figure: **Frequency-specific consumption components**. Each panel refers to a scale $j = 1, \ldots, J$ (scale j captures fluctuations between $2^{(j-1)}$ and 2^j quarters).



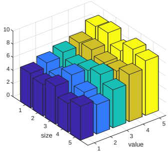
(a) $j = 1$



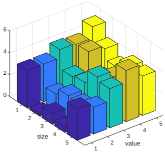
(b) $j = 2$



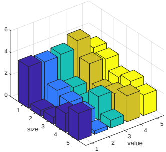
(c) $j = 3$



(d) $j = 4$

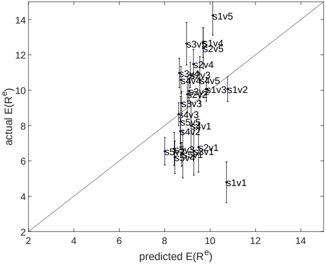


(e) $j = 5$

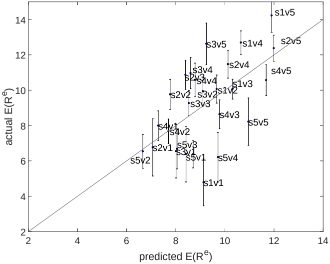


(f) $j = 6$

Figure: **Frequency-specific betas**. $\beta_i^{(j)}$s for size and book-to-market portfolios $i = 1, \ldots, 25$. Each panel refers to a scale $j = 1, \ldots, J$ (scale j captures fluctuations between $2^{(j-1)}$ and 2^j quarters).



(a) C-CAPM.



(b) Component $j = 4$.

Figure: **Cross-sectional fit**. **Panel (a)**: The figure plots fitted versus average excess returns (% per year) for the 25 size and book-to-market portfolios. **Panel (b)**: The figure plots fitted versus average excess returns (% per year) when the priced factor is the consumption component at scale $j = 4$ (a component associated with fluctuations between 2 and 4 years).

The prices of risk

Panel (a): $\mathbb{E}[R^{ei}_{t,t+1}] = \sum_{j=1}^{6} \lambda_j \beta_i^{(j)}$

Constant	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	$\sqrt{\alpha^2}$	$\|\alpha\|$	DoF	p-value	R^2
0	-0.585	0.093	-0.038	**1.413**	-0.614	0.665	1.79	1.49	21	0.000	
(-)	(0.390)	(0.233)	(0.449)	(0.483)	(0.563)	(0.674)					
0.512	-0.590	0.068	-0.045	**1.382**	-0.647	0.718	1.79	1.48	20	0.000	0.45
(0.823)	(0.380)	(0.246)	(0.462)	(0.584)	(0.551)	(0.408)					

Panel (b): $\mathbb{E}[R^{ei}_{t,t+1}] = \lambda_4 \beta_i^{(4)}$

Constant	λ_4	$\sqrt{\alpha^2}$	$\|\alpha\|$	DoF	p-value	R^2
0	**1.428**	1.94	1.64	25	0.000	
(-)	(0.617)					
1.709	**1.170**	1.91	1.63	24	0.000	0.38
(1.410)	(0.589)					

Conclusions

- Frequency is a dimension of risk.
- We provide a methodological framework designed to model and identify frequency-specific systematic risk.
- We do so in the time domain, thereby facilitating use and economic interpretability.
- We argue that emphasis on frequency may lead to economically-meaningful dimension reduction in cross-sectional pricing.

