Frequency-domain information for active portfolio management

Gonçalo Faria [1] Fabio Verona [2]

[1] Católica Porto Business School and CEGE, U.Vigo (RGEA)

[2] Bank of Finland - Monetary Policy and Research Department, U.Porto (*CEF.UP*)

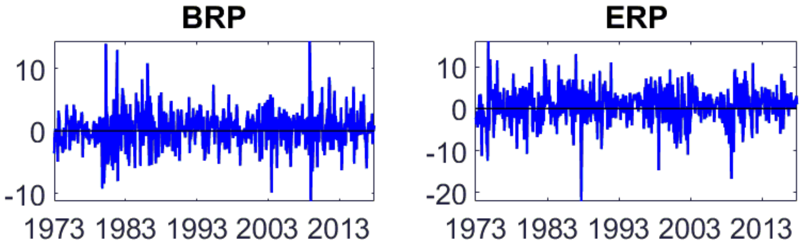
Big Data and Machine Learning in Finance Conference

Politecnico di Milano, June 10-11, 2021

Dynamic asset allocation and forecasting

- An investor implementing a **dynamic asset allocation** strategy is **actively** deviating from a benchmark with the aim of outperforming it
  - overweight (underweight) versus the benchmark on the asset classes that (s)he **expects** to outperform (underperform)
- For the construction of those **expectations** the crucial step is the process of **forecasting asset returns**
- **This paper**
  - proposes a methodology to forecast out-of-sample asset returns - bonds (10-year Treasury bond) and equities (S&P500 index) - using frequency-domain information (novelty for bond returns)
  - aims to understand the economic significance of frequency-domain information for active portfolio management (novelty)

Forecasting the bond and equity premium



- Bond risk premium (BRP) and equity risk premium (ERP): difference between the return on the 10-year Treasury bond and the return on the S&P500 index in month t, respectively, and the one-month Treasury bill known at the beginning of month t (*i.e.* the lagged risk free rate)

- Ludvigson and Ng (JFE 2007): factor analysis approach
- Rapach, Strauss, and Zhou (RFS 2010): combination forecast
- Ferreira and Santa-Clara (JFE 2011): sum-of-the-part method
- Dangl and Halling (JFE 2012): regressions with time-varying coefficients
- Pettenuzzo, Timmermann, and Valkanov (JFE 2014): impose economic constraints on the forecast
- Faria and Verona (JEF 2018): sum-of-the-part method in the time-frequency domain
- ⋮
- Our paper: forecast the ERP using frequency-decomposed predictors

Forecasting the bond premium

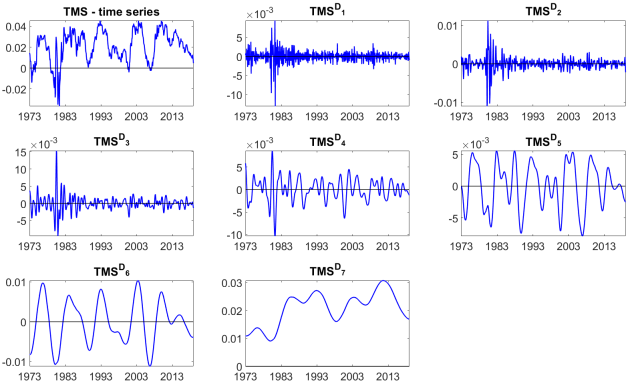
Forecasting methods

- Cochrane and Piazzesi (AER 2005): linear combination of forward rates as predictor
- Ludvigson and Ng (JFE 2007): factor analysis approach
- Sarno, Schneider, and Wagner (JEF 2016): novel estimation strategy for affine term structure models that jointly fits yields and bond excess returns
- Gargano, Pettenuzzo, and Timmermann (MS 2019): empirical modelling strategy that accounts for time-varying parameters, stochastic volatility and parameter estimation error
- Our paper: forecast the BRP using frequency-decomposed predictors

Forecasting the equity and bond premium

Why using frequency-domain information to forecast returns?

