



UNIVERSIDADE CATÓLICA PORTUGUESA

# International equity risk premium predictability in the frequency domain

Final Assignment in the modality of Dissertation presented to Universidade  
Católica Portuguesa for obtaining the degree of Master in Finance

by

Fernando Miguel Xavier Mariz

under the supervision of

(PhD) Gonçalo Faria and (PhD) Fabio Verona

Católica Porto Business School  
March of 2018



# Acknowledgments

I would like to thank everyone that supported me during this assignment.

To my parents, Fernando and Isabel, without whom I wouldn't have made it so far, I owe it all to them.

To my supervisors, Professor Gonçalo Faria and Professor Fábio Verona, for the knowledge transmitted, guidance and permanent availability along this challenging and rewarding journey.

To Benedita, for the help provided in this dissertation and for the constant presence and interest.

To all my closest family, especially my sister and friends for the motivation and support provided both directly and indirectly.



# Resumo

Nesta dissertação, estendemos o método SOPWAV de Faria e Verona (2018), que foi desenvolvido para prever o retorno das ações no mercado dos Estados Unidos da América, para mercados internacionais. O SOPWAV é baseado no método *sum of the parts* (SOP) proposto por Ferreira e Santa Clara (2011), que decompõe o retorno das ações em três partes diferentes, com o objetivo de, num primeiro momento, estimar cada uma destas separadamente e, num segundo momento, somá-las para obter a previsão do retorno das ações. O SOPWAV usa as frequências das três partes referidas, em vez das séries temporais originais (utilizadas no SOP), para as estimar, através de métodos de decomposição de ondas. O método SOPWAV permite isolar as frequências das partes dos retornos das ações que contêm maior poder preditivo. Através dos resultados obtidos, percebemos que este método melhora significativamente a previsão dos retornos das ações na amostra dos países analisados, nomeadamente, Japão, Reino Unido, França, Alemanha, Canadá, Suíça, Austrália e África do Sul. Para além de melhorar a previsão dos retornos, o método SOPWAV, de um modo geral, permite também maiores ganhos de utilidade para um investidor, quando comparados com a média histórica e o SOP; ganhos de utilidade, comprovados em estratégias de *trading*, que foram simuladas no decorrer desta investigação.

Palavras-chave: Previsão do Retorno das Ações, Mercados Internacionais, SOPWAV, Domínio da Frequência



# Abstract

In this dissertation, we extend the Faria and Verona (2018) SOPWAV method for forecasting US stock market returns to international stock markets. The SOPWAV has its roots in the sum of the parts (SOP) method proposed by Ferreira and Santa Clara (2011), that decomposes stock returns in three different parts, and forecasts them separately to obtain the estimated stock returns. The SOPWAV method uses the frequencies of those components, instead of their original time-series which are obtained through a wavelet decomposition. The SOPWAV method allows to isolate the frequencies of the stock return parts that have the highest predictive power. We found evidence that this method significantly improves predictability across the sample of countries tested (Japan, United Kingdom, France, Germany, Canada, Switzerland, Australia and South Africa) and delivered higher utility gains than the SOP method and the traditional benchmark (historical mean) in trading strategies simulated throughout this investigation.

Keywords: Stock return predictability, International stock markets, SOPWAV, frequency domain



# Contents

Acknowledgments .....	iii
Resumo .....	v
Abstract .....	vii
List of Figures .....	xi
List of Tables .....	xiii
Introduction .....	15
Literature Review .....	17
1.1. Forecasting stock returns .....	17
1.1.1. International stock returns .....	21
1.2. Frequency domain in economic and financial topics .....	24
Data and Methodology .....	29
2.1. Data description .....	29
2.2. Methodology .....	34
2.2.1. The SOP decomposition .....	34
2.2.2. Time-frequency decomposition of economic time series .....	36
2.2.3. Forecasting with the SOPWAV method .....	40
2.2.4. Out-of-sample forecasts .....	43
2.2.5. Forecast evaluation .....	44
Empirical Results .....	47
3.1. Statistical Performance .....	47
3.1.1. Out-of-sample period: 02/2000 to 10/2017 .....	48
3.1.2. Out-of-sample period: 01/1990 to 10/2017 .....	56
3.2. Robustness Tests .....	57
3.2.1. Alternative Multiple .....	57
3.2.2. Forecasting the U.S. stock returns .....	59
3.3. Trading Strategy .....	60
Conclusion .....	65
Bibliografy .....	67
Appendix .....	77



# List of Figures

**Figure 1.** MODWT MRA decomposition of the stock market return in UK..... 39



# List of Tables

<b>Table 1.</b> Out-of-sample R-squares (in percentages) for Japan .....	48
<b>Table 2.</b> Out-of-sample R-squares (in percentages) for the United Kingdom ..	49
<b>Table 3.</b> Out-of-sample R-squares (in percentages) for France .....	50
<b>Table 4.</b> Out-of-sample R-squares (in percentages) for Germany .....	51
<b>Table 5.</b> Out-of-sample R-squares (in percentages) for Canada .....	52
<b>Table 6.</b> Out-of-sample R-squares (in percentages) for Switzerland.....	53
<b>Table 7.</b> Out-of-sample R-squares (in percentages) for Australia.....	54
<b>Table 8.</b> Out-of-sample R-squares (in percentages) for South Africa.....	55
<b>Table 9.</b> Out-of-sample R-squares (in percentage) for the whole sample of countries in the forecast period of 01/1990 to 10/2017.....	56
<b>Table 10.</b> Out-of-sample R-squares (in percentage) for the whole sample of countries in the forecast period of 01/2000 to 10/2017 using the price-earnings growth rate. ....	58
<b>Table 11.</b> United States performance.....	60
<b>Table 12.</b> Annual percentage CER gains.....	63
<b>Table 13.</b> Annual Sharpe ratio gains .....	64



# Introduction

The highly debated topic of forecasting stock returns has been in the mind of researchers and practitioners of finance for quite some time.

For practitioners it is of utmost importance to understand the functioning of financial markets and especially real-time forecasted stock returns, as this provides a means to enhance asset allocation and ultimately the investor performance. From an academic standpoint, forecasting stock returns is important for many reasons, including the ability of using more realistic asset pricing models and improving the understanding of different factors that depend on stock returns.

But the task of forecasting stock returns is a very challenging one, translated in the fact that most of the existing forecasting models only explain a small part of the stock returns. Also, in line with the arbitrage pricing theory and the efficient market hypothesis, one of the problems in constructing models that can effectively forecast a small part of the stock returns is the fact that the forecast will be arbitrated away. More specifically, as investors start trading with the new information available to the public (in this case the hypothetical new model) the markets will adjust to incorporate that same information, which will lead to the disappearance of the model ability to predict stock returns (e.g., Timmermann and Granger (2004) and Timmermann (2008)).

Our work builds on the Faria and Verona (2018) SOPWAV method for forecasting stock returns, which is based on the Ferreira and Santa-Clara (2011) sum of the parts (SOP) method. In a nutshell, the SOP method consists in

decomposing the stock market return into three parts, forecasted separately, which leads to an improved forecast accuracy by exploiting the different time series persistence of the three parts.

The innovation of the SOPWAV method is that it forecasts the stock return parts by selectively summing some of the frequency-decomposed parts of the stock market returns, instead of summing the original parts (as in the SOP method).

It is shown that, for the US stock market, the SOPWAV method delivers statistically and economically significant gains for investors, outperforming both the historical mean benchmark and the SOP method. The objective of this dissertation is to extend the application of the SOPWAV methods to international stock markets (i.e. non U.S. equity markets) and to find if the out-of-sample predictability of the equity risk premium in international markets can also be improved by considering frequency-decomposed predictors. The countries selected for this analysis were Japan, United Kingdom, France, Germany, Canada, Switzerland, Australia and South Africa.

We obtained positive significant at the 1% level out-of-sample R-squares for every country under analysis as well as a demonstration of the utility gains for an investor that uses a trading strategy for those countries based on the SOPWAV.

The rest of the dissertation is structured as follows. In Chapter 1 we review related literature in forecasting returns (for both the United States and international markets), as well as using frequency decomposition methods in financial topics. In Chapter 2 we explain the data and the methodology used and in Chapter 3 we provide the empirical results, as well as the robustness tests and trading strategy

# Chapter 1

## Literature Review

The SOPWAV method relates with two branches of literature. The first one refers to out-of-sample (OOS) stock return predictability. The second one has to do with the application of wavelets decomposition methods and the use of frequency domain in financial topics. In the following two subsections we briefly review these two branches of literature.

### 1.1. Forecasting Stock Returns

Attempts to predict stock returns have a long tradition in finance, since a reliable forecast provides a crucial part in the computation of the cost of capital and the investment process in general. As pointed by Goyal and Welch (2008) “The literature is difficult to absorb. Different articles use different techniques, variables, and time periods. Results from articles that were written years ago may change when more recent data is used and some articles contradict the findings of others.”

It is possible to assume a general time-series of stock returns predictability, as pointed by Henkel *et al.* (2011), who presents the beginning of the time series as being the Fama (1965, 1970) random walk theory, in which stock returns were assumed to follow a random walk. After, the use of the short rate followed as a predictor of stock returns (Fama and Schwert (1977), Fama (1981), Geske and Roll

(1983)), then came the dividend yield predicts (Rozeff (1984), Shiller (1981)) followed by the term premium predicts (Campbell (1987), Fama (1984), Keim and Stambaugh (1986), Harvey (1988) and the default premium predicts (Chen *et al.* (1986), Keim and Stambaugh (1986)) which lead to some debate about predictability (Goetzmann and Jorion (1993), Hodrick (1992), Kim and Nelson (1993), Richardson and Stock (1989)) and eventually there was some illusion and disbelief in predictability present in the works of Bossaerts and Hillion (1999), Brennan and Xia (2005), Ang and Bekaert (2007), Cochrane (2008), Goyal and Welch (2003, 2008) and Valkanov (2003) who address some issues related to forecastability, namely lack of robustness out-of-sample.

To predict stock returns, usually a typical specification regresses an independent lagged predictor on the stock market rate of return:

$$\text{Stock Returns}_t = \beta_0 + \beta_1 x_{t-1} + \varepsilon_t \quad (1)$$

Where  $\beta_1$  is a measure of how significant is the  $x$  variable in predicting stock returns, that has been constantly explored through the years.

In the 1960s a series of studies examined the forecasting power of various technical indicators, including popular filter rules, moving averages, and momentum oscillator (Alexander (1961, 1964), Cootner (1962), Fama and Blume (1966) and Jensen and Benington (1970)).

Beginning in the late 1970s a search for predictors began, as a vast literature provided evidence that multiple economic variables predicted monthly, quarterly, and/or annual U.S. aggregate stock returns in predictive regressions. The most popular predictors used in these investigations were the dividend-price ratio (Rozeff (1984); Campbell and Shiller (1988a, 1998); Hodrick (1992); Fama and French (1988, 1989); Cochrane (2008); Lettau and Van Nieuwerburgh (2008); Pastor and Stambaugh (2009)), the earnings price-ratio (Campbell and Shiller

(1988b), (1998)), book-to-market ratio (Kothari and Shanken (1997); Pontiff and Schall (1998), nominal interest rates (Fama and Schwert (1977); Breen *et al.* (1989); Ang and Bekaert (2007)), interest rate spreads (Campbell, 1987; Fama and French, 1989), inflation (Nelson (1976); Campbell and Vuolteenaho (2004)), dividend-payout ratio (Lamont (1998)), corporate issuing activity (Baker and Wurgler (2000); Boudoukh *et al.* (2007)), consumption-wealth ratio (Lettau and Ludvigson (2001), stock market volatility (Guo (2006)), labor income (Santos and Veronesi (2006)), aggregate output (Rangvid (2006)), output gap (Cooper and Priestly (2009)), expected business conditions (Campbell and Diebold (2009)), oil prices (Driesprong *et al.* (2008)), Jagged industry portfolio returns (Hong *et al.* (2007)), and accruals (Hirshleifer *et al.* (2009)).<sup>1</sup>

Furthermore, Pesaran and Timmermann (1995) address the importance of model uncertainty and parameter instability when forecasting stock. This makes the selection of a consistently robust predictive model through time a very challenging task.

Moreover, although business cycles are not considered in this dissertation they are naturally linked with stock returns forecast. This results from the fact that asset returns are functions of an extensive set of macro and microeconomic variables linked with the business cycles fluctuations

Related with this, Henkel *et al.* (2011) investigated the stock return predictability in G7 countries and found a strong and robust link between the extent of aggregate return predictability and the business cycle in all of the G7 countries excepts Germany and also that “the short-horizon performance of aggregate return predictors such as the dividend yield and the short rate appears non-existent during business cycle expansions but sizable during contractions”.

---

<sup>1</sup> The authors considered are just a representative sample, as there are many more investigations that uses the predictors above mentioned

These results are in line with the ones presented in Rapach, Strauss and Zhou (2010), which found that stock return predictability is countercyclical.

Still, most readers are left with the impression that prediction works, or as Lettau and Ludvigson (2001) state: “It is now widely accepted that excess returns are predictable by variables such as dividend-price ratios, earnings-price ratios, dividend-earnings ratios, and an assortment of other financial indicators.”

Noting that predictive models require out-of-sample validation, Goyal and Welch (2008) demonstrated how most of the models up until that time performed poorly out-of-sample. As forecast models require out-of-sample predictability, to overcome this issue, Faria and Verona (2018) point out that “researchers have since then turned their attention to improving the out-of-sample forecastability of stock returns, exploring two different avenues”. The first avenue of research to overcome poor out-of-sample results focuses on developing and testing new predictors, while the second on improving existing forecast methods.

Works that relate with testing new predictors include Bollerslev *et al* (2009) that test the variance risk premium, Rapach *et al* explores the role of lagged U.S. market returns for the out-of-sample predictability for other countries, Li *et al* (2013) demonstrates that the implied cost of capital predicts stock returns predicts stock returns, Jordan *et al* (2014) uses two new technical variables (the ratio of rising stocks to fallers and the change in trading volume) and Neely *et al.* (2014) consider the relevance of several technical indicators to serve as complementary predictors for the common ones within the literature.

Some researches made that relate to improving existing models to overcome poor out-of-sample performance are for example Ferreira and Santa Clara (2011), who introduced the SOP method, Dangl and Halling (2012) who evaluated predictive regressions that considered the time-variation of coefficients, Pettenuzzo, Timmermann and Valkanov (2014) that proposed imposing constraints on time series forecasts of the equity premium, Lima and Meng (2017)

proposes a quantile combination approach and Faria and Verona (2018) who introduced the SOPWAV method.

### 1.1.1. International Stock Returns

Most of the literature regarding this subject is focused on the U.S. stock market returns, although there are exceptions.

Harvey (1991) examines the discrepancy in average stock returns across industrialized countries by measuring the conditional risk of 17 countries and concludes that “the results show that countries' risk exposures help explain differences in performance.” Meanwhile, Bekaert and Hodrick (1992) characterize the predictable components of returns on major international exchange markets using lagged excess returns, dividend yields, and forward premiums as instruments. Campbell and Hamao (1992) study the integration of long-term capital markets in the U.S and Japan, by using the monthly excess return predictability of those countries, and they find that U.S. variables (such as the lag-returns) help to forecast excess Japanese stock returns. This led to the consideration of U.S. lagged returns as being a good predictor of non-U.S. financial markets, as studied by Rapach, Strauss and Zhou (2013), who found that U.S. has a leading role, given that its lagged returns significantly predict returns in numerous non-U.S. industrialized countries, while lagged non-U.S. returns display limited predictive ability with respect to U.S. returns.

Ferson and Harvey (1993) investigated predictability in national equity market returns, and its relation to global economic risks in 18 equity markets and demonstrate how to consistently estimate the fraction of the predictive variation that is captured by an asset pricing model for the expected returns.

