HELLENIC
REPUBLIC

UNIVERSITY
OF MACEDONIA

ISSN 1791-3144

University of Macedonia
Department of Economics

Discussion Paper Series

Stock prices predictability at long horizons: Two tales from the time-frequency domain

Nikolaos Mitianoudis and Theologos Dergiades

Discussion Paper No. 4/2016

Department of Economics, University of Macedonia, 156 Egnatia str, 540 06
Thessaloniki, Greece, Fax: + 30 (0) 2310 891292
http://www.uom.gr/index.php?newlang=eng&tmima=3&categorymenu=2

Stock Prices Predictability at Long-horizons: Two Tales from the Time-Frequency Domain

Nikolaos Mitianoudis[a], Theologos Dergiades[b]

[a]*Image Processing and Multimedia Laboratory,*
Department of Electrical and Computer Engineering,
Democritus University of Thrace,
67100 Xanthi, Greece
[b]*Department of International and European Studies,*
University of Macedonia,
54006 Thessaloniki, Greece

November 29, 2016

Abstract

Accepting non-linearities as an endemic feature of financial data, this paper re-examines Cochrane's "new fact in finance" hypothesis (Cochrane, *Economic Perspectives*−FRB of Chicago 23, 36-58, 1999). By implementing two methods, frequently encountered in digital signal processing analysis, (*Undecimated Wavelet Transform* and *Empirical Mode Decomposition*−both methods extract components in the time-frequency domain), we decompose the real stock prices and the real dividends, for the US economy, into signals that correspond to distinctive frequency bands. Armed with the decomposed signals and acting within a non-linear framework, the predictability of stock prices through the use of dividends is assessed at alternative horizons. It is shown that the "new fact in finance" hypothesis is a valid proposition, provided that dividends contribute significantly to predicting stock prices at horizons spanning beyond 32 months. The identified predictability is entirely non-linear in nature.

Keywords: Stock prices and dividends, Time-frequency decomposition.
JEL: G10, C14, C22, C29

Email addresses: `nmitiano@ee.duth.gr` (Nikolaos Mitianoudis), `dergiades@uom.edu.gr` (Theologos Dergiades)

1. Introduction

During the last decade, the espousal of non-linear specifications, under the Cochrane's "new fact in finance" prism, is an empirical norm that attracts momentous attention in the literature (such as Psaradakis et al., 2004; Rapach and Wohar, 2005; McMillan, 2007). The theoretical underpinnings, stemming from the Present Value (PV) model (Shiller, 1981), nominate the utilisation of dividends (Campbell and Shiller, 1988) as a key determinant in explaining stock prices' future movements. The PV model, despite its simplicity and inherent intuitive soundness, is open to criticism mainly due to its conservative nature (e.g. it does not account for intangible assets like patents and brand names, or possible transaction costs). Unsurprisingly, the empirically stylised facts provide mixed evidence towards the adequacy of the linear PV model to describe in a satisfactory way the adjustment of stock prices to the steady state condition, implied by the underlying fundamentals.

Two strands of empirical literature have been unfolded. The first strand fails to justify the capacity of dividends (or valuation ratios) to contribute significantly to pricing stocks, while the second strand validates the efficacy of dividends (or valuation ratios) in pricing stocks adequately. The former strand of the literature is supported by more than a few studies (e.g. Shiller, 1981; Lanne, 2002; Fama and French, 2002; Torous et al., 2004), raising several justifications in order to: a) criticize past favorable inferences towards stock prices predictability or b) theoretically validate the observed systematic divergence from the long-run equilibrium. On the one hand, possible reasons for reaching a misleading statistical inference may be: a) the inadequacy of the adopted methodological framework (e.g. linear long-horizon regressions; see Wolf, 2000; Valkanov, 2003) or b) statistical inference issues (e.g. invalid selection of standard errors; see Ang and Bekaert, 2007), on the other hand, the theoretical reasoning for the inability of the fundamentals to explain assets pricing, includes: a) traders erroneous perceptions or more generally the existence of noise traders (Shiller, 1981; Kilian and Taylor, 2003), b) collapsing rational bubbles (Evans, 1991) and intrinsic bubbles (Froot and Obstfeld, 1991), c) transactions costs (Kapetanios et al., 2006), d) traders psychology (Summers, 1986; Cutler et al., 1991; Dergiades et al., 2015), or e) the existence of a time-varying discount rate (Shiller, 1989). Overall, it can be argued that all of the above-mentioned theoretical reasoning encompasses some sort of non-linear dynamic which result in equilibrium mis-pricing.

The second strand of the literature provides evidence in favor of stock prices' predictability (especially at long-horizons) by using valuation ratios. Early studies clung onto the linear paradigm (Campbell and Shiller, 1987; Phillips and Ouliaris, 1988), while in the most recent studies the basic proclivity is to adopt a non-linear methodological framework (Kanas, 2005; Rapach and Wohar, 2005; McMillan, 2007; Wu and Hu, 2012). No matter whether linear or non-linear methodology is implemented, a unanimous finding in the literature suggests that stock prices can be predictable at long-horizons without simultaneously validating predictability at short-horizons. The above systematically observed pattern gave birth to the "new fact in finance" hypothesis, as very eloquently coined by Cochrane (1999).

Campbell and Shiller (1998) find, at a short-horizon (one year), a trifling support for the predictability of stock returns by the price-dividend ratio, while at long-horizons (ten years) significant predictability is verified. Rapach and Wohar (2005), by concentrating on horizons spanning from one to ten years, identify a similar predictability pattern to that of Campbell and Shiller (1998). In particular, the price-dividend ratio as well as the price-earnings ratio, both contribute significantly to explaining stock prices growth at horizons spanning from six to ten years and from eight to ten years, respectively. Predictability at shorter horizons is not verified. Given the inborn inconsistency of the linear long-horizon regressions in separating between long-run and short-run predictive power (pointed out by Berkowitz and Giorgianni, 2001), the predictability pattern as shaped within the lines of the "new fact in finance" hypothesis is indirectly attributed to the presence of non-linear features in the data (Campbell and Shiller, 1998; Kilian, 1999). Therefore, a fertile ground for further research is the supplementary verification of non-linear dynamics (of a general form) at the long-horizon predictability of stock prices. Rapach and Wohar (2005) characteristically state that "further analysis of non-linear model specifications for valuation ratios is warranted and may help researchers to better understand long-horizon stock price predictability".

Building on the existing literature, the objective of our study is twofold: a) to re-examine the cogency of the "new fact in finance" hypothesis through the implementation of an alternative methodological framework (not previously implemented in similar exercise), and b) to authorise in a sound way the presence of non-linear dynamics in the predictability of real stock prices at long-horizons. The contribution of our study relies on the combination of signal processing methods along with non-linear and non-parametric causality techniques, framework that allows us to gain more insightful knowledge with respect to the non-linear factual linkage between stock prices and dividends at different horizons.

Utilising the extended dataset of Campbell and Shiller (1998) for the US economy, our approach allows us, firstly, to decompose the series into signals that correspond to dissimilar frequency bands and, secondly, to ascertain the existence of a non-linear predictability for the stock prices at short- and long-horizons based on dividends. The decomposition of both series is implemented through two alternative signal processing methods, that is, the *Empirical Mode Decomposition* (*EMD*, hereafter) and the *Undecimated Wavelet Transform* (*UWT*, hereafter). Once the series components have been extracted, we test the null hypothesis of no predictability for all the resulting pairs with identical frequency content. The null hypothesis is tested by two non-linear tests, that is, the Hiemstra and Jones (1994) test and the Diks and Panchenko (2006) test. Our findings confirm the "new fact in finance" hypothesis by identifying predictability only at horizons that expand beyond 32 months. Additionally, the identified predictability can be characterised as entirely non-linear in nature.

The rest of the paper is organised as follows: section 2 presents briefly the adopted methodological framework; section 3 illustrates the data sources and conducts the necessary preliminary analysis; section 4 discusses our empirical findings; while section 5 concludes.

2. Methodological Framework

2.1. Empirical Mode Decomposition (EMD)

The *EMD* method is a recent tool in time-series analysis, which was proposed by Huang et al. (1998). Originally, the method has been advanced with a purpose of decomposing a signal into what is known as *Intrinsic Mode Functions* (*IMFs*), so as to implement the Hilbert transform (Hilbert, 1953) at a second stage. Once *EMD* had been proposed, *IMFs* proved to be a valuable device for several other applications, without the necessity to perform the subsequent Hilbert Transform.

