UNIVERSITY OF OTTAWA

uOttawa

DOCTORAL THESIS

Essays in Empirical Asset Pricing and International Finance

Author	Supervisor	Co-Supervisor
Emile Herve NDOUMBE	Pr. Maral KICHIAN	Pr. Ba CHU

Thesis submitted to the University of Ottawa

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Economics

Department of Economics

Faculty of Social Sciences

University of Ottawa

Declaration of Authorship

I, Emile Herve NDOUMBE, declare that this thesis titled, "Essays in Empirical Asset Pricing and International Finance" and the work presented in it are my own.

Chapter 1 of this thesis was done jointly with Pr. Ba CHU. My contribution is equal to his.

Chapter 2 of this thesis is solo authored by myself.

Chapter 3 of this thesis was done jointly with Pr. Maral KICHIAN. My contribution is equal to hers.

I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: *Emile Herve Ndoumbe*

Date: December 17, 2025

Contents

List of Figures

List of Tables

Abstract

Chapter 1 introduces the Supervised Dynamic Orthogonal Components (sDOC) method as a novel framework for forecasting the equity risk premium out-of-sample. sDOC advances traditional linear dimension-reduction techniques—most notably Principal Component Analysis (PCA)—by integrating machine learning–based feature selection with the construction of dynamic, volatility-sensitive orthogonal factors. Unlike PCA, which does not capture nonlinear relationships or volatility interactions among predictors, sDOC explicitly models these elements, thereby achieving superior forecasting accuracy. The method is applied to 63 U.S. monthly industry equity portfolios and evaluated against several benchmark models: PCA, univariate GARCH(1,1), random forest (RF), and a three-layer neural network (NN3). Empirical evidence shows that, across most industries, sDOC outperforms PCA, univariate GARCH(1,1), and RF, while delivering predictive performance comparable to NN3. This superiority is evident across multiple evaluation criteria, including out-of-sample R^2, Mincer–Zarnowitz R^2, gains in average excess portfolio returns, and a range of risk-adjusted metrics such as the Sharpe ratio, Sortino ratio, Calmar ratio, and gain-to-pain ratio. Moreover, the rolling out-of-sample R^2 underscores sDOC's adaptability under both calm and turbulent market conditions, positioning it as a robust and interpretable tool for asset return prediction.

Chapter 2 proposes a modeling strategy to investigate the transmission of shocks—whether through interdependence, contagion, or decoupling—between two asset markets during periods of heightened volatility. The regime-switching model captures co-movements in both the mean and volatility processes of asset returns. The mean dynamics incorporate PCA-derived factors that reflect global and market-specific influences, while the variance–covariance structure accounts for common and idiosyncratic shocks, each governed by an independent Markov-switching process. Contagion or decoupling is defined as occurring when a high-volatility idiosyncratic shock originating in one market significantly alters the interdependence of mean returns across the considered assets. To statistically detect these

phenomena, we employ a novel bootstrap-based Student's t-test procedure. Applying this methodology to sovereign bond market pairs across three Latin American countries, we find evidence that, on average, decoupling has occurred in the Brazil–Mexico and Argentina–Mexico pairs. For the remaining combinations, our results suggest that shock transmission is characterized primarily by interdependence.

Chapter 3 investigates the impact of country-specific major export commodity shocks on domestic equity returns. Employing a Markov-switching framework, we jointly model and estimate commodity–equity return pairs, capturing co-movements in both the mean and the variance–covariance structure of these assets. Co-movement in the mean reflects the influence of global risk, while the variance–covariance structure accounts for common and idiosyncratic shocks, each governed by an independent Markov-switching process. We then examine the transmission of shocks during crises, focusing on cases where shocks originate in a country's commodity market (idiosyncratic shocks) and spill over into its equity market. Contagion is defined as the amplification of idiosyncratic commodity shocks within equity market returns during a crisis period. Using data from six major commodity-exporting countries between 2006 and 2024, and applying likelihood-based tests, we find evidence of contagion during the Global Financial Crisis—most notably in the oil–equity linkages of Norway and the United States. Out-of-sample forecasts further demonstrate that our proposed model consistently outperforms several widely used benchmarks, particularly in predicting equity returns.

Acknowledgements

I am deeply grateful to God for blessing me with motivation, patience, and good health, which have been indispensable in completing this thesis. I would also like to take this opportunity to express my heartfelt gratitude to the many individuals whose invaluable support and contributions have made my Ph.D. journey a successful and fulfilling experience.

First and foremost, I wish to express my deepest gratitude to my advisors, Professor Maral Kichian and Professor Ba Chu. Maral and Ba, thank you for your unwavering support, generous availability, extraordinary patience, and remarkable humility throughout this journey. Your guidance and encouragement have been invaluable at every stage, and I am truly grateful for the opportunity to learn from you.

My sincere thanks also go to the thesis committee members - Professors Francesca Rondina, François-Eric Racicot, Raúl Razo-Garcia - for their time, encouragement, insightful comments, and questions always aimed at improving the quality of all the chapters of my thesis. I would like to extend my appreciation to the external examiner, Professor James Morley, for his time and valuable comments on my thesis. This acknowledgment also goes to all faculty members in the Department of Economics at the University of Ottawa. Special mention and thanks to Professors Kathleen Day, Gamal Atallah, Rose Anne Devlin, Isabelle Salle, Catherine Deri Armstrong, Vicky Barham, and Louis-Philippe Morin. In addition, I would like to express my gratitude to the following university administrative staff: Alexie, Guilherme, Martine, and Nour.

I gratefully acknowledge the Thesis Completion Grant in memory of Professor Serge Nadeau, which provided crucial support during the writing of the final chapter of this thesis. This generous assistance afforded me the time and focus needed to bring the project to completion, and I deeply appreciate the recognition and encouragement it represents.

I would like to express my deepest gratitude to my mum, a truly remarkable woman whose strength, wisdom, and unwavering support shaped every step of this journey. Though she passed away in 2020, her belief in me continues to guide and inspire me. Her courage, love, and quiet resilience remain a source of strength in my life. This work is, in many ways, a tribute to her enduring presence in my heart.

To my sisters and younger brother—Alice, Francine, and Cyrille—your presence has been a source of boundless encouragement and motivation. I am deeply grateful for your love and support throughout this journey. I would also like to thank my new family in Canada, especially Ronal Pion, Michele Pion, Thérèse, Elba Dina, Youssouf Dina, Silvain Pion, Sarah Pion, Lise Gauvin, Cleomene, Marie-Alice, and Paulette Wilson. Your kindness, motivation, and unwavering encouragement helped me to persevere through challenges and stay focused on my goals. Your belief in me has meant more than words can express.

Last but certainly not least, I would like to extend my heartfelt thanks to all my friends, especially Soon Mo, Darcy, Dr Neba, Thierry Ngah, Axel, Kevin, Dhrupad, Abdoulaye Sako, Djamil, Pierre-Olivier, Jean-Luc, Ridenguer, and Nesly. Your support has been a source of strength throughout this long journey. I am also deeply grateful to my fellow economics doctoral colleagues, particularly Junior Sidie, Winjie Tian, David Valenta, Mahmoudou Alioum, and Yannick Ngongang for the mutual support and camaraderie we shared along the way.

To my late Mum: Djengue Ndoumbe,
whose sacrifices laid the foundation
for my doctoral journey
My late Grand Mum: Nyounga
My Sisters and Brother: Francine,
Alice, and Cyrille

General Introduction

Financial markets are inherently complex, shaped by a combination of interacting forces, including economic fundamentals, investor sentiment, institutional structures, and global macroeconomic trends. This complexity intensifies during crises, when volatility spikes, asset correlations change unpredictably, and conventional pricing relationships break down. In such turbulent environments, even well-established models often fail to capture the rapid and nonlinear changes in market dynamics. Understanding how asset prices behave under stress is, therefore, not only a theoretical challenge but a practical necessity for investors seeking to manage risk and for policymakers striving to maintain financial stability. This thesis explores different aspects of the dynamic behavior of asset classes, notably during turbulent times, offering novel empirical insights and methodological innovations that enhance our ability to forecast returns, understand the nature of shock transmissions, and assess cross-market spillovers.

Estimating the equity risk premium accurately is fundamental to constructing optimal portfolios. One of the ways in which this estimation can be accomplished is by forecasting excess stock returns using some underlying models. However, crises often disrupt historical patterns, making these models inadequate for producing good out-of-sample forecasts. Chapter 1 addresses this issue by introducing the Supervised Dynamic Orthogonal Components (sDOC) method, a machine learning-enhanced framework that captures nonlinear relationships and volatility dynamics among predictors. Applied to U.S. industry equity portfolios, sDOC demonstrates superior out-of-sample forecasting performance compared to popular alternatives like PCA, GARCH(1,1), and random forest, and it performs on par with neural network models. Its adaptability in stable and turbulent market regimes makes it a robust tool for investors seeking reliable return predictions under uncertainty.

Although forecasting excess returns is essential for portfolio construction, understanding how shocks propagate across markets is equally critical, especially during crises, when asset interdependencies sometimes intensify (contagion). In today's interconnected financial landscape, analyzing the nature of shock transmission between asset markets is increasingly complex. Assets co-move not only due to regional linkages, but also under the influence of global factors such as the Global Financial Cycle and the Global Trade Commodity Cycle—two macroeconomic forces often neglected in prior studies. Chapter 2 confronts this analytical challenge by proposing a testing strategy within a regime-switching model that can distinguish between interdependence, contagion, and decoupling amongst markets during crises. By incorporating PCA-derived global and local factors and modeling volatility through independent Markov-switching processes, the chapter reveals nuanced transmission patterns across various Latin American sovereign bond market pairs. These findings highlight the difficulty of isolating shocks in a globalized world and the importance of accounting for global and idiosyncratic influences.

Despite a growing literature on financial spillovers, the link between commodity prices and equity market returns remains underexplored, particularly the connection between a country's stock market and the world price of its main export commodity. Most existing studies adopt a global or panel-based approach, which can obscure country-specific transmission channels. Chapter 3 fills this gap by examining how shocks to the returns of major export commodities affect equity returns in different commodity-dependent economies. Using a Markov-switching framework, and examining several country cases during three different crisis periods, the chapter finds evidence of contagion effects, especially in oil-exporting countries like Norway and the U.S. Additionally, the model's demonstrated superior forecasting performance underscores its practical relevance, offering valuable insights for policymakers concerned with macro-financial stability and for investors navigating commodity-linked equity exposures.

Together, these essays contribute to a deeper understanding of asset pricing and market interactions during high-volatility periods and crises. For investors, the findings offer improved tools to manage risky asset returns and portfolio risk in volatile conditions. For policymakers, the research provides evidence on how global and commodity-specific shocks propagate through financial systems, forming strategies for crisis mitigation and economic resilience. By integrating advanced empirical methods with a crisis-focused lens, this thesis advances the frontier of asset pricing research and underscores the importance of context-sensitive analysis in a globalized financial world.

Chapter 1

Estimating Industry-level Equity Risk Premia using Information on Volatility of the Fundamentals: A Supervised Dynamic Orthogonal Component (sDOC) Approach

1.1 Introduction

The equity risk premium is a key quantity in finance, as it is the return on top of the amount obtained by investing in a risk-free asset, thus compensating investors for incurring risk when they invest in an asset that is not risk-free.

Estimating the equity risk premium accurately is fundamental to constructing optimal portfolios. This estimation can be achieved using the historical average of excess stock returns or by forecasting excess stock returns.[1] For example, in the latter case, studies such as Lettau and Ludvigson (2001a), Campbell and Thompson (2008) and Gu et al. (2020) show that relying on excess stock return forecasting can generate economic gains in asset portfolio allocation. However, while many have reported convincing evidence that their methods produce good *in-sample* forecast performance (Ferson & Harvey, 1993; Keim & Stambaugh, 1986; Lettau & Ludvigson, 2001a; Pontiff & Schall, 1998), forecasting excess stock returns *out-of-sample* remains a challenging task, notably due to the high volatility of stock returns (Dai et al., 2021; Gu et al., 2020; Zhang et al., 2019). The objective of this paper is to take on this challenge.

[1] The excess stock return is generally defined as the log return minus the 3-month Treasury bill rate.

1.1.1 Out-of-sample excess returns forecasting using a small number of features

The original Capital Asset Pricing Model (CAPM) argued that only a stock's systematic risk should matter for its returns. However, some time ago, this theory was challenged by numerous studies, providing evidence that additional factors also play a significant role in explaining expected stock returns. These empirical contributions typically use a relatively small set of predictors to forecast either the cross-section of excess stock returns out-of-sample (Fama & French, 2008; Lewellen, 2014) or the excess stock market return out-of-sample (Campbell & Thompson, 2008; Rapach & Zhou, 2013). For example, Campbell and Thompson (2008) employ a time-series regression of excess stock market returns on a small set of macroeconomic variables, such as earnings yields, the Treasury bill rate, and the term spread. They show evidence that, under certain constraints, these predictors improve out-of-sample forecasts of excess stock market returns relative to the historical average.[2] As another example, Lewellen (2014) applies the two-pass regression framework of Fama and MacBeth (1973) to the cross-section of excess stock returns, using three lagged firm characteristics—size (market capitalization), book-to-market ratio, and momentum. He finds that this model provides strong cross-sectional predictive power for realized returns.

There are also co-movements in the structure of a set of predictors,[3] which are driven by a few latent common factors not explained by the market (or not captured by systematic risk). Models with latent factors have become popular since, and despite few being included, they account for a substantial portion of the variation in portfolio stock returns (Giovannelli et al., 2021).

These latent common factors are typically estimated using dimension reduction methods.[4] The main motivation behind dimension reduction is to address the risk of overfitting noise, which arises when the number of predictors approaches or exceeds the number of observations. This issue leads to biased estimates, high variance, and unstable forecasts. Traditional methods—such as the Fama and MacBeth (1973) cross-sectional regression and time-series

[2]The constraints refer to weak restrictions imposed on the signs of coefficients and return forecasts. The historical average return serves as one of the benchmark indicators used to evaluate the performance of predictors in excess stock market return forecasting.

[3]Such datasets may contain macroeconomic variables, firm characteristics, textual data, etc.

[4]Dimension reduction methods (such as Principal Component Analysis, Partial Least Squares, and Dynamic Factor models) are useful for forecasting models that include many predictive variables.

regression—become infeasible or inconsistent as the number of predictors increases substantially. Technically, dimension reduction generates a small number of common factors either by using all predictors without reference to the target variable (as in Principal Component Analysis (PCA) or Dynamic Factor models), or by considering the linear dependence between the target variable and all predictors (as in Partial Least Squares (PLS)).[5]

1.1.2 Out-of-sample excess returns forecasting using machine learning and dimension reduction methods

The asset pricing literature has shown some favorable evidence when relying on dimension reduction methods and machine learning (ML) algorithms when they are used to forecast both the cross-section of excess stock returns and the time series excess stock market returns out-of-sample.