With the exception of Rapach *et al* (2013), the investigations provided above were among the first significant international stock return predictability studies conducted.

More recently, Schmeling (2009) examined whether consumer confidence affected expected stock returns in 18 industrialized countries and found that sentiment negatively forecasts aggregate stock market returns on average across countries and that when sentiment is high, future stock returns tend to be lower and vice versa. Engsted and Pedersen (2010) provide international evidence on stock returns and dividend growth predictability on long-term data from Denmark, Sweden and the United Kingdom and show that predictability patterns in three European stock markets are in many ways different from what characterize the US stock market. Hjalmarsson (2010) conducts tests in search of stock return predictability in the largest data set analysed so far, using four common forecasting variables (the dividend-price ratio, earnings-price ratio, short interest rate and the term spread), in a sample that comprises of 20000 monthly observations for 40 international markets. Across the sample, he concluded that traditional valuation measures such as the dividend-price and earnings-price ratios have very limited predictive ability in international data. However, he noted that using methods that do not account for the persistence and endogeneity of these variables would lead one to vastly misjudge their predictive powers and that interest rate variables are better predictors of stock returns, although their predictive power is mostly evident in developed markets. Narayan *et al* (2014) test for predictability of excess stock returns for 18 emerging markets and finds some evidence of in-sample predictability for 15 countries. And also shows that investors in most countries where short-selling is prohibited could make significant gains if limited borrowing and short-selling were allowed.

Henkel *et al.* (2011) carries out an investigation on the G7 countries to prove that “the short-horizon performance of aggregate return predictors such as the dividend yield and the short rate appears non-existent during business cycle expansions but sizable during contractions. This phenomenon appears related to

countercyclical risk premiums as well as the time-variation in the dynamics of predictors. Our empirical model outperforms the historical average out-of-sample in the US, but the results throughout the G7 are mixed.”

Thomadakis (2016) focuses on forecasting German stock returns and finds that the term spread has the in-sample ability to predict stock returns, delivering consistent out-of-sample forecast gains relative to the historical average, and also that combination forecasts do not appear to offer a significant evidence of consistently beating the historical average forecasts of the stock returns. This evidence of both in-sample and out-of-sample predictability of stock returns from the term spread is in line with some literature on the subject for the U.S. (e.g. Faria and Verona (2017)).

Jordan, Vivian and Wohar (2014) examines in-sample (INS) predictability but focuses on out-of-sample (OOS) forecasting of stock returns while using a big sample of 14 European countries with differing characteristics, that had none or very little prior OOS forecasting evidence. Jordan *et al.* (2014) also consider three types of predictor variables (fundamental ratios, macro variables, and technical variables) can forecast stock returns. Additionally, he examined if methods proposed to overcome poor return predictability can improve forecast accuracy or increase economic value. Lawrenz and Zorn (2017) conduct an analysis on 27 equity indices by taking the perspective of international asset allocation. The paper intends to test if predictive regressions conditional on time-series and cross-sectional information can improve forecasts of stock index returns by using different current price-to-fundamental ratios as predictors and by conditioning the sample on the indicator if time-series and cross-section deliver consistent or opposing signals.

Most of these studies estimate in-sample predictive regressions for non-U.S. country returns by using a variety of domestic and/or U.S. variables as predictors.

One of the main conclusions from these investigations is that stock returns offer the same degree of predictability in an international context as in the U.S.

However, not all studies hold up when assessing out-of-sample performance; of the ones previously cited, only Hjalmarrson (2010), Henkel *et al.* (2011), Rapach *et al.* (2013), Thomadakis (2016), Jordan, *et al.* (2014), Narayan *et al.* (2014) and Lawrenz and Zorn (2017) conduct out-of-sample tests with promising results in some of the countries analysed.

Given the literature on the subject, we can assume that in addition to the U.S. aggregate market, there is significant out-of-sample evidence of stock return predictability for other countries as well. Moreover, based on the out-of-sample predictability in international and cross-sectional returns, an investor can produce sizable utility gains from an asset allocation perspective.

We will analyse the out-of-sample forecasting performance of several international countries when using frequency domain based model (SOPWAV).

## 1.2. Frequency domain in economic and financial topics

The use of frequency domain, although highly popular in other topics, is still relatively unexplored in the financial/economical areas. A particular frequency domain technique is the wavelet transform, which has been particularly useful for analysing signals that can be considered as noisy, aperiodic or intermittent. This proves to be a valuable tool when addressing investigations in areas such as “geophysics, for the analysis of oceanic and atmospheric flow phenomena, seismic signals and climatic data; in medicine, for heart rate monitoring, breathing rate variability and blood flow and pressure; in engineering, for the assessment of machine process behaviour”, Rua (2011).

The attractiveness of using this method in the study of finance and economic topics comes from an idea presented in Adisson (2017) as the data sets (stock market indices, commodity prices, exchange rates, real estate prices, growth data,

mortgages, etc.) are typically highly non-stationary and demonstrate significant complexity and involve both random processes and intermittent processes. Therefore, as a way to overcome this “problem”, wavelet decomposition proves to be a very useful tool as Ranta (2010) states “Wavelet techniques possess an inherent ability to decompose this kind of time series into several sub-series which may be associated with a particular time scale. Processes at these different time-scales, which otherwise could not be distinguished, can be separated using wavelet methods and then subsequently analysed with ordinary time series methods”. This allows an investigator to unveil relationships between economic variables in the time-frequency space and to access simultaneously how variables are related at different frequencies and how that same relationship evolved through time. In this way, "Wavelets are treated as a 'lens' that enables the researcher to explore relationships that previously were unobservable", according to Ramsey (2002), and it also allows us to “say that with wavelet methods we are able to see both the forest and the trees” as pointed out by Ranta (2010).

Gençay *et al.* (2001a) argue that wavelet methods provide insight into the dynamics of economic/financial time series beyond that of standard time series methodologies. Wavelets also work naturally in the area of non-stationary time series, unlike the Fourier methods, which are crippled by the necessity of stationarity.

The main difference between both analyses (wavelet and Fourier) is that the wavelet transform has the ability to separate the dynamics in a time series over a variety of different time horizons, thus maintaining both frequency domain and time domain information, while in the Fourier spectral analyses, time domain information is lost. Howell and Mahrt (1997) noted that using a wavelet based multi-resolution analysis (MRA) decomposition in place of the Fourier analysis

for examining a quasi-stationary turbulent time series has the advantage of matching the MRA average length to the width of a localized event.

There are several examples of wavelet methods which have a lot of potential in economics and finance. The maximal overlap discrete wavelet transform (MODWT) (Percival & Walden (2000)) is one of them, which is a modification of the ordinary discrete wavelet transform, and different from the continuous wavelet transform as Rua (2011) describes “In the continuous wavelet transform, the wavelets used are not orthogonal and the data obtained by this transform are highly correlated. In fact, in a non-orthogonal wavelet analysis an arbitrary number of scales can be used to provide a complete picture. In contrast, the discrete wavelet transform decomposes the signal into a mutually orthogonal set of wavelets.”

In the fields of economics and finance, wavelet based methods have been used for example to study the relationship between several macroeconomic variables, namely money supply and output in the first case and consumption and income in the second (Ramsey and Lampart (1998a,b). Kim and In (2003, 2005) investigate the relationship between financial variables and industrial production and between stock returns and inflation respectively. Tiwari *et al.* (2013) assess oil price dependence via wavelet analysis. Reboredo and Riveira-Castro (2014) analyse dependency between oil prices and stock returns. Fernandez (2005, 2006) studies the Capital Asset Pricing Model at different scales. Lo Cascio (2007) decomposes UK real gross domestic product via wavelet filter so as to investigate the long-run structure of the data apart from external shocks and He *et al.* (2012) use a multivariate wavelet method to investigate the dynamics of correlations for international markets. Additional research using wavelets includes the study of the relationship between oil and stock markets in the G7 countries by Khalfaoui, Boutahar and Boubaker (2015) using maximal

overlap discrete wavelet methods and between shocks in crude oil prices and the stock market in China by Huang, An, Gao and Huang (2015).

When it comes to forecast purposes, which is the focus of this dissertation, Ramsey (1999) was the first to investigate the time series predictive power of wavelet based methods. Arino (1995) focused on car sales forecasts, Wong *et al.* (2003) provided an application to exchange rates, Conejo *et al.* (2005) forecasted electricity prices, Fernandez (2007) focused on forecasting shipments of US manufactured items, and more recently, Berger (2016) separated short-run noise from long-run trends and assessed the relevance of each frequency for volatility forecasting. Rua (2011) proposed a wavelet based multiscale principal component analysis to forecast GDP growth and inflation and found that significant predictive short-run improvements can be obtained with wavelet decomposition in combination with factor-augmented models.

Hsieh, Hsiao and Yeh (2010) used a wavelet decomposition to forecast U.S stock returns, and after applying several statistical methods and soft computing technics found that the proposed model with wavelet-based approach greatly outperforms the other methods tested, for which there was a robustness check to international markets (Japan and the United Kingdom).



# Chapter 2

## Data and Methodology

In this dissertation the focus is on the out-of-sample predictability of monthly stock returns for different international markets.

The methodology is the SOPWAV method proposed by Faria and Verona (2018), defined in the joint time-frequency domain, being a generalization of the Ferreira and Santa-Clara (2011) sum-of-the-parts method for forecasting stock market returns.

In sub-sections 2.1 and 2.2 we describe the data and the methodology used, respectively.

### 2.1 Data Description

The sample covers monthly data for 9 countries (Japan, United Kingdom, Germany, France, Canada, Switzerland, Australia, South Africa and United States<sup>2</sup>). The selection criteria was to include G20 countries for which there was data regarding the different predictors and time series of Thomson database Total Return index at least since January 1973.

The sample period is from January 1973 to October 2017 for every country except Switzerland, for which the range is January 1974 to October 2017.

---

<sup>2</sup> The United States were used to compare the results obtained in this dissertation with the ones obtained in Faria and Verona (2018) as a robustness exercise

The data is primarily from the Thomson Reuters Database<sup>3</sup> (data stream compiled stock indices for equity market data) for all variables except interest rate yields, for which the data is gathered from the Federal Reserve bank of St. Louis Economic Data (<https://fred.stlouisfed.org>), whose sources are the International Monetary Fund (Short-term rates) and the Organization for Economic Co-operation and Development (long term rates).

A description of the variables and their function in this investigation will be hereby provided.

- Stock returns (ri): The return index presents the growth in value of a stock holding, this way the stock returns for each country represent the total return which is adjusted for stock splits and dividends and given by the log of the “Total Return Index (RI)”. Thus  $r_t = \ln \left[ \frac{R_t}{R_{t-1}} \right]$ , where  $R_t$  are the stock total returns for country  $x$  on time  $t$ . The database codes are TOTMK\*\*(RI), where \*\* represent the country code.

- Dividend-Price ratio (dp): This is derived by calculating the total dividend amount and expressing it as a percentage of the market value for the constituents. This provides an average of the individual yields of the constituents weighted by market value. It is important to mention the difference in the empirical definition between dividend-price ratio and dividend yield. The first one is the ratio of dividend on time  $t$  and price on time  $t$ , while the second is the ratio of dividend on time  $t$  and price on  $t-1$ .

In this way the dividend-price ratio, is the log of the sum of dividends paid over the last 12 months on companies in the stock index divided by the current price of the stock index. The value from the database is given

---

<sup>3</sup> As in, for example, Jordan *et al.* (2014), McMillan and Wohar (2011), Guidolin *et al.* (2014) and Lawrenz and Zorn (2017).

in percentages and we convert into a decimal by dividing by 100, thus

$$dp = \ln \left[ 1 + \frac{\frac{\text{TOTMK}^{**}(\text{DY})}{100}}{12} \right].$$

- Price-Dividend Ratio Growth Rate (gmd):  $= \ln \left[ \frac{\text{Dividend Price}_t^{-1}}{\text{Dividend Price}_{t-1}^{-1}} \right]$

- Dividend Growth Rate (gd): The monthly growth rate of the dividends per share paid across the different nine different countries is obtained by multiplying the dividend price at time t by the index in the same time line to first obtain the dividend. Then, the growth rate is computed the same way as the gmd. The Thomson Reuters database provides the dividend price as TOTMK\*\*(DY) in percentages, so we need to divide by 100 and multiply by the market price index (PI), unadjusted stock returns (TOTMK\*\* (PI)).

- Price-earnings ratio growth rate (gm): Price earnings ratio is derived by dividing market value by total earnings, thus providing an earnings weighted average of the price-earnings ratio (PE) of the constituents. It is available on datastream with the code TOTMK\*\* (PE).

- Earnings growth rate (ge): The earnings per share of each international market is obtained by dividing the current price index over the price earnings ratio (to isolate the earnings per share) so the formula is  $\frac{\text{TOTMK}^{**}(\text{PI})}{\text{TOTMK}^{**}(\text{PE})}$  and the growth rate is constructed in the same way as the previously explained growth rates.

- Long-Term Yield (lty): This is the yield on governmental 10 year bonds obtained from the Organisation for Economic Co-operation and Development (OECD) database.
- Short-Term Yield (tbl): This is the yield on governmental 3 month three-month governmental bonds, which can be interpreted as the risk free rate. Data on short-term yields was obtained from the International Monetary Fund database of all the nine countries.
- Term Spread (tms): The difference between the long-term yield and the short-term yield.
- Earnings-Price Ratio (ep): This is the sum of earnings paid over the last 12 months on firms in the stock index divided by the current price of the stock index. The ratio is obtained by taking logs of the inverse of the price-earnings ratio, available in datastream as TOTMK\*\*(PE), so the formula is:  $\ln[1/\text{TOTMK}^{**}(\text{PE})]$
- Dividend Yield (dy): This is the sum of dividends paid over the last 12 months on firms in the stock index divided by the previous month's stock index price. This is calculated by taking the Dividend Price and multiplying it by its current price index (PI) and dividing by the previous month's price index.
- Dividend Payout Ratio (de): This variable represents the total amount of dividends paid to shareholders in comparison to the net income of the company and is computed as the difference between the log of dividends and the log of earnings.

- Price Pressure (pp): This variable was first used by Jordan *et al.* (2014) and represents the ratio of the number of companies in the index with positive market performance (winners) versus those with negative performance (losers). These values are obtained in the Thomson database as TOTMK\*\*(RS) for winners and TOTMK\*\*(FS) for losers.

Appendix I provide a preview of the univariate statistics for the main variables for every country as well as the time series for all the variables for every country.

Literature on forecasting stock returns, provides an extensive set of additional predictors such as: the stock variance (monthly sum of squared daily returns on the S&P 500 index), the book-to-market ratio (BM), net equity expansion (ratio of a 12-month moving sum of net equity issues by NYSE listed stocks to the total end-of-year market capitalization of NYSE stocks), the default yield spread (difference between BAA- and AAA-rated corporate bond yields), default return spread (long-term corporate bond return minus the long-term government bond return) and inflation.

Due to data unavailability either for some of the countries or for the full sample period, we are not using any of those potential additional predictors.

At last, Rapach *et al.* (2013) show that the U.S. lagged returns are a good predictor for foreign markets. However, we are not exploring this predictor in this dissertation because the SOP/SOPWAV framework makes it difficult to be used. Concretely, as this method is used, it implies that the stock returns are the sum of three different predictors that are forecasted separately in different ways, and the stock return component that is estimated using other literature predictors (price multiple growth rate) performed poorly when estimated with U.S. lagged returns.

## 2.2. Methodology

### 2.2.1. The SOP decomposition

The sum of the parts (SOP) method proposed by Ferreira and Santa-Clara (2011) is a method for predicting stock returns that consists in decomposing stock market returns into three different components (the dividend-price ratio, a price multiple growth rate and the price multiple denominator growth rate<sup>4</sup>) forecast those components separately, and later add them to retrieve the stock return forecast.