In order for a signal to be a valid *IMF*, it must satisfy the following two conditions: a) the number of local extrema and passings through zero must be equal or differ at most by one, and b) the average value of the *IMF* is locally almost zero everywhere. *EMD* encompasses four essential characteristics that render the method significantly powerful when dealing with non-linear and non-stationary signals. These characteristics are: completeness, orthogonality, locality, and adaptivity (Huang et al., 1998). In linear decompositions, completeness and orthogonality are considered as necessary conditions. Locality is a central characteristic when non-stationary signals are processed, and adaptivity is a vital characteristic for series that exhibit both non-linearities and non-stationarity.

The majority of time-series usually encountered cannot be characterised as *IMFs*. Nevertheless, it is possible to decompose any signal into *IMFs* plus a residual term, if we execute an algorithm called the *sifting process*. Given a time-series, say $x(t)$, the *sifting process* algorithm can be summarised into the following steps: a) identify all the existing extrema of $x(t)$, b) interpolate using cubic splines between the identified minima and maxima in order to produce two envelopes $e_{min}(t)$ and $e_{max}(t)$, c) compute the mean of the two envelopes $m(t) = (e_{min}(t)+e_{max}(t))/2$ and finally d) extract the *detail* $h(t) = x(t) - m(t)$. The *sifting process* is repeated on the *detail* $h(t)$, with some stopping criterion, until a first valid *IMF* signal $h_1(t)$ is received. To identify the second *IMF* signal, the previously extracted *IMF* signal is subtracted from the original time-series and the same process is repeated on the residual $r_1(t) = x(t) - h_1(t)$. The same process is repeatedly implemented up to the stage where the last residual term, $r_n(t)$, is strictly monotonic or if it contains at least one extrema.

$$r_1(t) = X(t) - h_1(t),\ r_2(t) = r_1(t) - h_2(t),\ \ldots,\ r_n(t) = r_{n-1}(t) - h_n(t) \quad (1)$$

Provided that we are interested in *IMFs* that occupy a distinct frequency band, we utilise the stopping criterion proposed by Rilling et al. (2003),[1] which is based on three control thresholds, namely a, θ_1 and θ_2. Once $m(t)$ has been computed, Rilling et al. (2003) introduce the *evaluation* function σ_t, which is given by Eq. (2):

[1]Huang et al. (1998) suggested a stopping criterion, which despite its simplicity exhibits several weaknesses. For a discussion see Rilling et al. (2003).

$$\sigma_t = \left| \frac{h(t)}{m(t)} \right| \tag{2}$$

For a fraction $(1-a)$ of the total duration of the signal, the sifting process is iterated until $\sigma_t < \theta_1$. For the rest of the signal, it is iterated until $\sigma_t < \theta_2$. Typical values for the thresholds a, θ_1 and θ_2 are 0.05, 0.05 and $10\theta_1$. The rationale for the inclusion of the two threshold values can be traced to the fact that the algorithm can take into account abnormally large fluctuations in the scrutinised time-series.

2.2. Discrete Wavelet Transform (DWT)

Non-stationarity and non-linearities are both endemic features of economic and financial data. The widespread use of wavelet transforms in economics and finance is mainly attributed to the capacity of the wavelet techniques to cope successfully with the above two features.[2] In contrast with the *Fourier Transform*, which decomposes a complex signal into a sum of sine and cosine functions at different frequencies but with infinite length in the time domain, wavelet analysis uses waves with various lengths of time frames. The length of the time frame is analogous to the level of frequency resolution. The *DWT* is based on the successive use of high-pass filters named *mother wavelets* h, and low-pass filters named *father wavelets* g. In the standard *DWT*, these filters must possess two properties. First, they are half-band filters implying that their cut-off frequency is in the middle of the frequency band of the starting time-series, and second, they are quadrature mirror filters. The latter property means that the power sum of the low and the high pass filter equals to one and also that their responses are symmetrical around their cut-off frequency. In addition, since half the frequencies of the signal are removed after the use of each filter, half the samples can be discarded according to Nyquist's theorem. Consequently, the filter outputs are each time sub-sampled by two.

[2] A discussion for the usefulness of wavelets in economic and financial analysis can be traced in Ramsey (1999). For a recent application of wavelet analysis to economic data see Michis (2014).

In cases where the low pass filter has an impulse response $g(t)$, the output $a(t)$ will be equal to its linear convolution (denoted by *) with the signal $x(t)$:

$$a(t) = g(t) * x(t) = \sum_{k=-\infty}^{+\infty} x(k)g(t-k) \tag{3}$$

Including the sub-sampling, the outputs $a[t]$ and $d[t]$ of the low and high pass filters will respectively be:

$$a(t) = \sum_{k=-\infty}^{+\infty} x(k)g(2t-k) \tag{4}$$

$$d(t) = \sum_{k=-\infty}^{+\infty} x(k)h(2t-k) \tag{5}$$

The appliance of the above two filters makes up for one level of the DWT, producing one *approximation* signal $a(t)$ and one *detail* signal $d(t)$. For more levels of decomposition, the filters are implemented in succession. In particular, at the first level, the input series $x(t)$ is passed through one high-pass filter giving the first level *detail* signal and one low-pass filter, providing the first level *approximation*. At each subsequent level, the *approximation* signal $a(t)$ is further analysed into two new signals using the same approach. At the end, our outputs consist of one final *approximation* signal and n *detail* signals, where n is the decomposition level.[3]

In many applications, the step of sub-sampling is omitted and as a result the output signals have the same length as the inputs (Starck et al., 2007). This alteration of the DWT can be identified in the literature by many terms, including the *Stationary Wavelet Transform* (SWT), the *Redundant Wavelet Transform* (RWT) and the *Undecimated Wavelet Transform* (UWT). In this paper, the term UWT is adopted. Thus, the outputs of the transform preserve the same length as the starting time-series, i.e. we retain time correspondence between the various approximation and *detail* outputs.

The adopted filtering procedure is similar to the one described above with the only difference being the fact that the sub-sampling part is omitted.

[3] The above described process can be depicted graphically through a standard Mallat-tree decomposition diagram.

Along with the *UWT* framework, we continue to receive *n detail* signals and one *approximation* signal as output of the transform. The transition from one level to the next is accomplished with the use of the filters below:[4]

$$a_{n+1}(t) = (h^{(j)} * a_n)(t) = \sum_{k=-\infty}^{+\infty} h(k) a_n(t + 2^j k) \tag{6}$$

$$d_{n+1}(t) = (g^{(j)} * a_n)(t) = \sum_{k=-\infty}^{+\infty} g(k) a_n(t + 2^j k) \tag{7}$$

where $h(t)$ and $g(t)$ are the low and high pass filters impulse responses, respectively. The star symbol (*) implies linear convolution and finally, $h^{(j)}$ is an indicator function defined by:

$$h^{(j)}(t) = \begin{cases} h(t), & \text{if } t/2^j \in \mathcal{Z} \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

3. Data Sources and Preliminary Analysis

To investigate the predictive power of dividends with respect to the stock prices, we utilise time-series data with monthly frequency spanning from January 1871 up to February 2013 (1706 observations). Focus of our analysis is the US economy. The monthly real stock prices (S) are approximated by monthly averages of daily closing prices for the S&P 500 composite index. The real dividend series (D), corresponding to the S&P 500 composite index, is built by the four quarter totals, while monthly observations are attained through linear interpolation. The exact construction details for both variables are reported in Shiller (1989), while the utilised time-series can be traced on Roberts J. Shiller's personal web site.[5] The evolution of the S and D over the examined sample is illustrated in Figure 1 and Figure 2 below, respectively.[6]

[4]For more details on wavelet decompositions, the reader is referred to Mallat (2008).

[5]Available at: http://www.econ.yale.edu/∼shiller/data.htm

[6]By examining Figure 1, one may argue in favor of different regimes. As wavelets can represent effectively complex series, *UWT* is capable to cope with "badly behaved data".