On the methodological side, dimension reduction methods as suggested by, for example, Stock and Watson (2002), employ PCA on large datasets to construct consistent common factors and generate predictions. Unlike PCA, which produces static factors, Forni et al. (2005) incorporate dynamics into principal components to develop a Generalized Dynamic Factor (GDF) model. Forni et al. (2015, 2017) further extend this framework by allowing for an infinite-dimensional factor space, thereby relaxing restrictions on lead-lag relationships among variables and common factors.

On the empirical side, dimension reduction based work has focused on both the time series and the cross-section of excess stock returns. For time series regression models, Ludvigson and Ng (2007) construct dynamic factors from a large number of quarterly macroeconomic and financial variables to forecast the excess stock market returns out-of-sample. They find evidence of predictability in quarterly excess stock market returns. Lettau and Ludvigson (2010) extend the previous work to forecast excess bond return out-of-sample. They also find evidence of forecasting excess bond return using dynamic factors. Çakmaklı and van Dijk (2016) use factor-augmented predictive regression models (with dynamic factors as interest variables) to predict monthly U.S. excess stock market returns out-of-sample, over the period 1975-2014. They find that dynamic common factors better forecast historical average and

[5]PCA and Dynamic Factor models are considered unsupervised methods. In these approaches, the common factors are linear combinations of the predictors, with Dynamic Factor models additionally incorporating lags. In contrast, PLS is a supervised method, where the common factors are derived from a linear regression framework and incorporate information from both the predictors and the target variable.

macroeconomic variables in terms of their market timing abilities. Giglio and Xiu (2021) study out-of-sample forecasting of excess stock market returns using common factor information extracted from a large sample of macroeconomic variables via the generalized dynamic factor model. They find that GDF model helps to forecast the excess stock market return out of sample.

For cross-sectional regression models, Light et al. (2017) apply the PLS method to extract a small number of common factors from a large set of firm characteristics to predict excess stock returns. They find evidence that PLS-derived factors have forecasting power in the cross-section of excess returns. Kelly et al. (2019) employ Instrumental Principal Component Analysis (IPCA) to estimate and test a factor pricing model with time-varying coefficients. Their results show that a small number of common factors explain the cross-section of excess returns more effectively than other leading factor models. Lettau and Pelger (2020) develop risk premia PCA to identify factors with limited time-series variation that nonetheless account for the cross-section of stock returns. Finally, Giglio and Xiu (2021) apply Supervised Principal Component Analysis (sPCA) to study cross-sectional stock returns. This approach involves first selecting a set of test assets and constructing common factors from the predictor sample, and then estimating and testing these factors using cross-sectional regressions with the chosen assets. They find that latent factors explain the cross-section of excess stock returns.

For machine learning algorithms applied to excess stock market returns, Rapach and Zhou (2013) provide evidence supporting the use of the Least Absolute Shrinkage and Selection Operator (LASSO) for forecasting aggregate excess stock returns, employing numerous lagged returns from multiple countries as predictors. Hutchinson et al. (1994) and Yao et al. (2000) provide evidence in favor of Neural Networks for forecasting derivatives prices. Khandani et al. (2010) and Butaru et al. (2016) show the effectiveness of regression trees in predicting consumer credit card delinquencies and defaults.

For the cross-section of stock returns, Moritz and Zimmermann (2016) show that tree-based models are effective for estimating portfolio sorting. Kozak et al. (2020) provide evidence in favor of shrinkage methods for approximating a stochastic discount factor for expected returns. Freyberger et al. (2020) use selection methods to approximate nonlinear functions for expected returns.

Most recent studies (Gu et al., 2020; Kelly et al., 2024; Kynigakis & Panopoulou, 2022; Xu & Liu, 2024) focus on comparative analyses of machine learning methods for forecasting excess stock returns out-of-sample. Gu et al. (2020) compare the performance of dimension reduction techniques (PCA and PLS) with nonlinear machine learning methods (Random Forest (RF) and Neural Networks) for both cross-sectional and aggregate excess return forecasts, using a large dataset of firm characteristics. They find that Random Forest and Neural Networks with three layers outperform the other methods, both statistically and economically, due to their ability to capture nonlinear interactions among predictors.

Kynigakis and Panopoulou (2022) evaluate the economic gains of dynamic dimension reduction methods compared to a wide range of machine learning algorithms and their combinations for stock return prediction using a large set of predictors. Their results show that shrinkage methods and Neural Networks with three layers (NN3) deliver the strongest individual forecasting performance, while combinations of machine learning models yield superior out-of-sample gains.

Xu and Liu (2024) evaluate several machine learning models—including those from Gu et al. (2020), as well as additional methods such as Ridge Regression, Support Vector Regression, K-Nearest Neighbors, and Extreme Gradient Boosted Trees—for forecasting the equity risk premium. Using a dataset comprising macroeconomic and technical indicators, they find that although tree-based models perform well in-sample, they generally fail to outperform the historical average benchmark out-of-sample, consistent with the findings of Welch and Goyal (2007). This underperformance is attributed to the limited sample size and the inherently low signal-to-noise ratio in equity premium forecasting. Their analysis highlights that interest rate-related variables are among the most influential predictors; however, the historical average still yields the highest returns and Sharpe ratio, confirming its robustness.

Kelly et al. (2024) examine the theoretical and empirical behavior of machine learning-based portfolios, particularly under conditions of high model complexity. They demonstrate that, contrary to the traditional preference for parsimonious models, highly parameterized approaches (such as ridgeless least squares) can yield positive Sharpe ratio improvements even when out-of-sample R^2 (R^2_{OoS}) is low or negative. This shifts the focus from predictive accuracy to economic value. Their findings suggest that employing rich, nonlinear models with many predictors can significantly enhance market timing performance, even in small-sample settings.

1.1.3 Motivation

In terms of out-of-sample R^2 (R^2_{OoS}), the consensus on the success of the above models in terms of out-of-sample forecasting is that this statistic is, at best, around 10%, although some authors—such as Welch and Goyal (2007)—argue that it is likely closer to zero. As a result, recent studies have focused on incorporating additional sources of information into the pool of predictors to enhance forecastability. Although Principal Component Aanalysis is widely used to parsimoniously capture the relevant information, it has certain limitations that may undermine its effectiveness in improving excess return predictability:

(1) The set of predictors may contain many irrelevant variables. As a result, dimension reduction methods may underperform compared to machine learning or deep learning approaches, since the presence of irrelevant variables introduces substantial variation unrelated to the target variable. In contrast, machine learning and deep learning models often mitigate this issue by automatically down-weighting or eliminating the influence of irrelevant features. Therefore, pre-selecting the most relevant predictors before applying a dimension reduction method is crucial to avoid this problem.

(2) "Principal components are contemporaneously uncorrelated; however, lagged cross-correlation may be nonzero, conditional correlation may be nonzero, and cross-correlation of nonlinear transformations, such as the squared process may be nonzero" (Matteson & Tsay, 2011). Failure to take this cross-dependence into account could result in redundant information that could bias estimates and misjudge volatility (Matteson & Tsay, 2011).

(3) Principal components capture only the information at the mean level of the predictors (Matteson & Tsay, 2011).

Since stock returns are highly volatile, it makes sense to incorporate information from both the mean and volatility levels of the predictors for out-of-sample return forecasting. However, it is important to do so while avoiding the aforementioned shortcomings. A key challenge is that volatilities are latent. The Dynamic Orthogonal Component (DOC) method (Matteson & Tsay, 2011) offers a way to construct common factors that capture the joint volatility structure of predictive variables. DOC is a dimension reduction technique that extracts uncorrelated common factors across all leads and lags, capturing co-movements not only at the mean and volatility levels, but also at higher moments of multiple time series variables.

Finally, in the recent studies mentioned above, dimension reduction and machine learning methods are primarily applied to cross-sectional excess stock return prediction. An important assumption in the cross-sectional setting is that the same model governs the relationship between predictors and future returns for each firm (Gu et al., 2020). This assumption simplifies computation when working with a relatively small number of stocks. However, applying rolling-window forecasts using machine learning methods to over 2,000 stocks can be extremely time-consuming. Therefore, aggregating stock data at the industry level becomes a practical solution for two key reasons: (1) it makes the problem computationally feasible, and (2) it helps avoid possible survivorship bias, as the dataset contains many inactive or "zombie" stocks that cannot be ignored in empirical analyses.

1.1.4 Contributions

Given the above, we propose a novel methodology—the Supervised Dynamic Orthogonal Component (sDOC) method—to obtain short-term excess industry-level stock return forecasts. We then compare its performance against several benchmark approaches, including Principal Component Analysis, univariate GARCH(1,1), and two machine learning models: a Neural Network with three layers and Random Forest regression.[6]

Our approach consists of two main stages. First, we apply supervised machine learning techniques—such as Random Forest regression—to identify variables that are useful and relevant for the forecasting exercise. Second, we apply the Dynamic Orthogonal Component method on the selected variables to extract uncorrelated common factors across time lags and leads, which parsimoniously capture their joint volatility structure. Each extracted factor is then modeled as a GARCH(1,1) process to account for its nonlinear dynamic properties. These volatility-aware components are subsequently used as predictors for out-of-sample excess industry-stock return forecasting. Finally, we compare the statistical and economic forecasting performance of the proposed sDOC method against PCA, NN3, Random Forest regression, and univariate GARCH(1,1), demonstrating its superior accuracy and robustness.

[6]We focus on short-horizon forecasts because investors prefer short-term stock return predictions, as high volatility and risk make long-term forecasts less reliable and harder to act on. We will explore the effectiveness of sDOC for longer-horizon forecasts in future work.

A great advantage of using the DOC method is the following. When using a regression model with multiple time series regressors, it is essential to address both temporal and cross-sectional dependencies among the predictors. In statistical analysis, orthogonal components are commonly employed to capture the underlying co-movements among time series by extracting the maximum shared information in a lower-dimensional representation (Matteson & Tsay, 2011). The more commonly used method, Principal Component Analysis, constructs common factors that primarily capture variation at the mean level. However, these factors are only contemporaneously orthogonal and may still exhibit temporal correlation. As a result, even when using multiple PCA-based factors, one must still model their dynamic interdependencies over time. By contrast, the Dynamic Orthogonal Component method is more flexible. It can extract common factors that reflect not only mean-level variation but also volatility and potentially higher-order moments of a large set of predictors. Crucially, the factors constructed via DOC are orthogonal both contemporaneously and across time, which significantly enhances forecast accuracy.

For practitioners, the flexibility of the DOC method offers a powerful tool for uncovering latent structures and time-varying dependencies in high-dimensional financial or economic data. For example, a portfolio or risk manager could use DOC to estimate common volatility components across a group of assets, such as equities or bonds. By identifying these underlying volatility drivers or systemic shocks, the manager gains deeper insights into market interdependencies and systemic risk exposure. This understanding can inform a range of strategic decisions, including improved asset allocation, enhanced diversification strategies, the design of robust stress-testing frameworks, and the development of early warning systems for financial instability. Moreover, the ability of DOC to generate temporally uncorrelated factors enhances the robustness of forecasting models, particularly in environments characterized by structural breaks or regime changes.

Matteson and Tsay (2011) evaluated the performance of the DOC method in financial and economic applications, including out-of-sample forecasts of multivariate volatility series for daily stock returns (S&P 500 Index, Cisco Systems, and Intel Corporation) and forecasts of multivariate GDP for five countries (South Korea, the United Kingdom, Canada, the United States, and Taiwan). Their statistical evaluations demonstrated that: (1) DOC outperforms PCA in forecasting multivariate GDP series, and (2) DOC outperforms PCA, Dynamic Conditional Correlation (DCC), and Conditional Uncorrelated Components (CUC) in predicting equity return volatilities.

After obtaining our sDOCs, we benchmark this method's performance against the widely used Principal Component Analysis. We show that sDOC consistently outperforms PCA in terms of out-of-sample R^2 across a majority of industries, with (positive) gains ranging from 0.08% to 3.94%. Unlike PCA, which ignores nonlinear and volatility-based interactions among predictors, sDOC explicitly models these relationships—allowing it to extract stronger predictive signals. In this way, sDOC addresses the limitations of linear factor models and enhances predictive power in a structured and interpretable way.

We note that our method differs fundamentally from the high-complexity, overparameterized strategies proposed by Kelly et al. (2024). While as Kelly et al. (2024) argue that model complexity alone—such as Ridgeless Least Squares—can improve economic performance even when R^2_{OoS} is negative, sDOC shows that structured feature selection combined with volatility-aware dynamic components can produce high statistical and economic performance without relying on over-parameterization. In contrast to their emphasis on theoretical complexity, sDOC provides practical robustness, interpretability, and empirical outperformance in both calm and turbulent periods.

Moreover, relative to the findings in Xu and Liu (2024), which show that many machine learning models fail to outperform the historical average due to low signal-to-noise ratios and small datasets, sDOC demonstrates statistical robustness even under these challenging conditions. Our model incorporates volatility information and predictor interactions to strengthen the signal, and the Mincer-Zarnowitz R^2 values confirm its consistent predictive accuracy across industries. This contrasts with the limited out-of-sample success observed in traditional tree-based or kernel methods used by Xu and Liu (2024).

We show that sDOC delivers strong economic gain through improved risk-adjusted returns. In most industries, sDOC outperforms PCA, RF, and GARCH(1,1) in terms of Sharpe ratio, Sortino ratio, gain to pain ratio, and Calmar ratio. Notably, sDOC and NN3 exhibit similar performance in many industries, suggesting that both methods effectively capture complex nonlinear relationships relevant to portfolio performance. These metrics reflect the methods' ability to manage different forms of risk, including downside risk and drawdowns. The robustness of these results across multiple return-based measures underscores the practical value of sDOC in portfolio construction and risk management.

Finally, we show that sDOC is flexible and reliable across different market regimes, including both tranquil periods and high-volatility phases such as the Covid-19 crisis. Its rolling R^2_{OoS} can reach 1.5%, confirming its ability to adapt to structural breaks and changing economic conditions—something many traditional models and even some machine learning methods fail to do.

The remainder of this paper is organized as follows. Section 1.2 describes the sDOC method used to forecast short-term excess industry-level stock return out-of-sample, defines the time series regression model, presents the data and the benchmark models, defines the rolling-window forecasting method, and provides the statistical and economic evaluation criteria. Section 1.3 provides the empirical results. Section 1.4 concludes.