In this dissertation we use the price-dividend ratio in our base case scenario and, as a robustness test, we alternatively use the price-earnings ratio.

The description of both the decomposition of stock returns and the method to forecast those components will be hereby provided, in the same way as they are presented in the original papers of Ferreira and Santa-Clara (2011) and Faria and Verona (2018).

First, the total return of the stock market index of each country is decomposed into dividend yield and capital gains:

$$1 + R_{t+1} = 1 + CG_{t+1} + DY_{t+1}, \quad (2)$$

where  $R_{t+1}$  is the return obtained from time  $t$  to time  $t + 1$ ,  $CG_{t+1}$  is the capital gain and  $DY_{t+1}$  is the dividend yield.

---

<sup>4</sup> The multiples that can be used include the price-earnings ratio, the price-dividend ratio, the price-to-book ratio and the price-to-sales ratio.

The capital gains component can be written as:

$$1 + CG_{t+1} = \frac{P_{t+1}}{P_t} = \frac{\frac{P_{t+1}}{D_{t+1}}}{\frac{P_t}{D_t}} \frac{D_{t+1}}{D_t} = \frac{M_{t+1}}{M_t} \frac{D_{t+1}}{D_t} = (1 + GMD_{t+1})(1 + GD_{t+1}), \quad (3)$$

Where  $P_{t+1}$  is the stock price at time  $t + 1$ ,  $D_{t+1}$  is the dividend per share paid during the return period,  $M_{t+1}$  is the price-dividend multiple,  $GMD_{t+1}$  is the price–dividend multiple growth rate, and  $GD_{t+1}$  is the dividends growth rate.<sup>5</sup>

The dividend yield can in turn be decomposed as:

$$DY_{t+1} = \frac{D_{t+1}}{P_t} = \frac{D_{t+1}}{P_{t+1}} \frac{P_{t+1}}{P_t} = DP_{t+1}(1 + GMD_{t+1})(1 + GD_{t+1}), \quad (4)$$

where  $DP_{t+1}$  is the dividend–price ratio.

By replacing the capital gain and the dividend yield in equation (2), it is possible to write the total return as the product of the dividend–price ratio and the growth rates of the price–dividend ratio and dividends:

$$\begin{aligned} 1 + R_{t+1} &= (1 + GMD_{t+1})(1 + GD_{t+1}) + DP_{t+1}(1 + GMD_{t+1})(1 + GD_{t+1}) \\ &= (1 + GMD_{t+1})(1 + GD_{t+1})(1 + DP_{t+1}) \end{aligned} \quad (5)$$

---

<sup>5</sup> Instead of the price–dividend ratio, another price multiple such as the ones described on <sup>3</sup> could have been used. In these alternatives, we must replace the growth in dividends with the growth rate of the denominator in the multiple (i.e., earnings, book value of equity, or sales).

Taking the logs of both sides the equation (5), it is obtained:

$$r_{t+1} = \log(1 + R_{t+1}) = \text{gmd}_{t+1} + \text{gd}_{t+1} + \text{dp}_{t+1} \quad (6)$$

In this way, the log stock returns can be written as the sum of the growth rate of the price-dividend ratio, the growth rate of dividends and the dividend price ratio.

### 2.2.2 Time-frequency decomposition of economic time series

The discrete wavelet transform (DWT) multiresolution analysis (MRA) has numerous advantages in comparison to the continuous wavelet transform, as pointed out in the chapter “Literature Review”, enabling the decomposition of a time series into its constituent multiresolution (frequency) components<sup>6</sup> And a detailed explanation of the DWT MRA will now be provided.

There are two types of wavelets: (i) mother wavelets ( $\psi$ ) that capture the detail and the high-frequency components of the series and (ii) father wavelets (or scaling function) ( $\phi$ ) which capture the smooth and low-frequency part of the series, and acts like a low-pass filter becoming this way associated with the smoothing of the signal.

The orthogonal wavelet series approximation to a time series  $y_t$  with a number of observations  $N$  is defined by:

$$y_t = \sum_k s_{J,k} \phi_{J,k}(t) + \sum_k d_{J,k} \psi_{J,k}(t) + \sum_k d_{J-1,k} \psi_{J-1,k}(t) + \dots + \sum_k d_{1,k} \psi_{1,k}(t), \quad (7)$$

---

<sup>6</sup> A detailed analysis on wavelet decomposition methods and in particular the DWT MRA and MODWT MRA can be found in Rua (2011)

where  $J$  represents the number of multiresolution levels (the scales or frequencies),  $k$  defines the length of the filter and ranges from one to the number of coefficients in the corresponding component. The coefficients  $s_{J,k}$ ,  $d_{J,k}$ ,  $d_{J-1,k}$ ,  $\dots$ ,  $d_{1,k}$  are the wavelet transform coefficients, which are given by:

$$s_{J,k} = \int y_t \phi_{J,k}(t) dt$$

$$d_{j,k} = \int y_t \psi_{j,k}(t) dt ,$$

These coefficients give a measure of the contribution of the corresponding wavelet function to the signal. The functions  $\phi_{J,k}(t)$  and  $\psi_{j,k}(t)$  are generated from the father  $\phi$  and mother  $\psi$  wavelets through scaling and translation as follows

$$\phi_{J,k}(t) = 2^{-\frac{J}{2}} \phi(2^{-J}t - k)$$

$$\psi_{j,k}(t) = 2^{-\frac{j}{2}} \psi(2^{-j}t - k) ,$$

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

In a more synthetic way, the wavelet multiresolution decomposition of  $y_t$  can be rewritten as

$$y_t = y_t^{S_J} + y_t^{D_J} + y_t^{D_{J-1}} + \dots + y_t^{D_1} , \quad (8)$$

where  $y_t^{S_J} = \sum_k s_{J,k} \phi_{J,k}(t)$  is the wavelet smooth component and  $y_t^{D_j} = \sum_k d_{j,k} \psi_{j,k}(t)$ ,  $j=1,2,\dots, J$ , are the  $J$  wavelet detail components.

This equation illustrates how an original time series  $y_t$  defined in the time-domain, can be decomposed in different components, each of them also defined

in the time-domain, that represents the fluctuation of the original series in different frequency bands.

For a small  $j$ , the  $j$  wavelet detail component represents the short term dynamics of the time series, representing its higher frequency characteristics. As  $j$  increases, the  $j$  wavelet detail component accounts for lower frequency movements of the series. Lastly, the wavelet smooth component captures the lowest frequency dynamics, or its long behaviour/trend.

As done in Faria and Verona (2018), due to practical limitations of DWT we performed the wavelet decomposition of the different time series using the maximal overlap discrete wavelet transform (MODWT) because “Unlike the DWT, the MODWT i) is not restricted to any sample size, ii) is translation invariant, so that it is not sensitive to the choice of the starting point of the examined series, and iii) does not introduce phase shifts in the wavelet coefficients, so that peaks or troughs in the original time series are correctly aligned with similar events in the MODWT MRA.”

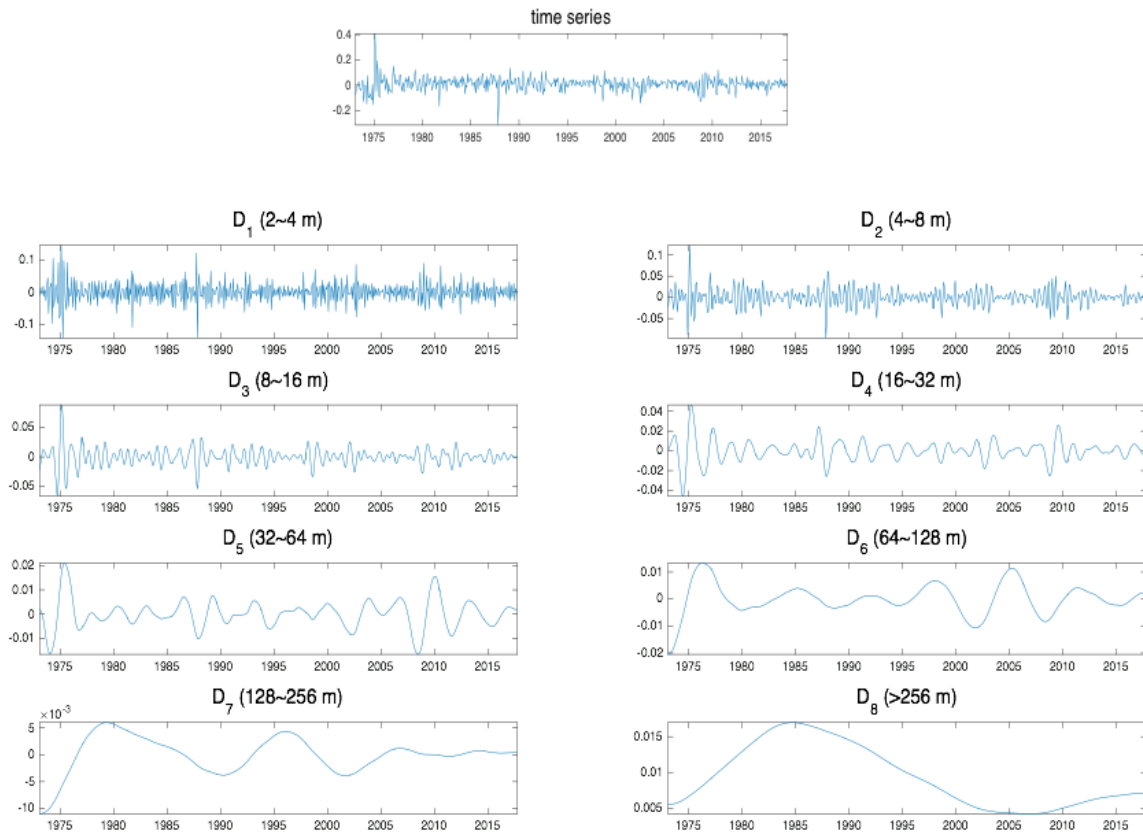
We applied the db2 wavelet filter with reflecting boundary conditions,<sup>7</sup> and applied a  $J=7$  and  $J=6$  levels of MRA:  $J=7$  for the higher in sample period ( $t_0 = 324$ ), delivering seven wavelet details ( $y_t^{D1}$  to  $y_t^{D7}$ ) and a wavelet smooth  $y_t^{S7}$ , and  $J=6$  for the shorter in-sample period ( $t_0 = 203$ ) delivering six components: ( $y_t^{D1}$  to  $y_t^{D6}$ ) and  $y_t^{S7}$ .

Since we are using monthly data, the first detail component  $y_t^{D1}$  captures the oscillations between 2 and 4 months, while detail components  $y_t^{D2}$ ,  $y_t^{D3}$ ,  $y_t^{D4}$ ,  $y_t^{D5}$ ,  $y_t^{D6}$ ,  $y_t^{D7}$  capture oscillations between 4-8, 8-16, 16-32, 32-64, 64-128 and 128-256 months respectively. Finally, the smooth component  $y_t^{S7}$ , now denoted as  $y_t^{D8}$ , captures the oscillations for a period greater than 256 months.

---

<sup>7</sup> Use of Daubechies family of filters can be problematic as they take future data into account (two-sided filter), see Murtagh *et al* (2004), in our analysis this is not a problem because we'll be doing an expanding window out-of-sample forecast.

To illustrate the decomposition of a time series through MODWT MRA, figure 1 represents the decomposition of United Kingdom's log stock returns. As expected, the lower the frequency the smoother the filtered time series. We can also observe that this method uncovers new dynamics within the time series that were not originally visible. The sample period runs from 02/1973 to 10/2017.



**Figure 1.** MODWT MRA decomposition of the stock market return in UK. It is possible to observe the original time series as well as the eight frequency-specific components ( $D_j$ ) equation (8) in which the log stock market return is decomposed.

### 2.2.3 Forecasting with the SOPWAV method

The SOPWAV method differs from the SOP when it comes to forecasting the parts of the market stock returns. While the SOP uses the time series of the different variables, the SOPWAV uses the frequency decomposed time series of the dividend-price ratio, the price multiple growth rate, the denominator of the price multiple growth rate and the seven different predictors that are used to estimate the price multiple in the extended version. The objective in using the frequency decomposed parts is to remove the noise of the time series and use the frequencies that have the highest predictive power. A detailed explanation of the SOPWAV econometric time-series method will be hereby provided, closely following the exposition in Faria and Verona (2018).

Firstly, the maximal overlap discrete wavelet transform multi-resolution analysis (MODWT MRA) decomposition (J=6 and J=7 levels)<sup>8</sup> is applied to the time series of all variables under analysis. For illustrative purposes, using the  $dp$  as an example we obtain:

$$dp_t = dp_t^{D_1} + dp_t^{D_2} + dp_t^{D_3} + dp_t^{D_4} + dp_t^{D_5} + dp_t^{D_6} + dp_t^{D_7} + dp_t^{D_8} \quad (9)$$

As it immediately results from equation (9), one of the characteristics of the MODWT MRA decomposition method is (as opposed to other frequency decomposition methods) the fact that it allows us to preserve the original time series. Since the sum of the components will give us the original series, this allows us to maintain the same information as in a standard time-series analysis.

---

<sup>8</sup> In this dissertation, as two out-of-sample periods are tested, there are two different J levels (number of frequencies that compose the time series) the j=6 level stands for a shorter in-sample period (until 12/1989) and the j=7 level is for the in-sample period until 01/2000.

Secondly, each frequency-decomposed part of the stock market return (dp, gmd and gd) is forecasted separately. The variables dp and gd are forecasted using an AR(1) process and using data only from the components themselves:

$$dp_{t+1}^{D_j} = \alpha_j + \beta_j dp_t^{D_j} + \varepsilon_{t+1} \quad \forall j = 1, \dots, 8 \quad (10)$$

$$gd_{t+1}^{D_j} = \gamma_j + \delta_j gd_t^{D_j} + \varepsilon_{t+1} \quad \forall j = 1, \dots, 8, \quad (11)$$

So that each frequency decomposed part forecast is given by:

$$E_t dp_{t+1}^{D_j} = \hat{\alpha}_j + \hat{\beta}_j dp_t^{D_j} \quad \forall j = 1, \dots, 8 \quad (12)$$

$$E_t gd_{t+1}^{D_j} = \hat{\gamma}_j + \hat{\delta}_j gd_t^{D_j} \quad \forall j = 1, \dots, 8, \quad (13)$$

Where  $\hat{\alpha}_j$ ,  $\hat{\beta}_j$ ,  $\hat{\gamma}_j$  and  $\hat{\delta}_j$  in equations (12) and (13) are the OLS estimates of  $\alpha_j$ ,  $\beta_j$ ,  $\gamma_j$  and  $\delta_j$  in equations (10) and (11), respectively.

The SOPWAV functions like the SOP in terms of the price multiple usage, so there is a baseline SOPWAV that assumes the price multiple growth as zero and the extended version, in which the multiple's frequency decomposed parts are forecasted. This is done by running OLS predictive regressions with the frequency component of the gmd as the dependent variable and each frequency component of the different predictors used (represented by  $x$ ), chosen one at a time, as the independent variable:

$$gmd_{t+1}^{D_j} = \eta_j + \lambda_j x_t^{D_j} + \varepsilon_{t+1} \quad \forall j = 1, \dots, 8, \quad (14)$$

And then we obtain the forecast of the price multiple growth as:

$$E_t gmd_{t+1}^{D_j} = \hat{\eta}_j + \hat{\lambda}_j x_t^{D_j} \quad \forall j = 1, \dots, 8, \quad (15)$$

where  $\hat{\eta}_j$  and  $\hat{\lambda}_j$  in equation (15) are the OLS estimates of  $\eta_j$  and  $\lambda_j$  in equation (14), respectively.