Figure 1: Real S&P 500 stock prices (S)

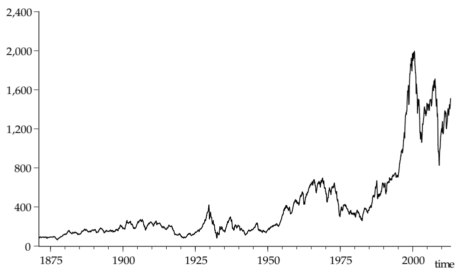
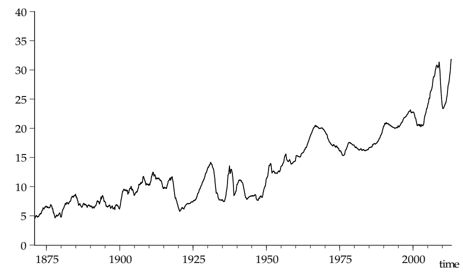


Figure 2: Real dividends (D)



To ensure that a possible verified predictability of S based on D is entirely non-linear in nature, a first-moment filtering is of major importance. An effective first-moment filtering dictates the adoption of a correctly specified model, provided that contrarily incorrect causal linkages may be inferred. Consequently, before the conduction of the first-moment filtering the verification or not of a long-run equilibrium is essential in terms of econometric modeling.[7] The Johansen (1988) approach to cointegration indicates the existence of a unique cointegrating vector between S and D. In particular, the trace statistic clearly rejects the null hypothesis of zero cointegrating vectors (the p-value is equal to 0.000) while it fails to reject the null hypothesis of at most one cointegrating vector (the p-value is equal to 0.353).[8]

The presence of cointegration dictates that the first-moment filtering procedure should be conducted through a VECM specification, since otherwise (using for example a standard VAR model) the long-run dynamics of the involved variables will be neglected. As soon as we estimate the optimal VECM specification, the first-moment filtered series that correspond to S and D can be recovered through the respective residuals, denoted from now on as S_r and D_r (see Figures 3 and 4). Acting in such a manner, we warrant that all the linear components of the series have been removed. Finally, once the VECM residuals have been recovered, we test the *identically and independently distributed* (i.i.d.) assumption through the BDS test as suggested by Brock et al. (1996). Rejection of the null i.i.d. hypothesis would be an

[7]Preliminary unit-root and stationarity testing indicates that both variables, stock prices and dividends, are integrated of order one. The tests implemented are: 1) ADF, 2) GLS-ADF, 3) Phillips−Perron and 4) KPSS

[8]The cointegration results are available upon request.

indication in favor of our adopted non-linear methodological framework. The BDS test which has been conducted into the VECM residuals, irrespective of the embedding dimension, consistently rejects, at the 0.01 significance level, the null hypothesis.[9]

Figure 3: Stock price residuals (S_r)

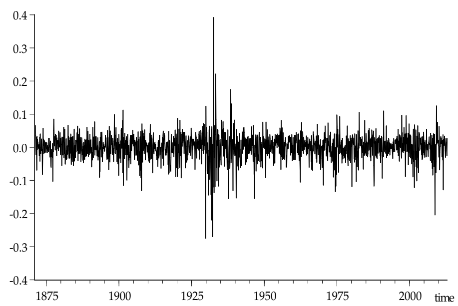
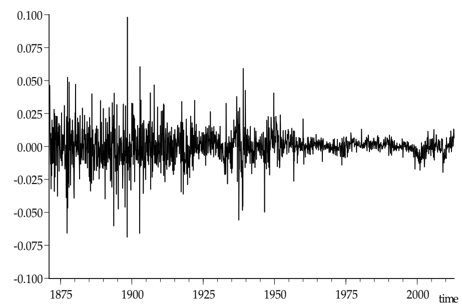


Figure 4: Dividends residuals (D_r)



4. Empirical Results

4.1. Time-Frequency Decomposition

First, we perform the *EMD* decomposition on the VECM residual series, that is S_r and D_r. We derive the first 5 *IMFs* for the S_r and D_r series (see Figures 3 and 4). We use the Rilling et al. (2003) stopping criterion with the recommended convergence values. The residual waveform is used to form the 6th component. In Figures 5 and 6, we depict the *EMD* decompositions for S_r and D_r, respectively. We observe that each extracted *IMF* features decreasing fluctuations from its previously extracted *IMFs*, as dictated by the decomposition process.[10]

Our first concern regarding the signal decomposition process is that each signal extracted from one time series, e.g. S_r, needs to encompass the same frequency content with its counterpart from the other series, e.g. D_r. This is a prerequisite before we explore causality relations between the extracted components of S_r and D_r. There needs to be consistent frequency correspondence between the respective components of S_r and D_r, i.e. they should offer

[9]The BDS test results are available upon request.

[10]The MATLAB code for the *EMD* decomposition is based on the code available at http://perso.ens-lyon.fr/patrick.flandrin/emd.html.

Figure 5: *EMD* Decomposition (S_r)

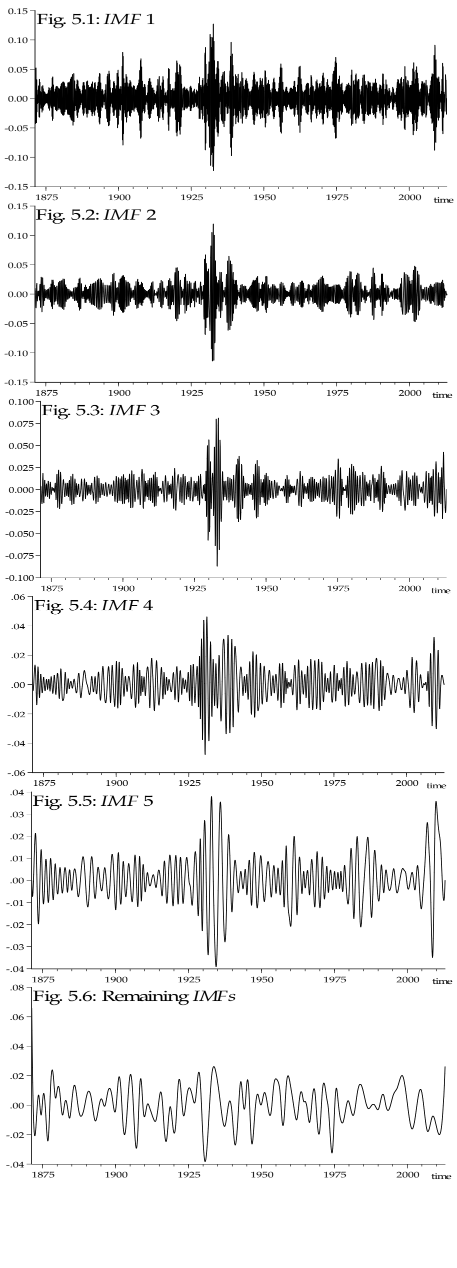
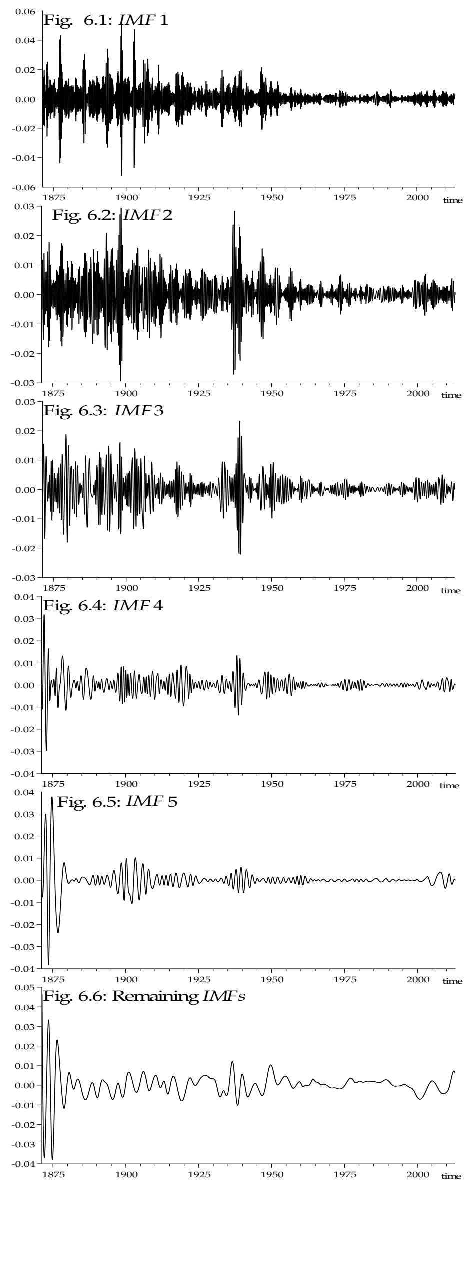


Figure 6: *EMD* Decomposition (D_r)



information about the same period of time. For example, if the 1st *IMF* of S_r has a frequency content of 2 to 4 months then the same should hold for the 1st *IMF* of D_r. Things are complicated with respect to the frequency content of the *IMFs*. *EMD* is an algorithmic procedure and there is no *a priori* any form of restriction on the *IMFs* frequency band. By construction, we can guarantee that each *IMF* corresponds to a lower frequency from its previous only locally. The frequency content of the *IMFs* depends on the type of signal we are decomposing as well as on the specific characteristics of the *EMD* algorithm used each time. Flandrin et al. (2004) observed that for Gaussian signals, the functions derived through *EMD* are similar to that from the *DWT*. The first *IMF* appears to be an output of a high-pass filter, whereas the rest are outputs of band-pass filters, each covering a different frequency band. Furthermore, Flandrin et al. (2005) reach the same conclusion, that is, the two methods work in a similar fashion for stationary signals with wide frequency content.

