CATÓLICA
CATÓLICA PORTO BUSINESS SCHOOL
PORTO

WORKING PAPERS

ECONOMICS

Nº 05/2016

FORECASTING STOCK MARKET RETURNS BY SUMMING THE FREQUENCY-DECOMPOSED PARTS

Gonçalo Faria
Católica Porto Business School

Fabio Verona
Bank of Finland

Forecasting stock market returns by summing the frequency-decomposed parts*

Gonçalo Faria[†] Fabio Verona[‡]

4 November 2016

Abstract

We forecast stock market returns by applying, within a Ferreira and Santa-Clara (2011) sum-of-the-parts framework, a frequency decomposition of several predictors of stock returns. The method delivers statistically and economically significant improvements over historical mean forecasts, with monthly out-of-sample R^2 of 3.27 % and annual utility gains of 403 basis points. The strong performance of this method comes from its ability to isolate the frequencies of the predictors with the highest predictive power from the noisy parts, and from the fact that the frequency-decomposed predictors carry complementary information that captures both the long-term trend and the higher frequency movements of stock market returns.

Keywords: predictability, stock returns, equity premium, asset allocation, frequency domain, wavelets

JEL classification: G11, G12, G14, G17

* The authors thank Amit Goyal and David Rapach for kindly providing codes and data on their webpages; Gene Ambrocio, João Correia da Silva, Eleonora Granziera, Adam Gulan, Esa Jokivuolle and seminar participants at the Bank of Finland 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/00731/2013).

† Universidade Católica Portuguesa – Católica Porto Business School and CEGE (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 has a long tradition in finance. A reliable forecast of stock market returns is a crucial input for the computation of the cost of capital and, therefore, for the investment process in general. It is hardly surprising then that stock return predictability has stimulated an immense literature and remains a very active research topic, as recent papers by Pettenuzzo, Timmermann and Valkanov (2014), Bollerslev, Todorov and Xu (2015), Huang et al. (2015) and Rapach, Ringgenberg and Zhou (2016) confirm. This paper contributes to this literature by using a method that allows to efficiently exploit the information resulting from the frequency decomposition of several potential predictors of stock market returns.

Our work builds on the Ferreira and Santa-Clara (2011) sum-of-the-parts (SOP) method for forecasting stock market returns. Conceptually, the SOP method consists in decomposing the stock market return into three components, which are first forecasted separately and then added in order to obtain the forecast of the stock market return. The SOP method improves the forecast accuracy (as compared to the historical mean benchmark) because it exploits the different time series persistence of the stock market returns components. Our proposed method provides an alternative way of forecasting those three components. Namely, we directly use (some of) the frequency-decomposed time series of a set of popular predictors from the literature. The frequency decomposition is implemented through a discrete wavelet transform multiresolution analysis, which is gaining popularity as a tool for econometric analysis (see *e.g.* Galagedera and Maharaj, 2008, Xue, Gencay and Fagan, 2013 and the references in section 2.2). In a nutshell, the method consists in decomposing a time series into n orthogonal time series components, each capturing the oscillations of the original variable within a specific frequency interval. Lower frequencies represent the long-term dynamics of the original time series, while the higher frequencies capture the short-term dynamics. As the n frequency-decomposed components are orthogonal, by adding them back together the original time series is recovered. We refer to our method as the SOPWAV method, as it consists in applying, within a SOP method for stock market returns forecasts, a wavelet decomposition approach (WAV). The SOPWAV method, by *explicitly* decomposing the different predictors of stock returns into their frequency time-series components, allows to identify the best predictors and to exclude the noisy parts. In other words, it only retains the components that have the greatest predictive power. This leads to expressive statistically and economically forecasting improvements.

As a preview of the results, figure 1 shows, for the full out-of-sample period considered in this paper, the realized S&P 500 index log return (black solid line) together with the forecasts based on the historical mean (HM) of returns (black dashed line) and the SOPWAV method (blue solid line). The out-of-sample forecast from the SOPWAV method clearly tracks the dynamics of stock market returns more closely than the HM forecast. Indeed, the correlation between the SOPWAV forecast and the realized S&P 500 index return is 0.25, while it is almost zero between the HM forecast and the realized return. The strength of the wavelet decomposition, which is at the core of the SOPWAV method, is that it permits the capture of both the low-frequency dynamics of stock market returns (long-run trend) as the HM does, and part of the higher-frequency movements of the stock market that are not captured by the HM.

Using the HM as a benchmark, the monthly out-of-sample R^2 for the SOPWAV method is 3.27 % for the full out-of-sample period (January 1950 to December 2015). When examining the economic significance of the SOPWAV predictive performance through an asset allocation analysis, we find that a mean-variance investor who allocates her wealth between equities and risk-free bills enjoys significant utility gains from using a SOPWAV-based trading strategy. Specifically, the rate of return that an investor would be willing to accept instead of holding the risky portfolio is 403 basis points. Furthermore, the annualized Sharpe ratio of the strategy based on the SOPWAV method is 0.65, which is about 1.8 times the Sharpe ratio generated by the HM-based strategy. These results are robust to the introduction of transaction costs, different sets of portfolio constraints and different out-of-sample forecasting periods.

The rest of the paper is structured as follows. Section 2 reviews the two strands of literature on which our work builds and places our contribution. Section 3 presents the data and the methodology. We introduce the two blocks of the SOPWAV method: the original SOP method and the discrete wavelet decomposition of the predictors. The out-of-sample forecasting procedure and asset allocation analysis are also described. Sections 4 and 5 present the out-of-sample and asset allocation results, respectively. Section 6 concludes.

2 Literature review

This paper draws on two strands of literature. The first deals with the out-of-sample stock return predictability using standard time series tools. The second involves the application of wavelet methods to economic and finance topics. In the next two sections, we provide a brief overview of these two strands of literature and we place the contribution of this paper with respect to both.

2.1 Forecasting stock returns

As evidenced in the reviews of Rapach and Zhou (2013) and Harvey, Liu and Zhu (2016), the literature on predicting stock returns and equity premium is vast. Several studies discuss the in-sample predictability using predictors such as the treasury bill rate, dividend yield, dividend-price ratio, term spread, equity market volatility or the consumption-wealth ratio. This is the case for the US stock market (see *e.g.* Fama and Schwert, 1977, Campbell, 1987, Ferson and Harvey, 1991, Lettau and Ludvigson, 2001, Cochrane, 2008, Goyal and Welch, 2008 and Pastor and Stambaugh, 2009), as well as for other stock markets (see *e.g.* Cutler, Poterba and Summers, 1991, Harvey, 1991, Bekaert and Hodrick, 1992, Ferson and Harvey, 1993, Ang and Bekaert, 2007 and Hjalmarsson, 2010). However, Goyal and Welch (2008) note how poorly the abovementioned predictors perform out-of-sample up to 2008. As predictive models require out-of-sample validation, and given this poor out-of-sample performance, researchers have then turned their attention to improving the out-of-sample forecastability of stock returns. In particular, the literature after Goyal and Welch (2008) has followed two main strategies.

The first strategy focuses on developing and testing new predictors. For instance, Bollerslev, Tauchen and Zhou (2009) test the variance risk premium, Cooper and Priestley (2009, 2013) use the output gap and the world business cycle, Rapach, Strauss and Zhou (2013) document the relevance of lagged US market returns for the out-of-sample predictability of stock returns of other industrialized countries, Li, Ng and Swaminathan (2013) study the aggregate implied cost of capital, Neely et al. (2014) study the relevance of some technical indicators (complementary predictors to the traditional set of variables), Huang et al. (2015) propose a new investor sentiment index, Moller and Rangvid (2015) study different macroeconomic variables by focusing on their fourth-quarter growth rate and Rapach, Ringgenberg and Zhou (2016) construct an aggregate short interest position indicator.

The second strategy focuses on improving existing forecasting methods. For example, Ludvigson and Ng (2007) and Kelly and Pruitt (2013) propose using dynamic factor analysis for large data sets to summarize a large amount of information by few estimated factors, Rapach, Strauss and Zhou (2010) suggest combining individual forecasts from different predictors, Ferreira and Santa-Clara (2011) introduce the SOP method, Dangl and Halling (2012) evaluate predictive regressions that explicitly consider the time-variation of coefficients, Pettenuzzo, Timmermann and Valkanov (2014) propose an approach to impose economic constraints on forecasts of the equity premium, Bollerslev, Todorov and Xu (2015) decompose the predictor (the variance risk premium) into a jump and a diffusion component, and Baetje and Menkhoff (2016) examine the time-instability of the standard set of economic and technical indicators.

We place our contribution in both strands of research, as the frequency decomposition of the predictors is not only a methodological contribution *per se*, but it also represents an expansion of the set of possible predictors, as each frequency of each predictor can be understood and potentially used as a new predictor.

2.2 Wavelets applications in economics and finance

Wavelets are a signal processing technique developed to overcome some of the limitations of traditional frequency domain methods (*i.e.* spectral analysis and Fourier transforms). The advantage of wavelets methodology is that it overcomes the weaknesses of traditional frequency domain tools while providing a more complete decomposition of the original time series. Unlike Fourier analysis, wavelets are defined over a finite support/window in the time domain, with the size of the window resized automatically according to the frequency of interest. In essence, using a short window allows to isolate the high frequency features of the time series, while looking at the same signal with a large window reveals the low frequency features. Hence, by varying the size of the time window, it is possible to capture simultaneously both time-varying and frequency-varying features of the time series. Wavelets are thus extremely useful when the time series have structural breaks or jumps, as well as with non-stationary time series.

Crowley (2007) and Aguiar-Conraria and Soares (2014) provide excellent reviews of economic and finance applications of wavelets.[1] The pioneering works in these fields are Ramsey and Lampart (1998*a*,*b*), in which wavelets are used to study the relationship between macroeconomic variables (consumption versus income and money supply versus income, respectively). More recently, wavelets methods have been applied to test for the (in-sample) frequency dependence between two (or more) variables (Kim and In, 2005, Gencay, Selcuk and Whitcher, 2005, Gallegati et al., 2011 and Gallegati and Ramsey, 2013), and to study the comovements and lead-lag relationship between variables at different frequencies (Rua and Nunes, 2009, Rua, 2010, Aguiar-Conraria and Soares, 2011 and Aguiar-Conraria, Martins and Soares, 2012). A relatively unexplored area of research is the application of wavelet methods for forecasting purposes. A notable exception is Rua (2011), who proposes a wavelet approach for factor-augmented forecasting of GDP growth and finds that significant predictive short-run improvements can be achieved by using wavelets (in combination with factor-augmented models).