1.2 Methodology

1.2.1 Description of the sDOC method

The objective of this section is to construct dynamic orthogonal components for the joint volatility (DOCs-vol) of the most relevant predictors, which are pre-selected from the original dataset. A GARCH(1,1) model is then applied to each component to extract the nonlinear dynamics information contained in these factors.[7] These estimated volatility processes are then used as strong predictors for excess industry-level returns.

Step 1: The selection of variables relevant for the forecasting exercise

We consider multiple industries, and for each, we use a stationary, standardized target variable and a large set of stationary, standardized predictors.[8] The target variable is the excess industry-level return, while the predictors are the aggregate firm characteristics at the industry level.[9] We begin by implementing the Random Forest regression method to pre-select

[7]We use the GARCH model developed by Bollerslev (1986) to describe volatility. However, other approaches—such as models from the GARCH family, stochastic volatility models, or nonparametric and semiparametric methods—can also be used to characterize volatility functions.

[8]See Table 1.13 of the Appendix for the meaning of industries.

[9]In Table 1.12 of the Appendix, we define the firm characteristics and in section 1.2.3.1, we explain the construction of the predictors and the target variable.

the most important predictors.[10] We then use a Python tool called GridSearchCV (from the Skearn library) to automatically test different model settings and find the ones that make the Random Forest model work best. Once we have the best settings, we build the final Random Forest model. This model helps us understand which predictors are the most useful for making predictions. Each predictor is assigned a score that shows how important it is to predict the target variable. Based on these scores, we choose the 30 most important predictors to use in our analysis.

Step 2: The Construction of uncorrelated processes

From the selected variables, we construct uncorrelated variables using Principal Component Analysis.[11] We use these uncorrelated processes instead of the pre-selected predictors directly because—even within a selected subset—the predictors may still exhibit high multicollinearity, which can lead to suboptimal forecasts. For example, consider forecasting a zero-mean variable Y using 10 predictors, each equal to Y plus independent and identically distributed *i.i.d* noise. In this case, the first principal component—essentially the average of these predictors—has lower variance than any individual predictor. Therefore, forecasting with all the predictors or even a subset may be inferior to using only the principal component, as the mean squared forecast error is lower when using the latter.

Step 3: The Construction of DOC factors

From the principal components, we apply the Dynamic Orthogonal Components (DOC) method by constructing two distinct sets of common factors: DOCs in mean (DOCs-mean) and DOCs in volatility (DOCs-vol). DOCs-mean are factors whose cross-covariance matrices are diagonal at all time lags, while DOCs-vol are factors whose cross-covariance matrices of squared values are diagonal at all time lags.

More precisely, we begin by decomposing each principal component into a linear relationship consisting of its conditional mean and an idiosyncratic error term, which may exhibit cross-correlation. Since the conditional mean of a principal component is not directly observable, we approximate it using the estimated DOC-mean. This DOC-mean

[10] The Random Forest regression combines multiple independent regression trees using sub-samples drawn from the original data set. Trees try to find groups of observations with similar behavior and thus have similar outcomes. For each tree, we compute the regression defined as the mean of outcomes within the partitions. The aggregation of these regressions is the output of the Random Forest regression. The mean idea behind the Random Forest is that, according to the central limit theorem, averaging multiple independent trees can reduce the variance which makes it less vulnerable to over-fitting. See section 1.5.1 of the Appendix for the algorithm description.

[11] See section 1.5.2 of the Appendix for a description of the PCA method

represents the latent common factor underlying the principal component. Accordingly, we can express the DOC-mean as a function of the principal component. As explained by Matteson and Tsay (2011), because the DOC-mean and the principal component are uncorrelated by construction, the relationship between the vector of DOCs-mean and the vector of principal components can be algebraically represented by an orthogonal matrix, which corresponds to a rotation in a space of dimension equal to the number of principal components. Thus, estimating the DOC-means amounts to estimating the parameters of this rotation matrix.

To do this, we construct a vector of all lagged cross-covariances between pairs of distinct DOCs-mean and build a diagonal weighting matrix corresponding to the lags. The weights are higher for shorter lags and lower for longer lags, reflecting the empirical observation that cross-correlations are typically stronger at shorter time horizons. We then combine the weighting matrix and the vector of lagged cross-covariances to construct the objective function as follows:

$$\underset{\gamma}{min} \overline{f}(\gamma)' \Xi \overline{f}(\gamma), \tag{1.1}$$

where $\overline{f}(\gamma)$ is a vector of lagged cross-covariances, Ξ is a weighting matrix and γ is a vector of parameters. This objective function means that the optimal $\widehat{\gamma}$ is determined by minimizing the linear dependence across components and across time. In other words, the objective function generates the optimal parameters such that the cross-covariance matrices of DOCs-mean are orthogonal. Then, the estimated DOCs-mean are:

$$\widehat{s}_{i,t} = s_{i,t}(\widehat{\gamma}). \tag{1.2}$$

Second, we compute the residual idiosyncratic errors and focus on their conditional covariance matrix to construct DOCs-vol. Specifically, we begin by applying PCA to the residual idiosyncratic errors in order to obtain uncorrelated residuals, which serve as the inputs for constructing DOCs-vol.

Following the same procedure used for DOCs-mean, each DOC-vol is expressed as a linear transformation of the uncorrelated residuals. This transformation is represented by an orthogonal matrix that performs a rotation in the vector space defined by the uncorrelated residuals. Consequently, estimating the DOCs-vol amounts to estimating the parameters of this rotation matrix.

To do so, we construct a vector of all lagged cross-covariances between pairs of squared DOCs-vol. We then build a diagonal weighting matrix in which the weights are higher for shorter lags and lower for longer lags, reflecting the empirical observation that lagged cross-correlations tend to be stronger at shorter horizons. Finally, we combine the weighting matrix with the vector of lagged cross-covariances to construct the following objective function:

$$\underset{\gamma}{min} \boldsymbol{C}(\gamma)' \Xi \boldsymbol{C}(\gamma), \tag{1.3}$$

where, $\boldsymbol{C}(\gamma)$ is a vector of lagged cross-covariances, Ξ is a weighting matrix and γ is a vector of parameters. This objective function means that the optimal value of γ is determined by minimizing the quadratic dependence across components and across time. In other words, the objective function generates the optimal parameters such that the cross-covariance matrices of squared DOCs-vol are orthogonal. Then, the estimated DOCs-vol are:

$$\widehat{z}_{i,t} = z_{i,t}(\widehat{\gamma}). \tag{1.4}$$

For a technical description of the DOCs-mean and DOCs-vol, see section 1.5.3 of the Appendix.

Step 4: The Construction of volatility functions as features for forecasting:

For each DOC-vol factor, we use a GARCH(1,1) model to capture how its variability changes over time.

$$h_{j,i,t} = \delta_{i,0} + \delta_{i,1} \widehat{z}^2_{j,i,t-1} + \delta_{i,2} h_{j,i,t-1}, j = 1, ..., c, \tag{1.5}$$

where, $\delta_{i,d} > 0$, $d \in \{0, 1, 2\}$ and $(\delta_{i,1} + \delta_{i,2}) < 1$ to ensure the positiveness and the stationary of the conditional variance. j represents the jth nonlinear function of a DOC-vol for the industry $i \in \{1, 2, ..., N\}$ and c is the number of DOC-vol. We call these volatility functions DOC-GARCH(1,1) functions.

1.2.2 Linear regression model with GARCH(1,1) error

Consider the one-period-ahead industry-level excess return $(r_{i,t+1})_{t=1,...,(T_i-1)}$, $i \in \{1, ..., N\}$ such that

$$r_{i,t+1} = \alpha_0 + \alpha_1' h_{i,t} + \varepsilon_{i,t+1} \tag{1.6}$$

$$\varepsilon_{i,t+1} = \underline{\varepsilon}_{i,t+1} V_{i,t+1}^{1/2} \tag{1.7}$$

$$V_{i,t+1} = a_0 + a_1 \varepsilon_{i,t}^2 + a_2 V_{i,t}, \tag{1.8}$$

where α_0 is the intercept, $h_{i,t}$ is a $(c \times 1)$ vector of non-linear functions corresponding to DOCs-vol for the industry i and α_1 is a $(c \times 1)$ vector of parameters associated with $h_{i,t}$. $\varepsilon_{i,t+1}$ is a forecast error term and $\underline{\varepsilon}_{i,t+1}$ is a standardized variable that follows a $N(0,1)$ distribution. $E(\varepsilon_{i,t+1}|h_{i,t}) = 0$, $Var(\varepsilon_{i,t+1}|h_{i,t}) = V_{i,t+1}$ is a conditional variance of the error term that follows a GARCH(1,1) process. $a_i > 0$ for $i \in \{0,1,2\}$ and $\Sigma_{i=1}^2 a_i < 1$.

1.2.3 Data and benchmark models

1.2.3.1 Data

We use U.S. monthly data on 2,026 firms collected from three sources—CRSP, Compustat, and IBES—accessed through WRDS.[12] Only firms with at least ten years of available data through 2020 are included. As a result, the sample spans the period from February 28, 1994, to December 31, 2020. We exclude firms in the bottom 5th percentile by size, stocks with very low prices in the prior period, and inactive firms.

For each firm, we obtain the excess short-term stock return, 87 firm-level characteristics, and 76 industry dummies, following Green et al. (2017). Industries are defined using the last two digits of the SIC code, which identifies a company's primary business activity. The description of firm characteristics is provided in Table 1.12 of the Appendix, and the industry classification is provided in Table 1.13 of the Appendix.

We then group the firms into industries. As shown in Table 1.13 of the Appendix, 63 industries contain at least one firm, while 13 industries have no firms and are excluded from the analysis. Within each industry that includes firms, we aggregate data as follows:

1. Lag firm characteristics and firm size (market capitalization) by one month to avoid look-ahead bias.

2. Compute weighted firm characteristics and weighted excess short-term firm-level returns by multiplying each variable by the firm's lagged size.

[12] CRSP: Center for Research in Security Prices. IBES: Institutional Brokers' Estimate System. WRDS: Wharton Research Data Services

3. Calculate weighted averages across firms in each industry. The weighted average excess short-term return represents the excess industry-level return, which serves as the target variable. The weighted average firm characteristics are used as predictors.

To ensure the validity of time series modeling, we test the stationarity of the aggregated series using both the Augmented Dickey-Fuller (ADF) test and the KPSS (Kwiatkowski-Phillips-Schmidt-Shin) test. The results from both tests are consistent. For series identified as non-stationary, we apply the first-difference transformation to achieve stationarity. Then, all variables are standardized before analysis.

Finally, in each industry, we select the most important predictors using the random forest regression model and include the one-month lagged excess industry-level return as an additional predictor to account for first-order autocorrelation in returns.

The dynamic orthogonal components are constructed using R and C++ packages, while the remainder of the analysis is conducted in Python 3.

1.2.3.2 Benchmark models

We compare the performance of our method with four widely used benchmark models: Principal Component Analysis (PCA), Random Forest (RF) regression, a Neural Network with 3 layers (NN3), and the univariate GARCH(1,1) model. The first model (PCA) is linear, the second and third models (RF and NN3) are nonlinear, while the last model (GARCH(1,1)) features a linear mean equation and a nonlinear variance equation.

Since DOC is an extension of PCA that generates uncorrelated factors across leads and lags, it is particularly relevant to compare the forecasting performance of sDOC with PCA in the context of excess industry-level return predictions. Additionally, prior studies such as Gu et al. (2020) have demonstrated the forecasting strength of RF and NN3 on aggregate excess stock returns. Therefore, it is informative to compare the performance of sDOC with that of RF and NN3 in forecasting excess industry-level returns.

Finally, the GARCH(1,1) model assumes a constant conditional mean. Comparing our method to GARCH(1,1) allows us to highlight the benefits of incorporating dynamic interactions in the conditional mean, a feature explicitly modeled by sDOC.

1.2.4 Rolling window forecasting

We use a rolling window approach to make one-period ahead forecasts, with a fixed window size $T_{e,i} \in [85, 300]$ (i represents an individual industry).[13] Using data from $T_{0,i}$ up to the time the forecast is made to generate one-period ahead forecasts as follows: if the first observation is at time $l_{e,i} = T_{0,i}$, then, the first forecast uses the data from the time interval $[l_{e,i}, l_{e,i} + (T_{e,i} - 1)]$ to obtain an out-of-sample forecast return at time $l_{e,i} + T_{e,i}$; the second forecast uses the data from the time interval $[l_{e,i} + 1, l_{e,i} + T_{e,i}]$ to produce a forecast return at time $l_{e,i} + T_{e,i} + 1$ and so on. We estimate our model using the maximum likelihood (ML) method for each data interval.

1.2.5 Performance evaluation

We analyze the forecasting performance of our method across 63 excess industry-level returns.

1.2.5.1 Statistical evaluation

1) Out-of-sample R^2 (R^2_{OoS})

We follow Campbell and Thompson (2008) to compute the out-of-sample R^2 statistic (R^2_{OoS}). In financial forecasting, R^2_{OoS} is widely used to evaluate predictive performance in out-of-sample settings (Campbell & Thompson, 2008; Gu et al., 2020). Unlike traditional metrics such as RMSE or MSE, which assess absolute prediction errors, R^2_{OoS} directly compares the performance of the forecasting model against a benchmark, typically the historical mean, offering a more intuitive measure of relative forecast effectiveness.

Let $\widehat{r}_{i,t+1}$, $t = l_{e,i} + T_{e,i}, ..., T_i - 1$ represent the 1-period forecasts associated with the actual observations $r_{i,t+1}, t = l_{e,i} + T_{e,i}, ..., T_i - 1$. So, we have:

$$R^2_{OoS} = 1 - \frac{\sum_{t=l_{e,i}+T_{e,i}}^{T_i-1} a^2_{i,t+1}}{\sum_{t=l_{e,i}+T_{e,i}}^{T_i-1} (r_{i,t+1} - \overline{r}_i)^2}, \tag{1.9}$$

where $a_{i,t+1} = r_{i,t+1} - \widehat{r}_{i,t+1}$ and $i \in \{1, 2, ..., N\}$. $\overline{r}_i$ is the rolling average excess return corresponding to the industry i.

[13]As in Timmermann (2008), we use a fix window length depending on the size of each industry to take into account the structural break when we predict the sign of excess stock return.

If $R^2_{OoS} > 0$, it indicates that the sum of squared prediction errors from the predictive regression is smaller than the sum of squared deviations between the actual excess returns and their rolling (historical) average return.[14]

It should be noted that a forecasting model can sometimes yield valuable economic insight even when its R^2_{OoS} is negative. For example, Kelly et al. (2024) show that a mean-variance trading strategy based on such forecasts can still be profitable. This highlights the importance of complementing statistical evaluation with measures of economic gain, especially when assessing the effectiveness of machine learning methods in financial applications.

