1 2 9 0

UNIVERSIDADE D
COIMBRA

Ana Sofia Melo Monteiro

Return Predictability and Portfolio Selection

Tese no âmbito do doutoramento em Economia orientada pelo Professor Doutor Nuno Miguel Barateiro Gonçalves Silva e pelo Professor Doutor Helder Miguel Correia Virtuoso Sebastião e apresentada à Faculdade de Economia da Universidade de Coimbra.

Fevereiro de 2022

FACULDADE DE ECONOMIA
UNIVERSIDADE DE COIMBRA

Ana Sofia Melo Monteiro

Return Predictability and Portfolio Selection

Tese no âmbito do doutoramento em Economia, apresentada à Faculdade de Economia da Universidade de Coimbra para obtenção do grau de Doutor

Orientadores:
Professor Doutor Nuno Miguel Barateiro Gonçalves Silva
e
Professor Doutor Helder Miguel Correia Virtuoso Sebastião

Fevereiro de 2022

Acknowledgements

First of all, I would like to express my sincere gratitude to my thesis supervisors Nuno Silva and Helder Sebastião for providing guidance and many insightful comments and suggestions throughout my Ph.D. studies. I also wish to thank them for always taking the time to discuss the research topics and for sharing their knowledge with me.

Finally, I would like to thank my parents, Áurea and Manuel, for giving me care and support during all stages of my academic life, and my boyfriend, Filipe, for his never-ending support and many inspiring conversations.

Funding

Financial support was provided by the SYLFF Program of Tokyo Foundation in November 2018 and by national funds through FCT – Fundação para a Ciência e a Tecnologia, I.P., Project UIDB/05037/2020.

Resumo

Nesta tese estudamos a relação entre a previsibilidade dos retornos e a seleção ótima de portfólio através de uma abordagem empírica. Para tal, em cada capítulo utilizamos diferentes métodos para prever os retornos de ativos e analisamos as implicações destas previsões na otimização de portefólio de investidores *CRRA*.

Em primeiro lugar, analisamos a precisão das previsões de diferentes modelos *Vector Autoregressive (VAR)* e modelos Bayesianos *Time-Varying Parameters Vector Autoregressive* (*TVP-VAR)* que integram a seleção/média dinâmica de modelos. Em particular, estes modelos são utilizados para prever conjuntamente os retornos mensais de índices de ações, de obrigações e dos *Real Estate Investment Trusts (REITs)* dos EUA, para o período de janeiro de 1976 a dezembro de 2017. Posteriormente, as previsões obtidas são utilizadas na seleção de portefólio. Os resultados obtidos sugerem que os métodos Bayesianos proporcionam ganhos significativos em termos de previsibilidade estatística, medida pelo pseudo-R^2 fora da amostra, e em termos de desempenho económico, que quantificamos pelo Equivalente Certo e pelos rácios de Sharpe e de Sortino. Por último, comparamos o desempenho dos modelos antes e depois da crise do *subprime* e concluímos que as abordagens Bayesianas são adequadas para acomodar a instabilidade do mercado.

Em segundo lugar, apresentamos um modelo assimétrico *Dynamic Conditional Correlation (DCC)* multivariado baseado em Machine-Learning para prever dinamicamente retornos e covariâncias que são posteriormente utilizados na otimização de portefólios. Este modelo é aplicado a retornos diários de 77 índices de ações e obrigações para o período de agosto de 2001 a setembro de 2020. Através desta aplicação concluímos que os modelos propostos levam a elevados ganhos económicos. Em particular, o modelo proposto assimétrico DCC multivariado que inclui Florestas Aleatórias aumenta consideravelmente o desempenho do portefólio e o Equivalente Certo de um investidor *CRRA*. Verificamos ainda que, investidores da América do Sul, da Europa, do Médio Oriente, da Ásia e da Oceânia beneficiariam amplamente de diversificar internacionalmente os seus portefólios, no período de 2012-2020.

Por último, analisamos as relações *lead-lag* entre onze indústrias de países desenvolvidos, no período de janeiro de 1973 a maio de 2021. Em particular, identificamos o papel de liderança internacional desempenhado pelos EUA. Nomeadamente, os retornos semanais das indústrias dos EUA e, em especial, as indústrias de Materiais Básicos e Energia, causam significativamente à *Granger* os retornos da maioria das indústrias de outros países. Este resultado sugere que as indústrias não americanas reagem com atraso a novas informações. Essa reação tardia é ainda mais percetível durante períodos de recessão nos EUA, quando as correlações entre países são mais elevadas. Assim, os retornos desfasados das indústrias americanas têm uma maior capacidade em prever os retornos das indústrias de outros países desenvolvidos, quando os EUA estão numa recessão económica. Identificamos ainda uma relação assimétrica semelhante entre a volatilidade das indústrias dos EUA e a volatilidade das indústrias de outros países. Por fim, a análise de causalidade na distribuição de retornos e volatilidade demonstra-nos, uma vez mais, que a causalidade se verifica principalmente dos EUA para outros países, especialmente na presença de choques negativos extremos.

Palavras-chave: Retornos; Previsibilidade; VAR; Modelos Bayesianos; Machine-Learning; DCC; Seleção de portefólios; Diversificação Internacional; Indústria; Lead-Lag.

Classificação JEL: C45, C61, G11, G12, G17.

Abstract

This thesis provides empirical evidence on the relationship between return predictability and optimal portfolio selection. Each chapter addresses this topic using different approaches to forecast asset returns and determine their implications for portfolio optimization by CRRA investors.

Firstly, we analyse the predictive accuracy of different multivariate VAR models and TVP-VAR Bayesian models with dynamic model selection/averaging to jointly forecast monthly returns of US stocks, bonds, and REITs indexes from January 1976 to December 2017. The forecasts are obtained by those models and then used as inputs for portfolio selection. We conclude that Bayesian-based approaches provide the most significant gains in terms of statistical predictability, as measured by out-of-sample pseudo-R^2, and in terms of economic performance, which we quantify through certainty equivalent returns, Sharpe ratios, and Sortino ratios. The comparison between the performance of the models before and after the subprime crisis supports the claim that Bayesian approaches can accommodate market instability.

Secondly, we use a Multivariate Machine Learning - Asymmetric Dynamic Conditional Correlation (DCC) model to dynamically forecast returns and covariances, which are then used in the portfolio optimization problem. We apply our model to daily returns of 77 national stock and bond indexes for the period from August 2001 to September 2020. We find that our methods lead to large economic gains. Most notably, we show that relative to the proposed Random Forest - Asymmetric DCC model considerably increases the portfolio performance and the certainty equivalent of a CRRA investor. We also show that international diversification is amply beneficial for investors from South America, Europe, the Middle East, Asia, and Oceania, between 2012-2020.

Thirdly, we analyse the lead-lag relationships within and across eleven industries of developed countries in the period from January 1973 to May 2021. We identify the international leading role played by the US, namely by showing that weekly returns of US industries, especially US Basic Materials and Energy industries, significantly Granger cause the returns of most of the industries of other countries, suggesting that non-US industries react with some delay to new information. This delayed reaction is even more noticeable during periods of recession in the US, when cross-country correlations are higher. This implies that the ability of lagged returns of US industries to predict returns of industries from other developed countries is even more pronounced when the US is experiencing an economic recession. A similar asymmetric relationship is also identified between the volatility of US-industries and the volatility of industries of other countries. The analysis of causality in the distribution of returns and volatility shows, once more, that causality runs mainly from the US to other countries, especially in the presence of extreme negative shocks.

Keywords: Return; Predictability; VAR; Bayesian models; Machine Learning; DCC; Portfolio Selection; International Diversification; Industry; Lead-Lag.

JEL codes: C45, C61, G11, G12, G17.

Acronyms

ADCC – Asymmetric Dynamic Conditional Correlation

AIC – Akaike Information Criterion

BIC – Bayesian Information Criterion

BM – Basic Materials

CA – Canada

CAViaR – Conditional Autoregressive Value-at-Risk

CCC – Constant Conditional Correlation

CD – Consumer Discretionary

CER – Certainty Equivalent Return

CH – China

CRRA – Constant Relative Risk Averse

CS – Consumer Staples

DCC – Dynamic Conditional Correlation

DM – Developed Markets

DMA – Dynamic Model Averaging

DMS – Dynamic Model Selection

ECB – European Central Bank

EM – Emerging Markets

EN – Energy

ES – Expected Shortfall

EU – European Union

EU RF-ADCC – European Union Random Forests - Asymmetric Dynamic Conditional Correlation

EU RF-DCC – European Union Random Forests - Dynamic Conditional Correlation

EU NN-ADCC – European Union Neural Networks - Asymmetric Dynamic Conditional Correlation

EU NN-DCC – European Union Neural Networks - Dynamic Conditional Correlation

EWMA – Exponentially Weighted Moving Average

FI – Financials

FR – France

GARCH – Generalized Autoregressive Conditional Heteroscedastic

GE – Germany

HC – Health Care

ICB – Industry Classification Benchmark

IN – Industrials

IS – In Sample

JP – Japan

MCMC – Markov Chain Monte Carlo

MSE – Mean-squared Error

MVP – Minimum-Variance Portfolio

NBER – National Bureau of Economic Research

NN – Neural Networks

NN-ADCC – Neural Networks - Asymmetric Dynamic Conditional Correlation

NN-DCC – Neural Networks - Dynamic Conditional Correlation

OOS – Out-of-sample

RE – Real Estate

REIT – Real Estate Investment Trusts

RF – Random Forests

RF-ADCC – Random Forests - Asymmetric Dynamic Conditional Correlation

RF-DCC – Random Forests - Dynamic Conditional Correlation

SOR – Sortino Ratio

SR – Sharpe ratio

STD – Standard Deviation

TEC – Technology

TEL – Telecommunications

TVP-VAR – Time-Varying Parameter – Vector Autoregressive

UK – United Kingdom

US – United States

UT – Utilities

VaR – Value-at-Risk

VAR – Vector Autoregressive

WMSFE – Weighted Mean Squared Forecast Error

List of Figures

List of Tables

Contents

Chapter 1 - Introduction

Return predictability is one of the most debated issues in modern financial literature. The ability to predict returns has been a subject of interest for both academics and finance practitioners due to its implications for strategic portfolio management.

Until the 1980s, it was believed that returns of bonds and stocks were utterly unpredictable, and this unpredictability was perceived as the essence of the Efficient Market Hypothesis. However, since the mid-1980s, academics have increasingly demonstrated that stocks and bonds were, to some extent, predictable from lagged valuation ratios, such as dividend-price ratios, earning-price ratios, and macroeconomic variables, such as nominal interest rates, interest rate spreads, labour income, stock market volatility, aggregate output, industry returns, output gap, expected business conditions, accruals, inflation rate, oil-related variables, technical indicators, manager sentiment, among others. Notably, return predictability maybe not be the result of irrationality and market inefficiency (mispricing, bubbles, noise traders, etc.) but, instead, may result from changes in aggregate risk-aversion and risk premia (Rapach et al. 2013). If that is the case, predictability is theoretically consistent with the Efficient Markets Hypothesis.

Historically, return predictability has been attributed to several sources. For instance, Rapach et al. (2013) argued that stock return predictability was closely related to the business cycles. According to the authors, investors become more risk-averse during recessions, when consumption and income levels are lower. Hence, during recessions, they will demand a higher expected return. Therefore, variables that measure and/or predict the state of the economy may help predict returns.

Another possible explanation for return predictability is that different investors use different forecasting models, resulting in different assessments of the level of financial uncertainty. When market instability rises, uncertainty becomes very high, and the level of disagreement between investors increases. According to Cujean and Hasler (2017), there is a positive relationship between investor disagreement and future returns, creating time-series momentum (which grows during bad times).

Return predictability has a great impact on portfolio selection (Campbell and Viceira, 2003; Dangl and Weissensteiner, 2020). An investor adapts her portfolio strategy to incorporate the predictions made by her forecasting models. If these models are effective, portfolios based on forecasts could perform better than naïve strategies solely aiming to achieve high diversification levels. If, in fact, returns are predictable, rational investors will incorporate that knowledge into their decisions related to strategic asset allocation, active portfolio management, and market timing.

Additionally, when stock returns are predictable, the optimal allocation becomes horizon-dependent and may originate changes in asset demands. For instance, an investor with a 10-year horizon tends to allocate more capital into stocks than an investor with a one-year horizon. Hence, stocks appear less risky to long-horizon than to short-term investors (Barberis, 2000; Xia, 2001; Ferson, 2003). A long-term investor is more interested in the consequences of predictability on the covariance structure of returns during the holding period. In other words, predictability makes the trade-off between risk and return depending on the investment horizon. Therefore, one part of the unconditional variance of asset returns is predictable, and it is no longer considered investment risk. This effect is known as time diversification. However, part of it is sometimes erased by estimation errors in the term structure of the trade-off between risk and return (Dangl et al., 2020).

According to the literature, the possibility of predicting asset returns creates an opportunity to construct dynamic trading strategies which offer superior expected return-risk trade-offs relative to standard portfolios. For instance, Xu et al. (2004) showed that a trading strategy, with monthly portfolio rebalancing, based on the observed small return predictability, from 1952 to 1998, would double the return of benchmark portfolios.

Despite the enormous advantages of predicting asset returns, predictability itself does not guarantee that an investor will obtain abnormal returns (Pesaran and Timmermann, 1995). This is due to the presence of transaction costs that may dilute the profit of dynamic trading strategies based on forecasting models, rendering these investment strategies less attractive than simple buy-and-hold strategies.

A question that often arises in predictability literature is: "What level of return predictability should we expect?". Several forecasting variables used to predict stock and bond returns have significant out-of-sample accuracy (Campbell and Thompson 2008; Welch and Goyal, 2008; Neely et al., 2014). However, the predictive power, typically measured via the in-sample R^2 or out-of-sample pseudo-R^2 statistics, is quite small. For instance, Fama (1990) reported monthly R^2 statistics of 1% for a model predicting US stocks from the dividend-price ratios. Zhou (2010) registered a monthly R^2 smaller than 1% for individual assets regressions based on ten popular economic variables. Campbell and Thompson (2008) indicated an even smaller pseudo-R^2 of less than 1%. Rapach et al. (2013) showed small and negative monthly pseudo-R^2. Zhu et al. (2013) reported pseudo-R^2 ranging from -4% to 4% for monthly stock returns using several typical predictors. Nevertheless, better results were achieved by Ludvigson and NG (2009) for forecasting bonds based on macro factors and Golez and Koudijs (2018), which reported annualized R^2 of around 10% for stocks in the US, the UK, and the Netherlands.

From the standpoint of return predictive regression models, we should only expect a limited degree of predictability in stocks and bonds returns. A monthly R^2 statistic of around 1% is a good indicator of return predictability or the presence of market inefficiencies in the perspective of traditional asset pricing models. However, many authors, such as Campbell and Thompson (2008) and Rapach et al. (2013) are suspicious about high predictability levels. According to these authors, predictive models that report high return predictability are most likely incorrectly designed, have several inaccuracies, or suffer from data mining problems, but nevertheless, there is the possibility that these positive results come from significant market inefficiencies or anomalies. Although one should expect a small degree of return predictability, it is crucial to note that even a minimal level of predictability may contain valuable information for a risk-averse investor who is planning her portfolio strategy, as it could translate into significant economic gains (Kandel et al., 1996; Xu, 2004; Campbell and Thompson, 2008; Rapach et al., 2013).

Despite the numerous advantages of return predictability, the literature has reported biases and limitations in the predictive regression models. The first one is the evidence of Stambaugh's bias (Stambaugh, 1999) that occurs when the predictor and return innovations are correlated, and the predictor is highly persistent in in-sample tests of return predictability. When applying a traditional t-statistic to test the null hypothesis of no predictability, this bias can generate significant size distortions. Several studies have developed approaches that improve the inference of predictive regressions with persistent predictors (e.g., Pástor and Stambaugh, 2009).

A second problem is related to the occasional unreliability of return predictions. This problem arises when academics implicitly use a strong prior on predictability for the econometric methods (Rapach et al., 2013). Predictive regression models have been associated with parameter instability and model uncertainty. Parameter instability means that the value of the coefficients can change over time. Model uncertainty occurs when the researcher does not know which model is correct (e.g., variables to include, functional specification, etc). Further, many regression models face overparameterization and cannot handle the uncertainty and instability that characterize financial data. These problems challenge the real predictability of the models and highlight the importance of providing flexible yet suitable methods that accommodate and improve out-of-sample forecasting performances.

A third concern that has for long divided academics and practitioners is the question of whether return predictability should be tested in-sample or out-of-sample, making an overall statement on return predictability a complex issue. Until the 2000s, most studies analysed the predictability of returns in-sample. However, in recent years, the analysis has shifted to out-of-sample evidence. The in-sample analysis is typically conducted using standard t- or F-tests, while out-of-sample tests focus mainly on forecasting accuracy. In many cases, in-sample tests are viewed as providing more power in detecting return predictability. This is usually justified by the fact that in-sample analysis uses all available data, delivers efficient parameter estimates, and translates into more precise estimates of the expected equity risk premium (Neely et al., 2014). However, in-sample tests are most likely subject to Stambaugh bias and usually tend to over reject the null hypothesis of "no predictability". This occurs because there is an artificial increase in the size of the predictive tests, which consequently leads to spurious rejections of the null hypothesis.

Also, typically, in-sample fits are more prone to data mining than out-of-sample tests predictability. Yet out-of-sample tests are also susceptible to data mining problems. Inoue (2005) argued that given appropriated critical values, both in-sample and out-of-sample tests are likely to face this problem. Obviously, the researcher can freely choose alternative predictor spaces until she finds a significant one before publication.

A common criticism of the out-of-sample analysis regards the estimation period. The choice of the periods over which a model is estimated and subsequently evaluated has important implications on the results, always involving the loss of some information (see, for instance, Inoue et al., 2005; Hansen et al., 2012; or Kolev et al., 2017). However, this choice should not be arbitrary. According to Welch and Goyal (2008), it is critical to have enough starting data to obtain solid estimates and a long enough evaluation period to be representative.

Out-of-sample evidence should not be viewed as a substitute for in-sample analysis but rather as an essential complement in determining the quality of the underlying model. Several authors compare the power of out-of-sample tests with in-sample tests; however, this is not correct. Under a well-specified model, an in-sample estimate is more efficient. A researcher who has confidence in her underlying model should rely on in-sample evidence. However, no one really knows a priori what a well-specified model is, hence there is model uncertainty. This is where the out-of-sample tests enter. Out-of-sample frameworks provide essential and valuable statistical diagnostics tools, helping the researcher to detect whether a model is well-specified and even stable through time. Both in- and out-of-sample evidence should be considered, and research should explore whether, conditional on observed in-sample significance, out-of-sample diagnostics are reasonably powerful (Welch and Goyal, 2008).

From the above discussion, we may undoubtedly state that the literature on return predictability has evolved over the years and has given itself room to accommodate several modifications that better explain the data. The increasing availability of data and computation power has proven effective in discovering more flexible procedures and methodologies. In this thesis, we attempt to contribute to this strand of literature. Particularly, this work consists of a compilation of three studies from which we attempt to provide empirical and conceptual evidence concerning the relationship between return predictability and portfolio selection. Each chapter addresses these topics using different approaches to assess the predictability of returns and their implication for portfolio allocation by Constant Relative Risk Averse (CRRA) investors.

First, in Chapter 2, "*Multi-asset return predictability using VARs*," we conduct a comprehensive analysis of return predictability on stocks, bonds, and real-estate investment trusts (REITs). Motivated by the limitations of existing literature in assessing which single model should be employed by investors, we compare the performance of different VAR models and time-varying Bayesian models with dynamic model averaging/ selection. Additionally, we highlight the importance for investors to have at their disposal adequate forecasting methods to better establish the empirical reliability of asset return predictability and to plan their investment choices.

In this sense, we make three main contributions. The first one lies in analysing different model specifications and features out-of-sample, such as time-varying parameters, model/forecast combinations, and dynamic Bayesian model selection/averaging. The critical element of our analysis is to consolidate several useful features employing simple but flexible computational methods, suitable for jointly forecasting multiple risky assets. Results point out that the Bayesian models are the best ones in terms of statistical and economic performances.

The second contribution of this chapter is to empirically analyse the interaction of these ensemble features in a multiple asset portfolio. More specifically, we use the forecasts obtained by the models as inputs in a dynamic optimizing investment portfolio. We focus on three distinct asset classes and assess the additional economic value of out-of-sample forecasts in a multi-asset investment strategy of an investor endowed with a power utility function.

Due to the frequent occurrence of shocks in financial markets, investors face high levels of uncertainty, therefore the need to have highly adaptive methods to construct optimal portfolios. Hence the third contribution of this chapter is related to the examination of whether portfolio performances based on a given model presented substantial differences before and after January 2008, the beginning of the subprime crisis.

In its most basic formulation, Markowitz portfolio selection requires estimates of the expected return vector and the covariance matrix of all assets in the investment universe. On the one hand, the expected return vector estimates resulting from traditional methods (such as the historical averaging or momentum) are prone to severe measurement errors. This is most likely caused by the high nonlinearity of the data and the inability of these methods to capture complex data interactions or structural breakdowns. On the other hand, the conditional covariance matrix estimation typically faces the problem of the curse of dimensionality. Therefore, in Chapter 3, "*International Portfolio Selection with Machine-Learning and a Multivariate Asymmetric DCC model*" we introduce a Multivariate Machine Learning Asymmetric DCC model that is well-suited to deal with the problems mentioned earlier.

Mainly, our approach builds on Random Forests and Artificial Neural Networks to forecast the returns and builds on an Asymmetric DCC model to estimate the covariance matrix. The key element of this approach is the ability to integrate several data features into a flexible method that is suitable for large predictive datasets and highly correlated predictors by reducing degrees of freedom and condensing redundant variations.

When applied to a diversified international portfolio, we find that our methods lead to significant economic gains. Notably, we show that the model considerably increases the certainty equivalent of CRRA investors. Finally, we examine the potential in diversification under different portfolio performance measures. Results reveal that international diversification is amply beneficial for risk-averse investors from five other regions, from 2012 to 2020.

The way that information is communicated between markets is essential for an investor planning her investment strategy. Nowadays, investors can invest in many classes of assets from various countries and regions. The decision on what assets to invest in benefits from the examination of which country plays the key role in international information transmission. The US is expected to have a prominent role since it is the world's largest open economy, harbouring the largest corporations in the world. Hence, events in the US are likely to impact other economies.

In Chapter 4, "*Industry Lead-lag relationships between the US and other developed countries*", we look at the information transmission role played by the US. More specifically, we analyse within and cross-industry interdependences and lead-lag relationships in a global context. We confirm the leading role played by the US, namely that weekly lagged returns of US industries Granger cause most of the industry returns of the non-US countries. This suggests that non-US industries react slowly to new information from US industries. This delayed reaction is even more pronounced during recession periods (when cross-country correlations are more substantial). This also suggests that the predictive power of US industry returns is much greater when the US is experiencing a recession. Lastly, we analyse the Granger causality in distribution for both industry returns and volatilities. Our results reveal that other countries do not timely incorporate shocks affecting the US industries.

This study differs from previous literature as we explicitly analyse the asymmetries in the leading role of the US industries in an international context. Previous research has focused primarily on the international stock index stock or firm-level returns, ignoring intra-industry information flows (see, for instance, Rapach et al., 2013; Bollerslev et al., 2013). And, to the best of our knowledge, this is one of the first studies to provide international empirical evidence supporting asymmetric reactions to news arriving from the US industries during expansionary and recessionary periods.

Chapter 5 provides a summary of the results reached in this thesis.

Chapter 2 – Multi-Asset Return Predictability Using VARs

2.1. Introduction

Return predictability and portfolio selection are two of the most relevant topics in financial markets. From the point of view of practitioners and investors, the ability to forecast returns has important implications on optimal long-term portfolio asset allocation, as highlighted by, e.g., Campbell and Viceira (2002) and Campbell et al. (2003). Furthermore, the possibility of return predictability motivates the use of more robust estimation techniques to detect all the pertinent information available in the data.

Recent research has analysed stock and bond return predictability considering several macroeconomic variables, such as inflation rates (Ludvigson and Ng, 2009), interest rates (Golez and Koudijs, 2018, and Bandi et al., 2019), valuation ratios, and the well-known dividend-price ratio (Cochrane, 2007), Typically, these studies have relied on specific sets of predictors to forecast multiple-asset returns (Gao and Nardari, 2018). The definition of the predictor space is of utmost importance in this type of studies, since the use of inadequate predictors reduces the predictive ability of forecasting regressions and, consequently, the performance of asset allocation strategies devised upon those models. One of the aims of this chapter is to analyse the adequacy of different sets of predictors for Stock, Bonds and REITs, using a pre-selection method.

Traditionally, research on return predictability has mainly been building up in-sample empirical evidence. However, more recent literature has highlighted the power and robustness of out-of-sample analyses (Welch and Goyal, 2008; Fisher et al., 2020). The debate between the advantages and disadvantages of in-sample versus out-of-sample analysis has focused on different aspects, such as data snooping, data mining, spurious regressions, and instability in return predictability (Wu et al., 2013; Dichtl et al., 2020). For instance, it is known that financial markets are subjected to permanent shocks and experience volatility clustering. Hence well-defined investment strategies should be designed upon flexible forecasting models that accommodate these sources of uncertainty.

There is an extensive body of empirical literature comparing the predictive power of different models. Typically, asset returns are forecasted by past values of predictive variables within a Vector Autoregressive (VAR) framework (Guidolin and Hyde, 2012). Despite its popularity, the flexibility of VAR models entails the danger of over-parameterization, leading to unreliable predictions. Nowadays, the toolbox of applied econometrics includes numerous efficient modelling tools to prevent the proliferation of parameters and reduce parameter and model uncertainty. These new modelling tools, such as time-varying parameters, forecast combinations, model averaging, and model selection techniques, have been fuelled by the noticeable advances in computational power. An example of this is the Bayesian approach and, more specifically, Bayesian model selection and model averaging. For instance, Koop and Korobilis (2013) applied these methods to estimate and forecast large VARs and time-varying VARs. Different from most methods presented in the literature of large multivariate models, which require Markov Chain Monte Carlo (MCMC) simulations, the authors estimate forgetting (discount) factors and allow for model switching between different restricted VARs to mitigate the probability of excessive parameterization.

Although the existing literature proposes numerous alternatives to VAR models, this framework still retains its interest when dealing with forecastability in a multivariate setup. According to Fisher et al. (2020), not a single feature alone, but an ensemble of them, is required to handle uncertainty and instability of financial markets, as well as making good predictions. Hence, our study analyses different model specifications and features, such as time-varying parameters, model/forecast combinations, and dynamic model selection/averaging, to jointly model multiple risky assets out-of-sample. The key element of our analysis is to consolidate several useful features employing a flexible yet straightforward computational method suitable to manage multiple risky assets. Therefore, this is our second contribution to the literature.

Despite some literature supporting the ensemble of features when modelling a single risky asset return, surprisingly, few studies have investigated how these features connect when jointly forecasting multiple risky asset returns. Nevertheless, most investors hold several risky assets in their portfolios, making this an empirically relevant issue. The third contribution of this study is to empirically analyse the interaction of these features in a configuration of multiple assets. We focus on three classes of assets: stocks, bonds, and Real Estate Investment Trusts (REITs). We also conduct an out-of-sample analysis of the economic value that the forecasts add to an investment strategy with multiple assets. More specifically, we apply a dynamic optimization technique to predict multiple assets returns and use these predictions as inputs for optimal portfolio allocation.

Investors face high levels of uncertainty stemming from the frequent shocks in financial markets and the need to have highly adaptive methods to construct optimal portfolios. Understanding how market conditions affect the variability of portfolio returns is extremely relevant for optimal portfolio choices. In this way, our fourth contribution is to examine whether portfolio returns present substantial performance differences before and after January 2008 (i.e., the beginning of the subprime crisis). We also analyse whether our models are appropriate under periods of market instability.

The main goal of this study is to obtain the single best model to forecast both the expected return vector and the covariance matrix of stocks, bonds, and REITs and employ it in portfolio selection. In sum, we proceed as follows: first, the predictive variables are pre-selected based on the correlations between asset returns and predictors. We begin by considering 19 predictors for stock returns, 125 predictors for bond returns, and 14 predictors for REITs that potentially contain relevant information to forecast these asset returns. Second, conventional Vector autoregressive (VAR) models are built, including the previously selected explanatory variables. Third, we combine the asset returns forecasts obtained through these VARs. These combinations are based on simple schemes such as the Mean, Median, Trimmed Mean, and the Weighted Mean Squared Forecast Error (WMSFE). Fourth, we estimate Bayesian models with Dynamic Model Selection (DMS) and Dynamic Model Averaging (DMA) to obtain another set of forecasts. The estimation process uses the well-known Kalman-filter method along with forgetting factors. These methods are robust since they choose the best model available or average different models over time.

Lastly, we evaluate the performance of the different ensemble features when jointly forecasting monthly excess returns on stocks, bonds, and REITs from January 1976 to December 2017. Further, we analyse the stability of these models before and after the beginning of the subprime crisis of 2008.

Among the features considered, we conclude that Bayesian-based models bring the largest gains in terms of statistical predictability, as measured by the R^2 out-of-sample and the MSFE-adjusted statistic of Clark and West (2007), and in terms of economic performance, which we quantify using Certainty Equivalent Returns (CERs), Sortino ratios, and Sharpe ratios. We also show that these models are stable under uncertainty periods. In sum, we emphasize that investors and practitioners need to have at their disposal adequate forecasting methods to better establish the empirical reliability of equity premium predictability and to formulate their investment choices.

The remaining of this chapter is structured into six sections. Section 2.2 presents a brief literature review. Section 2.3 describes the data and provides several descriptive statistics. Section 2.4 outlines the basic theoretical concepts and shows the specifications of the models. Section 2.5 shows the results obtained from different models. Section 2.6 conducts an analysis of the effects of the 2008 subprime crises on the return's predictability and portfolio performance. Finally, Section 2.7 highlights the main conclusions.

2.2. Literature Review

Throughout the years, several studies have analysed the predictability of financial returns using different predictor spaces (some recent examples are Kothari and Shanken, 1997, Campbell and Shiller, 1998, Pontiff and Schall, 1998, Baker and Wurgler, 2000, Goetzmann et al., 2001, Lettau and Ludvigson, 2001, Guo, 2006, van Binsbergen and Koijen, 2010, Ferreira and Santa-Clara, 2011, Rapach et al., 2013, Neely et al., 2014, Maio and Santa-Clara, 2015, Golez and Koudijs, 2018, Jagannathan and Liu, 2019, le Bris et al., 2019, Bandi et al., 2019, and Piatti and Trojani, 2019).

le Bris et al. (2019) studied the Bazacle company of Toulouse, the earliest documented shareholding corporation, using share prices and net dividends for a period of almost six centuries, from 1372 to 1946. According to the authors, a significant fraction of price variations resulted from changes in expectations regarding future dividends. Goetzmann et al. (2001) analysed the US aggregate stock market and found little evidence of stock return predictability during a period covering most of the XIX century until 1925. Ferreira and Santa Clara (2011) addressed the predictability of international stock returns using dividend-price ratios, earnings growths, and price-earnings ratio growths during the period from 1927 to 2007 finding substantial predictability, hence concluding that it would have been possible to profitably "time the market". Neely et al. (2014) examined the importance of technical analysis indicators in forecasting the equity risk premium in the US market from 1950 to 2011. They concluded that these indicators displayed statistically and economically significant in-sample and out-of-sample prediction power, matching or exceeding that of macroeconomic variables. Golez and Koudijs (2018) showed that dividend-price ratios did not predict US stock returns from 1871 to 1945, however, that forecastability seemed to exist afterward until 2015. Piatti and Trojani (2019) also reached a similar conclusion to the previous study.

Although the mainstream research has focused solely on stocks and bonds, more recently, other types of assets, such as Real Estate Investment Trusts (REITs), have attracted the attention of researchers. According to Bhuyan et al. (2014), REITs have been an alternative investment vehicle since the 1980s. Historically, REITs have been a desirable financial asset by providing diversification benefits, improving the risk-return trade-off, and supplying a non-negligible dividend income. Since REITs tend to adjust quickly to the cost of living, they provide an hedge against inflation, turning their real return relatively stable. Furthermore, REITs returns show high predictability since their income comes from the underlying commercial real estate with long-term lease periods (Bhuyan et al., 2014; Fugazza et al., 2015).

There is extensive research on which variables are more suitable to predict excess returns of stocks, bonds, and REITs. Research on the potential predictors of asset returns goes far back to 1933, with a seminal paper entitled "Can Stock Market Forecasters Forecast?" (Cowles, 1933). In that article, the author reported that portfolios based on broad market recommendations from 24 individual financial publications between 1928 and 1932 failed to outperform a passive investment in the DJIA index by 4% annually. The author also highlighted that the performance of the most successful portfolios was not substantially different from what would be expected from pure chance.

In the early 1960s, several studies examined the forecast power of several technical indicators such as moving averages, filter rules, and momentum oscillators. This line of research was recently recovered by some authors such as Neely et al. (2014), Gao et al. (2018), and Zhang et al. (2019). Besides these indicators, the literature has provided a broad list of predictors of stock and bond returns, such as the dividend-price ratio (Campbell and Shiller, 1988; Cochrane, 2007), earnings-price ratio (Campbell and Shiller, 1988b), book-to-market ratio (Kothari and Shanken, 1997), nominal interest rate and interest rate spread (Fama, 1990, Rapach et al., 2016), labour income (Santos and Veronesi, 2006), stock market volatility (Guo, 2006), aggregate output (Rangvid, 2006), lagged industry returns (Hong et al., 2007), oil prices (Driesprong et al., 2008) and oil-relative variables (Nonejad, 2018), output gap (Cooper and Priestly, 2009), expected business conditions (Campbell and Diebold, 2009), accruals (Hirshleifer et al., 2009), inflation rate (Ludvigson and Ng, 2009), and manager sentiment (Jiang, 2019). Welch and Goyal (2008) summarized a list of several variables that have been used in the literature with positive results. The present study considers not only the list of Welch and Goyal (2008) but also other variables that have also been used for predicting returns in a multi-asset framework.

As mentioned before, most studies report in-sample evidence on return predictability. The predominance of in-sample studies could be justified by using all available data which increases the power of econometric tests (Neely et al., 2014). As argued by the authors, in-sample estimations produced efficient and precise estimates of the parameters. However, in-sample tests may be biased if the predictor and return innovations are correlated, and the predictor is highly persistent (Stambaugh, 1999). That bias potentially leads to substantial size distortions in the usual t-tests on the significance of the variables.

The focus on in-sample predictability has been gradually shifting to out-of-sample predictability. For instance, Welch and Goyal (2008) and Thornton and Valente (2012) showed that, although some predictive variables successfully predicted returns in-sample, they were not significant out-of-sample. Predictions based on these variables failed to consistently outperform the simple historical average benchmark forecast in terms of Mean Squared Forecast Error (MSFE). In fact, the so-called “kitchen sink” forecast model, a multiple regression model that includes all potential predictors, also performed much worse than the historical average forecast. This is not surprising, as it is well-known that, due to in-sample over-fitting, highly parameterized models typically perform worst in out-of-sample configurations. In sum, the authors argued that forecasting regressions were not stable, and traditional forecasting methods performed worse than the historical average.

Whether returns are predictable out-of-sample or not is still an ongoing debate. According to Wu et al. (2013), the conflicting empirical results presented in the literature may be related to problems such as data mining, spurious regressions, and instability of return predictability. Hence, recent studies have provided adaptative methods that accommodate and improve forecasting in a dynamic setup. Amongst such methods, are those with time-varying parameters or time-varying volatility (Dangl and Halling, 2012), the adoption of a diffusion index approach to improve equity premium forecasting (Ludvigson and Ng, 2007), the combination of a large number of potential return predictors (Rapach et al., 2013, Fisher et al., 2020, Zhang et al., 2018, Bahrami et al., 2019, and Gargano et al., 2019), or the consideration of regime shifts (Hammerschmid and Lohre, 2018). Nevertheless, other recent studies have argued in favour of traditional predictive regressions, showing that these methods, updated with schemes to resolve parameter uncertainty and instability, outperform the historical average forecast in out-of-sample experiments (see, for instance, Rapach et al., 2013, Koop and Korobilis, 2013, or Fisher et al., 2020). Hence, this study mainly considers out-of-sample predictability by using up-to-date forecasting methods.

Typically, asset returns are forecasted by past values of predictor variables within a Vector Autoregressive (VAR) framework (Guidolin and Hyde, 2012). VAR models provide a coherent way to generate internally consistent multiperiod forecasts that account for concurrent and dynamic correlations across the variables (Elliott and Timmermann, 2008). The basic VAR model is a valuable tool for a small number of assets and a predictor set with low cardinality; however, as more variables are included, the number of parameters required to be estimated increases, possibly leading to a rise in estimation errors. Fortunately, different methodologies have been developed to deal with this issue, such as Bayesian methods, that make use of the high computational power that is now available to researchers. Koop and Korobilis (2013) and Dangl and Halling (2012) are two examples of applications of such methodologies with positive results (their models outperformed the forecasts based on the historical mean equity premium over a wide range of periods).

According to Parslow et al. (2013), the main advantage of Bayesian methods is the potential to systematically incorporate previous knowledge on models and parameters. Additionally, Bayesian frameworks allow for discounting or ignoring prior information through uninformative priors. Barberis (2000) examined optimal asset allocation for stocks and cash through a Bayesian framework by incorporating parameter uncertainty in the model specifications, i.e., the author specified uninformative prior beliefs to the parameters characterizing the linear relationships between asset returns and predictors. This method worked as follows: first, a posterior distribution of the parameters is obtained by applying Bayes' rule. Second, the resulting joint posterior distribution is used to generate the conditional predictive density of returns, a predictive distribution of future utility levels, and then the portfolio weights (Fugazza et al., 2015).

Bayesian models may also include numerous efficient modelling tools to prevent the proliferation of parameters and eliminate parameter/model uncertainty. An example is the Bayesian model averaging and model selection methods. For instance, Koop and Korobilis (2013) used Bayesian-based dynamic model averaging and model selection (DMA and DMS) methods for estimating and forecasting large VARs and time-varying VARs models.

Koop and Korobilis (2013) differed from the rest of the literature that typically worked with single VARs. These methods have three main advantages. First, they do not require Markov Chain Monte Carlo (MCMC) simulations. Instead, they rely on estimated discount factors that characterize the degree of variation of the VAR coefficients. Second, they allow switching between different models, which mitigates overparameterization. Basically, the method selects a model over a set of different dimensions based on the past predictive likelihoods of the dependent variables. Third, these methods consider time-varying parameters. Typically, forecasting models assume that coefficients are constant over time, although there is ample evidence of instability in the relationship between asset returns and predictors. This has led researchers, such as Dangl and Halling (2012) and Koop and Korobilis (2013), to consider time-varying parameters, highlighting that there is strong evidence on the existence of breaks in predictive regressions, which, if taken into account, may have a substantial impact on the optimal asset allocation.

The literature on asset allocation has shown substantial utility benefits when investors incorporate return predictability into their investment decisions. Nevertheless, it is important to account for model uncertainty in the asset allocation decision to realize such benefits (Rapach and Zhou, 2013; Diris et al., 2014). Hence, several studies have implemented Bayesian approaches to portfolio strategies. Johannes et al. (2014), Gargano et al. (2017), Gao and Nardari (2018), and Fisher et al. (2020) investigated the optimal asset allocation of a Bayesian investor endowed with a power utility function, i.e., a Constant Relative Risk Averse (CRRA) investor (Campbell and Viceira, 2003).

Most academic studies regarding asset allocation focus on solving recursive myopic portfolio optimization problems (DeMiguel, et al., 2009, Daskalaki and Skiadopoulos, 2011, and Cenesizoglu and Timmermann, 2012). However, an increasing number of studies has included dynamic optimization frameworks (Almadi, 2014; Johannes et al., 2014; and Fisher et al., 2020). For instance, Johannes et al. (2014) analysed the out-of-sample success of dynamic portfolio choice in a model where the stock market was the only risky asset over a two-year investment horizon. They concluded that dynamic approaches outperform static baseline approaches. Therefore, we also implement a dynamic optimization approach when evaluating portfolio strategies.

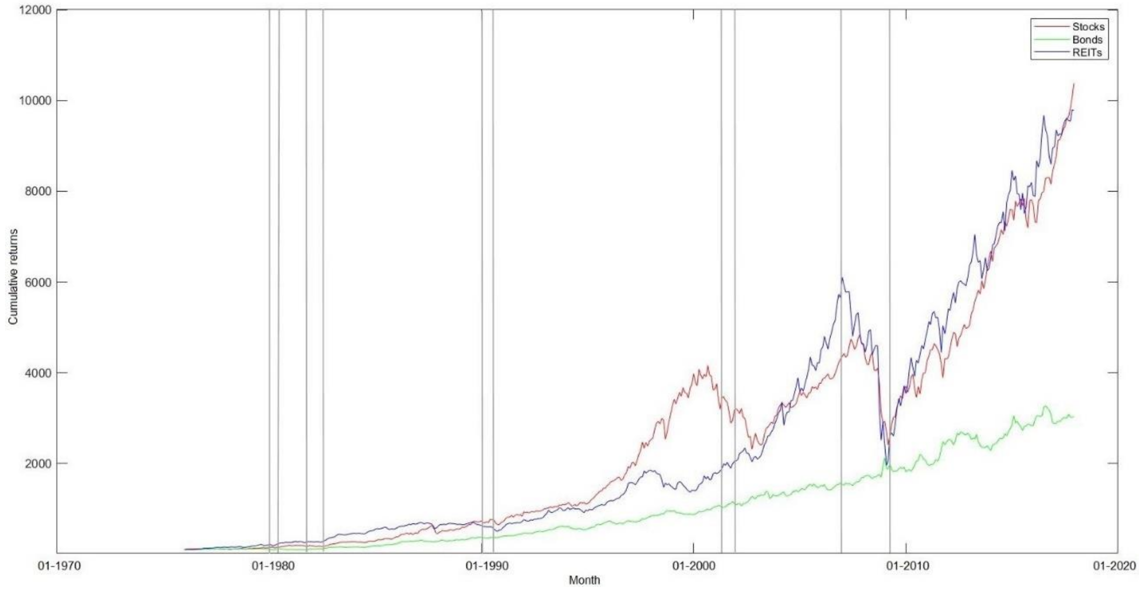
2.3. Data Description and Preliminary Analysis

2.3.1. Asset Classes

We consider three US-based asset classes: Stocks, Bonds, and REITs. Stocks are proxied by the S&P 500 total return index, Bonds are proxied by the Barclays Capital US aggregate bond index, and REITs refers to an index of US publicly traded Equity Real-estate Vehicles. Data on stocks and bonds were obtained from the Welch and Goyal (2008) database, while data on REITs were obtained from the NAREIT website (https://www.reit.com/data-research). The monthly log-returns of these indices were then used to compute excess returns, by subtracting the risk-free rate (proxied by the monthly yield-to-maturity of 3-month Treasury Bills, also obtained from the Welch and Goyal database). The total period spans from January 1976 to December 2017, covering 504 months.

Figure 2.1 plots the cumulative returns of the total return indexes of the Stocks, Bonds, and REITs (using a base value of 100). All series are notably more volatile after 2000, and Stocks and REITs are more sensitive to the business cycle than Bonds. However, the cumulative returns of Stocks and REITs dominate those of Bonds throughout the all sample.

Figure 2.1: Total return indexes



*Notes***:** This figure shows the path of the total return indexes of Stocks, Bonds and REITs computed using cumulative log-returns, i.e., $I_t = 100\exp\left(\sum_{\tau=1}^{t} r_{I,\tau}\right)$. The vertical lines bound the recession periods in the US according to the NBER (https://www.nber.org/research/data/us-business-cycle-expansions-and-contractions).