[1] Wavelet methods have long been popular in many fields, including astronomy, engineering, geology, medicine, meteorology and physics.

In this paper, we demonstrate the statistical and economic advantages of applying wavelet decomposition in the context of forecasting stock market returns.

3 Data and methodology

We focus on the out-of-sample (OOS) predictability of monthly stock returns, proxied by the S&P500 index total return. We use monthly data from January 1927 to December 2015 for a set of potential predictors from Ferreira and Santa-Clara (2011). Specifically, we use the dividend-price ratio, the growth rate of dividends, the book-to-market ratio, the default return spread, the default yield spread, the dividend-payout ratio, the earnings-price ratio, the inflation rate, the long-term government bond return, the long-term government bond yield, the net equity expansion, the return on equity, the stock variance, the term spread and the treasury bill rate. Appendix 1 provides a brief description of the predictors.

Our methodology to forecast stock market returns builds on two blocks: the SOP method proposed by Ferreira and Santa-Clara (2011) and the discrete wavelet transform decomposition of the different predictors. We describe these blocks in sections 3.1 and 3.2, respectively. The OOS procedure and asset allocation analysis are described in sections 3.3 and 3.4, respectively.

3.1 The sum-of-the-parts method: decomposition of the stock market return

Ferreira and Santa-Clara (2011) propose the SOP method for forecasting stock market returns. Conceptually, this consists in decomposing the stock market return into various components that are first forecasted separately and then added together to obtain the stock market return forecast.

The stock market total return from time t to time $t+1$, R_{t+1}, can be decomposed into capital gains, CG_{t+1}, and dividend yield, DY_{t+1}:

$$1 + R_{t+1} = 1 + CG_{t+1} + DY_{t+1} = \frac{P_{t+1}}{P_t} + \frac{D_{t+1}}{P_t} \quad , \tag{1}$$

where P_{t+1} is the stock price at time $t+1$ and D_{t+1} is the dividend per share paid between time t and $t+1$. Each component in equation (1) is then further decomposed.

Capital gains can be rewritten as

$$
\begin{aligned}
1+CG_{t+1} &= \frac{P_{t+1}}{P_t} = \frac{P_{t+1}/D_{t+1}}{P_t/D_t}\frac{D_{t+1}}{D_t} = \frac{M_{t+1}}{M_t}\frac{D_{t+1}}{D_t} = \\
&= (1+GM_{t+1})\,(1+GD_{t+1}) \quad , \qquad (2)
\end{aligned}
$$

where $M_{t+1} = P_{t+1}/D_{t+1}$ is the price-dividend multiple, GM_{t+1} is the price-dividend multiple growth rate, and GD_{t+1} is the dividend growth rate. At this point, the SOP method offers flexibility as to choice of the variable to be included in equation (2). For instance, instead of using the price-dividend multiple, other price multiples such as price-earnings, price-to-book or price-to-sales may be used. Analytically, in equation (2) one would just need to replace the growth rate of dividends by the growth rate of the denominator in the multiple used (earnings, book value of equity or sales). Ferreira and Santa-Clara (2011) obtain similar results using the price-earnings multiple or the price-dividend multiple. Here, we use the price-dividend multiple because it offers a clear forecasting advantage with the proposed method compared to the price-earnings multiple. This is mainly due to the huge earnings swing between 2008 and 2009, while dividends posted a much smoother behavior during that period.[2]

The dividend yield can be rewritten as:

$$
DY_{t+1} = \frac{D_{t+1}}{P_t} = \frac{D_{t+1}}{P_{t+1}}\frac{P_{t+1}}{P_t} = DP_{t+1}\,(1+GM_{t+1})\,(1+GD_{t+1}) \quad , \qquad (3)
$$

where the dividend-price ratio $DP_{t+1} = D_{t+1}/P_{t+1}$ differs from the dividend yield because of the contemporaneous timing of the dividend with respect to price.

Substituting equations (2) and (3) in (1), the total stock market return can then be rewritten as

$$
\begin{aligned}
1+R_{t+1} &= (1+GM_{t+1})\,(1+GD_{t+1}) + DP_{t+1}\,(1+GM_{t+1})\,(1+GD_{t+1}) \\
&= (1+DP_{t+1})\,(1+GD_{t+1})\,(1+GM_{t+1}) \quad , \qquad (4)
\end{aligned}
$$

that is, stock market return is the product of the dividend-price ratio, the growth rates of dividends and of the price-dividend ratio. Finally, by taking the logs on both sides of equation (4), the log stock return is given by the sum of the dividend-price ratio, the growth in dividends and the growth in the price-dividend

[2] Note that this large swing in earnings does not affect the results in Ferreira and Santa-Clara (2011) as their dataset ends in 2007.

multiple:

$$r_{t+1} = log(1 + R_{t+1}) = dp_{t+1} + gd_{t+1} + gm_{t+1} \quad , \tag{5}$$

where lowercase variables denote log rates.

3.2 Maximal Overlap Discrete Wavelet Transform MultiResolution Analysis

Wavelets tools allow for a decomposition of a given time series into different time series, each associated with a different time scale or frequency band. This decomposition process is known as multiresolution analysis (MRA). In a nutshell, by applying a maximal overlap discrete wavelet transform multiresolution analysis (MODWT MRA), a time series y_t can be decomposed as:

$$y_t = y\left(D_1\right)_t + ... + y\left(D_J\right)_t + y\left(S_J\right)_t \quad , \tag{6}$$

where the $y\left(D_j\right)_t$, $j = 1, 2, \ldots, J$ are the wavelet details and $y\left(S_J\right)_t$ is the wavelet smooth. The original time series is thus decomposed into orthogonal components ($y\left(D_1\right)_t$ to $y\left(D_J\right)_t$ and $y\left(S_J\right)_t$), called crystals, each 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 details represent the higher frequency characteristics of the time series (*i.e.* its short-term dynamics) and, as j increases, the j wavelet details represent lower frequencies movements of the series. The wavelet smooth captures the lowest frequency dynamics (*i.e.* its long-term behavior or trend).[3]

Given the availability of long data series, we apply a $J = 7$ levels MRA. Thus, the wavelet decomposition delivers eight orthogonal crystals: seven wavelet details ($y\left(D_1\right)_t$ to $y\left(D_7\right)_t$) and a wavelet smooth ($y\left(S_7\right)_t$). Since we employ monthly data in our analysis, the first detail level $y\left(D_1\right)_t$ captures oscillations between 2 and 4 months, the second detail level $y\left(D_2\right)_t$ captures oscillations between 4 and 8 months, while detail levels $y\left(D_3\right)_t$, $y\left(D_4\right)_t$, $y\left(D_5\right)_t$, $y\left(D_6\right)_t$ and $y\left(D_7\right)_t$ capture oscillations with a period of 8-16, 16-32, 32-64, 64-128 and 128-256 months, respectively. Finally, the smooth component $y\left(S_7\right)_t$ captures oscillations with a period longer than 256 months (21.3 years).[4]

[3] A detailed analysis of wavelet methods can be found in Appendix 2 and in Percival and Walden (2000).

[4] All the simulations were run using the WMTSA Wavelet Toolkit for Matlab, which is available at http://www.atmos.washington.edu/~wmtsa/. Here, we perform the MODWT MRA using the Haar wavelet filter (as in *e.g.* Manchaldore, Palit and Soloviev, 2010 and Malagon, Moreno and Rodriguez, 2015) with reflecting boundary conditions. As the wavelet family may influence the results, we also run the simulations using the Daubechies wavelet filter with the filter length L = 4 (as in Barunik and Vacha, 2015) and the Coiflet wavelet filter with the filter length L = 6 (as done by Galagedera and Maharaj, 2008). Our results are robust to changes in the wavelet family. As regards the choice of J, the number of observations dictates the maximum number of frequency bands that can be used. In particular, if N is the number of observations in the in-sample period, then J has to satisfy the constraint $N \geq 2^J$.

As an example, figure 2 plots the time series of the (log) stock market return (top graph) and its MODWT MRA decomposition (remaining graphs). Overall, the frequency-decomposed time series exhibit significantly different dynamics that are not visible from the original time series as all these different frequencies are in practice aggregated. As expected, the lower the frequency, the smoother the resulting filtered time series.

3.3 Out-of-sample forecasts

One-step-ahead OOS forecasts of stock market returns are generated using a sequence of expanding windows. We use an initial sample (January 1927 to December 1949) to make the first one-step-ahead OOS forecast. The sample is then increased by one observation and a new one-step-ahead OOS forecast is produced. We proceed in this way until the end of the sample, thus obtaining a sequence of 792 one-step-ahead OOS forecasts. The full OOS period spans the period from January 1950 to December 2015.

3.3.1 Forecast evaluation

We evaluate the forecast performance of the SOPWAV method in terms of the Campbell and Thompson (2008) OOS R-square$\left(R^2_{OS}\right)$. The R^2_{OS} statistic measures the proportional reduction in the mean squared forecast error (MSFE) for the predictive method relative to the historical sample mean and is given by

$$R^2_{OS} = 1 - \frac{\sum_{s=s_0}^{T-1}\left(r_{s+1} - \hat{\mu}_s\right)^2}{\sum_{s=s_0}^{T-1}\left(r_{s+1} - \overline{r}_s\right)^2} \ , \tag{7}$$

where $\hat{\mu}_s$ is the stock return forecast for $s{+}1$ from the SOPWAV method (see section 3.3.3), $\overline{r}_s$ is the historical mean of stock market returns up to time s, r_{s+1} is the realized stock market return in $s{+}1$, T is the total number of observations in the sample and s_0 is the number of observations in the initial sample. As standard in the literature, we choose the HM as the benchmark model. According to (7), a positive (negative) value of R^2_{OS} indicates that the SOPWAV method outperforms (underperforms) the HM in terms of MSFE.

