Out-of-sample time-frequency predictability of the equity premium*

Gonçalo Faria[†] Fabio Verona[‡]

Abstract

We propose a wavelet-based method to forecast the equity premium. The time series are decomposed into their frequency components, forecasted separately, and then summed to obtain the forecast of the equity premium. By extracting the relevant frequencies for equity premium forecasting purposes, this method significantly improves (in a statistical and economic way) upon traditional time series forecasting methods. This outperformance is robust regardless of the predictor used, the out-of-sample period considered, and the frequency of the data used.

Keywords: time-frequency predictability, equity premium, multiresolution analysis, discrete wavelet transform

JEL classification: C58, G11, G17

* This is a revised version of the paper that circulated as "Forecasting the equity risk premium with frequency-decomposed predictors". The authors thank Amit Goyal for providing data on his webpage; Gene Ambrocio, Theo Berger (discussant), Patrick Crowley, Eleonora Granziera, Mikael Juselius, Petri Jylhä, Francesco Ravazzolo, António Rua, Pedro Duarte Silva, Goufu Zhou as well as the participants at several conferences and seminars for useful comments; and Riccardo Verona for IT support. The views expressed are those of the authors and do not necessarily reflect those of the Bank of Finland. Faria gratefully acknowledges financial support from Fundação para a Ciência e Tecnologia through project UID/GES/00731/2016.

[†] Universidade Católica Portuguesa, Católica Porto Business School and CEGE, and University of Vigo – *RGEA* (gfaria@porto.ucp.pt)

[‡] Bank of Finland – Monetary Policy and Research Department, and University of Porto – *CEF.UP* (fabio.verona@bof.fi)

1 Introduction

Predicting stock market returns or the equity premium dates back almost a century to the work of Dow (1920). Despite this long tradition, the out-of-sample (OOS) empirical results – the relevant metric from the perspective of an investor – have been somewhat disappointing. In fact, for a large set of variables, predictability is either completely absent (Goyal and Welch, 2008) or concentrated in specific periods (Neely et al., 2014). Notwithstanding, in recent years there is increasing evidence that stock returns are, at least to some extent, predictable when using more advanced econometric techniques and novel macroeconomic and financial variables.

The interest on the spectral properties of financial asset returns (Dew-Becker and Giglio, 2016 and Chaudhuri and Lo, 2016) and of equity returns predictability (Bandi et al., 2018 and Faria and Verona, 2018) has recently increased. Along these lines, this paper contributes to the literature on the OOS predictability of the equity premium by proposing a method that exploits the time-frequency relationship between the equity premium and fourteen well-known variables from the equity premium predictability literature. Concretely, in the spirit of the trend-cycle decomposition of a time series (as proposed by *e.g.* Watson, 1986), we first use multiresolution wavelet analysis to decompose the time series of the equity premium and of its predictors into n time series components, each of them capturing the oscillations of the original variable within a specific frequency interval. A key feature of this decomposition is that, by adding those n time-frequency series components, the original time series is recovered. Then, considering one predicting variable at a time, we forecast separately each of the n time-frequency series component of the equity premium using the corresponding component of the predictor. At last, we aggregate those n forecasts to produce the forecast of the equity premium based on that predictor.

The main results of this paper can be summarized as follows. First, for *all* predictors considered, by selecting the proper time-frequency series components, the OOS forecasting performance is noticeably better than that using traditional time series forecasting methods. Furthermore, five variables (the earnings-price ratio, the dividend-payout ratio, the inflation rate, the long-term government bond return and the term spread) deliver positive and statistically significant OOS R-squares (R^2_{OS}), that is they outperform the historical mean (HM) of returns, which is the standard benchmark in the literature. This result thus unveils that some of the variables considered to be poor equity premium predictors are good predictors once the noisy components embedded in their original time series are filtered out and only the relevant frequencies for equity premium forecasting purposes are retained. Second, from an economic point of view, there are significant utility gains when making the forecasts using the proper time-frequency components of each predictor. Third, with the proposed method some predictors outperform the HM benchmark *also* during normal and good economic periods, which is typically not the case in the time series analysis.

The rest of the paper is organized as follows. In section 2 we review related literature to provide context for our contribution. Section 3 presents the data and the methodology. Section 4 presents the OOS results and section 5 the results of the robustness exercises. Finally, section 6 concludes.

2 Related literature

Multiresolution wavelet analysis allows to decompose any variable (regardless of its time series properties) into a trend, a cycle, and a noise component in a way which is similar to the traditional time series trend-cycle decomposition approach (Watson, 1986) or other filtering methods like the Hodrick and Prescott (1997) or the Baxter and King (1999) bandpass filter. In particular, the wavelet multiresolution decomposition allows to separately forecast each component of the time series in order to improve the forecast accuracy of the series as a whole.[1] As wavelet methods allow for a granular decomposition of a time series, they could in principle help making better forecast than traditional time series forecasting methods.

Wavelet-based forecasting methods have indeed been successfully used to forecast OOS economic and financial variables.[2] As regards forecasting economic variables, Rua (2011, 2017) proposes a wavelet-based multiscale principal component analysis to forecast GDP growth and inflation, while Kilponen and Verona (2016) forecast aggregate investment using the Tobin's Q theory of investment. As regards forecasting financial variables, Mitra and Mitra (2006) forecast exchange rates, while Zhang et al. (2017) and Faria and Verona (2018) focus on stock return predictability. In particular, Faria and Verona (2018) propose the SOPWAV method, which is a time-frequency forecast of stock market returns in the context of Ferreira and Santa-Clara (2011) sum-of-the-part method. In this paper we follow the idea of Faria and Verona (2018) but we run the time-frequency forecast of the equity premium within a more general OOS predictive setting instead of the sum-of-the-part method.

This paper is naturally related to the literature on the OOS forecasting of the equity premium, which was stimulated by Goyal and Welch (2008) findings that several equity premium predicting variables perform very poorly OOS. In the context of a single equation bivariate regression setup, recent methodological contributions that improved the OOS forecastability of the equity premium include regressions with time-varying coefficients (Dangl and Halling, 2012), with learning and time-varying volatility (Johannes et al., 2014), with economic constraints (Pettenuzzo et al., 2014), and with single predictor quantile combina-

[1] A related strand of literature debates whether it is better to forecast an aggregate variable directly (like GDP or inflation) or forecasting its components (like consumption, investment and government consumption, or each component of inflation) and then summing the component forecasts. Although there is not a full consensus, the empirical work on forecast aggregation is broadly supportive of the idea that aggregating forecasts can lead to improvements in accuracy when forecasting inflation and output (see Bermingham and D'Agostino, 2014 for a review).

[2] In the econometric literature, the recognition of frequency-specific modeling dates back at least to the work by Grether and Nerlove (1970) and to band-spectrum regression of Engle (1974). Crowley (2007) and Aguiar-Conraria and Soares (2014) provide excellent reviews of economic and finance applications of wavelets tools.

tion (Meligkotsidou et al., 2014). This paper contributes to this literature by proposing a different method for efficiently using the information embedded and aggregated in the time series of each individual variable.[3] In particular, after running the OOS equity premium forecast on a frequency-by-frequency basis, we show that statistically and economically OOS gains can be obtained by removing some frequencies (for each individual predictor) from the forecasting exercise.

Interestingly, we find that, for all predictors under analysis, their lowest frequency components are always selected as a relevant frequency for equity premium forecasting purposes. In some cases, *e.g.* the term spread, it is even the only relevant frequency. This finding adds to recent empirical evidence that the level and price of aggregate risk in equity markets are strongly linked to low-frequency economic fluctuations (see *e.g.* Dew-Becker and Giglio, 2016, Bianchi et al., 2017 and Gallegati and delli Gatti, 2018).

3 Data and methodology

We focus on the OOS predictability of monthly equity premium, measured by the difference between the log (total) return of the S&P500 index and the log return on a one-month Treasury bill. As it has been emphasized in the literature (*e.g.* Goyal and Welch, 2008 and Huang et al., 2015), the OOS exercise is more relevant to evaluate effective return predictability in real time while avoiding the in-sample over-fitting issue, eventual small-sample size distortions and the look-ahead bias concern. Moreover, we only focus on the one-month forecasting period as it has been documented that return predictability with a

[3] Methodological contributions that make use of several predictors to forecast the equity premium include dynamic factor models (Ludvigson and Ng, 2007, Kelly and Pruitt, 2013 and Neely et al., 2014), forecasts combination from different predictors (Rapach et al., 2010 and Pettenuzzo and Ravazzolo, 2016), regime-switching vector autoregression models (Henkel et al., 2011), the sum-of-the-parts method (Ferreira and Santa-Clara, 2011 and Faria and Verona, 2018), and Bayesian regime-switching combination or quantile combination approach (Zhu and Zhu, 2013 and Lima and Meng, 2017, respectively).

short horizon is usually magnified at longer horizons (see *e.g.* Cochrane, 2001).

We use monthly data from January 1973 to December 2016 for fourteen predictors from Goyal and Welch (2008) updated database. Specifically, we use the log dividend-price ratio (DP), the log dividend yield (DY), the log earnings-price ratio (EP), the log dividend-payout ratio (DE), the excess stock return volatility (RVOL), the book-to-market ratio (BM), the net equity expansion (NTIS), the Treasury bill rate (TBL), the long-term bond yield (LTY), the long-term bond return (LTR), the term spread (TMS), the default yield spread (DFY), the default return spread (DFR) and the lagged inflation rate (INFL). In appendix 1 these predictors are briefly explained and their time series are plotted. Table 1 reports summary statistics for the equity premium and its predictors. The average monthly equity premium is 0.41%, which, together with a monthly standard deviation of 4.44%, corresponds to an average monthly Sharpe ratio of 0.09 in the sample period.

Our methodology to forecast the equity premium is based on the wavelet multiresolution analysis, which is described in sub-section 3.1. The OOS procedure is then explained in sub-section 3.2.