To analyse the predictability of the returns of these three classes of assets, we divided the sample into an in-sample (IS) period (2/3 of the sample, corresponding to 336 months) and an out-of-sample (OOS) period (1/3 of the sample, corresponding to 168 months).

Table 2.1 reports the summary statistics of Stocks, Bonds, and REITs excess returns for the overall sample, in-sample, and out-of-sample periods. The mean excess stock return is approximately 0.54%, ranging from -24.82% to 12.23%. The mean is higher in the most recent data (OOS), 0.63%, and the range amplitude is lower, with the excess returns ranging between -18.33% to 10.35%. The same happens with the REITs excess returns, which are on average slightly higher in the full sample (0.55%) than those of Stocks, ranging from -36.05% to 24.65%.

Both Stocks and REITs offer higher monthly excess returns on average than the Long-term U.S. government Bonds for the full sample (the mean excess return of Bonds is only 0.30%). There is no significant difference in the standard deviations, except for REITs in the OOS period, where the standard deviation more than doubled relative to the previous period. The first-order autocorrelations indicate some persistence in excess returns, especially for Stocks in the OOS period and REITs in the overall sample. Bonds have positive skewness, whilst Stocks and especially REITs are negatively skewed. The three classes of asset present mild excess kurtosis, and REITs stand out as the one with the most leptokurtic distribution.

Table 2.1: Descriptive statistics of excess returns

Variable	**Sample**	**Stocks**	**Bonds**	**REITs**
Mean	**Full**	0.54	0.30	0.55
	IS	0.51	0.24	0.53
	OOS	0.63	0.42	0.62
Std	**Full**	4.28	3.12	4.82
	IS	4.46	3.11	3.72
	OOS	3.88	3.14	6.48
$\boldsymbol{\rho(1)}$	**Full**	4.30	4.84	10.61
	IS	-0.12	7.03	15.70
	OOS	15.83	0.36	07.24
Max	**Full**	12.23	13.46	24.65
	IS	12.23	13.08	13.20
	OOS	10.35	13.46	24.65
Min	**Full**	-24.82	-11.93	-36.05
	IS	-24.82	-10.41	-15.92
	OOS	-18.33	-11.93	-36.05
Skew	**Full**	-0.84	0.12	-1.41
	IS	-0.76	0.10	-0.51
	OOS	-1.02	0.1338	-1.56
Kurt	**Full**	6.20	4.96	13.07
	IS	6.04	4.69	5.50
	OOS	6.26	5.51	10.88

Notes: This table reports some summary statistics of Stocks, Bonds and REITs monthly excess log-returns over the period from January 1976 to December 2017 (504 monthly observations). *Full* refers to the full sample period, *IS* is the in-sample period ($2/3$ of the sample, corresponding to 336 months), and *OOS* is the out-of-sample period ($1/3$ of the sample, corresponding to 168 months). The statistics are the *Mean*, standard deviation, *Std*, first-order autocorrelation, $\rho(1)$, maximum and minimum values, *Max* and *Min*, respectively, skewness, *Skew*, and kurtosis, *Kurt*. All values are percentages, except *Skew* and *Kurt*.

2.3.2. Predictive Variables

We have collected a comprehensive set of predictive variables, which were documented in the literature on asset return predictability. We attempt to be as comprehensive as possible rather than arbitrarily selecting a few predictors indicated in previous studies. Appendix 2.1 lists these variables (19 for stocks, 125 for Bonds and 14 for REITs). To reduce the dimensionality of the predictor space, we proceed by selecting, for each asset class, the 5 predictors with the highest absolute correlation between excess returns at time t and predictors at time $t-1$ in the IS period. The chosen variables and the corresponding correlation coefficients are shown in Table 2.2.

Table 2.2: Selected predictors for each asset class

Stocks	Bonds	REITs
-0.0848 *(Net Equity Expansion)*	**-0.1686** *(Retail Sales)*	**-0.0843** *(Change in Employment)*
0.0895 *(Small minus Big Factor)*	**0.1805** *(1-Year Federal Fund Spread)*	**0.0855** *(Term Spread)*
-0.0954 *(High minus Low Factor)*	**-0.1828** *(New Order for Non-Defence Capital Goods)*	**-0.0867** *(Inflation)*
-0.0960 *(Inflation)*	**-0.1842** *(CPI Durables)*	**-0.0987** *(Mortgage Loan Amount)*
-0.1149 *(Conservative minus Aggressive Factor)*	**0.1847** *(Composite Federal Fund Spread)*	**0.1013** *(Default Yield Spread)*

Notes: This table reports, for each asset class, the 5 predictors with the highest absolute correlations. In bold are the correlation coefficients between the excess returns of Stocks, Bonds, and REITs and the lagged values of the predictive variables, in the in-sample period.

Notice that the variable *Inflation* is a common predictor of Stocks and REITS, so we end up with 14 predictive variables.

2.4. Methodology

This section presents the basic theoretical concepts and specifications. It begins by presenting simple Vector Autoregressive models (VAR) and the procedures used to obtain forecasts based on the combinations of various VAR models considering different combinations of the variables in the predictor space. Then we present the Time-Varying Parameter Vector Autoregressive model (TVP-VAR), the procedures used to estimate these models with forgetting factors, and the methods to select or combine these models. Finally, we present the measures of forecasting accuracy and portfolio performance from the perspective of a risk-averse investor.

2.4.1. Vector Autoregressive Models (VAR)

Since their introduction by Sims (1980), Vector autoregressive models (VARs) have become an important tool for predicting macroeconomic and financial time series. These models are a straightforward multivariate generalization of univariate autoregressions and can generate dynamic forecasts that ensure consistency across different equations and forecast horizons. For this reason, many researchers in macroeconomics and finance have used large VARs that involve dozens or even hundreds of dependent variables (see, among many others, Banbura et al., 2010, Carriero et al., 2009, and Koop and Korobilis, 2013).

This study implements first-order vector autoregressive models, VAR(1), to capture the dynamics of asset returns and predictors,[1] as in Campbell et al. (2003), such that:

$$\mathbf{z}_t = \mathbf{\Phi}_0 + \mathbf{\Phi}_1 \mathbf{z}_{t-1} + \boldsymbol{\epsilon}_t, \text{ where } \mathbf{z}_t \equiv \begin{bmatrix} \mathbf{r}_t \\ \mathbf{s}_t \end{bmatrix}, \tag{2.1}$$

where $\mathbf{z}_t$ denotes an $(m \times 1)$ column state vector at time t, with $m = m_1 + m_2$, which includes the $(m_1 \times 1)$ vector of excess log-returns at time t, $\mathbf{r}_t$, and the other forecasting variables into an $(m_2 \times 1)$ single state vector $\mathbf{s}_t$. In our empirical application, $\mathbf{r}_t = [r_{1,t}\ r_{2,t}\ r_{3,t}]'$, where $r_{1,t}$, $r_{2,t}$, and $r_{3,t}$ are the excess log-returns over the risk-free rate of Stocks, Bonds, and REITs, respectively. $\mathbf{\Phi}_0$ is the $(m \times 1)$ vector of intercepts, $\mathbf{\Phi}_1$ is the $(m \times m)$ matrix of slope coefficients, and $\boldsymbol{\epsilon}_t$ is the vector of shocks to the state variables satisfying the following distributional assumptions:

$$\boldsymbol{\epsilon}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}), \tag{2.2}$$

$$\mathbf{\Sigma} \equiv Cov_{t-1}(\boldsymbol{\epsilon}_t) = \begin{bmatrix} \mathbf{\Sigma}_{xx} & \mathbf{\Sigma}'_{xs} \\ \mathbf{\Sigma}_{xs} & \mathbf{\Sigma}_{ss} \end{bmatrix}. \tag{2.3}$$

Shocks are independently normal distributed, homoscedastic, with zero mean and covariance matrix $\mathbf{\Sigma}$, but they can be cross-sectionally correlated. Hence, the VAR captures the linear dependence structure between all the state variables in $\mathbf{z}_t$.

In the present study, we implement a VAR(1) with 3 assets and 3 predictors. This implies the estimation of $6 + 6 \times 6 = 42$ coefficients. Since there are 15 predictive

[1] The VAR(1) avoids additional lags that would require a larger state vector with a large number of parameters. Nevertheless, this representation is not restrictive, since any vector autoregression can be rewritten as a VAR(1) by increasing the variables in the state vector.

variables to choose from, 5 for each asset return, we would have $5^3 = 125$ possible combinations. However, it turns out that one predictor is common to both stocks and REITS, therefore the number of different combinations reduces to 120.

2.4.2. Forecast Combinations

In order to obtain the forecasts of the excess return vector, we use several methods drawn upon the combination of individual forecasts, $\mathbb{E}(\mathbf{r}_t | \mathcal{M}_j, \mathcal{F}_{t-1})$, and the forecasted covariance matrix, $\widehat{Cov}(\mathbf{r}_t | \mathcal{M}_j, \mathcal{F}_{t-1})$, considering the information set up to time $t-1$, $\mathcal{F}_{t-1}$, under the various models $\mathcal{M}_j$, with $j = 1, 2, \dots, N$. These combinations use the mean, median, trimmed-mean, and weighted mean squared forecasting errors (WMSFE) for the purposes of statistical analysis and the mean and WMSFE for the economic analysis. We benchmark the performance of these combinations against the historical average, which is simply the average of past asset returns up to the date on which the prediction is made.

The mean and median combinations forecasts are just the mean and median of the individual forecasts, respectively. The trimmed-mean forecast sets the weight of the individual forecasting models at time t as $\omega_{j,t} = 0$, for the lowest and highest forecasts, and $\omega_{j,t} = 1/(N-2)$ for the remaining forecasts. Due to its simplicity and the fact that it does not require any estimation procedure, these combinations are frequently used in the literature and often perform better than other, more complex, combination methods.

Despite the simplicity of the previous methods, it can be advantageous to give more emphasis to certain individual forecasts. A way to achieve this is by using the WMSFE (also known as the square Mahalanobis distance). The WMSFE of model j, is computed as follows:

$$WMSFE_j = \frac{1}{T-t_o} \sum_{t=t_o+1}^{T} \mathbf{e}'_{j,\tau} [\widehat{Cov}(\mathbf{r}_t)]^{-1} \mathbf{e}_{j,t}, \tag{2.4}$$

where $\mathbf{e}_{j,t}$ is the column vector of forecast errors at time t associated with model j, such that, $e_{j,i,t} = r_{i,t} - \hat{r}_{j,i,t}$ is the forecast error of the return of asset i (with $i = \{Stock, Bonds, REITs\}$ at time t computed as the difference between the asset excess return, $r_{i,t}$, and the one-step-ahead asset excess return forecast, $\hat{r}_{j,i,t}$, at time t using model j. $\widehat{Cov}(\mathbf{r}_t)$ denotes the sample estimate of the asset excess returns unconditional covariance matrix, computed over the evaluation period (in-sample), t_o denotes the end of the in-sample period, and T is the overall sample size.

The combination based on the WMSFE emphasizes certain individual forecasts since the covariance matrix weights the forecast errors of asset returns differently according to the variability and correlation of asset returns. This procedure attributes higher penalties to forecast errors with lower standard deviations, that is, those on which the investor is highly confident. Conversely, it penalizes more lightly diffuse forecasts. Similarly, there are higher penalties for forecast errors in contrary directions for correlated assets and high penalties for forecast errors in the same direction for negatively correlated assets.

Furthermore, combinations based on the WMSFE allow the weights on individual forecasting models to reflect their past predictive accuracy. More specifically, we compute model i weight at each point in time by looking at its WMSFE in the period before. This means that the weight combination based on the WMSFE uses the information up to the period in which the forecast is made (that is, in the return forecasts of t it uses the WMSFE calculated up to $t-1$). The weight of the model j at time t is given by:

$$\omega_{j,t} = \frac{\varphi_{j,t}^{-1}}{\sum_{j=1}^{n} \varphi_{j,t}^{-1}}, \tag{2.5}$$

Where $\varphi_{j,t}$ are the sorted WMSFE for the *n* models according to the WMSFE for the period t. The weight $\omega_{j,t}$ may be computed using all models or just a subset of these models. In this study we consider several values of n, corresponding to the 10%, 20%, 30%, 40%, and 50% best models. This forecasting scheme attaches greater weight to individual predictive forecasts with lower WMSFE (better forecasting performance) over period t.

Regardless of the method implemented, there are numerous advantages to using forecast combinations. These combinations allow the researcher to use information across individual forecasts, can be seen as a diversification strategy in asset allocation theory may capture different aspects of business conditions, and provide information signals to models and predictive power variations through time (Bates and Granger, 1969, Rapach et al., 2010). For instance, if the correlation between individual forecasts is weak, their combination may produce models that are less unstable and, in this way, stabilize forecasts, reduce forecast risk, and improve forecast performance under model instability and uncertainty (Rapach and Zhou, 2013).

2.4.3. Time-Varying Parameter Vector Autoregressive (TVP-VAR)

A Time-Varying Parameter-Vector Autoregressive model of order 1, TVP-VAR(1), may be represented as follows:

$$\mathbf{y}_t = \mathbf{X}_t \boldsymbol{\beta}_t + \boldsymbol{\varepsilon}_t, \tag{2.6}$$

and

$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \mathbf{u}_t. \tag{2.7}$$

Where $\boldsymbol{\varepsilon}_t$ is *i.i.d.* $\mathcal{N}(0, \boldsymbol{\Sigma}_t)$ and $\mathbf{u}_t$ is *i.i.d.* $\mathcal{N}(0, \mathbf{Q}_t)$. $\boldsymbol{\varepsilon}_s$ and $\mathbf{u}_t$ are independent of one another for all t and s. $\mathbf{y}_t$, for $t = 1, \dots, T$, is an $(M \times 1)$ vector containing the observations on M time series variables and $\mathbf{X}_t$ is a $(M \times k)$ matrix defined so that each TVP-VAR equation contains an intercept and the first lag of each of the M variables:

$$\mathbf{X}_t = \begin{pmatrix} \mathbf{x}'_t & 0 & \cdots & 0 \\ 0 & \mathbf{x}'_t & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \mathbf{x}'_t \end{pmatrix}, \tag{2.8}$$

where $\mathbf{x}_t$ is a vector containing an intercept and one lag of each of the M variables. Thus, $k = M(1 + M)$.

The TVP-VAR was commonly estimated using forgetting factors (also known as discount factors) in the past when computer capacities were small. However, this method is still implemented in recent applications due to its simplicity and fast-tracking (see, for instance, Dangl and Halling, 2012, and Koop and Korobilis, 2013).

Let first denote $y^s = (y_1, \dots, y_s)'$ as the observations through time s. The Bayesian inference for $\boldsymbol{\beta}_t$ involves the Kalman filter. The state vector distribution at $t - 1$, using information up to date $t - 1$, is

$$\boldsymbol{\beta}_{t-1} | y^{t-1} \sim \mathcal{N}\left(\boldsymbol{\beta}_{t-1|t-1}, \mathbf{V}_{t-1|t-1}\right), \tag{2.9}$$

where $\boldsymbol{\beta}_{t-1|t-1}$ and $\boldsymbol{V}_{t-1|t-1}$ formulations may be found, for instance, in Hamilton (1994). The distribution of the state vector in the next period, using the same information set, is

$$\boldsymbol{\beta}_t|y^{t-1} \sim \mathcal{N}\left(\boldsymbol{\beta}_{t|t-1}, \mathbf{V}_{t|t-1}\right), \tag{2.10}$$

for

$$\mathbf{V}_{t|t-1} = \mathbf{V}_{t-1|t-1} + \mathbf{Q}_t. \tag{2.11}$$

By replacing $\mathbf{Q}_t = (\lambda^{-1} - 1)\mathbf{V}_{t-1|t-1}$ in the previous equation we obtain

$$\mathbf{V}_{t|t-1} = \frac{1}{\lambda}\mathbf{V}_{t-1|t-1}, \tag{2.12}$$

where λ denotes the forgetting factor, with $0 < \lambda \leq 1$. This equation implies that observations h periods in the past have weight λ^h in the filtered estimate of $\boldsymbol{\beta}_t$. Raftery et al. (2010) simply set λ as a number very close to one, $\lambda = 0.99$. These authors state that this number leads to a stable model, where the coefficient changes gradually, and the parameter has properties similar to what Cogley and Sargent (2005) call a "business as usual" prior. Dangl and Halling (2012) conducted their estimations with two granularity choices for λ, more precisely $\lambda \in \{0.96, 0.98, 1.00\}$ and $\lambda \in \{0.96, 0.97, 0.98, 0.99, 1.00\}$. These authors refer that a value strictly lower than one corresponds to an increase in the variance of the coefficient vector by a factor of $\frac{1}{\lambda}$ per period, where λ represents the weight loss of past observations compared to the last one. For example, with $\lambda = 0.98$ an observation that occurred 20 periods ago is weighted in the variance estimate by $0.98^{20} = 66.7\%$ of the weight of the last observation, which implies that the weight of the last observation is $1/0.667 - 1 = 50\%$ greater than the weight of an observation 20 periods earlier. This characterizes a situation where coefficients are very unstable. That is why the authors consider 0.98 as the lower bound for λ. Koop and Korobilis (2013) implemented a more robust technique. Instead of simply setting it equal to a fixed value, they estimated λ using λ_t, such that

$$\lambda_t = \lambda_{min} + (1 - \lambda_{min})L^{f_t}, \tag{2.13}$$

where $f^t = -NINT(\tilde{\boldsymbol{\varepsilon}}'_{t-1}\tilde{\boldsymbol{\varepsilon}}_{t-1})$, and $\tilde{\boldsymbol{\varepsilon}}'_{t-1} = \mathbf{y}_t - \boldsymbol{\beta}_{t|t-1}\mathbf{X}_t$ is the one-step-ahead prediction error produced by the Kalman filter and *NINT* rounds to the nearest integer.

Following Koop and Korobilis (2013) we set $\lambda_{min} = 0.96$ and $L = 1.1$ to obtain values between 0.96 and 1 for the forgetting factor.

We proceed in the same way to eliminate the need to simulate the multivariate stochastic volatility in the measurement equation. We use an Exponentially Weighted Moving Average (EWMA) estimator for the error covariance matrix:

$$\widehat{\mathbf{\Sigma}}_t = \kappa\widehat{\mathbf{\Sigma}}_{t-1} + (1-\kappa)\tilde{\boldsymbol{\varepsilon}}_t\tilde{\boldsymbol{\varepsilon}}'_t, \tag{2.14}$$

where the decay factor κ is set to 0.96, as in Riskmetrics (1996). The computation of $\widehat{\mathbf{\Sigma}}_t$ also requires the choice of an initial condition for $\mathbf{\Sigma}_0$, that we set equal to the sample covariance matrix of $\mathbf{y}^{t_0}$, where $t_0 + 1$ is the initial period at which the forecasting estimations start.

2.4.3.1 TVP-VAR selection

Models such as TVP-VARs are designed to accommodate gradual changes in the coefficients (Koop and Korobilis, 2013). However, they are unable to adjust to abrupt changes, which reduces their performances. A way to deal with the possibility of significant changes is to allow the switching between different models. Thus, we enable the TVP-VAR(1) to change dimensions over time by using a Dynamic Model Selection (DMS) procedure. This procedure requires the estimation of the forgetting factors, the definition of the priors, and the definition of various dimensions of the combinations of the TVP-VARs models. We use the DMS to select the optimal values for the VAR shrinkage parameter in a time-varying manner.

To implement DMS, we consider TVP-VARs models with different sets of explanatory variables, that is, we consider different $j = 1, \ldots, N$ models.

DMS is a recursive algorithm where the important recursions are similar to the forecast and updating equations of the Kalman filter method. Following Koop and Korobilis (2013), the model prediction and updating equations using a forgetting factor α are derived from:

$$\pi_{t|t-1,j} = \frac{\pi^{\alpha}_{t-1|t-1,j}}{\sum_{l=1}^{N}\pi^{\alpha}_{t-1|t-1,l}}, \tag{2.15}$$

which is the probability that model j will be chosen, given the information up to $t-1$, and,

$$\pi_{t|t,j} = \frac{\pi_{t|t-1,j} p_j(y_t|y^{t-1})}{\sum_{l=1}^{N} \pi_{t|t-1,l}\, p_l(y_t|y^{t-1})}, \tag{2.16}$$

where $p_j(y_t|y^{t-1})$ is the predictive likelihood (the predictive density of model j calculated at y_t). We can also write the probability used to select models as:

$$\pi_{t|t-1,j} \propto \prod_{k=1}^{t-1} \left[p_j(y_{t-k}|y^{t-k-1})\right]^{\alpha^i}. \tag{2.17}$$

Model j will receive more weight at time t according to the accuracy of its forecasts in the recent past. The weight of past predictive densities is controlled by the forgetting factor, α, which has similar features to the forgetting factor defined before, λ.

In our study, we set $\alpha = 0.99$ as in Koop and Korobilis (2013), which implies that the forecast performance five years ago receives 55% as much weight as the forecast performance in the last period. We do not define any other values for α, as in Koop and Korobilis (2013). The authors choose a range of values between 0.95 and 1, however, results are not very different for these minor interval changes. We also consider the case of $\alpha = 1$, however, this matches the conventional model averaging using the marginal likelihood. Setting $\alpha = 0.95$ indicates that the forecast performance five years ago receives only 5% as much weight. In sum, we set α as a fixed value of 0.99, and we consider $\lambda = 0.99$ and $\kappa = 0.96$.

Our approach does not require the estimation of $\mathbf{Q}_t$, and, as we referred before, it uses an EWMA estimator of $\boldsymbol{\Sigma}_t$ that requires prior information on $\boldsymbol{\beta}_0$. Thus, as in Koop and Korobilis (2013), we use a tight Minnesota prior for $\boldsymbol{\beta}_0$. In the literature, it is popular to use training sample priors to produce hyperparameters that monitor the degree of shrinkage when working with large VARs or TVP-VARs (Banbura et al., 2010). Typically, a constant prior is used over each point in time. However, Koop and Korobilis (2013) used a different approach which allows for the estimation of the shrinkage hyperparameter in a time-varying manner. To do so, they used an automatic updating procedure. This approach is less computationally demanding since it does not require the re-estimation of the shrinkage priors or the model at each point in time.

In our study, we use a normal prior for $\boldsymbol{\beta}_0$ similar to a Minnesota prior, set the prior mean to be $\mathbb{E}(\boldsymbol{\beta}_0) = 0$, and the Minnesota prior covariance matrix of $\boldsymbol{\beta}_0$ is set to be diagonal, $cov(\boldsymbol{\beta}_0) = \underline{V}$, such that

$$\underline{V_i} = \begin{cases} \underline{a}, \text{for the intercepts} \\ \delta, \text{for the lag coefficients} \end{cases}, \tag{2.18}$$

where $\underline{V_i}$ is the i-th diagonal element of the covariance matrix. Hence, δ controls for the degree of shrinkage on the VAR coefficients,[2] and $\underline{a} = 10^3$, for the intercepts to be uninformative. We need a large degree of shrinkage to obtain a good forecasting performance in large VARs and TVP-VARs. To do so, we estimate δ at each point in time using a strategy similar to the estimation of the forgetting and decay factors. As in Koop and Korobilis (2013), we use a grid for δ, such that $\delta \in \{10^{-10}, 10^{-5}, 0.001, 0.005, 0.01, 0.05, 0.1\}$.

In sum, the DMS implies choosing the model with the highest value of $\pi_{t|t-1,j}$ to obtain the forecast at time t. Since $\pi_{t|t-1,j}$ varies over time, the forecasting model will change, allowing, in this way, the model switching feature.

Besides the DMS, we also consider Dynamic Model Averaging (DMA), which uses $\pi_{t|t-1,j}$ as the weighting scheme.

We also consider different dimensions (i.e., different cardinalities of the state space set) when implementing the DMS and DMA procedures. These models are:

- A "Small" model without any additional predictors, implying that it only considers the first-order lags of the excess returns.
- A "Medium" model with three additional predictors. These predictors are the ones with the highest absolute correlation with the excess returns.
- A "Large" model with all additional predictors, corresponding to 17 variables (3 lagged excess returns plus 14 predictive variables).
- A "Full" model, which is selected or averaged (for the DMS and DMA, respectively) across all small, medium, and large models.

A crucial point of TVP-VAR selection and averaging is the calculation of $\pi_{t|t-1,j}$. When forecasting at time t, we evaluate this probability for every model j and use the values of δ and the dimension of TVP-VAR(1) that maximizes it. We use the recursive algorithm presented in Equation (2.15) and Equation (2.16) and set the initial probability of selecting each model equal to $\pi_{0|0,j} = 1/N$ for all models. However, when dealing with TVP-VAR with different dimensions we have different predictive densities,

[2] Note that this differs from the Minnesota prior in that the latter contains two shrinkage parameters (corresponding to own lags and other lags) and these are set to be constant. For ease of computation, we only use one shrinkage parameter, as in Banbura et al. (2010).

$p_j(y_t|y^{t-1})$, since $\mathbf{y}_t$ has different dimensions, rendering them incomparable. A possible solution is to use the same predictive densities for all dimensions. So, we used the small model TVP-VAR predictive density since the variables in this model are common to all models.[3]

2.4.4. Measures of Forecasting Accuracy

The precision of the point forecasts of each model j is measured using the Mean-Squared Forecast Errors (MSFEs). This is the most common metric for assessing forecasting accuracy, especially in return predictability studies. To calculate the MSFE, firstly, we divide the total sample into an initial in-sample period comprising the first t_o observations and an out-of-sample period comprising of the last $T - t_o$ observations, and then we compute the mean value of the squared forecasting errors for the model and the historical averaging:

$$MSFE_{j,i} = \frac{1}{T - t_o} \sum_{t=t_o+1}^{T} e_{j,i,t}^2, \tag{2.19}$$

$$MSFE_{HA,i} = \frac{1}{T - t_o} \sum_{t=t_o+1}^{T} e_{HA,i,t}^2, \tag{2.20}$$

where t_o denotes the beginning of the out-of-sample period, j refers to the models under consideration, $e_{j,i,t}$ is the forecast error of asset return i at time t associated with model j, $r_{i,t}$ is the asset return for period t and $\hat{r}_{j,i,t}$ is the one-step ahead asset return forecast. $e_{HA,i,t} = r_{i,t} - \bar{r}_{i,t}$ is the forecast error in relation to the historical average of asset i at time t, i.e., $\bar{r}_{i,t}$ is the return average of asset i at time t, using information up to time $t - 1$.

Following Campbell and Thompson (2008), the additional predictive power of the models relative to the historical average can be measured through the pseudo R^2 out-of-sample:

$$R_{OOS,j,i}^2 = 1 - \frac{MSFE_{j,i}}{MSFE_{HA,i}}, \tag{2.21}$$

[3] This implies that the dynamic model selection is determined by the joint predictive likelihood of the three asset returns.

Model j produces better predictions than the reference model (historical average) if the $R^2_{OOS,j,i}$ is positive. We also analyse whether the models exhibit higher predictive ability than the historical average using the adjusted mean squared forecast errors statistic of Clark and West (2007) (hereafter adj-MSFE). This test is an approximately normal modified version of the MSFE statistic, which the authors show to be undersized. Its null hypothesis stipulates that the MSFE of both the model and the historical average are equal, whereas, according to the alternative hypothesis, the model predictions are more accurate. The most convenient way to implement this one-side test is to compute for each model i and each excess asset return i, at time t:

$$\hat{f}_{j,i,t} = \left(r_{i,t} - \bar{r}_{i,t-1}\right)^2 - \left[\left(r_{i,t} - \hat{r}_{j,i,t-1}\right)^2 - \left(\bar{r}_{i,t} - \hat{r}_{j,i,t-1}\right)^2\right] \tag{2.22}$$

and then regress $\hat{f}_{j,i,t}$ on a constant and use the resulting *t*-statistics. The null hypothesis of equal predictive ability is rejected, for example, at the 5% significance level, if the *t*-statistic exceeds 1.645.

2.4.5. Asset Allocation Problem

Here we present the asset allocation problem faced by a long-term investor with constant relative risk aversion (CRRA) utility function:

$$U(W_t) = \frac{W_t^{1-\gamma}}{1-\gamma}. \tag{2.23}$$

In that utility function, W_t denotes the investor's wealth at time t, and γ, with $\gamma > 1$, is the risk-aversion coefficient. At each point in time, the investor chooses the optimal allocation amongst the risky assets and a risk-free asset that maximizes her 1-period-ahead expected utility $\mathbb{E}_t[U(W_{t+1})]$. The optimal weights implied by model j are given by the solution of the following constrained maximization problem:

$$\begin{gathered} \arg\max_{\mathbf{x}_{j,t}} \mathbf{x}'_{j,t}\left(\hat{\boldsymbol{\mu}}_{j,t+1|t} + \frac{1}{2} diag\hat{\boldsymbol{\Sigma}}_{j,t+1|t}\right) - \frac{\gamma}{2}\mathbf{x}'_{j,t}\,\hat{\boldsymbol{\Sigma}}_{j,t+1|t}\,\mathbf{x}_{j,t} \\ s.t.\text{: (1) } \mathbf{x}'_{j,t}\,\boldsymbol{\iota} = 1 \text{ and (2) } \mathbf{x}_{j,t} \geq \mathbf{0} \end{gathered} \tag{2.24}$$

where $\mathbf{x}_{j,t}$ denotes the vector of portfolio weights, $\hat{\boldsymbol{\mu}}_{j,t+1|t} = \mathbb{E}(\mathbf{r}_{t+1}|\mathcal{M}_j, \mathcal{F}_t)$ is the mean of the predictive density of the vector of risky asset $\mathbf{r}_{t+1}$, computed using the information available at time t and under model j, $\widehat{\boldsymbol{\Sigma}}_{j,t+1|t} = \widehat{cov}(\mathbf{r}_{t+1}|\mathcal{M}_j, \mathcal{F}_t)$ is the risky assets' forecasted covariance matrix at time t based on the estimates given by models j and conditional on the information set at time t, and $\boldsymbol{\iota}$ is a vector of ones with the same length as $\mathbf{r}_{t+1}$. Following much of the asset allocation literature, we rule out short-selling (i.e., negative portfolio weights). Notice that the investment portfolio is rebalanced each month.

Assuming that the excess returns of the risky assets are log-normally distributed, the portfolio log-return implied by model j at time t is the following (Campbell et al., 2003):

$$r_{p,j,t+1} = r_{f,t} + \mathbf{x}'_{j,t}\left(\mathbf{r}_{t+1} - r_{f,t}\boldsymbol{\iota}\right) + \frac{1}{2}\mathbf{x}'_{j,t} diag\widehat{\boldsymbol{\Sigma}}_{j,t+1|t} - \frac{1}{2}\mathbf{x}'_{j,t}\widehat{\boldsymbol{\Sigma}}_{j,t+1|t}\mathbf{x}_{j,t} \quad (2.25)$$

where $r_{f,t}$ represents the continuously compounded risk-free rate.

2.4.6. Portfolio Performance Evaluation

Besides presenting the descriptive statistics of the returns of the portfolios (mean, standard-deviation, skewness, and kurtosis), we also use three metrics to assess the portfolio performance chosen by the risk-averse investor. The first metric is the Certainty Equivalent Return (CER) which is the risk-free return that would make the investor indifferent between following a certain investment strategy and accepting this risk-free return. The annualized CER can be expressed as follows:

$$CER_j = \left[\left(\frac{1}{T - t_0}\sum_{t=t_0+1}^{T} \widehat{W}_{j,t}^{1-\gamma}\right)^{\frac{12}{1-\gamma}} - 1\right] \quad (2.26)$$

where $\widehat{W}_{j,t} = \exp\{r_{p,j,t}\}$ denotes the realized wealth at time t as implied by model j.

The second measure is the annualized Sharpe ratio (SR), which measures the desirability of a risky investment strategy, by dividing the average by the standard deviation of the excess return. In other words, the SR measures the reward per unit of variability. It can be expressed as:

$$SR_j = \frac{\sqrt{12}\hat{\mu}_{r_{p,j}}}{\hat{\sigma}_{r_{p,j}}}, \tag{2.27}$$

where $\hat{\mu}_{r_{p,j}}$ and $\hat{\sigma}_{r_{p,j}}$ denote the mean and standard deviation of the portfolio excess return implied by model j at time t, respectively.

The third metric is the annualized Sortino ratio (SOR), which only considers the negative deviations from a certain target, B, i.e., the "downside risk".

$$SOR_j = \frac{\sqrt{12}\left(\hat{\mu}_{r_{p,j}} - B\right)}{\sqrt{\frac{1}{T - t_o}\sum_{t=t_o+1}^{T} min\left[\left(r_{p,j,t} - B\right), 0\right]^2}}, \tag{2.28}$$

where B is the reference point that constitutes the minimum acceptable rate of return, T is the total number of periods, and $t_0 + 1$ is the initial out-of-sample period. In the computation of SOR, we use $B = 0$. Notice that this is different from using the semi-standard deviation, which is obtained from the negative deviations relative to the endogenous mean, as the downside deviation is computed in relation to an exogenous reference point.

2.5. Empirical Results

This section presents the results on the forecast predictability of various model combination schemes for each asset class. Table 2.3 reports the pseudo-R^2 and the significance of the adj-MSFE test. Although there is only one model for small, medium, and large sets, the selection/average of δ is still made, among the 7 possible values reported in Section 2.4.3. The results of the DMS and DMS are the same for individual models (the same situation occurs in Koop and Korobilis, 2013).

Table 2.3: Out-of-sample pseudo-R^2

Forecasting schemes	Model	Stocks	Bonds	REITs
Mean		-4.06	5.00*	1.72
Median		-3.59	5.12*	1.32
Trimmed Mean		-3.95	5.03*	1.71
WMSFE		-3.85	3.99*	1.99
	Small	-1.79	1.83	0.56
DMS and DMA	Medium	3.69	1.96	2.47
	Large	0.63	-3.58	1.49**
DMS	Full	-0.43	-8.34	-0.60
DMA	Full	1.34	0.65	0.35

Notes: This table reports the out-of-sample pseudo-R^2 (Campbell and Thompson, 2008), in percentage, of the different forecasting schemes based on the mean, median, trimmed mean, WMSFE and DMS and DMA of the Small, Medium, Large and Full TVP-VAR(1) models. The DMS and DMA statistics are the same, except for the Full model. Associated with these statistics, the table also presents the significance of the adj-MSFE (Clark and West, 2007). Asterisks (*) and (**) indicate test statistics significant at the 10% and 5% levels.

Table 2.3 reports positive pseudo-R^2 for all VAR combinations to predict the excess returns of REITs returns, ranging from 1.32% to 1.99%, and, most notably, to predict excess returns of Bonds, ranging from 3.99% to 5.12%. However, only Bonds present statistical significance at the 10% level. For Stocks the VAR combinations have negative pseudo- R^2, hence the VAR combinations do not improve the forecasting ability relative to the historical mean return.

The DMS and DMA approaches provide better results than the VAR combinations for Stocks, although the DMS and DMA for the Small model and the DMS for the Full model still present negative pseudo-R^2. This is not the case for REITs, for which only the DMS and DMA for the Medium model show better results (although the highest significance is obtained for the DMS and DMA large model approaches), and, most especially, for Bonds, for which none of the DMS and DMA beat the VAR combinations.

These mixed results highlight the problem of model uncertainty faced when dealing with forecasting returns However, if one had to choose among all the approaches under study, the best candidate is the DMS (or DMA) using the Medium model. That is, selecting or averaging across TVP-VAR(1) a dimension given by the three excess returns plus three additional predictors, with the highest absolute correlation with the excess returns. The DMS Medium model presents a pseudo-R^2 of 3.69%, 1.96%, and 2.47% for Stocks, Bonds, and REITs, respectively. Although these pseudo-R^2 are lower than 5%, these are in fact, good results, as the literature reports pseudo-R^2 out-of-sample of around 1% for quarterly data and 0.5% for monthly data.

Figure 2.2 illustrates the cumulative differences between the squared prediction error of the historical average and the squared prediction error of several forecasting combinations for the three asset classes. For stocks, the series have negative values and are decreasing. The exception is the Full TVP-VAR DMA. This suggests that amongst these forecasting schemes, this is the only one that delivers out-of-sample forecasting gains on a consistent basis over time. For Bonds, all the forecasting schemes show mostly positive values, except the Full TVP-VAR DMS. Notably, there are jumps in the series in 2008, meaning that the forecasting schemes were able to better forecast the negative returns during the subprime crisis than the historical means. For REITs the combinations of VAR(1) based on the Mean, Median, Trimmed Mean, and WMSFE are clearly better than the Full TVP-VAR DMS and Full TVP-VAR DMA. For REITs those initial 4 combinations of VAR(1) were able to some extent forecast the effects of the subprime crisis, while the opposite happened in the latter 2 forecasting schemes.

Figure 2.2: Cumulative differences of squared forecast errors

Exhibit A: Stocks

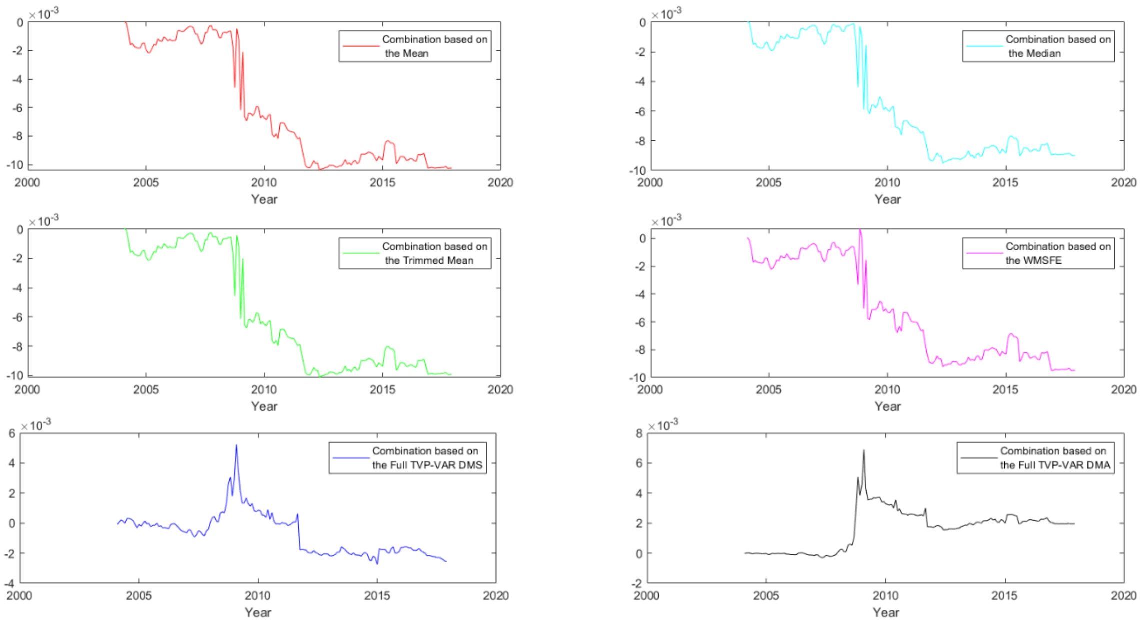


Exhibit B: Bonds

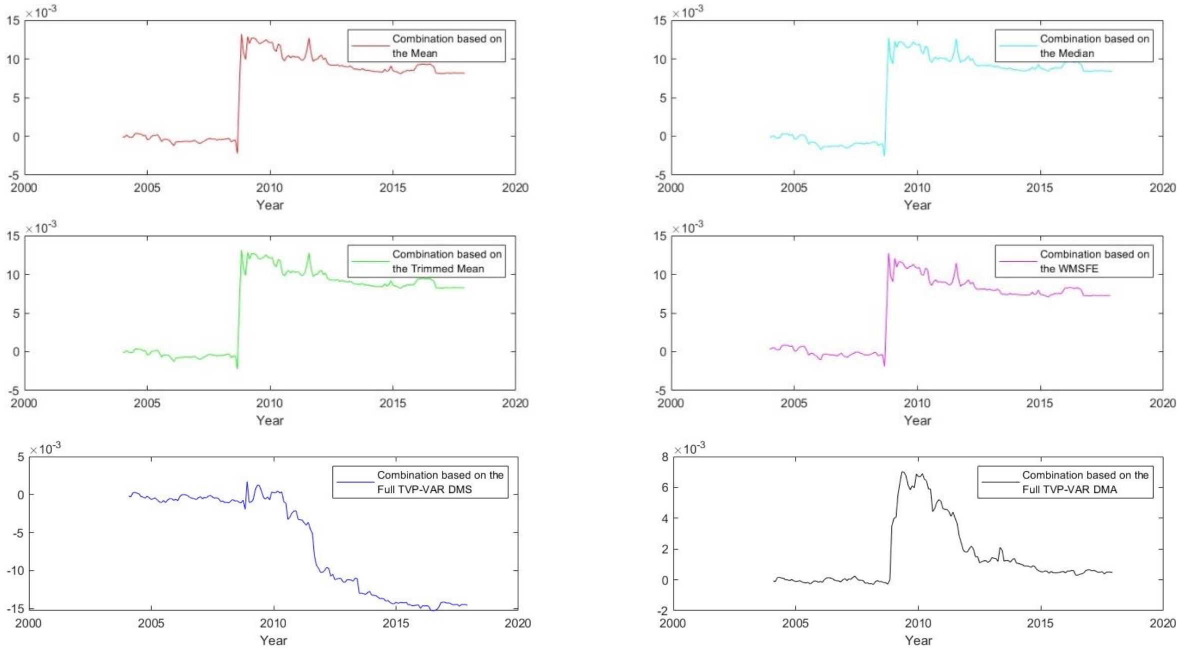
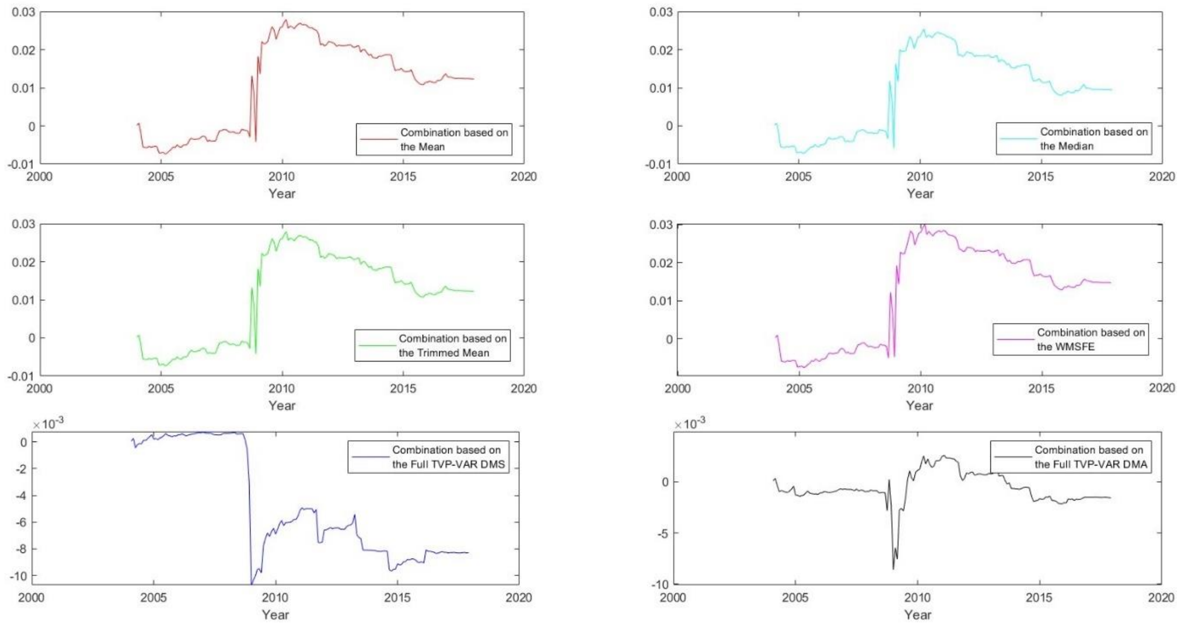


Exhibit C: REITs



Notes: This figure shows the cumulative differences between the squared prediction error of the historical average benchmark and the squared prediction error of a given forecasting combination (Mean, Median, Trimmed Mean, Full TVP-VAR DMS and DMA). Exhibit A, Exhibit B, and Exhibit C refer to Stocks, Bonds, and REITs, respectively.