2) Mincer-Zarnowitz R^2 (Forecast evaluation measure)

The Mincer-Zarnowitz R^2 is a commonly used metric to assess the quality of forecasts. It is based on the Mincer-Zarnowitz regression (Mincer & Zarnowitz, 1969), which regresses the actual outcomes in their corresponding forecasts:

$$y_t = \alpha + \beta \hat{y}_t + \varepsilon_t, \tag{1.10}$$

where y_t is the actual realized value, $\hat{y}_t$ is the forecasted value, and ε_t is the forecast error. The coefficient of determination (R^2) from this regression is known as the Mincer-Zarnowitz coefficient R^2. It measures how well the forecasts explain the variation in the actual outcomes.

There is no strict universal threshold for the Mincer-Zarnowitz R^2 that guarantees "good" forecasting performance — its interpretation depends on the context of the application, the data frequency, and the predictability of the variable. Generally, $R^2 > 0.5$ is often interpreted as reasonably good in many applied settings, particularly if the forecasting task is difficult (e.g., financial returns). An R^2 close to 1 indicates strong explanatory power of the forecasts and, thus, better forecasting performance.

1.2.5.2 Economic evaluation

Consider an investor with mean-variance utility who constructs a portfolio consisting of an aggregate risky asset and a risk-free asset. At each time t, the investor's objective function is:

$$\underset{w_{t+1}}{Max} E_t[r_{p,t+1}] - \frac{1}{2}\varrho Var_t[r_{p,t+1}], \tag{1.11}$$

[14]We use a rolling average to represent the historical mean, as it is a common benchmark in out-of-sample prediction studies on aggregate excess returns.

where, w_{t+1} is the proportion of the portfolio allocated to the risky asset, ϱ is the degree of relative risk aversion (RRA) and $r_{p,t+1}$ is the portfolio return defined as:

$$r_{p,t+1} = r_{f,t+1} + w_{t+1} r_{t+1}, \tag{1.12}$$

r_{t+1} denotes the excess industry-level return in period $t+1$, and $r_{f,t+1}$ is the risk-free return. $E_t[r_{p,t+1}]$ and $Var_t[r_{p,t+1}]$ are the expected value and the conditional variance (volatility) of the portfolio return, respectively. The optimal portfolio weight for the given investor is:

$$w^*_{t+1} = \frac{\widehat{r}_t}{\varrho \widehat{Var}_t}, \tag{1.13}$$

where, $\widehat{r}_t$ is the predicted one-period-ahead excess stock return based on a predictive regression model, and $\widehat{Var}_t$ is the predicted one-period-ahead volatility of the regression residuals using data up to time t.

I) Risk-adjusted performance metrics

1) Gain in terms of the Sharpe ratio

This section draws heavily from Campbell and Thompson (2008).

The Sharpe ratio is a widely used measure in finance that evaluates the performance of an investment by adjusting for its risk. It is defined as the ratio of the average excess return of the portfolio (i.e., average of return above the risk-free rate) to the standard deviation of that excess return.

To define the gain in terms of the Sharpe ratio, consider an investor comparing two investment strategies:

- The buy-and-hold strategy, where the investor passively holds the market portfolio over time without using any predictive signals.
- A predictive strategy, where the investor adjusts their portfolio based on information from a forecasting model (e.g., expected returns, volatility estimates).

If the investor chooses a buy-and-hold strategy, the Sharpe ratio at time t is:

$$SR_t = \frac{\overline{r}_t}{\overline{\sigma}_t}, \tag{1.14}$$

where, $\overline{r}_t$ and $\overline{\sigma}_t$ denote the average (or exponential-weighted average) and the realized volatility of the portfolio excess return, respectively, computed using data up to time t.

If the investor exploits the predictive information in its investment strategy, the Sharpe ratio is:

$$SR^* = \sqrt{\frac{SR^2 + R^2_{OoS}}{1 - R^2_{OoS}}}, \tag{1.15}$$

where SR is the average of equation (1.14).

The economic gain in terms of Sharpe ratio is then defined as the difference in Sharpe ratios between the predictive strategy and the buy-and-hold strategy. This gain reflects how much better (or worse) the predictive strategy performs relative to the passive benchmark after adjusting for risk: $SR^* - SR$.

A positive gain in terms of Sharpe ratio indicates that the predictive strategy offers superior risk-adjusted returns compared to simply holding the market.

2) Sortino ratio (SOR)

This section draws heavily from Sortino and Van Der Meer (1991).

The Sortino ratio is a risk-adjusted performance measure that evaluates the excess return of a portfolio relative to the downside risk, rather than total volatility. It refines the Sharpe ratio by considering only the standard deviation of negative returns (i.e., downside deviation), which provides a more accurate assessment of risk when investors are primarily concerned with losses. The formula is:

$$SOR = \frac{APR - rf}{ADSD}, \tag{1.16}$$

where APR is the annualized portfolio return, rf is the risk-free rate, and $ADSD$ is the annualized downside standard deviation. A Sortino ratio above 1.5 is generally considered indicative of strong risk-adjusted performance.

3) Calmar ratio (CR)

This section draws heavily from Young (1991).

The Calmar ratio is a measure of risk-adjusted performance, commonly used to evaluate investment strategies. It is defined as the ratio of the portfolio's annualized return (APR) to its maximum drawdown ($MaxDD$), capturing the trade-off between return and downside risk. The formula is:

$$CR = \frac{APR}{MaxDD}, \tag{1.17}$$

where the MaxDD measures the largest decline in the value of an investment over a given period. It is defined as the maximum observed loss from a peak to a subsequent trough before a new peak is attained. Formally, it is computed as:

$MaxDD = \underset{1 \le t_0 \le t_1}{Max}(ret_{t_0} - ret_{t_1})$. where ret_{t_0} and ret_{t_1} are the cumulative excess return of machine learning portfolio up to time t_0 and t_1 respectively. The excess return of machine learning portfolio is equal to $\frac{\widehat{r}_t r_{t+1}}{\varrho \widehat{Var}_t}$.

A higher Calmar ratio (≥ 0.5) indicates a more favorable risk-adjusted return profile, as it reflects higher returns relative to the worst peak-to-trough decline.

4) Gain to pain ratio (GPR)

This section draws heavily from Schwager (2012).

The gain to pain ratio is a risk-adjusted performance metric used to evaluate a trading strategy's efficiency. It is defined as the ratio of the sum of all positive returns to the absolute value of the sum of all negative returns over a given period.

II) Raw (absolute) performance metrics: The gain in terms of average portfolio return (GAPR)

This section draws heavily from Campbell and Thompson (2008).

The gain in terms of average portfolio return (GAPR) quantifies the improvement in investment performance when an investor uses a predictive signal compared to a passive buy-and-hold strategy. Specifically, it is defined as the difference between the average excess portfolio return earned by an investor who observes the predictor and the average excess portfolio return earned by a buy-and-hold investor who does not use the predictor.

When the investor uses the predictor, the excess portfolio return at time $t+1$ is given by:

$$a_{t+1} = \frac{\widehat{r}_t r_{t+1}}{\varrho \widehat{Var}_t}, \tag{1.18}$$

where $\widehat{r}_t$ is the forecasted excess return based on the predictor, r_{t+1} is the realized return, $\widehat{Var}_t$ is the forecasted variance, and γ is the investor's risk aversion coefficient.

When the investor does not use the predictor (buy-and-hold strategy), the excess portfolio return at time $t+1$ is given by:

$$b_{t+1} = \frac{\overline{r}_t r_{t+1}}{\varrho \overline{var}_t}, \tag{1.19}$$

where $\overline{r}_t$ and $\overline{var}_t$ are the historical average excess return and variance, respectively.

The gain in terms of average portfolio return is then calculated as:

$$GAPR = \overline{a}_{t+1} - \overline{b}_{t+1}. \tag{1.20}$$

1.3 Empirical results

1.3.1 Empirical analysis based on the full data sample

1.3.1.1 Pre-selection of important predictors and construction of volatility functions for the prediction of the excess industry-level return

Using the Random Forest selection approach, we retain—within each industry—the top 30 predictors with an importance score of at least 1.2. These scores reflect each feature's relative contribution to the model's predictive accuracy: the higher the score, the greater the predictor's influence. Intuitively, the variables selected by the Random Forest are those with stronger predictive power compared to the remaining features.

From the set of pre-selected variables, we include the first lag of excess industry-level stock return as an additional feature to capture the persistence in return dynamics. Next, we apply PCA to the feature set and retain the 10 most important principal components, which together explain 90% of the total variance. These orthogonal components serve as the input variables for estimating the DOCs-mean factors.

As described in Section 1.2.1, we express each principal component as a function of its conditional mean plus an error term. The conditional mean is estimated using the DOC approach, and we obtain the DOCs-mean. We then compute the residuals from these estimates and apply PCA again to obtain a new set of uncorrelated residual components, which are used as input variables for the DOCs-vol factors.

As with DOCs-mean factors, we estimate each DOCs-vol component $z_{i,t}$. To capture potential nonlinear and time-varying volatility dynamics in each component, we model them using univariate GARCH(1,1) processes. The resulting volatility series serve as informative predictors for forecasting excess industry-level stock returns.

1.3.1.2 Statistical performance measures

I) Analysis based on the out-of-sample R^2 (R^2_{OoS}) measure

Table 1.1: Industry-average R^2_{OoS} in %

Statistic	sDOC	RF	PCA	NN3	GARCH(1,1)
mean	0.344	-3.619	-0.055	0.637	-1.294
std	1.651	3.192	0.264	0.942	1.424
min	-3.700	-14.030	-1.040	-1.360	-6.900
25%	-0.825	-5.190	-0.160	0.045	-2.180
50%	0.320	-3.650	-0.090	0.720	-0.870
75%	1.050	-2.550	0.110	1.265	-0.215
max	3.940	4.220	0.600	3.060	1.030

This table reports the average (across all industries) R^2_{OoS} in %. Industry-specific R^2_{OoS} are presented in Table 1.6 of the Appendix.

Based on the summary statistics (Table 1.1), the sDOC method clearly outperforms both PCA and GARCH(1,1) on average, with a higher mean R^2_{OoS} (0.344% vs -0.055% for PCA and -1.294% for GARCH(1,1)), as well as superior median and upper-quartile values. Although sDOC exhibits greater variability (standard deviation of 1.651%), its performance is generally closer to that of NN3 (mean = 0.637%), suggesting that both methods capture useful predictive signals across industries. Moreover, both sDOC and NN3 significantly outperform the RF approach, which shows the lowest average performance (mean = -3.62%), indicating their stronger and more consistent forecasting ability.

The superior performance of sDOC compared to PCA can be attributed to sDOC's ability to capture nonlinear interactions among predictors during the feature selection stage, as well as its use of volatility interactions to forecast a volatile target variable. In contrast, PCA fails to account for volatility interactions when constructing predictive factors. The univariate GARCH(1,1) model assumes a constant conditional mean, whereas our forecasting approach extends beyond this by incorporating factors that capture dynamic interactions in the conditional mean, thereby enhancing predictive performance. While the random forest model does capture nonlinear relationships, its handling of predictor interactions may be too diffuse or insufficiently structured to effectively model the relevant dependencies.

To provide further detail, Table 1.6 in the Appendix reports the industry-specific R^2_{OoS} performance of the sDOC, PCA, univariate GARCH(1,1), RF, and NN3 models. The results show that sDOC and NN3 outperform the benchmark—the rolling average of the excess industry-level return—in forecasting excess returns. Specifically, the NN3 model achieves a positive R^2_{OoS} in 49 out of 63 industries, while the sDOC model does so in 38 industries. In contrast, PCA, RF, and GARCH(1,1) yield positive R^2_{OoS} values in only 24, 9, and 5 industries, respectively.

Additionally, Table 1.6 shows that sDOC outperforms GARCH(1,1), PCA, RF, and NN3 in 43, 38, 45, and 21 industries, respectively, based on higher R^2_{OoS} values. NN3 outperforms GARCH(1,1), PCA, RF, and sDOC in 62, 48, 56, and 42 industries, respectively. PCA outperforms GARCH(1,1) and RF in 53 industries, while RF outperforms GARCH(1,1) in only 18 industries. These results confirm that sDOC and NN3 statistically outperform the other methods, while RF and GARCH(1,1) exhibit the weakest overall performance.

II) Analysis based on the Mincer-Zarnowitz R^2 measure

Based on the summary statistics of the Mincer-Zarnowitz R^2 (Table 1.2), the sDOC method clearly outperforms all other approaches on average, with a mean R^2 of 0.243—an order of magnitude higher than PCA (0.023), NN3 (0.022), RF (0.015), and GARCH(1,1) (0.008). sDOC also exhibits the highest maximum (0.965) and median (0.122), indicating not only superior average performance but also greater consistency across industries. While PCA and NN3 display similar mean and dispersion values, their medians and upper quartiles remain substantially lower than those of sDOC. Compared to RF and GARCH(1,1), sDOC achieves both higher central tendency and wider variability, suggesting it captures more meaningful structure in the data. Despite its greater standard deviation (0.275), the strong median and upper-tail performance of sDOC reinforces its advantage over linear models like PCA, volatility-only models like GARCH(1,1), and machine learning alternatives such as RF and NN3.

Table 1.2: Industry-average Mincer-Zarnowitz R^2

Stataistic	sDOC	PCA	NN3	RF	GARCH(1,1)
mean	0.243	0.023	0.022	0.015	0.008
std	0.275	0.026	0.026	0.015	0.010
min	0.000	0.000	0.000	0.000	0.000
25%	0.035	0.005	0.005	0.004	0.001
50%	0.122	0.017	0.013	0.011	0.004
75%	0.365	0.034	0.028	0.020	0.011
max	0.965	0.121	0.151	0.059	0.052

This table reports the average (across all industries) Mincer-Zarnowitz R^2. Industry-specific Mincer-Zarnowitz R^2 are presented in Table 1.7 of the Appendix.

To provide further detail, Table 1.7 in the Appendix reports the Mincer-Zarnowitz R^2 performance of the sDOC, PCA, GARCH(1,1), RF, and NN3 models across all industries. Among these, sDOC outperforms GARCH(1,1), PCA, RF, and NN3 in 58, 52, 54, and 52 industries, respectively, based on higher R^2 values. NN3 performs better than GARCH(1,1), PCA, and RF in 45, 33, and 34 industries, respectively. PCA outperforms RF and GARCH(1,1) in 34 and 46 industries, respectively. RF surpasses GARCH(1,1) in 41 industries.

These results highlight the robustness and superior explanatory power of sDOC's forecasts, as reflected in its consistently higher Mincer-Zarnowitz R^2 values across the majority of industries.

1.3.1.3 Economic performance measures

For economic analysis, we consider an investor with a relatively low relative risk aversion, with the related parameter $\varrho = 3$.