Thirdly, the forecasts of  $dp_{t+1}$ ,  $gd_{t+1}$  and  $gmd_{t+1}$  are obtained by summing up the forecasts of their frequency components:

$$\begin{aligned} E_t dp_{t+1} &= E_t dp_{t+1}^{D_1} + E_t dp_{t+1}^{D_2} + \dots + E_t dp_{t+1}^{D_8} \\ &= \sum_{i=1}^8 (\hat{\alpha}_i + \hat{\beta}_i dp_t^{D_i}) \end{aligned} \quad (16)$$

$$\begin{aligned} E_t gd_{t+1} &= E_t gd_{t+1}^{D_1} + E_t gd_{t+1}^{D_2} + \dots + E_t gd_{t+1}^{D_8} \\ &= \sum_{i=1}^8 (\hat{\gamma}_i + \hat{\delta}_i gd_t^{D_i}) \end{aligned} \quad (17)$$

$$\begin{aligned} E_t gmd_{t+1} &= E_t gmd_{t+1}^{D_1} + E_t gmd_{t+1}^{D_2} + \dots + E_t gmd_{t+1}^{D_8} \\ &= \sum_{i=1}^8 (\hat{\eta}_i + \hat{\lambda}_i x_t^{D_i}) \end{aligned} \quad (18)$$

Finally, and following the approach of the SOP method, the forecast of the three parts that compose the stock market returns ( $dp$ ,  $gd$  and  $gmd$ ) are summed up to obtain the one-step ahead forecast of the stock market returns. For each different predictor  $x$  to forecast  $gmd$ , the SOPWAV forecasting method is given by:

$$\begin{aligned} E_t gmd_{t+1} &= E_t gmd_{t+1}^{D_1} + E_t gmd_{t+1}^{D_2} + \dots + E_t gmd_{t+1}^{D_8} \\ &= \sum_{i=1}^8 (\hat{\eta}_i + \hat{\lambda}_i x_t^{D_i}) \end{aligned} \quad (19)$$

In Faria and Verona (2018), the hypothesis of using all the frequencies of each part at the same time was first considered (SOPWAV\_ALL), but the results were poor and not statistically relevant.<sup>9</sup> In order to overcome this issue, Faria and Verona (2018) SOPWAV method only considers those frequencies that have the highest predictive power. This is achieved by properly selecting the weights  $\omega_i$ ,  $\omega_i = \{0,1\}$ , in the following equation:

$$E_t r_{t+1} = \sum_{i=1}^8 \omega_i (\hat{\alpha}_i + \hat{\beta}_i dp_t^{D_i}) + \sum_{i=1}^8 \omega_{i+8} (\hat{\gamma}_i + \hat{\delta}_i gd_t^{D_i}) + \sum_{i=1}^8 \omega_{i+16} (\hat{\eta}_i + \hat{\lambda}_i x_t^{D_i}), \quad (20)$$

with weights  $\omega_i$  being chosen so that they maximize the statistical performance of the predictive model.<sup>10</sup>

#### 2.2.4. Out-of-sample forecasts

There is in-sample evidence in existing literature (mostly for the U.S stock market) that stock returns contain sufficiently predicting components (Campbell (2000)).

However, Goyal and Welch (2008) showed that all stock returns predictors that performed well in-sample up until then, performed very poorly out-of-sample. By other words, they would not have helped an investor with access only to available information to profitably time the market.

Out-of-sample statistical significance is a very important tool when assessing the model effectiveness, as it demonstrates how the model would have performed during the months within the sample using only the information that

---

<sup>9</sup> In this dissertation, the inclusion of all frequencies was also tested in the international markets, but the results were also poor and therefore are not presented.

<sup>10</sup> The statistical performance is measured by the out-of-sample  $R_{os}^2$  from Campbell and Thompson (2008), explained in detail in section 2.2.5

would have been accessible to an investor. This way, it is generally accepted that predicting models require out-of-sample validation, as Campbell (2008) expressed: “The ultimate test of any predictive model is its out-of-sample performance”.

In this dissertation we are focused on the out-of-sample predictability performance of the models being tested. For robustness reasons we will consider two different out-of-sample periods. We generate one-step-ahead out-of-sample forecasts for stock returns in all the countries under analysis using a sequence of expanding windows. Two different initial samples (from 02/1973 to 12/1989 and from 02/1973 to 01/2000) were used to make the first one-step-ahead out-of-sample forecast (that begins the months after the last in-sample period). Then the process remains the same until the end of the sample, obtaining a total of 334 one-step-ahead forecasts for the 02/1973-12/1989 in-sample period and 213 one-step-ahead forecasts for the 02/1973 to 01/2000 in-sample period. The two out-of-sample periods are therefore from 01/1990 to 10/2017 and from 02/2000 to 10/2017. The popularity of the out-of-sample analysis has to do with the fact that once the forecast begins, the returns that are successively being estimated incorporate all the information from the past (at least from the beginning of the sample) which means that an investor at that time would have been able to carry out the forecast with all the available information that he could possess.

### 2.2.5 Forecast evaluation

The mean square forecast error (MSFE) is a popular metric for evaluating forecast accuracy, and it is not surprising that MSFE is constantly used in studies of stock return forecastability.

In this investigation the model’s forecasting performance is evaluated using the Campbell and Thomson (2008) out-of-sample R-square ( $R_{os}^2$ ), which measures the proportional reduction of the mean squared forecast error (MSFE)

for the SOPWAV method in comparison to the same MSFE of a benchmark model (in this case the model using the historical mean, which is the standard benchmark in the literature). The  $R_{os}^2$  is given by:

$$R_{os}^2 = 1 - \frac{\sum_{t=t_0}^{T-1} (r_{t+1} - E_t r_{t+1})^2}{\sum_{t=t_0}^{T-1} (r_{t+1} - \bar{r}_t)^2}, \quad (21)$$

where  $E_t r_{t+1}$  is the SOPWAV method stock return forecast for  $t+1$ ,  $\bar{r}_t$  is the historical mean of stock market returns up to time  $t$ ,  $r_{t+1}$  is the realized stock market return for the different indexes for  $t+1$ ,  $T$  is the total number of observations in the sample and  $t_0$  is the number of observations in the initial sample. From equation (21) results that a positive  $R_{os}^2$  (negative) means that the model being tested outperforms (underperforms) the benchmark model in terms of MSFE.

As in Ferreira and Santa-Clara (2011) and Faria and Verona (2018), the global statistical significance of the  $R_{os}^2$  results is addressed in terms of the MSFE-F statistic proposed by McCracken (2007)<sup>11</sup>. The statistical significance test formula is hereby presented:

$$MSFE - F = (T - t_0) \left[ \frac{\sum_{t=s_0}^{T-1} (r_{t+1} - \bar{r}_t)^2 - \sum_{t=t_0}^{T-1} (r_{t+1} - E_t r_{t+1})^2}{\sum_{t=t_0}^{T-1} (r_{t+1} - E_t r_{t+1})^2} \right] \quad (22)$$

---

<sup>11</sup> We are interested in testing  $H_0: MSFE_0 \leq MSFE_i$  against  $H_A: MSFE_0 > MSFE_i$ , where  $MSFE_0$  is the MSFE of the historical mean forecast, while  $MSFE_i$  is the MSFE of the predictive regression, which corresponds to  $H_0: R_{os}^2 \leq 0$ , against  $H_A: R_{os}^2 > 0$ .



# Chapter 3

## Empirical results

### 3.1. Statistical Performance

In this section we report the results regarding the statistical performance of the SOPWAV method applied in the sample of eight countries under analysis. For comparison purposes we also report the results obtained by the SOP method.

Both the SOP and the SOPWAV methods are considered in a specification where only  $dp$  and  $gd$  are used to forecast stock market returns (the baseline versions), as well as in the extended version where  $gm$  is additionally considered. In the following tables it is also reported the selected frequency-specific components ( $y_t^{D_1}$  to  $y_t^{D_8}$  or  $y_t^{D_1}$  to  $y_t^{D_7}$ )<sup>12</sup> of  $dp$ ,  $gd$  and  $gmd$  that maximize the  $R_{os}^2$  for both of the out-of-sample forecasting periods (02/2000 to 10/2017 and 01/1990 to 10/2017).

As stated in Rapach, Ringgenberg and Zhou (2016), due to the presence of a huge unpredictable component in stock returns, a  $R_{os}^2$  of approximately 0,5% represents an economically meaningful degree of return predictability

---

<sup>12</sup> Depending on which out-of-sample period is being considered

### 3.1.1. Out-of-sample period: 02/2000 to 10/2017

Table 1 reports the results for the statistical performance of the SOP and SOPWAV models for forecasting stock returns in Japan. The frequencies of dp, gd and gmd that maximize the  $R_{os}^2$  when using the SOPWAV method are also reported. It is immediate to conclude that the SOPWAV method outperforms the SOP both the baseline version and for all extended versions. All the  $R_{os}^2$  obtained by the SOPWAV models are positive and statistically significance at the 1% level. The higher  $R_{os}^2$  is 5,28% by using the price pressure to predict gmd. We can also see that the baseline model underperformed the extended version in all its variables.

Model	SOP	SOPWAV	Frequencies of dp, gd and gmd									
			D1	D2	D3	D4	D5	D6	D7	D8		
Baseline	0,83*	3,96***	dp	1	0	0	0	1	1	1	1	
			gd	0	1	0	0	0	1	0	0	
Extended	0,67*	4,60***	gmd	lty	0	0	0	0	1	1	0	0
				tbl	0	0	0	1	1	1	0	0
				tms	0	0	0	0	1	1	0	0
				dy	0	0	0	0	1	1	1	0
				ep	0	0	0	0	1	0	0	0
				de	0	1	1	0	0	0	0	0
				pp	0	1	0	0	0	1	1	0

**Table 1.** Out-of-sample R-squares (in percentages) for Japan. The sample period is monthly from 02/1973 to 10/2017, and the out-of-sample period under analysis is 02/2000 to 10/2017. Asterisks denote significance of the out-of-sample MSFE-F statistics of McCracken (2007). \*\*\*, \*\* and \* denote significance at the 1%, 5% and 10% level, respectively.

Table 2 provides the results in terms of  $R_{OS}^2$  for the United Kingdom when using the SOP and SOPWAV methods, as well as the frequencies of dp, gd and gmd that maximize the  $R_{OS}^2$ . It is possible to verify the effectiveness of the SOPWAV, when comparing to the SOP and also the better performance when using the extended SOPWAV over the baseline version. The variable that better forecasts gmd is the term spread, although all the variables in the extended version perform exceptionally well.

Model	SOP	SOPWAV	Frequencies of dp, gd and gmd									
			D1	D2	D3	D4	D5	D6	D7	D8		
Baseline	2,69***	6,09***	dp	0	1	0	0	0	1	1	1	
			gd	1	1	1	0	0	0	0	0	0
Extended	lty	2,36***	6,74***	0	0	0	1	1	0	0	0	
	tbl	2,45***	6,94***	0	1	0	1	1	0	1	0	
	tms	2,74***	7,39***	0	1	0	1	1	0	1	0	
	dy	3,28***	7,30***	gmd	1	0	0	0	1	1	1	0
	ep	2,89***	6,21***	0	0	0	0	1	0	0	0	
	de	2,61***	6,20***	0	0	0	0	1	1	0	0	
	pp	2,71***	7,21***	0	0	0	0	0	1	1	0	

**Table 2.** Out-of-sample R-squares (in percentages) for the United Kingdom. The sample period is monthly from 02/1973 to 10/2017, and the out-of-sample period under analysis is 02/2000 to 10/2017. Asterisks denote significance of the out-of-sample MSFE-F statistics of McCracken (2007). \*\*\*, \*\* and \* denote significance at the 1%, 5% and 10% level, respectively.

The results for France are presented in Table 3. Again the empirical evidence further reinforces the superior performance of the SOPWAV method versus the original SOP. It is also important to note that the frequency-specific components of  $gd$  were not used, which means that for France, the inclusion of  $gd$  is not relevant in the model, since its frequencies do not help to increase the  $R_{OS}^2$ .

Model	SOP	SOPWAV	Frequencies of dp, gd and gmd									
			D1	D2	D3	D4	D5	D6	D7	D8		
Baseline	0,93*	2,52***	dp	1	0	0	0	0	1	1	1	
			gd	0	0	0	0	0	0	0	0	0
Extended	lty	0,70*	5,02***	0	0	0	1	1	1	0	0	
	tbl	0,62*	4,92***	0	1	0	1	1	1	1	0	
	tms	1,04**	3,56***	1	1	1	0	1	0	1	0	
	dy	1,35**	3,03***	gmd	1	0	0	0	1	0	1	0
	ep	0,84*	2,80***	0	0	0	0	1	0	0	0	
	de	1,42**	5,04***	0	0	1	1	0	1	1	0	
	pp	1,20**	3,49***	0	0	0	0	0	1	0	0	

**Table 3.** Out-of-sample R-squares (in percentages) for France. The sample period is monthly from 02/1973 to 10/2017, and the out-of-sample period under analysis is 02/2000 to 10/2017. Asterisks denote significance of the out-of-sample MSFE-F statistics of McCracken (2007). \*\*\*, \*\* and \* denote significance at the 1%, 5% and 10% level, respectively.

Table 4 exhibits the results for the German stock market. As it happens with the forecasting of French stock market returns, there were no frequency-specific components of  $gd$  used to predict stock returns.

Nonetheless, when analysing Germany, we can conclude that the SOP method is not a good method for forecasting its results, because although presenting some positive  $R_{os}^2$ , none of them proved to be statistical significance at any level. On the contrary, both the baseline and the extended SOPWAV method provided positive  $R_{os}^2$ , either with 5% or 1% significant depending on the variables.

Model	SOP	SOPWAV	Frequencies of dp, gd and gmd								
			D1	D2	D3	D4	D5	D6	D7	D8	
Baseline	-0,12	1,02**	dp	1	0	0	0	0	1	1	1
			gd	0	0	0	0	0	0	0	0
Extended	lty	-0,35	2,45***	0	0	0	1	1	1	1	0
	tbl	-0,15	2,82***	1	0	1	1	1	1	1	0
	tms	0,12	1,58**	0	0	0	0	1	1	1	0
	dy	0,09	1,24**	0	0	0	0	1	0	1	0
	ep	-0,03	1,21**	0	0	0	0	0	1	0	0
	de	-0,10	1,76**	1	0	0	0	1	1	0	0
	pp	0,16	2,09***	0	0	1	0	0	1	0	0

**Table 4.** Out-of-sample R-squares (in percentages) for Germany. The sample period is monthly from 02/1973 to 10/2017, and the out-of-sample period under analysis is 02/2000 to 10/2017. Asterisks denote significance of the out-of-sample MSFE-F statistics of McCracken (2007). \*\*\*, \*\* and \* denote significance at the 1%, 5% and 10% level, respectively.

Table 5 display the SOP and SOPWAV statistical performance for the forecasting of Canada's stock market returns. The  $R_{os}^2$  for the SOP method are negative in all its versions, which means that a predictive regression based on the historical mean outperforms the SOP method for this stock market. In terms of SOPWAV it is clear that it outperforms the historical mean and SOP method, delivering positive  $R_{os}^2$  in the baseline and extended versions. Once again the baseline SOPWAV underperformed the extended in all its variables, and the ones that show higher predictive performance are de and pp, that have a statistical significance level of 1%.