- It unveils hidden information in original time series. E.g. the U.S. term spread:



Forecasting the bond and equity premium

The standard predictors

1. Log dividend-price ratio (DP)
2. Log dividend yield (DY)
3. Log earnings-price ratio (EP)
4. Excess stock return volatility (RVOL)
5. Book-to-market ratio (BM)
6. Net equity expansion (NTIS)
7. Long-term bond yield (LTY)
8. Long-term bond return (LTR)
9. Term spread (TMS)
10. Default yield spread (DFY)
11. Default return spread (DFR)
12. Inflation rate (INFL)

Forecasting the bond and equity premium

Maximal Overlap Discrete Wavelet Transform MultiResolution Analysis (MODWT MRA)

- Kim and In (JEF 2005), "The relationship between stock returns and inflation: new evidence from wavelet analysis"
- Rua (IJF 2017), "A wavelet-based multivariate multiscale approach for forecasting"
- Faria and Verona (JEF 2018), "Forecasting stock market returns by summing the frequency-decomposed parts"
- Faria and Verona (JFM 2020), "The yield curve and the stock market: mind the long run"
- Faria and Verona (QF forthcoming), "Time-frequency forecast of the equity premium"
- ⋮

- Using the MODWT MRA decomposition (Haar wavelet filter, reflecting boundary conditions), we consider 7 frequency bands:
  - D_1: $2 \sim 4$ months
  - D_2: $4 \sim 8$ months
  - D_3: $8 \sim 16$ months
  - D_4: $16 \sim 32$ months
  - D_5: $32 \sim 64$ months
  - D_6: $64 \sim 128$ months
  - D_7: >128 months
- We recompute the frequency components at each iteration of the OOS forecasting process to make sure that we only use current and past information when making the forecasts

One nice feature of the MODWT MRA decomposition is that

$$x_t = \underbrace{x_t^{D_1}}_{2m \sim 4m} + \underbrace{x_t^{D_2}}_{4m \sim 8m} + \underbrace{x_t^{D_3}}_{8m \sim 16m} + \underbrace{x_t^{D_4}}_{16m \sim 32m} + \underbrace{x_t^{D_5}}_{32m \sim 64m} + \underbrace{x_t^{D_6}}_{64m \sim 128m} + \underbrace{x_t^{D_7}}_{>128m}$$

Out-of-sample (OOS) forecasts

- We use Goyal and Welch (RFS 2008) updated dataset
- Monthly data, U.S., January 1973 to December 2018
- The 1-step ahead OOS forecasts of *BRP* and *ERP* are generated using a sequence of expanding windows
- We use an initial sample (1973M01 to 1989M12) to make the first 1-step ahead OOS forecast
- The sample is then increased by one observation and a new 1-step ahead OOS forecast is produced
- We proceed in this way until the end of the sample, thus obtaining a sequence of 348 1-step ahead OOS forecasts
- The full OOS period therefore spans from January 1990 to December 2018

Out-of-sample (OOS) forecasts

Predictive regression model

- Predictive regression model for *ERP*:

$$ERP_{t+1} = \alpha + \beta \boldsymbol{X}_t + \varepsilon_{t+1} \tag{1}$$

- One-step ahead OLS forecast:

$$E\hat{R}P_{t+1} = \hat{\alpha}_t + \hat{\beta}_t \boldsymbol{X}_t$$

  where $\hat{\alpha}$ and $\hat{\beta}$ are the OLS estimates of parameter α and vector of parameters β, respectively

- The same predictive regression model is used to forecast the *BRP*

We consider four cases when running model (1):

- ***X*** includes **one original predictor**, *i.e.* we run bi-variate regressions **using one original predictor at a time**. Model *single_ts*
- ***X*** includes **all original predictors**, *i.e.* we run multi-variate regressions **using several original predictors**. Model *multi_ts*
- ***X*** includes **the frequencies of one original predictor**, *i.e.* we run multi-variate regressions **using different frequencies of one original predictor at a time**. Model *single_wav*
- ***X*** includes **the frequencies of the original predictors**, *i.e.* we run multi-variate regressions **using several frequencies of different original predictors**. Model *multi_wav*

The comparison of:

- models *ts* and *wav* informs about the **value of using more granular data** from the frequency decomposition of the original predictors
- models *single* and *multi* informs about the **usefulness of using information from different original predictors**

Out-of-sample (OOS) forecasts

For the *ERP* (the same for the *BRP*):