I) Absolute performance metrics: Average excess portfolio return gain (GAPR)

Table 1.3: Industry-average GAPR

Statistic	sDOC	RF	PCA	NN3	GARCH(1,1)
mean	2.716	0.004	0.207	1.304	-0.319
std	2.314	0.614	1.833	3.650	0.900
min	-0.620	-1.480	-0.610	-1.020	-6.792
25%	0.175	-0.215	-0.245	-0.175	-0.322
50%	1.410	0.040	-0.050	0.060	-0.206
75%	3.125	0.290	0.105	0.265	-0.143
max	38.550	3.070	14.030	42.860	1.313

This table reports the average (across all industries) of average excess portfolio return gain (GAPR). Industry-specific GAPR are presented in Table 1.8 of the Appendix.

Table 1.3 shows that on average, the sDOC method delivers the highest average excess portfolio return gain (GAPR) among all predictive models, with a mean of 2.716, clearly outperforming PCA (0.207), RF (0.004), and GARCH(1,1) (-0.319). While NN3 has a higher maximum value (42.860 vs sDOC's 38.550), its average return (1.304) is lower than that of sDOC, making sDOC the best-performing model overall in terms of reward generation. GARCH(1,1) performs the worst, with a negative mean GAPR, suggesting that it often fails to improve upon the buy-and-hold benchmark. The consistently higher median (1.410) and upper quartile (3.125) of sDOC also reflect strong and stable performance across industries.

The intuition behind sDOC's superior performance likely lies in its ability to extract informative signals from the interaction of features and their joint volatility. Unlike PCA, which captures only linear combinations of variables, or GARCH(1,1), which assumes a constant conditional mean, sDOC employs volatility-driven variable selection to identify predictors most relevant for forecasting excess returns. This enables sDOC to generate more accurate forecasts, leading to improved investment timing and asset allocation, and ultimately resulting in greater realized gains. For investors, this translates into a strategically adaptive method that not only enhances average returns but also delivers more stable outperformance, making it a valuable tool for active portfolio management.

Table 1.8 in the Appendix shows that sDOC and NN3 generate higher average excess portfolio returns compared to the buy-and-hold strategy in the majority of industries. The average gains (GAPR) range from 0.02% to 38.55% for sDOC and from 0.04% to 42.86% for NN3. sDOC outperforms all other methods in the majority of industries, while NN3 delivers the highest performance in only 14 industries.

II) Risk-adjusted returns

1) Sharpe ratio

Table 1.4: Industry-average Sharpe ratio gain in %

Statistic	**sDOC**	**RF**	**PCA**	**NN3**	**GARCH(1,1)**
mean	0.423	-2.786	-0.052	0.554	-2.240
std	1.580	3.164	0.269	1.000	2.489
min	-2.380	-10.740	-1.100	-1.440	-8.460
25%	-1.025	-4.550	-0.160	0.000	-2.990
50%	0.190	-2.970	0.000	0.220	-1.600
75%	1.225	0.000	0.095	1.335	-0.560
max	4.030	4.330	0.610	3.120	3.350

This table reports the average (across all industries) Sharpe ratio gain in %. Industry-specific Sharpe ratio gains are presented in Table 1.9 of the Appendix.

The summary statistics of the Sharpe ratio gain (Table 1.4) indicate that the sDOC method outperforms both PCA and GARCH(1,1). With an average Sharpe ratio gain of 0.423%, sDOC significantly exceeds the performance of PCA (mean = -0.052%) and GARCH(1,1) (mean = -2.240%), demonstrating superior risk-adjusted return characteristics. The distribution of sDOC's gains, including a median of 0.190% and an upper quartile of 1.225%, highlights its consistent value across industries. Furthermore, sDOC's performance is quite close to NN3 (mean = 0.554%), with both methods exhibiting higher upper-tail outcomes and significantly better averages than RF (mean = -2.786%). This suggests that sDOC and NN3 are the most effective in enhancing risk-adjusted returns, with sDOC providing a robust balance between performance and variability across industry applications.

In more detail, Table 1.9 in the Appendix reports the Sharpe ratio gains achieved by the predictive methods—sDOC, PCA, RF, NN3, and GARCH(1,1)—relative to the buy-and-hold benchmark strategy across all industries. Among them, NN3, sDOC, PCA, and RF deliver higher risk-adjusted returns than the buy-and-hold strategy in 41, 34, 21, and 9 industries, respectively. Notably, in 9 industries—including educational services (ind_82), engineering, accounting, research, M., and R.S (ind_87), metal mining (ind_10), communications (ind_48), wholesale trade-durable goods (ind_50), electronic and other electrical equipment and C. (ind_36), industrial and commercial machinery and C.E (ind_35), measuring, analyzing and controlling instruments (ind_38), chemicals and allied products (ind_28)—the Sharpe ratio gain is zero for all predictive methods except GARCH(1,1). This indicates that in these industries, the predictive strategies do not outperform the buy-and-hold strategy in terms of risk-adjusted returns.

Table 1.9 also shows that the sDOC strategy outperforms PCA, NN3, RF, and GARCH(1,1) in 33, 22, 44, and 44 industries, respectively. Likewise, the NN3 model delivers higher Sharpe ratio gains than PCA, RF, sDOC, and GARCH(1,1) in 41, 47, 32, and 45 industries, respectively. Finally, PCA performs better than RF and GARCH(1,1) in 42 and 46 industries, while RF surpasses GARCH(1,1) in only 21 industries.

2) Gain to pain, Calmar, and Sortino ratios

Table 1.5: Industry-average GPR, CR, and SOR

Panel A: Gain-to-pain ratio (GPR)					
Statistic	sDOC	PCA	RF	NN3	GARCH(1,1)
mean	1.471	0.056	-0.046	0.094	0.007
std	1.432	0.866	0.312	0.800	0.017
min	-0.123	-0.649	-0.620	-1.000	-0.038
25%	0.027	-0.303	-0.274	-0.249	0.000
50%	1.810	-0.084	-0.065	-0.049	0.004
75%	2.749	0.121	0.127	0.161	0.013
max	4.010	6.126	0.990	3.500	0.071
Panel B: Calmar ratio (CR)					
Statistic	sDOC	PCA	RF	NN3	GARCH(1,1)
mean	1.321	-0.031	-0.023	0.081	-0.091
std	1.424	0.166	0.093	0.432	0.412
min	-1.300	-0.200	-0.170	-0.180	-0.963
25%	-0.002	-0.100	-0.087	-0.100	-0.326
50%	1.003	-0.080	-0.037	-0.030	-0.128
75%	2.750	0.001	0.013	0.044	0.070
max	3.930	0.770	0.280	1.880	2.150
Panel C: Sortino ratio (SOR)					
Statistic	sDOC	PCA	RF	NN3	GARCH(1,1)
mean	1.453	0.002	-0.077	0.112	-0.060
std	1.516	0.772	0.321	0.868	0.419
min	-0.251	-0.963	-0.961	-0.990	-0.571
25%	-0.009	-0.262	-0.257	-0.242	-0.265
50%	1.171	-0.126	-0.085	-0.099	-0.123
75%	2.769	0.116	0.120	0.203	0.065
max	5.020	5.295	0.623	3.860	2.604

This table reports average (across all industries) Gain-to-pain ratio (GPR), Calmar ratio (CR), and Sortino ratio (SOR). Industry-specific values of these risk-adjusted returns are presented in Tales 1.10 and 1.11 of the Appendix.

Based on Table 1.5, the sDOC method outperforms all other predictive models on average across the three risk-adjusted performance metrics—gain to pain ratio (GPR), Calmar ratio

(CR), and Sortino ratio (SOR). With mean values of 1.471 (GPR), 1.321 (CR), and 1.453 (SOR), sDOC significantly exceeds PCA, RF, NN3, and GARCH(1,1), all of which report either negative or near-zero average values. While NN3 is the second-best performer (mean SOR = 0.112), its performance remains substantially below that of sDOC. RF and GARCH(1,1) are the weakest overall, exhibiting consistently negative or negligible values across all three ratios. The strong and consistent performance of sDOC across these diverse risk-adjusted metrics—where the Calmar ratio captures drawdown risk, the Sortino ratio targets downside volatility, and the gain to pain ratio penalizes cumulative losses—highlights the robustness of the sDOC approach in delivering superior and stable returns across industries.

1.3.2 Time-varying forecasting performance across market regimes: sDOC during Covid-19 vs tranquil periods

This subsection evaluates the performance of the sDOC method by analyzing the Covid-19 period separately and comparing it to a tranquil period. The goal is to assess whether the model delivers measurable performance gains under high-volatility conditions compared to low-volatility environments. The Covid-19 period, which started in early 2020 and ended in the U.S. around May 2023, was marked by significant market turbulence and heightened volatility, as global financial markets experienced dramatic declines and rapid recoveries in response to the pandemic and associated economic shutdowns. The unprecedented nature of this shock makes it a critical stress test for assessing the robustness and adaptability of forecasting models such as sDOC.

Due to the relatively limited time span of our original dataset, we are unable to isolate and compare the long Covid-19 period against a tranquil period in a fully segmented manner. As a result, we adopt a dynamic evaluation approach using the rolling out-of-sample R^2 metric. This technique allows us to track the predictive performance of the sDOC model over time and observe how its effectiveness evolves, particularly during periods of financial stress such as the Covid-19 crisis.

In this analysis, we benchmark sDOC against the widely used models PCA, GARCH(1,1), and RF—to determine whether its performance gains are statistically significant. By focusing on this high-volatility sub-period, we aim to understand whether sDOC's integration of volatility interactions enables it to better adapted to extreme market conditions and to deliver superior performance.

1.3.2.1 The rolling R^2_{OoS} measure

The takeaways of this subsection is that, on average, sDOC outperforms other models in both turbulent and calm market conditions (the R^2_{OoS} can reach 1.5%), showcasing its flexibility and reliability. While PCA becomes more effective during tranquil periods, GARCH(1,1) and RF struggle to provide accurate and stable forecasts across both regimes.

Section 1.5.6 of the Appendix presents rolling R^2_{OoS} plots for the predictive methods sDOC, PCA, RF, and GARCH(1,1). To assess their predictive performance across different market regimes, we compare the behavior of the rolling R^2_{OoS} during the Covid-19 period and a tranquil period (e.g., pre-2020). This analysis is feasible in 32 industries, as in some cases the rolling R^2_{OoS} is available only for 2020. Specifically during 2020, we observe that sDOC has the best forecast performance on average, highlighting its effectiveness in periods of heightened market volatility.

During tranquil periods—characterized by low volatility conditions—all models exhibit relatively improved performance, but differences in predictive strength remained evident. The sDOC method consistently generates strong and positive out-of-sample R^2 values, indicating that it also performs well in low-volatility periods. PCA shows better performance during tranquil times, with more stable and moderately positive R^2_{OoS} values. Its reliance on static linear factor structures aligns better with market conditions when structural relationships remain relatively unchanged. GARCH(1,1) demonstrates slight performance, as volatility is smoother and more predictable in these periods. However, it is lagged behind sDOC and PCA due to its limited ability to model dynamic predictor interactions. RF struggles, producing noisy and inconsistent forecasts. In this stable conditions, its performance is weak, likely due to the absence of time-dependent structure and the challenges of capturing meaningful economic relationships without additional feature engineering.

In contrast, during the Covid-19 period, marked by extreme market volatility and structural breaks (especially in early 2020), the models diverges more sharply in performance. Once again, sDOC demonstrates on average, the strongest and most stable predictive power, consistently delivering positive R^2_{OoS} values. This performance highlights sDOC's ability to adapt to sudden regime changes, likely due to its flexible modeling of nonlinear interactions and volatility dynamics among predictors. PCA, on the other hand, shows modest and less consistent performance, with occasional dips in forecast accuracy. Its static structure struggles to adjust quickly to the rapidly shifting market environment. GARCH(1,1) performs poorly, frequently yielding negative R^2_{OoS} values. This result is expected, as its assumptions of persistent and smooth volatility are violated during this period of frequent and extreme shocks. RF shows the most unstable and largely negative performance. Despite being nonlinear and flexible, RF lacks a temporal structure, making it especially vulnerable in rapidly evolving conditions like the Covid-19 shock.

Overall, these results reinforce that sDOC is the most robust model, performing well in both stable and crisis periods, while other models tend to be regime-dependent and less adaptive to sharp shifts in market dynamics.

1.4 Conclusion

This study assesses the value of using Supervised Dynamic Orthogonal Components (sDOC) to forecast excess returns at the industry level. We construct 63 U.S. monthly industry equity portfolios and compare the forecasting performance of sDOC against PCA, univariate GARCH(1,1) model, and two widely used machine learning methods—Random Forest regression (RF) and a Neural Network with three layers (NN3). The statistical results show that, on average, sDOC outperforms PCA and GARCH(1,1) in terms of out-of-sample R^2. sDOC and NN3 demonstrate similar predictive performance and both outperform RF. Additionally, in terms of Mincer-Zarnowitz R^2, sDOC achieves on average, the highest performance among all methods. From an economic perspective, sDOC and NN3 perform similarly on average in terms of Sharpe ratio gains and both surpass the remaining methods. However, for the other risk-adjusted metrics—gain to pain ratio, Calmar ratio, Sortino ratio-as well as the gain in terms of average excess portfolio return, sDOC outperforms all other models, followed by NN3.

sDOC's ability to capture nonlinear interactions and volatility structures explains its consistent performance across both calm and volatile market regimes. Moreover, the strong risk-adjusted returns across diverse risk measures and the robust Mincer-Zarnowitz R^2 confirm the method's statistical and economic relevance.

Future research could explore the effectiveness of sDOC for longer-horizon forecasts of excess industry-level returns, such as 3-, 6-, or 12-month ahead predictions. Extending the framework to longer periods would allow for an assessment of how sDOC captures slower-moving risk-return dynamics and whether its volatility-aware structure remains effective beyond short-term horizons.

1.5 Appendix: Technical Details of sDOC, Tables, and Figures

1.5.1 Use of AI in research

In this study, OpenAI ChatGPT and Grammarly were used to check the grammar, catch typos, and enhance the writing to improve the clarity of the text.

1.5.2 Random forest regression for forecasting

Let the relation,

$$r_{i,t+1} = f(\underline{X}_{i,t}) + u_{i,t+1}, \tag{1.21}$$

where $r_{i,t+1}$ is a stationary excess industry-level (i) return at time $t+1$, $i \in \{1, ..., N\}$, $t \in \{1, ..., T_i - 1\}$ and $\underline{X}_{i,t} = (x_{i,t}^{(1)}, ..., x_{i,t}^{(l_0)})$, denote a stationary l_0-dimensional vector of predictors.[15] $f(\underline{X}_{i,t})$ is a function of predictors that can be linear or not and $u_{i,t+1}$ is an error term. We apply the random forest regression on this model in each rolling window as follows:

1. Split the original sample to training/validation/test set.
2. Fine-tuning hyperparameters for the random forest regression model.
3. Draw a bootstrap sample of size less than or equal to the entire sample size.
4. Grow a decision tree from each bootstrap sample.
5. At each node, randomly select a subset of predictors without replacement.
6. Split the node and the corresponding branches using the predictor that provides the best split using the threshold obtained by minimizing the impurity for each branch (minimizing the forecast error for each branch).
7. Repeat steps $\{3, 4, 5, 6\}$ l_1 times to create l_1 independent trees.
8. Aggregate the estimations produced by these l_1 independent trees.