Since there is no formal way to prove mathematically that the frequency content of the produced *IMFs* will form consistent and continuous frequent content areas, it is necessary to measure their frequency content via the *Fourier Transform*. After the implementation of the *EMD* algorithm, each series is decomposed into 6 valid *IMFs*. In order to explore the frequency content of the *EMD* components for S_r and D_r, we use the Welch spectral estimation with a Kaiser window of 64 points (Hayes, 1996). These are shown in Figures 7 and 8. We observe that although the frequency content of each component is relative band-pass, there is strong "leakage" between successive frequency bands. Thus, there is mixed frequency content between neighbouring *IMFs* and naturally causality inference may be misleading.

Figure 7: *EMD* frequency responses (S_r)

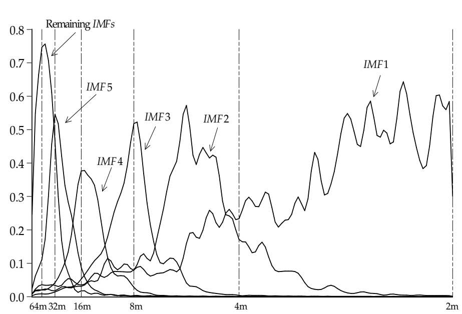
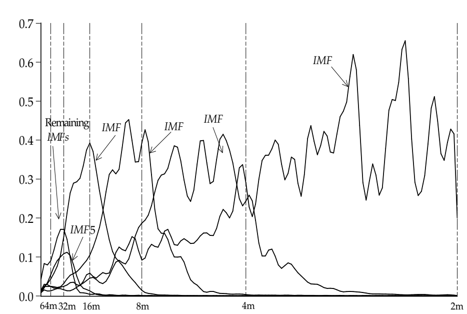


Figure 8: *EMD* frequency responses (D_r)



We gather that, instead of using each *IMF* separately, we utilise sums of neighboring *IMFs*, in order to receive more meaningful and coherent frequency bands. In more detail, we chose to use: a) the sum of the first and the second *IMF*, b) the sum of the third and fourth *IMF*, and finally c) the sum of all the remaining *IMFs*. This implies that we are essentially grouping neighbouring (both in terms of frequency and order of extraction) *IMFs*. The same aggregation applies to both S_r and D_r series. The constructed time-series are depicted in Figures 9 and 10. In order to explore the frequency content of the aggregated *EMD* components for S_r and D_r, we use the Welch spectral estimation with a Kaiser window of 64 points (Hayes, 1996). These are shown in Figures 11 and 12. We observe that the frequency content of the created components is more well-defined with clearly lower "leakage". Hence, the aggregated components will contain more consistent frequency (time-span) information, compared to the original components.

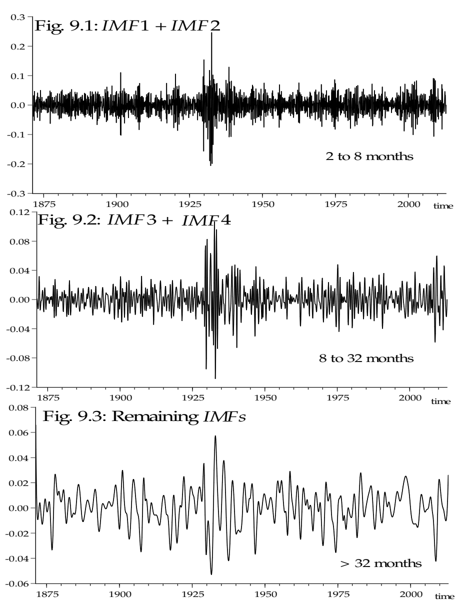


Figure 9: *EMD* components (S_r)

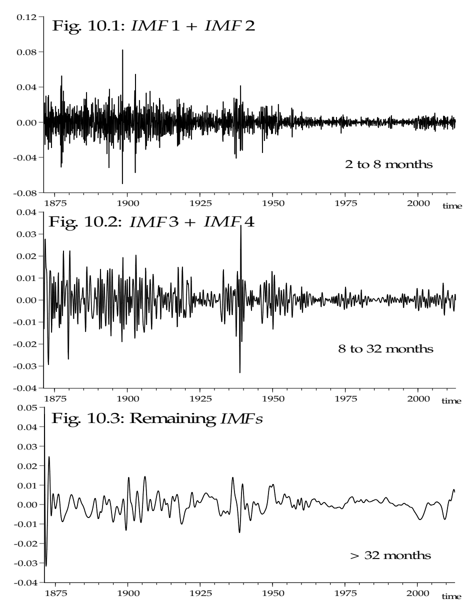
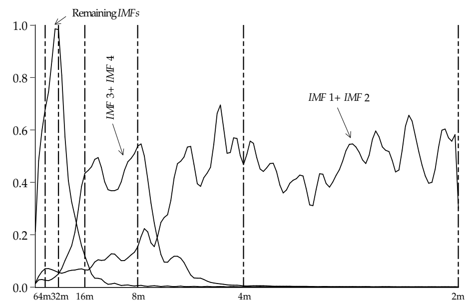
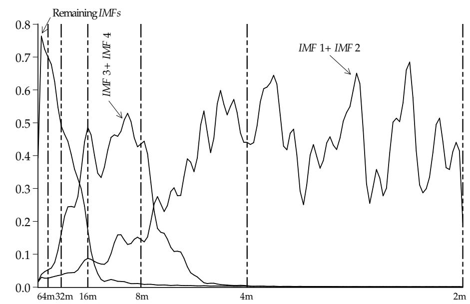


Figure 10: *EMD* components (D_r)

To validate and compare the analysis through the *EMD* components, we also perform the *UWT* decomposition on the VECM residual series, that is S_r and D_r. The "Symlet 8" mother wavelet is selected for its symmet-

Figure 11: *EMD* sums freq. responses (S_r) Figure 12: *EMD* sums freq. responses (D_r)





rical and accurate change localisation properties.[11] In Figures 13 and 14, we depict the 5-level UWT decomposition carried out on the S_r and D_r series, respectively. Unlike previous decimated wavelet approaches, the undecimated property of the UWT approach ensures that the five extracted *detail* series as well as the *approximation* series are all of equal length to the input. This will enable us to come up with a better time correspondence and accuracy between the components that we intend to analyze further. As expected, the *detail* components contain higher fluctuations compared to the *approximation* component. As we descend into the wavelet decomposition, *detail* components tend to be smoother, i.e. contain lower frequencies.

For the UWT approach, the same frequency content prerequisite is satisfied by the properties of the transform. Provided that the frequency bands of the outputs are increased in a dyadic manner as well as the fact that we use the same sampling frequency along with the same level of decomposition for our signals, the frequency bands in each level of decomposition will be the same. Specifically, given that our sampling frequency is 1 month, according to Nyquist sampling theorem (Nyquist, 1928); the smaller observable frequency will be 2 months, implying that the level one *detail* signal will have a frequency band of 2 to 4 months. Given the dyadic increase in the frequency content, it comes that the level two *detail* will contain frequencies in the 4 to 8 month band and so on. In Figures 15 and 16, we plot the frequency responses of the decomposition of S_r and D_r respectively, as a function of frequency, expressed in months. The frequency responses were estimated

[11]We used MATLAB R2012b Wavelet Toolbox for this decomposition.