Model	SOP	SOPWAV	Frequencies of dp, gd and gmd									
			D1	D2	D3	D4	D5	D6	D7	D8		
Baseline	-0,48	0,66*	dp	1	0	0	1	1	1	1	1	
			gd	0	0	1	0	0	0	0	0	0
Extended	lty	-0,61	0,81*	0	0	1	0	0	1	0	0	
	tbl	-0,84	0,95*	0	0	1	0	0	1	0	0	
	tms	-0,52	1,31**	0	1	0	0	0	0	1	0	
	dy	-0,23	1,21**	gmd	0	0	0	1	1	0	1	0
	ep	-0,45	1,03**	0	0	0	1	0	0	0	0	
	de	-0,56	2,27***	1	0	1	0	1	1	1	0	
	pp	-0,20	2,13***	0	0	1	0	0	1	1	0	

**Table 5.** Out-of-sample R-squares (in percentages) for Canada. The sample period is monthly from 02/1973 to 10/2017, and the out-of-sample period under analysis is 02/2000 to 10/2017. Asterisks denote significance of the out-of-sample MSFE-F statistics of McCracken (2007). \*\*\*, \*\* and \* denote significance at the 1%, 5% and 10% level, respectively.

The results for Switzerland are displayed in Table 6. This is a clear example of the superior ability for the SOPWAV to predict stock returns, with this country being the one for which the SOPWAV methods delivers the highest  $R_{os}^2$  in the whole sample under analysis (9,45%). This performance is particularly impressive considering that for Switzerland the the SOP method delivers negative  $R_{os}^2$  in all its forms. Once again it is proved that the inclusion of gmd enhances the ability for the SOPWAV to predict stock returns, as all its variables outperform the baseline version.

Model	SOP	SOPWAV	Frequencies of dp, gd and gmd									
			D1	D2	D3	D4	D5	D6	D7	D8		
Baseline	-1,14	4,94***	dp	1	0	0	0	0	1	1	1	
			gd	0	1	0	0	0	1	0	0	
Extended	lty	-1,92	7,53***	0	0	0	0	1	1	1	0	
	tbl	-1,75	7,58***	0	0	0	1	1	1	1	0	
	tms	-1,25	6,65***	0	0	1	1	1	1	1	0	
	dy	-1,08	5,66***	gmd	0	0	0	0	1	1	1	0
	ep	-0,94	5,35***	0	0	0	1	0	1	0	0	
	de	-1,86	6,58***	0	0	0	0	1	1	1	0	
	pp	-0,38	9,45***	0	1	0	1	1	0	0	0	

**Table 6.** Out-of-sample R-squares (in percentages) for Switzerland. The sample period is monthly from 01/1974 to 10/2017, and the out-of-sample period under analysis is 02/2000 to 10/2017. Asterisks denote significance of the out-of-sample MSFE-F statistics of McCracken (2007). \*\*\*, \*\* and \* denote significance at the 1%, 5% and 10% level, respectively.

Table 7 reports the results for Australia, where again it is possible to observe the power of the frequency-decomposed components of the stock market returns. All the results for Australia for the SOPWAV (baseline and extended versions) were significant at the 1% level. The variable that better forecasts stock returns is the dividend payout ratio.

Model	SOP	SOPWAV	Frequencies of dp, gd and gmd								
			D1	D2	D3	D4	D5	D6	D7	D8	
Baseline	0,40	2,25***	dp	1	0	0	0	1	1	1	1
			gd	1	1	0	0	0	0	0	0
Extended	lty	0,19	3,41***	0	0	0	0	1	1	0	0
	tbl	0,00	4,34***	0	1	0	1	1	1	0	0
	tms	0,72*	3,10***	0	0	0	1	1	0	0	0
	dy	0,83*	3,40***	0	0	0	0	1	1	0	0
	ep	0,53	2,55***	0	0	0	1	0	1	0	0
	de	0,27	4,60***	1	0	1	1	1	1	0	1
	pp	0,66*	4,27***	0	0	0	0	1	0	1	0

**Table 7.** Out-of-sample R-squares (in percentages) for Australia. The sample period is monthly from 02/1973 to 10/2017, and the out-of-sample period under analysis is 02/2000 to 10/2017. Asterisks denote significance of the out-of-sample MSFE-F statistics of McCracken (2007). \*\*\*, \*\* and \* denote significance at the 1%, 5% and 10% level, respectively.

At last, in Table 8 are reported the forecasting results for the South African stock market. This is the market where the statistical performance of the SOPWAV method is less robust. The baseline version delivers results that are not statistically significant. All extended versions of the SOPWAV deliver positive and statistically significant  $R_{os}^2$ . This performance is globally superior than the one attained by the SOP method.

Model	SOP	SOPWAV	Frequencies of dp, gd and gmd								
			D1	D2	D3	D4	D5	D6	D7	D8	
Baseline	0,45	0,42	dp	0	0	0	1	0	0	1	0
			gd	1	1	0	0	0	0	0	0
Extended	lty	0,43	0,82*	0	0	0	0	1	0	1	0
	tbl	0,13	1,88**	0	0	1	1	0	0	1	0
	tms	-0,02	2,85***	1	0	0	0	0	0	0	0
	dy	0,71*	0,83*	0	0	0	0	0	0	1	0
	ep	0,85*	0,77*	0	0	0	1	0	0	1	0
	de	0,17	0,75*	0	1	1	0	0	0	0	0
	pp	0,62*	1,64**	0	1	0	0	0	1	0	0

**Table 8.** Out-of-sample R-squares (in percentages) for South Africa. The sample period is monthly from 02/1973 to 10/2017, and the out-of-sample period under analysis is 02/2000 to 10/2017. Asterisks denote significance of the out-of-sample MSFE-F statistics of McCracken (2007). \*\*\*, \*\* and \* denote significance at the 1%, 5% and 10% level, respectively.

Overall, results disclosed for the markets under analysis support the strong performance of the SOPWAV model beyond the US market (as reported by Faria and Verona, 2018). There are clear forecasting gains when using frequency-specific components to predict the parts of the stock returns.

It is also curious to observe that across the sample of 8 countries under analysis, the frequencies of dp and gd that in most of the cases maximize the  $R_{os}^2$  are usually the higher ( $D_1$  and  $D_2$ ) and the lower ( $D_6$  to  $D_8$ ) ones which means that both the dp and gd are useful when predicting higher and lower frequencies of stock market returns. On the other hand, the relevant frequencies of gmd for

forecasting purposes, are generally those from  $D_3$  to  $D_6$  i.e., this part is helpful for predicting medium frequencies of stock market returns.

This provides valuable insight on the effectiveness of the SOPWAV method, as its components (dp, gd and gmd) follow the dynamics of the effective stock market returns much more efficiently than historical mean or SOP methods.

### 3.1.2. Out-of-sample period: 01/1990 to 10/2017

Overall, the results for this forecast period in terms of  $R_{os}^2$  are worse than the 02/2000 - 10/2017 period even though the SOPWAV still outperforms the SOP in most its variables, and still delivers positive  $R_{os}^2$  for every country on at least one variable. The results for the SOPWAV performance in this forecast period are presented in Table 9:

Country	Japan	United Kingdom	France	Germany	Canada	Switzerland	Australia	South Africa	
Baseline	3,09***	2,14***	0,28*	0,36*	-1,25	0,95**	-0,31	-1,30	
Extended	lty	3,48***	3,37***	4,05***	2,20***	-0,80	3,84***	3,57***	-1,40
	tbl	3,22***	2,54***	1,62***	2,44***	0,38*	3,73***	0,52**	-1,04
	tms	3,25***	2,52***	1,27***	1,13**	0,40*	2,71***	0,51*	0,85**
	dy	3,46***	2,29***	0,35*	0,73**	0,21*	2,30***	0,06	-1,15
	ep	3,35***	2,14***	0,59**	0,71**	0,27*	3,23***	-0,12	-0,81
	de	3,18***	2,19***	1,90***	1,14**	-0,58	1,68***	0,50*	0,73**
	pp	2,62***	2,46***	0,59**	1,70***	-0,10	4,14***	-0,45	-0,99

**Table 9.** Out-of-sample R-squares (in percentage) for the whole sample of countries in the out-of-sample period of 01/1990 to 10/2017. Asterisks denote significance of the out-of-sample MSFE-F statistics of McCracken (2007). \*\*\*, \*\* and \* denote significance at the 1%, 5% and 10% level, respectively.

As it is possible to see, the SOPWAV keeps its high predictive power for most countries, mainly Japan, United Kingdom, France, Germany and Switzerland, although the results for Canada, Australia and South Africa are not as strong as the out-of-sample period of 02/2000 to 10/2017.

## 3.2. Robustness Tests

### 3.2.1. Alternative Multiple

As stated before, the SOP and the SOPWAV methods can be implemented making use of various price multiples. In Faria and Verona (2018) and Ferreira and Santa Clara (2011) the first multiple used was price-earnings ratio, while their robustness tests involved the use of the price-dividend ratio.

In this dissertation it was the other way around, because earnings were more volatile than dividends during the 2007 – 2009 sample period, (mainly because of the 2008 financial crisis). Therefore, the main analysis of the different countries is with the price-dividend multiple, while the robustness test uses the price-earnings multiple. The SOP/SOPWAV formula for decomposing stock returns is the same as the one presented in equation (3) apart from the fact that price-dividend growth rate and dividends growth rate are replaced by price-earnings growth rate ( $gm$ ) and the earnings growth rate ( $ge$ ) respectively:

$$\begin{aligned} 1 + CG_{t+1} &= \frac{P_{t+1}}{P_t} = \frac{\frac{P_{t+1}}{E_{t+1}} E_{t+1}}{\frac{P_t}{E_t} E_t} = \frac{M_{t+1} E_{t+1}}{M_t E_t} \\ &= (1 + GM_{t+1})(1 + GE_{t+1}), \end{aligned} \tag{23}$$

where  $M_{t+1} = \frac{P_{t+1}}{E_{t+1}}$  is the price-earnings multiple,  $GM$  the price-earnings multiple growth rate and  $GE$  the earnings growth rate.

The process of forecasting both the  $gm$  and the  $ge$  is the same as using the  $gmd$  and  $gd$ , so the  $gm$  is predicted by running OLS regressions with the multiple itself as the dependant variable, and each of the different variables previously

presented as the independent variables. Table 9 provides the results for the  $R_{os}^2$  using the price-earnings multiple.

Country	Japan	United Kingdom	France	Germany	Canada	Switzerland	Australia	South Africa	
Baseline	1,09**	3,26***	3,13***	1,02**	0,46	2,50***	3,52***	0,93*	
Extended	lty	1,34**	4,23***	4,47***	1,72**	0,44	4,59***	4,57***	1,26**
	tbl	1,21**	4,11***	4,45***	2,2***	0,52	4,87***	4,97***	2,68***
	tms	1,44**	4,36***	4,07***	1,27**	1,11**	4,35***	3,85***	3,52***
	dy	0,91*	4,99***	4,06***	1,56**	2,44***	3,79***	4,78***	1,18**
	ep	1,37**	4,36***	3,86***	1,25**	1,36**	3,25***	4,37***	1,07**
	de	1,63**	3,87***	4,58***	1,45**	1,28**	6,02***	4,43***	0,53*
	pp	3,66***	4,64***	4,53***	2,09***	2,01***	7,38***	5,36***	1,92***

**Table 10.** Out-of-sample R-squares (in percentage) for the whole sample of countries in the forecast period of 01/2000 to 10/2017 using the price-earnings growth rate. Asterisks denote significance of the out-of-sample MSFE-F statistics of McCracken (2007). \*\*\*, \*\* and \* denote significance at the 1%, 5% and 10% level, respectively.

The results are mixed, with the price-earnings multiple based models outperforming<sup>13</sup> the dividend-price multiple (in terms of  $R_{os}^2$ ) based models in some countries like Australia, South Africa and Canada while for Japan, United Kingdom, France, Germany and Switzerland stock returns the dividend-price multiple based model delivers better results.

The alternative multiple was also tested in the out-of-sample period of 01/1990-10/2017 for all the countries and in general, similarly to the dividend-price multiple also performed much worse than the 02/2000-10/2017 out-of-sample period (Canada for example delivered negative  $R_{os}^2$  to all of the variables).

<sup>13</sup> In this case, outperforming is considered as long as one of the variables forecasting the price multiple delivers the highest  $R_{os}^2$  to the overall model of all the variables in consideration

Overall we can say that the price-dividend multiple had a better performance in the sample and periods analysed, although the price-earnings multiple also performed well, confirming the robustness of the SOPWAV model regarding the concrete multiple to be used in its construction.

### 3.2.2. Forecasting the U.S. stock returns

At last, we used the SOPWAV method to forecast stock returns in the United States. This was done in order to compare the results with the ones obtained in Faria and Verona (2018), and also with those obtained for the remaining countries under analysis.

The sample period in Faria and Verona (2018) was 1927-2015. Our findings<sup>14</sup> show that once again the 02/2000-10/2017 out-of-sample period performed much better than the 01/1990-10/2017 out-of-sample period.

In table 11 are disclosed the results in terms of  $R_{os}^2$  for the SOPWAV method using both price multiples (price-dividend and price-earnings) for the United States stock market. The results obtained by Faria and Verona (2018) with a sample period from 12/1927 to 12/2015 (out-of-sample period of 01/1950 to 12/2015) are the first two tables of the respective multiples, while the second tables provide the result obtained in this dissertation, which considers different data source and sample period (from 02/1974 to 10/2017, with the out-of-sample period of 02/2000 to 10/2017). Disclosed evidence clearly supports the strong statistical performance of the SOPWAV model for the forecasting of the US stock returns.

---

<sup>14</sup> to note that the price-earnings ratio was also tested

United States of America		Price-dividend multiple		Price-earnings multiple	
		Faria and Verona (2018)	SOPWAV	Faria and Verona (2018)	SOPWAV
Baseline		0,97***	2,17***	0,97***	1,90***
Extended	lty	2,37***	2,10***	2,6***	1,79***
	tbl	1,24***	2,23***	2,41***	1,93***
	tms	1,03***	4,02***	1,83***	3,74***
	dy	n.a.	2,76***	n.a.	3,12***
	ep	0,99***	2,57***	1,3***	2,45***
	de	1,01***	2,63***	1,56***	2,70***
	pp	n.a.	2,27***	n.a.	1,98***

**Table 11.** United States performance. Out-of-sample R-squares (in percentage) for the United States including the results of this dissertation and the ones in Faria and Verona using both the price-earnings growth rate and the price-dividend growth rate. The out-of-sample period is of 02/2000 to 10/2017 in this dissertation and 01/1950 to 12/2015 in Faria and Verona (2018). Asterisks denote significance of the out-of-sample MSFE-F statistics of McCracken (2007). \*\*\*, \*\* and \* denote significance at the 1%, 5% and 10% level, respectively.

### 3.3. Trading Strategy

In order to analyse the economic value of the SOPWAV method, we followed the approach of previous research (e.g. Marquering and Verbeek (2004), Campbell and Thompson (2008), Cooper and Priestly (2009) Rapach *et al.* (2010), Ferreira and Santa-Clara (2011), Neely *et al.* (2012), Dangl and Halling (2012), and Faria and Verona (2018)), and study the utility gains for a mean-variance investor that uses economic variables to predict stock return forecast based on the SOPWAV, against a homonymous who just uses the historical average of stock returns when constructing portfolios. We run the portfolio analysis assuming a monthly forecasting and rebalancing period.

Each period we use different estimates of the expected returns from the SOPWAV<sup>15</sup> to calculate the optimal weight on the stock market:

$$w_t = \frac{E_t r_{t+1} - r_{f,t+1}}{\gamma \hat{\sigma}_t^2}, \quad (24)$$

where  $E_t r_{t+1}$  denotes the historical mean/SOP/SOPWAV estimation of stock market return at time  $t$  for period  $t+1$ ,  $r_{f,t+1}$  is the risk-free rate<sup>16</sup> from time  $t$  to  $t+1$  but known at time  $t$ ,  $\gamma$  is the risk aversion coefficient<sup>17</sup> and  $\hat{\sigma}_t^2$  is the variance of stock market returns, at time  $t$ , that is estimated by calculating a moving average of the variance of returns for the previous 10 years of time  $t$ . The portfolio policies could have been implemented by an investor at real time, given that it uses all the available data until time  $t$ . To note that for a more realistic analysis, the weights were adjusted to lie within -1,5 and 1,5.