As in Ferreira and Santa-Clara (2011), we evaluate the statistical significance of the results using the *MSFE-F* statistic proposed by McCracken (2007).[5] The *MSFE-F* statistic tests for the equality of the MSFE of the HM and the SOPWAV method forecasts as follows:

$$MSFE-F = (T - s_0)\left[\frac{\sum_{s=s_0}^{T-1}\left(r_{s+1} - \overline{r}_s\right)^2 - \sum_{s=s_0}^{T-1}\left(r_{s+1} - \hat{\mu}_s\right)^2}{\sum_{s=s_0}^{T-1}\left(r_{s+1} - \hat{\mu}_s\right)^2}\right] .$$

[5] We thank Michael McCracken for providing us additional tables of critical values in order to evaluate the statistical significance of the R^2_{OS}.

3.3.2 Forecasting with the SOP method

To forecast stock market returns, the SOP method forecasts separately each component of the stock market return as derived in equation (5). Let $\hat{\mu}_s$ denote the expected stock market return at time s for period $s+1$, then forecasting with the SOP method implies that

$$E_s r_{s+1} = \hat{\mu}_s = \hat{\mu}_s^{dp} + \hat{\mu}_s^{gd} + \hat{\mu}_s^{gm} \quad . \tag{8}$$

The expected stock market return is the sum of the forecasts of three components: the expected dividend-price ratio ($\hat{\mu}_s^{dp}$), the expected dividend growth ($\hat{\mu}_s^{gd}$) and the expected price-dividend multiple growth ($\hat{\mu}_s^{gm}$). The equivalent of equation (8) in Ferreira and Santa-Clara (2011) (their equation 12) reflects the use of the earnings growth (instead of dividend growth). To forecast the expected dividend-price ratio $\hat{\mu}_s^{dp}$, Ferreira and Santa-Clara (2011) assume that *dp* follows a random walk such that the expected dividend-price ratio equals the current dividend-price ratio. As regards the expected earnings growth $\hat{\mu}_s^{ge}$, Ferreira and Santa-Clara (2011) assume that it is captured by the 20-year moving average of the growth in earnings up to time s. This is consistent with the view that earnings growth has a low-frequency predictable component. Finally, the expected multiple growth (price-earnings multiple in their case) is assumed to be zero in the simplest version of the SOP. In its extended version, it is predicted using bivariate regressions on a set of predictors from Goyal and Welch (2008).[6]

3.3.3 Forecasting with the SOPWAV method

We adopt a similar SOP approach to forecast stock market returns as given by equation (8). However, to forecast each component of the expected stock market return, we directly use the frequency-decomposed time series of the dividend-price ratio, of dividend growth and of a set of predictors from Ferreira and Santa-Clara (2011). In particular, we want to find and use the crystals of *dp* and *gd* that best forecast $\hat{\mu}_s^{dp}$ and $\hat{\mu}_s^{gd}$, respectively, as well as the set of crystals among the other predictors that best forecast $\hat{\mu}_s^{gm}$.

We begin our analysis by first applying the MODWT MRA decomposition to the time series of all predictors. Overall, we decompose 15 time series into $J = 7$ levels, so the total number of crystals that can be used as potential predictors is $15 \times (7 + 1) = 120$. We note that the amount of information we are using is exactly the same as in the standard time series analysis. This is because the obtained crystals for each predictor

[6] See Ferreira and Santa-Clara (2011, tables 2 and 3) for details.

are orthogonal, *i.e.* there is no overlap between them (so that their sum gives exactly the original time series of the corresponding predictor). Furthermore, as the MODWT MRA, at any given point in time, uses information of neighboring data points (both past and future), we recompute the crystals at each iteration of the OOS forecasting process to make sure that we only use current and past information when making the forecasts. Hence, our proposed SOPWAV method does not suffer from a look-ahead bias.

To make the forecast with the SOPWAV method, we assume that the time-s SOPWAV stock return forecast for period *s+1* is the sum of the time-s observations of some crystals. In particular, the forecasting equation specification of the SOPWAV method is given by

$$\hat{\mu}_s = \sum_{i=1}^{NC} crystal_{i,s} \quad , \tag{9}$$

where NC is the number of crystals that maximize the R^2_{OS} (for the entire OOS period) and $crystal_i$, $i = 1, \ldots, NC$ are the crystals selected.

In principle, we would like to search for the combination of crystals (among the 120 available) that delivers the highest possible R^2_{OS}. However, as this is computationally too intensive (given the large number of possible combinations of predictors), we implement the following sequential scheme selection. In the first step, we search for the crystal that gives the highest possible R^2_{OS}. We then save it and remove it from the list of possible predictors to use in the next step. In the second step, we sum the chosen crystal with all the other (119) crystals – chosen one at a time – and select the combination of crystals (the first one plus one among the remaining 119) that delivers the highest R^2_{OS}. We save these two crystals and repeat the same procedure until we use all the 120 crystals. As we show in section 4, the resulting R^2_{OS} function is an inverted U-shaped function of the number of crystals used. We can thus identify the (number of) crystals that maximize the R^2_{OS}. Beyond that point, adding more crystals leads to a deterioration of the forecasting performance as the information provided by these additional predictors adds noise to the forecast.[7]

3.4 Asset allocation analysis

To quantify the economic value of the SOPWAV method from an asset allocation perspective, we follow, among others, Kandel and Stambaugh (1996), Campbell and Thompson (2008) and Rapach, Ringgenberg

[7] As a robustness exercise, we estimate by OLS, for each crystal, an AR(1) process, *e.g.* $dp\,(S_7)_t = \hat{\alpha} + \hat{\beta} dp\,(S_7)_{t-1}$. We then make its one-step-ahead forecast as $E_t dp\,(S_7)_{t+1} = \hat{\alpha} + \hat{\beta} dp\,(S_7)_t$, and use the results to compute $\hat{\mu}_s$ (while our procedure implicitly assumes that $\hat{\alpha} = 0$ and $\hat{\beta} = 1$). Results are similar and available upon request.

and Zhou (2016), and consider a mean-variance investor who dynamically allocates her wealth between equities and risk-free bills. The asset allocation decision is made at the end of month s, and the optimal share allocated to equities during month *s+1* is given by

$$w_s \quad = \quad \frac{1}{\gamma}\frac{\hat{\mu}_s - rf_{s+1}}{\hat{\sigma}^2_{s+1}} \quad , \tag{10}$$

where rf_{s+1} denotes the risk-free return from time s to $s+1$ (*i.e.* the market rate, which is known at time s), γ is the investor's relative risk aversion coefficient, $\hat{\mu}_s$ is the SOPWAV predicted stock market return at time s for period *s+1*, and $\hat{\sigma}^2_{s+1}$ is the forecast of the variance of the excess return (difference between the stock market return and the risk-free return). As in Rapach, Ringgenberg and Zhou (2016), we assume a relative risk aversion coefficient of three, use a ten-year moving window of past excess returns to estimate the variance forecast and impose portfolio constraints by restricting the weights w_s to lie between -0.5 and 1.5. These constraints introduce realistic limits on the possibilities of short selling and leveraging the portfolio.[8]

The portfolio return at time *s+1*, rp_{s+1}, is given by $rp_{s+1} = w_s r_{s+1} + (1 - w_s)\, rf_{s+1}$. We first evaluate the average utility (or certainty equivalent return, CER) of an investor that uses the allocation rule (10). The CER 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 then report the average utility gain, which is defined as the difference between the CER for an investor that uses the SOPWAV method to forecast stock returns and the CER for an investor who uses the HM forecasting strategy. The CER gain is annualized and can be interpreted as the annual management fee that an investor would be willing to pay to be exposed to a trading strategy based on the SOPWAV method instead of one based on the HM forecast (Rapach, Ringgenberg and Zhou, 2016). We also compute the average monthly turnover (the percentage of wealth traded each month) and report the relative average turnover, calculated as the average monthly turnover of the SOPWAV-based portfolio divided by the average monthly turnover of the HM-based portfolio. Finally, we introduce transaction costs into the analysis, calculated using the monthly turnover measures and assuming a proportional cost of 50 basis points per transaction (as in *e.g.* Balduzzi and Lynch, 1999 and Huang et al., 2015).

[8] We obtain similar results using values from 1 to 5 for the relative risk aversion coefficient (as in Huang et al., 2015), and a five-year moving window of past excess returns or all available data up to time s (as in Huang et al., 2015 and Ferreira and Santa-Clara, 2011, respectively) to estimate the variance forecast.

4 Out-of-sample forecasting results

As figure 1 shows, the OOS forecast of the SOPWAV method follows the dynamics of effective stock market returns much more closely than the HM forecast. We use the R^2_{OS} statistic to evaluate the statistical performance of the SOPWAV method.

Figure 3 shows the R^2_{OS} as a function of the number of crystals used. With only one crystal, the R^2_{OS} is negative, so the HM still outperforms the SOPWAV method. However, the R^2_{OS} becomes positive if at least two crystals are included in the forecast exercise. The maximum monthly R^2_{OS} attainable with the SOPWAV method is 6.18 % using 24 crystals.[9] Moreover, the R^2_{OS} is an inverted U-shaped function of the number of crystals used as predictors. From a statistical point of view, this means that there are forecasting gains from using more crystals as predictors, but those gains are decreasing and at some point become negative. Convenience also dictates keeping the number of predictive variables as low as possible in the absence of any clear economic (not just statistical) justification for incorporating more variables. For these reasons, our proposed baseline SOPWAV method only uses the 4 most important crystals. These yield a monthly R^2_{OS} of 3.27 %. Despite the difference of 2.9 % between the R^2_{OS} with 24 crystals and with 4 crystals, in section 5 we show that the economic gains of using 24 crystals are not significantly larger than those using 4 crystals. In the subsequent analysis, we still report the results of the SOPWAV method with 24 crystals (SOPWAV$^+$) to document the maximum attainable performance with the SOPWAV method.