3.1 Wavelet multiresolution analysis

Wavelets are signal processing techniques that were developed to overcome some of the limitations of traditional frequency domain tools (spectral analysis and Fourier transforms), as they provide a more complete decomposition of the original time series without suffering their weaknesses. For instance, and differently from the Fourier analysis, wavelets are defined over a finite window in the time domain, with the size of that window being adjusted automatically according to the frequency of interest. This means that the high-frequency features of the time series can be captured by using a short window, whereas by looking at the same signal with a large window, the low-frequency features are revealed. Hence, wavelets allow to extract both time-varying and frequency-varying features simultaneously just by changing the size of the window. They are thus better to handle variables (like *e.g.* financial variables) that exhibit jumps, structural breaks, and time-varying volatility.

The wavelet multiresolution analysis (MRA) allows the decomposition of a time series into its constituent multiresolution (frequency) components.[4] Given a time series y_t, its wavelet multiresolution representation can be written as

$$y_t = y_t^{S_J} + y_t^{D_J} + y_t^{D_{J-1}} + \ldots + y_t^{D_1} \quad , \tag{1}$$

where $y_t^{S_J}$ is the wavelet smooth component and $y_t^{D_j}$, $j = 1, 2, \ldots, J$, are the J wavelet detail components. Equation (1) shows that the original series y_t, exclusively defined in the time domain, can be decomposed in different components, each also defined in the time domain and representing the fluctuation of the original time series in a specific frequency band. In particular, for small j, the j wavelet detail components represent the higher frequency characteristics of the time series (*i.e.* its short-term dynamics). As j increases, the j wavelet detail components represent lower frequencies movements of the series. Finally, the wavelet smooth component captures the lowest frequency dynamics (*i.e.* its long-term behavior).

In this paper, we use the maximal overlap discrete wavelet transform (MODWT) MRA with the Haar wavelet filter and reflecting boundary conditions.[5] Given the sufficiently long data series, we apply a J=6 level MRA so that the decomposition delivers seven time-frequency series: six wavelet detail components ($y_t^{D_1}$ to $y_t^{D_6}$) and the wavelet smooth component ($y_t^{S_6}$).[6]

[4] In this section, we closely follow the presentation of the wavelet multiresolution approach done in Faria and Verona (2018) and limit the description to the basic concepts which are directly useful to understand our empirical analysis. A more detailed analysis of wavelets methods can be found in Percival and Walden (2000) and in appendix 2.

[5] Examples of papers using the MODWT MRA decomposition include Galagedera and Maharaj (2008), Xue et al. (2013), Bekiros and Marcellino (2013), Barunik and Vacha (2015), Caraiani (2015), Bekiros et al. (2016), Berger (2016), Zhang et al. (2017) and Faria and Verona (2018). While the Haar filter is simple and widely used (see *e.g.* Manchaldore et al., 2010, Bandi et al., 2018 and Faria and Verona, 2018), the results in this paper are qualitatively the same using other wavelet filters (like *e.g.* Daubechies).

[6] As regards the choice of J, the number of observations dictates the maximum number of frequency bands

As we use monthly data, the first detail component $y_t^{D_1}$ captures oscillations between 2 and 4 months, while detail components $y_t^{D_2}$, $y_t^{D_3}$, $y_t^{D_4}$, $y_t^{D_5}$ and $y_t^{D_6}$ capture oscillations with a period of 4-8, 8-16, 16-32, 32-64 and 64-128 months, respectively. Finally, the smooth component $y_t^{S_6}$, which in what follows we re-denote $y_t^{D_7}$, captures oscillations with a period longer than 128 months (10.6 years).[7]

To illustrate the rich set of different dynamics aggregated (and therefore hidden) in the original time series, figure 1 plots the time series of the (log) equity premium (top left panel) and of its seven time-frequency series components (remaining panels). As expected, the lower the frequency, the smoother the resulting filtered time series.

Furthermore, wavelets allow to analyze the variability of a time series on a frequency-by-frequency basis. In particular, by running the so-called energy decomposition analysis, it is possible to compute the variance decomposition by frequency and, hence, to detect which frequency bands contribute relatively more to the overall volatility of the original time series. Table 2 reports the results of the energy decomposition analysis for the variables under analysis. For the variables with low persistence (equity premium, LTR and DFR), most of the volatility (more than 70%) is concentrated at higher frequencies (D_1 and D_2), whereas for the more persistent variables the lowest frequencies components (D_5 and above) account for the majority of the total variability of the series.

3.2 Out-of-sample forecasts

The one-step ahead OOS forecasts are generated using a sequence of expanding windows. We use an initial in-sample period (1973:01 to 1989:12) to make the first one-step ahead

that can be used. In particular, if t_0 is the number of observations in the in-sample period, then J has to satisfy the constraint $J \leq \log_2 t_0$.

[7] In the MODWT, each wavelet filter at frequency j approximates an ideal high-pass filter with passband $f \in \left[1/2^{j+1}, 1/2^{j}\right]$, while the smooth component is associated with frequencies $f \in \left[0, 1/2^{j+1}\right]$. The level j wavelet components are thus associated to fluctuations with periodicity $\left[2^{j}, 2^{j+1}\right]$ (months, in our case).

OOS forecast. The in-sample period is then increased by one observation and a new one-step ahead OOS forecast is produced. This is the procedure until the end of the sample. The full OOS period therefore spans from 1990:01 to 2016:12.

3.2.1 Predictive regression model: time series

Let r be the equity premium. For each individual predictor x_i, $i = 1, ..., 14$, the predictive regression model is

$$r_{t+1} = \alpha + \beta x_{i,t} + \varepsilon_{t+1} \ , \tag{2}$$

and the one-step ahead OOS forecast of the equity premium, $\hat{r}_{t+1}$, is given by:

$$\hat{r}_{t+1} = \hat{\alpha}_t + \hat{\beta}_t x_{i,t} \ , \tag{3}$$

where $\hat{\alpha}_t$ and $\hat{\beta}_t$ are the OLS estimates of α and β in equation (2), respectively, using data from the beginning of the sample until month t. We denote this forecast as the TS (time series) forecast.

3.2.2 Wavelet-based forecasting

To forecast with wavelets, we follow the method suggested by Rua (2011) and Faria and Verona (2018). In a nutshell, we fit a model like (2) to each time-frequency component of the MODWT MRA decomposition of r and x_i. The overall forecast for r can then be obtained by aggregating the forecasts of its time-frequency components.[8] Importantly, as the MODWT MRA at a given point in time uses information of neighboring data points (both past and

[8] This is also the spirit of the scale predictability in Bandi et al. (2018), who explore a model where returns and predictors are linear aggregates of components operating over different frequencies, and where predictability is frequency-specific. Similarly, by using a scale-by-scale selection of the series, Gallegati (2014), Voutilainen (2017), and Gallegati and delli Gatti (2018) construct a composite wavelet-based leading indicator of the business cycle, a proxy for the financial cycle, and a global crisis index, respectively.

future), we recompute the time-frequency series components at each iteration of the OOS forecasting process. This ensures that our method does not suffer from any look-ahead bias as the forecasts are made with current and past information only.

Let us explain in more detail the steps involved. Firstly, we apply the MODWT MRA decomposition to the variable to be forecasted (r) as well as to all predictors (x_i). Second, for each predictor x_i, we estimate a model like (2) for each frequency level $j = 1, ..., 7$. That is, we estimate – separately – each time-frequency component of the equity premium using the time-frequency component of the predictor at the same level j:[9]

$$r_{t+1}^{x_i,D_j} = \alpha_{t,j}^{x_i} + \beta_{t,j}^{x_i} x_{i,t}^{D_j} + \varepsilon_{t+1} \quad . \tag{4}$$

Third, we use the estimation results to produce the one-step ahead forecast of the corresponding time-frequency component of r:

$$\hat{r}_{t+1}^{x_i,D_j} = \hat{\alpha}_{t,j}^{x_i} + \hat{\beta}_{t,j}^{x_i} x_{i,t}^{D_j} \quad ,$$

where $\hat{\alpha}_{t,j}^{x_i}$ and $\hat{\beta}_{t,j}^{x_i}$ are the OLS estimates of $\alpha_{t,j}^{x_i}$ and $\beta_{t,j}^{x_i}$ in equation (4), respectively, using data from the beginning of the sample until month t. Fourth, the one-step ahead forecast for r using predictor x_i is obtained by summing those 7 forecasts ($j = 1, ..., 7$). As an example, the one-step ahead wavelet-based forecast of the equity premium using the dividend-payout ratio (DE) as a predictor, $\hat{r}_{t+1}^{DE}$, is given by:

$$\hat{r}_{t+1}^{DE} = \sum_{j=1}^{J+1} \hat{r}_{t+1}^{DE,D_j} = \sum_{j=1}^{J+1} \left(\hat{\alpha}_{t,j}^{DE} + \hat{\beta}_{t,j}^{DE} DE_t^{D_j} \right) \quad , \tag{5}$$

where $DE_t^{D_j}$, $j = 1, ..., 7$, are the time-frequency series components of DE. As in (5) we use

[9] In principle it is possible to fit different forecasting models for each frequency components. For instance, we could include more lags of the predictor and of the equity premium when forecasting the lowest frequency components of the latter. We leave this for future research.

all the frequency components to make the forecast of the equity premium, we denote this specification as WAV_ALL. After running the forecast with the forecasting model (5), we find that it does not outperform the HM benchmark. Our conjecture is that, by considering all frequency components, we are also including the more noisy ones, which in turn make the forecasting exercise too imprecise.