[15] In Section 1.2.3.1, we define and explain how to construct the vector of predictors and aggregate excess stock return.

1.5.3 Principal Component Analysis (PCA)

Given the most important features, we construct the uncorrelated variables as inputs of DOCs-mean.

Let $X_{i,t} = (x_{i,t}^{(1)}, ..., x_{i,t}^{(\bar{l})})$ the $(T_i - 1) \times \bar{l}$ matrix of pre-selected predictors and $M_i = (m_1, ..., m_c)$ is the $\bar{l} \times c$ matrix of weights for the industry i, with $c < \bar{l}$. m_c represents the vector of linear combination weights used to construct the c^{th} uncorrelated process. The $(T_i - 1) \times c$ matrix of uncorrelated processes is given by $Y_{i,t} = X_{i,t} M_i$ where the vector of linear combination of weights is defined as:

$$\underset{M}{\text{Argmax}}\, \text{Var}(X_t M), \;\; \text{s.t } M_i^\top M_i = 1, \;\; cov(X_{i,t} m_c, X_{i,t} m_{\underline{l}}) = 0, \;\; \underline{l} = 1, 2, .., c-1.$$

Intuitively, PCA seeks c linear combinations of $X_{i,t}$ such that the variables in $Y_{i,t}$ are uncorrelated and contain the maximum information from the pre-selected predictors.

1.5.4 Construction of DOCs-mean and DOCs-vol

We base this on the results of Section 2.3 of Matteson and Tsay (2011). We first decompose the uncorrelated features as a sum of conditional mean and the idiosyncratic vector of errors as:

$$Y_{i,t}^{'} = E(Y_{i,t}{'} \mid \underline{Y}_{i,t-1}^{'}) + e_{i,t}, \tag{1.22}$$

where $t = 1, ..., T$, $Y_{i,t}^{'} = (y_{1,t}, ..., y_{c,t})'$ is the c-dimensional vector of uncorrelated features, $E(Y_{i,t}^{'} \mid \underline{Y}_{i,t-1}^{'})$ is the conditional mean and $\underline{Y}_{i,t-1}^{'} = \{Y_{i,t-1}^{'}, Y_{i,t-2}^{'}, ...\}$. $e_{i,t} = (e_{1,i,t}, ..., e_{c,i,t})'$ is the vector of idiosyncratic errors and $\Sigma_t = cov(Y_{i,t}^{'} \mid \underline{Y}_{i,t-1}^{'}) = cov(e_{i,t} \mid \underline{Y}_{i,t-1}^{'})$ is the conditional covariance matrix. $E(Y_{i,t}^{'} \mid \underline{Y}_{i,t-1}^{'})$ is unknown and the idiosyncratic errors are cross-correlated.

Second, we estimate the conditional mean using DOCs-mean in three steps:

Step 1. Since DOCs-mean are common latent factors of uncorrelated features, we can write DOCs-mean as a function of uncorrelated features as:

$$s_{i,t} = DY_{i,t}^{'}, \tag{1.23}$$

where $s_{i,t} = (s_{1,i,t}, ..., s_{c,i,t})'$ is DOCs-mean vector. The problem now is to find the value of D. Without loss of generality, since $Y_{i,t}^{'}$ are uncorrelated and standardized, its variance-covariance matrix is equal to an identity matrix. This information is also applied to DOCs-mean and we have:

$$I_c = cov\{s_{i,t}\} = Dcov\{Y_{i,t}^{'}\}D^{'} = DI_cD^{'} = DD^{'}. \tag{1.24}$$

Equation (1.24) implies that D is an orthogonal matrix. This means D represents a rotation in the space of dimension c. Therefore, we can parameterized it by the length $b = \frac{c(c-1)}{2}$ vector γ of rotations angles as D_γ.

Step 2. The parameterization is as follow: for any orthogonal matrix D that belongs to the group of all $c \times c$ orthogonal matrices with determinant equal to +1 and for $\gamma \in (-\pi, +\pi]^b$, D is equivalent to:

$$D_\gamma = R^{(b)} ... R^{(2)}, \tag{1.25}$$

in which $R^{(o)} = R_{1,o}(\gamma_1^o)...R_{o-1,o}(\gamma_{o-1}^o)$, where for $i \neq j$, $R_{i,j}(\beta)$ is a Givens rotation matrix, i.e the identity matrix I_c such that the (i,i) and (j,j) values are substituted by $cos(\beta)$, the (i,j) value is substituted by $-sin(\beta)$, and the (j,i) value is substituted by $sin(\beta)$. The relation between s_t and Y_t becomes:

$$s_t(\gamma) = D_\gamma Y_t^{'}. \tag{1.26}$$

It remains to estimate the parameter vector γ.

Step 3. We first construct a vector of all lagged cross-covariances between two distinct DOCs-mean indexed by d and j as:

$$\overline{f}(\gamma) = (\overline{f}_{d,j}^{p}(\gamma))_{p,d,j}^{'}, \tag{1.27}$$

with

$$\overline{f}_{d,j}^{p}(\gamma) = \frac{1}{T}\Sigma_{t=1}^{T} s_{d,i,t}(\gamma)s_{j,i,t-p}(\gamma) - \frac{1}{T}\Sigma_{t=1}^{T} s_{d,i,t}(\gamma)\frac{1}{T}\Sigma_{t=1}^{T} s_{j,i,t-p}(\gamma), \tag{1.28}$$

where, the lag $p \in \underline{\mathbb{N}}_0 \subset \mathbb{N}_0$. $\underline{\mathbb{N}}_0$ including the subscript 0, which means that, equation (1.28) takes into account contemporaneous correlation between DOCs-mean. The length of $\overline{f}(\gamma)$ is $v = b(2 \mid \underline{\mathbb{N}}_0 \mid -1)$, where $\mid \underline{\mathbb{N}}_0 \mid$ is the cardinal of $\underline{\mathbb{N}}_0$. Example, if $c = 2$ and $\underline{\mathbb{N}}_0 = \{0, 1\}$, the length of $\overline{f}(\gamma)$ is 3 as $\overline{f} = (cov(s_{1,i,t}, s_{2,i,t}), cov(s_{1,i,t}, s_{2,i,t-1}), cov(s_{1,i,t-1}, s_{2,i,t}))'$.

Then, we build a diagonal weighting matrix associated with the lagged correlation such as weight is high for a small lag and low for a big lag, reflecting the fact that lagged cross-correlation is strong at low lag. This matrix of weights is defined as:

$$\Xi = diag\{\Xi_{p_1}, ..., \Xi_{p_1}, \Xi_{p_2}, ..., \Xi_{p_2}, ..., \Xi_{p_{|\mathbb{N}_0|}}, ..., \Xi_{p_{|\mathbb{N}_0|}}\}, \tag{1.29}$$

where $\Xi_p = \{\frac{1 - p/ \mid \underline{\mathbb{N}}_0 \mid}{\Sigma_p(1 - p/ \mid \underline{\mathbb{N}}_0 \mid)}\}/(b + bI_{\{p\neq 0\}})$, for $p \in \underline{\mathbb{N}}_0$ and $I_{\{.\}}$ the indicator function.

Finally, we define the objective function as:

$$\min_{\gamma} \overline{f}(\gamma)'\Xi\overline{f}(\gamma). \tag{1.30}$$

This objective function means that the optimal value of γ is determined by minimizing the linear dependence across components and across time. In other words, the objective function generates the optimal parameters such that the cross-covariance matrices of DOCs-mean are diagonal. Then, the estimated DOCs-mean are:

$$\widehat{s}_{i,t} = s_{i,t}(\widehat{\gamma}) = D_{\widehat{\gamma}}Y_{i,t}'. \tag{1.31}$$

Third, we focus on the conditional covariance matrix of the residual idiosyncratic errors $\widehat{e}_{i,t} = Y_{i,t}' - \widehat{s}_{i,t}$ to describe interactions between volatilities. But the volatility is not directly observable because it is always changing. So, from the residual idiosyncratic errors, DOC can construct DOCs-vol in two steps.

Step 1. We apply the PCA to $\widehat{e}_{i,t}$ to construct uncorrelated residual error $\widehat{v}_{i,t}$. Then, following the step 1 of the construction of DOCs-mean, the relation between $\widehat{v}_{i,t}$ and DOCs-vol $z_{i,t}$ can be written as:

$$z_{i,t} = D\widehat{v}_{i,t}. \tag{1.32}$$

Without loss of generality, $cov\{z_{i,t}\} = cov\{\widehat{v}_{i,t}\} = I_c$. This implies that D is a diagonal matrix since:

$$I_c = cov\{z_{i,t}\} = Dcov\{\widehat{v}_{i,t}\}D' = DD'. \tag{1.33}$$

Following step 2 of the parameterization of the matrix D, the relation between $z_{i,t}$ and $\widehat{v}_{i,t}$ becomes:

$$z_{i,t}(\gamma) = D_{\gamma}\widehat{v}_{i,t}. \tag{1.34}$$

Now, it remains to estimate the parameter vector γ for DOCs-vol.

Step 2. We first construct a vector of all lagged cross-covariances between two distinct squared DOCs-vol indexed by d and j as:

$$\boldsymbol{C}(\gamma) = (\boldsymbol{C}_{d,j}^{p}(\gamma))_{p,d,j}', \tag{1.35}$$

$$\boldsymbol{C}_{d,j}^{p}(\gamma) = \frac{1}{T}\Sigma_{t=1}^{T} z_{d,i,t}^{2}(\gamma) z_{j,i,t-p}^{2}(\gamma) - \frac{1}{T}\Sigma_{t=1}^{T} z_{d,i,t}^{2}(\gamma)\frac{1}{T}\Sigma_{t=1}^{T} z_{j,i,t-p}^{2}(\gamma), \tag{1.36}$$

where, the lag $p \in \underline{\mathbb{N}_0} \subset \mathbb{N}_0$. $\underline{\mathbb{N}_0}$ including the subscript 0, which means that, equation (1.36) takes into account contemporaneous correlation between DOCs-vol. The length of $\overline{k}(\gamma)$ is $v = b(2 \mid \underline{\mathbb{N}_0} \mid -1)$, where $\mid \underline{\mathbb{N}_0} \mid$ is the cardinal of $\underline{\mathbb{N}_0}$.

Then, we build a diagonal weighting matrix associated with the lagged correlation such as weight is high for a small lag and low for a big lag, reflecting the fact that lagged cross-correlation is strong at low lag. This matrix of weights is defined as:

$$\Xi = diag\{\Xi_{p_1}, ..., \Xi_{p_1}, \Xi_{p_2}, ..., \Xi_{p_2}, ..., \Xi_{p_{|\underline{\mathbb{N}_0}|}}, ..., \Xi_{p_{|\underline{\mathbb{N}_0}|}}\}, \tag{1.37}$$

where $\Xi_p = \{\frac{1 - p/\mid \underline{\mathbb{N}_0} \mid}{\Sigma_p(1 - p/\mid \underline{\mathbb{N}_0} \mid)}\}/(b + bI_{\{p\neq 0\}})$, for $p \in \underline{\mathbb{N}_0}$ and $I_{\{.\}}$ the indicator function.

Finally, we define the objective function as:

$$\underset{\gamma}{min} \boldsymbol{C}(\gamma)' \Xi \boldsymbol{C}(\gamma). \tag{1.38}$$

This objective function means that the optimal value of γ is determined by minimizing the quadratic dependence across components and across time. In other words, the objective function generates the optimal parameters such that the cross-covariance matrices of squared DOCs-vol are orthogonal. Then, the estimated DOCs-vol are:

$$\widehat{z}_{i,t} = z_{i,t}(\widehat{\gamma}) = D_{\widehat{\gamma}}\widehat{v}_{i,t}. \tag{1.39}$$

1.5.5 Nonlinear dynamic information from DOC-vol components

For each DOC-vol component within industry i, we estimate a GARCH(1,1) model to capture the nonlinear dynamics present in each factor as follows:

$$h_{j,i,t} = \delta_{i,0} + \delta_{i,1}\widehat{z}_{j,i,t-1}^{2} + \delta_{i,2}h_{j,i,t-1}, j = 1, ..., c, \tag{1.40}$$

where, $\delta_{i,d} > 0$, $d \in \{0, 1, 2\}$ and $(\delta_{i,1} + \delta_{i,2}) < 1$ to ensure the positiveness and the stationary of the conditional variance. j represents the jth nonlinear function of a DOC-vol for the industry.

1.5.6 Tables

1.5.6.1 Monthly industry-level R_{OoS}^2 in %

Table 1.6 reports the R_{OoS}^2 values across models and industries. A model is considered most successful if it achieves the highest average R_{OoS}^2 or outperforms others in the largest number of industries, reflecting superior out-of-sample predictive power. The most successful model for each industry is shown in bold.