Constant *spectral* betas

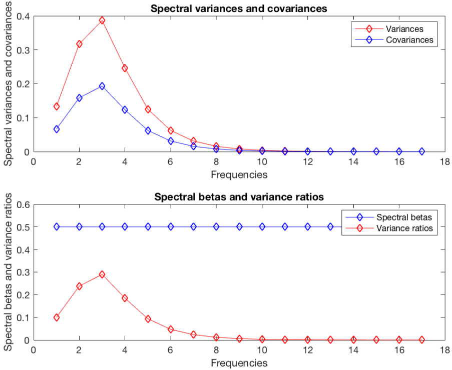


Figure: We report spectral covariances, variances and betas across frequencies. The values are derived from a bivariate VAR(1) with $\alpha^1 = 0.5$, $\alpha^2 = 0$, $\alpha^3 = 0$ and $\alpha^4 = 0.5$. The variance matrix of the bivariate shocks has $\sigma^1 = 1$, $\sigma^2 = 1$ and $\rho^{1,2} = 0.5$.

Decreasing *spectral* betas

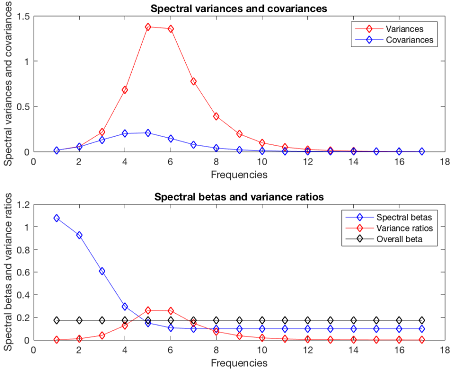


Figure: We report spectral covariances, variances and betas across frequencies. The values are derived from a bivariate VAR(1) with $\alpha^1 = 0.5$, $\alpha^2 = 0$, $\alpha^3 = 0$ and $\alpha^4 = 0.9$. The variance matrix of the bivariate shocks has $\sigma^1 = 1$, $\sigma^2 = 1$ and $\rho^{1,2} = 0.5$.

Increasing *spectral* betas

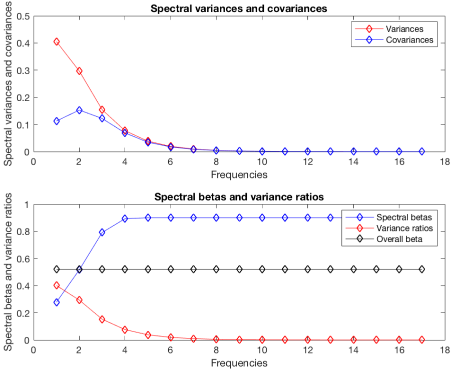


Figure: We report spectral covariances, variances and betas across frequencies. The values are derived from a bivariate VAR(1) with $\alpha^1 = 0.5$, $\alpha^2 = 0$, $\alpha^3 = 0$ and $\alpha^4 = 0.1$. The variance matrix of the bivariate shocks has $\sigma^1 = 1$, $\sigma^2 = 1$ and $\rho^{1,2} = 0.5$. back

A primer on Haar filtering

The case $J = 2$

- Consider the case $J = 1$. We have, by adding and subtracting $\frac{x_{t-1}}{2}$:

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

  which breaks the series into a "transitory" and a "persistent" component.

- For $J = 2$, by adding and subtracting $\frac{x_{t-2}+x_{t-3}}{4}$:

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

  which separates the "persistent" component $\pi_t^{(1)}$ into additional "transitory" and "persistent" components.

- This procedure can be generalized for $J > 2$. We now formalize the $J = 2$ case.

The case $J = 2$

- Let us focus on blocks of length $N = 2^J = 2^2$ and define the vector

$$X_t = [x_{t-3}, x_{t-2}, x_{t-1}, x_t]^\top.$$

- Consider, now, the orthogonal transform matrix $\mathcal{T}^{(2)}$ defined as

$$\mathcal{T}^{(2)} = \begin{pmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ -1/4 & -1/4 & 1/4 & 1/4 \\ -1/2 & 1/2 & 0 & 0 \\ 0 & 0 & -1/2 & 1/2 \end{pmatrix}.$$

- $\mathcal{T}^{(2)}(\mathcal{T}^{(2)})^\top$ is diagonal.