Hence, to improve the forecast, we exploit the flexibility and granularity of this method and propose a new way of improving the equity premium forecast. Namely, for each individual predictor, we search for the combination of its time-frequency series components that maximizes the Campbell and Thompson (2008) R^2_{OS} statistic (as explained in sub-section 3.2.3). Taking again DE as an example, the equity premium wavelet-based forecasting econometric model is given by:

$$\hat{r}^{DE}_{t+1} = \sum_{j=1}^{J+1} \delta_j \hat{r}^{DE,D_j}_{t+1} = \sum_{j=1}^{J+1} \delta_j \left[\hat{\alpha}^{DE}_{t,j} + \hat{\beta}^{DE}_{t,j} DE^{D_j}_t \right] \quad . \tag{6}$$

For each predictor, the weights of each frequency component are chosen in order to maximize the predictor's statistical performance. For computational reasons, in (6) we only consider five possible values for each weight δ_j: 0, 0.25, 0.50, 0.75 and 1. A weight of 0 excludes a particular frequency from the forecast, thus allowing to completely remove the information carried by that frequency to the forecast exercise.Although the results are likely to improve by using a finer grid, the main message of this exercise would be the same. Notwithstanding, the grid used in this paper represents an improvement with respect to Faria and Verona (2018), who only consider two possible weights, either inclusion ($\delta_j = 1$) or exclusion ($\delta_j = 0$) of a specific frequency.[10]

As we are interested in analyzing which frequencies of each predictor are, on average, relevant to forecast the equity premium, we use a time-invariant weighting scheme. However, several

[10] Similarly to Ludvigson and Ng (2007), our forecasting model selection is done on the basis of a search across several potential model specifications performed over the entire OOS period.

factors like *e.g.* market sentiment, monetary policies and uncertainty can motivate the use of time-varying schemes so that one can assess the importance of each frequency at each point in time. This is an interesting exercise that we leave for future research.

We denominate this specification as the WAV_BEST and should inform about the relevant frequencies of each predictor for the equity premium forecasting purposes.

3.2.3 Forecast evaluation

The forecasting performances of the time series (TS) and wavelet based (WAV) models are evaluated using the Campbell and Thompson (2008) R^2_{OS} statistic. As standard in the literature, the benchmark model is the prevailing mean forecast $\overline{r}_t$, which is the average excess return up to time t. The R^2_{OS} statistic measures the proportional reduction in the mean squared forecast error for the predictive model ($MSFE_{PRED}$) relative to the historical mean ($MSFE_{HM}$) and is given by

$$R^2_{OS} = 100\left(1 - \frac{MSFE_{PRED}}{MSFE_{HM}}\right) = 100\left[1 - \frac{\sum_{t=t_0}^{T-1}\left(r_{t+1} - \hat{r}_{t+1}\right)^2}{\sum_{t=t_0}^{T-1}\left(r_{t+1} - \overline{r}_t\right)^2}\right] \quad ,$$

where $\hat{r}_{t+1}$ is the excess return forecast for *t+1* from the TS or the WAV models considered and r_{t+1} is the realized stock market return from t to *t+1*. A positive (negative) R^2_{OS} indicates that the predictive model outperforms (underperforms) the HM in terms of MSFE.

The statistical significance of the results is evaluated using the Clark and West (2007) statistic. This statistic tests the null hypothesis that the MSFE of the HM model is less than or equal to the MSFE of the TS or specific WAV model against the alternative hypothesis that the MSFE of the HM model is greater than the MSFE of the TS or specific WAV model ($H_0 : R^2_{OS} \leq 0$ against $H_A : R^2_{OS} > 0$).

3.3 Asset allocation

We analyze the economic value of the different predictive models (TS and WAV) from an asset allocation perspective, considering a mean-variance investor who allocates his or her wealth between equities and risk-free bills. At the end of month t, the investor optimally allocates

$$w_t = \frac{1}{\gamma}\frac{\hat{R}_{t+1}}{\hat{\sigma}^2_{t+1}} \tag{7}$$

of the portfolio to equity for period *t+1*. In (7), γ is the investor's relative risk aversion coefficient, $\hat{R}_{t+1}$ is the time t (TS or WAV) model forecast of equity premium for *t+1*, and $\hat{\sigma}^2_{t+1}$ is the forecast of the variance of the equity premium. As in Rapach et al. (2016), we assume a relative risk aversion coefficient of three, use a ten-year moving window of past equity premium to estimate the variance forecast and constrain the weights w_t to lie between -0.5 and 1.5. These constraints limit the possibilities of short selling and leveraging the portfolio.

The realized portfolio return at time *t+1*, RP_{t+1}, is given by $RP_{t+1} = w_t R_{t+1} + RF_{t+1}$, where RF_{t+1} denotes the risk-free return from time t to *t+1* (*i.e.* the market rate, which is known at time t). The average utility (or certainty equivalent return, CER) of an investor that uses the portfolio rule (7) is given by $CER = \overline{RP} - 0.5\gamma\sigma^2_{RP}$, where $\overline{RP}$ and σ^2_{RP} are the sample mean and variance of the portfolio return, respectively. We report the annualized utility gain, which is computed as the difference between the CER for an investor that uses the TS/WAV model to forecast equity premium and the CER for an investor who uses the HM benchmark for forecasting. The difference is multiplied by 12, which allows to interpret it as the annual portfolio management fee that an investor would accept to pay to have access to the alternative forecasting model versus the historical average forecast.

4 Out-of-sample forecasting performance

4.1 Statistical performance

The second, fourth and fifth columns of table 3 report the R^2_{OS} statistics for each predictor using different model specifications versus the HM, for the entire OOS period (1990:01-2016:12).

The standard time series analysis (second column) confirms Goyal and Welch (2008) results, *i.e.* that traditional predictors perform badly OOS. As regards the wavelet-based forecasts, there is no value added by considering all frequencies (fourth column, WAV_ALL model). In fact, except for inflation (INFL), all the R^2_{OS}s are negative. This suggests that the forecasting exercise is too imprecise when considering the information from all frequencies. However, when the time-frequency series components are optimally chosen (fifth column, WAV_BEST), *all* R^2_{OS}s are higher than the respective TS R^2_{OS}. That is, the OOS forecasting performance of the WAV_BEST model is always better than that of the time series analysis. For some predictors their OOS performance is still not enough to outperform the HM benchmark ($R^2_{OS} < 0$). However, there are five variables for which the R^2_{OS}s are positive and statistically significant. This means that some of the equity premium predictors with reported poor performance in the literature have nevertheless predictability power, as long as their frequencies are properly chosen and used. Consider, for example, the dividend-payout ratio (DE) and inflation (INFL). In the time series analysis, their R^2_{OS}s are -2.17 and -0.69, respectively. However, using the WAV_BEST model, their R^2_{OS}s are 2.82 and 1.85, respectively, both statistically significant.

The weights of the frequency components ($\delta_1 - \delta_7$) are listed in the last seven columns of table 3. Regardless of the predictor used, the lowest frequency component is always included ($\delta_7 > 0$). This finding adds to recent empirical evidence that the level and price of aggregate risk in equity markets are strongly linked to low-frequency economic fluctuations (*e.g.* Dew-Becker and Giglio, 2016) and also that there are low-frequency, decades-long shifts in asset values relative to measures of macroeconomic fundamentals in the US (*e.g.* Bianchi et al. (2017)). For some predictors, it is also beneficial to include some high frequency fluctuations (δ_1), whereas intermediate frequencies (especially δ_3 and δ_4) are usually less important. Finally, considering the entire spectrum of predictors/frequencies, more than 50% of the frequencies have zero weight. This means that a lot of information needs to be removed from the predictive regressions in order to improve the forecastability of the equity premium.

To evaluate the consistency over time of the OOS performance of the forecasting model, we report the dynamics of the difference between the cumulative square forecasting error for the HM forecasting model and the cumulative square forecasting error when the TS or the WAV_BEST model for each predictor is used. Results, plotted in figure 2, should be read as follows. When the line increases/decreases, the predictive regression of the WAV model (in blue) or of the TS model (in black) outperforms/underperforms that of the HM. A forecasting model that consistently outperforms the HM will thus always have a positively sloped curve.

In the time series analysis (black lines), all predictors underperform the HM, so their corresponding lines are almost always below zero. Looking at the WAV_BEST models (blue lines), it is possible to broadly classify the predictors into four different groups as regards the consistency of their OOS performance over time. The first group includes predictors (DP, DY, NTIS, TBL and DFY) with an OOS performance close to that of the HM most of the time (*i.e.* the lines are relatively stable around zero). A second group includes predictors (RVOL, LTY and DFR) with an erratic forecasting performance, as the slopes of their plotted graphs swing between positive and negative values. A third group includes predictors (EP, DE, BM and INFL) which post a strong OOS outperformance versus the HM only during the last NBER-dated recession. Finally, two predictors (LTR and TMS) post a consistent positive outperformance throughout the entire OOS period (except for the first 5 years), with their corresponding lines featuring smooth upward-sloping trends.

4.2 Economic performance

In the previous sub-section we have shown that the proposed wavelet-based forecasting method delivers statistically significant gains. We now analize the performance of this method from an asset allocation perspective. Results are reported in the third and sixth columns of table 3.

The CER gains using the WAV_BEST forecasting method are usually larger than those in the TS analysis. Furthermore, 11 out of 14 predictors deliver positive CER gains with the WAV_BEST model, with the highest utility gains (548 basis points) obtained when using the term spread (TMS).

Figure 3 provides a dynamic perspective of the portfolio and cumulative wealth for an investor that uses the HM model, the WAV_BEST model for the dividend-payout ratio (DE, which obtains the highest R^2_{OS} in the OOS sample period under analysis) and the WAV_BEST model for the term spread (TMS, which obtains the highest CER gains in the OOS sample period under analysis).

Panel A presents the dynamic equity weights (constrained to lie between -0.5 and 1.5) for those three alternative portfolios. Two results stand out. First, the equity exposure of the HM portfolio (black line) is smoother than the alternative portfolios under analysis. Second, changes in the equity allocation in a portfolio based on the WAV_BEST TMS (blue line) are smoother than those on the WAV_BEST DE (red line). This can be explained by the fact that the WAV_BEST TMS only considers the lowest frequency (*i.e.* the long run) of the TMS, while the WAV_BEST DE considers both higher and lower frequencies of the DE.