Table 2.4 reports the mean, standard deviation, skewness, and kurtosis of the portfolio returns obtained using different approaches to forecast the asset returns considering three risk aversion coefficients ($\gamma = 3, 5,$ and 10).

The mean expected returns of the portfolios based on DMS and DMA are generally higher than those of all the other forecasting schemes. The only exceptions are the large individual TVP-VAR models. For instance, the mean return of a portfolio strategy for $\gamma = 5$ based on the Medium DMA/DMS is the largest, 5.99% per year. The equally weighted portfolio, 1/N, only achieves a mean return of 2.01%. Regarding the portfolio risk (measured by the standard deviation) for a level of risk aversion of $\gamma = 5$, we report a value of 14.28% per year for the medium DMA model, 16.25% for the full DMA approach, and only 11.49% for the 1/N portfolio.

It is not surprising that higher means are associated with higher risks. However, the standard deviations associated with portfolios based on Small, Medium, and Full DMS models are not substantially high relative to the average return they provide. Average returns for these strategies for an investor with $\gamma = 5$ range between 4.05% to 5.99% and the standard deviation between 14.28% and 18.65%, annually. The strategies based on WMSFE forecasting schemes have standard deviations ranging from 13.86% and 14.7%, but their average returns are very small, ranging between 3.36% to 3.60%. These results outline the superiority of Bayesian predictive models when formulating portfolio strategies.

Table 2.4 also reports the higher-order moments (skewness and kurtosis), which are also important for a power utility investor. Strategies based on WMSFE and Medium and Full DMA have positive skewness. In any case, the forecasting schemes provide a higher skewness than the 1/N portfolio. However, these models also produce relatively high kurtosis.

Table 2.4: Statistics of out-of-sample portfolio returns for different risk aversion coefficients

		Mean			Standard deviation			Skewness			Kurtosis		
Forecasting schemes		**γ = 3**	**γ = 5**	**γ = 10**	**γ = 3**	**γ = 5**	**γ = 10**	**γ = 3**	**γ = 5**	**γ = 10**	**γ = 3**	**γ = 5**	**γ = 10**
Mean		2.88	2.67	2.01	13.42	12.96	12.49	-0.88	-2.14	-3.69	16.44	16.69	21.62
WMSF	**10%**	3.78	3.36	2.60	15.49	13.86	11.49	2.75	0.16	-1.01	31.18	18.57	23.52
	20%	3.78	3.60	2.81	15.10	14.70	12.00	4.16	4.82	1.96	38.87	43.32	31.39
	30%	3.60	3.53	2.77	15.49	14.70	12.00	3.75	4.47	2.92	40.43	44.81	39.23
	40%	3.53	3.46	2.74	15.10	14.70	12.00	3.93	4.51	2.85	41.95	46.01	39.05
	50%	3.43	3.43	2.67	14.70	14.70	12.00	4.15	4.54	2.56	43.57	47.02	37.84
DMS and DMA	**Small**	5.51	5.65	5.79	14.70	14.28	12.96	-1.83	-1.01	0.95	29.76	25.75	20.04
	Medium	5.92	5.99	5.58	14.28	14.28	12.49	6.79	7.05	3.30	45.16	47.18	22.04
	Large	2.56	2.88	2.49	20.20	18.65	12.49	-2.60	-1.66	-2.81	34.13	37.94	25.17
DMS	**Full**	3.85	4.05	3.53	20.20	18.97	13.42	-1.95	-1.77	-1.34	34.88	36.66	26.89
DMA	**Full**	5.13	5.30	5.37	16.61	16.25	15.10	3.93	4.40	5.01	28.78	30.72	33.12
Historical		2.36	1.87	1.42	1.04	0.097	0.090	-1.57	-0.79	-0.18	34.21	33.87	37.00
1/N			2.01			11.49			-5.19			33.47	

Notes: This table reports the annualized mean, annualized standard deviation, skewness, and kurtosis of the portfolio returns based on the following forecasting schemes: Mean, WMSFE (10%, 20%, 30%, 40%, and 50% best models), DMS and DMA of the Small, Medium, Large and Full TVP-VAR(1). The portfolios were selected considering three risk aversion coefficients (γ = 3, γ = 5 and γ = 10). "Historical" refers to the portfolio based on the historical means and 1/N is the equally weighted portfolio. The mean and standard deviation are in percentage.

Table 2.5 shows some out-of-sample performance metrics of different portfolio strategies, considering different risk aversion coefficients. The portfolio strategy based on the historical means generates CER, SR, and SOR that are smaller than those of all forecasting schemes, except the strategy based on Large DMS/DMA for an investor with $\gamma = 3$. All the forecasting schemes are better than the 1/N, in all metrics and risk aversion coefficients, except the Large DMS/DMA for an investor with $\gamma = 3$ or $\gamma = 5$.

The Small and Medium DMS/DMA forecasting schemes provide the best performing strategies. For instance, the CER of portfolios based on the Medium DMS/DMA are twice as high as the CER of the other portfolios. For instance, an investor with $\gamma = 5$, who derives her portfolio strategy on a medium DMS/DMA model would have a CER of 19.16%, a SR of 1.47%, and a SOR of 4.04%.

Table 2.5: Portfolio performance for different risk aversion coefficients

		CER			SR			SOR		
Forecasting schemes		$\gamma = 3$	$\gamma = 5$	$\gamma = 10$	$\gamma = 3$	$\gamma = 5$	$\gamma = 10$	$\gamma = 3$	$\gamma = 5$	$\gamma = 10$
Mean		8.73	6.15	-0.87	0.74	0.71	0.57	1.18	1.08	0.81
WMSFE	**10%**	11.18	8.34	3.12	0.82	0.84	0.79	1.49	1.43	1.26
	20%	11.58	9.10	3.93	0.86	0.83	0.82	1.63	1.63	1.43
	30%	10.82	8.68	3.55	0.81	0.83	0.79	1.50	1.56	1.40
	40%	10.61	8.48	3.44	0.80	0.82	0.78	1.49	1.53	1.38
	50%	10.43	8.36	3.33	0.80	0.81	0.77	1.48	1.52	1.36
DMS and DMA	**Small**	18.43	16.71	13.89	1.31	1.37	1.55	2.30	2.51	3.22
	Medium	20.66	19.16	14.57	1.43	1.47	1.57	3.83	4.04	3.94
	Large	3.31	1.12	0.34	0.43	0.53	0.68	0.62	0.78	1.02
DMS	**Full**	8.28	5.08	2.78	0.66	0.74	0.90	1.00	1.16	1.47
DMA	**Full**	16.20	14.61	10.51	1.07	1.14	1.22	2.25	2.51	2.85
Historical		3.82	-1.39	-12.53	0.44	0.36	0.28	0.63	0.51	0.40
1/N		6.33	4.71	-0.12		0.62			0.86	

Notes: This table reports the annualized Certainty Equivalent Return (CER), Sharpe Ratio (SR), and Sortino ratio (SOR), in percentage, of the portfolio returns based on the following forecasting schemes: Mean, WMSFE (10%, 20%, 30%, 40%, and 50% best models), DMS and DMA of the Small, Medium, Large and Full TVP-VAR(1). The portfolios were selected considering three risk aversion coefficients ($\gamma = 3$, $\gamma = 5$ and $\gamma = 10$). "Historical" refers to the portfolio based on the historical means and 1/N is the equally weighted portfolio.

2.6. The impact of the subprime crisis

The question of how portfolio returns vary with economic activity is pertinent to the issue of optimal portfolio selection. In this section, we examine whether portfolio performances present substantial differences, before and after January 2008, the beginning of the subprime crisis. Hence, we investigate to what extent the models are suitable to accommodate market instability.

Table 2.6: Out-of-sample pseudo-R^2 before and after the subprime crisis (January 2008)

Forecasting schemes	**Model**	**Stocks**		**Bonds**		**REITs**	
		Pre	Post	Pre	Post	Pre	Post
Mean		-2.41	-4.27	-1.14	6.12	-1.43	2.40
Median		-1.49	-3.83	-2.90	6.66	-1.54	1.93
Trimmed Mean		-2.25	-4.15	-1.38	6.26	-1.46	2.39
WMSFE		-3.21	-4.78	-3.17	4.62	-1.84	2.02
DMS and DMA	Small	15.69	-2.24	-1.86	2.01	-3.43	0.72
	Medium	13.45	4.34	-1.48	1.96	-2.37	2.80
	Large	16.49	0.05	-3.49	-4.03	0.45	1.01
DMS	Full	17.83	0.83	-1.14	0.52	-0.84	-0.11
DMA	Full	16.64	-1.17	-3.60	-9.72	0.40	-1.54

Notes: This table presents a similar analysis to the one reported in Table 2.3, considering a partition of the data in January 2008, the beginning of the subprime crisis. "Pre" refers to the period from January 2004 to December 2007, and "Post" refers to the period from January 2008 to December 2017. This table reports the out-of-sample pseudo-R^2 (Campbell and Thompson, 2008), in percentage, of the different forecasting schemes based on the mean, median, trimmed mean, WMSFE and DMS and DMA of the Small, Medium, Large and Full TVP-VAR(1) models. The DMS and DMA statistics are the same, except for the Full model. Associated with these statistics, the table also presents the significance of the adj-MSFE (Clark and West, 2007). Asterisks (*) and (**) indicate test statistics significant at the 10% and 5% levels.

Table 2.6 reports the out-of-sample pseudo-R^2 of the different forecasting schemes considered in this study, over two sample periods, before the subprime crisis (from January 2004 to December 2007, we call it "Pre") and, during and after the crisis (January 2008 to December 2017, we call this period "Post").[4] Bayesian models perform

[4] We also analysed whether there was an asymmetric return predictability across business cycles. Following Sander (2018), we set up a recession dummy based on real-time indicators of recessions. We allowed the regression intercept and slope coefficients to change freely across recessions and expansions. We concluded

well in predicting Stocks excess returns before the crisis. The pseudo-R^2 for the period before the subprime crisis is surprisingly high for these schemes, varying between 13.45% to 17.83%. After 2008, the Medium DMS and DMA also present interesting results (the pseudo-$R^2$is 4.34%, 1.96%, and 2.80% for Stocks, Bonds, and REITs, respectively). This may indicate that the Medium DMS and DMA are suitable to accommodate periods with market turmoil. The simpler forecasting schemes also exhibit some predictive power for Bonds and REITs after 2008.

Table 2.7 reports the statistics of portfolios before and after 2008, while Table 2.8 present the annualized CERs, SR and SOR. Results for the skewness are negative before 2008 and positive after 2008 for every WMSFE strategy and Medium DMS/DMA. Standard deviations tend to be higher after 2008 due to the market turmoil. However, kurtosis is extremely high in both periods. Across all tables, almost all statistics and metrics exhibit better results after 2008 (the exception being the Mean forecasting scheme). For instance, for $\gamma = 3$, the CER of Medium DMS/DMA model after the crisis is 25.12% while before the crisis was only 10.55%. This implies that the investor would be better off holding this portfolio after the subprime crisis.

In sum, portfolios generally perform better after January 2008. Nevertheless, portfolios based on Bayesian models can produce good results in both sub-sample periods. This may indicate that these models are more able to accommodate market instability and therefore are more robust techniques.

---

that there were no significant differences in the predictability of returns between recession and expansion periods.

Table 2.7: Statistics of out-of-sample portfolio returns for different risk aversion coefficients, before and after the subprime crisis (January 2008)

Forecasting schemes			Mean $\gamma = 3$	Mean $\gamma = 5$	Mean $\gamma = 10$	Standard deviation $\gamma = 3$	Standard deviation $\gamma = 5$	Standard deviation $\gamma = 10$	Skewness $\gamma = 3$	Skewness $\gamma = 5$	Skewness $\gamma = 10$	Kurtosis $\gamma = 3$	Kurtosis $\gamma = 5$	Kurtosis $\gamma = 10$
Mean		Pre	2.49	2.29	1.73	32.86	10.95	11.49	-2.51	-3.41	-4.74	12.62	14.80	21.09
		Post	3.01	2.81	2.15	14.70	13.86	12.49	-0.68	-1.88	-3.39	15.64	16.45	21.55
WMSFE	**10%**	Pre	2.46	2.22	1.77	13.86	12.96	11.49	-4.22	-4.70	-5.45	18.45	19.61	24.37
		Post	4.30	3.85	2.94	16.61	14.28	11.49	4.27	1.51	0.62	32.12	17.51	22.62
	20%	Pre	2.84	2.70	2.18	11.49	11.49	10.39	-2.06	-2.71	-4.09	10.51	12.69	18.69
		Post	4.16	3.98	3.08	16.61	15.87	12.49	4.72	5.67	3.29	37.92	42.98	32.62
	30%	Pre	2.63	2.63	2.08	10.95	10.95	10.39	-1.99	-2.73	-4.16	11.06	13.02	19.23
		Post	3.98	3.91	3.08	16.61	15.87	12.96	4.13	5.18	4.16	38.63	43.80	40.13
	40%	Pre	2.60	2.60	2.08	10.39	10.95	9.80	-1.98	-2.37	-3.81	11.86	11.99	17.89
		Post	3.88	3.81	3.01	16.61	16.25	12.96	4.28	5.05	3.92	39.63	44.29	39.70
	50%	Pre	2.70	2.67	2.08	10.39	10.39	9.80	-1.84	-1.99	-3.51	12.76	12.14	17.17
		Post	3.78	3.74	2.91	16.61	16.25	12.96	4.42	4.99	3.52	40.42	44.85	38.44
DMS and DMA	**Small**	Pre	2.84	2.60	2.53	9.80	9.80	9.17	-1.75	-2.16	-2.18	8.22	8.50	8.50
		Post	6.62	6.86	7.10	16.25	15.49	13.86	-2.23	-1.40	0.71	28.42	24.68	18.97
	Medium	Pre	2.22	2.32	2.81	8.49	8.49	7.75	-1.84	-1.901	-1.50	10.07	10.65	9.33
		Post	7.41	2.22	6.72	15.87	8.49	13.42	6.44	-1.76	2.92	39.02	9.70	19.07
	Large	Pre	2.36	2.53	1.97	17.66	16.61	10.95	-3.47	-3.63	-3.27	13.08	14.43	12.23
		Post	2.63	3.01	2.67	21.35	19.60	12.96	-2.39	-1.19	-2.72	36.81	41.83	27.25
DMS	**Full**	Pre	2.36	2.53	2.04	17.66	16.61	11.49	-3.47	-3.62	-3.34	13.08	14.30	12.44
		Post	4.47	4.64	4.19	21.07	19.90	14.28	-1.71	-1.01	-0.94	38.01	40.62	27.43
DMA	**Full**	Pre	3.33	3.36	2.81	15.49	14.70	12.49	-1.93	-1.55	-1.74	8.81	8.05	7.90
		Post	5.82	6.10	6.44	16.97	16.61	16.25	5.53	5.86	6.23	32.59	34.47	35.74

Notes: This table presents a similar analysis to the one reported in Table 2.4, considering a partition of the data in January 2008, the beginning of the subprime crisis. "Pre" refers to the period from January 2004 to December 2007, and "Post" refers to the period from January 2008 to December 2017. This table reports the annualized mean, annualized standard deviation, skewness, and kurtosis of the portfolio returns based on the Mean, WMSFE (10%, 20%, 30%, 40%, and 50% best models), DMS and DMA of the Small, Medium, Large and Full TVP-VAR(1). The portfolios were selected considering three risk aversion coefficients ($\gamma = 3$, $\gamma = 5$ and $\gamma = 10$). The mean and standard deviation are in percentage.

Table 2.8: Portfolio performance for different risk aversion degrees before and after the subprime crisis (January 2008)

Forecasting schemes			CER			SR			SOR		
			$\gamma = 3$	$\gamma = 5$	$\gamma = 10$	$\gamma = 3$	$\gamma = 5$	$\gamma = 10$	$\gamma = 3$	$\gamma = 5$	$\gamma = 10$
Mean		Pre	10.97	8.41	1.48	0.83	0.71	0.52	1.29	1.04	0.70
		Post	7.84	5.25	-1.78	0.72	0.71	0.59	1.16	1.10	0.84
WMSFE	**10%**	Pre	9.24	8.82	7.84	0.61	0.59	0.54	0.85	0.82	0.73
		Post	11.98	11.13	8.85	0.90	0.93	0.88	1.80	1.75	1.55
	20%	Pre	11.92	11.35	9.65	0.86	0.82	0.71	1.36	1.25	1.02
		Post	11.44	11.19	9.10	0.87	0.87	0.85	1.72	1.79	1.62
	30%	Pre	11.23	11.28	9.39	0.82	0.82	0.70	1.29	1.26	1.00
		Post	10.65	10.70	8.96	0.82	0.84	0.82	1.57	1.67	1.58
	40%	Pre	11.20	9.90	4.95	0.83	0.83	0.70	1.32	1.30	1.02
		Post	10.38	7.92	2.85	0.81	0.82	0.82	1.55	1.62	1.53
	50%	Pre	11.75	10.33	5.37	0.90	0.88	0.73	1.47	1.40	1.07
		Post	9.90	7.58	2.53	0.79	0.81	0.79	1.50	1.57	1.47
DMS and DMA	**Small**	Pre	12.57	10.71	8.37	1.01	0.95	0.97	1.67	1.50	1.55
		Post	20.89	19.25	16.26	1.42	1.53	1.76	2.52	2.84	3.83
	Medium	Pre	10.55	10.32	10.67	0.91	0.99	1.27	1.46	1.63	2.2859
		Post	25.12	10.86	16.31	1.61	0.88	1.71	4.82	1.41	4.52
	Large	Pre	6.93	4.84	3.57	0.46	0.53	0.62	0.64	0.74	0.88
		Post	1.89	-0.33	-0.91	0.43	0.53	0.71	0.61	0.80	1.07
DMS	**Full**	Pre	6.92	4.77	3.60	0.46	0.53	0.63	0.63	0.73	0.90
		Post	8.84	5.22	2.68	0.73	0.81	1.00	1.14	1.32	1.72
DMA	**Full**	Pre	12.08	10.15	5.29	0.75	0.79	0.78	1.17	1.27	1.22
		Post	17.91	16.44	12.89	1.18	1.26	1.38	2.93	3.25	3.84

Notes: This table presents a similar analysis to the one reported in Table 2.5, considering a partition of the data in January 2008, the beginning of the subprime crisis. "Pre" refers to the period from January 2004 to December 2007, and "Post" refers to the period from January 2008 to December 2017. This table reports the annualized Certainty Equivalent Return (CER), Sharpe Ratio (SR), and Sortino ratio (SOR), in percentage, of the portfolio returns based on the Mean, WMSFE (10%, 20%, 30%, 40%, and 50% best models), DMS and DMA of the Small, Medium, Large and Full TVP-VAR(1). The portfolios were selected considering three risk aversion coefficients ($\gamma = 3$, $\gamma = 5$ and $\gamma = 10$).

2.7. Conclusion

This chapter aims to study the predictability of returns using VARs. It uses several schemes, such as Dynamic Model Selection (DMS) and Dynamic Model Averaging (DMA), to provide forecasts that are then used in portfolio selection problems. These approaches embed several useful features into a predictive system, namely, model and parameter uncertainty, time-varying parameters for modelling and forecasting multiple asset returns, combinations of predictors, and time-varying covariance matrices. For the period from 1976 to 2017, we have found large statistical and economic benefits from using these ensembles of features in predicting excess returns of stocks, bonds, and REITs. Notably, Bayesian DMS and DMA combinations deliver out-of-sample gains on a consistent basis over time, unlike simple predictive regressions. We conclude that a power utility investor benefits from holding a Bayesian portfolio rather than pursuing investment strategies based on other, simpler, forecasting schemes.

We have also compared the performance of different portfolios before and after January 2008, the beginning of the subprime crisis. We examined whether portfolio returns present substantial performance differences in both periods and the extent to which the implemented models were suitable under increased market instability. DMS and DMA portfolios have good results before and after the subprime crisis. Arguably, this suggests that Bayesian portfolios accommodate market instability in their specifications and can be seen as more robust methods.

Our dynamic optimizing strategies do not take transaction costs into account. These costs are likely to be relevant in the framework of dynamic portfolio selection and may reduce the performance of the dynamic models.

Chapter 3 - International Portfolio Selection with Machine-Learning and a Multivariate Asymmetric DCC model

3.1. Introduction

Since Markowitz (1952), it is unanimously recognized that an investor can significantly reduce her exposure to risk by properly diversifying the investment portfolio. While conservative investors previously depended mainly on low-volatility assets to reduce the risk of their portfolios, Markowitz showed that they might achieve identical volatility (or lower), together with higher (or the same) returns, through the combination of risky assets that have low or negative correlations. The initial Markowitz framework remains the basic research model for portfolio selection and diversification strategies, despite concerns regarding parameter sensitivity (Best and Grauer, 1991) and the numerous robust or more realistic extensions that have been proposed in the literature (Britten-Jones, 1999; Ang and Bekaert, 1999).

Higher levels of diversification may even be achieved by allocating the investment into assets of several international markets. The idea of international diversification is that investment opportunities within countries are often more correlated to each other than across countries, hence including assets from non-domestic countries may produce higher diversification benefits.

In its most basic formalization, the Markowitz portfolio selection model requires estimates of the expected return vector and the covariance matrix of all assets in the investment universe. Green et al. (2013) listed more than 300 papers that discuss the estimation problems relative to these inputs, and many more have likely been published since then. However, these issues continue to be studied, with methodologies that make use of recent technological and computational advances.

The expected return vector estimation resulting from traditional methods, such as the historical average or momentum, may incorporate a significant measurement error.

This is most likely caused by structural breaks and nonlinearities, which are not captured by those methods. The conditional covariance matrix, $\mathbf{H}_t = Cov(\mathbf{r}_t|\mathcal{F}_{t-1})$, where $\mathbf{r}_t$ is the vector of asset return at time t, subject to information up to time $t-1$, is the key object of interest in a multivariate environment. However, when the number of assets is fairly big, modelling $\mathbf{H}_t$ presents the problem usually known as the curse of dimensionality. The Dynamic Conditional Correlation (DCC) model of Engle et al. (2019) deals with this problem by combining the composite likelihood and a nonlinear shrinkage method. The DCC can deal with large dimensions of the asset space, which are frequently encountered in practical problems of portfolio selection and asset allocation. However, this model cannot capture some important features of the data, for instance, the leverage or asymmetric effects (Black, 1976). The common asymmetric effect in finance occurs when an unanticipated asset price decline (bad news) increases the volatility more than an unexpected rise (good news) of comparable magnitude. If this effect exists, assuming a symmetric function for the conditional variance may not be appropriate.

This study aims to combine new estimation schemes for the vector of expected returns and the covariance matrix in a large dimension problem. Firstly, we build on Machine Learning techniques to jointly forecast daily stock and bond excess returns. The key aspect of this approach is its ability to integrate several important data features into a flexible, yet computationally simple method, well-suited for dealing with large asset spaces. Secondly, we introduce asymmetric effects in the innovations of the Dynamic Conditional Correlation (DCC) model by Engle et al. (2019). The estimates of the return vector and covariance matrix obtained using these methodologies are then used as inputs in the Markowitz portfolio selection problem. Therefore, our first contribution is to empirically compare the performance of the portfolios using those more recent methodologies and the performance of portfolios derived from simple benchmark portfolios. The second contribution is to examine the potential contribution of different countries to diversification using several measures of portfolio performance. We rely on the proposed model to study the benefits of holding a diversified international portfolio of 77 stocks and bonds indexes.

Our results reveal that the proposed models produce sizable gains in terms of portfolio performances. We find that a power utility investor with moderate risk-aversion that uses the proposed Random Forest - Asymmetric Dynamic Conditional Correlation model (RF-ADCC) obtains an average return and a Sharpe ratio 10% and 48% higher than the ones obtained in a strategy based on a DCC model, respectively. This suggests that Random Forests provide an effective way to estimate the expected return. We also show that international diversification is economically and statistically beneficial for investors in South America, Europe, the Middle East, Asia, and Oceania, from 2012 to 2020. However, these benefits are time-varying.

The remaining of this study consists of five sections. Section 3.2 shows a brief literature review. Section 3.3 presents the data and provides some descriptive statistics. Section 3.4 outlines the basic theoretical concepts, presents the specifications of the models, the measures used to assess the performance of the portfolios out-of-sample, and the research setup. Section 3.5 shows the results, and Section 3.6 concludes the chapter.

3.2. Literature Review

Numerous studies have examined international diversification using different methodologies and datasets. Some studies are Grubel (1968), Levy and Sarnat (1970), Solnik (1974), Heston and Rouwenhorst (1994), Longin and Solnik (2001). More recent works are D'Ecclesia et al. (2006), Yang et al. (2006), Driessen et al. (2007), Bekaert et al. (2009), Christoffersen et al. (2014), Viceira and Wang (2018), and Kellner (2019).

Early research in the 1970s demonstrated that correlations between different national stock markets were small and that national markets primarily react to domestic economic fundamentals (Levy and Sarnat, 1970; and Solnik, 1974). In the 1980s, part of the literature revealed the existence of large and statistically significant interdependences between national economies. Cooper (1985) confirmed that economies were becoming more integrated worldwide. In the 1990s and 2000s, extreme global events, such as the 1987 market crash, the invasion of Kuwait in 1990, and the 9/11 attack in 2001, caused interconnected reactions in financial markets worldwide, resulting in a reduction of the effectiveness of international diversification. In the absence of extreme events, however, national financial markets are mostly prone to domestic fundamentals, therefore investing in foreign countries tends to enhance the advantages of diversification. For instance, Adjaouté and Danthine (2004) reported low portfolio diversification opportunities in Europe during the euro convergence period (1988 to 2001), which have significantly improved afterward. Yang et al. (2006), Driessen et al. (2007), and Bekaert et al. (2009)

reported decreasing diversification benefits in most of the developed markets in the 1990s and early 2000s.

So, despite the consensus on the benefits of having an internationally diversified long-term portfolio, many authors reported that in some periods these benefits decreased. Possible explanations for these phenomena are the evidence of increasing return correlation between international economies, the increasing market integration, and the fact that many investors still prefer to allocate their funds to local assets instead of broadening the geographical scope of their investment (see, for example, Bekaert et al., 2017; Viceira and Wang, 2018).

Since the 2000s, most studies have analysed separately developed or emerging markets (see, for instance, Yang et al., 2006; Driessen et al., 2007; Bekaert et al., 2009; Christoffersen et al., 2014; and Kellner et al., 2019), mainly concluding that emerging markets seem to be of higher interest due to lower interdependencies and higher diversification benefits. Additionally, the increasing level of integration among developed markets suggests that they are more susceptible to common factors, increasing their comonotonic behaviour and decreasing the diversification advantages. Christoffersen et al. (2014) showed that the correlations between emerging and developed markets increased significantly from 1989 to 2012. Kellner et al. (2019) analysed international investments between 2002 and 2016, highlighting that Chile, Jordan, Malaysia, Singapore, and Switzerland markets had high diversification potential, while Argentina, Brazil, and South Korea exhibited high individual risk levels and high market connections.

In more recent years, several studies have developed approaches to investigate the dynamics of assets' covariances. A model that typically performs well in identifying covariances fluctuations and volatility spillovers between financial time series is the multivariate Generalized Autoregressive Conditional Heteroscedastic (GARCH) model (Chang et al., 2013). When the asset space only considers a moderate number of assets (usually less than 25), this is one of the most used approaches. However, for a larger number of assets, the estimation of the conditional covariance becomes computationally difficult and faces the curse of dimensionality (Solnik and Roulet, 2000).

The fundamental cause of estimation problems, according to Pakel et al. (2020), is that the most commonly used approaches rely on a simple-to-implement estimator of the unconditional covariance/correlation matrix. This implies estimating $O(L^2)$ parameters using $O(LT)$ data points, where L is the number of assets and T is the length of the time series. Unless T is considerably larger than L, it will be subject to significant estimation errors. Consequently, the likelihood function would be contaminated, resulting in biased estimates. Furthermore, the Gaussian (quasi-) likelihood estimation procedure requires the inversion of the $(L \times L)$ conditional covariance matrix and since the estimation is based on numerical optimization, many matrix inversions are needed. Therefore, the multivariate GARCH may have an unsatisfactory performance for an investment universe with numerous assets.

Recently, several studies have proposed procedures to deal with the curse of dimensionality in the context of conditional covariance models. Engle et al. (2019) suggested a simple and efficient combination procedure based on composite likelihood and nonlinear shrinkage to estimate Dynamic Conditional Correlation (DCC) models. The former accounts for the complex dynamic correlation parameters (time series), while the latter accounts for the correlation targeting matrix (cross-section). This procedure overcomes large dimension problems, which are commonly encountered in practice.

Complex data interactions, nonlinearities and strongly correlated predictors usually characterize financial time series prediction problems (Gu et al., 2018), rendering traditional financial time series estimation methods inefficient. With the increasing computational ability, the focus has shifted to the use of Machine Learning techniques, such as, for instance, Random Forests (RF) and Artificial Neural Networks (NN). These methods have been successfully used in a wide variety of applications and enjoy considerable popularity in several disciplines (see, for instance, Cutler et al., 2012), and have also been successfully employed to forecast a wide range of financial time series (Cutler et al., 2012; Heaton et al., 2017; Krauss et al., 2017; Gu et al., 2018, Lohrmann et al., 2019). According to Cutler et al. (2012), these techniques are well-suited for challenging prediction problems, such as large predictive datasets and highly correlated predictors, because they increase the degrees of freedom and condense redundant variations among predictors. For instance, Krauss et al. (2017) employed a binary classification method based on RF and NN to predict one-day-ahead S&P 500 excess returns. They found sustainable profit opportunities in the short run. Gu et al. (2018) conducted a comparative study of approaches in the Machine Learning repertoire applied to return series. They concluded that these approaches, especially Random Forests and Artificial Neural Networks, are better than conventional forecasting methods. They argued that the predictive gains came from these techniques accounting for nonlinear predictor interactions that other techniques neglect.

3.3. Data Description and Preliminary Analysis

This study uses daily closing values, in US dollars, of 44 stocks and 33 bonds total return indexes for 16 countries of the European Union (EU) and 28 countries from different geographic regions around the world, from 01/08/2001 to 14/09/2020 (a total of 4989 days),[5] obtained from the "Thompson Reuters DataStream" database.

The EU countries are Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Netherlands, Poland, Portugal, Spain, and Sweden.

The other 28 countries are Canada, Mexico, and the US (North America), Argentina, Brazil and Chile (South America), Egypt, Israel, Pakistan, and Turkey (Middle East), China, India, Indonesia, Korea, Singapore, Japan, Hong Kong, Malaysia, Philippines, Taiwan and Thailand (Asia), Australia and New Zealand (Oceania), and Norway, Russia, Switzerland, and UK (Europe), and South Africa (Africa).

These countries were also grouped into emerging and developed markets according to the Market classification MSCI (https://www.msci.com/our-solutions/indexes/market-classification):

- Emerging Markets: Argentina, Brazil, Chile, China, Egypt, Hungary, India, Indonesia, Israel, Korea, Malaysia, Mexico, Pakistan, Philippines, Poland, Russia, South Africa, Singapore, Taiwan, Thailand, Turkey.

- Developed markets: Australia, Austria, Belgium, Canada, Czech Republic, Denmark, Finland, France, Germany, Greece, Hong Kong, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, the UK, and the US.

For the 44 stocks indexes, we used the Refinitiv Total Return Index, and for the 33 bonds, we used the Thomson Reuters 10 Year Government Benchmark, except for India, Indonesia, Korea, Malaysia, Philippines, Singapore, Taiwan, and Thailand, for which we use the iBoxx Government 10 year.

[5] Missing observations are replaced with the immediately preceding value, so that the return in that day is equal to zero.

Table 3.1: Descriptive statistics of excess returns

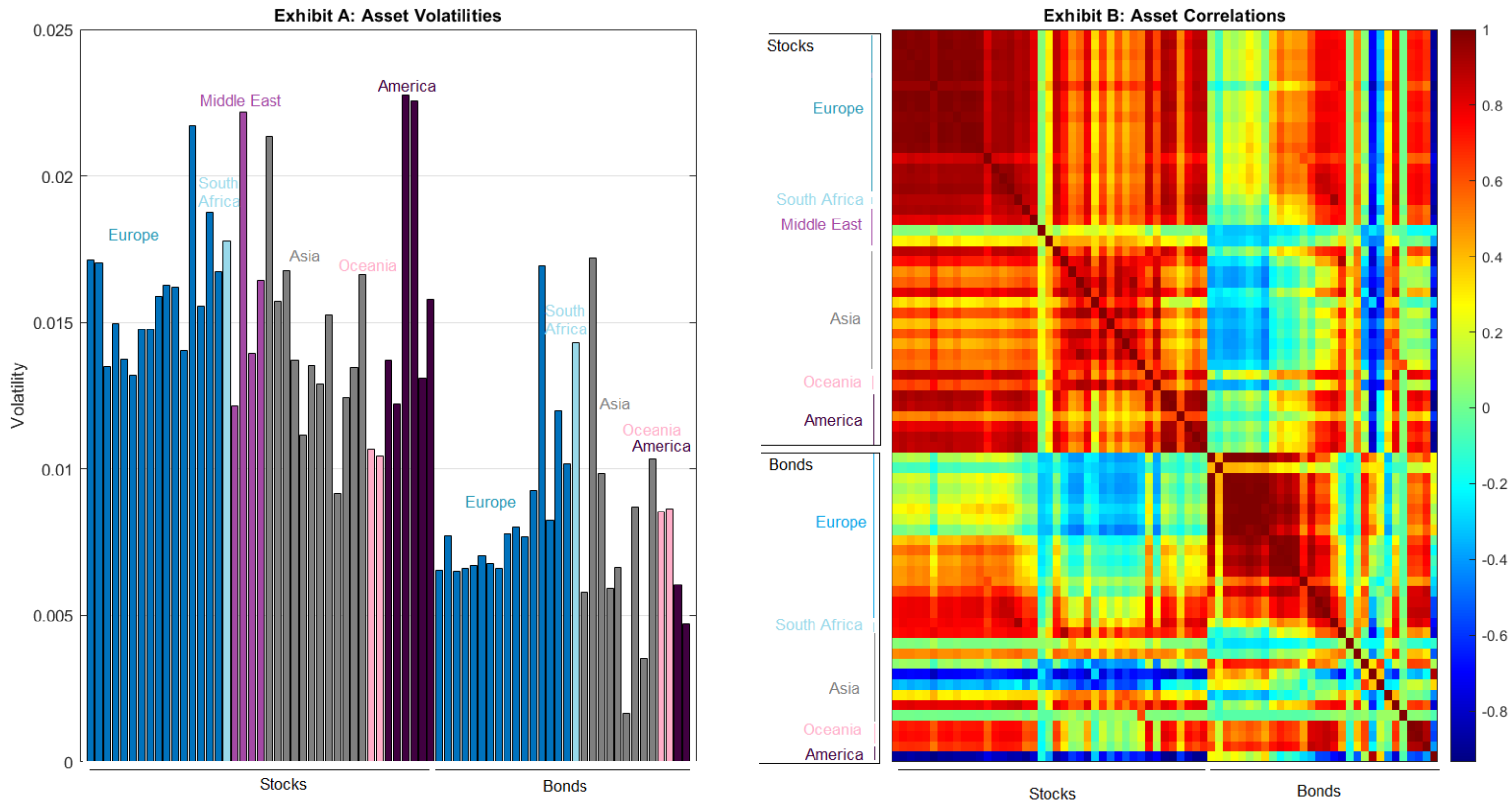
Panel A: European Union countries (EU)																
Country	**Mean**		**Std**		**Min**		**Max**		**Skew**		**Kurt**		$\boldsymbol{\rho(1)}$		**Q**	
	Stocks	***Bonds***	***Stocks***	***Bonds***	***Stocks***	***Bonds***	***Stocks***	***Bonds***	***Stocks***	***Bonds***	***Stocks***	***Bonds***	***Stocks***	***Bonds***	***Stocks***	***Bonds***
Austria	0.01	0.02	1.32	0.67	-13.51	-3.01	10.89	3.55	-0.65	-0.03	13.73	4.77	0.08	0.00	0.00	0.72
Belgium	0.02	0.03	1.38	0.70	-14.89	-3.95	9.98	4.25	-0.60	-0.01	11.28	5.24	0.06	0.02	0.00	0.37
Czech Rep	0.04	0.03	1.55	0.83	-17.00	-5.74	18.31	4.85	-0.56	-0.22	19.04	6.77	0.06	0.03	0.00	0.05
Denmark	0.04	0.02	1.35	0.65	-13.52	-3.26	10.94	3.50	-0.47	0.01	11.39	4.79	0.03	-0.01	0.00	0.86
Finland	0.02	0.02	1.71	0.65	-11.23	-3.23	10.32	3.62	-0.21	-0.01	7.70	4.91	0.01	-0.02	0.00	0.72
France	0.02	0.02	1.48	0.68	-13.75	-2.92	11.18	3.55	-0.28	-0.02	10.50	4.74	0.00	-0.01	0.00	0.53
Germany	0.02	0.02	1.48	0.66	-13.85	-2.91	13.78	3.58	-0.16	0.02	11.56	4.89	0.02	-0.02	0.00	0.86
Greece	-0.03	0.03	2.17	1.69	-23.21	-27.93	16.22	34.89	-0.55	1.12	12.80	91.46	0.08	0.16	0.00	0.00
Hungary	0.02	0.02	1.87	1.20	-21.46	-17.54	18.12	8.56	-0.29	-1.14	13.55	18.82	0.08	0.08	0.00	0.00
Italy	0.01	0.03	1.63	0.80	-18.81	-6.55	12.05	5.39	-0.59	-0.12	13.13	7.26	-0.01	0.05	0.00	0.00
Ireland	0.02	0.02	1.59	0.78	-16.63	-5.50	12.00	8.41	-0.70	0.12	12.43	10.95	0.04	0.09	0.00	0.00
Netherlands	0.02	0.02	1.49	0.66	-19.70	-2.89	20.32	3.61	-0.27	-0.01	23.93	4.74	-0.05	-0.01	0.00	0.79
Portugal	0.00	0.03	1.40	0.92	-13.00	-12.12	11.55	11.42	-0.33	-0.40	10.72	20.74	0.07	0.12	0.00	0.00
Poland	0.03	0.03	1.67	1.02	-15.87	-6.33	12.08	8.89	-0.53	-0.21	9.79	7.44	0.08	0.06	0.00	0.00
Spain	0.01	0.03	1.62	0.77	-15.80	-3.87	13.96	5.67	-0.35	0.06	11.91	5.98	0.02	0.08	0.00	0.00
Sweden	0.03	0.02	1.70	0.77	-14.02	-4.58	14.40	5.63	-0.21	0.01	9.51	6.08	0.01	-0.04	0.00	0.07
Mean	**0.02**	**0.02**	**1.59**	**0.84**	**-16.02**	**-7.02**	**13.51**	**7.46**	**-0.42**	**-0.05**	**12.69**	**13.10**	**0.04**	**0.04**	**0.00**	**0.31**
Median	**0.02**	**0.02**	**1.57**	**0.77**	**-15.35**	**-4.26**	**12.06**	**5.12**	**-0.41**	**-0.01**	**11.73**	**6.03**	**0.04**	**0.03**	**0.00**	**0.06**

Table 3.1: Descriptive statistics of excess returns (*continued*)

Country	Mean		Std		Min		Max		Skew		Kurt		$\rho(1)$		Q	
Panel B: Non-EU countries																
	Stocks	*Bonds*	*Stocks*	*Bonds*	*Stocks*	*Bonds*	*Stocks*	*Bonds*	*Stocks*	*Bonds*	*Stocks*	*Bonds*	*Stocks*	*Bonds*	*Stocks*	*Bonds*
Australia	0.03	0.03	1.04	0.86	-10.12	-7.21	6.69	7.12	-0.70	-0.33	11.81	10.26	-0.06	-0.10	0.00	0.00
Canada	0.03	0.02	1.37	0.60	-14.28	-3.56	11.48	3.57	-1.03	-0.05	17.01	5.45	-0.02	-0.05	0.00	0.00
Hong Kong	0.03	0.01	1.29	0.16	-12.62	-1.19	11.44	1.02	-0.31	-0.16	10.87	7.37	0.01	0.09	0.00	0.00
Israel	0.03	-	1.21	-	-8.21	-	9.62	-	-0.39	-	7.77	-	0.02	-	0.03	-
Japan	0.01	0.01	1.35	0.66	-11.94	-6.61	12.42	4.33	-0.26	0.01	9.44	8.21	-0.10	0.01	0.00	0.10
New Zealand	0.04	0.03	1.07	0.85	-7.68	-6.84	8.49	6.47	-0.60	-0.35	8.74	7.14	0.10	-0.01	0.00	0.00
Norway	0.00	0.00	0.02	0.01	-0.14	-0.08	0.12	0.05	-0.76	-0.37	10.43	8.26	0.01	-0.01	0.00	0.44
Singapore	0.02	0.02	1.11	0.59	-9.34	-4.10	7.79	5.48	-0.41	-0.11	10.36	7.39	0.04	0.09	0.00	0.00
Switzerland	0.03	0.02	1.19	0.75	-10.27	-8.85	10.50	17.98	-0.11	2.46	10.28	74.31	-0.03	0.01	0.00	0.02
UK	0.01	0.02	1.33	0.65	-13.32	-5.76	10.51	4.07	-0.52	-0.41	13.82	7.60	0.01	0.03	0.00	0.08
US	0.02	0.02	1.22	0.47	-13.09	-2.83	10.84	4.05	-0.51	-0.05	15.59	6.18	-0.11	-0.01	0.00	0.13
Argentina	0.00	-	2.28	-	-59.32	-	17.47	-	-5.86	-	138.61	-	0.05	-	0.00	-
Brazil	0.03	-	2.26	-	-17.48	-	19.88	-	-0.48	-	11.85	-	-0.04	-	0.00	-
Chile	0.02	-	1.31	-	-16.44	-	13.53	-	-0.68	-	19.09	-	0.11	-	0.00	-
China	0.01	-	1.53	-	-9.67	-	9.53	-	0.02	-	7.68	-	-0.01	-	0.00	-
Egypt	0.02	-	1.64	-	-46.06	-	9.30	-	-5.02	-	132.28	-	0.13	-	0.00	-
India	0.04	0.02	1.57	0.58	-14.78	-5.90	19.17	4.34	-0.49	-0.41	14.33	13.93	0.08	0.08	0.00	0.00
Indonesia	0.03	0.02	1.68	1.72	-16.31	-22.32	13.93	11.78	-0.69	-0.18	12.53	24.14	0.12	-0.12	0.00	0.00
Korea	0.03	0.02	1.37	0.98	-13.18	-28.10	11.44	11.95	-0.58	-4.76	10.95	147.71	0.01	-0.06	0.03	0.00
Mexico	0.02	-	1.58	-	-12.03	-	17.59	-	-0.31	-	12.74	-	0.06	-	0.00	-
Malaysia	0.03	-	0.92	-	-10.60	-	6.28	-	-0.74	-	12.39	-	0.12	-	0.00	-
Pakistan	0.04	-	1.39	-	-8.65	-	9.97	-	-0.36	-	7.15	-	0.11	-	0.00	-
Philippines	0.03	0.05	1.24	0.87	-13.20	-23.93	8.21	7.82	-0.98	-4.58	12.84	131.61	0.09	0.01	0.00	0.00
Russia	0.04	-	2.13	-	-20.91	-	23.70	-	-0.59	-	15.99	-	0.04	-	0.00	-
South Africa	0.03	0.02	1.78	1.43	-21.39	-16.21	10.31	16.63	-77.72	-0.63	11.29	15.64	0.04	0.05	0.00	0.00
Taiwan	0.03	0.02	1.35	0.35	-7.06	-2.12	8.12	2.72	-26.09	-0.01	6.60	6.65	0.05	0.00	0.00	0.00
Thailand	0.05	0.03	1.66	1.03	-16.93	-13.02	13.95	12.74	-42.17	0.09	18.19	66.45	-0.10	-0.27	0.00	0.00
Turkey	0.02	-	2.22	-	-17.28	-	14.79	-	-0.53	-	9.20	-	0.04	-	0.00	-
Mean	**0.03**	**0.02**	**1.48**	**0.79**	**-16.00**	**-9.91**	**12.11**	**7.63**	**-6.23**	**-0.59**	**21.09**	**33.75**	**0.03**	**-0.02**	**0.00**	**0.02**
Median	**0.03**	**0.02**	**1.37**	**0.71**	**-13.20**	**-6.72**	**11.44**	**5.98**	**-0.58**	**-0.17**	**12.39**	**11.07**	**0.04**	**0.00**	**0.00**	**0.00**

Notes: This table reports the summary statistics of daily excess log returns during the period from 01/08/2001 to 14/09/2020 (full sample), considering a risk-free rate proxied by the US 3 Month Treasury Bill obtained from Thomson Reuters. Panel A and Panel B show these statistics for 16 countries of the European Union (EU) and for 28 non-EU countries from different regions worldwide, respectively. The statistics are the mean, standard deviation (Std), maximum (Max) and minimum (Min) values, skewness (Skew), kurtosis (Kurt), first order autocorrelation ($\rho(1)$), and the p-value Ljung-Box Q-test for the autocorrelations up to lag 20, (Q). The rows *Mean* and *Median* refer to the mean and median of the statistics across countries, respectively. The values of Mean, Std, Min and Max are in percentage.