- One-step ahead forecast error: $e_{t+1} = ERP_{t+1} - E\hat{R}P_{t+1}$
- Squared forecast error: $e_{t+1}^2 = \left(ERP_{t+1} - E\hat{R}P_{t+1}\right)^2$
- Mean squared forecast error (MSFE): $\frac{1}{T-t_0}\sum_{t=t_0}^{T-1} e_{t+1}^2$

Out-of-sample (OOS) forecasts

- The forecast performance is evaluated using the Campbell and Thompson (RFS 2008) OOS $R^2 (R^2_{OS})$

- The R^2_{OS} measures the proportional reduction in the MSFE for the predictive model relative to the historical mean (HM). For the *ERP* (the same for the *BRP*):

$$R^2_{OS} = 100 \left[1 - \frac{\sum_{t=t_0}^{T-1} \left(ERP_{t+1} - E\hat{R}P_{t+1} \right)^2}{\sum_{t=t_0}^{T-1} \left(ERP_{t+1} - \overline{ERP}_t \right)^2} \right]$$

- A positive (negative) R^2_{OS} indicates that the predictive model outperforms (underperforms) the HM

R^2_{OS}: TIME SERIES

	single_ts		*multi_ts*	
	R^2_{OS}	Predictor	R^2_{OS}	Predictors
BRP	1.70**	TMS	3.40***	DP, DY, TMS
ERP	-0.29	LTR	-0.29	LTR

** and *** denote statistical significance at the 5% and 1% level, respectively, computed using the Clark and West (2007) statistic.

R^2_{OS}: TIME SERIES VS WAVELETS

- For the *BRP* and *ERP*, using several frequencies of one original predictor at a time (*single_wav*):

	single_ts		***single_wav***	
	R^2_{OS}	Predictor	R^2_{OS}	Predictor (frequency)
BRP	1.70**	TMS	5.45***	BM (D_1, D_2, D_5, D_7)
ERP	-0.29	LTR	1.77**	INFL (D_2, D_5)

** and *** denote statistical significance at the 5% and 1% level, respectively, computed using the Clark and West (2007) statistic.

R^2_{OS}: TIME SERIES VS WAVELETS

- For the *BRP* and *ERP*, using several frequencies of different original predictors (*multi_wav*):

	multi_ts		***multi_wav***	
	R^2_{OS}	Predictors	R^2_{OS}	Predictors (frequency)
BRP	3.40***	DP, DY, TMS	7.20***	BM (D_2), NTIS (D_1), TMS (D_5)
ERP	-0.29	LTR	3.97***	EP (D_3), RVOL (D_5), TMS (D_7)

** and *** denote statistical significance at the 5% and 1% level, respectively, computed using the Clark and West (2007) statistic.

R^2_{OS}: TIME SERIES VS WAVELETS

- Statistical performance: wrap-up

	single_ts	*multi_ts*	*single_wav*	*multi_wav*
	R^2_{OS}	R^2_{OS}	R^2_{OS}	R^2_{OS}
BRP	1.70**	3.40***	5.45***	7.20***
ERP	-0.29	-0.29	1.77**	3.97***

** and *** denote statistical significance at the 5% and 1% level, respectively, computed using the Clark and West (2007) statistic.

- Mean-variance investor who dynamically allocates his(er) wealth between bonds and equities. Weights ϖ_b and ϖ_e, respectively. Vector $\boldsymbol{\varpi} = (\varpi_b, \varpi_e)$
- The optimization problem backing the dynamic asset allocation (DAA) is:

$$\min_{\boldsymbol{\varpi}}[\gamma \Theta_P(\boldsymbol{\varpi}) - \boldsymbol{\varpi}' \hat{\boldsymbol{R}}] \tag{2}$$

where
  - $\gamma = 2$ is the investor's risk aversion coefficient
  - $\hat{\boldsymbol{R}} = \left(\hat{R}_{b,t+1}, \hat{R}_{e,t+1}\right)$: 1-step ahead return forecasts of bonds $(\hat{R}_{b,t+1})$ and equities $(\hat{R}_{e,t+1})$
  - $\Theta_P(\boldsymbol{\varpi}) = \sqrt{\boldsymbol{\varpi}' \hat{\boldsymbol{\Sigma}} \boldsymbol{\varpi}}$ is the portfolio risk function. $\hat{\boldsymbol{\Sigma}}$ is the estimated monthly returns covariance matrix