Panel B shows the log cumulative wealth for an investor that invests 1$ in January 1990 and reinvests all proceeds. Both strategies based on the WAV_BEST models clearly outperform the strategy based on the HM, with that outperformance being particularly strong during recession periods. This is essentially due to the improved market timing of both WAV_BEST model based strategies versus the HM based strategy, as illustrated in Panel A.

5 Robustness tests

We run two tests to evaluate the robustness of the wavelet-based forecast methodology. We first analyze the forecasting performance in different sample periods (sub-section 5.1), and then run the forecasting exercise using quarterly data (sub-section 5.2).

5.1 Different sample periods

5.1.1 Great moderation and great financial crisis

We divide the OOS period into two sub-periods: from January 1990 to December 2006, which broadly corresponds to the so-called great moderation period, and from January 2007 onward, which corresponds to the great financial crisis and aftermath.

Table 4 reports the R^2_{OS} and the CER gains for all predictors. Regardless of the forecasting method used (TS or WAV_BEST), the OOS predictability in the first period is usually weaker than in the second period.[11] In any case, in both sample periods there are significant OOS forecasting improvements for almost all predictors using the WAV_BEST forecasting model. In the first period, five variables yield positive and statistically significant R^2_{OS}s using the WAV_BEST model, while in the time series analysis no predictor outperforms the

[11] Regarding the WAV_BEST model, for each predictor and for each sub-sample period, we use the same weights for the frequencies as the ones in the full OOS period (reported in table 3). This is a conservative approach, as we would expect to improve the performance of the WAV_BEST models by choosing the optimal weights of different frequencies for each predictor and for each sub-sample period.

HM benchmark in a statistically significant way. A similar pattern is visible in the second period. Interestingly, with the WAV_BEST model four predictors (earnings-price ratio, dividend-payout ratio, long-term return and term spread) outperform the HM benchmark in both sub-sample periods. Very similar conclusions arise from the utility gains analysis. The maximum CER gains obtained are 500 and 811 basis points in the first and second sub-sample periods (term spread and inflation with WAV_BEST model, respectively), which are significantly higher than the gains achieved in the time series analysis (115 and 83 basis points for Treasury bill rate and earnings-price ratio, respectively).

5.1.2 Bad, normal, and good growth periods

A typical finding in the equity premium forecasting literature is that there is no predictability during expansions or good times (see *e.g.* Henkel et al., 2011 and Neely et al., 2014). However, Dangl and Halling (2012) and Huang et al. (2016) find positive and statistically significant levels of OOS predictability during expansions using time-varying coefficients regression and state-dependent predictive regression models, respectively.

Accordingly, and following Rapach et al. (2010), we evaluate the individual forecasts during periods of bad, normal, and good economic growth. Those regimes are defined as the bottom, middle, and top third of sorted growth rates of industrial production in the US, respectively.[12] We report the R^2_{OS}s and the CER gains for each regime in table 5.

Looking at the R^2_{OS} during bad growth periods, no predictor is statistically significant in the time series analysis, whereas five predictors are statistically significant and obtain expressive CER gains when using the WAV_BEST models. In particular, the maximum R^2_{OS} and CER gains are 7.09% and 1011 basis points, respectively, both achieved using the dividend-payout ratio.

[12] The data for the industrial production in the US was downloaded from Federal Reserve Economic Data at http://research.stlouisfed.org/fred2/.

The same qualitative conclusions can be extended to the normal growth period, even if for this regime only two predictors are statistically significant using the WAV_BEST model. Although the maximum R^2_{OS}s and CER gains using the WAV_BEST model during normal periods are usually lower than during bad periods, the levels are still quite high: the maximum R^2_{OS} and CER gains are 2.73% and 453 basis points using the stock return volatility and the term spread, respectively.

As regards the good period regime, three predictors (EP, LTR, and TMS) are statistically significant when using the WAV_BEST models. Moreover, the OOS performance is rather good, as their R^2_{OS}s are 2.93%, 1.03% and 0.93%, respectively. From an utility perspective, results are also strong, as their annualized CER gains are 606, 278 and 503 basis points, respectively.

Overall, for some predictors the wavelet-based forecasting method allows to improve the OOS forecast performance also when splitting the OOS period in bad, normal, and good growth periods.

5.2 Quarterly data

At last, we test the robustness of the wavelet-based forecasting method using quarterly data. As before, the OOS forecasts are made using a sequence of expanding windows. To have a sufficiently large initial sample period, we use data from 1952:Q1. The initial in-sample period is 1952:Q1 to 1989:Q4, and the full OOS period spans from 1990:Q1 to 2016:Q4. As in the analysis using monthly data, we apply a J=6 level MRA so that the decomposition delivers seven time-frequency series. We perform the MODWT MRA using the Daubechies filter with length 8 and reflecting boundary conditions. We adopt this filter, instead of the Haar filter used with monthly data, as it is more suited for (and more commonly used with) quarterly data.

We consider seventeen predictors: the same fourteen predictors used in the monthly data analysis plus the (lagged) investment to capital ratio (IK), the consumption-wealth ratio (CAY), and Tobin's Q (Q).[13] Table 6 reports the R^2_{OS}s for each predictor for both the time series analysis and the WAV_BEST model specification. The main conclusion is that the wavelet-based forecasting method is robust towards the use of quarterly data. There are indeed significant OOS forecasting improvements for almost all predictors using the WAV_BEST forecasting model. Furthermore, four variables (earnings-price ratio, dividend-payout ratio, long-term yield and investment rate) yield positive and statistically significant R^2_{OS}s using the WAV_BEST model, while in the time series analysis no predictor outperforms the HM benchmark in a statistically significant way.

6 Concluding remarks

Goyal and Welch (2008) and subsequent research have documented the poor out-of-sample (OOS) equity premium forecasting performance of an extensive list of predictors. In this paper we propose a wavelet-based method to forecast the equity premium. The series are decomposed into their time-frequency components, forecasted separately, and then aggregated to obtain the forecast of the equity premium. Regardless of the predictor used, the OOS period, and the frequency of the data considered, this method significantly improves upon the OOS forecast done using traditional time series tools. The proposed wavelet-based method allows for a more granular analysis, leading to its strong and robust empirical performance. In particular, the crucial step to improve the forecasting performance of the predictors is to exclude the noisy components embedded in the original time series so as to only retain the

[13] The quarterly time series of the IK and the CAY are available from the Goyal and Welch (2008) updated database, while the series for Tobin's Q is computed using the Federal Flow of Funds data. These three variables, which are briefly explained in appendix 1, have usually been used as equity premium predictors when using quarterly data (see e.g. Rapach et al., 2010, Lettau and Ludvigson, 2001, 2002 and Bianchi et al., 2017).

frequencies that are relevant for the equity premium forecasting exercise.

The proposed wavelet-based forecasting method could, in principle, be helpful to improve the forecast of other financial variables (*e.g.* equity market returns variance) and returns in other markets such as fixed income, currency and commodities. Moreover, given the role that the equity premium forecast has in asset allocation decisions, the proposed method may bring relevant insights about the frequency-domain implications in the optimal dynamic asset allocation decisions. At last, the proposed forecasting framework can also be useful for policymakers in their attempt of anticipating possible "over-heated" equity markets that could, ultimately, pose a threat to macroeconomic and financial stability.

References

Aguiar-Conraria, L. and M. J. Soares: 2014, 'The Continuous Wavelet Transform: Moving Beyond Uni- And Bivariate Analysis'. *Journal of Economic Surveys* **28**(2), 344–375.

Bandi, F., B. Perron, A. Tamoni, and C. Tebaldi: 2018, 'The Scale of Predictability'. *Journal of Econometrics*. forthcoming.

Barunik, J. and L. Vacha: 2015, 'Realized wavelet-based estimation of integrated variance and jumps in the presence of noise'. *Quantitative Finance* **15**(8), 1347–1364.

Baxter, M. and R. King: 1999, 'Measuring Business Cycles: Approximate Band-Pass Filters For Economic Time Series'. *Review of Economics and Statistics* **81**(4), 575–593.

Bekiros, S. and M. Marcellino: 2013, 'The multiscale causal dynamics of foreign exchange markets'. *Journal of International Money and Finance* **33**(C), 282–305.

Bekiros, S., D. K. Nguyen, G. S. Uddin, and B. Sjo: 2016, 'On the time scale behavior of equity-commodity links: Implications for portfolio management'. *Journal of International Financial Markets, Institutions and Money* **41**(C), 30–46.

Berger, T.: 2016, 'Forecasting Based on Decomposed Financial Return Series: A Wavelet Analysis'. *Journal of Forecasting* **35**(5), 419–433.

Bermingham, C. and A. D'Agostino: 2014, 'Understanding and forecasting aggregate and disaggregate price dynamics'. *Empirical Economics* **46**(2), 765–788.

Bianchi, F., M. Lettau, and S. Ludvigson: 2017, 'Monetary Policy and Asset Valuation'. CEPR Discussion Papers 12275.

Campbell, J. Y. and S. B. Thompson: 2008, 'Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?'. *Review of Financial Studies* **21**(4), 1509–1531.

Caraiani, P.: 2015, 'Estimating DSGE models across time and frequency'. *Journal of Macroeconomics* **44**(C), 33–49.

Chaudhuri, S. and A. Lo: 2016, 'Spectral Portfolio Theory'. mimeo.

Clark, T. and K. West: 2007, 'Approximately normal tests for equal predictive accuracy in nested models'. *Journal of Econometrics* **138**(1), 291 – 311.

Cochrane, J.: 2001, *Asset Pricing*. Princeton University Press.

Crowley, P. M.: 2007, 'A Guide To Wavelets For Economists'. *Journal of Economic Surveys* **21**(2), 207–267.

Dangl, T. and M. Halling: 2012, 'Predictive regressions with time-varying coefficients'. *Journal of Financial Economics* **106**(1), 157–181.