Figure 3.1: Asset volatilities and correlations



Notes: This figure shows the volatility (Exhibit A) and the correlations (Exhibit B) of daily excess log returns, for both stocks and bonds indexes, during the period from 01/08/2001 to 14/09/2020 (full sample), grouped according to their geographical region. Europe refers to European countries independently of belonging to the EU or not. America considers both North and South America counties. On Exhibit B, on the right side, there is the correlation scales associated to each colour. The dark red corresponds to a correlation of 1.

Table 3.1 contains the descriptive statistics of the stocks and bonds indexes daily log excess returns for EU markets (Panel A) and non-EU markets (Panel B) for the period from 01/08/2001 to 14/09/2020 (full sample). While there are significant cross-country variations, Table 3.1 shows that the mean and median returns in EU countries were the same for stocks and bonds, while they were higher for stocks than bonds in non-EU countries (plus 0.01%). Stock indexes present higher variability, measured by standard deviation and range (maximum minus minimum), than bond indexes in EU and non-EU countries. Also, on average, stocks are more negatively skewed than bonds, while bonds present higher excess kurtosis, especially for non-EU countries. First-order autocorrelations of EU countries and non-EU countries are most times positive (75% and approximately 65% of the series, respectively). The Ljung-Box test rejected the null of autocorrelations jointly equal to zero for all EU stock indexes, while 9 out of 16 EU bond indexes are not significant autocorrelated up to lag 20. For non-EU countries, the null is rejected for all stocks at a 5% significance level, while 4 out of 17 non-EU bond indexes show no significant autocorrelations at the 5% significance level.

Figure 3.1 shows the full sample volatilities and correlations. As documented before, in general terms, the volatility of stock indexes is higher than the volatility of bond returns, although the variation of volatility across countries in the same geographic area are visible for stocks and bonds. The correlations map makes it possible to see a higher degree of correlation between stocks across all regions than bonds, except for some countries in the Middle East. European stock and bond indexes are more correlated than stock and bond indexes of non-European countries, which is not surprising as most of these countries belong to the Euro zone.

3.4. Methodology

This section presents the basic theoretical concepts and the specifications of the models used in the empirical application. Subsections 3.4.1. presents the Machine Learning techniques used to forecast the expected returns vector. Subsection 3.4.2 explains the methods and procedures used to estimate the multivariate Asymmetric DCC model. Subsection 3.4.3 presents the research setup. Finally, Subsection 3.4.4 presents the portfolio selection problem, the basic benchmark portfolios and the metrics used to assess the performance out-of-sample.

3.4.1. Machine Learning Methods: Random Forests and Neural Networks

This study employs Random Forests (RF) and Artificial Neural Networks (NN) to obtain the vector of expected excess return one-day ahead.

A Random Forest is an ensemble learning technique with good performances in terms of generalization and overfitting avoidance. Random Forests were firstly proposed by Brieman (2001). In our study, we focus on Random Regression Forests, a combination of randomly trained regression trees that collectively are used to output a prediction. When “growing” each individual tree, we randomly choose a subset of the training data for training the individual regression tree, choose a random subset of features (predictors) at each node, and only consider splitting on these features. The prediction of the RF is the average of the predictions of all individual regression trees in the model. Because the baseline procedure draws different bootstrap data samples, individual predictions are uncorrelated.

Arguably, Artificial Neural Networks (NN) are one of the most powerful modelling devices in Machine Learning. NNs have theoretical underpinnings as universal approximators for any smooth predictive association (Gu et al., 2018) and may incorporate a wide range of models. One of the simplest and broadly used models in predictive problems is the “feed-forward” network. These models consist of an input layer of raw predictors, one or more “hidden layers” that interact and nonlinearly transform the predictors, and an “output layer” that aggregates hidden layers into an ultimate outcome prediction. Equivalent to axons in the biological brain, layers of the network correspond to groups of “neurons”, with each layer linked by “synapses” that convey signals among neurons of different layers. The number of units in the input layer is equal to the dimension of the predictor space. Each predictor signal is amplified or attenuated based on a parameter vector that contains an intercept and one weight parameter per predictor. The output layer aggregate includes the weighted signals in the forecast.

3.4.2. The multivariate Asymmetric DCC Model

Let $\mathbf{r}_t = (r_{1,t}, \ldots, r_{N,t})'$ be the vector of excess log returns, such that $r_{i,t}$ is the excess log return of the *i-th* asset at time t, for $i = 1, \ldots, N$, and $t = 1, \ldots, T$. N is the number of assets under consideration to build the portfolio and T indicates the sample size. For simplicity, let us consider that $\mathbb{E}(\mathbf{r}_t|\mathcal{F}_{t-1}) = \mathbf{0}$, where $\mathcal{F}_{t-1}$ denotes the information available at time $t-1$. Let us consider that $\mathbf{H}_t = Cov(\mathbf{r}_t|\mathcal{F}_{t-1})$ is the conditional covariance matrix with elements $h_{i,k,t} = Cov(r_{i,t}, r_{k,t}|\mathcal{F}_{t-1})$. At time $t-1$, the investor is interested in forecasting $\mathbf{H}_t$ aiming to select a portfolio for period $(t-1, t]$.

The conditional covariance matrix is usually estimated in three stages:

(1) Univariate GARCH-type models are fitted to each series.

(2) The unconditional correlation matrix is estimated and an estimator of the correlation target matrix **C** is selected using the devolatilised residuals.

(3) The DCC model is estimated, and the conditional correlation and conditional covariance matrices are obtained.

In the first stage of the covariance matrix estimation, we estimate an Asymmetric GARCH(1,1) for each return series individually.[6] Engle et al. (2019) use a GARCH(1,1) model; however, this method cannot capture some important features of the data. For instance, it does not address the leverage or asymmetric effects reported by Black (1976). We introduce asymmetric effects in the innovations of the GARCH(1,1) model, as follows:

$$d_{i,t}^2 = \omega_i + [a_i + \gamma_i I(\varepsilon_{i,t-1} > 0)]{\varepsilon_{i,t-1}}^2 + b_i d_{i,t-1}^2, \tag{3.1}$$

where $I(\varepsilon_{i,t-1} > 0)$ is the indicator function that assumes the unity value if the lagged error term is positive. Hence, negative values of γ_i, the asymmetry coefficient, implies that positive shocks result in smaller increases in future volatility than negative shocks of the same magnitude, all else being equal.

The second stage consists in estimating the unconditional correlation matrix. Usually, this estimation is conducted using the sample correlation (and covariance)

[6] It is important to note that we introduce asymmetric effects in the innovations of the univariate GARCH model and not in the time-varying conditional correlations as it is the case of the Asymmetric Dynamic Conditional Correlations model of Cappiello et al. (2006).

matrix. However, if the number of parameters is large (it can have the same dimension of the series length), this method is prone to overfitting and thus is inefficient out-of-sample. Recent methodologies provide a solution to this problem, by correcting the in-sample bias of the correlation (or covariance) matrix. This correction uses the eigenvalues, namely by considering that small (large) sample eigenvalues are too small (large). Hence the shrinkage procedure consists in moving the small ones up and pull the large ones down. In our study, we use the nonlinear shrinkage method of Ledoit and Wolf (2012) and Ledoit and Wolf (2015).

Consider that $\mathbf{S} := [s_{i,t}]$ is the $N \times T$ matrix of devolatilized returns, $\boldsymbol{\Sigma}$ is its population covariance matrix, and $\hat{\mathbf{C}}$ is its sample counterpart, which can be decomposed as:

$$\hat{\mathbf{C}} = \frac{1}{T}\mathbf{SS}' = \sum_{i=1}^{N} \lambda_i \, u_i u_i^{'}, \tag{3.2}$$

where $(\lambda_1, \ldots, \lambda_N)$ is the set of eigenvalues, sorted in descending order, and $(u_1, \ldots, u_N)$ represents the corresponding eigenvectors. Following Engle et al. (2019), we estimate the population eigenvalues, $\tilde{\boldsymbol{\lambda}}(\tilde{\tau}) := (\tilde{\lambda}_1(\tilde{\tau}), \ldots, \tilde{\lambda}_N(\tilde{\tau}))$, using the method proposed by Ledoit and Wolf (2015). Finally, we replace the sample eigenvalues by the shrunk eigenvalues in Equation (3.2), and rebuild the shrinkage estimator of the covariance matrix:

$$\tilde{\mathbf{C}} := \sum_{i=1}^{N} \tilde{\lambda}_i(\tilde{\tau}) \, u_i u_i^{'}. \tag{3.3}$$

The last stage is the estimation of the DCC model. The dynamics of the correlation matrix over time are modelled through the correlation targeting, as in Engle et al. (2019). Several different formulations of the DCC model are available, but a popular specification is due to Engle (2002) and Engle et al. (2019). The model is related to the Constant Conditional Correlation (CCC) formulation, but the correlations are allowed to vary over time. The formulation is as follows:

$$\mathbf{Q}_t = (1 - a - b)\mathbf{C} + a\mathbf{s}_{t-1}\mathbf{s}'_{t-1} + b\mathbf{Q}_{t-1}, \tag{3.4}$$

where $\mathbf{s}_t = \mathbf{D}_t^{-1}\mathbf{r}_t$ is the vector of standardised residuals, such that $\mathbf{D}_t$ denotes the diagonal matrix whose *i-th* diagonal element is $d_{i,t}$, and $\mathbf{C} = Corr(\mathbf{r}_t) = Cov(\mathbf{s}_t)$. The parameters a and b are non-negative, with $a + b < 1$, hence $\mathbf{Q}_t$ is an $N \times N$ symmetric positive definite conditional pseudo-correlation matrix. From this formulation we obtain the conditional correlation matrix and the conditional covariance matrix as:

$$\mathbf{R}_t = diag\{\mathbf{Q}_t^*\}^{-1}\mathbf{Q}_t diag\{\mathbf{Q}_t^*\}^{-1}. \tag{3.5}$$

and

$$\mathbf{H}_t = \mathbf{D}_t\mathbf{R}_t\mathbf{D}_t, \tag{3.6}$$

where $diag(\cdot)$ denotes a resulting matrix comprising the main diagonal elements and $\mathbf{Q}_t^*$ is a matrix that takes the square roots of each element in $\mathbf{Q}_t$. This operation divides the covariances in $\mathbf{Q}_t$ by the product of the appropriate standard deviations in $\mathbf{Q}_t^*$ to create a matrix of correlations.

Estimating DCC models with many assets is challenging. One of the difficulties is related to the inversion of the conditional covariance matrix $\mathbf{H}_t$, for each $t = 1,\ldots,T$, to compute the (log)likelihood. One way to achieve this is by using the composite (log)likelihood method of Pakel et al. (2020). This method sums up the (log)likelihoods of each pair of assets instead of dealing with all assets at once. The authors proposed different methodologies, but the most used one is the "2MSCLE", which only required the computation of $N-1$ bivariate (log)likelihoods. This method can be used to estimate models with large dimensions, nevertheless providing consistent estimators of a and b, in Equation (3.4).

3.4.3. Portfolio Performance Evaluation

We assume that the investor has constant relative risk aversion (CRRA), such that her utility function is given by:

$$U(W_{t+1}) = \begin{cases} \dfrac{W_{t+1}^{1-\gamma}}{1-\gamma} & if\ \gamma > 1 \\ \log(W_{t+1}) & if\ \gamma = 1 \end{cases}, \tag{3.7}$$

where W_{t+1} denotes the investor's wealth at time $t+1$, and γ is the coefficient representing the investor's degree of risk aversion. At each point in time, the investor chooses her optimal portfolio, i.e., band trades N risky assets and a risk-free asset that maximizes her 1-period-ahead expected utility $\mathbb{E}_t[U(W_{t+1})]$. The portfolio is rebalanced each day. The optimal weights implied by model j are given by the solution of the following constrained quadratic maximization problem:

$$\arg\max_{\mathbf{x}_{j,t}} \mathbf{x}'_{j,t}\left(\widehat{\boldsymbol{\mu}}_{j,t+1|t} + \frac{1}{2}diag\widehat{\boldsymbol{\Sigma}}_{j,t+1|t}\right) - \frac{\gamma}{2}\mathbf{x}'_{j,t}\,\widehat{\boldsymbol{\Sigma}}_{j,t+1|t}\,\mathbf{x}_{j,t} \tag{3.8}$$

$$s.t.: (1)\ \mathbf{x}'_{j,t}\ \boldsymbol{\iota} = 1 \text{ and } (2)\ \mathbf{x}_{j,t} \geq \mathbf{0}$$

where $\boldsymbol{\iota}$ is a vector of ones with the same length as $\mathbf{r}_{t+1}$. The vector $\mathbf{x}_{j,t}$ denotes the portfolio weights, $\widehat{\boldsymbol{\mu}}_{j,t+1|t} = \mathbb{E}(\mathbf{r}_{t+1}|\mathcal{M}_j, \mathcal{F}_t)$ is the mean of the predictive density of the vector of risky asset $\mathbf{r}_{t+1}$, computed using model j on the information available at time t, and $\widehat{\boldsymbol{\Sigma}}_{j,t+1|t} = \widehat{Cov}(\mathbf{r}_{t+1}|\mathcal{M}_i, \mathcal{F}_t)$ is the forecasted covariance matrix of risky assets at time t based on the estimates given by model j, conditional on the information set at time t. Following most of the asset allocation literature, we rule out short-selling (i.e., negative portfolio weights).

If the excess returns of the N risky assets are log-normal, the portfolio log return implied by model j at time t is defined as (Campbell et al., 2003 and Fisher et al. 2020):

$$r_{p,j,t+1} = r_{f,t} + \mathbf{x}'_{j,t}\left(\mathbf{r}_{t+1} - r_{f,t}\boldsymbol{\iota}\right) + \frac{1}{2}\mathbf{x}'_{j,t} diag \widehat{\boldsymbol{\Sigma}}_{j,t+1|t} - \frac{1}{2}\mathbf{x}'_{j,t}\widehat{\boldsymbol{\Sigma}}_{j,t+1|t}\mathbf{x}_{j,t}. \qquad (3.9)$$

where $r_{f,t}$ represents the continuously compounded risk-free rate.

In our comparative analysis, we use three basic benchmark portfolios: The equally weighted portfolio (1/N), the portfolio that replicates the Refinitiv Europe Total Return for stocks, and that replicates the iBoxx Eurozone index for bonds, which is an equally weighted combination of both for stocks and bonds (we denote this benchmark portfolio as European Index) and the Minimum Variance Portfolio (MVP).

The excess return of any equally weighted portfolio is given by $r_{p,t+1} = \frac{1}{N}\boldsymbol{\iota}'\mathbf{r}_{t+1}$, where N is the number of assets in the portfolio. The MVP (allowing for short selling), is the solution of the following problem:

$$\min_{\mathbf{x}_t} \mathbf{x}'_t\, \widehat{\boldsymbol{\Sigma}}_{t+1|t}\, \mathbf{x}_t \qquad (3.10)$$
$$s.t.: \mathbf{x}'_t\, \boldsymbol{\iota} = 1$$

This problem has an analytical solution, given by:

$$\hat{\mathbf{x}}_t := \frac{\widehat{\boldsymbol{\Sigma}}^{-1}_{t+1|t}\boldsymbol{\iota}}{\boldsymbol{\iota}'\widehat{\boldsymbol{\Sigma}}^{-1}_{t+1|t}\boldsymbol{\iota}}. \qquad (3.11)$$

Where $\widehat{\boldsymbol{\Sigma}}_{t+1|t}$ is usually estimated using directly the historical covariance matrix. And the resulting portfolio excess returns are given by:

$$r_{p,t+1} = \hat{\mathbf{x}}'_t \mathbf{r}_{t+1}. \tag{3.12}$$

We now turn our attention to the metrics used to assess the performance of the portfolios. Besides the annualized mean and standard deviation of the excess returns, we compute the Certainty Equivalent Return (CER), which is the risk-free return that would make the investor indifferent between following a certain strategy or accepting this risk-free return. The annualized CER, considering daily data, can be expressed as follows:

$$CER_j = \begin{cases} \left[\left(\frac{1}{T - t_0} \sum_{t=t_0+1}^{T} \widehat{W}_{j,t}^{1-\gamma} \right)^{\frac{252}{1-\gamma}} - 1 \right] & if\ \gamma \neq 1 \\ \left[\frac{1}{T - t_0} \sum_{t=t_0+1}^{T} \log\left(\widehat{W}_{j,t}\right) \right] \times 252 & if\ \gamma = 1 \end{cases} \tag{3.13}$$

where $\widehat{W}_{j,t} = \exp\{r_{p,j,t}\}$ denotes the realized wealth at time t as implied by model j.

Another metric used is the annualized Sharpe ratio (SR),

$$SR_j = \frac{\sqrt{252}\hat{\mu}_{r_{p,j}}}{\hat{\sigma}_{r_{p,j}}}, \tag{3.14}$$

where $\hat{\mu}_{r_{p,j}}$ and $\hat{\sigma}_{r_{p,j}}$ denote the mean and standard deviation of the portfolio daily excess returns implied by model j, respectively.

The annualized Sortino ratio (SOR), which only considers the negative deviations from a certain target, B, i.e., the "downside risk".

$$SOR_j = \frac{\sqrt{252}\left(\hat{\mu}_{r_{p,j}} - B\right)}{\sqrt{\frac{1}{T - t_o} \sum_{t=t_o+1}^{T} min\left[\left(r_{p,j,t} - B\right), 0\right]^2}}, \tag{3.15}$$

where B is the reference point that constitutes the minimum acceptable rate of return, T is the total number of periods, and $t_0 + 1$ is the initial out-of-sample period. In the computation of SOR, we use $B = 0$.

The last metric is the annualized Expected Shortfall (ES) which assesses the average portfolio loss that happens with α probability. The ES is defined as:

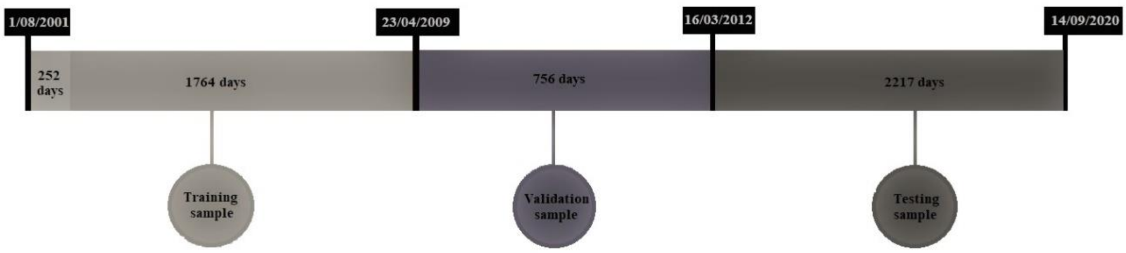
$$ES_j^{\alpha}(r_{p,j,t}) = -\sqrt{252}\mathbb{E}[r_{p,j,t}|r_{p,j,t} \leq F_{p,j,t}^{-1}(\alpha)]. \tag{3.16}$$

In the above expression, $F_{p,j,t}^{-1}(\alpha)$ is the inverse cumulative distribution function of the returns of portfolio p given by model j at time t assessed at a probability α (commonly set to 1% or 5%). This last metric is only used to assess the benefits of international diversification across geographic and economic regions (Section 3.5.3.).

3.4.4. Research Setup

As referred in Subsection 3.4.1, we use two Machine Learning techniques to forecast the return vector, which then will be used as input of the portfolio selection problem. As usual in Machine Learning applications, we partitioned the data into training, validation, and test samples. Figure 3.2 presents the sample partition, using 41%+15%+44% of the data.

Figure 3.2: Partition of the data into training, validation, and testing samples



The training sample begins on 01/08/2001 and ends on 22/03/2009. The training sample is used to compute the initial series of predictors associated with the initial series of the dependent variable (which are the daily returns), that is the first window. The predictor space has 14 variables: the previous day return, the previous business week return (5 days), the previous 1-month return (21 days), previous 2-months return (42 days), previous 3-months return (63 days), and so on, until the previous 12-months returns (252 days). Because the initial 252 observations are needed to compute the initial realizations of the vector of predictors, the training sample will have 1764 days. The validation sample has 756 days, from 23/04/2009 to 15/03/2012, and is used to choose the best models, that is, to choose the best combination of hyperparameters for the two Machine Learning techniques, Random Forests (RF) and Artificial Neural Networks (NN).[7] The testing sample has 2217 days, from 16/03/2012 to 14/09/2020, and is used to compute the forecasts of daily returns using the best RF and NN models. In the terminology commonly used in portfolio selection literature, this testing period corresponds to the out-of-sample period.

Model selection is performed using rolling windows with fixed length, equal to the number of days in the training sample, i.e., 1764 days. The criterium to choose the best model within each class for each asset is the minimization of the Mean Squared Forecasting Error.

We proceed in the following manner to choose the best within-class model. We begin by using the information of the previous 1764 days of the dependent variable and predictor set (the information in the training sample), to forecast the first 21 days (a business month) of the validation sample, we make the forecasts for that month, and then move the window forward 21 days, and make forecasts for the second business month of the validation sample, and so on, until the returns of all days of the validation sample are forecasted. For each asset we only use its own returns (with different time scales) and make forecasts for each possible combination of hyperparameters.

For the RF, we consider the following hyperparameter: (a) 500 or 1000 trees, (b) 1/3 or 2/3 of predictors to sample, and (c) 3, 5, or 10 nodes. This implies testing 12 possible combinations of hyperparameters. For NNs, we consider (a) 5 or 10 hidden layers and (b) learning rates of 0.01 or 0.001. This implies testing 4 possible combinations of hyperparameters. The specific hyperparameters tested are the most used in the literature and have usually produced good results in predicting asset returns (Ishwaran, 2015; Gu et al., 2018; Probst et al., 2019). For each asset, the sets of hyperparameters of RF and NN that lead to the best forecasting performances are chosen to forecast the daily returns in the testing sample.[8]

[7] RF and NN are implemented in MATLAB (R2020b), using the Machine Learning and Deep Learning Packages.

[8] We choose to move the window every 21 days instead of daily due to the huge computation time needed to perform the forecasts on a daily basis. That would imply running the models for 77 indexes, 16 combinations of hyperparameters and 756 days, that is running 943,488 models. Moving the window every period of 21 days reduces the number of models to be estimated to 44,928.

In the testing sample, the best RF and NN models for each asset are used to forecast the daily returns, which are then used as inputs of the portfolio selection problem. These returns are also forecasted using rolling windows with a fixed length of 2520 days, therefore making use of all the information in the training (1764 days) and validation (756 days) samples. The window is rolled forward every day.

For the portfolio selection problems that do not use the vector of returns forecasted by RF and NN, (hereafter denoted by DCC and ADCC) we use the historical geometric averages of the previous 252 daily returns, excluding the 21 most recent ones, as in Engle et al. (2019) and Jegadeesh and Titman (1993). Thus, the historical means are computed using rolling windows with a fixed length of 231 days lagged 21 days.

For all models, the dynamic conditional covariance matrixes, symmetric and asymmetric, are estimated using the most recent 1250 days, which roughly corresponds to using 5 years of past data, as in Engle et al. (2019). So, these matrixes are estimated using a rolling window with a fixed length of 1250 days. This is also the rolling window used to compute the historical covariance matrix in the Minimum Variance Portfolio (MVP)

In sum, all portfolios resulting from the different models are updated on a daily basis but using sometimes different data sets.

3.5. Empirical Results

This section presents a comparative analysis of the out-of-sample performance of the portfolios derived from different benchmarks and models. It also reports some robustness checks, and an analysis on the potential of international diversification.

3.5.1. Model Performances

Besides the three benchmark portfolios already presented in Subsection 3.4.3., we consider ten more portfolios, based on Machine Learning and conditional covariances matrixes. Table 3.2 characterizes all models used in the comparative analysis. These models are applied to stock indexes (Panel A), bond indexes (Panel B), and to stock and bond indexes (Panel C). Table 3.3 presents the annualized out-of-sample performance metrics for a moderately risk-averse investor ($\gamma = 3$).

Table 3.2: Summary description of the models

Model Acronym	Expected returns	Covariance Matrix	Countries
1/N	--	--	All countries
European index	--	--	European countries
MVP	--	Historical Covariances	All countries
DCC	Historical Means	Dynamic Conditional Covariances forecasts	All countries
ADCC	Historical Means	Asymmetric Dynamic Conditional Covariances forecasts	All countries
RF-DCC	Random Forest forecasts	Dynamic Conditional Covariances forecasts	All countries
RF-ADCC	Random Forest forecasts	Asymmetric Dynamic Conditional Covariances forecasts	All countries
NN-DCC	Artificial Neural Network forecasts	Dynamic Conditional Covariances forecasts	All countries
NN-ADCC	Artificial Neural Network forecasts	Asymmetric Dynamic Conditional Covariances forecasts	All countries
EU RF-DCC	Random Forest forecasts	Dynamic Conditional Covariances forecasts	European Union countries
EU RF-ADCC	Random Forest forecasts	Asymmetric Dynamic Conditional Covariances forecasts	European Union countries
EU NN-DCC	Artificial Neural Network forecasts	Dynamic Conditional Covariances forecasts	European Union countries
EU NN-ADCC	Artificial Neural Network forecasts	Asymmetric Dynamic Conditional Covariances forecasts	European Union countries

The results show that for stocks, the NN-ADCC is the best model which consistently outperforms the benchmarks and alternative portfolios. For this model, the mean excess return is 8.27%, the certainty equivalent return is 3.44%, the Sharpe ratio is 43.12%, and the Sortino ratio is 61.17%. Considering just EU stock indexes reduces the performance of all model classes.

The results show that for bonds, the DCC is the best model. For this model, the mean excess return is 13.29%, the certainty equivalent return is 11.09%, the Sharpe ratio is 87.89%, and the Sortino ratio is 130.22%. However, ADCC and RF-ADCC, and EU RF-ADCC have close results. Models based on NN do not perform well. The models for bonds do not show a significant performance reduction when the asset space is limited to EU bond indexes and, generally speaking, they present better results than the models for stock indexes.

Table 3.3: Portfolio performance out-of-sample

	1/N	European Index	MVP	DCC	ADCC	RF-DCC	RF-ADCC	NN-DCC	NN-ADCC	EU RF-DCC	EU RF-ADCC	EU NN-DCC	EU NN-ADCC
Panel A: Stocks													
Mean	5.03	6.27	1.77	-3.67	-3.53	5.70	5.42	8.11	8.27	2.12	2.72	7.82	7.94
Std	12.98	16.89	4.57	22.19	22.12	18.21	18.28	19.32	19.18	16.06	15.98	21.82	21.83
CER	3.16	2.60	2.13	-10.14	-9.97	1.30	0.98	3.21	3.44	-1.10	-0.46	1.51	1.63
SR	38.77	37.11	38.62	-16.53	-15.96	31.28	29.64	41.98	43.12	13.20	17.06	35.84	36.40
SOR	51.82	50.52	51.55	-21.46	-20.70	43.63	41.35	60.00	61.17	18.27	23.69	53.67	54.60
Panel B: Bonds													
Mean	4.44	6.60	2.74	13.29	12.75	10.73	10.87	2.65	2.45	12.47	12.62	8.44	8.45
Std	6.69	10.45	3.68	15.12	15.15	12.80	12.80	10.43	10.47	14.32	14.36	16.05	16.00
CER	4.52	5.78	3.24	11.09	10.49	9.34	9.50	1.68	1.47	10.54	10.69	5.34	5.37
SR	66.40	63.21	74.37	87.89	84.18	83.88	84.98	25.38	23.40	87.04	87.87	52.59	52.79
SOR	94.42	90.47	105.73	130.22	125.19	123.43	125.46	35.67	32.90	123.45	125.52	73.79	74.01
Panel C: Stocks & Bonds													
Mean	4.78	6.44	3.03	2.83	2.77	9.70	10.07	4.99	5.27	8.22	8.74	6.89	6.54
Std	9.08	10.04	5.34	14.50	14.51	14.82	14.76	13.10	13.08	14.30	14.22	19.10	18.97
CER	4.27	5.73	3.30	0.26	0.20	7.31	7.73	3.11	3.40	5.99	6.57	2.14	1.86
SR	52.66	64.12	56.72	19.52	19.09	65.44	68.24	38.11	40.28	57.52	61.45	36.05	34.50
SOR	72.26	89.19	77.56	25.67	25.12	94.84	98.70	53.66	56.87	83.67	89.73	52.59	50.05

Notes: This table reports the out-of-sample performance of portfolios constructed according to several models. Table 3.2 provides a summary description of these models. These portfolios are constructed assuming a power utility function with a risk aversion coefficient $\gamma = 3$. The performance metrics are the mean and standard deviation (Std) of excess returns, the certainty equivalent return (CER), Sharpe ratio (SR), and Sortino ratio (SOR). The models are applied to just stock indexes (Panel A), just bond indexes (Panel B), and to stock and bond indexes (Panel C). EU refers to indexes from European Union countries only (Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Netherlands, Poland, Portugal, Spain, and Sweden). The out-of-sample (testing sample) is from 16/03/2012 to 14/09/2020. All measures are annualized and presented in percentage.

Panel C of Table 3.3 presents the results for portfolios constructed using stock and bond indexes. The performances of these portfolios are generally better than stock portfolios, but worse than bond portfolios. The best model is RF-ADCC, which achieves a mean excess return of 10.07%, a certainty equivalent return of 7.73%, a Sharpe ratio of 68.24%, and a Sortino ratio of 98.70%. This model also produces a good performance for a portfolio with only EU indexes.

In sum, the results shown in Table 3.3. allow us to retrieve the following overall conclusions: (1) Covariances matrixes forecasted by ADCC improve the portfolio performances, (2) Machine Learning technics are valuable in predicting the vector of mean returns, namely NN for stocks and RF for bonds, (3) including international assets in the portfolio (not only EU assets), improves the performance of the portfolios, not only by reducing the risk but also by increasing the mean return, rendering high CERs, SRs and SORs. If one had to choose just one model for all three asset spaces, the best candidate would be the RF-ADCC.

Figure 3.3 shows the out-of-sample paths of the Sharpe ratios of RF-ADCC (Exhibit A) and NN-ADCC (Exhibit B) models and the benchmark 1/N (Exhibit C) considering stocks, bonds, and stocks and bonds for all countries. The main feature that we may observe from the figure is the variability of the moving averages of the Sharpe ratios. We highlight the significant drops in the Sharpe ratios in 2017-2018 due to the European sovereign crisis and in late 2019 until early 2020 due to the Covid-19 epidemic crisis. The worse performance for stocks using the RF-ADCC model in comparison with the NN-ADCC model seems to be related to the lower ability of the former model to accommodate the negative shocks resulting from the European sovereign crisis.

Figure 3.3: Moving averages of Sharpe ratios out-of-sample

Exhibit A: RF-ADCC

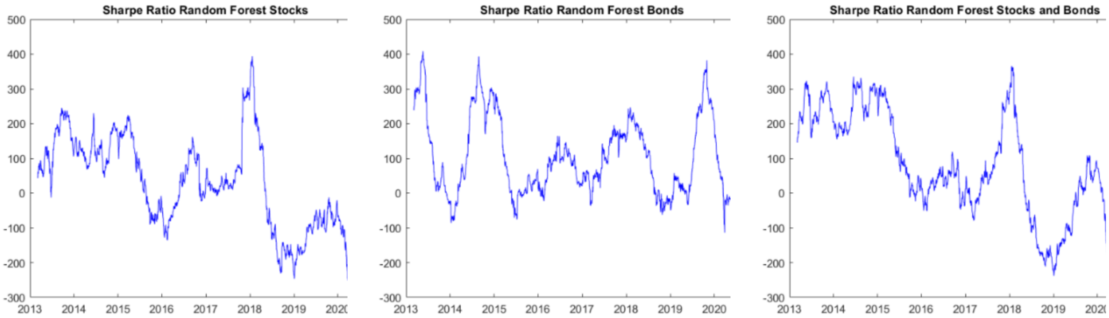


Exhibit B: NN-ADCC

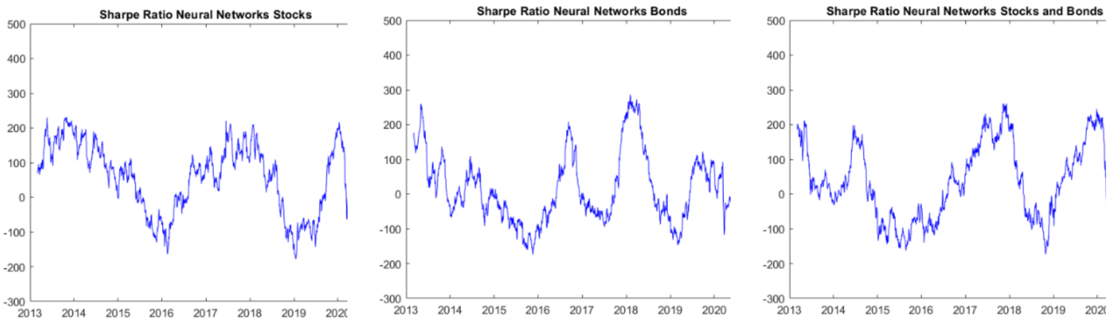
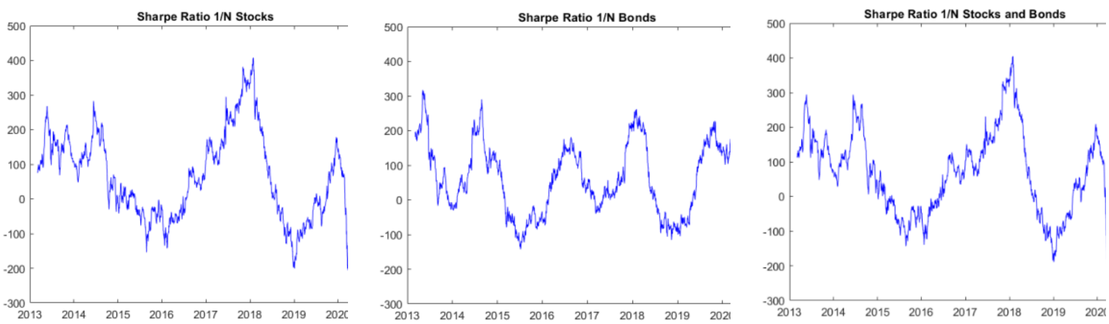


Exhibit C: 1/N



Notes: This figure shows the out-of-sample paths of the Sharpe ratios of an investor with moderate risk aversion, $\gamma = 3$, that bases her portfolio strategy on RF-ADCC (Exhibit A) and NN-ADCC (Exhibit B) models and the benchmark 1/N (Exhibit C) considering stocks, bonds and stocks and bonds for all countries. The presented SR values are moving averages, with a length of 252 days (a business year) of annualized Sharpe ratios (in percentage). Notice that the first reported value corresponds to the day after 252 days after the beginning of the period out-of-sample.

3.5.2. Sensitivity to the Risk Aversion Coefficient

In this subsection, we perform a sensitivity analysis of the out-of-sample performance to the risk aversion coefficient. We analyse the portfolio out-of-sample performance measures for low, $\gamma = 1$, and highly, $\gamma = 5$, risk-averse investors. Results are shown in Table 3.4.

A lowly risk-averse investor ($\gamma = 1$) would prefer to use a stock portfolio based on the NN-ADCC, a bond portfolio based on ADCC (closely followed by RF-ADCC), and a stock and bond portfolio based on RF-ADCC. So, for the lowly risk-averse investor, the optimal models are the same ones reported earlier for a moderately risk-averse investor ($\gamma = 3$).

An investor who is highly risk-averse ($\gamma = 5$) would prefer to use a stock portfolio based on the NN-ADCC (but including only EU stocks), a bond portfolio based on RF-ADCC (but including only EU bonds), and an EU stock and bond portfolio based on RF-ADCC.

The above portfolios beat the best benchmark portfolio for each category of assets at any metric, except in terms of CER, for a highly risk-averse investor.

Therefore, the main conclusion to retain here is that although the models to be chosen are the same independently of the degree of risk-aversion, highly risk-averse EU investors are less keen to diversify internationally their portfolios.

Table 3.4: Sensitivity of out-of-sample performance to the risk aversion coefficients ($\gamma = 1$ and $\gamma = 5$)

	γ	1/N	½ European	MVP	DCC	ADCC	RF-DCC	RF-ADCC	NN-DCC	NN-ADCC	EU RF-DCC	EU RF-ADCC	EU NN-DCC	EU NN-ADCC
Panel A: Stocks														
Mean	*1*	5.03	6.27	1.77	-3.82	-2.75	4.77	4.77	9.06	9.57	2.65	2.66	2.62	2.68
	5	5.03	6.27	1.77	-1.92	-1.95	3.68	3.90	6.97	6.95	2.47	3.39	9.07	9.04
Std	*1*	12.98	16.89	4.57	26.52	26.61	24.02	24.02	20.22	20.31	19.04	19.11	20.12	20.10
	5	12.98	16.89	4.57	18.67	18.73	16.33	16.42	18.73	18.53	13.91	13.90	20.91	20.82
CER	*1*	4.83	5.48	2.32	-6.79	-5.75	2.43	4.96	7.67	8.16	1.49	1.48	1.24	1.30
	5	1.37	-0.41	1.91	-9.81	-9.90	-2.44	-2.33	-1.22	-1.08	-1.72	-0.80	-0.86	-0.80
SR	*1*	38.77	37.11	38.62	-14.41	-10.34	19.86	19.86	44.82	47.10	13.93	13.93	13.04	13.32
	5	38.77	37.11	38.62	-10.26	-10.42	22.57	23.73	37.19	37.49	17.77	24.41	43.37	43.43
SOR	*1*	51.82	50.52	51.55	-18.63	-13.37	26.61	26.61	64.40	67.60	19.37	23.69	17.75	18.13
	5	51.82	50.52	51.55	-13.33	-13.56	31.23	32.75	52.35	52.28	24.61	33.98	65.36	65.42
Panel B: Bonds														
Mean	*1*	4.44	6.60	2.74	14.58	14.56	13.37	13.69	3.60	3.36	17.72	17.66	8.34	8.13
	5	4.44	6.60	2.74	12.94	12.49	10.63	10.46	2.04	1.86	11.94	12.06	8.52	8.40
Std	*1*	6.69	10.45	3.68	17.28	17.42	17.35	17.29	12.03	12.29	21.05	20.99	21.52	21.31
	5	6.69	10.45	3.68	14.22	14.22	11.45	11.42	10.08	10.09	12.59	12.60	15.14	15.09
CER	*1*	4.87	6.70	3.32	13.75	13.70	12.56	12.89	3.53	3.27	8.89	8.84	6.65	6.49
	5	4.05	4.62	3.10	8.92	8.43	8.32	8.16	0.14	-0.04	8.99	9.12	3.45	3.36
SR	*1*	66.40	63.21	74.37	84.41	83.59	77.05	79.20	29.88	27.37	84.17	84.11	38.75	38.16
	5	66.40	63.21	74.37	91.02	87.87	92.80	91.61	20.20	18.48	94.81	95.72	56.29	55.70
SOR	*1*	94.42	90.47	105.73	128.92	128.14	123.45	127.01	44.11	40.43	124.21	123.89	54.68	53.53
	5	94.42	90.47	105.73	134.98	130.51	135.81	133.88	28.18	25.76	137.24	139.05	79.34	78.26
Panel C: Stock & Bonds														
Mean	*1*	4.78	6.44	3.03	4.03	4.48	13.88	14.32	6.17	6.63	9.88	9.94	10.71	10.63
	5	4.78	6.44	3.03	3.26	3.27	8.42	8.15	4.96	5.06	8.87	8.89	5.40	5.36
Std	*1*	9.08	10.04	5.34	17.08	17.37	18.62	18.50	14.48	14.68	16.37	16.46	25.20	25.17
	5	9.08	10.04	5.34	13.43	13.39	13.44	13.41	12.59	12.57	12.49	12.46	17.43	17.25
CER	*1*	5.02	6.59	3.54	3.21	3.61	12.81	13.27	5.77	6.20	9.19	9.24	8.38	8.30
	5	3.40	4.65	3.01	-0.72	-0.69	4.64	4.39	1.64	1.76	5.78	5.82	-1.53	-1.41
SR	*1*	52.66	64.12	56.72	23.59	25.79	74.57	77.39	42.64	45.13	60.38	60.41	42.51	42.22
	5	52.66	64.12	56.72	24.31	24.42	62.65	60.81	39.37	40.24	70.99	71.32	31.00	31.10
SOR	*1*	72.26	89.19	77.56	31.72	34.84	114.18	118.68	59.97	63.99	87.49	87.62	70.68	70.35
	5	72.26	89.19	77.56	32.10	32.27	89.96	87.32	55.76	57.05	103.96	104.37	44.20	44.36

Notes: This table reports the out-of-sample performance of portfolios constructed according to several models. Table 3.2 provides a summary description of these models. These portfolios are constructed assuming a power utility function with risk aversion coefficients $\gamma = 1$ and $\gamma = 5$. The performance metrics are the mean and standard deviation (Std) of excess returns, the certainty equivalent return (CER), Sharpe ratio (SR), and Sortino ratio (SOR). The models are applied to just stock indexes (Panel A), just bond indexes (Panel B), and to stock and bond indexes (Panel C). EU refers to indexes from European Union countries only (Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Netherlands, Poland, Portugal, Spain, and Sweden). The out-of-sample (testing sample) is from 16/03/2012 to 14/09/2020. All measures are annualized and presented in percentage.