The 4 most important crystals are the smooth wavelet of the dividend-price ratio ($dp\,(S_7)$), the smooth wavelet of dividends growth ($gd\,(S_7)$), the second wavelet detail of the return on equity ($roe\,(D_2)$) and the fifth wavelet detail of the long-term bond return ($ltr\,(D_5)$). Remarkably, even without imposing any specific crystal (or set of crystals), two of the most important crystals turn out to be related with *dp* and *gd*, which are the two key variables in the SOP method. Furthermore, and similarly to Ferreira and Santa-Clara (2011), we find that the low-frequency component of *gd* is quite important in forecasting stock returns.[10]

The baseline SOPWAV method forecast at time *s* for the stock market return at time *s+1* is therefore given by

$$\hat{\mu}_s \quad = \hat{\mu}_s^{dp} + \hat{\mu}_s^{gd} + [\hat{\mu}_s^{gm}] = \quad dp\,(S_7)_s + gd\,(S_7)_s + [roe\,(D_2)_s + ltr\,(D_5)_s] \quad , \tag{11}$$

[9] See Appendix 3 for the list of crystals.

[10] The combination of 4 crystals (among the 8 millions possible combinations) that delivers the highest R^2_{OS} ($R^2_{OS} = 3.29\%$) is $dp\,(S_7)$, $gd\,(S_7)$, $roe\,(D_3)$ and $ltr\,(D_5)$, which is very similar to the combination obtained under the adopted sequential scheme selection.

that is, we forecast i) the expected dividend-price ratio $\hat{\mu}_s^{dp}$ and expected dividends growth $\hat{\mu}_s^{gd}$ using the lowest frequency components of the dividend-price ratio ($dp\,(S_7)$) and dividends growth ($gd\,(S_7)$), respectively, and ii) the price-dividends multiple growth $\hat{\mu}_s^{gm}$ using a high frequency of the return on equity ($roe\,(D_2)$) and a medium frequency of the long-term government bond return ($ltr\,(D_5)$).

Table 1 and figure 4 present the summary statistics of the 4 predictors of the baseline SOPWAV method and their time series dynamics, respectively. From the correlation coefficients (panel B of table 1), it is clear that these predictors carry complementary information for the OOS forecast. Except for the pair $gd\,(S_7)$ and $dp\,(S_7)$, the correlation coefficients are indeed around zero or slightly negative. Moreover, the graphs in figure 4 suggest that the predictors $gd\,(S_7)$ and $dp\,(S_7)$ capture the long-term movements of the stock market returns, while $roe\,(D_2)$ and $ltr\,(D_5)$ capture the higher frequencies of the stock returns dynamics. The wavelet decomposition thus enables the SOPWAV method to capture both the low-frequency dynamics of the stock market returns (as the HM) and some of the higher frequency turbulence that characterizes the dynamics of stock market returns.

Table 2 presents the statistical results of the forecasting performance of the SOPWAV method versus the HM, for the entire OOS period (second column) and for different subsamples (columns 3 to 5), corresponding to the Great Inflation period (1950:01-1983:12), the Great Moderation period (1984:01-2006:12) and the Global Financial Crisis period and its aftermath (2007:01-2015:12).[11] Several results stand out. First, for the full OOS period, the positive and statistically significant R^2_{OS} of 3.27 % indicates that the SOPWAV method outperforms the HM benchmark. Second, the superior performance of the SOPWAV method is robust throughout the sample period, *i.e.* the $R^2_{OS}s$ for all subsamples are always positive, statistically significant and range from 2.48 % to 4.15 %. Moreover, the $R^2_{OS}s$ for the Great Inflation and the Global Financial Crisis periods are higher than for the Great Moderation period, which is consistent with the findings in the literature that return predictability is higher in subsamples with deep recessions (see *e.g.* Henkel, Martin and Nardari, 2011, Huang et al., 2015 and Rapach, Ringgenberg and Zhou, 2016). Third, when comparing the $R^2_{OS}s$ of the SOPWAV and the SOPWAV$^+$ methods, it is clear that the statistical performance improves substantially by using more (24) crystals, with the improvement most apparent in the last subsample. However, as we show in the next section, these statistical gains are not economically so expressive.

[11] McConnell and Perez-Quiros (2000) document a structural decline in the volatility of real US GDP growth in the first quarter of 1984. This supports the break of the time windows for our first two subsamples.

5 Asset allocation

In the previous section, we demonstrated that the SOPWAV method delivers statistically significant forecasting gains. We now quantify the economic value of the SOPWAV method for stock return forecasting from an asset allocation perspective.

Columns 2 through 4 in table 3 report the results of the CER analysis. The baseline SOPWAV method clearly outperforms the HM benchmark. For the full OOS period (panel A), the annualized CER gain for an investor who trades using the baseline SOPWAV method is 403 basis points (before transaction costs). Following a trading strategy based on the baseline SOPWAV method implies trading, on average, about seven times more than using a trading strategy based on the HM forecast. Nonetheless, net of transaction costs, the CER gain still stands at 282 basis points. Looking at the subsample periods (panels B, C and D), the CER gains tend to increase over time (334 basis points in the Great Inflation period rising to 406 basis points during the Great Moderation period and to 652 basis points during the Global Financial Crisis period). Except for the last OOS period, the CER gains obtained with the SOPWAV$^+$ method do not seem to be noticeably higher (and are actually lower in some cases) than the CER gains with the baseline SOPWAV method. So, although the R^2_{OS} of the SOPWAV$^+$ method is higher than the baseline SOPWAV R^2_{OS}, the economic advantage of using more variables (24 rather than 4) is quite modest.

We also compute the Sharpe ratio (SR) of the portfolio, *i.e.* the mean portfolio return in excess of the risk-free rate divided by the standard deviation of the excess portfolio return. Annualized Sharpe ratios are reported in columns 5 to 8 in table 3. We report the absolute value of the SOPWAV SR, as well as the ratio between the SOPWAV SR and the HM SR (before and after transaction costs). These results further confirm that the SOPWAV method clearly outperforms the HM benchmark. For the full OOS period (panel A), the annualized SRs of the strategy based on the baseline SOPWAV method are 0.65 and 0.54 before and after transaction costs, respectively. This compares with SRs of 0.35 and 0.32 for the strategy based on the HM, respectively (not reported in the table). While the SOPWAV's superior performance stands for all subsample periods (panels B, C and D), it is particularly evident during the Great Moderation and Global Financial Crisis periods, when the SRs delivered by the baseline SOPWAV strategy are 3.2 and 2.3 times larger than those of the HM, respectively. As with the CER gains analysis, the only period when the SOPWAV$^+$ clearly outperforms the baseline SOPWAV in terms of SR is during and after the Global Financial Crisis.

So far we have assumed that transaction costs are 50 basis points per transaction, which is generally considered to be a relatively high number. In fact, according to Balduzzi and Lynch (1999), 50 basis points might be a realistic value for an investor who trades in individual stocks directly, while a value of 2 basis points might be appropriate for an investor who trades futures contracts on the S&P500 index. We thus evaluate the sensitivity of our results when transaction costs are reduced up to 0 basis points. Figure 5 illustrates the sensitivity of the CER gains (left graph) and the SRs (right graph) with respect to transaction costs, for the full OOS period and for three scenarios involving portfolio constraints: the baseline case ($-0.5 \leq w \leq 1.5$), a scenario with no short sales ($0 \leq w \leq 1.5$, as in Neely et al., 2014 and Huang et al., 2015) and one with no restrictions on weights (as in Ferreira and Santa-Clara, 2011). As expected, given a set of portfolio constraints, the lower the transaction costs, the higher the CER gains and the SRs. More importantly, the effective portfolio constraints to which the investor is exposed to play a major role. In fact, regardless of the magnitude of the transaction costs, the scenario where there are no portfolio constraints (blue lines) always dominates the scenario where constrained leverage and short selling is allowed (black lines), which in turn dominates the scenario where short selling is not allowed (red lines).

Finally, figure 6 provides, for the full OOS period, a dynamic perspective of the portfolio and cumulative wealth for an investor using the SOPWAV method or the HM forecast. Panel A presents the equity weights for the SOPWAV and HM portfolios. In the baseline scenario, the equity weight is constrained to lie between -0.5 and 1.5 (black solid line for the SOPWAV method and black dashed line for the HM), while the blue line represents the equity weight for an investor using the SOPWAV method without portfolio constraints. The first result that stands out is the substantially different dynamics of the equity exposure between the SOPWAV portfolios and the HM portfolio. The SOPWAV method implies significant and frequent changes in equity weights, whereas under the HM portfolio, the equity exposure changes much more smoothly. Interestingly, during the two recessions in the 2000s, the equity exposure of the HM portfolio tracks a slightly upward trend, while for the SOPWAV method the equity exposure is strongly reduced at the beginning of the recessions (including short selling) and the average exposure during the recession period is lower than in the HM portfolio.

The second notable result from panel A is that the constraints on the weights ($-0.5 \leq w \leq 1.5$) are strongly binding for the SOPWAV portfolio, yet they are almost irrelevant for the HM portfolio. Limits on leveraging the portfolio are particularly effective during almost the entire 1950s, immediately before the 2000 recession and after the 2008 recession. Limits on short selling are particularly binding during the mid-1970s recession, around the 1987 crash, and during and after both recessions in the 2000s. In fact, in these specific periods, the unconstrained SOPWAV method implies levels of short selling and leverage sometimes exceeding 500%. This seems somewhat unrealistic as regards its implementation, and supports our decision to impose weight constraints in the baseline scenario.