Dew-Becker, I. and S. Giglio: 2016, 'Asset Pricing in the Frequency Domain: Theory and Empirics'. *Review of Financial Studies* **29**(8), 2029–2068.

Dow, C. H.: 1920, *Scientific Stock Speculation.* The Magazine of Wall Street.

Engle, R. F.: 1974, 'Band Spectrum Regression'. *International Economic Review* **15**(1), 1–11.

Faria, G. and F. Verona: 2018, 'Forecasting stock market returns by summing the frequency-decomposed parts'. *Journal of Empirical Finance* **45**, 228 – 242.

Ferreira, M. A. and P. Santa-Clara: 2011, 'Forecasting stock market returns: the sum of the parts is more than the whole'. *Journal of Financial Economics* **100**(3), 514–537.

Galagedera, D. and E. Maharaj: 2008, 'Wavelet timescales and conditional relationship between higher-order systematic co-moments and portfolio returns'. *Quantitative Finance* **8**(2), 201–215.

Gallegati, M.: 2014, 'Making leading indicators more leading: A wavelet-based method for the construction of composite leading indexes'. *OECD Journal: Journal of Business Cycle Measurement and Analysis* **2014**(1), 1–21.

Gallegati, M. and D. delli Gatti: 2018, 'Macrofinancial imbalances in historical perspective: A global crisis index'. *Journal of Economic Dynamics and Control* **xx**(x), yy–yy.

Goyal, A. and I. Welch: 2008, 'A Comprehensive Look at The Empirical Performance of Equity Premium Prediction'. *Review of Financial Studies* **21**(4), 1455–1508.

Grether, D. M. and M. Nerlove: 1970, 'Some Properties of 'Optimal' Seasonal Adjustment'. *Econometrica* **38**(5), 682–703.

Henkel, S. J., J. S. Martin, and F. Nardari: 2011, 'Time-varying short-horizon predictability'. *Journal of Financial Economics* **99**(3), 560–580.

Hodrick, R. J. and E. C. Prescott: 1997, 'Postwar U.S. Business Cycles: An Empirical Investigation'. *Journal of Money, Credit and Banking* **29**(1), 1–16.

Huang, D., F. Jiang, J. Tu, and G. Zhou: 2015, 'Investor Sentiment Aligned: A Powerful Predictor of Stock Returns'. *Review of Financial Studies* **28**(3), 791–837.

Huang, D., F. Jiang, J. Tu, and G. Zhou: 2016, 'Forecasting Stock Returns in Good and Bad Times: The Role of Market States'. mimeo.

Johannes, M., A. Korteweg, and N. Polson: 2014, 'Sequential Learning, Predictability, and Optimal Portfolio Returns'. *Journal of Finance* **69**(2), 611–644.

Kelly, B. and S. Pruitt: 2013, 'Market Expectations in the Cross-Section of Present Values'. *Journal of Finance* **68**(5), 1721–1756.

Kilponen, J. and F. Verona: 2016, 'Testing the Q theory of investment in the frequency domain'. Research Discussion Papers 32/2016, Bank of Finland.

Lettau, M. and S. Ludvigson: 2001, 'Consumption, aggregate wealth, and expected stock returns'. *Journal of Finance* **56**(3), 815–849.

Lettau, M. and S. Ludvigson: 2002, 'Time-varying risk premia and the cost of capital: An alternative implication of the Q theory of investment'. *Journal of Monetary Economics* **49**(1), 31–66.

Lima, L. R. and F. Meng: 2017, 'Out-of-Sample Return Predictability: A Quantile Combination Approach'. *Journal of Applied Econometrics* **32**(4), 877–895.

Ludvigson, S. C. and S. Ng: 2007, 'The empirical risk-return relation: A factor analysis approach'. *Journal of Financial Economics* **83**(1), 171 – 222.

Manchaldore, J., I. Palit, and O. Soloviev: 2010, 'Wavelet decomposition for intra-day volume dynamics'. *Quantitative Finance* **10**(8), 917–930.

Meligkotsidou, L., E. Panopoulou, I. D. Vrontos, and S. D. Vrontos: 2014, 'A Quantile Regression Approach to Equity Premium Prediction'. *Journal of Forecasting* **33**(7), 558–576.

Mitra, S. and A. Mitra: 2006, 'Modeling exchange rates using wavelet decomposed genetic neural networks'. *Statistical Methodology* **3**(2), 103 – 124.

Neely, C., D. Rapach, J. Tu, and G. Zhou: 2014, 'Forecasting the Equity Risk Premium: The Role of Technical Indicators'. *Management Science* **60**(7), 1772–1791.

Percival, D. and A. Walden: 2000, *Wavelet methods for time series analysis.* Cambridge University Press.

Pettenuzzo, D. and F. Ravazzolo: 2016, 'Optimal Portfolio Choice Under Decision-Based Model Combinations'. *Journal of Applied Econometrics* **31**(7), 1312–1332.

Pettenuzzo, D., A. Timmermann, and R. Valkanov: 2014, 'Forecasting stock returns under economic constraints'. *Journal of Financial Economics* **114**(3), 517–553.

Rapach, D. E., M. C. Ringgenberg, and G. Zhou: 2016, 'Short interest and aggregate stock returns'. *Journal of Financial Economics* **121**(1), 46 – 65.

Rapach, D. E., J. K. Strauss, and G. Zhou: 2010, 'Out-of-Sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy'. *Review of Financial Studies* **23**(2), 821–862.

Rua, A.: 2011, 'A wavelet approach for factor-augmented forecasting'. *Journal of Forecasting* **30**(7), 666–678.

Rua, A.: 2017, 'A wavelet-based multivariate multiscale approach for forecasting'. *International Journal of Forecasting* **33**(3), 581–590.

Voutilainen, V.: 2017, 'Wavelet decomposition of the financial cycle: an early warning system for financial tsunamis'. Research Discussion Papers 11/2017, Bank of Finland.

Watson, M. W.: 1986, 'Univariate detrending methods with stochastic trends'. *Journal of Monetary Economics* **18**(1), 49–75.

Xue, Y., R. Gencay, and S. Fagan: 2013, 'Jump detection with wavelets for high-frequency financial time series'. *Quantitative Finance* **14**(8), 1427–1444.

Zhang, K., R. Gencay, and M. E. Yazgan: 2017, 'Application of wavelet decomposition in time-series forecasting'. *Economics Letters* **158**, 41–46.

Zhu, X. and J. Zhu: 2013, 'Predicting stock returns: A regime-switching combination approach and economic links'. *Journal of Banking and Finance* **37**(11), 4120 – 4133.

	mean	median	min	max	std. dev.	AR(1)
Equity premium (%)	0.41	0.84	-24.8	14.9	4.44	0.04
DP	-3.63	-3.60	-4.52	-2.75	0.44	1.00
DY	-3.62	-3.60	-4.53	-2.75	0.44	1.00
EP	-2.83	-2.86	-4.84	-1.90	0.48	0.99
DE	-0.80	-0.85	-1.24	1.38	0.34	0.99
RVOL (ann.)	0.15	0.14	0.05	0.32	0.05	0.96
BM	0.48	0.35	0.12	1.21	0.29	1.00
NTIS	0.01	0.01	-0.06	0.05	0.02	0.98
TBL (%. ann.)	4.83	5.00	0.01	16.3	3.51	0.99
LTY (%. ann.)	6.95	6.93	1.75	14.8	2.84	1.00
LTR (%)	0.71	0.75	-11.2	15.2	3.13	0.05
TMS (%. ann.)	2.12	2.33	-3.65	4.55	1.48	0.95
DFY (%. ann.)	1.10	0.96	0.55	3.38	0.46	0.96
DFR (%)	0.01	0.05	-9.75	7.37	1.48	-0.04
INFL (%)	0.33	0.30	-1.92	1.81	0.39	0.62

Table 1: Summary statistics

This table reports summary statistics for the log equity premium and for the set of predictive variables. The sample period is from 1973:01 to 2016:12. Equity premium, LTR, DFR, and INFL (TBL, LTY, TMS, and DFY) are measured in percent (annual percent).

	D_1	D_2	D_3	D_4	D_5	D_6	D_7
Equity premium	48	26	12	7	4	2	2
DP	0	0	1	2	3	5	89
DY	0	0	1	2	3	5	89
EP	1	1	3	8	11	11	65
DE	1	2	8	19	29	26	15
RVOL	2	3	6	15	25	29	20
BM	0	0	1	2	3	4	91
NTIS	1	2	4	11	20	18	44
TBL	0	1	2	3	6	10	78
LTY	0	1	1	2	3	5	89
LTR	47	28	13	6	3	1	1
TMS	3	4	7	11	20	30	25
DFY	2	4	7	13	18	18	37
DFR	52	26	11	6	3	1	0
INFL	19	19	16	7	5	5	29

Table 2: Energy decomposition (%)

This table reports the variance decomposition by frequency for the time series under analysis. The sample period is from 1973:01 to 2016:12. Percentages may not add up to 100 because of rounding.