- We estimate $\boldsymbol{\Sigma}$ using an exponentially weighted moving average approach (decay parameter $= 0.97$)
- Portfolio constraints
  - Upper bound to the sum of the portfolio weights, $\boldsymbol{\varpi}' \boldsymbol{I}_2 = h$. h denotes the maximum leverage. We set $h = 1.5$
  - Lower bound l to the weight of each asset, $w_i \geq l$, $i = bond, equity$. We set $l = 0$
- The active portfolio return at $t+1$, $R_{p,t+1}$, is given by

$$R_{p,t+1} = \hat{\boldsymbol{\varpi}}'_t \boldsymbol{R}_{t+1} + \left(1 - \hat{\boldsymbol{\varpi}}'_t \mathbf{1}_2\right) r_f$$

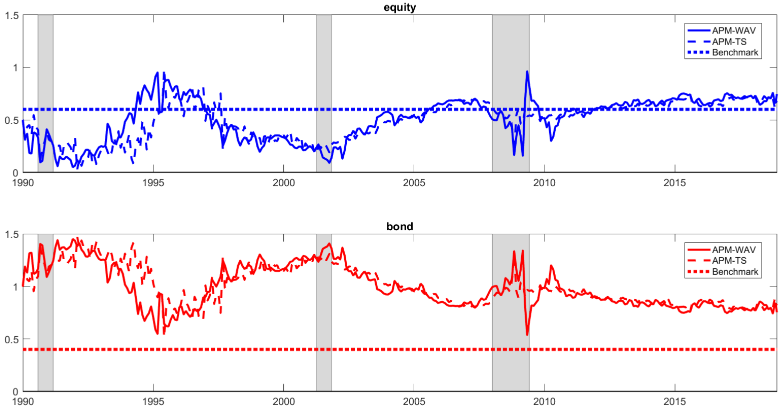
Economic analysis

Three strategies under analysis

- *APM_WAV*: uses best *BRP* and *ERP* forecasts using frequency-domain information (*multi_wav*) to feed portfolio-optimization
- *APM_TS*: uses best *BRP* and *ERP* forecasts using original time-series information (*multi_ts*) to feed portfolio-optimization
- $Benchmark_{60-40}$: conventional allocation of 60% to stocks and 40% to bonds as the benchmark

Economic analysis

Portfolio weights

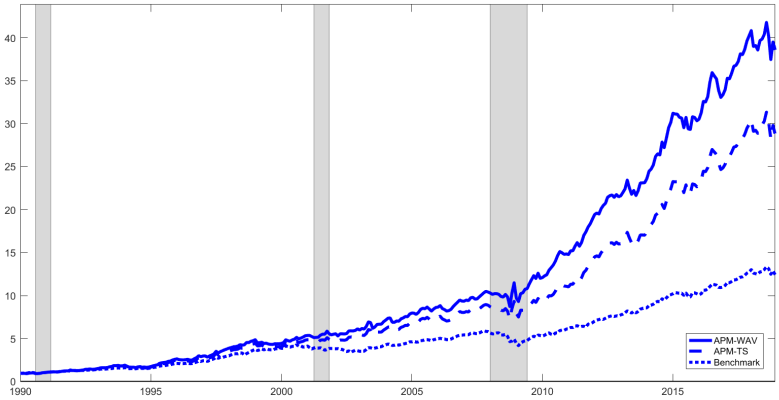


- Performance is evaluated using standing portfolio performance metrics and utility analysis (CER gain)

	Average return	CAGR	Sharpe ratio	Maximum drawdown	Information ratio	CER gain
APM_WAV	14.2 %	13.4 %	1.28	19.1 %	0.57	1.12
APM_TS	13.0 %	12.3 %	1.18	19.6 %	0.45	0.81
$Benchmark_{60-40}$	9.5 %	9.1 %	1.13	29.1 %	-	-

Economic analysis

Cumulative wealth over the OOS period



Robustness analysis

Alternative portfolio-optimization framework

- It is known that the mean-variance optimization has high input-sensitivity and produces extreme weights that fluctuate substantially over time
- An alternative optimization setting used by practitioners is the Black-Litterman model
  - We consider a modified version proposed by Da Silva, Lee, and Pornrojnangkool (JPM 2009) and Almadi, Rapach, and Suri (JPM 2014)
- Results are qualitatively similar to those in the mean-variance setting