Figure 13: UWT components (S)

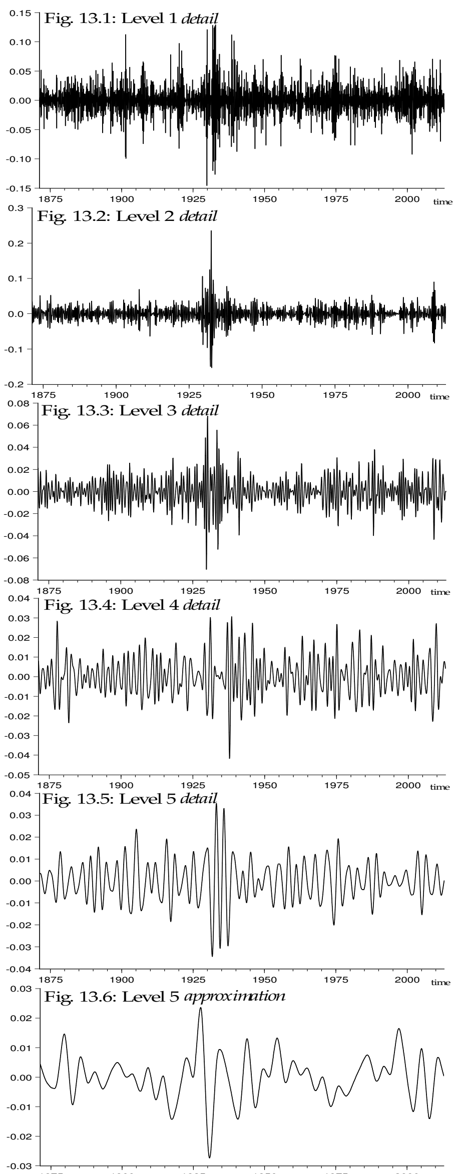
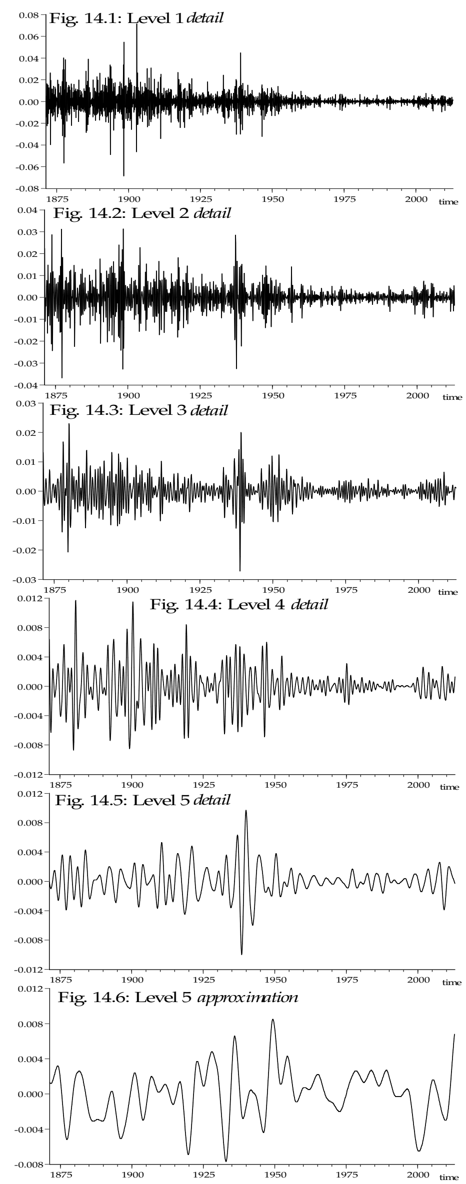


Figure 14: UWT components (D)



using the Welch method and a Kaiser window of 64 points (Hayes, 1996). It appears that the corresponding *detail* and *approximation* components of S_r and D_r have similar frequency range. Nonetheless, we should devise a method to create components of similar frequency content to the aggregates created by the *IMF* components.

Figure 15: *UWT* frequency responses (S_r)

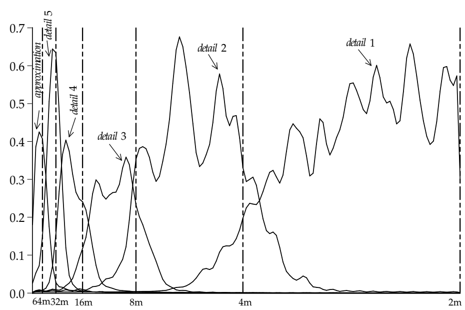
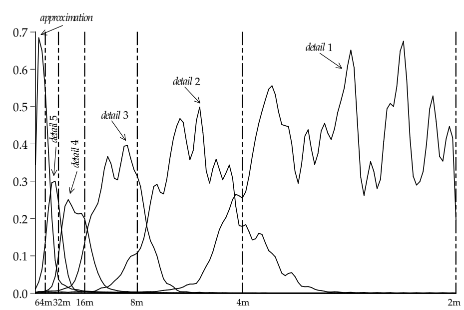


Figure 16: *UWT* frequency responses (D_r)



After we measure the frequency content of the *IMF* sums, we observe that their frequency bands increase in a triadic manner. Consequently, the first sum illustrates a frequency content of 2 to 8 months, the second sum illustrates a frequency content of 8 to 32 months, and the final sum illustrates a frequency content above 32 months. Under these circumstances, it is possible to create *UWT* aggregates with approximately the same frequency bands as those obtained through the *EMD* method, thus allowing a direct comparison between the two methods. The corresponding *UWT* sums are: the sum of the 1st and the 2nd *detail*, the sum of 3rd and the 4th *detail* and finally the sum of the 5th *detail* with the *approximation*. These sums of *UWT* components are shown in Figures 17 and 18 for the S_r and D_r series, respectively. Their frequency response is estimated by the Welch spectral estimation technique and is depicted in Figures 19 and 20.

In order to demonstrate that the new components, created from grouping of previous components satisfy the necessary conditions, we estimate the percentage of their power that correspond into the desirable frequency band. The spectral power of a discrete signal $x(t)$, under a discrete *Fourier Transform* $X(\omega)$, can be approximated generally by the following formula:

Figure 17: UWT sums (S_r)

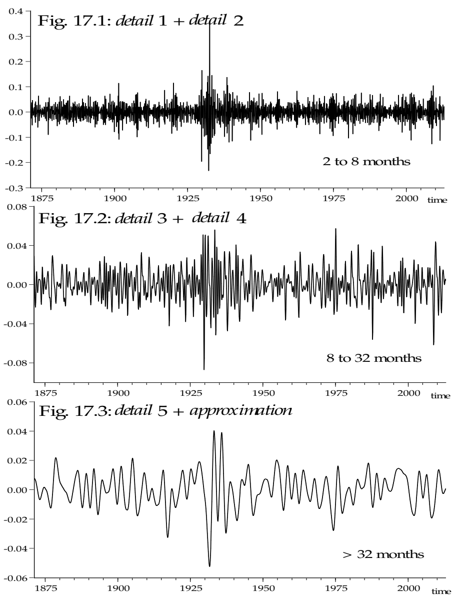


Figure 18: UWT sums (D_r)

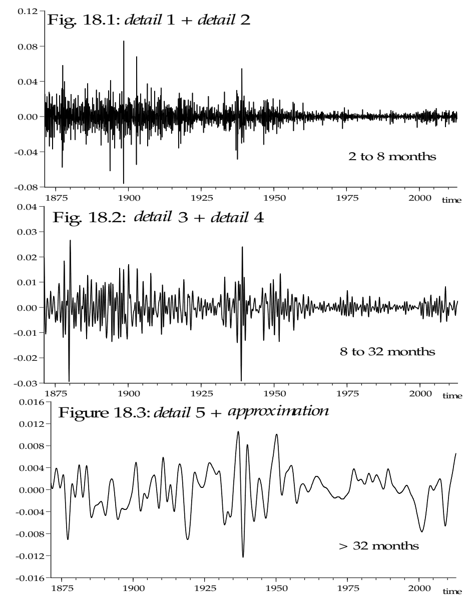


Figure 19: UWT sums freq. responses (S_r)

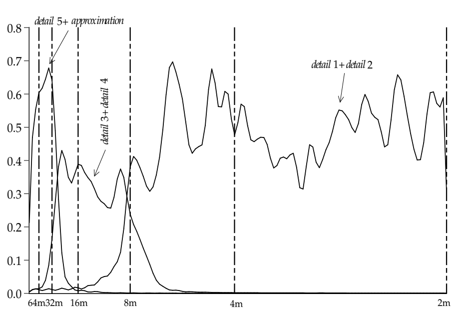
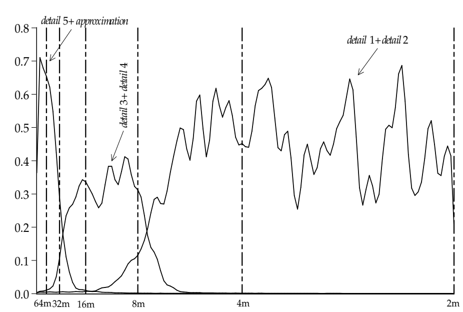


Figure 20: UWT sums freq. responses (D_r)



$$P_x \simeq \Delta T \left| \sum_{t=-\infty}^{+\infty} x(t) e^{i\omega t} \right|^2 = \Delta T X(\omega) X^*(\omega) \tag{9}$$

where ΔT is the sampling period and $X^*(\omega)$ the conjugate of $X(\omega)$. The results are summarised in Table 1. The calculated signal power for the new components varies between 79.7% and 98.5%. This denotes that these components have their energy greatly concentrated in the frequency range they represent. Hence, these decompositions can be employed to explore predictability in these three time periods (2 to 8 m., 8 to 32 m., > 32 m.).