3.5.3. Gains of International Diversification

This subsection examines whether investors from seven different regions, North America, South America, America, EU, Middle East, Asia, and Oceania, will profit from having an internationally worldwide diversified portfolio rather than only investing in their home region. We also analyse this issue by segmenting the 44 countries into emerging markets (EM) and developed markets (DM). This analysis is conducted considering that investors, with different risk aversion levels, choose their stock and bond portfolios based on RF-ADCC models, which, as reported previously, produced good results in terms of out-of-sample performance. Table 3.5 reports the results.

North American portfolios stand out as the best ones in terms of several performance metrics, although American investors may improve the mean return by internationally diversifying their portfolios. In fact, North American portfolios achieve the best CER, SR, SOR and ES 5%, at any level of risk aversion (except the CER of a lowly risk-averse investor, which is lower than the corresponding figure of the portfolio with all countries). These good performances of North American portfolios have been duly documented in the literature over the last decade. Middle Eastern portfolios have the worst performance, implying that international diversification results in higher benefits for investors located in that region.

The results show that for other regions, low or moderately risk-averse investors, i.e., with $\gamma = 1$ and $\gamma = 3$, benefit from international diversification, even sometimes at the expense of an increased risk (measured by the standard deviation and expected shortfall). For regions other than North America, the portfolio "All countries" achieves the CER, SR and SOR. Precisely because of that increased risk, highly risk-averse investors may be less prone to diversify their portfolios worldwide, especially those investors from EU countries.

Overall, low and moderately risk-averse investors from South America, Europe, the Middle East, Asia, and Oceania, who invest outside their regions, typically benefit from diversifying their portfolio. Notably, investors from emerging (developed) markets benefit from including in their portfolios, assets from developed (emerging) markets.

Table 3.5: Comparative analysis of portfolio performance by geographic and economic regions

	γ	**North America**	**South America**	**America**	**EU**	**Middle East**	**Asia**	**Oceania**	**EM**	**DM**	**All countries**
	1	11.26	6.29	5.56	9.94	-5.22	9.65	3.78	9.90	8.52	***14.32***
Mean	**3**	9.14	-1.03	6.85	8.74	-4.93	6.35	1.78	6.42	6.47	***10.07***
	5	7.94	-1.05	6.32	**8.89***	-4.15	5.59	1.43	6.33	5.95	*8.15*
	1	12.70*	25.68	19.73	16.46*	23.45	13.34*	**11.55***	18.28*	15.66*	*18.50*
Std	**3**	**8.56***	19.42	11.81*	14.22*	17.45	10.37*	10.14*	13.46*	11.38*	*14.76*
	5	**7.56***	16.92	10.02*	12.46*	13.92	9.32*	9.33*	12.02*	10.33*	*13.41*
	1	11,11	3.65	4.27	9.24	-7.60	9.41	3.77	8.88	7.96	***13.27***
CER	**3**	**9.08***	-5.86	5.56	6.57	-9.00	5.54	0.89	4.45	5.32	*7.73*
	5	**7.43***	-7.14	4.57*	5.82*	-8.36	4.16	-0.10	3.41	4.01	*4.39*
	1	**88.67***	24.50	28.18	60.41	-22.26	72.35	32.75	54.18	54.44	*77.39*
SR	**3**	**106.82***	-5.32	57.98	61.45	-28.23	61.25	17.57	47.72	56.82	*68.24*
	5	**105.07***	-6.20	63.08*	71.32*	-29.83	60.01	15.35	52.67	57.57	*60.81*
	1	**128.38***	35.19	40.19	87.62	-26.82	107.91	45.55	76.08	82.24	*118.68*
SOR	**3**	**158.53***	-7.55	85.69	89.73	-34.10	90.63	24.05	66.41	82.76	*98.70*
	5	**156.30***	-9.17	93.88*	104.37*	-36.93	88.87*	21.21	74.08	83.65	*87.32*
	1	**14.94***	46.68	35.14	24.00	53.58	17.86*	20.04*	27.81	23.77*	*23.85*
ES 5%	**3**	**8.51***	41.09	17.51*	20.59	40.92	15.04*	19.13*	21.34	17.01*	*20.38*
	5	**7.65***	35.95	14.35*	16.82*	32.87	13.63*	17.81*	18.46*	15.36*	*19.50*

Notes: This table reports the out-of-sample performance of stock and bond portfolios constructed according to the RF-ADCC model for each geographic region and considering the grouping of countries into emerging markets (EM) and developed markets (DM). These portfolios are constructed assuming a power utility functions with risk aversion coefficients $\gamma = 1$, $\gamma = 3$ and $\gamma = 5$. The performance metrics are the mean and standard deviation (Std) of excess returns, the Certainty Equivalent Return (CER), Sharpe ratio (SR), Sortino ratio (SOR), and Expected Shortfall at 5% (ES 5%). The geographic regions are: North America (Canada, Mexico, and the US), South America (Argentina, Brazil and Chile), America, which is the ensemble of North and South America countries, EU (Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Netherlands, Poland, Portugal, Spain, and Sweden), Middle East (Egypt, Israel, Pakistan, and Turkey, Asia (China, India, Indonesia, Korea, Singapore, Japan, Hong Kong, Malaysia, Philippines, Taiwan and Thailand), and Oceania (Australia and New Zealand (Oceania). All countries consider all 44 countries in the sample, on which overall international diversification is achieved. The distinction between EM and DM was done according to the Market classification MSCI (https://www.msci.com/our-solutions/indexes/market-classification). Values in Bold represent the best value of each statistic (maximum for Mean, CER, SR and SOR; minimum for Std and ES 5%). Values with an asterisk represent those statistics that are better than those of All countries. The out-of-sample (testing sample) is from 16/03/2012 to 14/09/2020. All measures are annualized and presented in percentage.

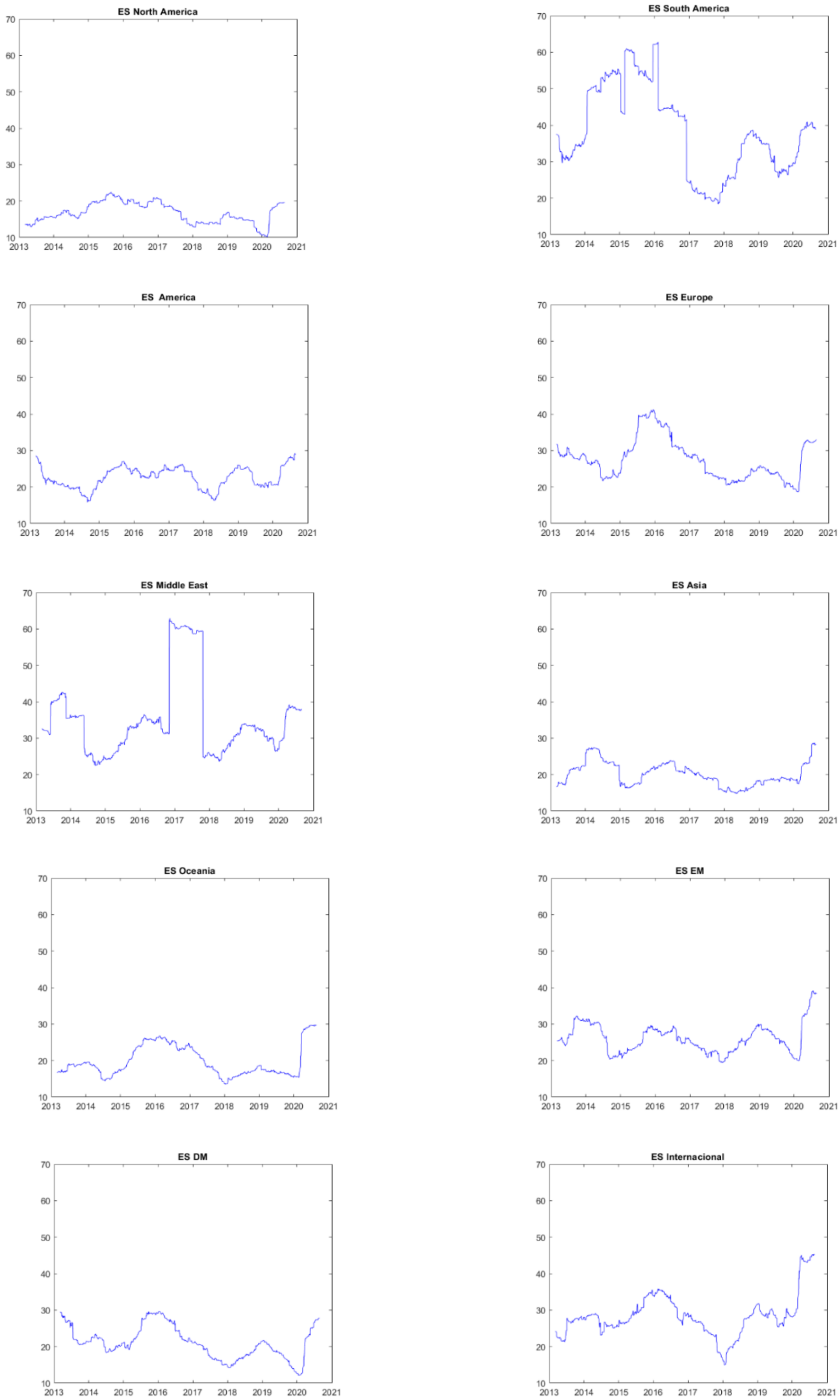
Figure 3.4 reports the out-of-sample annual moving averages of Expected Shortfall (ES) at 5% for nine different geographic regions, Emerging and Developed markets, and considering all countries (ES international). The ES were computed considering an investor with moderate risk aversion ($\gamma = 3$), which bases her allocation strategy on the RF-ADCC model considering both stocks and bonds.

Expected Shortfall (ES) is high, especially in South America and the Middle East, and varies substantially over time and across geographic regions. Some events may explain the differences between different regions. For instance, the 2013-2016 recession in the Latin American countries and the 2017 conflicts in Turkey and Egypt may justify the 30% increase in the ES for both regions. The 2015 sovereign crisis in Europe may explain the rise of 15% in the EU expected shortfall. And the outburst of Covid-19 almost surely justifies the high spike in the ES at the end of all series.

The tail risk in America, especially in North America, is the lowest across all regions, varying between 10% and 20%, followed by Asia, and Oceania, where the values range between 10% and 30%. Hence, in terms of tail risk, an investor from the Middle East or South America would substantially benefit from diversifying her portfolio internationally, especially for 2016-2017. This is justified by the tail risk measure of the international portfolio trending downwards and reaching a value of 10% in this period.

Lastly, emerging and developed markets ES present very interesting results. Both markets show similar trends at the beginning of the sample, suggesting an increase in country dependence when markets are experiencing extreme negative returns, possibly due to market integration or cross-country trade agreements. However, from 2016 to 2019, the patterns have changed. The ES of DM declined more rapidly, from 30% to 10%, while the ES of EM varied from 30% to 20%. This suggests that the benefits of diversifying over EM have been diluted over time, while the benefits of investing in DM have grown from a portfolio tail-risk perspective.

Figure 3.4: Moving average of Expected Shortfall out-of-sample



Notes: This figure presents the out-of-sample paths of Expected Shortfall (ES) at the 5% level of a moderately risk-averse risk aversion ($\gamma = 3$) who based her asset allocation strategy on a RF-ADCC model. The geographic regions considered are North America, South America, America, EU, Middle East, Asia, Oceania, the segmentation into Emerging and Developed Markets, and for the overall diversified portfolio (ES international). The presented ES values are moving averages, with a length of 252 days (a business year) of annualized ES (in percentage). Notice that the first reported value corresponds to the day after 252 days after the beginning of the period out-of-sample.

3.6. Conclusion

This chapter assesses the performance of internationally diversified portfolios built on new estimation schemes for the vector of expected returns and covariance matrix in a large dimension asset space. Firstly, we use Machine Learning techniques, namely Random Forests and Artificial Neural Networks, to forecast daily excess stock and bond returns. Secondly, we generalize the DCC model by Engle et al. (2019) by allowing for asymmetric effects in innovations. We empirically test the roles played by these features when compared to simple benchmark portfolios, and other simpler models. We conclude that the proposed models (the Asymmetric DCC with Random Forests or Neural Networks) consistently outperform the benchmarks and other simple model constructions. We find that a power utility investor with moderate risk aversion that bases her allocation strategy on the proposed RF-ADCC model would experience a good performance. These good results may be attributed to the ability of the model to effectively capture nonlinearities in data.

Further, we also study the international diversification potential under different portfolio performance measures. We rely on the proposed model to explore the benefits of holding an internationally diversified portfolio of 77 stocks and bonds indexes. We conclude that investors from South America, the EU, the Middle East, Asia, and Oceania would benefit from further diversifying their portfolios. Additionally, we observe a decrease in the benefits of diversifying over Emerging countries and a rise for Developed countries from March 2012 until September 2020.

Chapter 4 - Industry Lead-Lag Relationships Between the US and Other Developed Countries

4.1. Introduction

In a complete and frictionless market, conventional asset pricing theory assumes that information dissemination across related markets occurs immediately. In a frictionless market, with rational expectations, a shock in one asset can be rapidly recognized by investors in other related assets. Consequently, equity prices adjust promptly and completely to any information shock. However, there is compelling empirical evidence that investors face limitations in processing information and non-trivial market frictions, therefore information does not spread across markets as assumed by the conventional theory (see, for instance, Shiller, 2000; Hong et al., 2007). In a more realistic framework, equity prices may adjust to new information with some delay. For instance, industry-specialized investors may fail to fully assimilate new information from shocks in other industries. Hence, at the industry level, this may imply the existence of significant lead-lag relationships between industry indexes, and, therefore, industry return predictability.

This study analyses the interdependence between industry returns in an international context. To the best of our knowledge, this study is one of the firsts to directly examine lead-lag relationships across countries at the industry level. Previous research has mainly focused on firm-level or intra-industry information flows (see, for instance, Rapach et al., 2013; Bollerslev et al., 2013). However, as argued by Hou et al., 2007, due to market segmentation, industry information only gradually spreads out over related industries, and thus returns of an industry can be predicted by returns of related industries.

At the international level, the US fulfils a key role. According to the World Bank (2021), the US has the world's largest national economy, accounting for almost a quarter of global GDP; it is the most important export destination of most countries worldwide and represents more than one third of global stock market capitalization. Because of its size and interconnectivity, events in the US economy are likely to have a global impact. So, our main research hypothesis is that lagged returns of US industries may help predict returns of industries of other countries.

We show that weekly lagged returns of US industries have a strong and significant causal relationship with most other countries considered in the study, while lagged returns of other countries have limited ability to predict US returns at industry levels. Notably, we highlight that lagged returns of US Basic Materials and Energy industries have strong and significant predictive power and causality to industries of other countries. This finding is highly plausible since firms in other industries rely heavily on commodities and fuels, and hence lagged returns of these later industries, which are placed earlier in the production chain, should impact the returns of industries positioned later in the production chain.

The leading role of the US is even more pronounced during recession periods, when cross-country correlations are stronger. This implies that the ability of the US lagged returns to predict current returns of other countries is much greater when the US experienced a recession in the week before. Results also suggest that past values of US industries' volatilities contain information that helps predict the volatility of other countries. Lastly, we analyse the Granger causality in distribution for returns and volatilities at the industry level. Our results suggest that other countries did not timely incorporate shocks affecting the US industries, meaning that countries react with a delay to news from the US.

The remaining of this study is structured into five sections. Section 4.2 presents a brief literature review. Section 4.3 presents the data and provides some descriptive statistics. Section 4.4 outlines the basic theoretical concepts and presents the tests specifications. Section 4.5 shows the main results, and Section 4.6 concludes the study.

4.2. Literature Review

The complexity of the relationships between asset prices has long been subject of analysis from many academics and practitioners. For instance, Lo and MacKinlay (1990) wrote one of the most influential and earliest works in the lead-lag literature. The authors showed that returns of large stocks led returns of small stocks in the US from 1962 to 1987. In the early nineties, many studies analysed the lead-lag relationship between various asset prices and industries (see, e.g., Roll et al., 1992; Arshanapalli and Doukas, 1993; Brennan et al., 1993; Boudoukh et al., 1994; Jegadeesh et al., 1995; Copeland et al., 1998; Moskowitz and Grinblatt, 1999). For instance, Copeland et al. (1998) found that the US had a statistically significant one-day lead over markets in Europe and Asia in the early nineties. However, this lead did not extend over one day. The authors also found that internationalized industries (e.g., airlines) were significantly more sensitive to leads than local ones (e.g., casinos). Moskowitz and Grinblatt (1999) showed that cross-sectional industry momentum accounted for the cross-sectional momentum in individual firm returns, reinforcing the idea that industries had important interconnections with each other.

In the mid-2000s, many other studies have analysed lead-lag relationships between various industries. However, they mainly focused on information flows at the firm level in the US market. For instance, Hou et al. (2007) studied the transmission of information between big and small firms. The authors identified a lead-lag effect between stock returns of these firms in the US between July 1963 and December 2001. According to the authors, this slow information transmission could result from many sources, including incomplete markets, limited stock market participation, asymmetric information, noise trading, limited investor attention, transaction costs, short-sale restrictions, legal constraints imposed to institutional investors, and other market frictions. Hong et al. (2007) investigated the transmission of information between US industries and the overall market from January 1946 to December 2002. They concluded that the US stock market reacted with a delay to the information contained in the industry returns about their fundamentals. As a result, industry returns that incorporated information on macroeconomic fundamentals tended to lead the aggregate market. Hence, a substantial number of US industries, such as retail services, commercial real estate, metal, and petroleum, could anticipate the stock market by up to two months.

Menzly and Ozbas (2010) have found economic links between certain specific firms and industries, which contributed significantly to cross-firm and cross-industry return predictability. In line with the work of Hong et al. (2007), the authors interpreted their findings as evidence of delayed information transmission across economically connected firms and industries. According to Rapach et al. (2015), an industry has an economic link to another if its returns can be predicted by the lagged returns of the other one. This suggests, for example, that shocks in the technology industry might impact returns in the manufacturing industry, even if these industries are not directly involved with each other. Industries can also be indirectly connected along the production chain, resulting in valuable economic connections that extend beyond the direct customer-supplier link. The authors argued that complex industry interdependences increase the potential for delayed responses to information and produce cross-industry return predictability.

Several studies have reported cross-industry linkages. For instance, Rapach et al. (2019) showed that lagged returns of financial and commodities industries had forecasting ability on most industries. Jacobsen et al. (2019) demonstrated that industrial metal returns led the stock market even after adjusting for other widely used predictors. Additionally, the authors showed that during recessions, there was a direct relationship between the stock market and past industrial metal returns and an inverse relationship during expansions. Khalfaoui et al. (2021) analysed the lead-lag relationships between oil and several metal prices and concluded that gold and platinum were highly connected with oil. They showed that these metals were strongly influenced by oil prices, especially during turmoil periods in global markets. Further, Jiang et al. (2020) verified that there was a lead-lag linkage implying that oil prices led six stock market indices (China, India, Japan, Saudi Arabia, Russia, and Canada) from 27 to 30 weeks.

More broadly, Lee et al. (2019) studied the impact of technological proximity on the lead-lag relationship between stock returns. They showed that businesses with a positive peer group return in the previous month outperformed negative ones. Parsons et al. (2020) documented lead-lag effects on returns between cohead quartered firms in different industries and showed the existence of geographic lead-lags that imply a risk-adjusted return of 5%-6% annually, which was half the value observed for industry lead-lag effects. Whereas industry lead-lag effects were stronger among small, thinly traded stocks with low analyst coverage, geographic lead-lags were unrelated to these proxies for investor scrutiny. More recently, Zeng et al. (2021) reported that economic links such as customer-supplier relationships and peer effects only accounted for a small share of the frequently observed cross-firm lead-lag relations. Instead, most of the cross-firm lead-lag ties were driven by several classical factors.

Although some studies have analysed lead-lag relationships across industries, they mainly focus on the US market, which is not surprising given that the US is the world's largest stock market. Most of these studies point out that, even for a highly liquid mark such as the US, the existence of lead-lag relationships can be interpreted as evidence of information frictions resulting from limited investor attention and limited information-processing capabilities.

An exception to this trend is Rapach et al. (2013), that besides the US also reported strong predictability of lagged US monthly index return over international markets from 1980 until 2010. Additionally, since the US is a large trading partner for many countries and has the world's largest stock market, investors are likely to focus more on the US. As a result, information on the US macroeconomic fundamentals is relevant to foreign stock markets. Rapach et al. (2013) also reported that Swedish returns showed in-sample predictability power on other foreign returns. This can be justified by the high institutional ownership of Sweden and the fact that institutional investors are more able to collect and process information, which contributes to a higher pricing efficiency in the Swedish market.

Wen et al. (2015) also studied international markets and demonstrated that the US stock returns could predict South African returns from 1973 to 2014. Cambón et al. (2017) showed that Spanish industries, that provided valuable and important economic information, did not drive the equity market or economic activity. The hypothesis which Hong et al. (2007) presented, was not supported in the case of Spain, where company characteristics, especially size, may be more relevant in understanding lead-lag patterns. Sarwar et al. (2017) studied the impact of US stock market uncertainty (proxied by VIX) on Latin American and aggregate emerging markets before, during, and after the 2008 crisis. The authors found that increases in VIX led to significant immediate and delayed declines in emerging market returns in all periods. Tse (2018) examined the lead-lag relationships between 11 industrialized countries using international futures prices. He concluded that the futures markets were more contemporaneously correlated in market downturns, while the lead-lag relationships were more significant in market upturns. These findings suggest that investors react quicker to negative than positive news. When one national stock market falls, investors in other countries sell their domestic stocks in the same month. In contrast, investors are less likely to buy stocks from other countries when one market is performing well.

Understanding whether one company or industry leads another has important implications for investment planning. Croce et al. (2019) showed that firms in leading industries (i.e., industries whose cash flows contained information relevant for future aggregate growth) pay a 4% higher average annualized return than firms in lagging industries. Rapach et al. (2019) reported that a long investment strategy in highly forecastable industries and a short position in the lowest forecastable ones would generate an annualized alpha of at least 8%.

In sum, previous literature has focused mainly on firm-level returns or intra-industry information flow. However, industry or firm information only gradually spreads out over the related firms and industries. Thus, returns of an industry can be predicted by its related industries (Hou et al., 2007). Also, studies primarily focus on the US market rather than having a global worldwide scope. And, as we know, international markets are now, more than ever, highly connected due to globalization, technological advances, and increasing market integration.

4.3. Data Description and Preliminary Analysis

The data consists of closing daily values of Total Return Equity Indexes, in US dollars,[9] of eleven industries for seven countries, from 01/01/1973 to 17/05/2021, retrieved from the Thomson Reuters DataStream database. These industries, corresponding level 1 Industry Classification Benchmark (ICB), are: Basic Materials (BM), Consumer Discretionary (CD), Consumer Staples (CS), Energy (EN), Financials (FI), Health Care (HC), Industrials (IN), Real Estate (RE), Technology (TEC), Telecommunications (TEL), and Utilities (UT). We selected the top 7 countries according to the MSCI ACWI Country Index, namely Canada, France, Germany, Japan, China, the UK, and the US. Daily data was then converted into weekly data using Wednesday-to-

[9] Series that were in the domestic currency were converted to US dollars using the series of exchange rates, also obtained from the Thomson Reuters DataStream database.

Wednesday values. We work with weekly data to avoid nonsynchronous trading due to different time zones, while using data from Wednesdays avoids eventual Monday effects.

Table 4.1 shows the data availability, presenting the starting dates when the series do not cover the overall period. China has the shorter times series, as most of the series start after 1993, and, for the Technology sector, it started only in 2015. Nevertheless, we choose to include China in our study because it is the largest and one of the fastest-growing emerging economies worldwide. Germany also presents two series that begin quite late, but for the same reasons, we include this country in our study (the German market represents 2.5% of the world market value, according to the MSCI ACWI index, on May 14, 2021).

Table 4.1: Data availability

	BM	CD	CS	EN	FI	HC	IN	RE	TEC	TEL	UT
Canada	C	C	C	C	C	C	C	**01/07/1985**	C	**05/02/1976**	C
France	C	C	C	C	C	C	C	C	C	**17/10/1997**	**18/07/2000**
Germany	C	C	C	**03/05/2006**	C	C	C	**23/09/1993**	**03/11/1988**	C	C
Japan	C	C	C	C	C	C	C	C	C	C	C
China	**26/07/1993**	**26/07/1993**	**26/07/1993**	**02/12/1994**	**26/07/1993**	**27/02/2004**	**26/07/1993**	**26/07/1993**	**29/06/2015**	**15/11/2002**	**28/06/1995**
UK	C	C	C	C	C	C	C	C	C	**04/11/1981**	**05/12/1986**
US	C	C	C	C	C	C	C	C	C	C	C

Notes: This table presents the availability of data for the 11 industries: Basic Materials (BM), Consumer Discretionary (CD), Consumer Staples (CS), Energy (EN), Financials (FI), Health Care (HC), Industrials (IN), Real Estate (RE), Technology (TEC), Telecommunications (TEL), and Utilities (UT). "C" indicates that the series is complete, for the overall sample (01/01/1973 to 17/05/2021). Dates indicate the beginning of the series if they are only available after 01/01/1973.

Table 4.2 contains the descriptive statistics of weekly logarithmic returns of 11 industries for seven countries. While cross-industry variations are large, the table shows that, on average, the US presents the highest mean (0.13%) and the lowest risk level (2.76%). France offers the second-highest mean return (0.12%); however, it has a relatively high risk (3.35%). Canada, Japan, and the UK have the same mean return (0.11%). The highest mean return is reported for the Canadian Consumer Staples (0.19%), and the lowest one corresponds to the German Energy sector (-0.05%). One should recall that data for this sector only starts in 2015. The two sectors that present the highest trade-off between risk-return are the Consumer Staples (CS), with an average return of 0.15%

and a risk of 2.97%, and the Health Care (HC) sector, with an average of 0.15% and a risk of 2.90%.

Table 4.2: Descriptive statistics of weekly returns

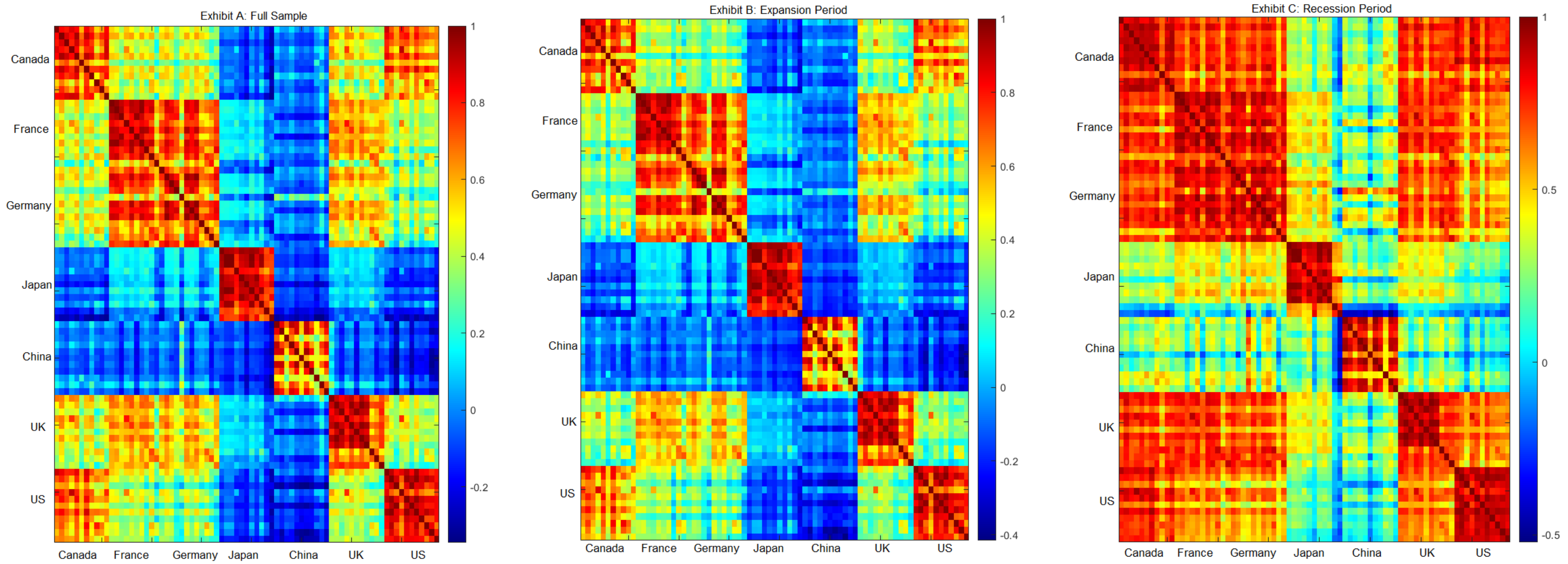
		BM	CD	CS	EN	FI	HC	IN	RE	TEC	TEL	UT	*Average*
	Canada	0.07	0.10	0.19	0.08	0.14	0.14	0.14	0.07	0.11	0.12	0.09	**0.11**
	France	0.17	0.15	0.18	0.11	0.12	0.15	0.16	0.09	0.15	0.00	-0.01	**0.12**
	Germany	0.10	0.14	0.17	-0.05	0.08	0.12	0.10	0.10	0.20	0.02	0.07	**0.09**
Mean	**Japan**	0.09	0.13	0.13	0.07	0.07	0.15	0.13	0.10	0.13	0.18	0.07	**0.11**
	China	0.07	0.16	0.10	0.05	0.10	0.18	0.06	0.18	-0.02	0.02	0.05	**0.09**
	UK	0.10	0.10	0.14	0.10	0.08	0.14	0.12	0.06	0.15	0.10	0.08	**0.11**
	US	0.13	0.15	0.16	0.10	0.14	0.18	0.16	0.13	0.17	0.08	0.07	**0.13**
	Average	0.10	0.13	0.15	0.07	0.10	0.15	0.12	0.10	0.13	0.07	0.06	**0.11**
	Canada	3.66	2.81	2.29	3.65	2.69	4.13	3.35	2.71	4.31	2.56	2.21	**3.12**
	France	3.21	3.52	3.07	3.86	3.61	2.99	3.45	2.75	4.41	3.19	2.78	**3.35**
	Germany	2.95	3.86	2.96	3.25	3.24	2.37	3.07	2.77	3.82	3.71	2.56	**3.14**
Std	**Japan**	3.23	2.83	2.92	4.10	3.50	2.59	3.06	3.99	3.73	4.20	3.33	**3.41**
	China	4.43	4.62	4.75	3.91	3.79	3.12	3.96	4.90	1.72	2.68	3.61	**3.77**
	UK	3.94	3.28	2.70	3.65	3.39	2.81	3.30	3.85	4.18	3.37	2.29	**3.34**
	US	3.10	2.66	2.10	3.15	2.91	2.26	2.68	3.57	3.31	2.49	2.10	**2.76**
	Average	3.50	3.37	2.97	3.65	3.30	2.90	3.27	3.51	3.64	3.17	2.70	**3.27**
	Canada	-0.58	-1.61	-0.30	-1.27	-0.49	-1.75	-0.52	-2.48	-0.81	-0.58	-1.01	**-1.04**
	France	-0.47	0.04	0.91	-0.15	-0.41	-0.28	-1.07	-1.41	-0.27	-0.38	-1.79	**-0.48**
	Germany	-0.63	3.25	7.69	-1.17	-0.76	-0.51	-0.88	-0.60	-0.39	-0.12	-0.71	**0.47**
Skew	**Japan**	-0.09	-0.12	-0.06	0.13	0.20	0.02	-0.21	0.07	-0.03	0.79	0.31	**0.09**
	China	0.10	0.41	1.60	0.24	0.40	0.17	0.39	0.07	-0.49	0.13	-0.74	**0.21**
	UK	-0.49	-0.78	-0.18	-0.33	-0.45	-0.11	-0.68	-0.50	-0.24	0.05	-0.06	**-0.34**
	US	-0.66	-0.59	-0.60	-1.11	-0.70	-0.48	-0.75	-0.43	-0.51	-0.42	-0.60	**-0.62**
	Canada	7.71	29.72	5.93	17.14	8.82	27.60	6.04	33.10	10.98	9.86	12.00	**15.35**
	France	5.79	11.86	19.65	13.10	9.07	4.78	13.82	25.91	6.89	13.42	26.16	**13.68**
	Germany	7.42	81.84	217.5	23.00	13.29	6.65	10.40	21.27	11.84	6.72	8.66	**37.14**
Kurt	**Japan**	6.40	5.45	5.63	5.51	7.18	5.01	5.18	6.90	5.41	7.43	7.77	**6.17**
	China	10.01	14.27	24.84	11.52	13.72	22.53	10.81	14.95	38.70	15.05	13.17	**17.23**
	UK	8.11	14.44	6.53	15.23	9.91	6.37	9.72	11.42	10.75	7.24	7.68	**9.76**
	US	7.82	7.00	6.34	17.06	10.26	5.83	7.64	9.58	6.14	7.85	9.05	**8.60**

Notes: This table presents the mean, standard deviation (Std), skewness (Skew), and kurtosis (Kurt) of weekly logarithmic returns of the 11 industries (Basic Materials (BM), Consumer Discretionary (CD), Consumer Staples (CS), Energy (EN), Financials (FI), Health Care (HC), Industrials (IN), Real Estate (RE), Technology (TEC), Telecommunications (TEL), and Utilities (UT)) for Canada, France, Germany, Japan, China, the UK, and the US. Data covers the period from 03/01/1973 to 12/05/2021. Mean and standard deviation values are in percentage.

On average, the skewness is moderate across all countries. Weekly returns are left-skewed for Canada, France, the UK, and the US. All countries present excess kurtosis, especially Germany, mainly due to Consumer Staples (CS) and Consumer Discretionary (CD) industries.

Figure 4.1 shows the correlation maps for the series industry/country for the full sample period (Exhibit A), the expansion periods (Exhibit B), and the recession periods (Exhibit C). The partition of the data into expansion and recession periods is based on the NBER business cycle classification in the US. For all samples, we can observe high and positive cross-country correlations between France and Germany, and between Canada and the US. Interestingly, most industries in China and Japan are negatively correlated to industries of other countries in the overall sample and during expansion periods. There are two main illations to withdraw from Figure 4.1: Firstly, as expected, correlations between industries of the same country present the higher correlations and, secondly, correlations increase significantly in recession periods, especially between European and American countries.

Figure 4.1: Correlation maps



Notes: This figure shows the correlation heat maps between the industries of 7 countries, for the full sample (Exhibit A), in expansion periods (Exhibit B), and in recession periods (Exhibit C) in the US according to the NBER (https://www.nber.org/research/data/us-business-cycle-expansions-and-contractions). The countries are Canada, France, Germany, Japan, the UK, and the US. The industries, from the left (top) to the right (bottom), are Basic Materials (BM), Consumer Discretionary (CD), Consumer Staples (CS), Energy (EN), Financials (FI), Health Care (HC), Industrials (IN), Real Estate (RE), Technology (TEC), Telecommunications (TEL), and Utilities (UT). On the right side of each panel, the table reports the correlation scales associated to each colour. The dark red corresponds to a correlation of 1.

4.4. Methodology

In this section, we present the econometric tests implemented in the empirical application. Our analysis of the lead-lag relationships in international industries proceeds in the following way: First, we estimate bivariate VAR(1) for all the series (11 industries, 7 countries). The use of VARs with just one lag was suggested by the Akaike and Schwartz information criteria (AIC and BIC). Using the estimated VAR(1) we compute pairwise Granger causality tests and feedback measures. Next, we partitioned the data into expansion and recession periods according to the business cycles classification for the US and conduct a similar analysis in these periods. Besides the analyses of causality and feedback in the mean, using the same methodology we also study causality and feedback in volatility for the overall sample, and the expansion and recession periods. Lastly, we also study Granger causality in distribution of returns and volatility.

4.4.1. Granger Causality

The lead-lag relationship is firstly identified via Granger causality tests (Granger, 1969). To establish the general result, suppose that we have two time series of returns $r_{1,t}$ and $r_{2,t}$. Supposing that their dynamics follow a bivariate VAR(1), then:

$$\begin{bmatrix} r_{1,t} \\ r_{2,t} \end{bmatrix} = \begin{bmatrix} \alpha_y \\ \alpha_x \end{bmatrix} + \begin{bmatrix} \phi_1 & \phi_{1,2} \\ \phi_{2,1} & \phi_2 \end{bmatrix} \begin{bmatrix} r_{1,t-1} \\ r_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{r_1 t} \\ \varepsilon_{r_2,t} \end{bmatrix}, \tag{4.1}$$

$$\mathbf{\Omega} = Cov \begin{bmatrix} \varepsilon_{r_1 t} \\ \varepsilon_{r_2,t} \end{bmatrix} = \begin{bmatrix} \sigma^2_{r_{1,t}} & \sigma^2_{r_{1,t},r_{2,t}} \\ \sigma^2_{r_{2,t},r_{1,t}} & \sigma^2_{r_{2,t}} \end{bmatrix}. \tag{4.2}$$

where $\mathbf{\Phi} = \begin{bmatrix} \phi_1 & \phi_{1,2} \\ \phi_{2,1} & \phi_2 \end{bmatrix}$ is the coefficient matrix. It is usual to assume that innovations are gaussian and serially uncorrelated.

To assess the causality from r_2 to r_1 we test the hypothesis $\phi_{1,2} = 0$. Similarly, r_1 does not Granger cause r_2 if $\phi_{2,1} = 0$. The absence of Granger causality in either direction implies the coefficient matrix $\mathbf{\Phi}$ is diagonal. Under this hypothesis, the VAR simplifies to:

$$\begin{bmatrix} r_{1,t} \\ r_{2,t} \end{bmatrix} = \begin{bmatrix} \alpha_y \\ \alpha_x \end{bmatrix} + \begin{bmatrix} \phi_1 & 0 \\ 0 & \phi_2 \end{bmatrix} \begin{bmatrix} r_{1,t-1} \\ r_{2,t-1} \end{bmatrix} + \begin{bmatrix} \xi_{r_1,t} \\ \xi_{r_2,t} \end{bmatrix}. \tag{4.3}$$

In the present case of a VAR(1), testing Granger causality from r_2 to r_1 requires computing the sum of squared residuals of the regression of $r_{1,t}$ on $r_{1,t-1}$ and $r_{2,t-1}$, RSS_1, computing the sum of squared residuals of the regression of $r_{1,t}$ only on $r_{1,t-1}$, RSS_0 and computing the test:

$$S = \frac{(RSS_1 - RSS_0)}{RSS_0/(T-3)}. \tag{4.4}$$

The test S follows a F distribution with 1 and $T-3$ degrees of freedom, i.e., $F(1, T-3)$. The Granger causality from r_1 to r_2 is tested analogously.

4.4.2. Geweke Measures of Feedback

To assess the information transmission between returns, we use the feedback measures of Geweke (1982). These measures are applied to each pair of industry/country returns. They can be used to test the degree of feedback in both directions, contemporaneously and overall linear dependence.

Measure of lagged feedback from r_1 to r_2:

$$F_{r_1 \to r_2} = \ln\left(\frac{\sigma^2_{\xi_{r_2}}}{\sigma^2_{\varepsilon_{r_2}}}\right). \tag{4.5}$$

Measure of lagged feedback from r_2 to r_1:

$$F_{r_2 \to r_1} = \ln\left(\frac{\sigma^2_{\xi_{r_1}}}{\sigma^2_{\varepsilon_{r_1}}}\right). \tag{4.6}$$

Measure of contemporaneous feedback between r_1 and r_2:

$$F_{r_1 \leftrightarrow r_2} = \ln\left(\frac{\sigma^2_{\varepsilon_{r_1}} \sigma^2_{\varepsilon_{r_2}}}{|\boldsymbol{\Omega}|}\right). \tag{4.7}$$

Measure of total feedback (total linear dependence) between r_1 and r_2:

$$F_{r_1.r_2} = \ln\left(\frac{\sigma^2_{\xi_{r_1}}\sigma^2_{\xi_{r_2}}}{|\boldsymbol{\Omega}|}\right). \tag{4.8}$$

$|\boldsymbol{\Omega}|$ denotes the determinant of the innovations' covariance matrix in the unrestricted model (Equation 4.2). Under the null hypotheses, these measures, multiplied by the number of observations, T, are asymptotically independent and follow chi-squared distributions with degrees of freedom 1, 1, 1 and 3, respectively.

Because the feedback measures are just log-likelihood ratio statistics under the null hypotheses, their asymptotic distributions are well defined. Also under the alternative hypotheses, these measures, multiplied by the number of observations, follow asymptotically non-central chi-squared distributions.

$$T\hat{F}_{r_1\to r_2} \sim \mathcal{X}'^2(1, TF_{r_1\to r_2}), \tag{4.9}$$

$$T\hat{F}_{r_2\to r_1} \sim \mathcal{X}'^2(1, TF_{r_2\to r_1}), \tag{4.10}$$

$$T\hat{F}_{r_1\leftrightarrow r_2} \sim \mathcal{X}'^2(1, TF_{r_1\leftrightarrow r_2}), \text{ and} \tag{4.11}$$

$$T\hat{F}_{r_1.r_2} \sim \mathcal{X}'^2(3, TF_{r_1.r_2}). \tag{4.12}$$

The measures presented above are additive, that is $F_{r_1.r_2} = F_{r_1\to r_2} + F_{r_1\leftrightarrow r_2} + F_{r_2\to r_1}$.

4.4.3. Granger Causality in Distribution

This subsection presents the test of Granger causality in distribution proposed by Candelon and Tokpavi (2016), from which the description bellow is heavily drawn. This test is based on the Value-at-Risk (VaR), which is a measure often used to assess the extent of loss of an asset or portfolio over a specific time frame. Considering $r_i = r_1, r_2$, the VaR at the α% confidence level is given by

$$\Pr\left[r_{i,t} < VaR_t^{r_i}(\boldsymbol{\theta}^0_{r_i}) \middle| \mathcal{F}^{r_i}_{t-1}\right] = \alpha. \tag{4.13}$$

$VaR_t^{r_i}$ is the Value-at-Risk of asset i at time t, $\boldsymbol{\theta}^0_{r_i}$ is the vector of true unknown finite-dimensional parameters related to the specification of the Value-at-Risk models for r_i, and $\mathcal{F}^{r_i}_{t-1}$ are the information sets at time $t-1$, defined as $\mathcal{F}^{r_i}_{t-1} = [r_{i,l},\ l \le t-1]$.