To wrap up

- Fama and French (JFE 1989) find that different financial variables are useful to predict equity returns as they track different frequency components of the equity premium

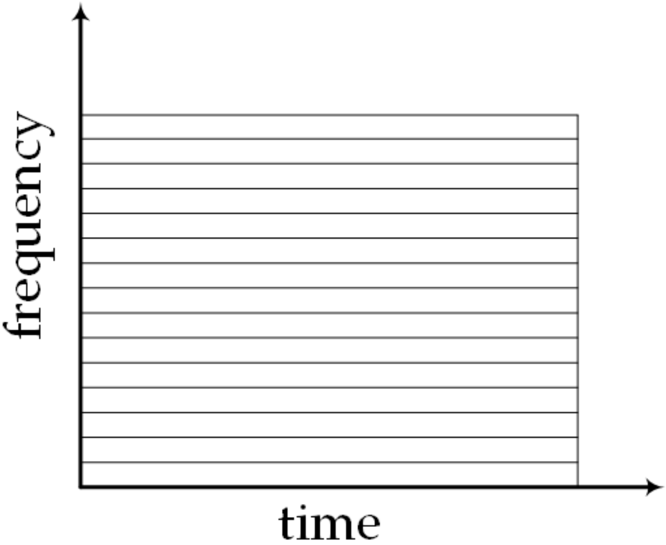
- We show that using information from different frequencies of different variables helps to improve forecasts of bond and equity returns

- **Main conclusion**: when used in the context of active portfolio management, these forecasts made with frequency-domain information lead to superior portfolio performances by several measures

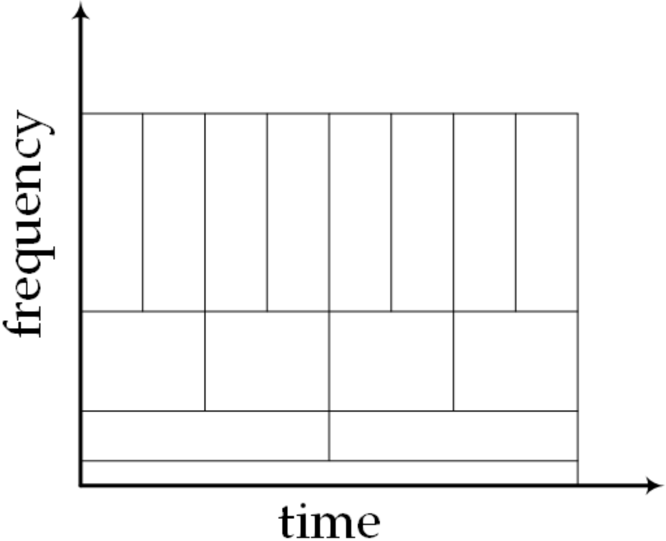
Extra slides

Frequency decomposition

Fourier



Frequency decomposition
MODWT



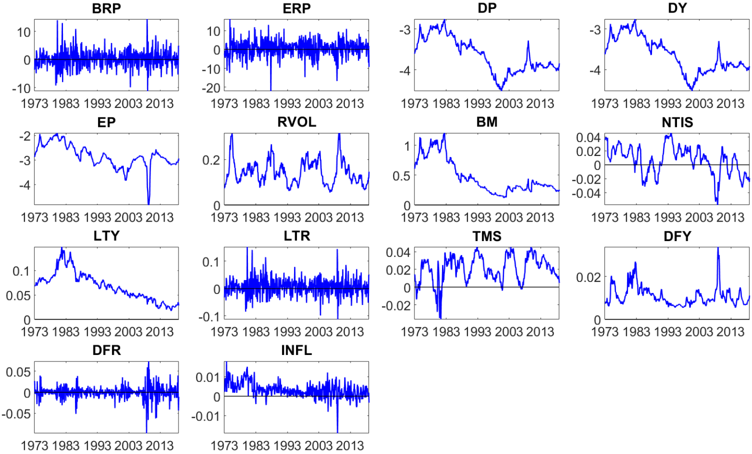
Frequency decomposition

Band-pass filter vs MODWT

- The Baxter and King (REStat 99) band-pass filter is a combination of a moving average in the time domain with a Fourier decomposition in the frequency domain optimized by minimizing the distance between the Fourier transform and an ideal filter
  - It applies a kind of optimal Fourier filtering on a sliding window (in the time domain), keeping the size of the window constant
  - Similar to the so-called short-time Fourier transform (also known as Gabor or windowed Fourier transform)