4.2. Predictability testing at alternative horizons

Once, the spectral coherence between the extracted components of S_r and D_r, via both *UWT* and *EMD* methods, has been satisfied, the non-linear predictive content of D with respect to S is tested by conducting two tests proposed by Hiemstra and Jones (1994) (H&J, hereafter) and Diks and Panchenko (2006) (D&P, hereafter). The predictability testing is performed on the respective three created components of D_r and S_r, components that encompass the same frequency content. The non-linear causality tests (H&J and D&P) illustrated in Tables 2 and 3, are executed for eight different lags ($\ell_x = \ell_y = i$ with i ranging from 1 to 8). The subscripts used in the causality inference column represent the dominant significance/insignificance causality inference. In particular, the subscripts show the number of the calculated p-values (out of 8 in total) that belong to the indicated causality inference symbol. For example, $**_{(7/8)}$ implies that 7 out of the 8 calculated p-values point out significance at the 5% level, with the remaining p-values (1 in our example) to be higher or lower. Finally, the p-value range column shows for those p-values that belong to the dominant significance/insignificance predictability inference (7 in the previous example), the range of their values (max-min). In the previous example, if a p-value range equals to 0.015-0.036, implies that the 7 p-values belonging to the dominant set of the 5% significance level, receive values that vary from 0.015 up to 0.036.

Additionally, for simplicity as well as for presentation purposes, each and every extracted component of the S_r and D_r series is signified by a composite subscript. The first part of the subscript implies the method used for the decomposition (*se* for the *EMD* sums and *su* for the *UWT* sums), while the second part implies the decomposed component that corresponds to a distinct frequency content and therefore to a distinct time period (the second part ranges from $c1$ to $c3$ for both the *EMD* sums method and the *UWT* sums method). For instance, the notation S_{su_c1} refers to the first extracted component ($c1$) for the S_r series using the *UWT* sums method (su). All the extracted components for both time series, S_r and D_r, along with the corresponding time length and the associated power are analytically presented in Table 1.

Table 1: The power of the extracted components for the *UWT* and *EMD* methods

decomposition method	corresponding time frame	components for S	signal power	components for D	signal power
EMD sums	2 to 8 m.	S_{e_c1}	90.1%	D_{e_c1}	91.6%
	8 to 32 m.	S_{e_c2}	79.7%	D_{e_c2}	84.7%
	>32 m.	S_{e_c3}	94.5%	D_{e_c3}	92.1%
UWT sums	2 to 8 m.	S_{su_c1}	97.6%	D_{su_c1}	96.6%
	8 to 32 m.	S_{su_c2}	87.5%	D_{su_c2}	85.5%
	>32 m.	S_{su_c3}	91.3%	D_{su_c3}	98.5%

Notes: *EMD* sums and *UWT* sums denote the sums of the Empirical Mode Decomposition method and the sums of the Undecimated Wavelet Transform method, respectively.

The results of both causality tests conducted on the respective pairs of signals for S_r and D_r, produced by the *EMD* method are illustrated in Table 2. In more detail, the majority of both tests fail to reject the null hypothesis of no predictability for the short-run ($c1-2$ to 8 months) and the medium-run ($c2-8$ to 32 months) components for the series of interest. Though, this is not the case for the long-run component ($c3-$above 32 months). In the long-run, D appears to non-linearly Granger cause S at the 0.05 significance level for both tests H&J and D&P.

Finally, in the case where the testing procedure is reapplied into the sums of *UWT*, where the aggregated signals contain the same frequency content as the derived *IMFs* after the *EMD* decomposition, the resulting causal inference is qualitatively similar to that of the *EMD* method (see Table 3). The revealed inference suggests that for the short-run ($c1-2$ to 8 months) as well as for medium-run ($c2-8$ to 32 months) frequency components, we clearly fail to reject the null hypothesis of no-causality at the conventional levels of significance. Interestingly, it appears that both tests support, at the 0.05 significance level, the non-linear predictability of real stock prices

Table 2: Non-linear predictability of S under the *EMD* sums method

Causality direction	H&J test $p-$values range	H&J test inference	D&P test $p-$values range	D&P test inference
$D_{se\_c1} \rightarrow S_{se\_c1}$	0.128−0.250	$\notin_{(5/8)}$	0.173−0.248	$\notin_{(5/8)}$
$D_{se\_c2} \rightarrow S_{se\_c2}$	0.139−0.817	$\notin_{(7/8)}$	0.132−0.797	$\notin_{(7/8)}$
$D_{se\_c3} \rightarrow S_{se\_c3}$	0.011−0.016	$**_{(4/8)}$	0.012−0.020	$**_{(8/8)}$

Notes: The arrow ($\rightarrow$) denotes that the tested predictability runs from the left-hand side variable to the right-hand side variable. The symbols *, ** and *** denote existence of causality at the 10%, 5% and 1% significance level, respectively. The symbol ($\notin$) signifies no causality at the conventional levels of significance. The analytical results for each lag are presented in the Appendix A.1 (see Table 4).

through real dividends for the long-run extracted components ($c3-$above 32 months). A clear conclusion after the comparison of the two alternative time-frequency decomposition methods, UWT and *EMD*, is that a non-linear predictability of S through D can be affirmed at long-horizons.

Table 3: Non-linear predictability for the UWT sums components

Causality direction	H&J test $p-$values range	H&J test inference	D&P test $p-$values range	D&P test inference
$D_{su\_c1} \rightarrow S_{su\_c1}$	0.441−0.909	$\notin_{(8/8)}$	0.430−0.917	$\notin_{(8/8)}$
$D_{su\_c2} \rightarrow S_{su\_c2}$	0.380−0.762	$\notin_{(8/8)}$	0.197−0.699	$\notin_{(8/8)}$
$D_{su\_c3} \rightarrow S_{su\_c3}$	0.012−0.027	$**_{(4/8)}$	0.012−0.021	$**_{(4/8)}$

Notes: The arrow ($\rightarrow$) denotes that the tested predictability runs from the left-hand side variable to the right-hand side variable. The symbols *, ** and *** denote existence of causality at the 10%, 5% and 1% significance level, respectively. The symbol ($\notin$) signifies no causality at the conventional levels of significance. The analytical results for each lag are presented in the Appendix A.1 (see Table 5).

Overall, our findings (as these are summarized in Tables 2 and 3) support the long-run predictability of stock returns based on dividends and are in accordance with other studies (see Campbell and Shiller, 1998; Rapach and Wohar, 2005). In contrast to Rapach and Wohar (2005), who verify predictability at horizons spanning from six to ten years, our analysis using a different methodological framework finds that predictability appears after the first 2.7 years. The observed differences in the identified forecasting horizons may be attributed to several reasons. For example, one such reason may be the adopted methodological framework, when that it does not consider the presence of non-linear features in the data. Our study by relying on the joint use of signal processing methods along with non-linear and non-parametric techniques, allows us to identify possible non-linear linkages between stock prices and dividends at different horizons. Our findings support further the usage of non-linear specifications when it comes to testing the validity of the "new fact in finance" hypothesis.