Panel B in figure 6 shows the log cumulative wealth for an investor with the HM portfolio (black dashed line) and with the SOPWAV portfolio when the equity weights are constrained to lie between -0.5 and 1.5 (black solid line), when short selling is not possible (red line) and when there are no equity holdings constraints (blue line). In the simulation, we assume that the investor begins with \$1 and reinvests all proceeds. Comparing the HM and the baseline SOPWAV portfolios (black lines), there is a continuous divergence of the cumulative wealth of the investors under these alternative portfolios that clearly favors the SOPWAV approach. The much higher rotation of the SOPWAV portfolio, as illustrated in panel A, reflects an enhanced market timing of this strategy. The divergence on the cumulative wealth between the two portfolios is particularly strong during the 1950s and since the early 2000s. During the 1950s, the SOPWAV portfolio benefits from a leveraged exposure to the equity market, while during the 2000s the HM-based portfolio is negatively affected by two severe drawdowns (the first occurring around the recession at the beginning of the 2000s and the second during the 2008-2009 crisis). The cumulative wealth of the investor with an HM-based strategy, in fact, only returns to the level of the early 2000s in 2014. In contrast, an investor adopting a SOPWAV-based strategy quickly recovers from the (smaller) drawdowns suffered during the 2000s recessions and benefits from a strong upward trend in her wealth until the end of the sample period.

This figure complements the results from figure 5 on the economic impact of portfolio constraints. In particular, the cumulative wealth of an investor trading with the SOPWAV method with no constraints (blue line) is clearly higher than for cases where some constraints are imposed (black and red lines). The possibility of freely leveraging the portfolio enables a significant increase in cumulative wealth during the 1950s. The same happens during and after the two recessions in the 2000s, when leveraging and shorting freely bring significant wealth gains. This figure also shows that there is only a small negative wealth impact from impeding short selling as compared to a scenario of limited short selling.

6 Concluding remarks

This paper proposes a new method for forecasting stock market returns (SOPWAV) that is easy to implement and parameter free (*i.e.* no parameters to calibrate or estimate). The method involves, within the Ferreira and Santa-Clara (2011) sum-of-the-parts framework of stock return forecasts, a wavelet decomposition of several predictors of stock market returns. We forecast out-of-sample stock market returns (S&P500 index) for the period running from January 1950 to December 2015, as well as for subsamples corresponding to the periods of the Great Inflation, the Great Moderation and the Global Financial Crisis. No matter which sample period is used, the SOPWAV method delivers statistically and economically significant gains for investors that clearly outperform the Historical Mean (HM) benchmark. The results are also robust to the introduction of transaction costs and different sets of portfolio constraints.

The strong performance of the SOPWAV method is mainly due to two reasons. First, the use of wavelet decomposition makes it possible to isolate the frequencies of the predictors that have the greatest predictive power from the noisier parts. Second, the frequency-decomposed predictors carry complementary information for the forecasting exercise that captures both the long-term trend and the higher frequency movements of stock market returns. As a result, the SOPWAV method anticipates well the future dynamics of stock market returns.

References

Aguiar-Conraria, Luis, and Maria Joana Soares. 2011. "Business cycle synchronization and the Euro: A wavelet analysis." *Journal of Macroeconomics*, 33(3): 477–489.

Aguiar-Conraria, Luis, and Maria Joana Soares. 2014. "The Continuous Wavelet Transform: Moving Beyond Uni- And Bivariate Analysis." *Journal of Economic Surveys*, 28(2): 344–375.

Aguiar-Conraria, Luis, Manuel M.F. Martins, and Maria Joana Soares. 2012. "The yield curve and the macro-economy across time and frequencies." *Journal of Economic Dynamics and Control*, 36(12): 1950–1970.

Ang, Andrew, and Geert Bekaert. 2007. "Stock Return Predictability: Is it There?" *Review of Financial Studies*, 20(3): 651–707.

Baetje, Fabian, and Lukas Menkhoff. 2016. "Equity premium prediction: Are economic and technical indicators unstable?" *International Journal of Forecasting*, 32(4): 1193–1207.

Balduzzi, P., and A.W. Lynch. 1999. "Transaction costs and predictability: some utility cost calculations." *Journal of Financial Economics*, 52(1): 47–78.

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.

Bekaert, Geert, and Robert J. Hodrick. 1992. "Characterizing Predictable Components in Excess Returns on Equity and Foreign Exchange Markets." *The Journal of Finance*, 47(2): 467–509.

Bollerslev, Tim, George Tauchen, and Hao Zhou. 2009. "Expected Stock Returns and Variance Risk Premia." *Review of Financial Studies*, 22(11): 4463–4492.

Bollerslev, Tim, Viktor Todorov, and Lai Xu. 2015. "Tail risk premia and return predictability." *Journal of Financial Economics*, 118(1): 113–134.

Campbell, John Y. 1987. "Stock returns and the term structure." *Journal of Financial Economics*, 18(2): 373–399.

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

Cochrane, John H. 2008. "The Dog That Did Not Bark: A Defense of Return Predictability." *Review of Financial Studies*, 21(4): 1533–1575.

Cooper, Ilan, and Richard Priestley. 2009. "Time-Varying Risk Premiums and the Output Gap." *Review of Financial Studies*, 22(7): 2601–2633.

Cooper, Ilan, and Richard Priestley. 2013. "The World Business Cycle and Expected Returns." *Review of Finance*, 17(3): 1029–1064.

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

Cutler, David M., James M. Poterba, and Lawrence H. Summers. 1991. "Speculative Dynamics." *Review of Economic Studies*, 58(3): 529–546.

Dangl, Thomas, and Michael Halling. 2012. "Predictive regressions with time-varying coefficients." *Journal of Financial Economics*, 106(1): 157–181.

Fama, Eugene F., and G. Schwert. 1977. "Asset returns and inflation." *Journal of Financial Economics*, 5(2): 115–146.

Ferreira, Miguel A., and Pedro 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.

Ferson, Wayne, and Campbell Harvey. 1991. "The Variation of Economic Risk Premiums." *Journal of Political Economy*, 99(2): 385–415.

Ferson, Wayne E., and Campbell R. Harvey. 1993. "The Risk and Predictability of International Equity Returns." *Review of Financial Studies*, 6(3): 527–566.

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, Marco, and James B. Ramsey. 2013. "Bond vs stock market's Q: Testing for stability across frequencies and over time." *Journal of Empirical Finance*, 24(C): 138–150.

Gallegati, Marco, Mauro Gallegati, James B. Ramsey, and Willi Semmler. 2011. "The US Wage Phillips Curve across Frequencies and over Time." *Oxford Bulletin of Economics and Statistics*, 73(4): 489–508.

Gencay, Ramazan, Faruk Selcuk, and Brandon Whitcher. 2005. "Multiscale systematic risk." *Journal of International Money and Finance*, 24(1): 55–70.

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.

Harvey, Campbell R. 1991. "The World Price of Covariance Risk." *The Journal of Finance*, 46(1): 111–157.

Harvey, Campbell R., Yan Liu, and Heqing Zhu. 2016. "... and the Cross-Section of Expected Returns." *Review of Financial Studies*, 29(1): 5–68.

Henkel, Sam James, J. Spencer Martin, and Federico Nardari. 2011. "Time-varying short-horizon predictability." *Journal of Financial Economics*, 99(3): 560–580.

Hjalmarsson, Erik. 2010. "Predicting Global Stock Returns." *Journal of Financial and Quantitative Analysis*, 45: 49–80.

Huang, Dashan, Fuwei Jiang, Jun Tu, and Guofu Zhou. 2015. "Investor Sentiment Aligned: A Powerful Predictor of Stock Returns." *Review of Financial Studies*, 28(3): 791–837.

Kandel, Shmuel, and Robert F Stambaugh. 1996. "On the Predictability of Stock Returns: An Asset-Allocation Perspective." *Journal of Finance*, 51(2): 385–424.

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

Kim, Sangbae, and Francis In. 2005. "The relationship between stock returns and inflation: new evidence from wavelet analysis." *Journal of Empirical Finance*, 12(3): 435–444.

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

Li, Yan, David T. Ng, and Bhaskaran Swaminathan. 2013. "Predicting market returns using aggregate implied cost of capital." *Journal of Financial Economics*, 110(2): 419–436.

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

Malagon, J., D. Moreno, and R. Rodriguez. 2015. "Time horizon trading and the idiosyncratic risk puzzle." *Quantitative Finance*, 15(2): 327–343.

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

McConnell, M., and G. Perez-Quiros. 2000. "Output Fluctuations in the United States: What Has Changed since the Early 1980's?" *American Economic Review*, 90(5): 1464–1476.

McCracken, M.W. 2007. "Asymptotics for out of sample tests of Granger causality." *Journal of Econometrics*, 140(2): 719 – 752.

Mele, Antonio. 2007. "Asymmetric stock market volatility and the cyclical behavior of expected returns." *Journal of Financial Economics*, 86(2): 446–478.

Moller, Stig V., and Jesper Rangvid. 2015. "End-of-the-year economic growth and time-varying expected returns." *Journal of Financial Economics*, 115(1): 136 – 154.

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.

Pastor, Lubos, and Robert Stambaugh. 2009. "Predictive Systems: Living with Imperfect Predictors." *The Journal of Finance*, 64(4): 1583–1628.

Percival, DB, and AT Walden. 2000. *Wavelet methods for time series analysis.* Cambridge University Press.

Pettenuzzo, Davide, Allan Timmermann, and Rossen Valkanov. 2014. "Forecasting stock returns under economic constraints." *Journal of Financial Economics*, 114(3): 517–553.

Ramsey, James B., and Camille Lampart. 1998*a*. "The Decomposition of Economic Relationships by Time Scale Using Wavelets: Expenditure and Income." *Studies in Nonlinear Dynamics & Econometrics*, 3(1): 1–22.

Ramsey, James B., and Camille Lampart. 1998*b*. "Decomposition Of Economic Relationships By Timescale Using Wavelets: Money and Income." *Macroeconomic Dynamics*, 2(1): 49–71.

Rapach, David, and Guofu Zhou. 2013. "Forecasting Stock Returns." Vol. 2 of *Handbook of Economic Forecasting*, 328–383. Elsevier.

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

Rapach, David E., Jack K. Strauss, and Guofu Zhou. 2013. "International Stock Return Predictability: What Is the Role of the United States?" *Journal of Finance*, 68(4): 1633–1662.

Rapach, David E., Matthew C. Ringgenberg, and Guofu Zhou. 2016. "Short interest and aggregate stock returns." *Journal of Financial Economics*, 121(1): 46 – 65.