We then calculate the portfolio return at the end of each month as:

$$rp_{t+1} = w_t E_t r_{t+1} + (1 - w_t) r_{f,t+1} \quad (25)$$

We run this process until the end of the sample, in order to obtain a time series of returns for each trading strategy (historical mean, SOP and SOPWAV) and for each country only considering the out-of-sample period of 02/2000-10/2017, meaning that the trading strategy lasts for 213 months.

---

<sup>15</sup> For each country there are 8 different estimates of the stock returns (one for the baseline and seven for the variables that predict  $gmd$ ), but for the trading strategy, it was only considered the estimated returns from the variable whose forecast provided the higher  $R_{os}^2$ , unlike Faria and Verona (2018) that tested the utility gains for all variables.

<sup>16</sup> Which was considered to be the 3-month governmental bond for each country.

<sup>17</sup> That is considered to be 2 in this dissertation, as common within the literature.

To evaluate the performance of the different strategies, and to give means of comparison between them, we calculated the certainty equivalent return (CER):

$$\text{CER} = \bar{r}_p - \frac{\gamma}{2} \sigma_{r_p}^2 \quad (26)$$

Where  $\bar{r}_p$  is the mean of the portfolio return and  $\sigma_{r_p}^2$  is the variance of the portfolio return, the CER can be interpreted as “the risk-free return that a mean-variance investor with a risk-aversion coefficient  $\gamma$  would consider equivalent to investing in the strategy” as Ferreira and Santa-Clara (2011) described.

The Sharpe ratio<sup>18</sup> (SR) of the portfolios was also calculated for every country and for the different strategies using the SOPWAV, the SOP and the historical mean.

Both the Sharpe ratio and the CER gains were computed in a monthly basis and then annualized. We present the Sharpe ratio gains as the difference between the SR of a trading strategy based on the SOPWAV/SOP and the SR of a trading strategy based on the historical mean forecast.

Table 12 provides the results for the CER gains while Table 13 for the Sharpe ratio gains.

---

<sup>18</sup> The mean of the portfolio returns in excess of the risk-free rate (3-month governmental bond) divided by the standard deviation of the excess portfolio return.

Country	CER gains	
	SOPWAV	SOP
Japan	10,90	0,30
United Kingdom	6,13	6,23
France	4,88	4,10
Germany	0,73	1,98
Canada	1,30	-1,10
Switzerland	6,25	1,40
Australia	0,09	0,00
South Africa	0,00	0,00

**Table 12.** Annual percentage CER gains. The presented values represent the annual certainty equivalent return percentage gains for all the countries under analysis for a trading strategy using the SOPWAV and SOP. The CER gain is being presented as the difference between the CER of one of the SOPWAV or SOPW methods and the CER of a trading strategy using predicted returns from the historical mean. The trading strategies with the different models use a risk aversion coefficient of 2.

It is possible to observe that the SOPWAV method provides higher monthly CER gains than the historical mean for all countries except South Africa.<sup>19</sup> When comparing to the SOP the results are mixed, with the SOPWAV outperforming the SOP in terms of CER gains in Japan, France, Canada, Switzerland and Australia.

---

<sup>19</sup> Given the weight limit of -1,5 to 1,5, as the South African risk free was always much higher than the expected stock returns, the weights on the risky portfolio were always -1,5 for all months and for all the trading strategies, which meant that the trading strategies using the SOPWAV, SOP and historical mean, provided the same results (for Australia there were only a few months for which the weights were not always -1,5)

Country	Sharpe Ratio gains	
	SOPWAV	SOP
Japan	0,73	0,12
United Kingdom	0,12	0,28
France	0,16	0,14
Germany	-0,01	0,06
Canada	0,12	0,01
Switzerland	0,41	0,09
Australia	0,00	0,00
South Africa	0,00	0,00

**Table 13.** Annual Sharpe ratio gains. The annual Sharpe ratio gains for all the countries under analysis for a trading strategy using the SOPWAV and the SOP. The Sharpe ratio gain is being presented as the difference between the SR of one of the SOPWAV or SOPW methods and the SR of a trading strategy using predicted returns from the historical mean. The trading strategies with the different models use a risk aversion coefficient of 2.

The Sharpe ratio gains are in line with the CER gains, although the historical mean outperformed the SOPWAV in terms of Sharpe ratio gains for Germany, the rest of the countries (except Australia and South Africa) demonstrate the annual Sharpe ratio gains for an investor trading with the SOPWAV.

The SOPWAV outperforms the SOP in Japan, France, Canada and Switzerland and is outperformed in the United Kingdom and Germany markets.

## Conclusion

In this dissertation we explore out-of-sample stock return predictability in international markets by considering a frequency domain analysis.

To conduct our investigation, we used the SOPWAV method proposed by Faria and Verona (2018), that generalizes the Ferreira and Santa-Clara (2011) sum of the parts method in which stock returns are divided in three parts, forecasted separately and then summed up to obtain the stock return forecast. More specifically, the SOPWAV proposes a selective sum of the frequency-decomposed parts of the stock market returns.

The SOPWAV method was used to predict stock returns in 8 countries (Japan, United Kingdom, France, Germany, Canada, Switzerland, Australia and South Africa) beyond the US, for the sample period between 02/1974 and 10/2015.

Two out-of-sample periods were tested (02/2000 to 10/2015 and 01/1990 to 10/2015) and across the sample of countries, there was evidence of out-of-sample predictability for all stock markets analysed for the forecasting period of 02/2000 to 10/2015, measured by the out-of-sample R-squares, who were positive and significant at the 1% level for every country.

The results are robust when performing the SOPWAV using a different price multiple and a different out-of-sample period.

At last we evaluated the economic gains associated with the use of SOPWAV method in different markets under analysis. We found relevant utility gains for

an investor using the SOPWAV method for portfolio management, in most of the markets analysed.

Two factors justify the strong statistical and economic performance of the SOPWAV method. The first one relates with the use of wavelet decomposition which makes it possible to extract and isolate the frequencies of the parts that have the highest predictive power and use them to perform the forecasts. The second one is that by only considering the relevant frequencies (those at which stock returns are predictable) of the relevant parts (those that have predictive power at relevant frequencies), the SOPWAV method is able to track the future dynamics of stock market returns.

The main limitation of this dissertation had to do with the search for data. The objective in the beginning was to perform the analysis for all the G20 countries, but as there was no data from the same source for some of the variables and for the desired sample period, the sample of countries had to be restricted.

Faria and Verona (2018) already proposed extending the wavelet-based SOPWAV forecasting method to predict other macroeconomic and financial variables. Now we add that it would be interesting to extend the set of variables being used as predictors, namely by including options, futures and other derivatives that potentially contain relevant information for forecasting stock returns.

# Bibliography

Addison, P. S. (2017). *The illustrated wavelet transform handbook: introductory theory and applications in science, engineering, medicine and finance*. CRC press.

Alexander, S. S. (1961). Price movements in speculative markets: Trends or random walks. *Industrial Management Review (pre-1986)*, 2(2), 7.

Alexander, S. S. (1964). Price Movements in Speculative Markets--Trends or Random Walks, Number 2. *IMR; Industrial Management Review (pre-1986)*, 5(2), 25.

Ang, A., & Bekaert, G. (2006). Stock return predictability: Is it there?. *The Review of Financial Studies*, 20(3), 651-707.

Baker, M., & Wurgler, J. (2000). The equity share in new issues and aggregate stock returns. *the Journal of Finance*, 55(5), 2219-2257.

Bekaert, G., & Hodrick, R. J. (1992). Characterizing predictable components in excess returns on equity and foreign exchange markets. *The Journal of Finance*, 47(2), 467-509.

Binsbergen, V., Jules, H., & Koijen, R. S. (2010). Predictive regressions: A present-value approach. *The Journal of Finance*, 65(4), 1439-1471.

Bollerslev, T., Tauchen, G., & Zhou, H. (2009). Expected stock returns and variance risk premia. *The Review of Financial Studies*, 22(11), 4463-4492.

Bossaerts, P., & Hillion, P. (1999). Implementing statistical criteria to select return forecasting models: what do we learn?. *The Review of Financial Studies*, 12(2), 405-428.

Boudoukh, J., Michaely, R., Richardson, M., & Roberts, M. R. (2007). On the importance of measuring payout yield: Implications for empirical asset pricing. *The Journal of Finance*, 62(2), 877-915.

Breen, W., Glosten, L. R., & Jagannathan, R. (1989). Economic significance of predictable variations in stock index returns. *The Journal of Finance*, 44(5), 1177-1189.

Brennan, M. J., & Xia, Y. (2005). Tay's as good as cay. *Finance Research Letters*, 2(1), 1-14.

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

Campbell, J. Y. (1987). Stock returns and the term structure. *Journal of financial economics*, 18(2), 373-399.

Campbell, J. Y., & Hamao, Y. (1992). Predictable stock returns in the United States and Japan: A study of long-term capital market integration. *The Journal of Finance*, 47(1), 43-69.

Campbell, J. Y., & Shiller, R. J. (1988). The dividend-price ratio and expectations of future dividends and discount factors. *The Review of Financial Studies*, 1(3), 195-228.

Campbell, J. Y., & Thompson, S. B. (2007). Predicting excess stock returns out of sample: Can anything beat the historical average?. *The Review of Financial Studies*, 21(4), 1509-1531.

Campbell, J. Y., & Vuolteenaho, T. (2004). Inflation illusion and stock prices. *American Economic Review*, 94(2), 19-23.

Campbell, S. D., & Diebold, F. X. (2009). Stock returns and expected business conditions: Half a century of direct evidence. *Journal of Business & Economic Statistics*, 27(2), 266-278.

Cascio, I. L. (2007). *Wavelet analysis and denoising: New tools for economists* (No. 600). Working Paper, Department of Economics, Queen Mary, University of London.

Chakrabarty, A., De, A., Gunasekaran, A., & Dubey, R. (2015). Investment horizon heterogeneity and wavelet: Overview and further research directions. *Physica A: Statistical Mechanics and its Applications*, 429, 45-61.

Chen, N. F., Roll, R., & Ross, S. A. (1986). Economic forces and the stock market. *Journal of business*, 383-403.

Cochrane, J. H. (2001). *A rehabilitation of stochastic discount factor methodology* (No. w8533). National Bureau of Economic Research.

Cochrane, J. H. (2007). The dog that did not bark: A defense of return predictability. *The Review of Financial Studies*, 21(4), 1533-1575.

Conejo, A. J., Contreras, J., Espinola, R., & Plazas, M. A. (2005). Forecasting electricity prices for a day-ahead pool-based electric energy market. *International Journal of Forecasting*, 21(3), 435-462.

Cooper, I., & Priestley, R. (2008). Time-varying risk premiums and the output gap. *The Review of Financial Studies*, 22(7), 2801-2833.

Cooper, I., & Priestley, R. (2008). Time-varying risk premiums and the output gap. *The Review of Financial Studies*, 22(7), 2801-2833.

Cootner, P. H. (1962). Stock prices: Random vs. systematic changes. *Industrial Management Review (pre-1986)*, 3(2), 24.

Dangl, T., & Halling, M. (2012). Predictive regressions with time-varying coefficients. *Journal of Financial Economics*, 106(1), 157-181.

Driesprong, G., Jacobsen, B., & Maat, B. (2008). Striking oil: another puzzle?. *Journal of Financial Economics*, 89(2), 307-327.

Engsted, T., & Pedersen, T. Q. (2010). The dividend-price ratio does predict dividend growth: International evidence. *Journal of Empirical Finance*, 17(4), 585-605.

Fama, E. F. (1965). The behavior of stock-market prices. *The journal of Business*, 38(1), 34-105.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, 25(2), 383-417.

Fama, E. F. (1981). Stock returns, real activity, inflation, and money. *The American economic review*, 71(4), 545-565.

Fama, E. F. (1984). The information in the term structure. *Journal of financial economics*, 13(4), 509-528.

Fama, E. F., & Blume, M. E. (1966). Filter rules and stock-market trading. *The Journal of Business*, 39(1), 226-241.

Fama, E. F., & French, K. R. (1988). Dividend yields and expected stock returns. *Journal of financial economics*, 22(1), 3-25.

Fama, E. F., & French, K. R. (1989). Business conditions and expected returns on stocks and bonds. *Journal of financial economics*, 25(1), 23-49.

Fama, E. F., & Schwert, G. W. (1977). Asset returns and inflation. *Journal of financial economics*, 5(2), 115-146.

Faria, G., & Verona, F. (2017). The Equity Risk Premium and the Low Frequency of the Term Spread.

Faria, G., & Verona, F. (2018). Forecasting stock market returns by summing the frequency-decomposed parts. *Journal of Empirical Finance*, 45, 228-242.

Fernandez, V. (2007). Wavelet-and SVM-based forecasts: An analysis of the US metal and materials manufacturing industry. *Resources Policy*, 32(1-2), 80-89.

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

Ferson, W. E., & Harvey, C. R. (1993). The risk and predictability of international equity returns. *Review of financial Studies*, 6(3), 527-566.

Gençay, R., Selçuk, F., & Whitcher, B. (2001). Scaling properties of foreign exchange volatility. *Physica A: Statistical Mechanics and its Applications*, 289(1-2), 249-266.

Geske, R., & Roll, R. (1983). The fiscal and monetary linkage between stock returns and inflation. *The Journal of Finance*, 38(1), 1-33.

Goetzmann, W. N., & Jorion, P. (1993). Testing the predictive power of dividend yields. *The Journal of Finance*, 48(2), 663-679.

Goyal, A., & Welch, I. (2003). Predicting the equity premium with dividend ratios. *Management Science*, 49(5), 639-654.

Guidolin, M., Hyde, S., McMillan, D., & Ono, S. (2014). Does the Macroeconomy Predict UK Asset Returns in a Nonlinear Fashion? Comprehensive Out-of-Sample Evidence. *Oxford Bulletin of Economics and Statistics*, 76(4), 510-535.

Guo, H. (2006). On the out-of-sample predictability of stock market returns. *Journal of Business* 79, 645-670.

Harvey, C. R. (1988). The real term structure and consumption growth. *Journal of Financial Economics*, 22(2), 305-333.

Harvey, C. R. (1991). The world price of covariance risk. *The Journal of Finance*, 46(1), 111-157.

He, K., Lai, K. K., & Xiang, G. (2012). Portfolio value at risk estimate for crude oil markets: A multivariate wavelet denoising approach. *Energies*, 5(4), 1018-1043.

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

Hirshleifer, D., Hou, K., & Teoh, S. H. (2009). Accruals, cash flows, and aggregate stock returns. *Journal of Financial Economics*, 91(3), 389-406.

Hsieh, T. J., Hsiao, H. F., & Yeh, W. C. (2011). Forecasting stock markets using wavelet transforms and recurrent neural networks: An integrated system based on artificial bee colony algorithm. *Applied soft computing*, 11(2), 2510-2525.

Hjalmarsson, E. (2010). Predicting global stock returns. *Journal of Financial and Quantitative Analysis*, 45(1), 49-80.

Hodrick, R. J. (1992). Dividend yields and expected stock returns: Alternative procedures for inference and measurement. *The Review of Financial Studies*, 5(3), 357-386.

Hodrick, R. J. (1992). Dividend yields and expected stock returns: Alternative procedures for inference and measurement. *The Review of Financial Studies*, 5(3), 357-386.

Hong, H., Torous, W., & Valkanov, R. (2007). Do industries lead stock markets?. *Journal of Financial Economics*, 83(2), 367-396.