5. Summary and Conclusions

The use of long-horizon regression equations consists the major workhorse when stock prices' predictability comes as an empirical question (Wu and Hu, 2012). In light of this, this study re-examines the Cochrane's (1999) "new fact in finance" hypothesis under an alternative methodological framework not previously used in a suchlike research inquiry. On condition that the identified long-horizon predictability of S is attributed to the presence of non-linearities underlying the data, we work in the time-frequency domain and we adopt a non-linear predictability framework. Acting so, we are in a position not only to disentangle the series (S and D) in frequency components which correspond to different distinct horizons, but also to gain insightful knowledge with respect to the nature of the predictive content encompassed in the D series with respect to the S series.

Using Shiller's (1998) extended dataset on real stock prices and real dividends for the US economy, spanning from 1891:1 to 2013:2 (monthly frequency), and after an appropriate linear filtering (using a VECM specification), we decompose the series into signals that correspond to dissimilar frequency bands. The decomposition is carried out via two methods frequently implemented in the digital signal processing analysis, that is the *Empirical Mode Decomposition* (*EMD*) and the *Undecimated Wavelet Transform* (*UWT*). Once the decompositions are accomplished, the existence of a non-linear predictability for S, at short- and long-horizons, is ascertained. When the two different decomposition methods are compared, our findings confirm the "new fact in finance" hypothesis by identifying a significant non-linear predictability for S, at long-horizons. In particular, the null hypothesis of no predictability is rejected at the 0.05 significance level, when the investigated horizons expand beyond 32 months. On the contrary, for frequencies below 32 months, there is a systematic failure to reject the null hypothesis.

This paper's contribution is twofold. First, we confirm the "new fact in finance" hypothesis under a different not previously implemented methodological framework. Second, we provide evidence that the observed pattern in stock prices predictability (long-horizon predictability) is mainly attributed to the inherent non-linear structure of the data. Our results are suggestive towards the adoption of a non-linear framework when the stock prices' predictability at long-horizons is examined. Finally, future empirical research may provide additional evidence towards the validity of the "new fact in finance" hypothesis by focusing on the predictive capacity of the valuation ratios, such as the dividend-price ratio or the earnings-price ratio.

Furthermore, apart from the adoption of a non-linear methodological framework, another appealing direction that worth special investigation is the robustness of the "new fact in finance" hypothesis, not only in different markets but also over-time. In other words, it would be interesting to whether stock price predictability has increased over the most recent years or vice versa. The detection of regimes in which the predictive content of dividends (or other valuation ratios) against the real stock prices is altering (reduces, increases or even disappears), may lead to a significant and valuable inferences. A natural approach of analysis, to shed some light on the above question, is each adopted methodology to be conducted on a rolling basis using sample windows with several lengths. This way, potential shifts in the predictive content may be identified, and new inferences may be accomplished.

Appendix A.1

Table 4: Non-linear predictability of S under the *EMD* sums method (all lags)

Causality direction	H&J test			D&P test		
	lag	$p-$value	inference	lag	$p-$values	inference
$D_{se\_c1} \rightarrow S_{se\_c1}$	1	0.069	*	1	0.067	*
	2	0.013	**	2	0.019	**
	3	0.064	*	3	0.081	*
	4	**0.128**	$\notin$	4	**0.173**	$\notin$
	5	0.171	$\notin$	5	0.227	$\notin$
	6	0.147	$\notin$	6	0.188	$\notin$
	7	0.228	$\notin$	7	0.258	$\notin$
	8	**0.250**	$\notin$	8	**0.248**	$\notin$
$D_{se\_c2} \rightarrow S_{se\_c2}$	1	0.029	**	1	0.031	**
	2	**0.139**	$\notin$	2	**0.132**	$\notin$
	3	0.458	$\notin$	3	0.434	$\notin$
	4	0.716	$\notin$	4	0.721	$\notin$
	5	0.707	$\notin$	5	0.700	$\notin$
	6	0.761	$\notin$	6	0.731	$\notin$
	7	0.802	$\notin$	7	0.779	$\notin$
	8	**0.817**	$\notin$	8	**0.797**	$\notin$
$D_{se\_c3} \rightarrow S_{se\_c3}$	1	0.012	**	1	**0.020**	**
	2	**0.011**	**	2	0.017	**
	3	0.001	***	3	0.013	**
	4	0.001	***	4	0.013	**
	5	0.001	***	5	0.012	**
	6	0.001	***	6	**0.012**	**
	7	0.011	**	7	0.013	**
	8	**0.016**	**	8	0.017	**

Notes: The arrow ($\rightarrow$) denotes that the tested predictability runs from the left-hand side variable to the right-hand side variable. The symbols *, ** and *** denote existence of causality at the 10%, 5% and 1% significance level, respectively. The symbol ($\notin$) signifies no causality at the conventional levels of significance.

Table 5: Non-linear predictability for the UWT sums components (all lags)

Causality direction	H&J test			D&P test		
	lag	$p-$value	inference	lag	$p-$values	inference
$D_{su\_c1} \rightarrow S_{su\_c1}$	1	**0.909**	$\notin$	1	**0.917**	$\notin$
	2	0.854	$\notin$	2	0.844	$\notin$
	3	0.725	$\notin$	3	0.704	$\notin$
	4	0.695	$\notin$	4	0.698	$\notin$
	5	0.533	$\notin$	5	0.526	$\notin$
	6	0.490	$\notin$	6	0.491	$\notin$
	7	**0.441**	$\notin$	7	0.446	$\notin$
	8	0.448	$\notin$	8	**0.430**	$\notin$
$D_{su\_c2} \rightarrow S_{su\_c2}$	1	0.689	$\notin$	1	0.597	$\notin$
	2	**0.762**	$\notin$	2	**0.699**	$\notin$
	3	0.361	$\notin$	3	0.246	$\notin$
	4	0.304	$\notin$	4	**0.197**	$\notin$
	5	0.408	$\notin$	5	0.303	$\notin$
	6	**0.380**	$\notin$	6	0.295	$\notin$
	7	0.385	$\notin$	7	0.308	$\notin$
	8	0.472	$\notin$	8	0.355	$\notin$
$D_{su\_c3} \rightarrow S_{su\_c3}$	1	**0.027**	**	1	**0.021**	**
	2	0.022	**	2	0.018	**
	3	0.017	**	3	0.017	**
	4	**0.012**	**	4	**0.012**	**
	5	0.009	***	5	0.009	***
	6	0.007	***	6	0.007	***
	7	0.005	***	7	0.008	***
	8	0.004	***	8	0.007	***

Notes: The arrow ($\rightarrow$) denotes that the tested predictability runs from the left-hand side variable to the right-hand side variable. The symbols *, ** and *** denote existence of causality at the 10%, 5% and 1% significance level, respectively. The symbol ($\notin$) signifies no causality at the conventional levels of significance.

Appendix A.2

Non-linear Causality

The theoretical underpinnings of the non-linear causality test proposed by Diks and Panchenko (2006) (D&P, hereafter) can be traced back to the work of Hiemstra and Jones (1994). D&P pointed out that the test statistic suggested by Hiemstra and Jones (1994) tends to over-reject the null hypothesis when this is actually true. D&P, to remedy the observed inconsistency, introduce a modified statistic, to reduce the risk of over-rejecting the null. To illustrate the D&P testing procedure we introduce two delay vectors $\mathbf{D}_t^{l_D}$ and $\mathbf{S}_t^{l_S}$, with $\mathbf{D}_t^{l_D} = (D_{t-l_D+1}, ..., D_t)$, $\mathbf{S}_t^{l_S} = (S_{t-l_S+1}, ..., S_t)$, and $l_D, l_S \geq 1$. Under the above-introduced notation the standard Granger non-causality hypothesis running from D_t to S_t is stated as:

$$S_{t+1} \middle| \left(\mathbf{D}_t^{l_D}; \mathbf{S}_t^{l_S}\right) \sim S_{t+1} \middle| \mathbf{S}_t^{l_S} \tag{10}$$

Assuming that D_t and S_t are strictly stationary and weakly dependent, the non-causality hypothesis is a statement about the invariant distribution of the $(l_D + l_S + 1)-$dimensional vector $\mathbf{w}_t = \left(\mathbf{D}_t^{l_D}, \mathbf{S}_t^{l_S}, Z_t\right)$, with $Z_t = S_{t+1}$.[12] Therefore, the validity of Eq. (10) implies that the conditional distribution of Z given $(D, S) = (d, s)$ is equivalent to Z provided that $S = s$. Given the null hypothesis, the joint probability density function $\phi_{D,S,Z}(d, s, z)$, along with its marginals, should satisfy:

$$\frac{\phi_{D,S,Z}(d, s, z)}{\phi_S(s)} = \frac{\phi_{D,S}(d, s)}{\phi_S(s)} \frac{\phi_{S,Z}(s, z)}{\phi_S(s)} \tag{11}$$