For each return series, a vector of VaRs at time t based on the previous equation may be defined as follows: First, let A be a set of risk levels, such that $A = \{\alpha_1, \alpha_2, \dots, \alpha_{m+1}\}$, for $0 < \alpha_1 < \alpha_2 <\dots \alpha_{m+1} < 1$, hence

$$VaR_{t,1}^{r_i}(\boldsymbol{\theta}_{r_i}^0, \alpha_1) < VaR_{t,2}^{r_i}(\boldsymbol{\theta}_{r_i}^0, \alpha_2) < \cdots < VaR_{t,m+1}^{r_i}(\boldsymbol{\theta}_{r_i}^0, \alpha_{m+1}). \tag{4.14}$$

Next, the variables $r_{i,t}$ are divided into m disjoint regions according to indicator variables that identify the events covering two consecutive VaR levels.

$$Z_{t,s}^{r_i}(\boldsymbol{\theta}_{r_i}^0) = \begin{cases} 1 & if\ VaR_{t,s}^{r_i}(\boldsymbol{\theta}_{r_i}^0, \alpha_s) \le r_{i,t} < VaR_{t,s+1}^{r_i}(\boldsymbol{\theta}_{r_i}^0, \alpha_{s+1}) \\ 0 & otherwise \end{cases} \ for\ s = 1, 2, \dots m. \tag{4.15}$$

Hence, to test the Granger causality in distribution we first define $\mathbf{H}_t^{r_i}$ as the vector containing the m indicator variables as defined in Equation (4.15).

$$\mathbf{H}_t^{r_i}(\boldsymbol{\theta}_{r_i}^0) = \left(Z_{t,1}^{r_i}(\boldsymbol{\theta}_{r_i}^0), Z_{t,2}^{r_i}(\boldsymbol{\theta}_{r_i}^0) \dots, Z_{t,m}^{r_i}(\boldsymbol{\theta}_{r_i}^0)\right)^T \tag{4.16}$$

Formally, $r_{2,t}$ does not Granger cause $r_{1,t}$ in distribution if the following hypothesis holds:

$$\mathbb{H}_0: \mathbb{E}\left[\mathbf{H}_t^{r_1}(\boldsymbol{\theta}_{r_1}^0) | \mathcal{F}_{t-1}^{r_1\,\&\,r_2}\right] = \mathbb{E}\left[\mathbf{H}_t^{r_1}(\boldsymbol{\theta}_{r_1}^0) | \mathcal{F}_{t-1}^{r_1}\right] \tag{4.17}$$

where $\mathcal{F}_{t-1}^{r_1\,\&\,r_2} = \{(r_{1,l}, r_{2,l}), l \le t-1\}$. Under the null, the information set related to variable r_{2t} does not provide any additional information to predict $\mathbf{H}_t^{r_1}(\boldsymbol{\theta}_{r_1}^0)$, beyond the information present in the distribution support of $r_{1,t}$.

To test the hypothesis presented in Equation (4.17), we estimate the conditional autoregressive Value-at-Risk (CAViaR) by Engle and Manganelli (2004) for each series and risk level, which is defined as follows:

$$VaR_{t,s}^{r_i}(\boldsymbol{\theta}_{r_i}^0, \alpha_s) = \theta_{s0}^{r_i} + \theta_{s1}^{r_i} VaR_{t-1}^{r_i}(\boldsymbol{\theta}_{r_i}^0) + \theta_{s2}^{r_i} r_{i,t-1}^{+} + \theta_{s3}^{r_i} r_{i,t-1}^{-} + \varepsilon_{t,s} \tag{4.18}$$

for which $r_{i,t-1}^{+} = \max(r_{i,t-1}, 0)$, $r_{i,t-1}^{-} = -\min(r_{i,t-1}, 0)$, for s levels of risk. The error terms, $\varepsilon_{t,s}$, conditional on all past information form a stationary process, with continuous conditional density. The parameters of these CAViaR models are estimated by quantile regression. Using the estimated VaRs, next the empirical counterparts of $\mathbf{H}_t^{r_i}$ can be computed, obtaining $\widehat{\mathbf{H}}_t^{r_i} \equiv H_t^{r_i}(\hat{\theta}_1^{r_i}, \dots, \hat{\theta}_m^{r_i})$.

The test statistic is obtained through the following 4 steps:

First, the sample cross-correlation matrix between $\widehat{\mathbf{H}}_t^{r_1}$ and $\widehat{\mathbf{H}}_t^{r_2}$ is computed as:

$$\widehat{\mathbf{\Lambda}}(j) \equiv \begin{cases} T^{-1} \sum_{t=1+j}^{T} \left(\widehat{\mathbf{H}}_t^{r_1} - \widehat{\mathbf{\Pi}}_{r_1}\right)\left(\widehat{\mathbf{H}}_{t-1}^{r_2} - \widehat{\mathbf{\Pi}}_{r_2}\right)^T & 0 \leq j \leq T-1 \\ T^{-1} \sum_{t=1-j}^{T} \left(\widehat{\mathbf{H}}_{t+j}^{r_1} - \widehat{\mathbf{\Pi}}_{r_1}\right)\left(\widehat{\mathbf{H}}_t^{r_2} - \widehat{\mathbf{\Pi}}_{r_2}\right)^T & 1-T \leq j \leq 0 \end{cases}, \tag{4.19}$$

where $\widehat{\mathbf{\Pi}}_{r_1}$ and $\widehat{\mathbf{\Pi}}_{r_2}$ represent the sample means of $\widehat{\mathbf{H}}_t^{r_1}$ and $\widehat{\mathbf{H}}_t^{r_2}$, respectively. Second, the corresponding sample cross-correlation matrix is computed as:

$$\widehat{\mathbf{R}}(j) = diag(\widehat{\mathbf{\Sigma}}_{r_1})^{-\frac{1}{2}}\widehat{\mathbf{\Lambda}}(j)diag(\widehat{\Sigma}_{r_2})^{-1/2}, \tag{4.20}$$

where $diag(.)$ is the diagonal form of a matrix, and $\widehat{\mathbf{\Sigma}}_{r_1}$and $\widehat{\mathbf{\Sigma}}_{r_2}$ are the sample covariance matrices of $\widehat{\mathbf{H}}_t^{r_1}$ and $\widehat{\mathbf{H}}_t^{r_2}$, respectively. Third, quadratic form that accounts for the dependence between current values of $\widehat{\mathbf{H}}_t^{r_1}$ and the lagged values of $\widehat{\mathbf{H}}_t^{r_2}$ is calculated by:

$$\hat{\mathcal{T}} = \sum_{j=1}^{T-1} \kappa^2 \left(\frac{j}{M}\right) \hat{Q}(j), \tag{4.21}$$

where κ is a kernel function, M is a truncation parameter and $\hat{Q}(j)$ is defined as:

$$\hat{Q}(j) = Tvec(\widehat{\mathbf{R}}(j))^T\left(\widehat{\mathbf{\Gamma}}_{r_1}^{-1} \otimes \widehat{\mathbf{\Gamma}}_{r_2}^{-1}\right)vec\left(\widehat{\mathbf{R}}(j)\right), \tag{4.22}$$

where $\widehat{\mathbf{\Gamma}}_{r_1}$ and $\widehat{\mathbf{\Gamma}}_{r_2}$ represent the sample correlation matrices of $\widehat{\mathbf{H}}_t^{r_1}$ and $\widehat{\mathbf{H}}_t^{r_2}$, respectively. Lastly, the test statistic - a centred and scaled version of the quadratic form $\hat{\mathcal{T}}$ - is given by:

$$V_{Y \to X} = \frac{\hat{\mathcal{T}} - m^2 C_T(M)}{(m^2 D_T(M))^{\frac{1}{2}}} \xrightarrow{d} \mathcal{N}(0,1) \tag{4.23}$$

where $C_T(M)$ and $D_T(M)$ are defined as:

$$C_T(M) = \sum_{j=1}^{T-1} (1 - \frac{j}{T})\kappa^2 \left(\frac{j}{M}\right), \tag{4.24}$$

$$D_T(M) = 2\sum_{j=1}^{T-1} (1 - \frac{j}{T})(1 - \frac{j+1}{T})\kappa^4 \left(\frac{j}{M}\right). \tag{4.25}$$

In the empirical application we use $M = ln(T)$ and the Barlett kernel.[10]

[10] As suggested in Candelon and Tokpavi (2016), we have tested the sensitivity of the results using $M = ln(T)$, $M = 1.5T^{0.3}$, and $M = 2T^{0.3}$, and different kernels functions (Bartlett, Daniell, and Parzen). Results remained

4.5. Empirical Results

This section presents the results on the relationship between industry returns from and to the US and other countries.

4.5.1. Causality and Feedback in the Mean

Table 4.3 shows the p-values of the Granger causality bivariate tests on returns considering unrestricted and restricted VAR(1) models.

Table 4.3: Granger causality in the mean

	BM	CD	CS	EN	FI	HC	IN	RE	TEC	TEL	UT
$US \rightarrow CN$	**0.001***	**0.000***	**0.005***	**0.004***	**0.009***	**0.009***	0.059	**0.002***	0.888	**0.001***	**0.000***
$CN \rightarrow US$	0.071	0.290	0.063	0.195	0.885	0.142	0.807	0.527	**0.001***	**0.006***	**0.025**
$US \rightarrow FR$	**0.000***	**0.043**	0.246	**0.000***	**0.048**	0.411	0.269	**0.008***	0.285	0.991	**0.000***
$FR \rightarrow US$	0.397	0.139	0.229	0.428	0.068	0.173	0.632	0.554	**0.003***	0.216	0.127
$US \rightarrow GE$	**0.000***	**0.000***	**0.017**	**0.001***	**0.026**	**0.001***	**0.013**	0.291	0.487	0.411	**0.008***
$GE \rightarrow US$	0.947	0.334	0.143	0.652	0.119	0.174	0.482	0.936	**0.001***	0.590	0.973
$US \rightarrow JP$	**0.000***	**0.000***	**0.001***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.001***	**0.001***
$JP \rightarrow US$	0.349	**0.010**	0.381	0.167	**0.035**	0.340	0.306	0.939	0.504	0.055	0.316
$US \rightarrow CH$	**0.000***	0.329	0.460	**0.000***	0.057	**0.046**	0.059	0.385	0.676	**0.007***	0.627
$CH \rightarrow US$	0.259	0.441	0.761	0.156	**0.001***	0.133	0.061	0.095	0.299	0.500	0.794
$US \rightarrow UK$	**0.000***	**0.000***	**0.001***	**0.000***	**0.002***	**0.006***	**0.001***	**0.000***	**0.000***	**0.043**	**0.049**
$UK \rightarrow US$	0.783	0.860	**0.000***	**0.001***	0.130	**0.010**	0.499	0.379	0.441	0.063	0.563

Notes: This table presents the p-values of the Granger causality test resulting from a bivariate VAR(1) applied to returns using data from 03/01/1973 to 12/05/2021. The tests are conducted pairwise between the US and other six developed countries (Canada (CN), France (FR), Germany (GE), Japan (JP), China CH), and the UK) for 11 industries (Basic Materials (BM), Consumer Discretionary (CD), Consumer Staples (CS), Energy (EN), Financials (FI), Health Care (HC), Industrials (IN), Real Estate (RE), Technology (TEC), Telecommunications (TEL), and Utilities (UT)). $US \rightarrow k$ indicates Granger causality from the US to the returns of country k, and $k \rightarrow US$ indicates Granger causality from country k to the US, for each industry. Numbers in bold indicate the rejection of the null hypothesis of no Granger causality at the 5% level. One asterisk, "*", denotes significance at the 1% level.

Results in Table 4.3 show that the US exhibit the leading role within each industry. From 66 causality tests $US \rightarrow k$, 49 are significant at the 5% level, from which 41 present a p-value less than 1%; while from the other 66 causality tests $k \rightarrow US$, only 11 are significant

broadly unchanged, hence we used $M = ln(T)$ and the Barlett kernel as these are the specifications usually used in the literature.

at the 5% level. This suggests that lagged returns in the US industries may contain relevant information to predict the returns of industries of other countries. This is especially visible for Japan, the UK, and Canada.

At the industry level, US Basic Materials (BM) and Energy (EN) show causality to all countries at the 1% significance level. This is possibly justified by the type of commodities produced by these industries, such as oil, metals, and coal which are highly export-oriented and whose shocks have historically led the global economy into a downturn (Venditti et al., 2020). The lagged US returns also contain relevant information to predict returns of non-US countries for Financials (FI), except for China. This is expected due to the high degree of financial sector integration worldwide and the fact that firms in many industries rely heavily on financial services and financial intermediaries, therefore, it is expected that they have a large impact on companies around the world (Rapach et al., 2015).

Table 4.4 reports the estimated pairwise Geweke feedback measures within each industry. Results indicate that there is a linear dependence between the US and the other countries for all the industries. The contemporaneous feedback is the major contributor to the total feedback, where the percentages range from 72%, for Utilities (UT) in Japan, to 99.5%, for Industrials (IN) in France, with an overall average value of 94.1%. These results suggest that most markets are highly integrated and that, on average, 94.1% of the return variability is transmitted within one week. The level of integration is weaker in Asian markets, which present the lowest values for the contemporaneous feedback.

For all industries the percentage of lagged feedback from the US to non-US countries is substantially higher than the feedback in the opposite direction. Therefore, the lagged feedback is asymmetrical and runs dominantly from the US to other countries, and in most cases is even unidirectional. We highlight the results for the lagged feedbacks from the US to Japan, which are significant at a 1% level and show high weights for all industries. The lagged feedback to Japan in Utilities, 26%, presents the highest value of lagged feedback across all countries and industries. In the opposite direction, the lagged feedbacks from non-US countries to the US are very marginal and, in most cases, not significant at a 5% level. We report a maximum significant relative value of 8.6% (0.005 in absolute terms) from China for the Financials Industry (FI).

Table 4.4: Geweke feedback measures in the mean

		BM	CD	CS	EN	FI	HC	IN	RE	TEC	TEL	UT
$F_{US\rightarrow k}$	**CN**	**0.007***	**0.016***	**0.004***	**0.003**	**0.003**	**0.005***	**0.004***	**0.005***	0.000	**0.004***	**0.008***
		1.0%	*3.8%*	*1.9%*	*0.5%*	*0.5%*	*2.7%*	*0.6%*	*2.4%*	*0.0%*	*3.2%*	*3.8%*
	FR	**0.007***	**0.003**	**0.002**	**0.009***	**0.002**	0.002	0.001	**0.003**	0.001	0.000	**0.006***
		2.0%	*1.2%*	*1.5%*	*2.4%*	*0.6%*	*0.9%*	*0.3%*	*3.3%*	*0.2%*	*0.0%*	*4.2%*
	GE	**0.009***	**0.009***	**0.004***	**0.005***	**0.002**	**0.009***	**0.003**	0.001	0.000	0.001	**0.003**
		2.5%	*5.1%*	*6.6%*	*3.9%*	*0.7%*	*5.2%*	*1.0%*	*0.9%*	*0.1%*	*0.6%*	*3.3%*
	JP	**0.012***	**0.015***	**0.005***	**0.012***	**0.007***	**0.014***	**0.020***	**0.008***	**0.020***	**0.005***	**0.006***
		9.7%	*10.1%*	*7.7%*	*10.4%*	*7.0%*	*19.2%*	*12.1%*	*15.9%*	*12.4%*	*13.9%*	*26.0%*
	CH	**0.016***	0.000	0.000	**0.009***	0.002	0.002	0.001	0.000	0.000	**0.003**	0.000
		12.9%	*1.2%*	*4.2%*	*10.1%*	*3.0%*	*7.4%*	*3.2%*	*4.1%*	*0.8%*	*8.5%*	*1.3%*
	UK	**0.009***	**0.008***	**0.005***	**0.011***	**0.005***	**0.004***	**0.007***	**0.009***	**0.008***	**0.002**	**0.003**
		1.9%	*3.3%*	*2.5%*	*1.9%*	*1.2%*	*1.5%*	*2.2%*	*8.4%*	*3.9%*	*1.4%*	*2.6%*
$F_{k\rightarrow US}$	**CN**	0.001	0.001	0.001	0.001	0.000	0.001	0.001	0.000	**0.005***	**0.003**	**0.002**
		0.2%	*0.2%*	*0.6%*	*0.2%*	*0.0%*	*0.7%*	*0.2%*	*0.1%*	*1.2%*	*2.0%*	*1.1%*
	FR	0.000	0.001	0.002	0.000	0.001	0.001	0.000	0.000	**0.004***	0.001	0.001
		0.1%	*0.5%*	*1.1%*	*0.1%*	*0.4%*	*0.4%*	*0.1%*	*0.1%*	*1.6%*	*0.6%*	*0.7%*
	GE	0.000	0.001	0.002	0.000	0.001	0.001	0.000	0.001	**0.004***	0.000	0.000
		0.0%	*0.3%*	*2.8%*	*0.4%*	*0.4%*	*0.4%*	*0.1%*	*0.8%*	*2.0%*	*0.2%*	*0.2%*
	JP	0.000	0.003	**0.000**	0.001	0.002	0.000	0.001	0.001	0.000	0.002	0.000
		0.2%	*1.9%*	*0.6%*	*1.0%*	*1.8%*	*0.6%*	*0.4%*	*1.9%*	*0.2%*	*4.5%*	*2.0%*
	CH	0.001	0.000	0.001	0.001	**0.005***	0.001	0.001	0.001	0.001	0.001	0.000
		0.7%	*1.5%*	*6.2%*	*1.5%*	*8.6%*	*5.9%*	*3.2%*	*11.1%*	*5.2%*	*3.6%*	*0.6%*
	UK	0.000	0.001	**0.006***	**0.005***	0.002	**0.003**	0.000	0.000	0.000	0.002	0.000
		0.0%	*0.4%*	*3.2%*	*0.9%*	*0.4%*	*1.3%*	*0.2%*	*0.4%*	*0.3%*	*1.1%*	*0.4%*
$F_{US\leftrightarrow k}$	**CN**	**0.729***	**0.407***	**0.195***	**0.729***	**0.591***	**0.170***	**0.584***	**0.219***	**0.397***	**0.134***	**0.207***
		98.8%	*95.9%*	*97.5%*	*99.3%*	*99.5%*	*96.6%*	*99.2%*	*97.5%*	*98.8%*	*94.8%*	*95.1%*
	FR	**0.325***	**0.259***	**0.141***	**0.367***	**0.354***	**0.194***	**0.304***	**0.095***	**0.251***	**0.110***	**0.142***
		97.9%	*98.3%*	*97.4%*	*97.5%*	*99.0%*	*98.7%*	*99.5%*	*96.5%*	*98.2%*	*99.4%*	*95.1%*
	GE	**0.366***	**0.166***	**0.055***	**0.122***	**0.327***	**0.167***	**0.339***	**0.068***	**0.219***	**0.115***	**0.083***
		97.5%	*94.6%*	*90.7%*	*95.7%*	*98.9%*	*94.4%*	*99.0%*	*98.3%*	*97.9%*	*99.2%*	*96.5%*
	JP	**0.110***	**0.130***	**0.064***	**0.099***	**0.093***	**0.059***	**0.142***	**0.041***	**0.140***	**0.029***	**0.017***
		90.1%	*88.1%*	*91.7%*	*88.6%*	*91.2%*	*80.2%*	*87.5%*	*82.2%*	*87.4%*	*81.6%*	*72.0%*
	CH	**0.105***	**0.033***	**0.008***	**0.078***	**0.052***	**0.020***	**0.041***	**0.009***	**0.020***	**0.030***	**0.025***
		86.3%	*97.4%*	*89.6%*	*88.4%*	*88.4%*	*86.8%*	*93.6%*	*84.7%*	*94.0%*	*87.9%*	*98.1%*
	UK	**0.451***	**0.238***	**0.189***	**0.569***	**0.408***	**0.242***	**0.292***	**0.099***	**0.184***	**0.162***	**0.097***
		98.0%	*96.4%*	*94.3%*	*97.2%*	*98.4%*	*97.2%*	*97.7%*	*91.3%*	*95.8%*	*97.6%*	*97.0%*
$F_{US,k}$	**CN**	**0.738***	**0.424***	**0.200***	**0.734***	**0.594***	**0.176***	**0.589***	**0.225***	**0.402***	**0.141***	**0.218***
	FR	**0.332***	**0.264***	**0.145***	**0.376***	**0.358***	**0.196***	**0.306***	**0.099***	**0.256***	**0.110***	**0.149***
	GE	**0.375***	**0.175***	**0.061***	**0.128***	**0.330***	**0.177***	**0.343***	**0.069***	**0.223***	**0.116***	**0.086***
	JP	**0.122***	**0.148***	**0.070***	**0.111***	**0.102***	**0.074***	**0.162***	**0.050***	**0.160***	**0.036***	**0.024***
	CH	**0.122***	**0.033***	**0.009***	**0.088***	**0.059***	**0.023***	**0.044***	**0.011***	**0.022***	**0.034***	**0.026***
	UK	**0.460***	**0.247***	**0.201***	**0.586***	**0.415***	**0.249***	**0.299***	**0.109***	**0.192***	**0.166***	**0.100***

Notes: This table presents the Geweke feedback measures resulting from a bivariate VAR(1) using data from 03/01/1973 to 12/05/2021. The tests are conducted pairwise between the US and other six developed countries (Canada (CN), France (FR), Germany (GE), Japan (JP), China CH), and the UK) for 11 industries (Basic Materials (BM), Consumer Discretionary (CD), Consumer Staples (CS), Energy (EN), Financials (FI), Health Care (HC), Industrials (IN), Real Estate (RE), Technology (TEC), Telecommunications (TEL) and Utilities (UT)). $\boldsymbol{F_{US\rightarrow k}}$ is the measure of lagged feedback from the US to country k, $\boldsymbol{F_{k\rightarrow US}}$ is the measure of lagged feedback from country k to the US, $\boldsymbol{F_{US\leftrightarrow k}}$ is the measure of contemporaneous feedback, and $\boldsymbol{F_{US,k}}$ is the measure of total feedback. Numbers in bold indicate the rejection of the null of no feedback at the 5% level. One asterisk, "*", denotes significance at the 1% level. Numbers in italic represent the weight of the lagged and contemporaneous feedbacks to the total feedback.

Given the relevance of US Basic Materials (BM) and Energy (EN) industries reported in Table 4.3, we present in Table 4.5 the cross-industry Granger causality tests between these two US industries and all industries of other countries.[11]

Table 4.5: Cross-industry Granger causality in the mean from the US to other countries

	BM	**CD**	**CS**	**EN**	**FI**	**HC**	**IN**	**RE**	**TEC**	**TEL**	**UT**
BM	**0.001***	**0.000***	0.197	**0.000***	**0.001***	**0.004***	**0.023**	**0.000***	0.882	**0.000***	**0.000***
EN	0.478	**0.000***	0.340	**0.004***	**0.045**	**0.020**	0.143	**0.001***	0.872	**0.010***	**0.003***
BM	**0.000***	0.080	0.773	**0.024**	**0.007***	0.203	**0.020**	**0.001***	0.412	**0.010**	**0.000***
EN	**0.001***	0.912	0.305	**0.000***	**0.002***	0.369	0.097	**0.001***	0.229	0.213	**0.000***
BM	**0.000***	**0.000***	0.117	**0.001***	**0.003***	**0.001***	**0.000***	0.055	0.510	0.787	**0.001***
EN	**0.000***	**0.000***	0.083	**0.001***	**0.008***	**0.000***	**0.002***	0.199	0.912	0.926	**0.003***
BM	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.001***	**0.000***	0.207	0.220
EN	**0.000***	**0.011**	**0.014**	**0.000***	**0.004***	**0.001***	**0.000***	**0.011**	**0.000***	**0.043**	0.284
BM	**0.000***	**0.006***	0.473	**0.000***	**0.000***	**0.000***	**0.000***	**0.001***	0.094	**0.000***	0.053
EN	**0.000***	0.052	0.531	**0.000***	**0.000***	**0.000***	**0.001***	**0.000***	0.159	**0.000***	**0.003***
BM	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.028**	0.130	**0.015**
EN	**0.008***	**0.001***	**0.000***	**0.000***	**0.000***	**0.001***	**0.000***	**0.001***	**0.025**	**0.001***	**0.002***

Notes: This table presents the p-values of the Granger causality test resulting from a bivariate VAR(1) using data from 03/01/1973 to 12/05/2021. The tests are conducted pairwise considering the causality from the US Basic Materials (BM) and Energy (EN) industries to the industries of other six developed countries (Canada (CN), France (FR), Germany (GE), Japan (JP), China CH), and the UK). There are 11 industries of the other countries: Basic Materials (BM), Consumer Discretionary (CD), Consumer Staples (CS), Energy (EN), Financials (FI), Health Care (HC), Industrials (IN), Real Estate (RE), Technology (TEC), Telecommunications (TEL), and Utilities (UT). Numbers in bold indicate the rejection of the null hypothesis of no Granger causality at the 5% level. One asterisk, "*", denotes significance at the 1% level.

From Table 4.5, we can observe that the US returns of Basic Materials (BM) Granger cause 50 out of 66 series industry/country series, i.e., causality runs in more than 75%, while the US returns of Energy (EN) Granger cause 48 series, corresponding to more than 72% of the times. These findings are highly plausible since firms in other industries rely heavily on commodities and fuels (Venditti et al., 2020; Khalfaoui et al., 2021). Further, the lagged returns for commodity- and material-producing sectors placed earlier in the production chain are frequently strongly connected to returns of industries positioned later in the production chain (Rapach et al., 2015). This outcome is consistent with commodity positive price shocks increasing product prices and returns for industries in earlier stages of production while reducing profit margins and dropping returns for industries positioned in later production phases. Due to the overall economic interdependence, a positive cash flow shock in one industry has implications on cash flows in other industries. However,

[11] We have verified if, in fact, these two industries presented the most important results for cross-industry causality, which was confirmed. For example, the lagged US returns in the US Financial or the Technology industry only Granger cause, on average, 5 out of 11 industries.

information-processing limitations inhibit investors in other industries from quickly adjusting equity prices to the full impact of the cash flows, leading to cross-industry return predictability.

In sum, our findings suggest that the US is the dominant market in terms of transmission of information in most industries, except for China. The leading role of the US is justified by the US economy being the world's largest in terms of GDP and an important trading partner for many countries. Also, the US financial market exhibits the world's largest market capitalization. According to data from the World Bank (2021), in 2019, the market capitalization of listed domestic companies was 33.890 trillion dollars, around 41% of the worldwide total. Furthermore, the US market is scrutinized by investors worldwide. For instance, the US indexes are often used as a benchmark in the fixed income markets since they offer both great breadths of coverage and length of historical data. These high coverage and attention from investors and analysts make the macroeconomic fundamentals from the US market to impact gradually across international markets (Rizova et al., 2010, and Rapach et al., 2013).

Other possible explanations to the key role of US industries may relate to institutional holdings, market share, and trading volume.

Badrinath et al. (1995) have found that institutional ownership of firms influences the lead-lag role of a firm. This relates to the "prudent man" rule that governs the investment behaviour of institutional portfolio managers. According to this rule, portfolio managers are required to make "prudent" investments. As a result, institutional investors are compelled to invest in just a subset of tradable assets. Badrinath et al. (1995) found that when firms are owned by institutions, they normally have a leading role over non-institutional firms. According to the World Bank (2021), in 2019, in the US market, most firms are institutionally owned, which may influence the leading role of this country.

Generally, new information has a greater influence on industry leaders with a large market share. Because of market frictions, this information may not be immediately incorporated into the prices of other firms. As a result, there is a lead-lag relationship between industry leaders and followers. According to Lo and MacKinlay (1990), Brennan et al. (1993), and Hou et al. (2007), this slow transmission of information can be attributed to a variety of factors, such as incomplete markets and constrained stock market participation, information asymmetries, noise traders, limited investor attention, transaction costs, short-sale constraints, legal constraints faced by institutional investors, and other forms of market frictions and institutional constraints.

Finally, Chordia and Swaminathan (2000) argued that the trading volume is a key driver of the lead-lag pattern detected in stock markets, as low-volume stocks tend to adjust more slowly to information than high-volume stocks. According to the World Bank (2021), the US reported in 2019 a stock trading volume of around 23.192 trillion dollars (the highest national value), while China reported 18.248 trillion dollars, Japan 5.097 trillion dollars, the UK 2.357 trillion dollars, Canada 1.432 trillion dollars, Germany 1.350 trillion dollars, and France, 1.168 trillion dollars.

4.5.2. Causality and Feedback in Volatility

This subsection examines the causal and lead-lag relationships between the volatilities of industries from and to the US and other countries (Canada, France, Germany, Japan, China, and the UK). The weekly series of volatilities are constructed using the standard deviation of daily returns within a week. The metrics are obtained from unrestricted and restricted VAR(1).

Table 4.6: Granger causality in volatility

	BM	**CD**	**CS**	**EN**	**FI**	**HC**	**IN**	**RE**	**TEC**	**TEL**	**UT**
$US \rightarrow CN$	**0,000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.004***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***
$CN \rightarrow US$	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.012**	**0.000***	0.060	**0.000***	**0.000***	**0.000***
$US \rightarrow FR$	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***
$FR \rightarrow US$	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***
$US \rightarrow GE$	**0.000***	**0.000***	**0.000***	**0.038**	**0.000***	**0.000***	**0.000***	0.465	**0.000***	**0.000***	**0.000***
$GE \rightarrow US$	**0.000***	**0.001***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	0.367	**0.000***	**0.000***	**0.000***
$US \rightarrow JP$	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.002***
$JP \rightarrow US$	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.003***	**0.000***	**0.001***	0.640
$US \rightarrow CH$	**0.000***	**0.001***	0.056	**0.000***	**0.000***	0.881	**0.001***	0.494	**0.037**	0.094	**0.000***
$CH \rightarrow US$	**0.000***	**0.000***	0.256	**0.000***	**0.000***	0.843	**0.000***	0.423	0.432	**0.000***	**0.000***
$US \rightarrow UK$	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***
$UK \rightarrow US$	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***

Notes: This table presents the p-values of the Granger causality test resulting from a bivariate VAR(1) applied to weekly volatilities using data from 03/01/1973 to 12/05/2021. The tests are conducted pairwise between the US and other six developed countries (Canada (CN), France (FR), Germany (GE), Japan (JP), China CH), and the UK) for 11 industries (Basic Materials (BM), Consumer Discretionary (CD), Consumer Staples (CS), Energy (EN), Financials (FI), Health Care (HC), Industrials (IN), Real Estate (RE), Technology (TEC), Telecommunications (TEL), and Utilities (UT)). $US \rightarrow k$ indicates Granger causality from the US to the returns of country k, and $k \rightarrow US$ indicates Granger causality from country k to the US, for each industry. Numbers in bold indicate the rejection of the null hypothesis of no Granger causality at the 5% level. One asterisk, “*”, denotes significance at the 1% level.

Table 4.6 reports the p-values of the Granger causality tests for bivariate VAR(1) applied to weekly standard deviations of eleven industries. As expected, causality in volatility is more pronounced than causality in the mean. Only 12 out of 132 tests are not significant at the 5% level, and for the $US \rightarrow k$ tests only 5 are not statistically significant at the 5% level. Hence, the causal relationship from the US is very strong, with most statistics showing a p-value lower than 1%. Thus, the US volatility is an important leading indicator of the industry volatility in other countries. However, the causal relationship is less asymmetric than in the mean. For instance, for France and the UK, there is causality in both directions for all industries with a significance level of 1%. Also, Canadian and Japanese volatilities cause the US volatilities, except for Real Estate (RE) and Utilities (UT). In Germany, there is causality from and to the US except for Real Estate (RE). Once again, China is the country that presents the lowest number of significant causal relationships (in both directions), nevertheless, the volatility in the US Granger causes 7 out of 11 Chinese industries.

The pairwise relations in volatility between the US and non-US countries are further analysed by the Geweke feedback measures reported in Table 4.7.

As expected, feedback measures applied to volatilities are more significant than the correspondent figures for returns (see Table 4.4). But once again, the percentage lagged feedback in the volatility from the US to the other six countries is, in general, higher than the feedback in the opposite direction. The only exception is China, where the lagged feedback is higher for most industries than the feedback from the US, with an average value of 21%, in relative terms (average value across Chinese industries). This suggests that information on volatility flows mainly from the Chinese market to the US market. This situation also appears in some industries for France, Germany, and the UK. Nevertheless, results show that lagged feedback in volatility is globally asymmetric and predominantly dominated by the US.

Table 4.7: Geweke feedback measures in volatility

		BM	CD	CS	EN	FI	HC	IN	RE	TEC	TEL	UT
	CN	**0.039***	**0.090***	**0.068***	**0.031***	**0.086***	**0.003***	**0.064***	**0.010***	**0.033***	**0.043***	**0.068***
		8.8%	27.4%	26.9%	5.3%	17.1%	3.6%	14.4%	13.5%	11.7%	24.2%	27.9%
	FR	**0.093***	**0.075***	**0.054***	**0.063***	**0.084***	**0.059***	**0.079***	**0.052***	**0.064***	**0.007***	**0.014***
		36.4%	35.3%	39.1%	24.3%	31.0%	33.5%	33.2%	46.4%	37.4%	7.6%	10.0%
	GE	**0.084***	**0.060***	**0.029***	0.002	**0.079***	**0.047***	**0.060***	0.000	**0.007***	**0.058***	**0.038***
		28.6%	49.1%	38.9%	2.0%	29.3%	23.5%	21.8%	2.9%	6.8%	30.3%	30.6%
$\boldsymbol{F_{US\to k}}$	**JP**	**0.042***	**0.038***	**0.029***	**0.034***	**0.030***	**0.030***	**0.041***	**0.014***	**0.041***	**0.013***	**0.004***
		26.8%	26.0%	24.9%	37.8%	28.6%	23.3%	26.1%	33.6%	31.7%	35.4%	9.0%
	CH	**0.011***	**0.004***	0.001	**0.011***	**0.016***	0.000	**0.004***	0.000	0.002	0.001	**0.017***
		14.9%	14.7%	14.1%	21.7%	31.4%	0.1%	12.5%	7.7%	13.8%	5.0%	27.5%
	UK	**0.037***	**0.042***	**0.035***	**0.067***	**0.101***	**0.046***	**0.062***	**0.084***	**0.035***	**0.040***	**0.039***
		12.6%	22.1%	17.9%	18.8%	31.9%	22.7%	31.1%	41.0%	29.6%	21.6%	28.0%
	CN	**0.034***	**0.012***	**0.007***	**0.026***	**0.010***	**0.003***	**0.014***	0.001	**0.017***	**0.022***	**0.017***
		7.7%	3.6%	2.6%	4.4%	2.0%	2.8%	3.3%	1.9%	5.9%	12.8%	7.0%
	FR	**0.022***	**0.025***	**0.012***	**0.027***	**0.018***	**0.018***	**0.014***	**0.005***	**0.026***	**0.035***	**0.056***
		8.5%	11.7%	9.0%	10.2%	6.6%	10.1%	6.0%	4.3%	15.2%	37.5%	40.7%
	GE	**0.018***	**0.004***	**0.007***	**0.028***	**0.013***	**0.013***	**0.024***	0.000	**0.026***	**0.037***	**0.013***
		6.2%	3.4%	9.0%	33.5%	4.9%	6.5%	8.8%	4.5%	25.0%	19.4%	10.6%
$\boldsymbol{F_{k\to US}}$	**JP**	**0.009***	**0.015***	**0.012***	**0.009***	**0.006***	**0.007***	**0.014***	**0.004***	**0.023***	**0.004***	0.000
		5.9%	10.5%	10.0%	10.2%	5.5%	5.5%	8.7%	8.4%	17.7%	10.9%	0.2%
	CH	**0.015***	**0.006***	0.001	**0.016***	**0.015***	0.000	**0.013***	0.000	0.000	**0.010***	**0.022***
		18.8%	20.9%	5.0%	31.2%	29.4%	0.1%	37.0%	10.5%	2.0%	43.2%	36.8%
	UK	**0.033***	**0.024***	**0.016***	**0.043***	**0.023***	**0.018***	**0.007***	**0.049***	**0.021***	**0.041***	**0.028***
		11.3%	12.3%	8.3%	12.0%	7.2%	9.2%	3.3%	23.8%	17.5%	22.1%	19.9%
	CN	**0.371***	**0.228***	**0.179***	**0.521***	**0.407***	**0.084***	**0.365***	**0.063***	**0.234***	**0.111***	**0.158***
		83.5%	69.0%	70.5%	90.2%	80.9%	93.6%	82.3%	84.6%	82.5%	63.0%	65.2%
	FR	**0.141***	**0.112***	**0.072***	**0.171***	**0.169***	**0.099***	**0.144***	**0.056***	**0.081***	**0.051***	**0.067***
		55.1%	53.0%	51.9%	65.5%	62.4%	56.4%	60.9%	49.3%	47.4%	54.9%	49.3%
	GE	**0.192***	**0.058***	**0.038***	**0.055***	**0.179***	**0.140***	**0.190***	**0.007***	**0.071***	**0.097***	**0.072***
		65.1%	47.5%	52.1%	64.4%	65.9%	70.0%	69.4%	92.6%	68.2%	50.4%	58.8%
$\boldsymbol{F_{US\leftrightarrow k}}$	**JP**	**0.105***	**0.093***	**0.077***	**0.047***	**0.070***	**0.093***	**0.103***	**0.025***	**0.065***	**0.020***	**0.039***
		67.2%	63.5%	65.1%	52.0%	65.9%	71.2%	65.2%	58.0%	50.6%	53.7%	90.8%
	CH	**0.051***	**0.019***	**0.008***	**0.024***	**0.020***	**0.017***	**0.018***	0.002	**0.011***	**0.012***	**0.021***
		66.3%	64.4%	80.9%	47.2%	39.3%	99.9%	50.5%	81.8%	84.3%	51.8%	35.7%
	UK	**0.223***	**0.126***	**0.145***	**0.245***	**0.192***	**0.137***	**0.130***	**0.072***	**0.062***	**0.104***	**0.073***
		76.1%	65.6%	73.8%	69.2%	60.9%	68.2%	65.6%	35.2%	52.9%	56.3%	52.1%
	CN	**0.445***	**0.330***	**0.254***	**0.577***	**0.503***	**0.090***	**0.443***	**0.075***	**0.283***	**0.176***	**0.242***
	FR	**0.257***	**0.211***	**0.138***	**0.261***	**0.272***	**0.176***	**0.237***	**0.113***	**0.170***	**0.093***	**0.137***
	GE	**0.295***	**0.122***	**0.073***	**0.085***	**0.271***	**0.200***	**0.274***	**0.007***	**0.104***	**0.192***	**0.123***
$\boldsymbol{F_{US,k}}$	**JP**	**0.156***	**0.147***	**0.118***	**0.091***	**0.106***	**0.130***	**0.158***	**0.043***	**0.128***	**0.037***	**0.043***
	CH	**0.077***	**0.029***	**0.010***	**0.050***	**0.050***	**0.017***	**0.036***	0.002	**0.013***	**0.022***	**0.060***
	UK	**0.294***	**0.192***	**0.196***	**0.354***	**0.316***	**0.201***	**0.199***	**0.206***	**0.117***	**0.185***	**0.141***

Notes: This table presents the Geweke feedback measures resulting from a bivariate VAR(1) applied to weekly volatilities using data from 03/01/1973 to 12/05/2021. The tests are conducted pairwise between the US and other six developed countries (Canada (CN), France (FR), Germany (GE), Japan (JP), China CH), and the UK) for 11 industries (Basic Materials (BM), Consumer Discretionary (CD), Consumer Staples (CS), Energy (EN), Financials (FI), Health Care (HC), Industrials (IN), Real Estate (RE), Technology (TEC), Telecommunications (TEL) and Utilities (UT)). $\boldsymbol{F_{US\to k}}$ is the measure of lagged feedback from the US to country k, $\boldsymbol{F_{k\to US}}$ is the measure of lagged feedback from country k to the US, $\boldsymbol{F_{US\leftrightarrow k}}$ is the measure of contemporaneous feedback, and $\boldsymbol{F_{US,k}}$ is the measure of total feedback. Numbers in bold indicate the rejection of the null of no feedback at the 5% level. One asterisk, "*", denotes significance at the 1% level. Numbers in italic represent the weight of the lagged and contemporaneous feedbacks to the total feedback.

The contemporaneous feedback is the main contributor to the total feedback, where percentages range from 35.2%, for the Real Estate (RE) in the UK, to 99.9%, for the Health Care (HC) in China, with a global average value of 65%. These results suggest that most markets are integrated and that, on average, 65% of the volatility is communicated within one week.

At the industry level, we highlight that the Basic Materials (BM), Energy (EN), and Financials (FI) industries report high levels of feedback transmission. This is justified by these industries containing the largest companies in the world where volatility shocks are more rapidly spread out (World Bank, 2021).

4.5.3. Causality and Feedback during Expansions and Recessions

This subsection performs an analysis on causality and feedback during expansion and recession periods in the US. These periods are identified using the NBER business cycle classification. Table 4.8 reports the p-values of Granger causality tests.

The transmission of information mainly flows from the US to other countries during expansion and recession periods, but it is visibly less pronounced during expansion periods, although the US continues to dominate other countries during expansionary periods. The decrease in causality extends across all countries. For instance, the US returns only Granger cause the Canadian returns in less than half of the industries, while, for the full sample, Granger cause Canadian returns in 10 out of 11 industries. A similar situation occurs for Germany and the UK, where the US returns only Granger cause 3 and 6 industries, respectively. The exception to this pattern is Japan, for which the US returns Granger cause most industry returns in expansion and recession periods. The differences in the causal relationships between countries during expansions and recessions are notorious, implying that more information transmission occurs during recession periods.

Table 4.8: Granger causality in the mean during expansions and recessions

	BM	CD	CS	EN	FI	HC	IN	RE	TEC	TEL	UT
					Expansions						
$US \rightarrow CN$	0.059	**0.000***	0.060	0.556	0.422	0.312	0.756	0.054	0.754	**0.021**	0.059
$CN \rightarrow US$	0.648	0.347	0.982	**0.041**	0.731	**0.034**	0.381	0.781	0.133	**0.014**	0.648
$US \rightarrow FR$	0.072	0.502	0.627	**0.004***	0.255	0.805	0.547	0.645	0.884	0.844	0.072
$FR \rightarrow US$	0.239	0.964	0.588	0.261	0.208	0.784	0.768	0.869	**0.004***	0.424	0.239
$US \rightarrow GE$	**0.006***	0.301	**0.042**	0.937	0.948	0.057	0.252	0.926	0.488	0.498	**0.006***
$GE \rightarrow US$	0.320	0.185	0.393	0.437	0.093	0.310	0.373	0.592	**0.000***	0.246	0.320
$US \rightarrow JP$	**0.001***	**0.000***	**0.032**	**0.000***	**0.028**	**0.000***	**0.000***	**0.024***	**0.000***	**0.016**	**0.001***
$JP \rightarrow US$	0.122	**0.002***	**0.038**	**0.007***	0.553	0.162	0.369	0.764	0.917	0.114	0.122
$US \rightarrow CH$	**0.000***	0.830	0.703	**0.001***	0.114	0.222	0.324	0.712	0.674	0.058	**0.000***
$CH \rightarrow US$	0.617	0.534	0.424	**0.020**	0.051	0.171	0.752	0.555	0.189	0.801	0.617
$US \rightarrow UK$	**0.010**	**0.009***	**0.031**	**0.000***	0.234	0.118	0.138	0.091	**0.000***	0.487	**0.010**
$UK \rightarrow US$	0.369	0.811	0.193	**0.001***	0.441	0.408	0.859	0.130	0.275	0.276	0.369
					Recessions						
$US \rightarrow CN$	**0,020**	**0,000***	0,069	**0.022**	0.056	**0.005***	**0.017**	0.101	0.894	**0.034**	**0.037**
$CN \rightarrow US$	0.071	0.568	**0.000***	0.641	0.408	0.364	0.680	0.360	**0.013**	0.227	0.172
$US \rightarrow FR$	**0.001***	0.095	0.642	**0.016**	**0.001***	0.674	0.087	**0.002***	0.208	0.970	**0.003***
$FR \rightarrow US$	0.768	**0.017**	**0.002***	0.929	0.279	**0.033**	0.604	0.873	0.318	0.441	0.167
$US \rightarrow GE$	**0.001***	**0.000***	0.372	**0.001***	**0.005***	**0.006***	0.098	0.203	0.977	**0.013**	**0.017**
$GE \rightarrow US$	0.428	0.941	0.073	0.569	0.967	0.260	0.957	0.759	0.554	0.849	0.551
$US \rightarrow JP$	**0.000***	**0.000***	**0.027**	**0.001***	**0.000***	**0.000***	**0.000***	**0.000***	**0.000***	**0.008***	**0.050**
$JP \rightarrow US$	0.477	0.795	0.353	0.329	**0.018**	0.618	0.837	0.678	0.529	0.284	0.800
$US \rightarrow CH$	**0.000***	0.279	0.397	**0.004***	0.197	0.131	0.106	0.110	0.816	**0.047**	0.917
$CH \rightarrow US$	**0.030**	0.569	0.308	0.782	**0.006***	0.692	**0.037**	**0.005***	0.758	0.116	0.104
$US \rightarrow UK$	**0.003***	**0.009***	**0.047**	**0.015**	**0.024**	**0.040**	**0.014**	**0.000***	0.058	**0.012**	0.126
$UK \rightarrow US$	0.590	0.989	**0.001***	0.346	0.274	**0.007***	0.922	0.689	0.605	0.101	0.660

Notes: This table presents the p-values of the Granger causality test resulting from a bivariate VAR(1) applied to weekly returns using data in periods identified as expansions and recessions in the US according to the NBER business cycle classification (https://www.nber.org/research/data/us-business-cycle-expansions-and-contractions). The tests are conducted pairwise between the US and other six developed countries (Canada (CN), France (FR), Germany (GE), Japan (JP), China CH), and the UK) for 11 industries (Basic Materials (BM), Consumer Discretionary (CD), Consumer Staples (CS), Energy (EN), Financials (FI), Health Care (HC), Industrials (IN), Real Estate (RE), Technology (TEC), Telecommunications (TEL) and Utilities (UT)). $US \rightarrow k$ indicates Granger causality from the US to the returns of country k, and $k \rightarrow US$ indicates Granger causality from country k to the US, for each industry. Numbers in bold indicate the rejection of the null hypothesis of no Granger causality at the 5% level. One asterisk, "*", denotes significance at the 1% level.