	TS		WAV_ALL	WAV_BEST								
	R^2_{OS}	CER gains	R^2_{OS}	R^2_{OS}	CER gains	δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7
DP	-1.98	-2.98	-39.6	-0.37	-0.47	0	0	0	0	0	0	0.5
DY	-2.12	-2.78	-24.6	-0.38	-0.47	0	0	0	0	0	0.25	0.5
EP	-1.09	-0.33	-48.4	2.62**	3.40	0	0.5	1	0.25	0	0	0.75
DE	-2.17	-1.04	-3.24	2.82***	3.62	1	0	0.25	0.5	0	0	1
RVOL	-0.61	-1.87	-4.05	-0.02	0.65	0	0.25	0	0	0.75	0	0.75
BM	-0.54	-0.72	-15.0	0.18	0.70	0	0	0	0	1	1	0.5
NTIS	-3.07	-2.32	-3.75	-0.01	0.21	0.75	1	0	0	0	0	0.5
TBL	-0.41	0.69	-3.39	-0.26	0.16	0	0	0	0	0	1	0.75
LTY	-0.31	0.00	-1.49	-0.17	0.09	0	0	0	1	0.25	1	0.5
LTR	-0.34	-0.74	-2.04	1.05**	2.41	0.25	0.5	0	0.25	0	0	1
TMS	-0.79	0.11	-0.62	1.81***	5.48	0	0	0	0	0	0	1
DFY	-2.93	-4.56	-12.7	-0.64	-0.75	0	0	0	0	0	0	0.75
DFR	-2.05	0.59	-7.48	0.34	0.15	1	0	0	0	0.75	1	0.25
INFL	-0.69	-0.65	0.98**	1.85**	3.65	1	1	0	0.25	1	1	0.75

Table 3: Out-of-sample R-squares (R^2_{OS}) and annualized CER gains

This table reports the out-of-sample R-squares (in percentage) for the equity premium forecasts at monthly (non-overlapping) frequencies from the model as given by equation (3) for each of the original predictors (TS, second column), from WAV_ALL model specification (equation 5, fourth column) and from the WAV_BEST model (equation 6, fifth column) for each predictor, where the frequency components used and corresponding weights (δ_j, $j = 1, 2, \ldots, 7$) are listed in the last seven columns. The out-of-sample R-squares $\left(R^2_{OS}\right)$ measures the proportional reduction in the mean squared forecast error for the predictive model relative to the forecast based on the historical mean (HM). The 1-month ahead out-of-sample forecast of equity premium is generated using a sequence of expanding windows. In columns three and six are reported the annualized certainty equivalent return (CER) gain (in percent) for an investor who allocates his or her wealth between equities and risk free bills according to the rule (7), using stock return forecasts from models in equations (3) and (6) instead of the forecasts based on the HM. The sample period is from 1973:01 to 2016:12. The full out-of-sample forecasting period is from 1990:01 to 2016:12, monthly frequency. Asterisks denote significance of the out-of-sample *MSFE*-adjusted statistic of Clark and West (2007). ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

	1990:01 - 2006:12				2007:01 - 2016:12			
	TS		WAV_BEST		TS		WAV_BEST	
	R^2_{OS}	CER gains	R^2_{OS}	CER gains	R^2_{OS}	CER gains	R^2_{OS}	CER gains
DP	-3.29	-4.49	-0.37	-0.36	-0.13	-0.43	-0.37	-0.65
DY	-3.61	-4.36	-0.33	-0.25	-0.01	-0.09	-0.45	-0.82
EP	-1.05	-1.02	0.94*	2.39	-1.14	0.83	4.99*	5.12
DE	-1.83	-1.60	0.88*	2.27	-2.65	-0.10	5.53**	5.91
RVOL	-1.73	-3.02	1.87**	3.72	0.96	0.07	-2.68	-4.50
BM	-0.86	-1.14	0.06	0.60	-0.09	-0.02	0.36	0.88
NTIS	-2.84	-1.59	0.12	0.67	-3.38	-3.57	-0.20	-0.55
TBL	-0.49	1.15	-0.38	0.00	-0.29	-0.11	-0.10	0.45
LTY	-0.41	-0.33	-0.62	-0.65	-0.18	0.57	0.47	1.34
LTR	-0.83	-0.98	0.67*	1.95	0.36	-0.36	1.58*	3.20
TMS	-1.11	1.05	1.52***	5.00	-0.34	-1.50	2.22***	6.26
DFY	-3.37	-4.43	-0.42	-0.35	-2.31	-4.80	-0.94	-1.42
DFR	-2.64	0.96	0.20	0.29	-1.22	-0.04	0.53	-0.09
INFL	-0.37	0.56	-0.63	1.01	-1.15	-2.68	5.31**	8.11

Table 4: Out-of-sample R-squares (R^2_{OS}) and annualized CER gains

This table reports the out-of-sample R-squares (in percentage) for equity premium forecasts at monthly (non-overlapping) frequencies from the model as given by equation (3) for each of the original predictors (TS) and from the WAV_BEST model in equation (6) for each predictor, where the frequency components used and corresponding weights (δ_j, $j = 1, 2, \ldots, 7$) are listed in the last seven columns of Table 3. The out-of-sample R-squares $\left(R^2_{OS}\right)$ measures the proportional reduction in the mean squared forecast error for the predictive model relative to the forecast based on the historical mean (HM). The 1-month ahead out-of-sample forecast of equity premium is generated using a sequence of expanding windows. It is also reported the annualized certainty equivalent return (CER) gain (in percent) for an investor who allocates his or her wealth between equities and risk free bills according to the rule (7), using stock return forecasts from above mentioned models in equations (3) and (6) instead of forecasts based on the HM. The sample period is from 1973:01 to 2016:12. Two out-of-sample forecasting periods are considered: from 1990:01 to 2006:12 and from 2007:01 to 2016:12, monthly frequency. Asterisks denote significance of the out-of-sample *MSFE*-adjusted statistic of Clark and West (2007). ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

	Bad				Normal				Good			
	TS		WAV_BEST		TS		WAV_BEST		TS		WAV_BEST	
	R^2_{OS}	CER gains	R^2_{OS}	CER gains	R^2_{OS}	CER gains	R^2_{OS}	CER gains	R^2_{OS}	CER gains	R^2_{OS}	CER gains
DP	-0.99	-1.85	-0.50	-0.34	-2.90	-3.85	-0.17	-1.39	-2.41	-3.23	-0.36	0.33
DY	-1.08	-1.50	-0.48	-0.28	-3.08	-3.85	-0.04	-1.14	-2.58	-2.98	-0.52	0.03
EP	-1.87	-0.80	6.90*	8.38	0.28	0.54	-4.38	-4.21	-1.20	-0.71	2.93**	6.06
DE	-2.56	1.08	7.09***	10.11	-3.15	-3.71	-2.75	-3.52	-1.02	-0.51	2.07	4.30
RVOL	-0.62	-2.52	-1.07	-0.85	-0.50	-2.16	2.73*	3.37	-0.67	-0.96	-0.85	-0.52
BM	-0.81	-1.70	0.35	1.55	0.17	-0.18	0.00	-0.96	-0.75	-0.28	0.13	1.51
NTIS	-4.72	-4.73	1.29	2.57	-1.18	-0.35	-1.60	-2.95	-2.58	-1.86	-0.31	1.03
TBL	-0.31	0.18	-0.06	0.89	0.59	1.79	-0.24	-0.87	-1.24	0.05	-0.51	0.48
LTY	-0.41	-0.54	0.04	0.92	0.45	0.54	-0.09	-1.09	-0.75	0.00	-0.45	0.46
LTR	-0.11	-0.25	1.33*	4.04	-0.45	-1.08	0.65	0.41	-0.52	-0.92	1.03*	2.78
TMS	0.39	0.30	2.31***	6.86	-2.04	-0.09	2.25***	4.53	-1.22	0.10	0.93*	5.03
DFY	-2.38	-3.85	-0.84	-0.88	-4.72	-6.25	-0.12	-0.94	-2.25	-3.59	-0.78	-0.42
DFR	-10.63	-1.90	-1.45	0.72	-0.46	-0.31	-0.12	-1.70	6.41	3.98	2.68	1.43
INFL	-0.96	-1.00	2.92**	5.99	0.42	0.45	-0.08	-0.49	-1.20	-1.40	2.04	5.40

Table 5: Out-of-sample R-squares (R^2_{OS}) and annualized CER gains

This table reports the out-of-sample R-squares (in percentage) for equity premium forecasts at monthly (non-overlapping) frequencies from the model as given by equation (3) for each of the original predictors (TS) and from the WAV_BEST model in equation (6) for each predictor, where the frequency components used and corresponding weights (δ_j, $j = 1, 2, \ldots, 7$) are listed in the last seven columns of Table 3. The out-of-sample R-squares $\left(R^2_{OS}\right)$ measures the proportional reduction in the mean squared forecast error for the predictive model relative to the forecast based on the historical mean (HM). The 1-month ahead out-of-sample forecast of equity premium is generated using a sequence of expanding windows. It is also reported the annualized certainty equivalent return (CER) gain (in percent) for an investor who allocates his or her wealth between equities and risk free bills according to the rule (7), using stock return forecasts from above mentioned models in equations (3) and (6) instead of forecasts based on the HM. The sample period is from 1973:01 to 2016:12. Three out-of-sample forecasting periods are considered, each with 108 monthly observations: bad growth, normal growth and good growth. Those regimes are defined as the bottom, middle and top third of sorted growth rates of industrial production in the US, respectively. Asterisks denote significance of the out-of-sample *MSFE*-adjusted statistic of Clark and West (2007). ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

	TS	WAV_BEST							
	R^2_{OS}	R^2_{OS}	δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7
DP	-4.83	-1.49	0	0	0	0	1	0.25	0.5
DY	-6.68	-1.28	0	0	0	0	1	0.25	0.5
EP	-2.46	6.60**	0.75	0	0	0	0.25	0	0.5
DE	-4.56	2.45*	0	0.5	0	0	0	0	1
RVOL	-4.93	-2.75	0	0	0	0	0	0	0.5
BM	-1.09	0.37	0	0	1	1	0	0	0.5
NTIS	-5.44	-1.32	0	0.75	0	0	0	0	0.75
TBL	-1.28	-0.58	1	0.25	0	0	0	0	0.75
LTY	0.06	1.99*	1	1	0	1	0	0	0.75
LTR	-2.60	0.81	1	0	0.25	0	1	0	1
TMS	-3.47	0.03	0	0	0	0	0	0.25	0.75
DFY	-1.92	-1.34	1	0	0	0	0	0	0.5
DFR	-1.12	-0.55	0	0	0.25	0.25	0	0	0.5
INFL	1.64	1.47	0	0	0.25	1	1	0.5	0.75
IK	1.49	4.49**	0	1	0.5	0	0	0	1
CAY	-2.81	1.29	0	0	0	0	1	1	0.75
Q	-3.07	-1.11	0	0	0	1	0.75	0	0.75

Table 6: Out-of-sample R-squares (R^2_{OS}) using quarterly data

This table reports the out-of-sample R-squares (in percentage) for the equity premium forecasts at quarterly (non-overlapping) frequencies from the model as given by equation (3) for each of the original predictors (TS, second column) and from the WAV_BEST model (equation 6, third column) for each predictor where the frequency components used and corresponding weights (δ_j, $j = 1, 2, \ldots, 7$) are listed in the last seven columns. The out-of-sample R-squares $\left(R^2_{OS}\right)$ measures the proportional reduction in the mean squared forecast error for the predictive model relative to the forecast based on the historical mean (HM). The 1-quarter ahead out-of-sample forecast of equity premium is generated using a sequence of expanding windows. The sample period is from 1952:Q1 to 2016:Q4. The full out-of-sample forecasting period is from 1990:Q1 to 2016:Q4, quarterly frequency. Asterisks denote significance of the out-of-sample *MSFE*-adjusted statistic of Clark and West (2007). ** and * denote significance at the 5% and 10% levels, respectively.