Rua, Antonio. 2010. "Measuring comovement in the time-frequency space." *Journal of Macroeconomics*, 32(2): 685–691.

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

Rua, Antonio, and Luis C. Nunes. 2009. "International comovement of stock market returns: A wavelet analysis." *Journal of Empirical Finance*, 16(4): 632–639.

Xue, Yi, Ramazan Gencay, and Stephen Fagan. 2013. "Jump detection with wavelets for high-frequency financial time series." *Quantitative Finance*, 14(8): 1427–1444.

Panel A: Univariate statistics								
Return predictors	Mean	Median	Standard deviation	Min	Max	Skewness	Kurtosis	AR(1)
$dp\,(S_7)$	0.29	0.28	0.10	0.13	0.49	0.22	2.15	1.00
$gd\,(S_7)$	0.47	0.47	0.18	0.13	1.08	0.41	3.31	1.00
$roe\,(D_2)$	0.00	0.02	0.36	-2.78	3.31	0.23	30.65	0.70
$ltr\,(D_5)$	0.00	-0.02	0.40	-1.36	1.46	-0.01	3.85	0.90
r	0.88	1.20	4.17	-24.25	15.51	-0.65	5.43	0.05

Panel B: Correlations					
Return predictors	$dp\,(S_7)$	$gd\,(S_7)$	$roe\,(D_2)$	$ltr\,(D_5)$	r
$dp\,(S_7)$	1.00				
$gd\,(S_7)$	0.55	1.00			
$roe\,(D_2)$	0.00	-0.12	1.00		
$ltr\,(D_5)$	0.02	0.01	-0.17	1.00	
r	0.09	0.05	0.11	0.10	1.00

Table 1: Stock market return predictors

Panel A reports the mean, median, standard deviation, minimum, maximum, skewness, kurtosis, and first-order autocorrelation coefficient of the predictors of stocks market returns as given by equation (11). $dp\,(S_7)$ is the wavelet smooth of the dividend-price ratio, $gd\,(S_7)$ is the wavelet smooth of the dividend growth, $roe\,(D_2)$ is the second wavelet detail of the return on equity, $ltr\,(D_5)$ is the fifth wavelet detail of the long-term government bond return and r is the stock market return. Panel B reports the correlations between the predictors and between each predictor and stock market returns. The $J = 7$ level wavelet decomposition of each potential predictor implies their decomposition into eight orthogonal crystals, *i.e.* seven wavelet details (D_1, $D_2, \ldots$, D_7) representing the higher-frequency characteristics of the predictor, and a wavelet smooth (S_7) that captures the low-frequency dynamics of the predictor. See section 3.2 and Appendix 2 for technical details on wavelet decomposition. The sample period runs from 1950:01 to 2015:12, monthly frequency.

Method	Sample period			
	1950:01 2015:12	1950:01 1983:12	1984:01 2006:12	2007:01 2015:12
SOPWAV	3.27***	3.60***	2.48***	4.15***
SOPWAV$^+$	6.18***	6.78***	2.47***	12.79***

Table 2: Out-of-sample R-squares $\left(R^2_{OS}\right)$

This table reports the out-of-sample R-squares (in percentages) for stock market return forecasts at monthly (nonoverlapping) frequencies from the baseline SOPWAV method specification (equation (11)) and from the SOPWAV$^+$ method specification (equation (9)) using the 24 crystals listed in Appendix 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 method relative to the forecast based on the historical mean (HM). The one-month-ahead out-of-sample forecast of stock market return is generated using a sequence of expanding windows. The sample period is from 1927:01 to 2015:12. The full out-of-sample forecasting period runs from 1950:01 to 2015:12, with R^2_{OS} reported in the second column. The R^2_{OS} for three out-of-sample forecasting subperiods are reported in the remaining columns. Asterisks denote significance of the out-of-sample *MSFE-F* statistic of McCracken (2007). *** denotes significance at the 1 % level.

Method	CER gain (%) no costs	Relative average turnover	CER gain (%) cost = 50 bps	SR no costs absolute	SR no costs relative	SR cost = 50 bps absolute	SR cost = 50 bps relative
			Panel A. Sample period: 1950:01-2015:12				
SOPWAV	4.03	6.7	2.82	0.65	1.8	0.54	1.7
SOPWAV^+	4.68	8.7	3.02	0.67	1.9	0.54	1.7
			Panel B. Sample period: 1950:01-1983:12				
SOPWAV	3.34	4.3	2.44	0.67	1.3	0.59	1.2
SOPWAV^+	3.94	7.2	2.22	0.70	1.4	0.56	1.2
			Panel C. Sample period: 1984:01-2006:12				
SOPWAV	4.06	10.6	2.56	0.56	3.2	0.42	2.7
SOPWAV^+	3.59	11.5	1.94	0.51	2.9	0.37	2.4
			Panel D. Sample period: 2007:01-2015:12				
SOPWAV	6.52	14.5	4.90	0.73	2.3	0.62	2.1
SOPWAV^+	10.21	13.2	8.74	0.94	3.0	0.85	2.8

Table 3: Out-of-sample CER gains and Sharpe ratios

The second column reports the annualized certainty equivalent return (CER) percentage gain for an investor allocating her wealth between equities and risk-free bills according to the rule (10), using stock return forecasts from the baseline SOPWAV specification (equation (11)) and from the SOPWAV^+ specification using the 24 crystals listed in Appendix 3. The third column shows the relative average turnover, calculated as the average monthly turnover of the SOPWAV and SOPWAV^+ -based portfolios divided by the average monthly turnover of the HM-based portfolio. The fourth column reports the CER gain after transaction costs. Columns 5 to 8 give the annualized Sharpe ratios (SR). Columns 5 and 7 present the absolute SR without and with transactions costs, respectively. Columns 6 and 8 present the ratio of the SR using the SOPWAV (or the SOPWAV^+) method and the SR using forecasts based on the HM without and with transactions costs, respectively. The investor is assumed to have a relative risk aversion coefficient of three and the equity weight in the portfolio is constrained to lie between -0.5 and 1.5. Transaction costs are calculated using the monthly turnover measures and assume a proportional cost equal to 50 basis points per transaction. Panel A reports the results for the full out-of-sample period (1950:01 to 2015:12) and panels B, C and D report the results for three out-of-sample forecasting subperiods.

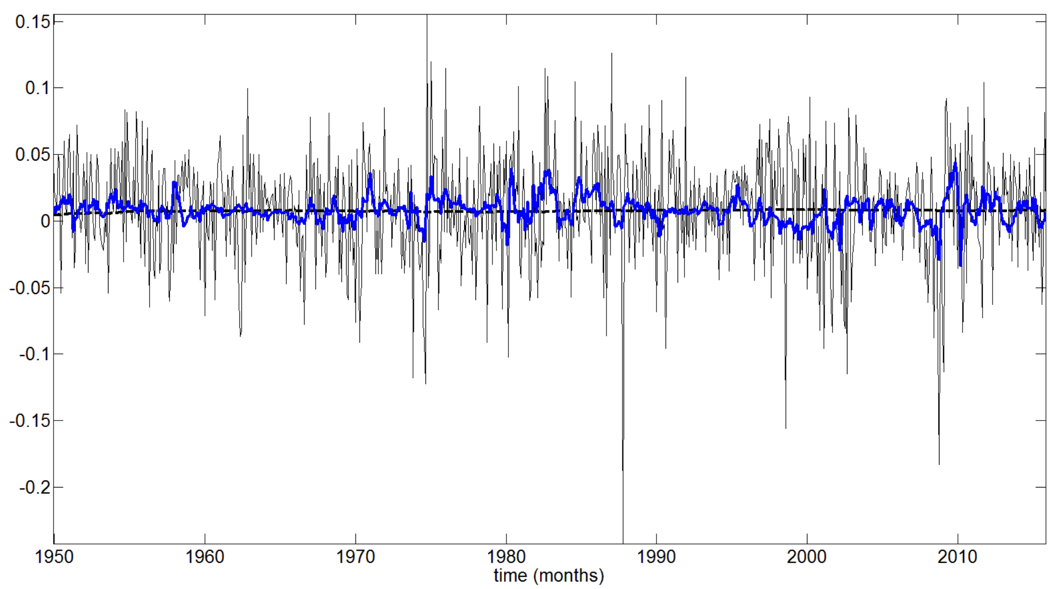


Figure 1: Realized and predicted stock market returns

The black solid line corresponds to the (log) realized stock market return as proxied by the (log) S&P 500 index return. The blue line plots the one-month ahead out-of-sample predictive regression forecast for the stock market return based on the SOPWAV method. The black dashed line represents the one-month-ahead out-of-sample stock market return forecast based on the historical mean (HM) of returns. The sample period runs from 1950:01 to 2015:12, monthly frequency.

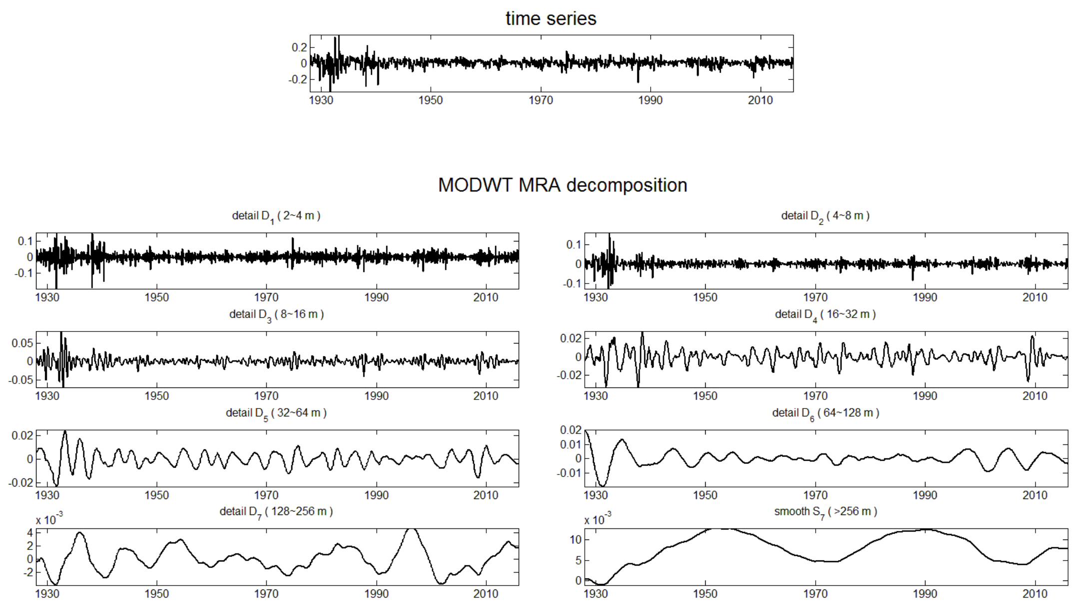


Figure 2: Stock market return, time series and MODWT MRA decomposition

The time series of the (log) stock market return as proxied by the S&P 500 index returns is presented in the top graph. The eight orthogonal crystals in which the stock market return time series is decomposed are presented in the remaining graphs. The $J = 7$ levels wavelet decomposition leads to seven wavelet details (D_1, $D_2, \ldots,$ D_7), representing the higher-frequency characteristics of the series, as well as a wavelet smooth (S_7) that captures the low-frequency dynamics of the series. See section 3.2 and Appendix 2 for technical details on the wavelet decomposition. The sample period runs from 1927:01 to 2015:12, monthly frequency.