Table 4.9 and Table 4.10 show the feedback measures during expansion and recession periods, respectively. Total feedback is lower during expansions periods. The total feedback during expansions is on average 0.166, while during recessions, it is on average 0.316. During recessions, the average unidirectional feedback is also higher than during expansions (5.57% and 2.02%, respectively). However, there is a different pattern in the contemporaneous feedback. Despite remaining the dominant contributor to the total feedback, we observe that this feedback is higher during expansions than during recessions, in relative terms. The average relative contemporaneous feedback during expansions is 96%. It is 89% during recessions and 94% for the full sample period. This indicates that during an expansionary period, there is an increase of 2% in the transmission of information communicated between markets within one week in relation to a recession period.

In conclusion, we observe that, during a recession, the linear dependence increases but the time that countries take to adjust to new information is higher than during an expansion, suggesting that investors react with a larger delay. Arguably, during a recession, the levels of uncertainty tend to be higher, and the confidence of investors on the information signals tends to decrease.

Table 4.9: Geweke feedback measures in the mean during expansions

		BM	CD	CS	EN	FI	HC	IN	RE	TEC	TEL	UT
$\boldsymbol{F_{US\to k}}$	**CN**	0.002	**0.007***	0.002	0.000	0.000	0.000	0.000	0.002	0.000	**0.002**	**0.005***
		0.2%	*1.9%*	*0.8%*	*0.0%*	*0.1%*	*0.3%*	*0.0%*	*1.0%*	*0.0%*	*2.0%*	*3.9%*
	FR	**0.002**	0.000	0.000	**0.004***	0.001	0.000	0.000	0.000	0.000	0.000	0.001
		0.5%	*0.1%*	*0.1%*	*1.3%*	*0.2%*	*0.0%*	*0.1%*	*0.2%*	*0.0%*	*0.0%*	*1.1%*
	GE	**0.004***	0.000	0.002	0.000	0.000	**0.002**	0.001	0.000	0.000	0.000	0.000
		1.2%	*0.3%*	*3.6%*	*0.0%*	*0.0%*	*1.0%*	*0.2%*	*0.0%*	*0.1%*	*0.2%*	*0.9%*
	JP	**0.005***	**0.006***	**0.002**	**0.008***	**0.002**	**0.009***	**0.011***	**0.002**	**0.014***	**0.003**	**0.004**
		5.2%	*5.6%*	*4.1%*	*9.7%*	*3.5%*	*14.8%*	*8.6%*	*9.7%*	*9.6%*	*8.6%*	*18.0%*
	CH	**0.010***	0.000	0.000	**0.006***	**0.001**	0.001	0.000	0.000	0.000	**0.002**	0.000
		11.4%	*0.1%*	*1.0%*	*7.7%*	*4.3%*	*5.1%*	*1.6%*	*1.5%*	*0.7%*	*10.4%*	*1.3%*
	UK	**0.003***	**0.003***	**0.002***	**0.006**	0.001	0.001	0.001	0.001	**0.007***	0.000	0.000
		0.9%	*1.9%*	*1.3%*	*1.3%*	*0.2%*	*0.5%*	*0.5%*	*2.5%*	*4.6%*	*0.2%*	*0.5%*
$\boldsymbol{F_{k\to US}}$	**CN**	0.000	0.000	0.000	**0.002**	0.000	**0.002**	0.000	0.000	0.001	**0.003**	0.001
		0.0%	*0.1%*	*0.0%*	*0.3%*	*0.0%*	*1.2%*	*0.1%*	*0.0%*	*0.3%*	*2.3%*	*0.4%*
	FR	0.001	0.000	0.000	0.001	0.001	0.000	0.000	0.000	0.004	0.000	**0.008***
		0.2%	*0.0%*	*0.1%*	*0.2%*	*0.3%*	*0.0%*	*0.0%*	*0.0%*	*1.9%*	*0.3%*	*7.2%*
	GE	0.000	0.001	0.000	0.000	**0.001**	0.000	0.000	0.000	**0.006***	0.001	0.000
		0.2%	*0.5%*	*0.6%*	*0.5%*	*0.5%*	*0.3%*	*0.1%*	*0.4%*	*3.6%*	*0.6%*	*0.5%*
	JP	0.001	**0.004***	**0.002**	**0.003***	0.000	0.001	0.000	0.000	0.000	0.001	0.000
		1.2%	*3.8%*	*3.8%*	*4.3%*	*0.3%*	*1.6%*	*0.3%*	*0.2%*	*0.0%*	*3.7%*	*2.1%*
	CH	0.000	0.000	0.000	**0.003**	**0.002**	0.001	0.000	0.000	0.001	0.000	**0.001**
		0.1%	*0.8%*	*4.3%*	*3.5%*	*6.5%*	*6.4%*	*0.2%*	*3.8%*	*6.4%*	*0.2%*	*7.2%*
	UK	0.000	0.000	0.001	**0.005***	0.000	0.000	0.000	0.001	0.001	0.001	0.000
		0.1%	*0.0%*	*0.5%*	*1.0%*	*0.1%*	*0.1%*	*0.0%*	*2.0%*	*0.4%*	*0.4%*	*0.0%*
$\boldsymbol{F_{US\leftrightarrow k}}$	**CN**	**0.680***	**0.374***	**0.192***	**0.577***	**0.494***	**0.167***	**0.502***	**0.176***	**0.398***	**0.118***	**0.126***
		99.7%	*98.0%*	*99.2%*	*99.6%*	*99.9%*	*98.5%*	*99.9%*	*99.0%*	*99.7%*	*95.7%*	*95.7%*
	FR	**0.279***	**0.213***	**0.122***	**0.279***	**0.268***	**0.180***	**0.228***	**0.049***	**0.202***	**0.101***	**0.102***
		99.2%	*99.9%*	*99.8%*	*98.5%*	*99.5%*	*100%*	*99.9%*	*99.8%*	*98.1%*	*99.7%*	*91.7%*
	GE	**0.289***	**0.150***	**0.052***	**0.062***	**0.248***	**0.168***	**0.261***	**0.035***	**0.165***	**0.104***	**0.045***
		98.6%	*99.1%*	*95.8%*	*99.5%*	*99.5%*	*98.7%*	*99.6%*	*99.6%*	*96.3%*	*99.2%*	*98.7%*
	JP	**0.089***	**0.103***	**0.049***	**0.067***	**0.061***	**0.049***	**0.121***	**0.022***	**0.133***	**0.028***	**0.017***
		93.7%	*90.6%*	*92.1%*	*86.0%*	*96.2%*	*83.6%*	*91.1%*	*90.2%*	*90.3%*	*87.7%*	*79.9%*
	CH	**0.077***	**0.024***	**0.007***	**0.064***	**0.024***	**0.012***	**0.027***	**0.004***	**0.012***	**0.014***	**0.017***
		88.5%	*99.2%*	*94.8%*	*88.8%*	*89.3%*	*88.4%*	*98.2%*	*94.8%*	*92.9%*	*89.4%*	*91.5%*
	UK	**0.349***	**0.167***	**0.169***	**0.460***	**0.313***	**0.231***	**0.206***	**0.051***	**0.139***	**0.138***	**0.069***
		99.0%	*98.1%*	*98.3%*	*97.6%*	*99.7%*	*99.4%*	*99.5%*	*95.5%*	*95.0%*	*99.4%*	*99.5%*
$\boldsymbol{F_{US,k}}$	**CN**	**0.682***	**0.382***	**0.194***	**0.580***	**0.495***	**0.170***	**0.503***	**0.178***	**0.399***	**0.124***	**0.131***
	FR	**0.281***	**0.213***	**0.122***	**0.284***	**0.270***	**0.181***	**0.228***	**0.049***	**0.206***	**0.101***	**0.111***
	GE	**0.293***	**0.152***	**0.054***	**0.062***	**0.249***	**0.170***	**0.262***	**0.035***	**0.172***	**0.105***	**0.046***
	JP	**0.095***	**0.114***	**0.053***	**0.078***	**0.063***	**0.058***	**0.132***	**0.025***	**0.147***	**0.032***	**0.022***
	CH	**0.087***	**0.024***	**0.007***	**0.072***	**0.027***	**0.014***	**0.028***	**0.004***	**0.013***	**0.016***	**0.019***
	UK	**0.352***	**0.170***	**0.171***	**0.472***	**0.314***	**0.232***	**0.207***	**0.053***	**0.146***	**0.139***	**0.070***

Notes: This table presents the Geweke feedback measures resulting from a bivariate VAR(1) applied to weekly returns using data in periods identified as expansions in the US according to the NBER business cycle classification (https://www.nber.org/research/data/us-business-cycle-expansions-and-contractions). The tests are conducted pairwise between the US and other six developed countries (Canada (CN), France (FR), Germany (GE), Japan (JP), China CH), and the UK) for 11 industries (Basic Materials (BM), Consumer Discretionary (CD), Consumer Staples (CS), Energy (EN), Financials (FI), Health Care (HC), Industrials (IN), Real Estate (RE), Technology (TEC), Telecommunications (TEL) and Utilities (UT)). $\boldsymbol{F_{US\to k}}$ is the measure of lagged feedback from the US to country k, $\boldsymbol{F_{k\to US}}$ is the measure of lagged feedback from country k to the US, $\boldsymbol{F_{US\leftrightarrow k}}$ is the measure of contemporaneous feedback, and $\boldsymbol{F_{US,k}}$ is the measure of total feedback. Numbers in bold indicate the rejection of the null of no feedback at the 5% level. One asterisk, "*", denotes significance at the 1% level. Numbers in italic represent the weight of the lagged and contemporaneous feedbacks to the total feedback.

Table 4.10: Geweke feedback measures in the mean during recessions

		BM	CD	CS	EN	FI	HC	IN	RE	TEC	TEL	UT
$F_{US\rightarrow k}$	CN	**0.015**	**0.039***	**0.009**	**0.014**	**0.010**	**0.021***	**0.015**	**0.007**	**0.000**	**0.012**	**0.012**
		1.8%	*7.7%*	*3.8%*	*1.4%*	*1.3%*	*10.9%*	*1.9%*	*2.3%*	*0.0%*	*5.9%*	*2.7%*
	FR	**0.030***	**0.008**	0.001	**0.016**	**0.030***	0.000	0.008	**0.026***	0.004	0.000	0.024
		6.6%	*2.0%*	*0.3%*	*2.7%*	*5.0%*	*0.2%*	*1.6%*	*11.9%*	*1.1%*	*0.0%*	*9.8%*
	GE	**0.028***	**0.040***	0.002	**0.030***	**0.022***	**0.020***	**0.007**	0.004	0.000	**0.017**	**0.016**
		4.6%	*17.1%*	*2.9%*	*10.3%*	*4.0%*	*12.1%*	*1.4%*	*2.5%*	*0.0%*	*11.3%*	*7.1%*
	JP	**0.050***	**0.038***	**0.013**	**0.028***	**0.043***	**0.035***	**0.038***	**0.033***	**0.039***	**0.019***	0.010
		22.4%	*15.6%*	*10.6%*	*13.1%*	*16.0%*	*25.6%*	*16.8%*	*22.3%*	*20.7%*	*37.6%*	*31.2%*
	CH	**0.039***	0.003	0.002	**0.022***	0.005	0.006	0.007	0.007	0.000	**0.011**	0.000
		17.3%	*4.8%*	*10.9%*	*15.8%*	*2.1%*	*12.4%*	*6.6%*	*11.6%*	*0.2%*	*10.7%*	*0.0%*
	UK	**0.023***	**0.019***	**0.011**	**0.016**	**0.014**	**0.011**	**0.016**	**0.040***	**0.010**	**0.017**	0.006
		3.3%	*4.3%*	*3.8%*	*2.1%*	*2.3%*	*3.9%*	*2.8%*	*16.3%*	*2.6%*	*6.4%*	*2.9%*
$F_{k\rightarrow US}$	CN	**0.009**	0.001	**0.036***	0.001	0.002	0.002	0.000	0.002	**0.017**	0.004	0.005
		1.1%	*0.2%*	*15.2%*	*0.1%*	*0.2%*	*1.2%*	*0.1%*	*0.7%*	*4.1%*	*1.9%*	*1.2%*
	FR	0.000	**0.015**	**0.026***	0.000	0.003	**0.012**	0.001	0.000	0.003	0.002	0.005
		0.1%	*4.1%*	*12.2%*	*0.0%*	*0.5%*	*5.6%*	*0.1%*	*0.0%*	*0.7%*	*1.0%*	*2.1%*
	GE	0.002	0.000	**0.009**	0.001	0.000	0.003	0.000	0.000	0.001	0.000	0.001
		0.3%	*0.0%*	*11.7%*	*0.3%*	*0.0%*	*2.1%*	*0.0%*	*0.1%*	*0.2%*	*0.1%*	*0.4%*
	JP	0.001	0.000	0.002	0.003	**0.015**	0.001	0.000	0.000	0.001	0.003	0.000
		0.6%	*0.1%*	*1.9%*	*1.2%*	*5.6%*	*0.5%*	*0.1%*	*0.3%*	*0.6%*	*6.1%*	*0.5%*
	CH	**0.013**	0.001	0.003	0.000	**0.021***	0.000	**0.012**	**0.022***	0.000	0.007	0.007
		5.7%	*1.3%*	*15.8%*	*0.1%*	*9.6%*	*0.9%*	*11.1%*	*35.7%*	*0.4%*	*6.7%*	*11.9%*
	UK	0.001	0.000	**0.031***	0.002	0.003	**0.020***	0.000	0.000	0.001	0.007	0.001
		0.1%	*0.0%*	*10.9%*	*0.3%*	*0.5%*	*6.9%*	*0.0%*	*0.2%*	*0.2%*	*2.8%*	*0.2%*
$F_{US\leftrightarrow k}$	CN	**0.800***	**0.474***	**0.190***	**0.974***	**0.774***	**0.170***	**0.772***	**0.316***	**0.385***	**0.190***	**0.423***
		97.2%	*92.2%*	*81.0%*	*98.5%*	*98.5%*	*87.9%*	*98.0%*	*97.0%*	*95.9%*	*92.2%*	*96.2%*
	FR	**0.432***	**0.353***	**0.188***	**0.554***	**0.565***	**0.210***	**0.500***	**0.194***	**0.403***	**0.154***	**0.214***
		93.4%	*93.9%*	*87.6%*	*97.3%*	*94.4%*	*94.2%*	*98.3%*	*88.1%*	*98.3%*	*99.0%*	*88.0%*
	GE	**0.567***	**0.194***	**0.064***	**0.256***	**0.523***	**0.143***	**0.524***	**0.171***	**0.479***	**0.131***	**0.201***
		95.1%	*82.9%*	*85.4%*	*89.4%*	*96.0%*	*85.8%*	*98.6%*	*97.3%*	*99.8%*	*88.6%*	*92.4%*
	JP	**0.174***	**0.207***	**0.109***	**0.179***	**0.213***	**0.102***	**0.188***	**0.115***	**0.147***	**0.029***	**0.023***
		77.0%	*84.3%*	*87.5%*	*85.6%*	*78.4%*	*73.9%*	*83.2%*	*77.4%*	*78.7%*	*56.3%*	*68.3%*
	CH	**0.174***	**0.062***	**0.013***	**0.119***	**0.189***	**0.043***	**0.088***	**0.032***	**0.063***	**0.083***	**0.054***
		77.0%	*93.9%*	*73.3%*	*84.1%*	*88.3%*	*86.8%*	*82.3%*	*52.7%*	*99.4%*	*82.6%*	*88.1%*
	UK	**0.670***	**0.413***	**0.242***	**0.735***	**0.594***	**0.259***	**0.562***	**0.208***	**0.363***	**0.241***	**0.212***
		96.5%	*95.7%*	*85.3%*	*97.6%*	*97.2%*	*89.2%*	*97.2%*	*83.6%*	*97.2%*	*90.8%*	*96.8%*
$F_{US,k}$	CN	**0.824***	**0.515***	**0.235***	**0.989***	**0.786***	**0.194***	**0.788***	**0.326***	**0.402***	**0.207***	**0.440***
	FR	**0.463***	**0.376***	**0.215***	**0.569***	**0.599***	**0.223***	**0.509***	**0.220***	**0.410***	**0.156***	**0.244***
	GE	**0.596***	**0.234***	**0.075***	**0.287***	**0.544***	**0.167***	**0.531***	**0.176***	**0.480***	**0.147***	**0.217***
	JP	**0.225***	**0.245***	**0.124***	**0.210***	**0.271***	**0.138***	**0.227***	**0.149***	**0.187***	**0.051***	**0.033***
	CH	**0.226***	**0.066***	**0.018***	**0.142***	**0.214***	**0.050***	**0.107***	**0.060***	**0.063***	**0.100***	**0.061***
	UK	**0.694***	**0.431***	**0.283***	**0.753***	**0.611***	**0.290***	**0.578***	**0.249***	**0.374***	**0.265***	**0.218***

Notes: This table presents the Geweke feedback measures resulting from a bivariate VAR(1) applied to weekly returns using data in periods identified as recessions in the US according to the NBER business cycle classification (https://www.nber.org/research/data/us-business-cycle-expansions-and-contractions). The tests are conducted pairwise between the US and other six developed countries (Canada (CN), France (FR), Germany (GE), Japan (JP), China CH), and the UK) for 11 industries (Basic Materials (BM), Consumer Discretionary (CD), Consumer Staples (CS), Energy (EN), Financials (FI), Health Care (HC), Industrials (IN), Real Estate (RE), Technology (TEC), Telecommunications (TEL) and Utilities (UT)). $\boldsymbol{F_{US\rightarrow k}}$ is the measure of lagged feedback from the US to country k, $\boldsymbol{F_{k\rightarrow US}}$ is the measure of lagged feedback from country k to the US, $\boldsymbol{F_{US\leftrightarrow k}}$ is the measure of contemporaneous feedback, and $\boldsymbol{F_{US,k}}$ is the measure of total feedback. Numbers in bold indicate the rejection of the null of no feedback at the 5% level. One asterisk, "*", denotes significance at the 1% level. Numbers in italic represent the weight of the lagged and contemporaneous feedbacks to the total feedback.

4.5.4. Causality in Distribution

This subsection examines Granger causality in distribution, using the procedures proposed by Candelon and Tokpavi (2016).

Table 4.11 shows the p-values for the tests applied to the left and right tails of the distribution of returns. The tests in the left tail are conducted considering α = 1%, 5% and 10%, while the tests in the right tail consider α = 90%, 95% and 99%.

For the left-tail of the distribution, there is causality from US industries to other countries, at a 5% significance level, in 32 industries, while causality from other countries to the US happens in 22 industries. Japan is the country that exhibits the highest level of reaction to information coming from the US (7 Japanese industries are Granger caused by the corresponding US industries, at a 1% significance level). For the other countries, the number of industries that cause and are caused by US returns are almost always not very different. At the industry level, the Technology (TEC) industry is the one that presents more significant causalities between the US and other countries

Results on the causality in the right tail reveal less causality. Japan and Canada are the countries that exhibit the highest number of significant causalities from and to the US.

Table 4.12 reports the p-values for the left and right tails of the distribution in the volatilities. Generally speaking, in the left tail of the distribution causality is low, mainly coming from the US to the other countries. On average, the US leads 2 out of 11 industries for each country. This evidence suggests a small information transmission when volatility is low across economies. However, causality is higher in the right tail of the distribution, and mainly flows from the US to other countries. For instance, in France and Germany, 8 out of 11 industries react significantly to high volatilities in the US industries. In the case of Canada and the UK, these results also show that other countries do not timely incorporate high volatility shocks that affect the US industries.

Table 4.11: Granger causality in distribution of returns

	BM	CD	CS	EN	FI	HC	IN	RE	TEC	TEL	UT
					Left tail						
$US \rightarrow CN$	**0.000***	0.794	0.776	**0.002***	0.373	0.581	0.225	0.563	**0.036**	**0.022**	**0.027**
$CN \rightarrow US$	0.121	0.736	0.063	0.972	**0.002***	**0.032**	0.359	**0.001***	**0.000***	0.684	0.172
$US \rightarrow FR$	0.855	**0.005***	0.376	**0.005***	0.692	**0.000***	**0.000***	0.835	**0.000***	0.547	0.521
$FR \rightarrow US$	0.084	**0.000***	**0.014**	0.365	0.785	0.628	**0.006***	**0.000***	**0.047**	0.969	0.653
$US \rightarrow GE$	**0.000***	**0.000***	0.701	0.542	0.219	**0.000***	**0.000***	0.686	0.192	0.525	**0.035**
$GE \rightarrow US$	**0.000***	0.058	**0.000***	0.329	0.751	0.059	0.744	0.175	**0.000***	**0.000***	0.362
$US \rightarrow JP$	**0.009***	0.774	0.180	**0.000***	**0.001***	**0.000***	0.112	**0.011**	**0.001***	0.709	**0.001***
$JP \rightarrow US$	0.575	0.287	0.149	0.816	0.145	0.676	0.065	**0.001***	**0.000***	0.200	0.690
$US \rightarrow CH$	**0.000***	**0.031**	0.461	**0.045**	**0.001***	**0.027**	0.721	0.704	0.555	0.213	0.902
$CH \rightarrow US$	0.256	0.441	0.541	0.167	0.524	**0.021**	0.996	0.169	0.660	0.393	0.811
$US \rightarrow UK$	0.059	0.253	**0.002***	0.077	0.083	**0.000***	0.266	**0.001***	**0.045**	0.248	**0.043**
$UK \rightarrow US$	0.079	**0.019**	**0.000***	**0.001***	0.057	0.182	**0.045**	0.343	**0.000***	0.777	**0.045**
					Right tail						
$US \rightarrow CN$	**0.000***	0.482	0.420	0.367	0.842	0.890	**0.003***	0.433	0.554	**0.000***	0.543
$CN \rightarrow US$	**0.018**	0.435	0.191	0.244	0.470	**0.001***	0.230	0.660	0.251	0.282	**0.011**
$US \rightarrow FR$	**0.040**	0.988	0.073	0.148	0.979	0.148	0.730	0.265	**0.022**	0.969	0.694
$FR \rightarrow US$	0.440	0.134	0.797	0.781	0.850	0.239	0.307	0.434	0.790	0.547	0.693
$US \rightarrow GE$	0.255	0.412	0.076	0.960	0.563	0.253	0.327	0.675	0.287	0.080	0.712
$GE \rightarrow US$	**0.045**	0.464	0.332	0.477	0.466	0.168	0.077	0.211	**0.001***	**0.001***	0.731
$US \rightarrow JP$	0.619	**0.000***	0.139	0.993	0.913	**0.000***	**0.009***	0.778	0.116	0.659	0.231
$JP \rightarrow US$	**0.042**	0.958	0.672	0.147	0.741	**0.016**	**0.000***	0.833	0.594	0.312	0.302
$US \rightarrow CH$	0.411	0.861	0.278	0.055	0.957	**0.025**	0.145	0.855	0.999	**0.000***	0.820
$CH \rightarrow US$	0.717	0.159	0.944	**0.049**	0.376	0.135	0.874	0.303	0.110	0.147	0.087
$US \rightarrow UK$	0.117	0.280	0.529	**0.018**	0.174	0.086	0.761	0.366	0.814	0.660	0.345
$UK \rightarrow US$	0.335	0.777	0.845	0.395	0.795	0.618	0.217	0.750	0.159	0.996	0.346

Notes: This table presents the p-values of the Granger causality in distribution of returns using data from 03/01/1973 to 12/05/2021. The tests in the left tail are conducted considering α = 1%, 5% and 10%, while the tests in the right tail consider α = 90%, 95% and 99%. The tests are conducted pairwise between the US and other six developed countries (Canada (CN), France (FR), Germany (GE), Japan (JP), China CH), and the UK) for 11 industries (Basic Materials (BM), Consumer Discretionary (CD), Consumer Staples (CS), Energy (EN), Financials (FI), Health Care (HC), Industrials (IN), Real Estate (RE), Technology (TEC), Telecommunications (TEL), and Utilities (UT)). $US \rightarrow k$ indicates Granger causality from the US to the returns of country k, and $k \rightarrow US$ indicates Granger causality from country k to the US, for each industry. Numbers in bold indicate the rejection of the null hypothesis of no Granger causality at the 5% level. One asterisk, "*", denotes significance at the 1% level.

Table 4.12: Granger causality in distribution of the volatility

	BM	CD	CS	EN	FI	HC	IN	RE	TEC	TEL	UT
					Left tail						
$US \rightarrow CN$	0.220	0.243	0.922	**0.000***	0.178	0.609	0.332	0.346	0.223	0.601	**0.020**
$CN \rightarrow US$	0.176	0.157	0.061	**0.041**	0.161	0.366	0.098	0.731	0.766	0.621	**0.046**
$US \rightarrow FR$	**0.022**	0.912	**0.019**	0.059	0.133	0.562	0.648	0.050	0.323	0.094	**0.000***
$FR \rightarrow US$	0.794	0.702	**0.034**	0.066	0.633	0.875	0.404	**0.044**	0.773	0.351	**0.001***
$US \rightarrow GE$	0.117	0.846	**0.010**	-	0.543	0.364	0.083	0.933	0.579	**0.001***	**0.000***
$GE \rightarrow US$	0.794	0.778	0.859	-	0.285	0.174	0.212	0.467	0.078	0.102	0.878
$US \rightarrow JP$	0.125	0.533	0.124	0.668	0.119	0.307	0.546	0.233	0.883	**0.001***	0.782
$JP \rightarrow US$	0.289	0.537	**0.007***	0.504	0.598	0.812	0.455	**0.018**	0.495	**0.001***	0.469
$US \rightarrow CH$	0.579	**0.044**	0.773	**0.000***	0.052	-	0.323	0.288	-	-	0.212
$CH \rightarrow US$	**0.018**	0.420	0.178	0.415	0.156	-	0.768	0.611	-	-	0.057
$US \rightarrow UK$	0.209	0.720	0.782	0.581	**0.016**	**0.036**	0.172	0.160	0.428	0.352	**0.047**
$UK \rightarrow US$	**0.036**	0.314	0.747	0.169	**0.003***	**0.024**	0.085	0.117	0.490	**0.018**	0.108
					Right tail						
$US \rightarrow CN$	**0.000***	**0.083**	**0.000***	0.785	**0.000***	**0.000***	**0.000***	0.291	0.191	0.993	**0.046**
$CN \rightarrow US$	0.097	0.361	0.396	0.750	0.667	0.853	0.069	0.062	**0.007***	**0.027**	0.407
$US \rightarrow FR$	**0.000***	**0.030**	**0.003***	**0.036**	**0.021**	0.210	0.469	**0.000***	**0.000***	**0.002***	0.232
$FR \rightarrow US$	0.798	0.637	0.134	0.148	**0.026**	0.322	0.151	**0.033**	0.179	**0.001***	**0.002***
$US \rightarrow GE$	**0.000***	**0.000***	**0.002***	-	**0.000***	0.299	**0.000***	**0.000***	0.955	**0.000***	**0.000***
$GE \rightarrow US$	0.950	0.146	0.843	-	0.148	0.471	0.251	0.778	**0.012**	**0.014**	0.118
$US \rightarrow JP$	0.804	**0.000***	**0.001***	0.648	0.676	**0.024**	0.632	0.479	**0.000***	0.612	0.687
$JP \rightarrow US$	0.250	**0.006***	0.451	0.941	0.882	0.058	0.784	**0.000***	0.532	0.121	0.223
$US \rightarrow CH$	0.055	**0.001***	**0.000***	0.331	**0.000***	-	**0.046**	**0.000***	-	-	0.187
$CH \rightarrow US$	0.317	**0.039**	0.445	0.120	**0.000***	-	0.598	**0.001***	-	-	0.367
$US \rightarrow UK$	0.859	**0.010***	**0.009***	0.659	**0.006***	**0.028**	**0.000***	0.834	**0.027**	0.532	0.626
$UK \rightarrow US$	0.951	0.298	0.569	0.903	0.786	0.535	0.471	0.805	0.089	0.720	0.317

Notes: This table presents the p-values of the Granger causality in distribution of volatility using data from 03/01/1973 to 12/05/2021. The tests in the left tail are conducted considering α = 1%, 5% and 10%, while the tests in the right tail consider α = 90%, 95% and 99%. The tests are conducted pairwise between the US and other six developed countries (Canada (CN), France (FR), Germany (GE), Japan (JP), China CH), and the UK) for 11 industries (Basic Materials (BM), Consumer Discretionary (CD), Consumer Staples (CS), Energy (EN), Financials (FI), Health Care (HC), Industrials (IN), Real Estate (RE), Technology (TEC), Telecommunications (TEL), and Utilities (UT)). $US \rightarrow k$ indicates Granger causality from the US to the returns of country k, and $k \rightarrow US$ indicates Granger causality from country k to the US, for each industry. Numbers in bold indicate the rejection of the null of no Granger causality at the 5% level. One asterisk, "*", denotes significance at the 1% level. "-" indicates that it was not possible to obtain reliable estimates of the Conditional Autoregressive Value-at-Risk (CAViaR) due to small sample size.

4.5.5. Robustness Checks

To assess the sensitivity of our results to different model specifications and to the data, we performed several robustness checks.

First, we derive the causality and feedback statistics from VAR(2). Causality results were not significantly different from the ones obtained with a VAR(1), except that fewer countries showed a lower relationship with the US. The Geweke feedback measures remain almost the same up to the last significant digit.

Since some series start later than others, we also analysed the sensitivity of our results to the inclusion of only complete series for each industry. For instance, in the case of the Basic Materials (BM) industry, we had complete data for all countries after 1993. Therefore, we excluded 20 years of data for some countries, which forced us to discard a significant part of the series that potentially contains relevant information for our study. Results showed some differences in the Geweke feedback measures. For instance, the US reported a smaller percentage of unilateral feedback than before, while France, Germany, Canada, and the UK reported larger percentages. This is possibly justified by their level of integration which was larger in recent years. Naturally, there were no significant differences for countries that had small data availability before (for example, China).

Lastly, we performed an analysis using daily and monthly data. There was a more and less pronounced causality for daily and weekly data between the US and most countries respectively, than for weekly data. This pattern also exists considering data partition into expansion and recession periods.

4.6. Conclusion

This chapter examines the linear interdependences between international industries, with a focus on the relationships between the US and other 6 countries (Canada, France, Germany, Japan, China, and the UK). This work differs from previous research, which mainly focused on international stock indexes or firm-level returns, ignoring international inter-industry information transmission.

Our results based on the bivariate Granger causality test show that the weekly returns of US industries have a strong and significant causal relationship with most of the countries under scrutiny. The only exception is China, possibly due to international trading constraints. We also find that returns of non-US countries have limited predictive ability over the returns of US industries, implying that the causal relationship is mainly asymmetric. This asymmetry is also supported by the feedback measures. Furthermore, contemporaneous feedback is the major contributor to the total feedback, with an average weight of 94%. These results imply that most markets are highly integrated and that, on average, most of the information transmission occurs within one week. Causality in volatility is stronger and less asymmetric.

US Basic Materials and Energy industries have the strongest and significant causality to industries of other countries. This finding is highly plausible since firms in other industries rely heavily on commodities and fuels. Additionally, returns of commodity- and material-producing industries, placed earlier in the production chain, are frequently strongly connected to returns of industries positioned later in the production chain.

During expansionary periods, the US dominates other countries, but there is less causality than in recession periods. During recessions, there is a high linear dependence between countries. However, they react with a larger delay to information in the other market. This is possible due to the high levels of uncertainty reported during recessions. To the best of our knowledge, this is the first study to directly document international evidence supporting asymmetric reactions to the US industries, during recessions.

Lastly, we analyse the Granger causality in distribution for both industry returns and volatilities. Our results reveal that other countries do not timely incorporate US industries shocks. In particular, countries react with a delay to news from the US, especially in the left tail of the distribution of returns and the right tail of the distribution of volatilities.

In sum, we may conclude that the US plays a leadership role in international markets. This is justified by the US being the world's largest economy largest in terms of GDP and an important trading partner for many countries. Also, the US market is analysed and scrutinized by investors worldwide. This deep analyst coverage and investors' attention increases the impact of US macroeconomic fundamentals on other international markets.

Some possible extensions to this work could be the inclusion of exogenous variables in the VAR models, such as proxies of specific industry characteristics (e.g., institutional holdings, market share, firm size, and trading volume).

5. Final Conclusions

In this thesis, we investigated asset return predictability and its implications for portfolio selection. We have pursued this global aim by analysing several modelling frameworks applied to different datasets. These datasets differ on the asset spaces, predictor spaces, and data frequency. We assessed the effect of predictability on the portfolios of Constant Relative Risk-Averse (CRRA) investors.

We begin our study by looking at how different Vector Autoregressions (VARs) and Time-Varying Parameter – Vector Autoregressions (TVP-VARs), used in different dynamic forecasting schemes, namely Bayesian approaches such Dynamic Model Selection (DMS) and Dynamic Model Averaging (DMA), performed in forecasting weekly excess returns of US stocks, bonds, and REITs indexes in the period from January 1976 to December 2017. These approaches allow the integration of several useful features into a flexible predictive system, namely model and parameter uncertainty, time-varying parameters, combinations of predictors, and time-varying covariance matrices.

Bayesian DMS and DMA combinations of TVP-VAR(1) presented the best results, in statistical terms, but also produced significant benefits measured by out-of-sample pseudo-R^2. We also show that a CRRA investor would benefit from using these frameworks, instead of using other simpler models proposed in the literature, such as equally weighted portfolios based on historical mean returns. Notably, the results, in terms of Certainty Equivalents, Sharpe ratios and Sortino ratios, were quite promising for different risk aversion coefficients, when the investor used DMS or DMA in a predictor space only formed by the first-order lags of the excess returns, and most especially if that predictor space is enlarged by including more three exogenous variables for each asset.

Additionally, we also argue that Bayesian portfolios could better accommodate market instability in their specifications and, hence, can be seen as more robust forecasting methods. This conclusion is backed up by the examination of the performance of different portfolios based on VARs and Bayesian models before and after January 2008, i.e., the beginning of the subprime crisis.

The results presented above provide empirical evidence on the importance of the relationship between asset prediction and portfolio selection. Accordingly, we proceed to analyse the benefits of using more recent methodologies to forecast returns and covariances matrixes, the inputs to the portfolio optimization problem. Firstly, we use two Machine Learning techniques, Random Forests (RF) and Artificial Neural Networks (NN), to forecast daily excess stock and bond returns. Secondly, we introduce more flexibility to the Dynamic Conditional Correlations (DCC) model of Engle et al. (2019) by allowing for asymmetric effects in innovations. This is a relevant issue due to the compelling evidence on the nonlinearities that characterizes asset returns.

We applied this framework to the daily returns of 77 national stock and bond indexes from 44 countries for the period from August 2001 to September 2020. The results show that NN-ADCC for stock indexes, and most especially RF-ADCC models for bond indexes and for both stock and bond indexes, consistently outperformed the benchmarks and models that do not make use of Machine Learning techniques, namely 1/N, MVP, the European index Portfolio, DCC, and ADCC. In particular, we found that a CRRA investor with moderate risk aversion using the proposed RF-ADCC would have experienced a substantial increase in the economic performance of her portfolio. The good performance of this model could be attributed to its ability to capture nonlinearities in data and suitability to deal with large datasets.

Additionally, considering the best model, RF-ADCC, we analyse the benefits of international diversification. We showed that investors from South America, European Union (EU), the Middle East, Asia, and Oceania would benefit from amplifying the geographical scope of their portfolios. From another perspective, we also show that investors from emerging (developed) markets benefit from including in their portfolios assets from developed (emerging) markets, especially in the former case.

Given the evidence on return predictability at an aggregate level (US asset classes and stock and bond international indexes), we continue our study by examining how international markets transmitted information at an industry level. The focus was on the expected key role of the US, as it is the world's biggest economy and an important commercial partner for many countries around the world. Knowing what the leading industries are, allows investors to act on that knowledge, improving the forecastability of the vector of returns, hence increasing the performance of their portfolios.

Accordingly, we studied the interdependences (Granger causality and Geweke feedback measures) within and across international eleven industries between the US and a group of 6 countries (Canada, France, Germany, Japan, China, and the UK) using weekly returns from January 1973 to May 2021.

Our results showed that weekly lagged returns of US industries significantly Granger cause most industries of the other countries, while the reverse causality was far less important. The analysis of the pairwise Geweke feedback measures reinforced that, although most industry variability was transmitted within a week (the contemporaneous feedback accounted, on average, to 94% of the total linear feedback), the lead-lag relationship between the US and other countries was asymmetrical, and information mainly flowed from the US to other countries. The same conclusions were obtained for causality and feedback in the volatility.

Additionally, we identified that returns of US Basic Materials and Energy industries significantly caused the returns of other industries and were the main leading international industries. This finding is highly plausible since firms in other industries rely heavily on commodities, especially fuels. Industries that produce these commodities are situated at the beginning of the production chain, hence price shocks tend to be communicated to other industries situated later in that chain.

After analysing the role played by the US internationally, we assessed whether US dominance prevailed during the country's expansion and recession periods. The US still dominated other countries during expansionary periods, but there was less causal relation than in the full sample or the recession periods. During recession periods, the high linear dependence between countries is higher, but the returns of the industries of other countries reacted with a larger delay to US information. This was possible due to the high levels of uncertainty reported during recessions. To the best of our knowledge, this is the first study to directly document empirical evidence supporting asymmetric reactions to the US industries in the different stages of the business cycle.

Finally, we analysed the Granger causality in distribution for both returns and volatilities. Our results revealed that causality mainly flowed from the US in the left tail of the returns' distribution and in the right tail of the volatility's distribution, meaning that non-US countries did not timely incorporate US industries shocks.

Overall, our thesis supports the main conclusion that, if one does not a priori preclude the existence of return predictability, then its consideration in the selection of well-diversified portfolios may result in important performance gains for international investors.

Understanding how information is transmitted across markets and what are the best and more robust methodologies to forecast returns are crucial.

Nevertheless, all the suggested methodologies present limitations and can be improved in future works. For instance, instead of modelling conditional covariance matrixes, we can use realized co-volatilities to forecast risk within and across markets, use other Machine Learning models to produce forecasts or base forecasts on ensembles of these models, introduce exogenous variables in the lead-lag regressions, etc.

Also, one should keep in mind that we have not proven the profitability of investment strategies based on the predictability of returns. The latter does not necessarily imply the former, as implicit and explicit transaction costs were not taking into account when measuring the portfolios performances. It might be the case that the costs of rebalancing the portfolios, at least on a daily or weekly basis, dilute the potential profitability assessed before transaction costs.

Finally, a word of caution when interpreting these results. We do not claim that we found the best models. Our main objective is to contribute to the literature by highlighting several relevant insights that can be used holistically in the design of a good investment strategy.

References

Adjaouté, K., & Danthine, J. P. (2004). Equity returns and integration: is Europe changing? *Oxford Review of Economic Policy, 20(*4), 555-570. https://doi.org/10.1093/oxrep/grh032

Almadi, H., Rapach, D. E., & Suri, A. (2014). Return predictability and dynamic asset allocation: How often should investors rebalance? *The Journal of Portfolio Management, 40*(4), 16-27. https://doi.org/10.3905/jpm.2014.40.4.016

Ang, A., & Bekaert, G. (1999). *International asset allocation with time-varying correlations* (No. w7056). National Bureau of Economic Research. https://www.nber.org/papers/w7056

Arshanapalli, B., & Doukas, J. (1993). International stock market linkages: Evidence from the pre-and post-October 1987 period. *Journal of Banking & Finance, 17*(1), 193-208. https://doi.org/10.1016/0378-4266(93)90088-U

Badrinath, S. G., Kale, J. R., & Noe, T. H. (1995). Of shepherds, sheep, and the cross-autocorrelations in equity returns. *The Review of Financial Studies, 8*(2), 401-430. https://doi.org/10.1093/rfs/8.2.401

Bahrami, A., Shamsuddin, A., & Uylangco, K. (2019). Are advanced emerging market stock returns predictable? A regime-switching forecast combination approach. *Pacific-Basin Finance Journal, 55,* 142-160. https://doi.org/10.1016/j.pacfin.2019.02.003

Baker, M., & Wurgler, J. (2000). The equity share in new issues and aggregate stock returns. *The Journal of Finance, 55*(5), 2219-2257. https://doi.org/10.1111/0022-1082.00285

Banbura, M., Giannone, D., & Reichlin, L. (2010). Large Bayesian vector auto regressions. *Journal of Applied Econometrics, 25,* 71-92. https://doi.org/10.1002/jae.1137

Bandi, F. M., Perron, B., Tamoni, A., & Tebaldi, C. (2019). The scale of predictability. *Journal of Econometrics, 208*(1), 120-140. https://doi.org/10.1016/j.jeconom.2018.09.008

Barberis, N. (2000). Investing for the long run when returns are predictable. *The Journal of Finance, 55*(1), 225-264. https://doi.org/10.1111/0022-1082.00205

Bates, J.M., & Granger, C.W.J. (1969). The combination of forecasts. *Operational Research Quarterly, 20,* 451-468. https://doi.org/10.1057/jors.1969.103

Bekaert, G., Hodrick, R. J., & Zhang, X. (2009). International stock return comovements. *The Journal of Finance, 64*(6), 2591-2626. https://doi.org/10.1111/j.1540-6261.2009.01512.x

Bekaert, G., & Hodrick, R. (2017). *International Financial Management*. Cambridge University Press.

Best, M. J., & Grauer, R. R. (1991). On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results. *The Review of Financial Studies, 4*(2), 315-342. https://doi.org/10.1093/rfs/4.2.315

Bhuyan, R., Kuhle, J., Ikromov, N., & Chiemeke, C. (2014). Optimal portfolio allocation among REITs, stocks, and long-term bonds: An empirical analysis of US financial markets. *Journal of Mathematical Finance, 4*(2), 104-112. https://doi.org/10.4236/jmf.2014.42010

Black, F. (1976). Studies in stock price volatility changes. *1976 Proceedings of the American Statistical Association,* pp.177-181.