Eq. (11) suggests that D and S are two independent variables that are conditional on $S = s$ for every fixed value of s. The restated null hypothesis suggested by D&P implies that:

$$q \equiv E\left[\phi_{D,S,Z}(d, s, z)\, \phi_S(s) - \phi_{D,S}(d, s)\, \phi_{S,Z}(s, z)\right] = 0 \tag{12}$$

Moreover, let us denote as $\hat{\phi}_W(\mathbf{w}_i)$ the local density estimator of the vector $\mathbf{w}$ at $\mathbf{w}_i$, provided by $\hat{\phi}_W(\mathbf{w}_i) = (2\theta_n)^{-d_W}(n-1)^{-1}\sum_{j,j\neq i} I_{ij}^W$, where $I_{ij}^W = I(\|W_i - W_j\| \leq \theta)$, $I(\cdot)$ to be the indicator function and, finally, θ_n the

[12] Up to this point as a common presentation practice, the time subscript is dropped and we set $l_D = l_S = 1$.

bandwidth which depends on the sample size n. The $\hat{\phi}_W(\mathbf{w}_i)$ estimator allowed D&P to propose the subsequent test statistic which is actually the sample version of Eq. (12):

$$T_n(\theta_n) = \frac{(n-1)}{n(n-2)} \sum_i \left(\hat{\phi}_{D,S,Z}(D_i, Z_i, S_i)\, \hat{\phi}_S(S_i) - \hat{\phi}_{D,S}(F_i, S_i)\, \hat{\phi}_{S,Z}(S_i, Z_i) \right) \quad (13)$$

D&P showed that if $\theta_n = Cn^{-\beta}$ with $C > 0$ and $1/4 < \beta < 1/3$, then $T_n(\theta_n)$ converges to the standard normal distribution:

$$\sqrt{n}\frac{(T_n(\theta_n) - q)}{S_n} \xrightarrow{D} N(0,1) \quad (14)$$

where S_n is the estimated standard error of $T_n(\cdot)$. Summing up, the D&P approach minimizes the risk of over-rejecting the null hypothesis with respect to the testing approach suggested by Hiemstra and Jones (1994).

References

Ang, A., Bekaert, G., 2007. Stock return predictability: Is it there? Review of Financial Studies 20, 651–707.

Berkowitz, J., Giorgianni, L., 2001. Long-horizon exchange rate predictability? The Review of Economics and Statistics 83, 81–91.

Brock, W.A., Scheinkman, J.A., Dechert, W.D., LeBaron, B., 1996. A test for independence based on the correlation dimension. Econometric Reviews 15, 197–235.

Campbell, J.Y., Shiller, R.J., 1987. Cointegration and tests of present value models. Journal of Political Economy 95, 1062–1088.

Campbell, J.Y., Shiller, R.J., 1988. The dividend-price ratio and expectations of future dividends and discount factors. Review of Financial Studies 1, 195–228.

Campbell, J.Y., Shiller, R.J., 1998. Valuation ratios and the long-run stock market outlook. Journal of Portfolio Management 24, 11–26.

Cochrane, J.H., 1999. New facts in finance. Economic Perspectives: Federal Reserve Bank of Chicago 23, 36–58.

Cutler, D.M., Poterba, J.M., Summers, L.H., 1991. Speculative dynamics. Review of Economic Studies 58, 529–546.

Dergiades, T., Milas, C., Panagiotidis, T., 2015. Tweets, google trends, and sovereign spreads in the giips. Oxford Economic Papers 67, 406–432.

Diks, C., Panchenko, V., 2006. A new statistic and practical guidelines for nonparametric granger causality testing. Journal of Economic Dynamics and Control 30, 1647–1669.

Evans, G.W., 1991. Pitfalls in testing for explosive bubbles in asset prices. The American Economic Review 81, 922–930.

Fama, E.F., French, K.R., 2002. The equity premium. The Journal of Finance 57, 637–659.

Flandrin, P., Goncalves, P., Rilling, G., 2005. EMD equivalent filter bank, from interpretation to applications. World Scientific Publishing Co., Singapore. chapter 3. pp. 54–74.

Flandrin, P., Rilling, G., Goncalves, P., 2004. Empirical mode decomposition as a filter bank. Signal Processing Letters, IEEE 11, 112–114.

Froot, K.A., Obstfeld, M., 1991. Intrinsic bubbles: The case of stock prices. The American Economic Review 81, 1189–1214.

Hayes, M., 1996. Statistical Digital Signal Processing and Modeling. Wiley.

Hiemstra, C., Jones, J.D., 1994. Testing for linear and nonlinear granger causality in the stock price-volume relation. Journal of Finance 49, 1639–1664.

Hilbert, D., 1953. Grundzge einer allgemeinen Theorie der linearen Integralgleichungen. Chelsea Publishing Company, New York.

Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.C., Tung, C.C., Liu, H.H., 1998. The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis, in: Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, pp. 903–995.

Johansen, S., 1988. Statistical analysis of cointegration vectors. Journal of Economic Dynamics and Control 47, 231–254.

Kanas, A., 2005. Nonlinearity in the stock price-dividend relation. Journal of International Money and Finance 24, 583–606.

Kapetanios, G., Shin, Y., Snell, A., 2006. Testing for cointegration in nonlinear smooth transition error correction models. Econometric Theory 22, 279–303.

Kilian, L., 1999. Exchange rates and monetary fundamentals: What do we learn from long-horizon regressions? Journal of Applied Econometrics 14, 491–510.

Kilian, L., Taylor, M.P., 2003. Why is it so difficult to beat the random walk forecast of exchange rates? Journal of International Economics 60, 85–107.

Lanne, M., 2002. Testing the predictability of stock returns. Review of Economics and Statistics 84, 407–415.

Mallat, S., 2008. A Wavelet Tour of Signal Processing: The Sparse Way. Academic Press, New York.

McMillan, D.G., 2007. Bubbles in the dividend-price ratio? evidence from an asymmetric exponential smooth-transition model. Journal of Banking and Finance 31, 787–804.

Michis, A., 2014. Time scale evaluation of economic forecasts. Economics Letters 123, 330–333.

Nyquist, H., 1928. Certain topics in telegraph transmission theory. AIEE transactions 47, 617–644.

Phillips, P., Ouliaris, S., 1988. Testing for cointegration using principal components methods. Journal of Economic Dynamics and Control 12, 205–230.

Psaradakis, Z., Sola, M., Spagnolo, F., 2004. On markov error-correction models, with an application to stock prices and dividends. Journal of Applied Econometrics 19, 69–88.

Ramsey, J.B., 1999. The contribution of wavelets to the analysis of economic and financial data. Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences 357, 2593–2606.

Rapach, D.E., Wohar, M.E., 2005. Valuation ratios and long-horizon stock price predictability. Journal of Applied Econometrics 20, 327–344.

Rilling, G., Flandrin, P., Gonçalvès, P., 2003. On empirical mode decomposition and its algorithms, in: Proceedings of the 6th IEEE/EURASIP Workshop on Nonlinear Signal and Image Processing (NSIP '03), Grado, Italy.

Shiller, R.J., 1981. Do stock prices move too much to be justified by subsequent changes in dividends? American Economic Review 71, 421–436.

Shiller, R.J., 1989. Market volatility. Cambridge, MA: MIT Press.

Starck, J., Fadili, J., Murtagh, F., 2007. The undecimated wavelet decomposition and its reconstruction. IEEE Transactions on Image Processing 16, 297–309.

Summers, L.H., 1986. Does the stock market rationally reflect fundamental values? The Journal of Finance 41, 591–601.

Torous, W., Valkanov, R., Yan, S., 2004. On predicting stock returns with nearly integrated explanatory variables. The Journal of Business 77, 937–966.

Valkanov, R., 2003. Long-horizon regressions: theoretical results and applications. Journal of Financial Economics 68, 201–232.

Wolf, M., 2000. Stock returns and dividend yields revisited: A new way to look at an old problem. Journal of Business and Economic Statistics 18, 18–30.

Wu, J.L., Hu, Y.H., 2012. Price-dividend ratios and stock price predictability. Journal of Forecasting 31, 423–442.