Howell, J. F., & Mahrt, L. (1997). Multiresolution flux decomposition. *Boundary-Layer Meteorology*, 83(1), 117-137.

Huang, S., An, H., Gao, X., & Huang, X. (2015). Identifying the multiscale impacts of crude oil price shocks on the stock market in China at the sector level. *Physica A: Statistical Mechanics and its Applications*, 434, 13-24.

Jensen, M. C., & Benington, G. A. (1970). Random walks and technical theories: Some additional evidence. *The Journal of Finance*, 25(2), 469-482.

Jordan, S. J., Vivian, A. J., & Wohar, M. E. (2014). Forecasting returns: new European evidence. *Journal of Empirical Finance*, 26, 76-95.

Keim, D. B., & Stambaugh, R. F. (1986). Predicting returns in the stock and bond markets. *Journal of financial Economics*, 17(2), 357-390.

Khalifaoui, R., Boutahar, M., & Boubaker, H. (2015). Analyzing volatility spillovers and hedging between oil and stock markets: Evidence from wavelet analysis. *Energy Economics*, 49, 540-549.

Kim, S., & In, F. (2005). The relationship between stock returns and inflation: new evidence from wavelet analysis. *Journal of Empirical Finance*, 12(3), 435-444.

Kim, S., & In, F. H. (2003). The relationship between financial variables and real economic activity: evidence from spectral and wavelet analyses. *Studies in Nonlinear Dynamics & Econometrics*, 7(4).

Kothari, S. P., & Shanken, J. (1997). Book-to-market, dividend yield, and expected market returns: A time-series analysis. *Journal of Financial Economics*, 44(2), 169-203.

Lamont, O. (1998). Earnings and expected returns. *The Journal of Finance*, 53(5), 1563-1587.

Lawrenz, J., & Zorn, J. (2017). Predicting international stock returns with conditional price-to-fundamental ratios. *Journal of Empirical Finance*, 43, 159-184.

Leitch, G., & Tanner, J. E. (1991). Economic forecast evaluation: profits versus the conventional error measures. *The American Economic Review*, 81, 580-590.

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

Lettau, M., & Van Nieuwerburgh, S. (2007). Reconciling the return predictability evidence. *The Review of Financial Studies*, 21(4), 1607-1652.

Li, Y., Ng, D. T., & Swaminathan, B. (2013). Predicting market returns using aggregate implied cost of capital. *Journal of Financial Economics*, 110(2), 419-436.

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

Marquering, W., & Verbeek, M. (2004). The economic value of predicting stock index returns and volatility. *Journal of Financial and Quantitative Analysis*, 39(2), 407-429.

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

McMillan, D. G., & Wohar, M. E. (2011). Sum of the parts stock return forecasting: international evidence. *Applied Financial Economics*, 21(12), 837-845.

Narayan, P. K., Narayan, S., & Thuraisamy, K. S. (2014). Can institutions and macroeconomic factors predict stock returns in emerging markets?. *Emerging Markets Review*, 19, 77-95.

Neely, C. J., Rapach, D. E., Tu, J., & Zhou, G. (2014). Forecasting the equity risk premium: the role of technical indicators. *Management Science*, 60(7), 1772-1791.

Nelson, C. R. (1976). Inflation and rates of return on common stocks. *The Journal of Finance*, 31(2), 471-483.

Nelson, C. R., & Kim, M. J. (1993). Predictable stock returns: The role of small sample bias. *The Journal of Finance*, 48(2), 641-661.

Pástor, L., & Stambaugh, R. F. (2009). Predictive systems: Living with imperfect predictors. *The Journal of Finance*, 64(4), 1583-1628.

Percival, D. B., & Walden, A. T. (2006). *Wavelet methods for time series analysis* (Vol. 4). Cambridge university press.

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

Pontiff, J., & Schall, L. D. (1998). Book-to-market ratios as predictors of market returns<sup>1</sup>. *Journal of Financial Economics*, 49(2), 141-160.

Ramsey, J. B. (1999). The contribution of wavelets to the analysis of economic and financial data. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 357(1760), 2593-2606.

Ramsey, J. B. (2002). Wavelets in economics and finance: Past and future. *Studies in Nonlinear Dynamics & Econometrics*, 6(3).

Ramsey, J. B., & Lampart, C. (1998). Decomposition of economic relationships by timescale using wavelets. *Macroeconomic dynamics*, 2(1), 49-71.

Ramsey, J. B., & Lampart, C. (1998). The decomposition of economic relationships by time scale using wavelets: expenditure and income. *Studies in Nonlinear Dynamics & Econometrics*, 3(1).

Rangvid, J. (2006). Output and expected returns. *Journal of Financial Economics*, 81(3), 595-624.

Ranta, M. (2010). *Wavelet multiresolution analysis of financial time series*. Helsinki: Universitas Wasaensis.

Rapach, D. E., & Wohar, M. E. (2009). Multi-period portfolio choice and the intertemporal hedging demands for stocks and bonds: International evidence. *Journal of International Money and Finance*, 28(3), 427-453.

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

Rapach, D. E., Strauss, J. K., & Zhou, G. (2010). Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *The Review of Financial Studies*, 23(2), 821-862.

Rapach, D. E., Strauss, J. K., & Zhou, G. (2013). International stock return predictability: what is the role of the United States?. *The Journal of Finance*, 68(4), 1633-1662.

Rapach, D., & Zhou, G. (2013). Forecasting stock returns. In *Handbook of economic forecasting* (Vol. 2, pp. 328-383). Elsevier.

Reboredo, J. C., & Rivera-Castro, M. A. (2014). Wavelet-based evidence of the impact of oil prices on stock returns. *International Review of Economics & Finance*, 29, 145-176.

Richardson, M., & Stock, J. H. (1989). Drawing inferences from statistics based on multiyear asset returns. *Journal of Financial Economics*, 25(2), 323-348.

Rozeff, M. S. (1984). Dividend yields are equity risk premiums. *Journal of Portfolio Management* 11, 68-75.

Rua, A. M. P. D. O. (2011). *Essays on Wavelets in Economics*.

Santos, T., & Veronesi, P. (2006). Labor income and predictable stock returns. *Review of Financial Studies*, 19(1), 44.

Schmeling, M. (2009). Investor sentiment and stock returns: Some international evidence. *Journal of empirical finance*, 16(3), 394-408.

Shiller, R. J. (1980). Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review* 71 (3), 421-436.

Thomadakis, A. (2016). Do Combination Forecasts Outperform the Historical Average? *Economic and Statistical Evidence*.

Timmermann, A., & Granger, C. W. (2004). Efficient market hypothesis and forecasting. *International Journal of forecasting*, 20(1), 15-27.

Timmermann, A. (2008). Elusive return predictability. *International Journal of Forecasting*, 24(1), 1-18.

Tiwari, A. K., Dar, A. B., & Bhanja, N. (2013). Oil price and exchange rates: A wavelet based analysis for India. *Economic Modelling*, 31, 414-422.

Valkanov, R. (2003). Long-horizon regressions: theoretical results and applications. *Journal of Financial Economics*, 68(2), 201-232.

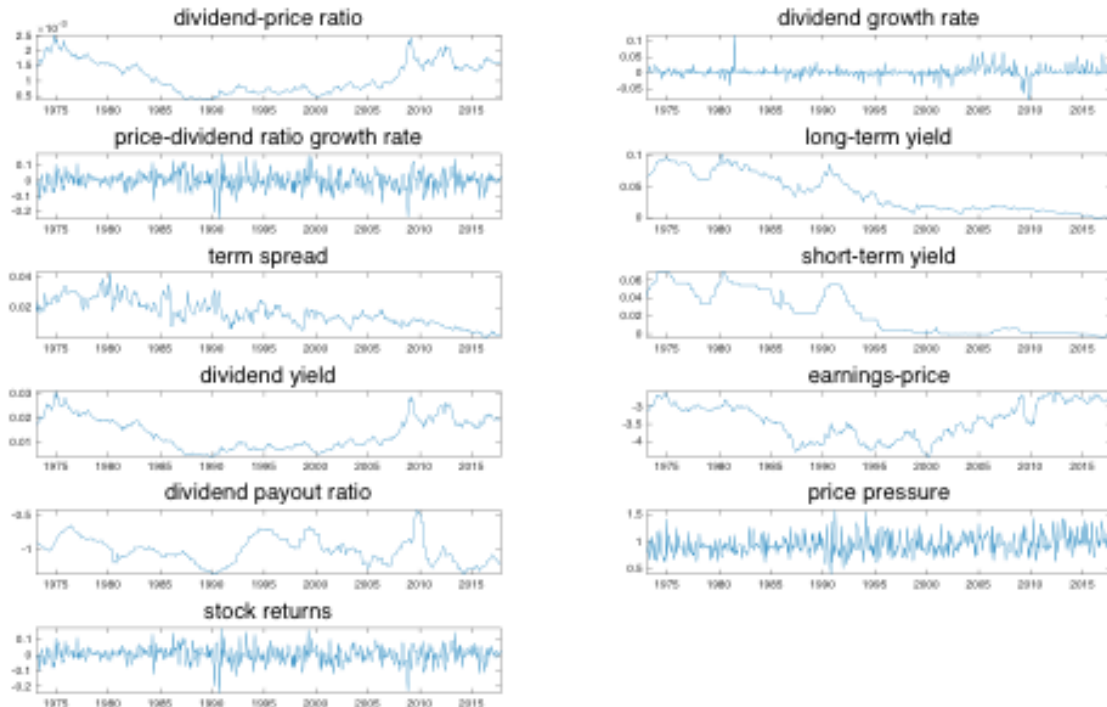
Welch, I., & Goyal, A. (2007). A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies*, 21(4), 1455-1508.

Wong, H., Ip, W. C., Xie, Z., & Lui, X. (2003). Modelling and forecasting by wavelets, and the application to exchange rates. *Journal of Applied Statistics*, 30(5), 537-553.

# Appendix

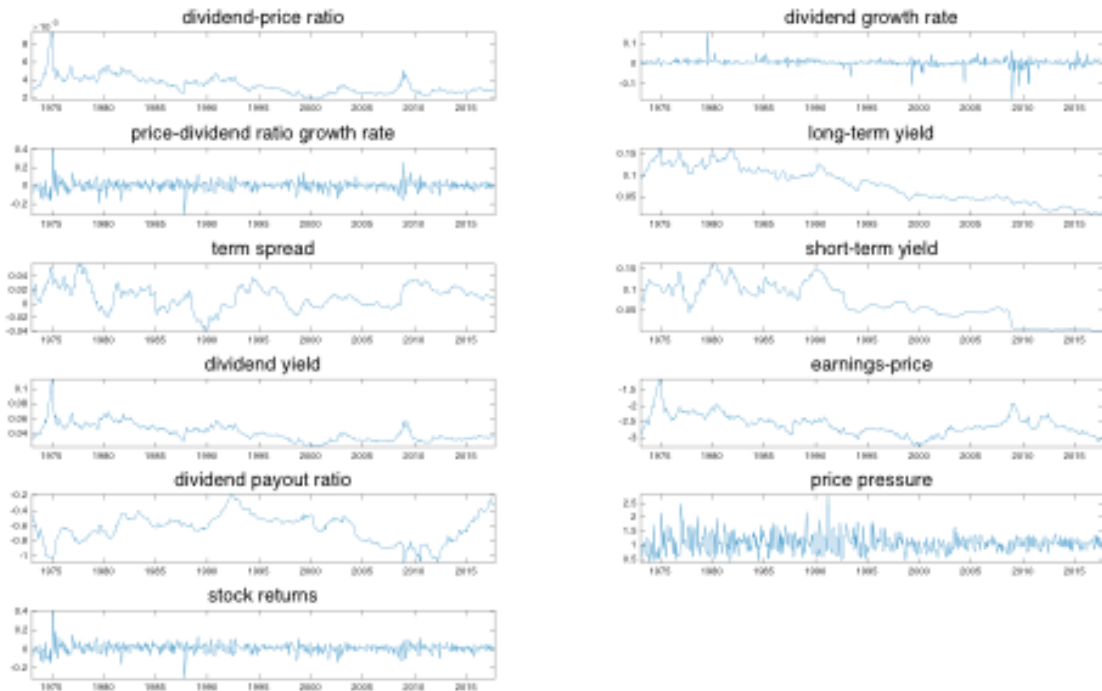
## Univariate statistics and time series from Japan's variables

Return Component:	Mean	Median	Standard Deviation	Minimum	Maximum
Monthly frequency (February 1973 - October 2017)					
gmd	-0.0166	0	5.4346	-24.4453	17.0345
gd	0.3242	0.1794	1.5746	-8.5550	11.9447
dp	0.1153	0.0983	0.0542	0.0358	0.2538
r	0.4236	0.6603	5.2534	-24.3768	17.5104



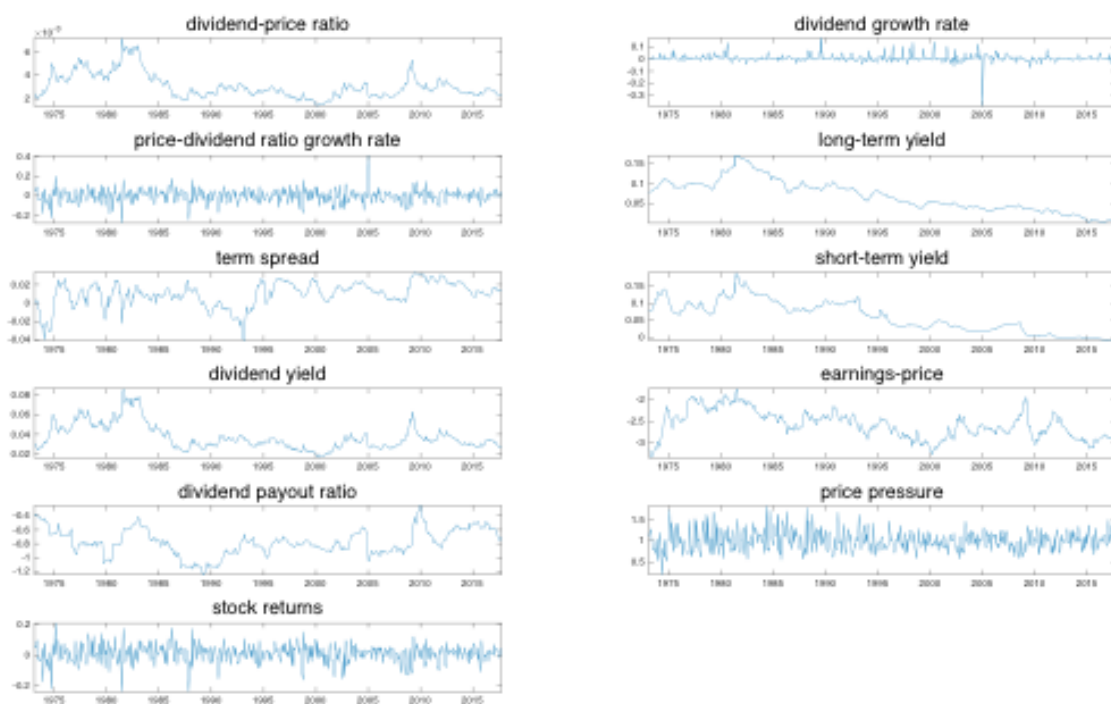
# Univariate statistics from the United Kingdom's main variables

Return Component:	Mean	Median	Standard Deviation	Minimum	Maximum
Monthly frequency (February 1973 - October 2017)					
gmd	-0.0202	0.3082	5.7508	-32.7735	41.3953
gd	0.6092	0.6120	1.9167	-17.8371	15.6450
dp	0.3441	0.3195	0.0982	0.1857	0.9340
r	0.9348	1.3678	5.3300	-31.5138	40.9148