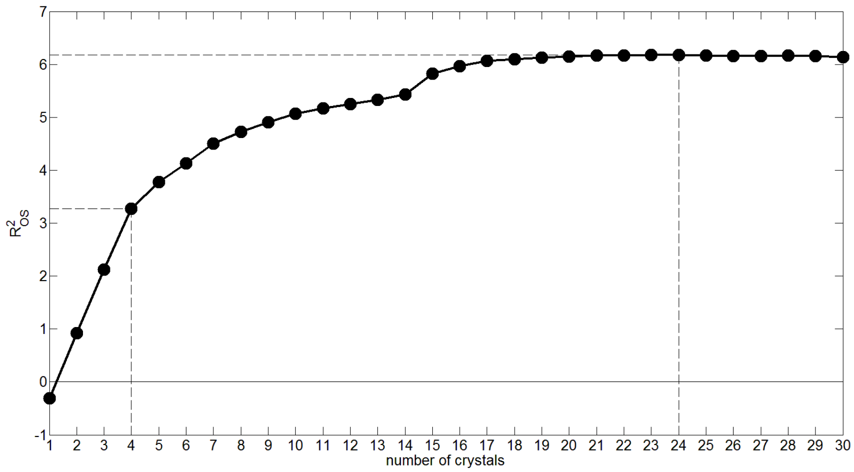


Figure 3: Out-of-sample R-squares of the SOPWAV method versus the HM

This figure reports the out-of-sample R-squares (in percentage) for stock market returns forecasts at monthly (non-overlapping) frequencies from the SOPWAV method (equation (9)) using an increasing number of crystals obtained from the wavelet decomposition of the set of 15 predictors considered (described in Appendix 1). The out-of-sample R-squares measures the proportional reduction in the mean squared forecast error for the predictive SOPWAV method relative to the forecast based on the historical mean (HM). The $J = 7$ levels wavelet decomposition of each of the 15 predictors considered implies their decomposition into eight orthogonal crystals, *i.e.* 120 crystals. See section 3.2 and Appendix 2 for technical details on wavelet decomposition. The out-of-sample period runs from 1950:01 to 2015:12, monthly frequency.

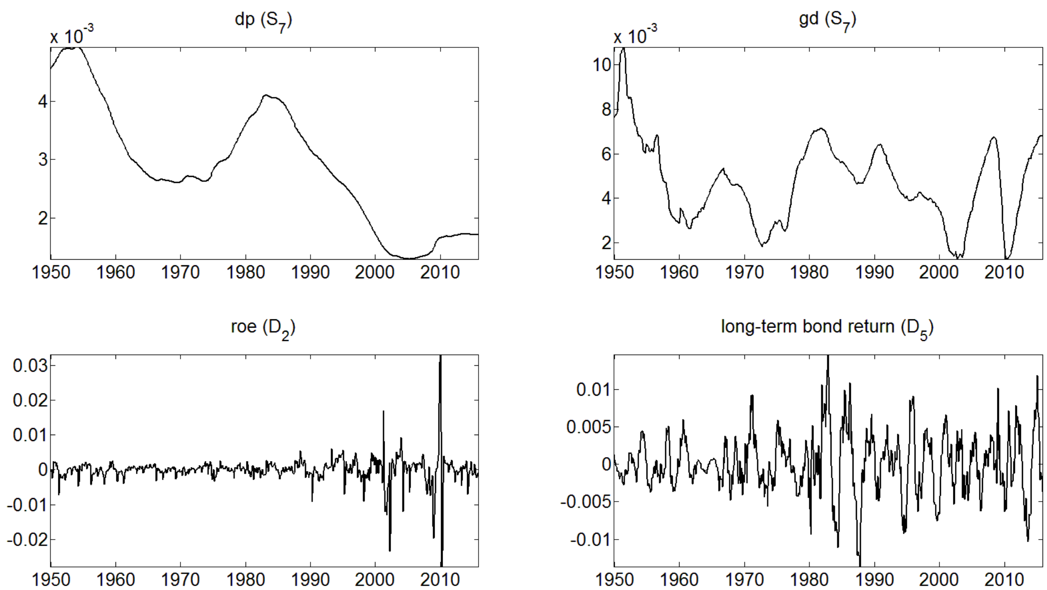


Figure 4: Stock market return predictors dynamics

This figure plots the time series of the 4 predictors used in the baseline SOPWAV method (equation (11)), for the period 1950:01 to 2015:12, monthly frequency. $dp\,(S_7)$ is the wavelet smooth of the dividend-price ratio, $gd\,(S_7)$ the wavelet smooth of dividend growth, $roe\,(D_2)$ the second wavelet detail of return on equity and $ltr\,(D_5)$ is the fifth wavelet detail of long-term government bond return. The $J = 7$ levels wavelet decomposition of each potential predictor considered implies their decomposition into eight orthogonal crystals, *i.e.* seven wavelet details (D_1, $D_2, \ldots,$ D_7) representing the higher-frequency characteristics of the predictor, as well as a wavelet smooth (S_7) that captures the low-frequency dynamics of the predictor. See section 3.2 and Appendix 2 for technical details on wavelet decomposition.

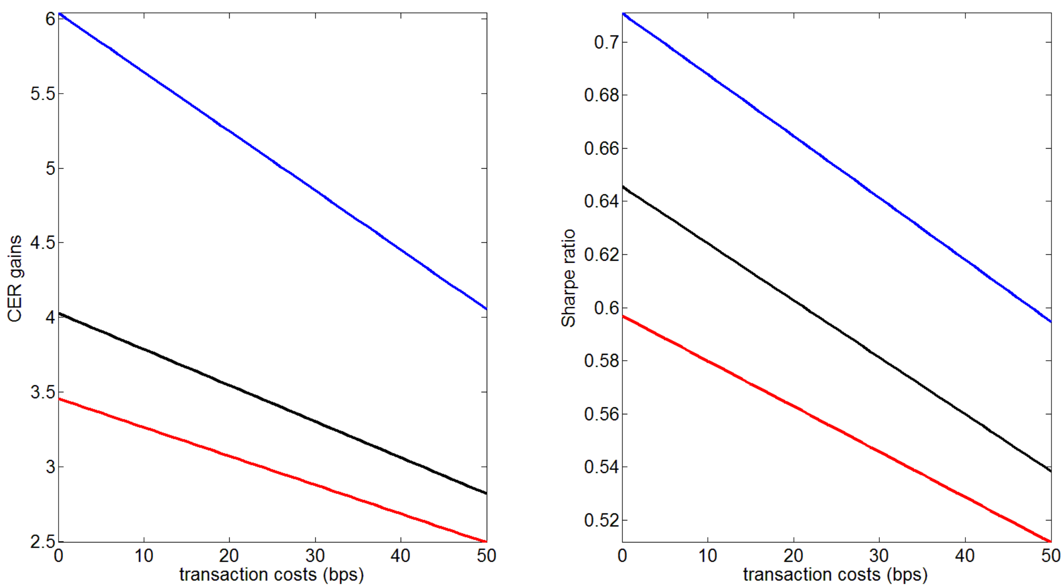
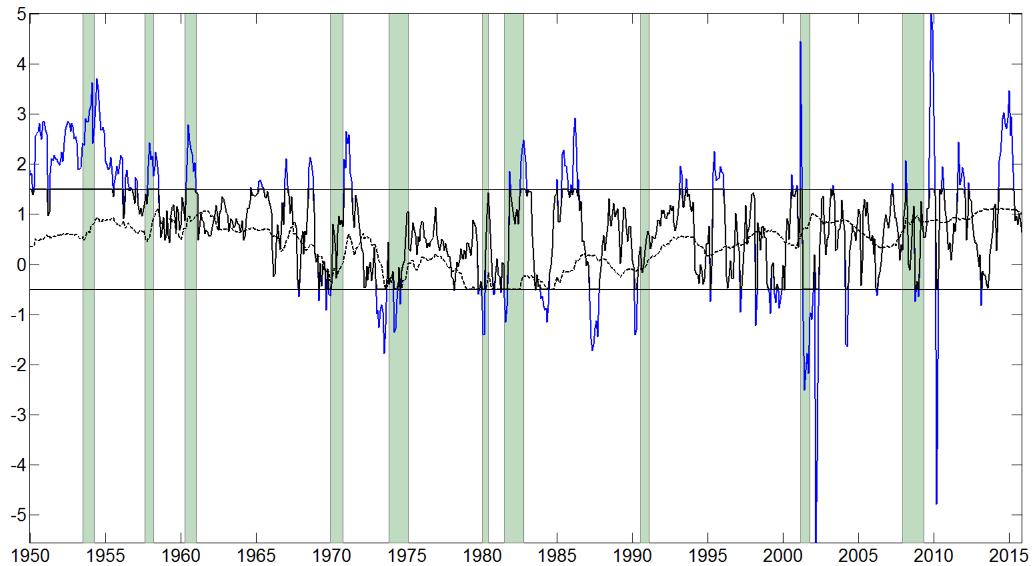


Figure 5: Out-of-sample CER gains and Sharpe ratios - sensitivity analysis
The left and right panels show the annualized certainty equivalent return (CER) percentage gain and the annualized Sharpe ratio (SR) for an investor allocating her wealth between equities and risk-free bills according to the rule (10), using stock return forecasts from the baseline SOPWAV specification (equation (11)). The investor is assumed to have a relative risk aversion coefficient of three. The CER gains and the SR are computed for (i) different levels of transaction costs, calculated using monthly turnover measures and assuming a proportional cost from 0 to 50 basis points per transaction, and (ii) different constraints on equity weights (10), with the blue line representing no restriction, the black line weights constrained to between -0.5 to 1.5, and a red line with weights between 0 and 1.5 (no short sales). The sample period runs from 1950:01 to 2015:12, monthly frequency.