Table 1.6: R^2_{OoS} across models and industries (Part 1)

Industry	**RF**	**PCA**	**sDOC**	**NN3**	**GARCH(1,1)**
Apparel and other finished products (ind_23)	-3.980	-0.110	-1.180	**0.160**	-0.050
Wholesale trade-nondurable goods (ind_51)	**1.050**	-0.170	0.520	0.350	-0.064
Chemicals and allied products (ind_28)	-2.780	-0.090	**0.320**	0.190	-0.130
Printing, publishing, and allied industries (ind_27)	-8.930	-0.090	0.320	**1.390**	-0.830
Rubber and miscellaneous plastic products (ind_30)	-5.260	**0.150**	-2.280	-0.190	-0.810
Oil and gas extraction (ind_13)	-5.680	**0.040**	-0.150	-0.780	-2.210
Petroleum refining and related industries (ind_29)	-3.320	0.350	**2.480**	1.360	-0.600
Food and kinred products (ind_20)	-5.120	0.040	2.990	**3.060**	-3.000
Heavy construction other than B.C.C (ind_16)	-5.280	0.120	0.500	**1.140**	-0.750
Fabricatel metal products, except machinery and T.P (ind_34)	-3.900	0.100	0.520	**1.810**	-1.330
Amusement and recreation services (ind_79)	-2.940	0.190	**1.860**	1.080	-0.870
Transportation by air (ind_45)	-1.790	0.250	-0.540	**1.370**	-1.220
Measuring, analyzing, and controlling instruments (ind_38)	-6.590	-0.140	0.470	**0.720**	-0.730
Stone, clay, glass, and concrete products (ind_32)	-1.300	-0.010	**1.800**	0.050	-2.740
Primary metal industries (ind_33)	-4.070	-0.330	**1.920**	1.040	-1.550
Industrial and commercial machinery and C.E (ind_35)	-14.030	0.120	0.630	**0.750**	-0.230
Transportation Equipment (ind_37)	-7.520	-0.090	**1.080**	0.930	-0.380
Electronic and other electrical equipment and C. (ind_36)	-4.110	0.170	0.690	**0.720**	-0.670
Miscellaneous manufacturing Industries (ind_39)	-1.740	0.060	**1.920**	0.280	-4.260
General merchandise stores (ind_53)	-3.440	0.100	**1.910**	0.650	-1.340
Apparel and accessory stores (ind_56)	-5.440	0.060	**1.430**	0.720	-0.520
Wholesale trade-durable goods (ind_50)	-4.190	-0.160	0.200	**0.780**	-0.870
Holding and other Investment offices (ind_67)	-3.170	0.000	0.240	**0.990**	-0.750
Home furniture, furnishings, and E.S (ind_57)	-4.290	0.010	**3.600**	-0.890	-6.900
Security and commodity Brokers, D., E., and S. (ind_62)	-3.010	-0.160	0.940	**2.250**	-2.780
Communications (ind_48)	-4.720	0.310	**0.990**	0.910	-1.110
Electric, gas, and S.S (ind_49)	-0.720	0.350	0.670	**1.430**	-1.170
Health services (ind_80)	-6.890	**-0.100**	-0.130	-0.110	-0.120
Depository institutions (ind_60)	-3.070	-0.170	-1.760	**-1.060**	-3.170
Non-depository credit institutions (ind_61)	-3.650	-0.070	**2.130**	1.530	0.050

Table 1.6: R^2_{OoS} across models and industries (Continued)

Industry	**RF**	**PCA**	**sDOC**	**NN3**	**GARCH(1,1)**
Eating and drinking places (ind_58)	-3.030	-0.080	-1.180	**0.470**	-1.700
Insurance carriers (ind_63)	-7.740	**-0.440**	-3.380	-0.960	-3.740
Motion pictures (ind_78)	**3.090**	-0.130	-1.420	0.190	-0.820
Miscellaneous retail (ind_59)	0.020	0.250	**3.180**	2.260	-2.290
Business services (ind_73)	-3.580	-0.050	0.810	**0.890**	-0.100
Educational services (ind_82)	-4.540	-0.250	-0.320	**-0.010**	-0.180
Engineering, accounting, research, M., and R.S (ind_87)	-7.130	-0.090	0.300	**0.530**	0.340
Nonclassification establishment (ind_99)	**4.220**	-0.320	0.930	1.600	-1.710
Metal mining (ind_10)	-4.510	-0.050	**0.740**	0.080	-1.290
Bituminous coal and L.M (ind_12)	-0.470	-0.140	-1.610	**0.280**	-0.430
Mining and quarrying of nonmetallic minerals, except fuels (ind_14)	-2.900	0.140	**3.940**	2.530	-3.100
Building construction general contractors and O.B (ind_15)	-2.750	0.090	1.740	**2.140**	1.030
Construction special trade contractors (ind_17)	-6.530	-0.270	0.080	**0.430**	-0.410
Tobacco products (ind_21)	-4.470	0.320	1.020	**1.950**	-2.360
Textile mill products (ind_22)	-4.330	-0.100	**3.630**	0.000	-3.700
Lumber and wood products, except furniture (ind_24)	0.930	0.600	0.560	**1.590**	0.510
Furniture and fixtures (ind_25)	1.500	0.300	-3.700	**0.220**	-1.310
Paper and allied products (ind_26)	-4.160	-0.100	-1.190	**0.820**	-1.280
Leather and Leather products (ind_31)	-2.670	-0.110	**3.850**	-0.540	-2.470
Railroad transportation (ind_40)	-4.890	-0.340	0.000	**1.250**	-0.870
Motor freight transportation, and W. (ind_42)	-6.720	-0.420	-1.630	-0.780	**0.500**
Water transportation (ind_44)	**1.200**	-0.170	0.000	0.930	-2.760
Transportation services (ind_47)	-9.620	-0.380	-1.400	**0.050**	-0.700
Building materials, hardware, garden supply, and M.H.D (ind_52)	-3.910	-0.160	-1.110	**0.060**	-0.120
food stores (ind_54)	-2.430	-0.720	**-0.130**	-0.780	-3.190
Automotive dealers and gasoline S.S (ind_55)	-2.760	-0.160	0.200	**1.790**	0.190
Insurance agents, brokers and service (ind_64)	-10.270	**-0.150**	-1.250	-0.620	-2.150
Real estate (ind_65)	1.700	0.300	-1.870	**1.280**	-0.180
Hotels, rooming houses, camps, and other lodging places (ind_70)	-0.530	-0.190	-0.060	**0.040**	-0.890
Personal services (ind_72)	-5.420	-0.050	-0.360	**1.160**	0.280
Automobile repairs, services, and parking (ind_75)	-2.850	-0.470	-1.360	**0.010**	-1.290
Social services (ind_83)	**0.760**	0.190	0.190	0.000	-4.030
Miscellaneous services (ind_89)	-3.350	-1.040	-1.600	-1.360	**-0.200**

1.5.6.2 Mincer-Zarnowitz (M-Z) R^2 values

Table 1.7 reports the M-Z R^2 values across models and industries. A model is considered most successful if it achieves the highest average R^2 or outperforms others in the largest number of industries. The most successful model for each industry is shown in bold.

Table 1.7: Mincer-Zarnowitz R^2 across models and industries (Part 1)

Industry	**sDOC**	**PCA**	**NN3**	RF	**GARCH(1,1)**
Metal mining (ind_10)	**0.322**	0.053	0.013	0.014	0.001
Bituminous coal and L.M (ind_12)	**0.411**	0.041	0.013	0.013	0.011
Oil and gas extraction (ind_13)	0.003	0.017	**0.019**	0.000	0.004
Mining and quarrying of nonmetallic minerals, except fuels (ind_14)	0.004	**0.034**	0.017	0.015	0.006
Building construction general contraction and O.B (ind_15)	0.003	0.000	0.000	**0.015**	0.001
Heavy construction other than B.C.C (ind_16)	**0.050**	0.003	0.007	0.013	0.001
Construction special trade contractors (ind_17)	**0.845**	0.007	0.002	0.000	0.022
Food and kinred products (ind_20)	**0.445**	0.001	0.008	0.010	0.003
Tobacco products (ind_21)	**0.068**	0.007	0.009	0.016	0.001
Textile mill products (ind_22)	**0.231**	0.004	0.001	0.018	0.009
Apparel and other finished products (ind_23)	**0.043**	0.002	0.003	0.012	0.010
Lumber and wood products, except furniture (ind_24)	**0.757**	0.020	0.040	0.004	0.009
Furniture and fixtures (ind_25)	**0.100**	0.019	0.007	0.023	0.000
Paper and allied products (ind_26)	**0.034**	0.008	0.020	0.004	0.000
Printing, publishing, and allied industries (ind_27)	**0.161**	0.018	0.019	0.018	0.004
Chemicals and allied products (ind_28)	**0.897**	0.015	0.010	0.044	0.003
Petroleum refining and related industries (ind_29)	0.052	0.023	**0.057**	0.042	0.011
Rubber and miscellaneous plastic products (ind_30)	**0.327**	0.022	0.006	0.000	0.002
Leather and leather products (ind_31)	**0.009**	0.000	0.004	0.000	0.006
Stone, clay, glass, and concrete products (ind_32)	**0.159**	0.006	0.007	0.008	0.001
Primary metal industries (ind_33)	**0.140**	0.002	0.015	0.005	0.018
Fabricated metal products (ind_34)	0.011	**0.038**	0.025	0.009	0.016
Industrial and commercial machinery (ind_35)	0.010	0.021	0.007	0.004	**0.026**
Electronic and electrical equipment (ind_36)	**0.118**	0.000	0.000	0.017	0.015
Transportation equipment (ind_37)	**0.379**	0.048	0.025	0.004	0.005
Measuring, analyzing instruments (ind_38)	**0.086**	0.007	0.008	0.004	0.005
Miscellaneous manufacturing industries (ind_39)	**0.487**	0.005	0.029	0.000	0.017
Railroad transportation (ind_40)	0.000	0.121	**0.151**	0.059	0.000
Motor freight transportation (ind_42)	**0.217**	0.011	0.016	0.004	0.009
Water transportation (ind_44)	**0.068**	0.014	0.046	0.025	0.001

Table 1.7: Mincer-Zarnowitz R^2 across methods and industries (continued)

Industry	sDOC	PCA	NN3	RF	GARCH(1,1)
Transportation by air (ind_45)	0.000	**0.039**	0.026	0.003	0.000
Transportation services (ind_47)	**0.036**	0.001	0.001	0.029	0.003
Communications (ind_48)	**0.061**	0.028	0.027	0.029	0.052
Electric, gas, and sanitary services (ind_49)	**0.056**	0.008	0.005	0.001	0.000
Wholesale trade-durable goods (ind_50)	**0.205**	0.068	0.002	0.013	0.009
Wholesale trade-nondurable goods (ind_51)	0.031	**0.035**	0.034	0.007	0.004
Building materials, hardware, garden supply, and M.H.D (ind_52)	**0.122**	0.048	0.044	0.034	0.001
General merchandise stores (ind_53)	0.008	0.017	0.019	**0.034**	0.003
Food stores (ind_54)	**0.033**	0.002	0.003	0.008	0.002
Automotive dealers and gasoline S.S (ind_55)	**0.135**	0.004	0.001	0.000	0.034
Apparel and accessory stores (ind_56)	0.016	0.017	**0.036**	0.019	0.034
Furniture, furnishings, and E.S (ind_57)	**0.041**	0.027	0.022	0.006	0.009
Eating and drinking places (ind_58)	**0.094**	0.012	0.015	0.056	0.011
Miscellaneous retail (ind_59)	**0.605**	0.018	0.032	0.044	0.020
Depository institutions (ind_60)	**0.834**	0.003	0.004	0.001	0.018
Non-depository credit institutions (ind_61)	**0.122**	0.001	0.009	0.011	0.000
Security and commodity brokers, D., E., and S. (ind_62)	**0.719**	0.004	0.040	0.010	0.003
Insurance carriers (ind_63)	**0.360**	0.005	0.006	0.007	0.001
Insurance agents, brokers and service (ind_64)	**0.648**	0.034	0.026	0.019	0.003
Real estate (ind_65)	**0.247**	0.040	0.013	0.021	0.021
Holding and other investment offices (ind_67)	0.110	0.014	0.036	**0.037**	0.001
Hotels, rooming houses, camps, and other lodging places (ind_70)	**0.228**	0.017	0.013	0.001	0.000
Personal services (ind_72)	**0.965**	0.004	0.001	0.031	0.000
Business services (ind_73)	**0.346**	0.073	0.036	0.022	0.001
Automobile repairs, services, and parking (ind_75)	**0.899**	0.051	0.089	0.001	0.006
Motion pictures (ind_78)	0.301	0.066	**0.091**	0.015	0.014
Amusement and recreation services (ind_79)	0.009	0.002	0.000	0.014	**0.026**
Health services (ind_80)	**0.407**	0.011	0.000	0.008	0.004
Educational services (ind_82)	0.013	**0.027**	0.005	0.000	0.000
Social services (ind_83)	0.010	0.023	0.040	**0.043**	0.002
Engineering, accounting, research, M., and R.S (ind_87)	**0.115**	0.016	0.021	0.002	0.000
Miscellaneous services (ind_89)	**0.369**	0.120	0.001	0.002	0.001
Nonclassifiable establishments (ind_99)	**0.705**	0.070	0.076	0.001	0.001

1.5.6.3 Average excess portfolio return gain (GAPR)

Table 1.8 reports the GAPR values across models and industries. A model is considered most successful if it achieves the highest average GAPR or outperforms others in the largest number of industries. The most successful model for each industry is shown in bold.

Table 1.8: GAPR values across models and industries (Part 1)

Industry	**sDOC**	**PCA**	**RF**	**NN3**	**GARCH(1,1)**
Apparel and other finished products (ind_23)	**1.050**	0.210	0.040	0.200	-0.273
Wholesale non-durable goods (ind_51)	**1.120**	-0.270	0.080	-0.160	-6.792
Chemicals and allied products (ind_28)	2.070	0.360	0.420	**4.800**	-0.224
Printings, publishing and allied industries (ind_27)	**2.720**	-0.050	0.320	0.160	-0.483
Rubber and miscellaneous plastic products (ind_30)	**0.770**	-0.250	-0.580	-0.240	-0.192
Oil and gas extraction (ind_13)	**0.160**	-0.090	-0.380	-0.180	-0.196
Petroleum refining and related industries (ind_29)	0.980	**14.030**	0.030	-0.140	0.320
Food and kinred products (ind_20)	**0.480**	-0.520	-0.940	-0.770	-0.175
Heavy construction and other B.C.C (ind_16)	**1.410**	0.050	0.040	0.160	-0.453
Fabricated metal products, except machinery and T.P (ind_34)	**1.340**	-0.010	0.070	0.140	-0.290
Amusement and recreation services (ind_79)	**1.330**	0.770	0.690	0.710	-0.229
Transportation by air (ind_45)	-0.060	0.240	0.250	**0.420**	-0.165
Measuring, analyzing, and controlling instruments (ind_38)	**12.990**	0.280	0.260	0.350	-0.084
Stone, clay, glass, and concrete products (ind_32)	**2.750**	-0.160	0.440	-0.150	-0.172
Primary metal industries (ind_33)	**2.290**	-0.130	0.070	0.110	-0.403
Industrial and commercial machinery and C.E (ind_35)	**4.140**	3.100	3.070	3.160	-0.083
Transportation equipment (ind_37)	**4.300**	0.040	-0.200	0.260	-0.066
Electronic and other electrical equipment and C. (ind_36)	**5.180**	-0.120	-0.170	-0.640	-0.591
Miscellaneous manufacturing industries (ind_39)	**4.460**	0.090	0.130	0.090	-0.325
General merchandise stores (ind_53)	**5.780**	0.110	0.520	0.270	-0.255
Apparel and accessory stores (ind_56)	0.150	-0.590	-0.570	**0.360**	-0.257
Wholesale trade-durable goods (ind_50)	-0.070	-0.430	-0.530	**1.900**	-0.330
Holding and other investment offices (ind_67)	**1.050**	-0.530	-0.070	-0.310	-0.206
Home furniture, furnishings, and E.S (ind_57)	**3.250**	-0.050	0.004	-0.360	-0.142
Security and commodity Brokers, D., E., and S. (ind_62)	**0.070**	-0.210	-0.050	-0.050	-0.244
Communications (ind_48)	**15.080**	-0.610	0.330	0.120	-0.355