The case $J = 2$

- We have:

$$\mathcal{T}^{(2)} X_t = \begin{pmatrix} \frac{x_t + x_{t-1} + x_{t-2} + x_{t-3}}{4} \\ \frac{x_t + x_{t-1} - x_{t-2} - x_{t-3}}{4} \\ \frac{x_{t-2} - x_{t-3}}{2} \\ \frac{x_t - x_{t-1}}{2} \end{pmatrix} = \begin{pmatrix} \pi_t^{(2)} \\ x_t^{(2)} \\ x_{t-2}^{(1)} \\ x_t^{(1)} \end{pmatrix}.$$

- Clearly,

$$x_t = x_t^{(1)} + x_t^{(2)} + \pi_t^{(2)}.$$

Interpretation

- The generic j-th detail can be represented as follows:

$$x_t^{(j)} = \underbrace{\frac{\sum_{i=0}^{2^{(j-1)}-1} x_{t-i}}{2^{(j-1)}}}_{\pi_t^{(j-1)}} - \underbrace{\frac{\sum_{i=0}^{2^j-1} x_{t-i}}{2^j}}_{\pi_t^{(j)}},$$

where the elements $\pi_t^{(j)}$ satisfy the recursion

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

- Thus, we have

$$x_t = \sum_{j=1}^{J} x_t^{(j)} + \pi_t^{(J)} = \sum_{j=1}^{J} \left\{ \pi_t^{(j-1)} - \pi_t^{(j)} \right\} + \pi_t^{(J)} = \pi_t^{(0)}.$$

- $x_t^{(j)}$ represents changes at scale $j-1$ and $\pi_t^{(J)}$ is a long-run trend.

Decimation

- As shown, the details can be obtained in calendar time.
- They can also be obtained in their corresponding scale time:

$$\left\{x_t^{(j)},\ t = k2^j \text{ with } k \in \mathbb{Z}\right\},$$
$$\left\{\pi_t^{(J)},\ t = k2^J \text{ with } k \in \mathbb{Z}\right\}.$$

- Let us return to the case $J = 2$ and the matrix $\mathcal{T}^{(2)}$ so that

$$\mathcal{T}^{(2)} \begin{pmatrix} x_{t-3} \\ x_{t-2} \\ x_{t-1} \\ x_t \end{pmatrix} = \begin{pmatrix} \pi_t^{(2)} \\ x_t^{(2)} \\ x_{t-2}^{(1)} \\ x_t^{(1)} \end{pmatrix}.$$

- By letting t vary in the set $\left\{t = k2^2 \text{ with } k \in \mathbb{Z}\right\}$ we can construct the *decimated* counterparts $\left\{x_t^{(j)},\ t = k2^j \text{ with } k \in \mathbb{Z}\right\}$ for $j = 1, 2$ and $\left\{\pi_t^{(2)},\ t = k2^2 \text{ with } k \in \mathbb{Z}\right\}$. Back

An empirical evaluation of the β representation

Identifying the components

We run market model-style regressions on two portfolios: a high book-to-market (value) portfolio and a low book-to-market (growth) portfolio:

$$R_{value,t} = \text{const} + 1.149 \times R_{mkt,t} + \varepsilon_t, \quad R^2 = 0.90,$$
$$(\text{t-stat} = 31.69)$$
$$R_{growth,t} = \text{const} + 0.973 \times R_{mkt,t} + \varepsilon_t, \quad R^2 = 0.90.$$
$$(\text{t-stat} = 38.19)$$

Since the volatility of the excess market return series in our sample is 19.06% per annum, the beta estimates imply ...

- a covariance equal to $1.149 \times \left(19.06/\sqrt{12}\right)^2 = 34.784$ for value.
- a covariance equal to $0.973 \times \left(19.06/\sqrt{12}\right)^2 = 29.456$ for growth.