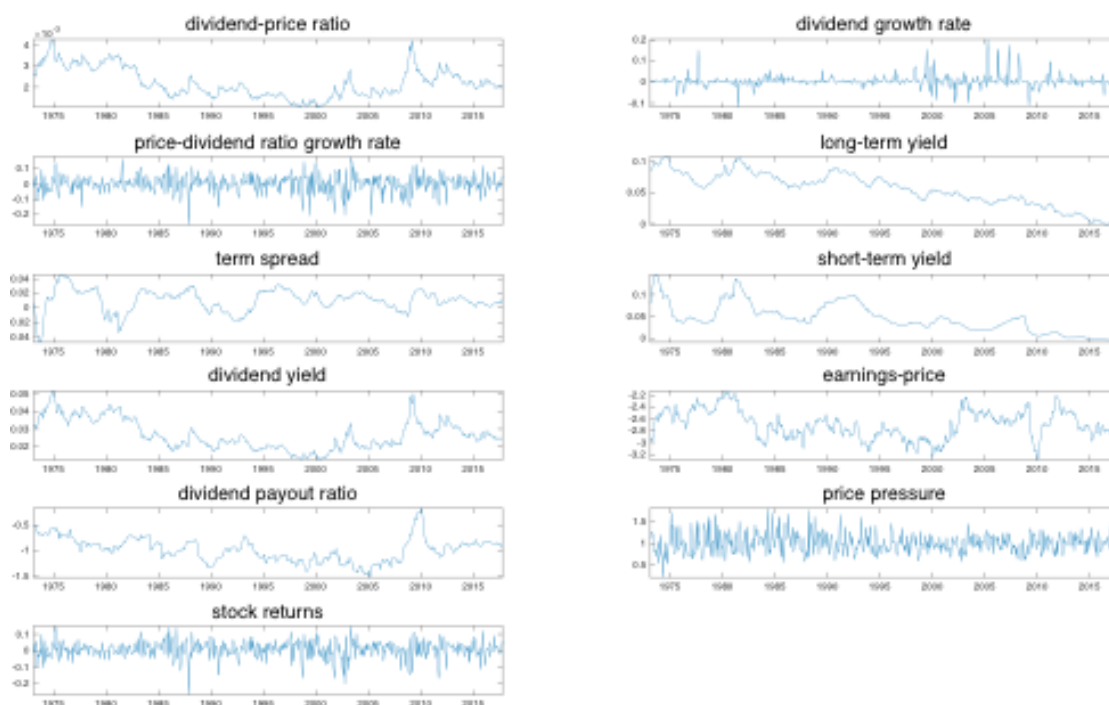
# Univariate statistics and time series from France's variables

Return Component	Mean	Median	Standard Deviation	Minimum	Maximum
Monthly frequency (February 1973 - October 2017)					
gmd	0.0279	0.3515	6.6511	-27.6140	40.9232
gd	0.6055	0.1858	3.1122	-38.8453	17.2044
dp	0.3063	0.2763	0.1068	0.1357	0.7158
r	0.9418	1.3123	5.8752	-23.9682	20.8562



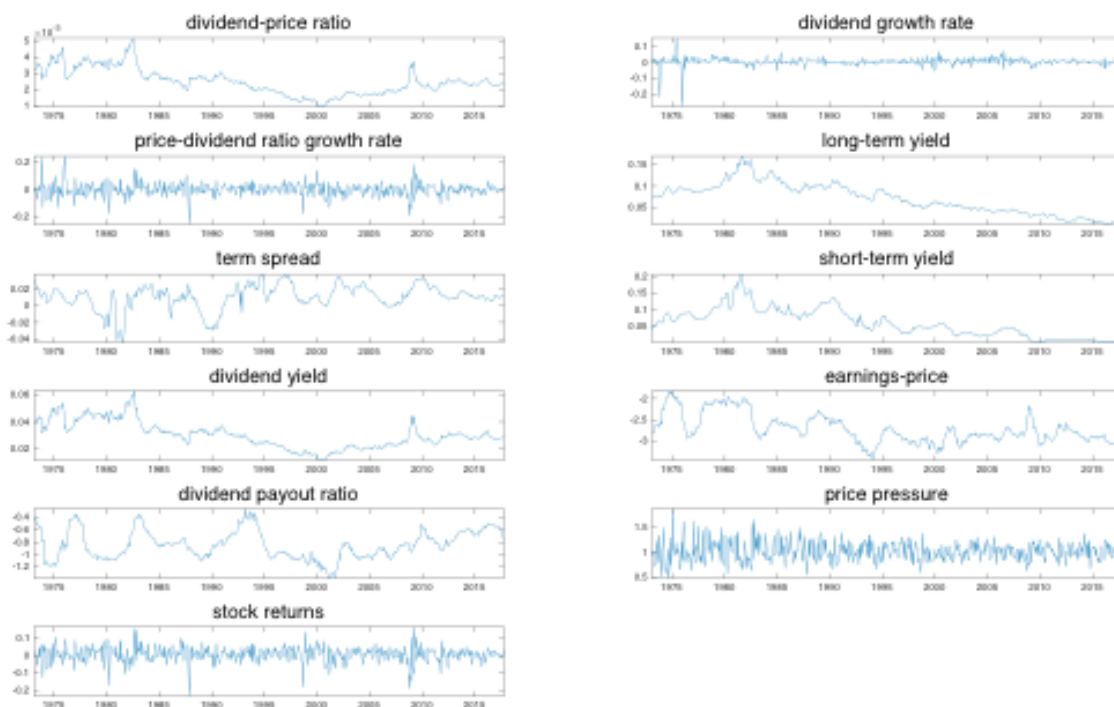
# Univariate statistics and time series from Germany's main variables

Return Component:	Mean	Median	Standard Deviation	Minimum	Maximum
Monthly frequency (February 1973 - October 2017)					
gmd	0.0592	0.3008	5.8332	-26.6524	17.7026
gd	0.3975	0.1562	2.7043	-11.9467	20.1399
dp	0.2215	0.2048	0.0716	0.1024	0.4307
r	0.6794	0.9996	5.1910	-26.6334	14.9973



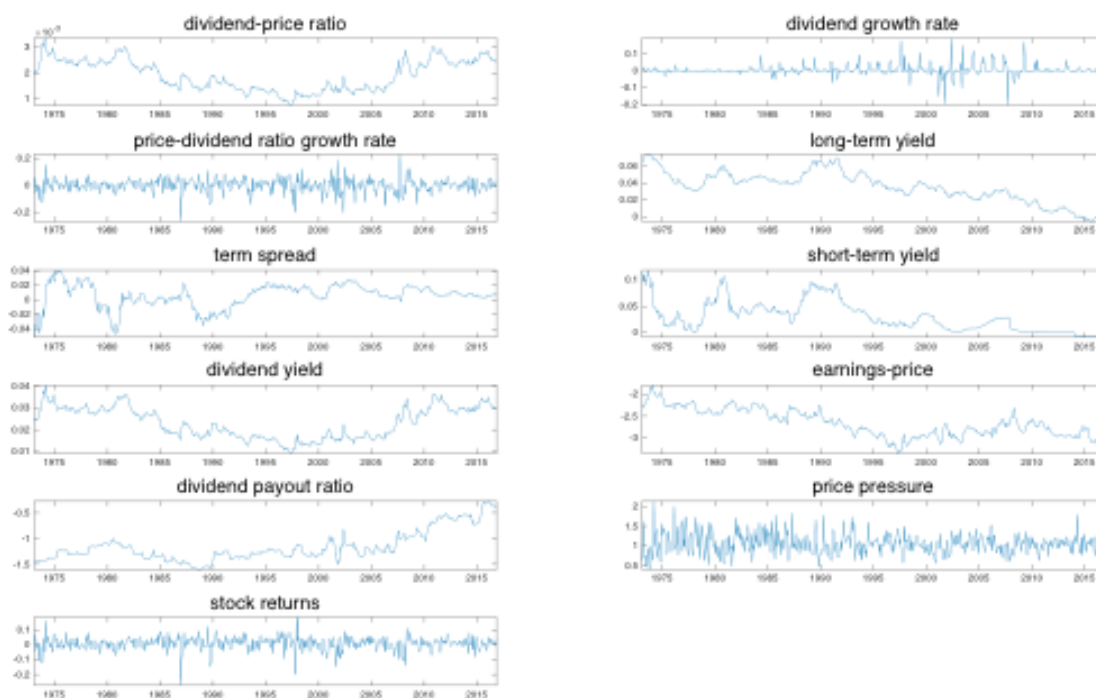
# Univariate statistics and time series from Canada's main variables

Return Component:	Mean	Median	Standard Deviation	Minimum	Maximum
Monthly frequency (February 1973 - October 2017)					
gmd	0.0644	0.2805	4.9124	-25.6878	24.8571
gd	0.4741	0.4379	2.4273	-27.4757	15.2502
dp	0.2509	0.2439	0.0807	0.0983	0.5187
r	0.7909	1.0748	4.4361	-23.2532	16.5789



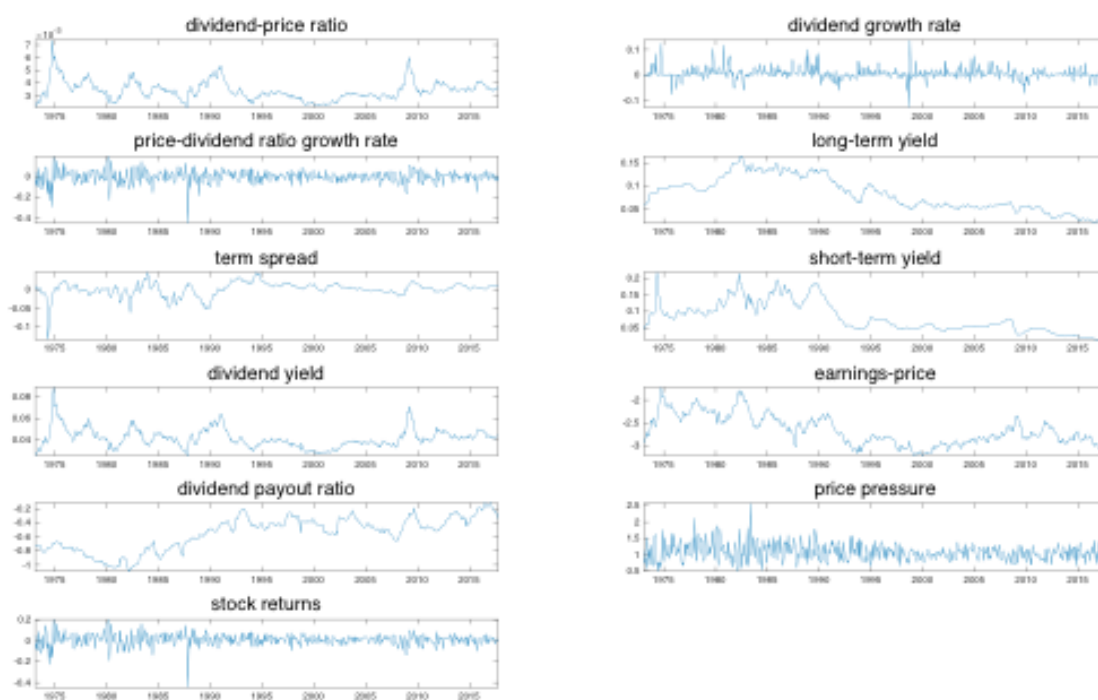
# Univariate statistics and time series from Switzerland's main variables

Return Component:	Mean	Median	Standard Deviation	Minimum	Maximum
Monthly frequency (February 1973 - October 2017)					
gmd	-0.0384	0.5658	5.2492	-26.9781	23.3094
gd	0.5688	0.1176	3.0302	-20.2797	19.5627
dp	0.1921	0.1894	0.0598	0.0741	0.3361
r	0.7234	1.1921	4.4723	-27.3141	18.3894



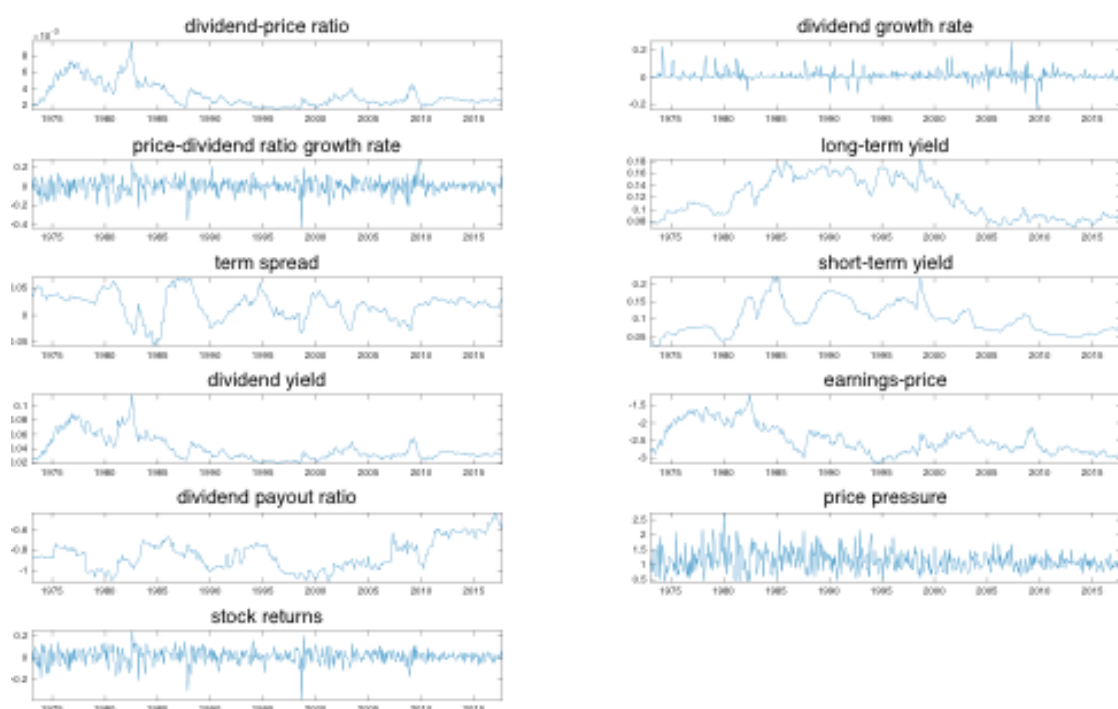
# Univariate statistics and time series from Australia's main variables

Return Component:	Mean	Median	Standard Deviation	Minimum	Maximum
Monthly frequency (February 1973 - October 2017)					
gmd	-0.0926	0.5141	5.9967	-45.0898	19.7656
gd	0.6461	0.3065	2.3046	-12.3704	13.5437
dp	0.3410	0.3303	0.0735	0.2089	0.7373
r	0.8960	1.3580	5.5598	-44.5150	20.3958



# Univariate statistics and time series from South Africa's main variables

Return Component:	Mean	Median	Standard Deviation	Minimum	Maximum
Monthly frequency (February 1973 - October 2017)					
gmd	-0.0150	0.4491	7.5609	-43.1441	27.6253
gd	1.0942	0.3679	3.7540	-24.4452	26.6275
dp	0.3260	0.2692	0.1443	0.1516	0.9703
r	1.4071	1.7292	6.5172	-39.1791	25.0116



# Univariate statistics and time series from the main variables of the United States

Return Component:	Mean	Median	Standard Deviation	Minimum	Maximum
Monthly frequency (February 1973 - October 2017)					
gmd	0.0343	0	4.5521	-25.5594	17.6243
gd	0.5456	0.5422	1.0649	-10.5109	4.2830
dp	0.2436	0.2148	0.1133	0.0791	0.5443
r	0.8248	1.1908	4.5395	-24.1111	17.5797