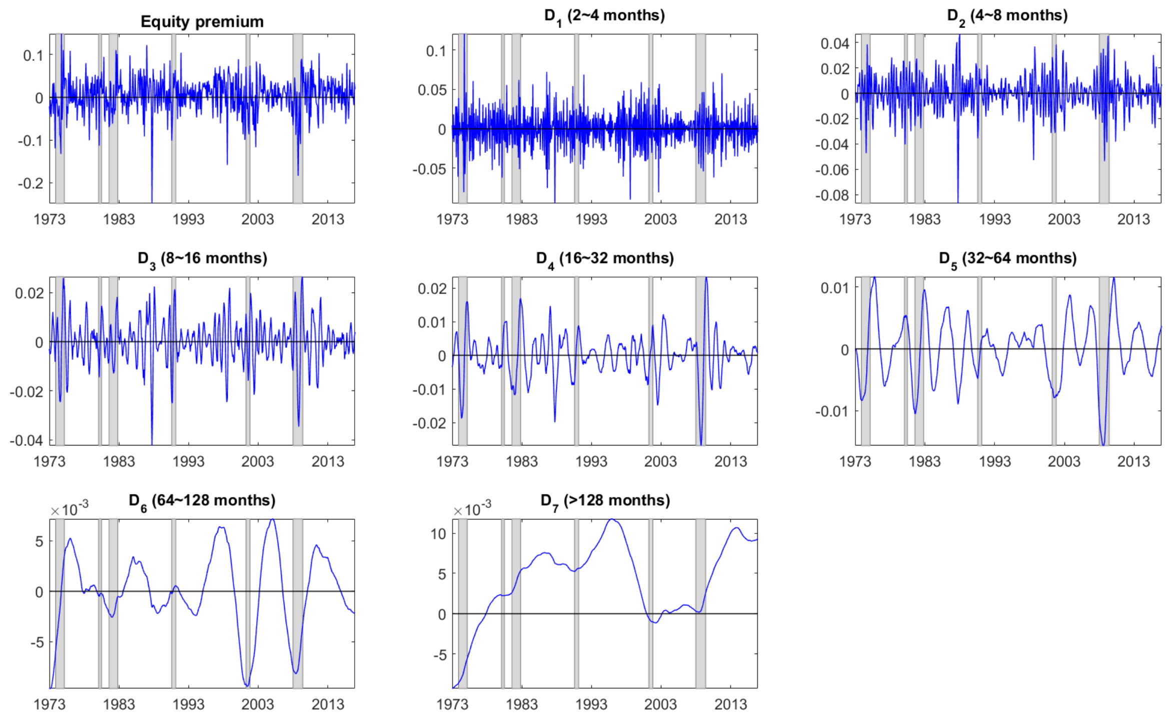


Figure 1: Equity premium, time series and wavelet decomposition

The time series of the (log) equity premium as proxied by the log S&P 500 index total return minus the log return on a one-month Treasury bill is presented in the top left panel. In the remaining panels are plotted the seven frequency components into which the equity premium time series is decomposed. It is applied a $J = 6$ level wavelet decomposition which leads to six wavelet details $(D_1, D_2, \ldots, D_6)$, representing the higher-frequency characteristics of the series, and a wavelet smooth (D_7), that captures the low-frequency dynamics of the series. Sample period from 1973:01 to 2016:12, monthly frequency.

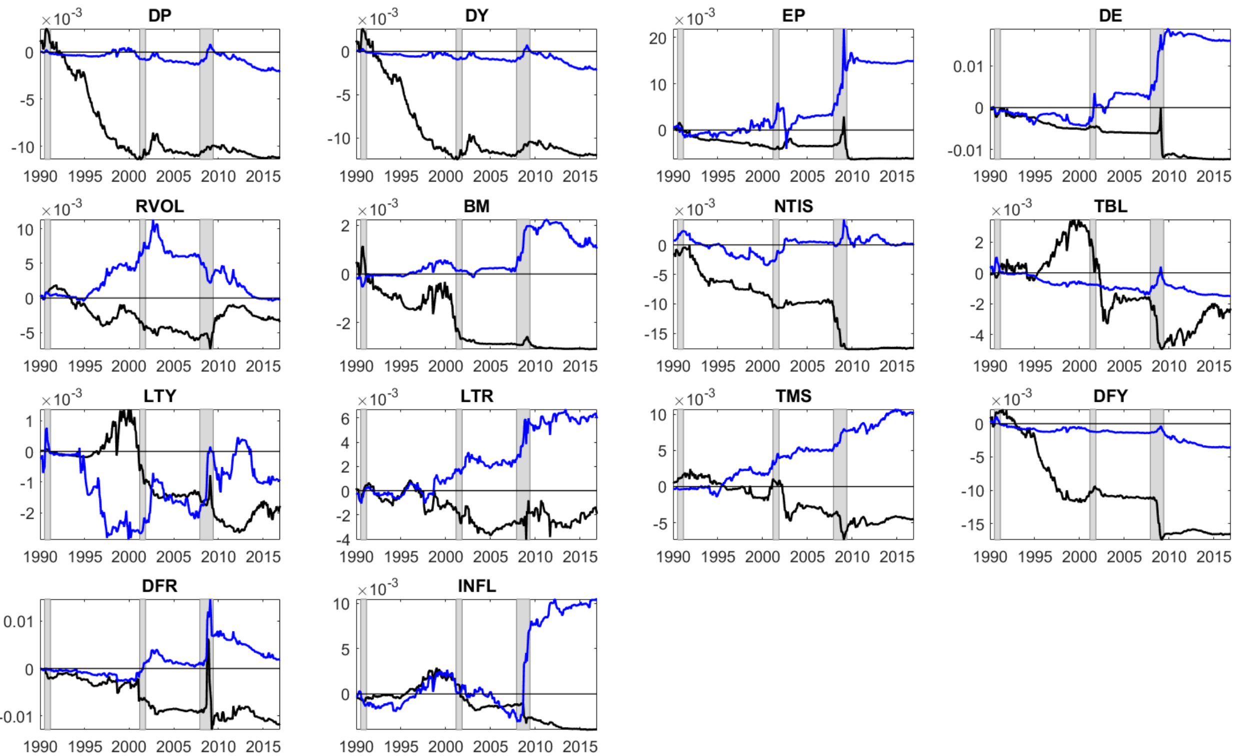
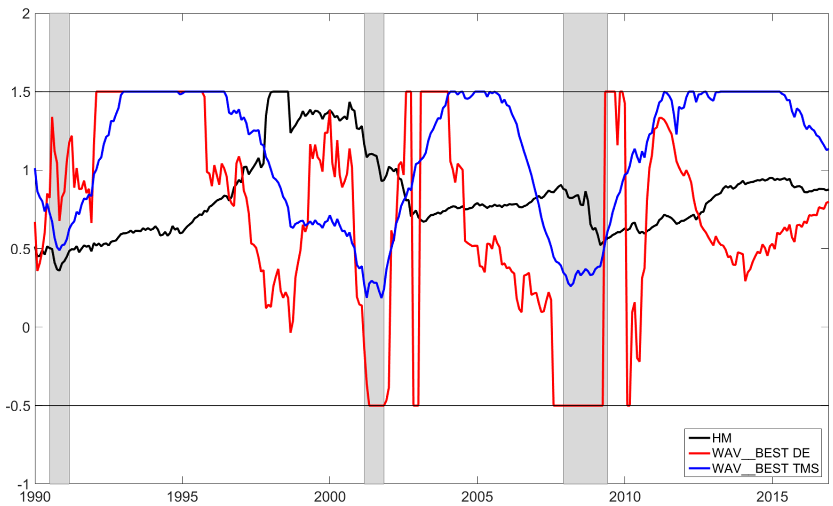


Figure 2: Difference between cumulative square forecasting error for the HM forecasting model and the cumulative square forecasting error for the individual predictive regression forecasting model

This figure reports the dynamics of the difference between the cumulative square forecasting error for the HM forecasting model and the cumulative square forecasting error for the predictive regression forecasting based on the WAV_BEST model (6) for each predictor with frequency components reported in table 3 (blue line), and when each predictor is considered in its original monthly time series (TS, black line). The sample period is from 1973:01 to 2016:12. The full out-of-sample forecasting period is from 1990:01 to 2016:12, monthly frequency.