Panel A. Equity weights



Panel B. Log cumulative wealth

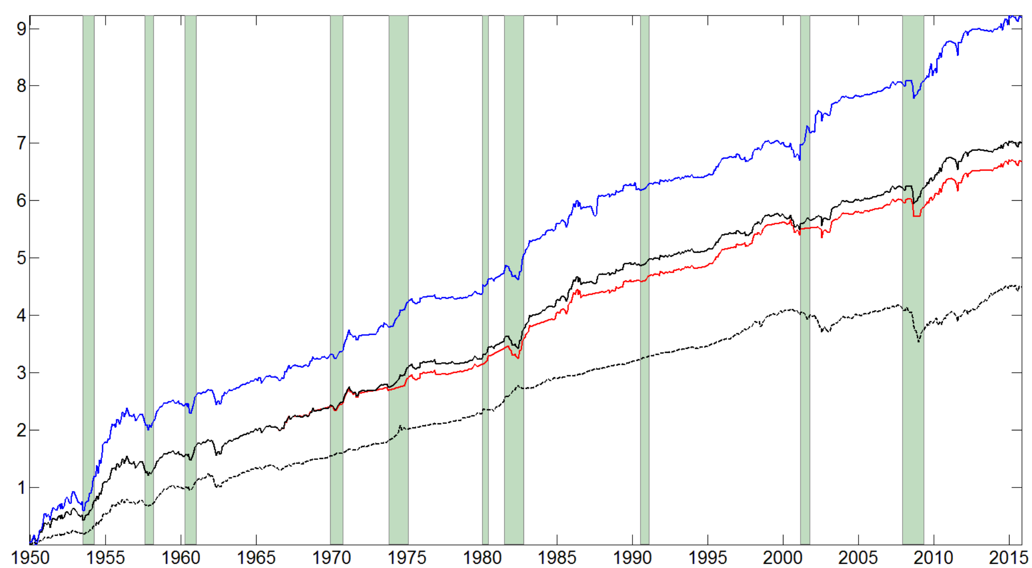


Figure 6: Equity weights and log cumulative wealth

Panel A plots the dynamics of the equity weight for a mean-variance investor allocating her wealth each month between equities and risk-free bills according to the rule (10), using stock return forecasts based on the SOPWAV specification (equation (11)) and the HM benchmark. The baseline scenario is the equity weight constrained to lie between -0.5 and 1.5 (black line for SOPWAV and dashed black line for HM). The blue line represents the equity weight for the investor using the SOPWAV method without portfolio constraints. Panel B delineates the corresponding log cumulative wealth for the investor, assuming she begins with \$1 and reinvests all proceeds. The red line represents the log cumulative wealth for the investor using the SOPWAV method with no short sales (equity weight constrained to be between 0 and 1.5). Gray bars denote NBER-dated recessions. The investor is assumed to have a relative risk aversion coefficient of three. Transaction costs are calculated using the monthly turnover measures and assuming a proportional cost equal to 50 basis points per transaction. The sample period runs from 1950:01 to 2015:12, monthly frequency.

Appendix 1. Definition of predictors of stock market returns

- Book-to-market (BM): ratio of book value to market value for the Dow Jones Industrial Average.
- Default return spread (DFR): difference between long-term corporate bond and long-term bond returns.
- Default yield spread (DFY): difference between BAA- and AAA-rated corporate bond yields.
- Dividend-payout ratio (DE): difference between the log of dividends (12-month moving sums of dividends paid on the S&P 500) and the log of earnings (12-month moving sums of earnings on the S&P 500).
- 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 price).
- Dividend growth (GD): monthly growth rate of dividends paid on the S&P 500.
- Earnings-price ratio (EP): difference between the log of earnings (12-month moving sums of earnings on the S&P 500) and the log of prices (S&P 500 index price).
- Inflation rate (INFL): growth in the consumer price index with a one-month lag.
- Long-term return (LTR): long-term government bond return.
- Long-term yield (LTY): long-term government bond yield.
- Net equity expansion (NTIS): ratio of 12-month moving sums of net issues by NYSE-listed stocks to NYSE market capitalization.
- Return on equity (ROE): ratio of 12-month moving sums of earnings to book value of equity for the S&P 500.
- Stock variance (SVAR): sum of squared daily stock market returns on the S&P 500.[12]
- Term spread (TMS): difference between the long-term government bond yield and the T-bill.
- Treasury bill rate (TBL): three-month Treasury bill rate.

[12] We find very similar results using the excess stock return volatility (computed using a 12-month moving standard deviation estimator), as suggested by Mele (2007), instead of the SVAR.

Appendix 2. A brief introduction to the discrete wavelet transform (DWT)

A wavelet is a function of finite length that oscillates around a time axis and loses power as it moves away from the center. The term wavelet originates from the admissibility condition, which requires the mother wavelet to be of finite support (small) and of oscillatory (wavy) behavior. Hence, a small wave, or wavelet.

The DWT allows decomposition of a time series into its constituent multiresolution components. High-frequency components reflect the short-term behavior of the variable, whereas the low-frequency component captures its long-term dynamics. There are two kinds of wavelets: father wavelets, ϕ, that capture the smooth and low-frequency part of the series, and mother wavelets, ψ, that capture the detail and high-frequency components of the series:

$$\int \phi_t dt = 1 \quad and \quad \int \psi_t dt = 0 \ .$$

Given a time series y_t, with the number of observations equal to N, its orthogonal wavelet approximation is defined by

$$y_t = \sum_k s_{J,k} \phi_{J,k,t} + \sum_k d_{J,k} \psi_{J,k,t} + \sum_k d_{J-1,k} \psi_{J-1,k,t} + \cdots + \sum_k d_{1,k} \psi_{1,k,t} \ , \tag{12}$$

with J representing the number of multi-resolution levels (or frequencies) and k ranging between one and the number of coefficients in the corresponding component. The maximum number of frequencies that can be considered in the analysis is determined by the number of observations, with $N \geq 2^J$.

In equation (12), there are two families of inputs: the approximating wavelet functions $\phi_{J,k,t}$ and $\psi_{j,k,t}$, and the wavelet transform coefficients $s_{J,k}$, $d_{J,k}$, $d_{J-1,k}, \ldots,$ $d_{1,k}$. The approximating wavelet functions are generated from the father ϕ and mother ψ wavelets through scaling and translation in the following way:

$$\begin{aligned} \phi_{J,k,t} &= 2^{-J/2} \phi \left(\frac{t - 2^J k}{2^J} \right) , \\ \psi_{j,k,t} &= 2^{-j/2} \psi \left(\frac{t - 2^j k}{2^j} \right) , \quad j = 1, 2, \ldots J \ . \end{aligned}$$

The wavelet transform coefficients represent the contribution of the respective wavelet function to the signal and are given by

$$
\begin{aligned}
s_{J,k} &= \int y_t \phi_{J,k,t} dt, \\
d_{j,k} &= \int y_t \psi_{j,k,t} dt \quad j = 1, 2, \ldots, J \ .
\end{aligned}
$$

Those wavelet transform coefficients are obtained using the DWT method, which maps the vector $y = (y_1, y_2, \ldots, y_N)'$ to a vector of N wavelet coefficients that includes the smooth coefficients $s_{J,k}$ and the detail coefficients $d_{j,k}$. The DWT method thus maps the original time series y_t in the time domain to a representation in the time-frequency domain $(y_{1,t}, y_{2,t}, \ldots, y_{N,t})'$. Equation (12) can therefore be rewritten as:

$$
y_t = S_{J,t} + D_{J,t} + D_{J-1,t} + \ldots + D_{1,t} \ , \tag{13}
$$

where $S_{J,t} = \sum_k s_{J,k} \phi_{J,k,t}$ is the wavelet smooth and $D_{j,t} = \sum_k d_{j,k} \psi_{j,k,t}$ for $j = 1, 2, \ldots, J$ are the J wavelet details. A J-level wavelet decomposition of the variable y_t thus consists of J wavelet details, which represent the higher-frequency characteristics of y_t, and the wavelet smooth that captures the low-frequency dynamics. Equation (13) represents the time-frequency decomposition of y_t and is the wavelet multiresolution decomposition, *i.e.* the original series y_t, exclusively defined in the time domain, is decomposed in orthogonal components (or crystals), $S_{J,t}$, $D_{J,t}$, $D_{J-1,t}, \ldots, D_{1,t}$, each defined in the time domain and each representing the fluctuation of the original time series in a specific frequency band.

Appendix 3. List of crystals that maximize the R^2_{OS}

The crystals that maximize the R^2_{OS} for the full forecasting period (1950:01 to 2015:12) of the SOPWAV method versus the historical mean benchmark are the following:

Variable	**Crystals**
default return spread	D_4, D_5, D_6, D_7
default yield spread	D_1, D_4, D_5
dividend-price ratio	$S_7, D_1, D_2, D_3, D_5, D_6$
growth dividend	S_7
inflation rate	D_1, D_2, D_3
long-term bond return	D_4, D_5, D_6
long-term bond yield	D_3
return on equity	D_2
term spread	D_1, D_7