Panel A: Parametric (AR(p) based)

	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$	$j=6$	$j>6$	$\sum_{j=1}^{J+1}\mathbb{C}(R_{mkt}^{(j)}, R_p^{(j)})$
Covariance decomposition								
Value	10.273	12.533	6.535	2.994	1.516	0.558	0.219	34.626
Growth	10.020	9.796	5.204	2.515	1.222	0.346	0.139	29.242

	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$	$j=6$	$j>6$	$\sum_{j=1}^{J+1} w_p^{(j)}\beta_p^{(j)}$
Beta decomposition and reweighting								
Value	1.027	1.196	1.234	1.207	1.229	1.311	1.389	1.143
weight (rel. variance)	0.330	0.346	0.175	0.082	0.041	0.014	0.005	
Growth	1.003	0.950	0.961	1.000	0.987	0.813	0.883	0.965
weight (rel. variance)	0.330	0.340	0.179	0.083	0.041	0.014	0.005	

Panel B: Nonparametric (Haar based)

	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$	$j=6$	$j>6$	$\sum_{j=1}^{J+1}\mathbb{C}(R_{mkt}^{(j)}, R_p^{(j)})$
Covariance decomposition								
Value	13.887	10.081	5.762	2.830	1.462	0.542	0.237	34.802
Growth	12.417	8.278	4.612	2.389	1.185	0.337	0.180	29.398

	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$	$j=6$	$j>6$	$\sum_{j=1}^{J+1} w_p^{(j)}\beta_p^{(j)}$
Beta decomposition and reweighting								
Value	1.101	1.164	1.208	1.191	1.221	1.305	1.387	1.148
weight (rel. variance)	0.416	0.286	0.158	0.078	0.040	0.014	0.005	
Growth	0.984	0.956	0.967	1.005	0.989	0.812	0.882	0.969
weight (rel. variance)	0.416	0.286	0.158	0.078	0.040	0.014	0.005	

An empirical evaluation of the β representation

Orthogonality of the components

- We bundle together frequencies below 16 months and above 16 months.
- In other words, for each return series, we sum all of the components up to scale 4 (included) and dub this new component "the high-frequency component" (HF).
- Analogously, for each return series, we sum all of the components higher than scale 4 and dub this new component "the low-frequency component" (LF).

We then run the following simple regressions:

$$R_{p,t}^{LF} = \mathrm{const} + \beta_p^{LF} \times R_{mkt,t}^{LF} + \varepsilon_t,$$
$$R_{p,t}^{HF} = \mathrm{const} + \beta_p^{HF} \times R_{mkt,t}^{HF} + \varepsilon_t.$$

By the orthogonality of the components, the corresponding multiple regression, i.e.,

$$R_{p,t} = \mathrm{const} + \beta_p^{HF} \times R_{mkt,t}^{HF} + \beta_p^{LF} \times R_{mkt,t}^{LF} + \varepsilon_t,$$

should deliver analogous beta estimates.

	Panel A: Parametric (AR(p) based)				Panel B: Nonparametric (Haar based)			
	Simple Regression		Multiple Regression		Simple Regression		Multiple Regression	
	β^{LF}	β^{HF}	β^{LF}	β^{HF}	β^{LF}	β^{HF}	β^{LF}	β^{HF}
	(t-stat)	(t-stat)	(t-stat)	(t-stat)	(t-stat)	(t-stat)	(t-stat)	(t-stat)
Value	1.287	1.137	1.237	1.141	1.228	1.144	1.231	1.142
	(22.14)	(27.96)	(15.43)	(26.54)	(18.53)	(28.35)	(15.12)	(26.48)
Growth	0.937	0.986	0.874	0.986	0.947	0.985	0.891	0.983
	(18.03)	(34.38)	(11.24)	(33.48)	(20.50)	(34.39)	(11.78)	(32.84)

Table: Simple and multiple regression on high- and low-frequency betas.

back