- The MODWT automatically adjusts the size of the window according to the frequency

Time series - BRP, ERP, and predictors



Certainty equivalent return (CER)

- The CER of a power-utility investor who has access to the *DAA* and benchmark portfolios is given by:

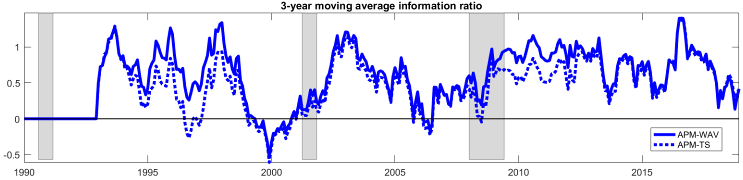
$$CER_j = \left[(1-\gamma)\,\overline{U}_j\right]^{1/(1-\gamma)} - 1\,,\, j = DAA, benchmark$$

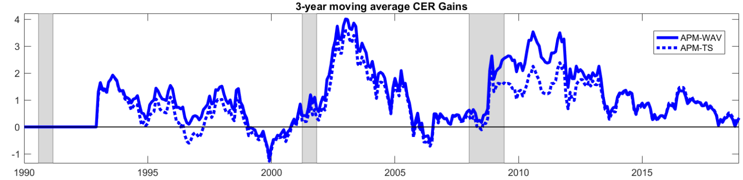
  where $\overline{U}_j\,, j = DAA, benchmark$ denote the average utility of an investor and $U(x) = \left[\frac{1}{(1-\gamma)}\right] x^{(1-y)}, x = 1 + R_p$ and R_p is the portfolio return

- We report the annualized utility gain, computed as $12 \cdot (CER_{DAA} - CER_{benchmark})$

- This difference can be understood as the annual portfolio management fee that an investor would accept to pay to have access to the active portfolio versus the benchmark portfolio

3-year moving average IR and CER gain





Robustness analysis

Alternative portfolio-optimization framework

- It is known that the mean-variance optimization has high input-sensitivity and produces extreme weights that fluctuate substantially over time
- An alternative optimization setting used by practitioners is the Black-Litterman model
  - We consider a modified version proposed by Da Silva, Lee, and Pornrojnangkool (JPM 2009) and Almadi, Rapach, and Suri (JPM 2014)
- The Black-Litterman model (BLM) is useful within a context of active portfolio management: the final purpose is to outperform the benchmark within a certain tracking error

  $\Rightarrow$ Maximize the information ratio and not the Sharpe ratio

- Power-utility maximizing investor with $\gamma = 2$ and the $Benchmark_{60-40}$
- We assume the investor will neither leverage nor short sell available assets ($h = 1$ and $l = 0$)
- *APM_BLMWAV*: active strategy based on asset return forecasts from wavelet-based methodologies used in the context of BLM
- *APM_BLMTS*: active strategy based on asset return forecasts from original time series methodologies used in the context of BLM

Robustness analysis

Portfolio performance statistics

	Average return	CAGR	Sharpe ratio	Maximum drawdown	Information ratio	CER gain
APM_BLMWAV	12.4 %	11.7 %	1.23	24.2 %	0.44	0.68
APM_BLMTS	10.2 %	9.6 %	1.05	26.1 %	0.11	0.12
$Benchmark_{60-40}$	9.5 %	9.1 %	1.13	29.1 %	-	-

Robustness analysis

Additional tests

- **Alternative benchmarks**
  - Naive diversification rule 1/N (50% equity and 50% bond)
  - Allocation of 40% equity and 60% bonds
- **Alternative set of portfolio constraints**
  - No leverage and short-selling possibilities
  - No leverage possibilities but with allowed short-selling
  - Both leverage and short-selling possibilities
- **Main result:** outperformance of the APM_WAV strategy versus the APM_TS strategy (and versus the $Benchmark_{60-40}$) **persists**