A. Equity weights



B. Log cumulative wealth

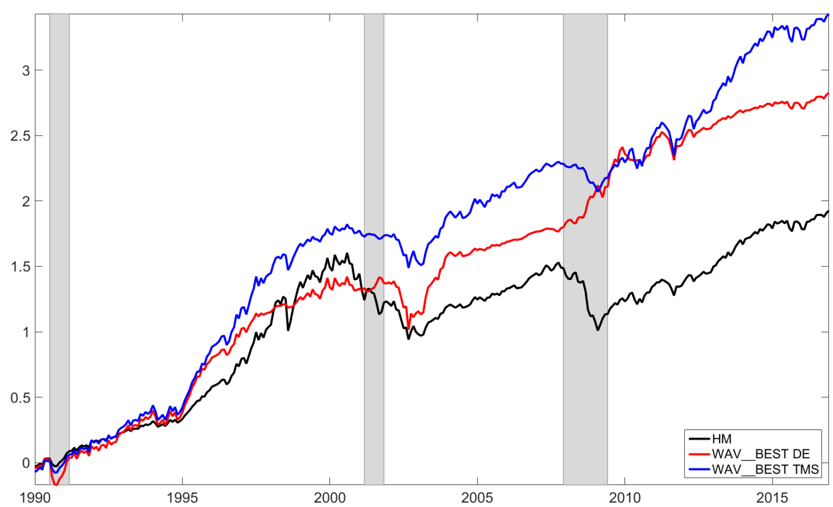


Figure 3: Equity weights and log cumulative wealth

Panel A plots the dynamics of the equity weight for a mean-variance investor who allocates monthly his or her wealth between equities and risk free bills according to the rule (7), using stock return forecasts based on the HM benchmark (black line), on the wavelet forecast with the WAV_BEST model (6) for the TMS (WAV_BEST TMS, blue line) and the DE (WAV_BEST_DE). The equity weight is constrained to lie between -0.5 and 1.5. Panel B delineates the corresponding log cumulative wealth for the investor, assuming that he or she begins with 1$ and reinvests all proceeds. Grey bars denote NBER-dated recessions. The investor is assumed to have a relative risk aversion coefficient of three. Sample period from 1990:01 to 2016:12, monthly frequency.

Appendix 1. Definition of equity premium predictors

- Log dividend-price ratio (DP): difference between the log of dividends (12-month moving sums of dividends paid on S&P 500) and the log of prices (S&P 500 index).
- Log dividend yield (DY): difference between the log of dividends (12-month moving sums of dividends paid on S&P 500) and the log of lagged prices (S&P 500 index).
- Log earnings-price ratio (EP): difference between the log of earnings (12-month moving sums of earnings on S&P 500) and the log of prices (S&P 500 index price).
- Log dividend-payout ratio (DE): difference between the log of dividends (12-month moving sums of dividends paid on S&P 500) and the log of earnings (12-month moving sums of earnings on S&P 500).
- Excess stock return volatility (RVOL): calculated using a 12-month moving standard deviation estimator.
- Book-to-market ratio (BM): ratio of book value to market value for the Dow Jones Industrial Average.
- Net equity expansion (NTIS): ratio of 12-month moving sums of net equity issues by NYSE-listed stocks to the total end-of-year NYSE market capitalization.
- Treasury bill rate (TBL): three-month Treasury bill rate.
- Long-term yield (LTY): long-term government bond yield.
- Long-term return (LTR): long-term government bond return.

- Term spread (TMS): difference between the long-term government bond yield and the T-bill.
- Default yield spread (DFY): difference between Moody's BAA- and AAA-rated corporate bond yields.
- Default return spread (DFR): difference between long-term corporate bond and long-term government bond returns.
- Inflation rate (INFL): calculated from the Consumer Price Index (CPI) for all urban consumers.

For quarterly data, we also use:

- Investment to capital ratio (IK): ratio of aggregate (private nonresidential fixed) investment to aggregate capital for the whole economy.
- Consumption-wealth ratio (CAY): log consumer spending minus log asset wealth (total household net worth) and minus log labor income, all measured on an aggregate basis.
- Tobin's Q (Q): the data is from the Flow of Funds Table B103 – Balance Sheet of Non-financial Corporate Business; Q is calculated as the ratio of non-financial corporate business, corporate equities, liability (level) to non-financial corporate business, net worth (level).

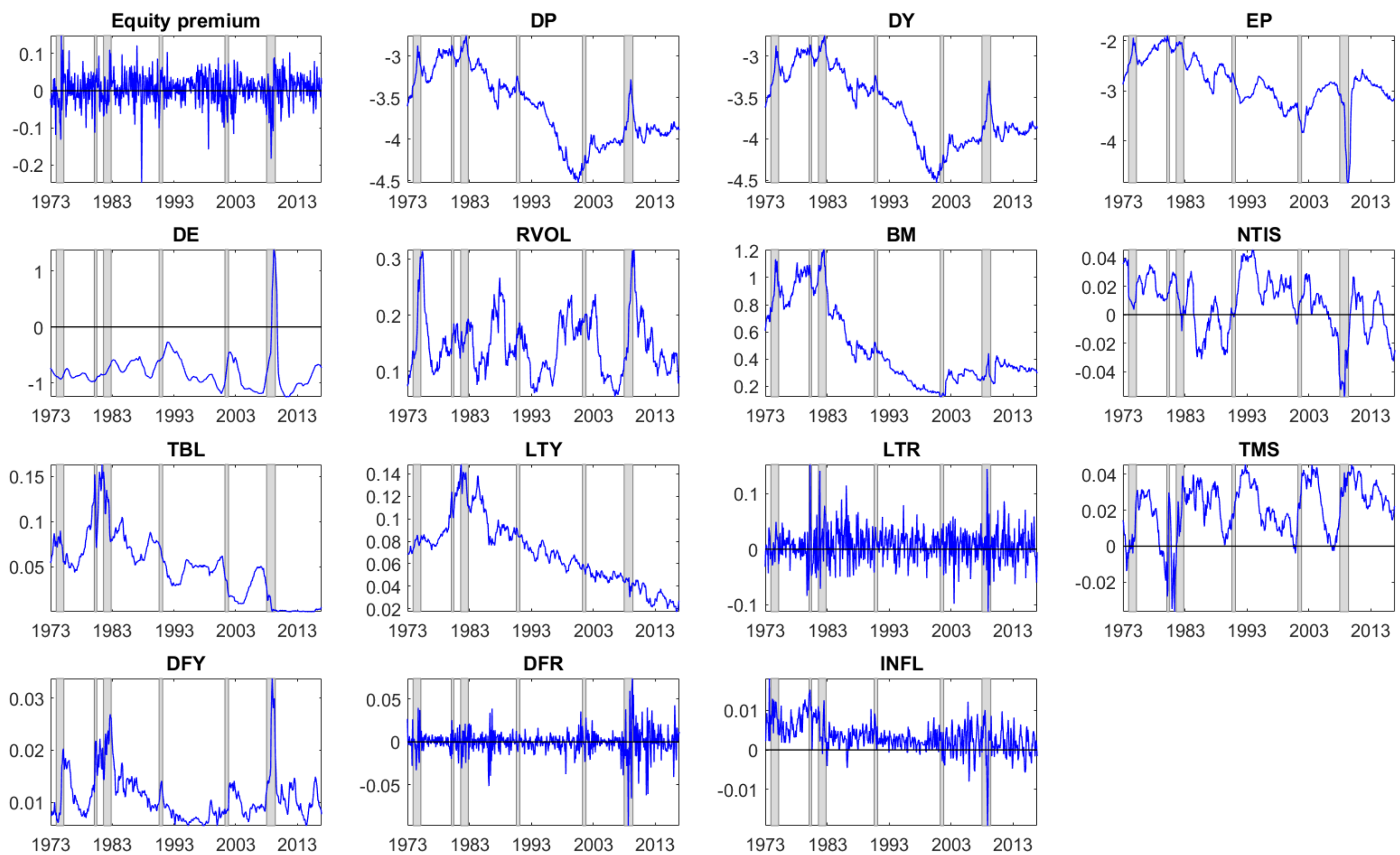


Figure 4: Monthly time series of the equity premium and the predictors

Appendix 2.

The discrete wavelet transform (DWT) multiresolution analysis (MRA) allows the decomposition of a time series into its constituent multiresolution (frequency) components. There are two types of wavelets: father wavelets (ϕ), which capture the smooth and low frequency part of the series, and mother wavelets (ψ), which capture the high frequency components of the series, where $\int \phi(t)\,dt = 1$ and $\int \psi(t)\,dt = 0$.

Given a time series y_t with a certain number of observations N, its wavelet multiresolution representation is given by

$$y_t = \sum_k s_{J,k}\phi_{J,k}(t) + \sum_k d_{J,k}\psi_{J,k}(t) + \sum_k d_{J-1,k}\psi_{J-1,k}(t) + \cdots + \sum_k d_{1,k}\psi_{1,k}(t) \quad , \qquad (8)$$

where J represents the number of multiresolution levels (or frequencies), k defines the length of the filter, $\phi_{J,k}(t)$ and $\psi_{j,k}(t)$ are the wavelet functions and $s_{J,k}$, $d_{J,k}$, $d_{J-1,k}, \ldots, d_{1,k}$ are the wavelet coefficients.

The wavelet functions are generated from the father and mother wavelets through scaling and translation as follows

$$\begin{aligned} \phi_{J,k}(t) &= 2^{-J/2}\phi\left(2^{-J}t - k\right) \\ \psi_{j,k}(t) &= 2^{-j/2}\psi\left(2^{-j}t - k\right) \quad , \end{aligned}$$

while the wavelet coefficients are given by

$$\begin{aligned} s_{J,k} &= \int y_t \phi_{J,k}(t)\,dt \\ d_{j,k} &= \int y_t \psi_{j,k}(t)\,dt \quad , \end{aligned}$$

where $j = 1, 2, ..., J$.

Due to the practical limitations of DWT in empirical applications, we perform wavelet decomposition analysis here by applying the maximal overlap discrete wavelet transform (MODWT). The MODWT is not restricted to a particular sample size, is translation-invariant so that it is not sensitive to the choice of the starting point of the examined time series, and does not introduce phase shifts in the wavelet coefficients (so peaks or troughs in the original time series are correctly aligned with similar events in the MODWT MRA). This last property is especially relevant in the forecasting exercise.