Bollerslev, T., Osterrieder, D., Sizova, N., & Tauchen, G. (2013). Risk and return: Long-run relations, fractional cointegration, and return predictability. *Journal of Financial Economics, 108*(2), 409-424. https://doi.org/10.1016/j.jfineco.2013.01.002

Boudoukh, J., Richardson, M., & Whitelaw, R. F. (1994). Industry's returns and the Fisher effect. *The Journal of Finance, 49*(5), 1595-1615. https://doi.org/10.1111/j.1540-6261.1994.tb04774.x

Brennan, M. J., Jegadeesh, N., & Swaminathan, B. (1993). Investment analysis and the adjustment of stock prices to common information. *The Review of Financial Studies, 6*(4), 799-824. https://doi.org/10.1093/rfs/6.4.799

Britten-Jones, M. (1999). The sampling error in estimates of mean-variance efficient portfolio weights. *The Journal of Finance, 54*(2), 655-671. https://doi.org/10.1111/0022-1082.00120

Cambón, M. I., & Vaduva, M. A. (2017). Lead-lag patterns in the Spanish and other European equity markets. *The Spanish Review of Financial Economics, 15*(2), 63-77. https://doi.org/10.1016/j.srfe.2017.05.001

Campbell, J. Y., & Shiller, R. (1988). The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies, 1*(3), 195-228. https://doi.org/10.1093/rfs/1.3.195

Campbell, J. Y., Viceira, L. M., & Viceira, L. M. (2002). *Strategic asset allocation: portfolio choice for long-term investors.* Clarendon Lectures in Economic.

Campbell, J. Y., Chan, Y. L., & Viceira, L. M. (2003). A multivariate model of strategic asset allocation. *Journal of Financial Economics, 67*(1), 41-80. https://doi.org/10.1016/s0304-405x(02)00231-3

Campbell, J. Y., & Thompson, S. B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies, 21,* 1509-1531. https://doi.org/10.1093/rfs/hhm055

Campbell, S.D. & Diebold, F.X. (2009). Stock returns and expected business conditions: a half century of direct evidence. *Journal of Business and Economic Statistics, 27,* 266-278. https://doi.org/10.1198/jbes.2009.0025

Candelon, B., & Tokpavi, S. (2016). A nonparametric test for Granger causality in distribution with application to financial contagion. *Journal of Business & Economic Statistics, 34*(2), 240-253. https://doi.org/10.1080/07350015.2015.1026774

Cappiello, L., Engle, R. F., & Sheppard, K. (2006). Asymmetric dynamics in the correlations of global equity and bond returns. *Journal of Financial Econometrics, 4*(4), 537-572. https://doi.org/10.1093/jjfinec/nbl005

Carriero, A., Kapetanios, G., & Marcellino, M. (2009). Forecasting exchange rates with a large Bayesian VAR. *International Journal of Forecasting, 25,* 400-417. https://doi.org/10.1016/j.ijforecast.2009.01.007

Cenesizoglu, T., & Timmermann, A. (2012). Do return prediction models add economic value? *Journal of Banking & Finance, 36*(11), 2974-2987. https://doi.org/10.1016/j.jbankfin.2012.06.008

Chang, C. L., McAleer, M., & Tansuchat, R. (2013). Conditional correlations and volatility spillovers between crude oil and stock index returns. *The North American Journal of Economics and Finance, 25*, 116-138. https://doi.org/10.1016/j.najef.2012.06.002

Chordia, T., & Swaminathan, B. (2000). Trading volume and cross-autocorrelations in stock returns. *The Journal of Finance, 55*(2), 913-935. https://doi.org/10.1111/0022-1082.00231

Clark, T. E., & West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics, 138*, 291-311. https://doi.org/10.1016/j.jeconom.2006.05.023

Cochrane, J. H. (2007). The dog that did not bark: A defense of return predictability. *The Review of Financial Studies, 21*(4), 1533-1575. https://doi.org/10.1093/rfs/hhm046

Cogley, T., & Sargent, T. J. (2005). Drifts and volatilities: monetary policies and outcomes in the post WWII US. *Review of Economic Dynamics, 8*(2), 262-302. https://doi.org/10.1016/j.red.2004.10.009

Cooper, I., & Priestly, R. (2009). Time-varying risk premiums and the output gap. *Review of Financial Studies, 22*, 2801-2833. https://doi.org/10.1093/rfs/hhn087

Cooper, R. N. (1985). Economic interdependence and coordination of economic policies. *Handbook of International Economics, 2,* 1195-1234. https://doi.org/10.1016/S1573-4404(85)02014-7

Copeland, M., & Copeland, T. (1998). Leads, lags, and trading in global markets. *Financial Analysts Journal, 54*(4), 70-80. https://doi.org/10.2469/faj.v54.n4.2199

Cowles, A. (1933). Can stock market forecasters forecast? *Econometrica, 1,* 309-324. https://doi.org/10.2307/1907042

Croce, M. M., Marchuk, T., & Schlag, C. (2019). *The leading premium* (No. w25633). National Bureau of Economic Research. https://www.nber.org/system/files/working_papers/w25633/w25633.pdf

Cujean, J., & Hasler, M. (2017). Why does return predictability concentrate in bad times? *The Journal of Finance, 72*(6), 2717-2758. https://doi.org/10.1111/jofi.12544

Cutler, A., Cutler, D. R., & Stevens, J. R. (2012). Random forests. *Ensemble Machine Learning*, 157-175. Springer, Boston, MA. http://dx.doi.org/10.1007/978-1-4419-9326-7_5

D'Ecclesia, R. L., & Costantini, M. (2006). Comovements and correlations in international stock markets. *The European Journal of Finance, 12*(6-7), 567-582. https://doi.org/10.1080/13518470500531135

Dangl, T. & Halling, M. (2012). Predictive regressions with time-varying coefficients. *Journal of Financial Economics, 106*(1), 157-181. https://doi.org/10.1016/j.jfineco.2012.04.003

Dangl, T., & Weissensteiner, A. (2020). Optimal portfolios under time-varying investment opportunities, parameter uncertainty, and ambiguity aversion. *Journal of Financial and Quantitative Analysis, 55*(4), 1163-1198. https://doi.org/10.1017/s0022109019000425

Daskalaki, C., & Skiadopoulos, G. (2011). Should investors include commodities in their portfolios after all? New evidence. *Journal of Banking & Finance, 35*(10), 2606-2626. https://doi.org/10.1016/j.jbankfin.2011.02.022

DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy?. *The Review of Financial Studies, 22*(5), 1915-1953. https://doi.org/10.1093/rfs/hhm075

Diris, B., Palm, F., & Schotman, P. (2014). Long-term strategic asset allocation: An out-of-sample evaluation. *Management Science, 61*(9), 2185-2202. https://doi.org/10.1287/mnsc.2014.1924

Dichtl, H., Drobetz, W., Neuhierl, A., & Wendt, V. S. (2021). Data snooping in equity premium prediction. *International Journal of Forecasting, 37*(1), 72-94. https://doi.org/10.1016/j.ijforecast.2020.03.002

Driesprong, G., Jacobsen, B., & Maat, B. (2008). Striking oil: another puzzle? *Journal of Financial Economics, 89*, 307-327. https://doi.org/10.1016/j.jfineco.2007.07.008

Driessen, J., & Laeven, L. (2007). International portfolio diversification benefits: Cross-country evidence from a local perspective. *Journal of Banking & Finance, 31(*6), 1693-1712. https://doi.org/10.1016/j.jbankfin.2006.11.006

Elliott, G., & Timmermann, A. (2008). Economic forecasting. *Journal of Economic Literature, 46*(1), 3-56. https://doi.org/10.1257/jel.46.1.3

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics, 20*(3), 339-350. https://doi.org/10.1198/073500102288618487

Engle, R. F., Ledoit, O., & Wolf, M. (2019). Large dynamic covariance matrices. *Journal of Business & Economic Statistics, 37*(2), 363-375. https://doi.org/10.1080/07350015.2017.1345683

Engle, R. F., & Manganelli, S. (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics, 22*(4), 367-381. https://doi.org/10.1198/073500104000000370

Fama, E. F. (1990). Term-structure forecasts of interest rates, inflation and real returns. *Journal of Monetary Economics, 25*(1), 59-76. https://doi.org/10.1016/0304-3932(90)90045-6

Ferson, W., Sarkissian, S., & Simin, T. (2003). Is stock return predictability spurious. *Journal of Investment Management, 1*(3), 1-10.

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. https://doi.org/10.1016/j.jfineco.2011.02.003

Fisher, J. D., Pettenuzzo, D., & Carvalho, C. M. (2020). Optimal asset allocation with multivariate Bayesian dynamic linear models. *Annals of Applied Statistics*, *14*(1), 299-338. https://doi.org/10.1214/19-AOAS1303

Fugazza, C., Guidolin, M., & Nicodano, G. (2015). Equally weighted vs. long-run optimal portfolios. *European Financial Management*, *21*(4), 742-789. https://doi.org/10.1111/eufm.12042

Gao, L., Han, Y., Li, S. Z., & Zhou, G. (2018). Market intraday momentum. *Journal of Financial Economics, 129,* 394-414. https://doi.org/10.1016/j.jfineco.2018.05.009

Gao, X., & Nardari, F. (2018). Do commodities add economic value in asset allocation? New evidence from time-varying moments. *Journal of Financial and Quantitative Analysis, 53*(1), 365-393. https://doi.org/10.1017/S002210901700103X

Gargano, A., Pettenuzzo, D., & Timmermann, A. (2019). Bond return predictability: Economic value and links to the macroeconomy. *Management Science, 65*(2), 508-540. https://doi.org/10.1287/mnsc.2017.2829

Geweke, J. (1982). Measurement of linear dependence and feedback between multiple time series. *Journal of the American Statistical Association, 77*(378), 304-313. https://doi.org/10.2307/2287238

Goetzmann, W. N., Li, L., & Rouwenhorst, K. G. (2001). *Long-term global market correlations* (No. w8612). National Bureau of Economic Research. https://www.nber.org/papers/w8612

Golez, B., & Koudijs, P. (2018). Four centuries of return predictability. J*ournal of Financial Economics, 127*(2), 248-263. https://doi.org/10.1016/j.jfineco.2017.12.007

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, 424-438. https://doi.org/10.2307/1912791

Green, J., Hand, J. R. M., and Zhang, X. F. (2013). The supraview of return predictive signals. *Review of Accounting Studies, 18,* 692-730. https://doi.org/10.1007/s11142-013-9231-1

Grubel, H. G. (1968). Internationally diversified portfolios: welfare gains and capital flows. *The American Economic Review*, 1299-1314.

Gu, S., Kelly, B., & Xiu, D. (2018). *Empirical asset pricing via machine learning* (No. w25398). National Bureau of Economic Research. https://www.nber.org/papers/w25398

Guidolin, M., & Hyde, S. (2012). Can VAR models capture regime shifts in asset returns? A long horizon strategic asset allocation perspective. *Journal of Banking & Finance, 36*(3), 695-716. https://doi.org/10.1016/j.jbankfin.2011.10.011

Guo, H. (2006). On the out-of-sample predictability of stock market returns. *Journal of Business, 79,* 645-670. https://doi.org/10.1086/499134

Hamilton, J. D. (1994). *Time series analysis*. Princeton University Press.

Hammerschmid, R., & Lohre, H. (2018). Regime shifts and stock return predictability. *International Review of Economics & Finance, 56,* 138-160. https://doi.org/10.1016/j.iref.2017.10.021

Hansen, P. R., & Timmermann, A. (2012). Choice of sample Split in out-of-sample forecast evaluation. *European University Institute. Economics Working Papers* ECO2012/10. https://EconPapers.repec.org/RePEc:eui:euiwps:eco2012/10

Heaton, J. B., Polson, N. G., & Witte, J. H. (2017). Deep learning for finance: deep portfolios. *Applied Stochastic Models in Business and Industry, 33*(1), 3-12. https://doi.org/10.1002/asmb.2209

Heston, S. L., & Rouwenhorst, K. G. (1994). Does industrial structure explain the benefits of international diversification? *Journal of Financial Economics, 36*(1), 3-27. https://doi.org/10.1016/0304-405X(94)90028-0

Hirshleifer, D., Hou, K., & Teoh, S.H. (2009). Accruals, cash flows, and aggregate stock returns. *Journal of Financial Economics, 91,* 389-406. https://doi.org/10.1016/j.jfineco.2007.11.009

Hong, H., Torous, W., & Valkanov, R. (2007). Do industries lead stock markets? *Journal of Financial Economics, 83,* 367-396. https://doi.org/10.1016/j.jfineco.2005.09.010

Hou, K. (2007). Industry information diffusion and the lead-lag effect in stock returns. The *Review of Financial Studies, 20*(4), 1113-1138. https://doi.org/10.1093/revfin/hhm003

Inoue, A., & Kilian, L. (2005). In-sample or out-of-sample tests of predictability: Which one should we use? *Econometric Reviews, 23*(4), 371-402. https://doi.org/10.1081/ETC-200040785

Ishwaran, H. (2015). The effect of splitting on random forests. *Machine Learning, 99*(1), 75-118. https://doi.org/10.1007/s10994-014-5451-2

Jacobs, H. (2016). Market maturity and mispricing. *Journal of Financial Economics, 122*(2), 270-287. https://doi.org/10.1016/j.jfineco.2016.01.030

Jacobsen, B., Marshall, B. R., & Visaltanachoti, N. (2019). Stock market predictability and industrial metal returns. *Management Science, 65*(7), 3026-3042. https://doi.org/10.1287/mnsc.2017.2933

Jagannathan, R., & Liu, B. (2019). Dividend dynamics, learning, and expected stock index returns. *The Journal of Finance, 74*(1), 401-448. https://doi.org/10.1111/jofi.12731

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, *48*(1), 65-91, https://doi.org/10.1111/j.1540-6261.1993.tb04702.x

Jegadeesh, N., & Titman, S. (1995). Overreaction, delayed reaction, and contrarian profits. *The Review of Financial Studies, 8*(4), 973-993. https://doi.org/10.1093/rfs/8.4.973

Jiang, F., Lee, J., Martin, X., & Zhou, G. (2019). Manager sentiment and stock returns. *Journal of Financial Economics, 132*(1), 126-149. https://doi.org/10.1016/j.jfineco.2018.10.001

Jiang, Z., & Yoon, S. M. (2020). Dynamic co-movement between oil and stock markets in oil-importing and oil-exporting countries: Two types of wavelet analysis. *Energy Economics, 90*, 104835. https://doi.org/10.1016/j.eneco.2020.104835

Johannes, M., Korteweg, A., & Polson, N. (2014). Sequential learning, predictability, and optimal portfolio returns. *The Journal of Finance, 69*(2), 611-644. https://doi.org/10.1111/jofi.12121

Kandel, S., & Stambaugh, R. F. (1996). On the predictability of stock returns: an asset-allocation perspective. *The Journal of Finance, 51*(2), 385-424. https://doi.org/10.1111/j.1540-6261.1996.tb02689.x

Kellner, R., & Rösch, D. (2019). A country specific point of view on international diversification. *Journal of International Money and Finance, 98,* 102064. https://doi.org/10.1016/j.jimonfin.2019.102064

Khalfaoui, R., Tiwari, A. K., Kablan, S., & Hammoudeh, S. (2021). Interdependence and lead-lag relationships between the oil price and metal markets: Fresh insights from the wavelet and quantile coherency approaches. *Energy Economics, 101*, 105421. https://doi.org/10.1016/j.eneco.2021.105421

Kolev, G. I., & Karapandza, R. (2017). Out-of-sample equity premium predictability and sample split–invariant inference. *Journal of Banking & Finance, 84*, 188-201. https://doi.org/10.1016/j.jbankfin.2016.07.017

Koop, G., & Korobilis, D. (2013). Large time-varying parameter VARs. *Journal of Econometrics, 177*(2), 185-198. https://doi.org/10.1016/j.jeconom.2013.04.007

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. https://doi.org/10.1016/S0304-405X(97)00002-0

Krauss, C., Do, X. A., & Huck, N. (2017). Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500. *European Journal of Operational Research, 259*(2), 689-702. https://doi.org/10.1016/j.ejor.2016.10.031

Le Bris, D., Goetzmann, W. N., & Pouget, S. (2019). The present value relation over six centuries: The case of the Bazacle company. *Journal of Financial Economics, 132*(1), 248-265. https://doi.org/10.1016/j.jfineco.2017.03.011

Ledoit, O., & Péché, S. (2011). Eigenvectors of some large sample covariance matrix ensembles. *Probability Theory and Related Fields, 151*(1), 233-264. https://doi.org/10.1007/s00440-010-0298-3

Ledoit, O., & Wolf, M. (2012). Nonlinear shrinkage estimation of large-dimensional covariance matrices. The Annals of Statistics, 40(2), 1024-1060. https://doi.org/10.1214/12-AOS989

Ledoit, O., & Wolf, M. (2015). Spectrum estimation: A unified framework for covariance matrix estimation and PCA in large dimensions. *Journal of Multivariate Analysis, 139*, 360-384. https://doi.org/10.1016/j.jmva.2015.04.006

Lee, C. M., Sun, S. T., Wang, R., & Zhang, R. (2019). Technological links and predictable returns. *Journal of Financial Economics, 132*(3), 76-96. https://doi.org/10.1016/j.jfineco.2018.11.008

Lettau, M. & Ludvigson, S. (2001). Consumption, aggregate wealth, and expected stock returns. *The Journal of Finance*, *56*(3), 815-849. https://doi.org/10.1111/0022-1082.00347

Levy, H., & Sarnat, M. (1970). International diversification of investment portfolios. *The American Economic Review*, *60*(4), 668-675. https://econpapers.repec.org/article/aeaaecrev/v_3a60_3ay_3a1970_3ai_3a4_3ap_3a668-75.htm

Lo, A. W., & MacKinlay, A. C. (1990). When are contrarian profits due to stock market overreaction? *The Review of Financial Studies, 3*(2), 175-205. https://doi.org/10.1093/rfs/3.2.175

Lohrmann, C., & Luukka, P. (2019). Classification of intraday S&P500 returns with a Random Forest. *International Journal of Forecasting, 35*(1), 390-407. https://doi.org/10.1016/j.ijforecast.2018.08.004

Longin, F., & Solnik, B. (2001). Extreme correlation of international equity markets. *The Journal of Finance, 56*(2), 649-676. https://doi.org/10.1111/0022-1082.00340

Ludvigson, S. C., & Ng, S. (2009). Macro factors in bond risk premia. *The Review of Financial Studies*, *22*(12), 5027-5067. https://doi.org/10.1093/rfs/hhp081

Maio, P., & Santa-Clara, P. (2015). Dividend yields, dividend growth, and return predictability in the cross section of stocks. *Journal of Financial and Quantitative Analysis, 50*(1-2), 33-60. https://doi.org/10.1017/S0022109015000058

Menzly, L., & Ozbas, O. (2010). Market segmentation and cross-predictability of returns. *The Journal of Finance, 65*(4), 1555-1580. https://doi.org/10.1111/j.1540-6261.2010.01578.x

Moskowitz, T. J., & Grinblatt, M. (1999). Do industries explain momentum? *The Journal of Finance, 54*(4), 1249-1290. https://doi.org/10.1111/0022-1082.00146

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. https://doi.org/10.1287/mnsc.2013.1838

Nonejad, N. (2018). Déjà vol oil? Predicting S&P 500 equity premium using crude oil price volatility: Evidence from old and recent time-series data. *International Review of Financial Analysis, 58*, 260-270. https://doi.org/10.1016/j.irfa.2018.03.012

Pakel, C., Shephard, N., Sheppard, K., & Engle, R. F. (2020). Fitting vast dimensional time-varying covariance models. *Journal of Business & Economic Statistics*, 1-17. https://doi.org/10.1080/07350015.2020.1713795

Parslow, J., Cressie, N., Campbell, E. P., Jones, E., & Murray, L. (2013). Bayesian learning and predictability in a stochastic nonlinear dynamical model. *Ecological applications, 23*(4), 679-698. https://doi.org/10.1890/12-0312.1

Parsons, C. A., Sabbatucci, R., & Titman, S. (2020). Geographic lead-lag effects. *The Review of Financial Studies, 33*(10), 4721-4770. https://doi.org/10.1093/rfs/hhz145

Pástor, Ľ., & Stambaugh, R. F. (2009). Predictive systems: Living with imperfect predictors. *The Journal of Finance, 64*(4), 1583-1628. https://doi.org/10.1111/j.1540-6261.2009.01474.x

Pesaran, M. H., & Timmermann, A. (1995). Predictability of stock returns: Robustness and economic significance. *The Journal of Finance, 50*(4), 1201-1228. https://doi.org/10.1111/j.1540-6261.1995.tb04055.x

Pontiff, J., & Schall, L. D. (1998). Book-to-market ratios as predictors of market returns. *Journal of Financial Economics, 49*(2), 141-160. https://doi.org/10.1016/S0304-405X(98)00020-8

Probst, P., Boulesteix, A. L., & Bischl, B. (2019). Tunability: importance of hyperparameters of machine learning algorithms. *The Journal of Machine Learning Research, 20*(1), 1934-1965. https://doi.org/10.1007/978-3-030-33778-0_10

Raftery, A., Karny, M., & Ettler, P. (2010). Online prediction under model uncertainty via dynamic model averaging: application to a cold rolling mill. *Technometrics, 52*(1), 52-66. https://doi.org/10.1198/TECH.2009.08104

Rangvid, J. (2006). Output and expected returns. *Journal of Financial Economics, 81,* 595-624. https://doi.org/10.1016/j.jfineco.2005.07.010

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. https://doi.org/10.1093/rfs/hhp063

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. https://doi.org/10.1111/jofi.12041

Rapach, D. E., & Zhou, G. (2013). Forecasting stock returns. *In Handbook of economic forecasting* (Vol. 2, pp. 328-383). Elsevier. https://doi.org/10.1016/B978-0-444-53683-9.00006-2

Rapach, D. E., Ringgenberg, M. C., & Zhou, G. (2016). Short interest and aggregate stock returns. *Journal of Financial Economics, 121*, 46-65. https://doi.org/10.1016/j.jfineco.2016.03.004

Rapach, D. E., Strauss, J., Tu, J., & Zhou, G. (2015). *Industry interdependencies and cross-industry return predictability.* Research Collection Lee Kong Chian School Of Business. https://doi.org/10.2139/ssrn.2566541

Rapach, D. E., Strauss, J. K., Tu, J., & Zhou, G. (2019). Industry return predictability: A machine learning approach. *The Journal of Financial Data Science, 1*(3), 9-28. https://doi.org/10.3905/jfds.2019.1.3.009

RiskMetrics, 1996. Technical Document, fourth edition. http://www.riskmetrics.com/system/files/private/td4e.pdf.

Rizova, S. (2010). Predictable returns of trade-linked countries: Evidence and explanations. *University of Chicago Booth School of Business.*

Roll, R. (1992). Industrial structure and the comparative behavior of international stock market indices. *The Journal of Finance, 47*(1), 3-41. https://doi.org/10.1111/j.1540-6261.1992.tb03977.x

Santos, T. & Veronesi, P. (2006). Labor income and predictable stock returns. *Review of Financial Studies, 19,* 1-44. https://doi.org/10.1093/rfs/hhj006

Sander, M. (2018). Market timing over the business cycle. *Journal of Empirical Finance, 46,* 130-145. https://doi.org/10.1016/j.jempfin.2017.12.002

Sarwar, G., & Khan, W. (2017). The effect of US stock market uncertainty on emerging market returns. *Emerging Markets Finance and Trade, 53*(8), 1796-1811. https://doi.org/10.1080/1540496X.2016.1180592

Shiller, R. J. (2000). Measuring bubble expectations and investor confidence. *The Journal of Psychology and Financial Markets, 1*(1), 49-60. https://doi.org/10.1207/S15327760JPFM0101_05

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica: journal of the Econometric Society,* 1-48. https://doi.org/10.2307/1912017

Solnik, B. (1974). Why not diversify internationally rather than domestically? *Financial Analysts Journal, 30*(4), 48-54. https://doi.org/10.2469/faj.v30.n4.48

Solnik, B., & Roulet, J. (2000). Dispersion as cross-sectional correlation. *Financial Analysts Journal*, *56*(1), 54-61. https://doi.org/10.2469/faj.v56.n1.2330

Stambaugh, R. F. (1999). Predictive regressions. *Journal of Financial Economics, 54*(3), 375-421. https://doi.org/10.1016/S0304-405X(99)00041-0

Thornton, D. L. & Valente, G. (2012). Out-of-sample predictions of bond excess returns and forward rates: An asset allocation perspective. *Review of Financial Studies, 25*(10), 3141-3168. https://doi.org/10.1093/rfs/hhs069

Tse, Y. (2018). Return predictability and contrarian profits of international index futures. *Journal of Futures Markets, 38*(7), 788-803. https://doi.org/10.1002/fut.21892

Van Binsbergen, J. H., & Koijen, R. S. (2010). Predictive regressions: A present-value approach. *The Journal of Finance, 65*(4), 1439-1471. https://doi.org/10.1111/j.1540-6261.2010.01575.x

Venditti, F., & Veronese, G. (2020). *Global financial markets and oil price shocks in real time.* European Central Bank (ECB) Working Paper Series No 2472 / September 2020. https://www.ecb.europa.eu/pub/pdf/scpwps/ecb.wp2472~611f104931.en.pdf

Viceira, L. M., & Wang, Z. K. (2018). *Globcargal portfolio diversification for long-horizon investors (No. w24646).* National Bureau of Economic Research. https://www.nber.org/papers/w24646

Wang, Y., Liu, L., Ma, F., & Diao, X. (2018). Momentum of return predictability. *Journal of Empirical Finance, 45,* 141-156. https://doi.org/10.1016/j.jempfin.2017.11.003

Welch, I., & Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies, 21*(4), 1455-1508. https://doi.org/10.1093/rfs/hhm014

Wen, Y. C., Lin, P. T., Li, B., & Roca, E. (2015). Stock return predictability in South Africa: The role of major developed markets. *Finance Research Letters*, *15*, 257-265. https://doi.org/10.1016/j.frl.2015.10.014

World Bank (2021, August). World Bank Open Data Website. Retrieved on May 16, 2021, from http://data.worldbank.org

Wu, X., Zhu, X., Wu, G. Q., & Ding, W. (2013). Data mining with big data. *IEEE transactions on knowledge and data engineering*, *26*(1), 97-107. https://doi.org/10.1109/TKDE.2013.109

Xia, Y. (2001). Long Term Bond Markets and Investor Welfare. *University of Pennsylvania Finance Working Paper.* http://dx.doi.org/10.2139/ssrn.271089

Xu, Y. (2004). Small levels of predictability and large economic gains. *Journal of Empirical Financ*e, *11*(2), 247-275. https://doi.org/10.1016/j.jempfin.2003.04.002

Yang, J., Hsiao, C., Li, Q., & Wang, Z. (2006). The emerging market crisis and stock market linkages: further evidence. *Journal of Applied Econometrics*, *21*(6), 727-744. https://doi.org/10.1002/jae.889

Zeng, K., & Atta Mills, E. F. E. (2021). Can economic links explain lead–lag relations across firms? *International Journal of Finance & Economics*. https://doi.org/10.1002/ijfe.2480

Zhang, Y., Ma, F., Shi, B., & Huang, D. (2018). Forecasting the prices of crude oil: an iterated combination approach. *Energy Economics, 70,* 472-483. https://doi.org/10.1016/j.eneco.2018.01.027

Zhang, Y., Zeng, Q., Ma, F., & Shi, B. (2019). Forecasting stock returns: Do less powerful predictors help? *Economic Modeling*, *78,* 32-39. https://doi.org/10.1016/j.econmod.2018.09.014

Zhu, X., & Zhu, J. (2013). Predicting stock returns: A regime-switching combination approach and economic links. *Journal of Banking & Finance*, *37*(11), 4120-4133. https://doi.org/10.1016/j.jbankfin.2013.07.016

Appendix

Table A1: Description of Stocks predictors

Variable	Description	Source	Transformation
DP	Dividend Price ratio	Goyal and Welch (2008)	--
Inf	Inflation	Goyal and Welch (2008)	--
Tbl	Treasury Bills	Goyal and Welch (2008)	--
Lty	Long Term Yield	Goyal and Welch (2008)	--
Tms	Term Spread	Goyal and Welch (2008)	--
Rf	Risk free rate	Goyal and Welch (2008)	--
b/m	Book to Market Value	Goyal and Welch (2008)	--
EP Ratio	Earning Price Ratio	Goyal and Welch (2008)	--
DP Ratio	Dividend Payout Ratio	Goyal and Welch (2008)	--
Svar	Stock Variance	Goyal and Welch (2008)	--
D/y	Dividend yield	Goyal and Welch (2008)	--
Dfy	Default yield spread	Goyal and Welch (2008)	--
CAPE	Cyclically Adjusted Price-to-earnings ratio	Robert Shiller Database	--
Factors: SMB	Small minus Big factor	Fama and French (2018)	--
Factors: HML	High Minus Low factor	Fama and French (2018)	--
Factors: RMW	Robust Minus Weak	Fama and French (2018)	--
Factors: CMA	Conservative Minus Aggressive	Fama and French (2018)	--
Ntis	Net Equity Expansion	Goyal and Welch (2008)	--
Ltr	Long Term Rate	Goyal and Welch (2008)	--

Notes: In this table it is presented the 19 variables used to predict stock returns and their sources. We do not transform the data (to correct stationarity) as it is common practice for the predictors considered. "--" indicates no transformations. Goyal and Welch (2008) database is available on the website: https://sites.google.com/view/agoyal145, Robert Shiller database is available on: http://www.econ.yale.edu/~shiller/data.htm, Fama and French (2018) database is available on: https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

Table A2: Description of Bonds predictors

Variable	Description	Source	Transformation
Piy	Personal Income	Ludvigson and Ng (2009)	Δln
Piless transfers	Personal Income minus Transfers	Ludvigson and Ng (2009)	Δln
RealConsumption expensures	Real Consumption Expensures	Ludvigson and Ng (2009)	Δln
M&Tsales	Manufacturing and Trade Sales	Ludvigson and Ng (2009)	Δln
Retail sales	Retail sales	Ludvigson and Ng (2009)	Δln
IP:total	Industrial Total Production Index	Ludvigson and Ng (2009)	Δln
IP:products	Industrial Production Index	Ludvigson and Ng (2009)	Δln
IP:final products	Industrial Final Production Index	Ludvigson and Ng (2009)	Δln
IP:cons goods	Industrial Consumption Production Index	Ludvigson and Ng (2009)	Δln
IP:cons durable goods	Industrial Consumption Durables Production Index	Ludvigson and Ng (2009)	Δln
IP:cons non-durable goods	Industrial Consumption non-Durables Production Index	Ludvigson and Ng (2009)	Δln
IP:bus equipament	Industrial Bus equipament Production Index	Ludvigson and Ng (2009)	Δln
IP:matls	Industrial Materials Production Index	Ludvigson and Ng (2009)	Δln
IP:dble material	Industrial Durable materials Production Index	Ludvigson and Ng (2009)	Δln
IP:nondble materials	Industrial non-Durable materials Production Index	Ludvigson and Ng (2009)	Δln
IP:mfg manufacturing	Industrial manufacturing Production Index	Ludvigson and Ng (2009)	Δln
IP:residental utilitites	Industrial residental utilitites Production Index	Ludvigson and Ng (2009)	Δln
IP:fuels	Industrial fuels Production Index	Ludvigson and Ng (2009)	Δln
Caputil	Capacity Utilization	Ludvigson and Ng (2009)	Δ
Helpwanted	Index Of Help-Wanted Advertising In Newspapers	Ludvigson and Ng (2009)	Δ
Helpwanted/emp	Employment Ratio	Ludvigson and Ng (2009)	Δln
EmpCPS total	Civilian Labor Force: Employed, Total	Ludvigson and Ng (2009)	Δln
EmpCPS	Civilian Labor Force: Employed	Ludvigson and Ng (2009)	Δln
wps1	PPI Finish Goods	Ludvigson and Ng (2009)	Δ2ln
wps2	PPI Finish Consumer Goods	Ludvigson and Ng (2009)	Δ2ln
wps3	PPI Intermediate Goods	Ludvigson and Ng (2009)	Δ2ln
wps4	PPI Crude Price	Ludvigson and Ng (2009)	Δ2ln
NAPMempl	Employment Index	Ludvigson and Ng (2009)	--
inflation	Inflation	Goyal and Welch (2008)	--
Macro Factors	8 Macro Factors	Ludvigson and Ng (2009)	--

Table A2: Description of Bonds predictors (continued)

Variable	Description	Source	Transformation
U:all	Unemployment rate	Ludvigson and Ng (2009)	Δ
U:mean	Mean Unemployment duration	Ludvigson and Ng (2009)	Δ
U<5	Less than 5 years Unemployment	Ludvigson and Ng (2009)	Δln
U5-4	5 to 4 years Unemployment	Ludvigson and Ng (2009)	Δln
U15+	More than 15 years Unemployment	Ludvigson and Ng (2009)	Δln
U15-26	15 to 26 years Unemployment	Ludvigson and Ng (2009)	Δln
U27+	More than 27 years Unemployment	Ludvigson and Ng (2009)	Δln
UIclaims	Average Weekly Initial Claims	Ludvigson and Ng (2009)	Δln
Emp:total	Total Employees on Nonfarm Payrolls	Ludvigson and Ng (2009)	Δln
Emp:gds	Employees on Goods sector	Ludvigson and Ng (2009)	Δln
Emp:mining	Employees on Mining sector	Ludvigson and Ng (2009)	Δln
Emp:const	Employees on Construction sector	Ludvigson and Ng (2009)	Δln
Emp:mfg	Employees on manufacturing goods sector	Ludvigson and Ng (2009)	Δln
Emp:dble	Employees on durables goods sector	Ludvigson and Ng (2009)	Δln
Emp:nondbles	Employees on non durables goods sector	Ludvigson and Ng (2009)	Δln
Emp:services	Employees on services sector	Ludvigson and Ng (2009)	Δln
Emp:TTU	Employees on transports sector	Ludvigson and Ng (2009)	Δln
Emp:wholesale	Employees on wholesale sector	Ludvigson and Ng (2009)	Δln
Emp:retail	Employees on retail sector	Ludvigson and Ng (2009)	Δln
Emp:FIRE	Employees on financial sector	Ludvigson and Ng (2009)	Δln
Emp:Govt	Employees on government sector	Ludvigson and Ng (2009)	Δln
Avghrs	Average Weekly Hours of Production	Ludvigson and Ng (2009)	--
Overtime:mfg	Overtime: Average Weekly Hourrs of Production on manufactory goods	Ludvigson and Ng (2009)	Δ
Avghrs: mfg	Average Weekly Hourrs of Production on manufactory goods	Ludvigson and Ng (2009)	--
1-Year Forward Rate	1-Year Forward Rate	US Department of the Treasury	--
2-Year Forward Rate	2-Year Forward Rate	US Department of the Treasury	--
5-Year Forward Rate	5-Year Forward Rate	US Department of the Treasury	--
10-Year Forward Rate	10-Year Forward Rate	US Department of the Treasury	--
T10Y2YM	10-Year Treasury Constant Maturity Minus 2-Month Treasury Constant Maturity	FRED	--
T10Y3YM	10-Year Treasury Constant Maturity Minus 3-Month Treasury Constant Maturity	FRED	--

Table A2: Description of Bonds predictors (continued)

Variable	Description	Source	Transformation
Permit	New House Permits	Ludvigson and Ng (2009)	ln
Permitne	New House Permits, northeast	Ludvigson and Ng (2009)	ln
permitmw	New House Permits, midwest	Ludvigson and Ng (2009)	ln
permits	New House Permits, south	Ludvigson and Ng (2009)	ln
permitw	New House Permits, west	Ludvigson and Ng (2009)	ln
amd	New Order for Consumer goods	Ludvigson and Ng (2009)	Δln
and	New Order for Durable goods	Ludvigson and Ng (2009)	Δln
amdmu	New Order for Non defense capital goods	Ludvigson and Ng (2009)	Δln
Businv	Total Business Inventories	Ludvigson and Ng (2009)	Δln
isratio	Inventories to Sales Ratio	Ludvigson and Ng (2009)	Δ
m1	M1 Money Stock	Ludvigson and Ng (2009)	Δ2ln
m2	M2 Money Stock	Ludvigson and Ng (2009)	Δ2ln
m2real	M2 Real Money Stock	Ludvigson and Ng (2009)	Δ2ln
Amb	St luis Adjusted Monetary Base	Ludvigson and Ng (2009)	Δ2ln
Totresns	Total Reserves	Ludvigson and Ng (2009)	Δ2ln
nonborr	Reserves of Depository Institutions	Ludvigson and Ng (2009)	Δ
busloans	Commercial and Industrial Loans	Ludvigson and Ng (2009)	Δ2ln
Realln	Real Estate Loans	Ludvigson and Ng (2009)	Δln
nonrev	Nonrevolving Credit	Ludvigson and Ng (2009)	Δ2ln
conspi	Consumer Credit to Personal Income	Ludvigson and Ng (2009)	ln
Ff	Federal Fund rate	Ludvigson and Ng (2009)	ln
cp3m	Commercial Paper rate	Ludvigson and Ng (2009)	ln
tb3m	3-month Treasury Bill	Ludvigson and Ng (2009)	ln
tb6m	6-month Treasury Bill	Ludvigson and Ng (2009)	ln
gs1	1-year Treasury rate	Ludvigson and Ng (2009)	ln
gs5	5-years Treasury rate	Ludvigson and Ng (2009)	ln
gs10	10-years Treasury rate	Ludvigson and Ng (2009)	ln
Aaa	Aaa Bond Yield	Ludvigson and Ng (2009)	ln
Baa	Baa Bond Yield	Ludvigson and Ng (2009)	ln
Comp	Composite Federal Fund spread	Ludvigson and Ng (2009)	--
tb3	3-month Federal Fund spread	Ludvigson and Ng (2009)	--
tb6	6-month Federal Fund spread	Ludvigson and Ng (2009)	--
t1yffm	1-year Federal Fund spread	Ludvigson and Ng (2009)	--
t5yffm	5-year Federal Fund spread	Ludvigson and Ng (2009)	--
t10yffm	10-year Federal Fund spread	Ludvigson and Ng (2009)	--
Aaaf	Aaa Federal Fund spread	Ludvigson and Ng (2009)	--
baaf	Bbb Federal Fund spread	Ludvigson and Ng (2009)	--
tw	Trade Weight for Us Dollar	Ludvigson and Ng (2009)	Δln
Exs	Exchange rate Switzland	Ludvigson and Ng (2009)	Δln
Exj	Exchange rate Japan	Ludvigson and Ng (2009)	Δln
Exu	Exchange rate UK	Ludvigson and Ng (2009)	Δln
Exc	Exchange rate Canada	Ludvigson and Ng (2009)	Δln

Table A2: Description of Bonds predictors (continued)

Variable	Description	Source	Transformation
oil	Crude oil	Ludvigson and Ng (2009)	Δ2ln
Ppic	PPI Metals	Ludvigson and Ng (2009)	Δ2ln
Cpiau	CPI all Items	Ludvigson and Ng (2009)	Δ2ln
Cpiap	CPI Apparel	Ludvigson and Ng (2009)	Δ2ln
Cpitr	CPI Transports	Ludvigson and Ng (2009)	Δ2ln
Cpimed	CPI Medical Care	Ludvigson and Ng (2009)	Δ2ln
cus1	CPI Commodities	Ludvigson and Ng (2009)	Δ2ln
cus2	CPI Durables	Ludvigson and Ng (2009)	Δ2ln
cus3	CPI services	Ludvigson and Ng (2009)	Δ2ln
Cpiul	CPI all but Food	Ludvigson and Ng (2009)	Δ2ln
cus4	CPI all but Shelter	Ludvigson and Ng (2009)	Δ2ln
cus5	CPI all but Medical Care	Ludvigson and Ng (2009)	Δ2ln
Pcepi	Personal Consumption Chain	Ludvigson and Ng (2009)	Δ2ln
ddur	Personal Consuption Durable	Ludvigson and Ng (2009)	Δ2ln
Dnd	Personal Consuption Non Durable	Ludvigson and Ng (2009)	Δ2ln
dse	Personal Consuption Services	Ludvigson and Ng (2009)	Δ2ln
ces7	Average Hours of Earnings in Goods sector	Ludvigson and Ng (2009)	Δ2ln
ces8	Average Hours of Earnings in Construction sector	Ludvigson and Ng (2009)	Δ2ln
ces9	Average Hours of Earnings in Manufacturing sector	Ludvigson and Ng (2009)	Δ2ln
Mzms	Money Stock	Ludvigson and Ng (2009)	Δ2ln
Invest	Securities in Bank Credit	Ludvigson and Ng (2009)	Δ2ln
vxocls	VXO's	Ludvigson and Ng (2009)	--
CP	Forward Rate Factor	Cochrane and Piazzesi (2009)	--

Notes: In this table it is presented the 125 variables used to predict bond returns, the sources, the transformations made in Excel or MATLAB to obtain the format used in the literature and to correct non stationarity series ("Δ" is the first difference, "Δln" is the first log difference, "Δ2ln" is the second log difference, "ln" is the log transformation, "--" indicates no transformations). Ludvigson and Ng (2009) database is available on: https://www.sydneyludvigson.com/data-and-appendixes, Cochrane and Piazzesi (2009) database is available on: https://www.aeaweb.org/articles?id=10.1257/0002828053828581, US Department of the Treasury database is available on: https://home.treasury.gov/data/treasury-open-data, Federal Reserve Bank of St. Louis (FRED) database is available on: http://research.stlouisfed.org/fred2/categories/22, Goyal and Welch (2008) database is available on the website: https://sites.google.com/view/agoyal145.

Table A3: Description of REITs Predictors

Variable	Description	Source	Transformation
Dividend Price ratio	Dividend Price ratio	Goyal and Welch (2008)	--
inflation	inflation	Goyal and Welch (2008)	--
Tms	Term Spread	Goyal and Welch (2008)	--
Dfy	Default yield spread	Goyal and Welch (2008)	--
chng income	change in income	FRED	--
chng adult population	change in adult population	FRED	Δln
chng employment	change in employment	FRED	--
housing starts	Rate of growth of housing units	Federal Housing Finance Agency	%Δ
mortgage loan amount	Amount of mortgage loan	Federal Housing Finance Agency	%Δ
purchase price	Houses purchase price	Federal Housing Finance Agency	%Δ
loan to price ratio	loan to price ratio	Federal Housing Finance Agency	Δln
rent vacancy rate	rent vacancy rate	Federal Housing Finance Agency	ln
change in mortgage loan	percentage change mortgage loan	Federal Housing Finance Agency	--
30-Year Conventional Mortgage Rate	30-Year Conventional Mortgage	Federal Housing Finance Agency	Δln

Notes: In this table it is presented the 14 variables used to predict REITs returns, the sources, the transformations made in excel or MATLAB to obtain the format used in the literature and to correct non stationarity series ("%Δ" is the first difference, "Δln" is the first log difference, "ln" is the log transformation, "--" indicates no transformations). Goyal and Welch (2008) database is available on https://sites.google.com/view/agoyal145 and Federal Reserve Bank of St. Louis (FRED) database is available on: http://research.stlouisfed.org/fred2/categories/22, Federal Housing Finance Agency: https://www.fhfa.gov/DataTools/Downloads/Pages/National-Mortgage-Database-Aggregate-Data.aspx and http://www.freddiemac.com/pmms/pmms30.html.