Table 1.8: GAPR values across models and industries (Continued)

Industry	**sDOC**	**PCA**	**RF**	**NN3**	**GARCH(1,1)**
Electric, gas, and S.S. (ind_49)	-0.080	**0.210**	-0.460	-0.460	-0.181
Health services (ind_80)	**3.160**	-0.030	-0.090	0.040	-0.170
Depository institutions (ind_60)	-0.070	-0.350	-0.820	-0.870	**0.380**
Non-depository credit institutions (ind_61)	**3.010**	-0.400	0.450	0.330	0.079
Eating and drinking places (ind_58)	**3.450**	-0.280	0.900	0.210	-0.039
Insurance carriers (ind_63)	**0.850**	0.150	-0.750	-0.670	-0.090
Motion pictures (ind_78)	0.190	0.020	0.180	**0.250**	-0.304
Miscellaneous retail (ind_59)	**3.870**	-0.100	0.390	0.340	-0.327
Business services (ind_73)	**2.160**	0.040	0.100	0.230	-0.319
Educational services (ind_82)	-0.620	-0.240	-1.140	**31.120**	-0.418
Engineering, accounting, research, M., and R.S (ind_87)	**38.550**	-0.110	0.090	0.380	-0.187
Nonclassification establishment (ind_99)	0.140	-0.050	**0.220**	-0.160	-0.205
Metal mining (ind_10)	2.040	0.420	0.060	**42.860**	-0.193
Bituminous coal and L.M (ind_12)	0.110	0.270	-0.030	**0.600**	-0.128
Mining and quarrying of nonmetallic minerals (ind_14)	3.010	-0.130	**0.640**	0.350	-0.270
Building construction general contractors and O.B (ind_15)	**2.640**	-0.430	-0.130	-0.010	-0.365
Construction special trade contractors (ind_17)	**2.250**	-0.020	-0.230	-0.030	-0.184
Tobacco products (ind_21)	-0.060	-0.160	-0.220	**0.060**	-2.048
Textile mill products (ind_22)	**2.060**	0.190	-0.160	-1.020	-0.019
Lumber and wood products (ind_24)	**3.110**	0.400	0.070	-0.340	0.317
Furniture and fixtures (ind_25)	-0.050	0.040	**0.200**	0.080	-0.144
Paper and allied products (ind_26)	-0.470	0.020	-0.550	-0.120	**-0.333**
Leather and leather products (ind_31)	**3.760**	-0.040	-0.030	-0.480	-0.141
Railroad transportation (ind_40)	**3.610**	-0.110	-0.150	0.130	-0.135
Motor freight transportation and W. (ind_42)	**1.170**	-0.040	-0.490	-0.170	-0.286
Water transportation (ind_44)	**3.300**	-0.420	0.490	-0.330	-0.238
Transportation services (ind_47)	-0.420	**0.090**	-1.480	-0.490	-0.195
Building materials, hardware, garden supply and M.H.D (ind_52)	**1.230**	-0.320	-0.790	0.030	-0.195
Food stores (ind_54)	**0.790**	-0.340	0.090	-0.050	-0.136
Automative dealers and gasoline S.S (ind_55)	**0.060**	-0.160	-0.210	0.020	-0.257
Insurance agents, brokers and service (ind_64)	-0.050	**0.100**	-1.060	-0.260	-0.552
Real estate (ind_65)	**2.930**	0.170	0.510	0.340	1.313
Hotels, rooming houses, camps and lodging places (ind_70)	**3.070**	-0.080	0.500	-0.080	-0.249
Personal services (ind_72)	**1.050**	0.220	-0.100	0.220	-0.195
Automobile repairs, services, and parking (ind_75)	**3.140**	-0.390	0.030	-0.240	-0.326
Social services (ind_83)	**3.020**	-0.340	0.570	0.160	-0.270
Miscellaneous services (ind_89)	-0.080	-0.370	**0.380**	-0.050	-0.407

1.5.6.4 Risk-adjusted performance measures

Table 1.9 reports the Sharpe ratio gain across models and industries. A model is considered most successful if it achieves the highest average Sharpe ratio gain or outperforms others in the largest number of industries. The most successful model for each industry is shown in bold.

Table 1.9: Sharpe ratio gain across models and industries (Part 1)

Industry	**RF**	**PCA**	**sDOC**	**NN3**	**GARCH(1,1)**
Apparel and other finished products (ind_23)	-3.910	0.110	**1.170**	0.160	-0.050
Wholesale trade-nondurable good (ind_51)	**1.110**	-0.180	0.550	0.370	-6.970
Chemicals and allied products (ind_28)	**0.000**	**0.000**	**0.000**	**0.000**	-0.150
Printing, publishing, and allied industries (ind_27)	-9.040	-0.090	0.330	**1.450**	-2.370
Rubber and miscellaneous plastic products (ind_30)	-5.260	0.150	**1.280**	-0.190	-0.940
Oil and gas extraction (ind_13)	-5.740	0.040	**0.160**	-0.810	-8.460
Petroleum refining and related industries (ind_29)	-3.510	0.380	**2.840**	1.480	-7.850
Food and kinred products (ind_20)	-5.050	0.040	3.050	**3.120**	-3.000
Heavy construction other than B.C.C (ind_16)	-5.260	0.120	**0.510**	0.000	-1.790
Fabricatel metal products, except machinery and T.E (ind_34)	-3.880	0.100	0.540	**1.840**	-4.270
Amusement and recreation services (ind_79)	-2.970	0.200	**1.900**	1.120	-0.870
Transportation by air (ind_45)	-1.780	0.250	0.540	**1.380**	-1.970
Measuring, analyzing, and controlling instruments (ind_38)	**0.000**	**0.000**	**0.000**	**0.000**	-1.090
Stone, clay, glass, and concrete products (ind_32)	-1.330	-0.010	**1.870**	0.050	-2.980
Primary metal industries (ind_33)	-4.210	-0.340	**2.020**	1.050	-2.904
Industrial and commercial machinery and C.E (ind_35)	**0.000**	**0.000**	**0.000**	**0.000**	-0.390
Transportation equipment (ind_37)	-7.320	-0.090	**1.110**	0.950	-0.740
Electronic and other electrical equipment and C. (ind_36)	**0.000**	**0.000**	**0.000**	**0.000**	-2.780
Miscellaneous manufacturing industries (ind_39)	-1.740	0.060	**1.960**	0.280	-6.350
General merchandise stores (ind_53)	-3.380	0.100	**1.930**	0.650	-1.420
Apparel and accessory stores (ind_56)	-5.300	0.060	**1.450**	0.720	-1.040
Wholesale trade-durable goods (ind_50)	**0.000**	**0.000**	**0.000**	**0.000**	-0.980
Holding and other Investment offices (ind_67)	-4.720	0.000	0.540	**1.510**	-0.750
Home furniture, furnishings, and E.S (ind_57)	-4.380	0.010	**3.780**	-0.920	-8.400
Security and commodity Brokers, D., E., and S. (ind_62)	-3.260	-0.170	1.030	**2.460**	-2.800

Table 1.9: SR gain across models and industries (continued)

Industry	**RF**	**PCA**	**sDOC**	**NN3**	**GARCH(1,1)**
Communications (ind_48)	0.000	0.000	0.000	0.000	0.000
Electric, gas, and sanitary services (ind_49)	-0.730	0.360	0.680	**1.460**	-1.200
Health services (ind_80)	-7.000	**-0.110**	-0.140	-0.120	-0.140
Depository institutions (ind_60)	-3.030	**-0.170**	-1.750	-1.050	-1.708
Non-depository credit institutions (ind_61)	-3.590	-0.070	**2.160**	1.550	0.180
Eating and drinking places (ind_58)	-3.030	-0.080	-1.200	**0.480**	-2.660
Insurance carriers (ind_63)	-10.180	**-0.560**	-1.380	-1.220	-5.560
Motion pictures (ind_78)	**3.140**	-0.130	-1.010	0.190	-1.600
Miscellaneous retail (ind_59)	0.020	0.250	**3.240**	2.290	-3.550
Business services (ind_73)	-3.560	-0.050	0.830	**0.910**	-0.150
Educationel services (ind_82)	**0.000**	**0.000**	**0.000**	**0.000**	-0.280
Engineering, accounting, research, M., and R.S (ind_87)	0.000	0.000	0.000	0.000	**0.340**
Nonclassifiable establishment (ind_99)	**4.330**	-0.320	0.940	1.620	-1.790
Metal mining (ind_10)	**0.000**	**0.000**	**0.000**	**0.000**	-1.590
Bituminous coal and L.M (ind_12)	-0.470	-0.140	-1.040	**0.280**	-0.670
Mining and quarrying of nonmetallic minerals (ind_14)	-2.860	0.140	**4.030**	2.560	-4.650
Building construction general contractors and O.B (ind_15)	-2.720	0.090	1.760	**2.170**	1.160
Construction special trade contractors (ind_17)	-6.340	-0.270	**0.080**	0.000	-0.040
Tobacco products (ind_21)	0.000	0.320	1.030	**1.970**	-7.480
Textile mill products (ind_22)	-4.270	-0.100	**3.700**	0.000	-6.920
Lumber and wood products, except furniture (ind_24)	0.950	0.610	0.570	1.630	**3.350**
Furniture and fixtures (ind_25)	**1.510**	0.300	-1.640	0.220	-2.300
Paper and allied products (ind_26)	-4.310	-0.100	-1.360	**0.860**	-2.800
Leather and leather products (ind_31)	-2.640	-0.110	**3.930**	-0.540	-2.530
Railroad transportation (ind_40)	-4.900	-0.350	0.000	**1.280**	-0.890
Motor freight transportation and W. (ind_42)	-6.890	-0.440	-1.820	-0.810	**0.660**
Water transportation (ind_44)	**1.210**	-0.170	0.000	0.940	-2.920
Transportation services (ind_47)	-9.580	-0.390	-1.240	**0.050**	-0.730
Building materials, hardware, garden supply, and M.H.D (ind_52)	-3.870	-0.160	-1.100	**0.060**	-0.130
Food stores (ind_54)	-2.480	-0.740	-1.180	**-0.800**	-4.000
Automotive dealers and gasoline S.S (ind_55)	-2.740	-0.160	0.200	**1.810**	0.380
Insurance agents, brokers and service (ind_64)	-10.740	-0.160	-1.760	**-0.660**	-5.000
Real estate (ind_65)	**1.720**	0.300	-1.860	1.290	-2.930
Hotels, rooming houses, camps, and other lodging places (ind_70)	-0.530	-0.190	-1.880	**0.040**	-1.050
Personal services (ind_72)	-5.340	-0.050	-2.380	**1.190**	0.280
Automobile repairs, services, and parking (ind_75)	-2.890	-0.480	-1.390	**0.010**	-4.300
Social services (ind_83)	**0.770**	0.190	0.190	0.000	-4.830
Miscellaneous services (ind_89)	-3.540	-1.100	-1.130	-1.440	**-0.450**

Tables 1.10 and 1.11 report the Calmar, Sortino, and Gain-to-Pain ratios across models and industries. For each ratio, a model is considered the most successful if it achieves the highest average value or outperforms the others in the greatest number of industries. The most successful model for each industry and each ratio is shown in bold.

Table 1.10: GPR, CR, and SOR values across models and industries (Part 1)

Industry	**sDOC**			**PCA**			**RF**			**NN3**		
	GPR	*CR*	*SOR*	*GPR*	*CR*	*SOR*	*GPR*	*CR*	*SOR*	*GPR*	*CR*	*SOR*
Apparel and other finished products (ind_23)	-0.039	-0.010	-0.051	**0.085**	**0.770**	0.109	-0.163	-0.050	-0.160	0.072	0.008	**0.120**
Wholesale non-durable good (ind_51)	**2.122**	**2.020**	**2.125**	-0.197	-0.009	-0.219	0.111	0.008	0.063	-0.164	-0.070	-0.171
Chemicals and allied products (ind_28)	-0.009	**2.006**	-0.070	-0.649	0.000	-0.224	-0.167	-0.030	-0.108	**2.500**	1.500	2.500
Printings, publishing and allied industries (ind_27)	**2.418**	**3.100**	**2.691**	-0.442	-0.100	-0.546	0.214	0.030	0.245	-0.264	-0.060	-0.262
Rubber and miscellaneous plastic products (ind_30)	-0.010	-0.100	-0.068	0.152	-0.010	0.120	-0.047	-0.020	-0.072	**0.245**	**0.070**	**0.218**
Oil and gas extraction (ind_13)	**3.220**	**3.100**	**3.317**	0.520	0.200	0.648	-0.270	-0.007	-0.210	0.499	0.130	0.327
Petroleum refining and related industries (ind_29)	2.397	**2.070**	2.492	**6.126**	0.400	**5.295**	0.057	0.001	0.069	-0.204	-0.006	-0.241
Food and kinred products (ind_20)	**2.096**	**2.030**	**2.208**	0.058	-0.070	0.086	-0.206	-0.090	-0.242	-0.023	-0.030	-0.029
Heavy construction and other B.C.C (ind_16)	**2.219**	**2.080**	**2.373**	-0.169	-0.090	-0.195	-0.205	-0.100	-0.234	-0.205	-0.080	-0.202
Fabricatel metal products, except machinery and T.P (ind_34)	0.052	**0.050**	0.073	-0.136	-0.100	-0.152	-0.380	-0.100	-0.397	-0.190	-0.090	-0.243
Amusement and recreation services (ind_79)	**2.249**	**2.020**	**2.475**	0.047	-0.090	-0.041	0.099	0.020	0.104	0.077	-0.003	0.063
Transportation by air (ind_45)	-0.021	-0.020	-0.019	-0.054	-0.060	-0.062	-0.163	-0.060	-0.189	0.084	**0.020**	0.117
Measuring, analyzing, and controlling instruments (ind_38)	**3.500**	**3.510**	**3.800**	-0.503	-0.070	-0.184	-0.377	-0.070	-0.250	-0.999	0.000	-0.144
Stone, clay, glass, and concrete products (ind_32)	**2.531**	**2.200**	**2.946**	-0.433	-0.100	-0.471	0.349	0.070	0.487	-0.456	-0.100	-0.496
Primary metal industries (ind_33)	**2.109**	**2.040**	**2.187**	-0.405	-0.100	-0.672	-0.402	-0.100	-0.708	-0.307	-0.100	-0.398
Industrial and commercial machinery and C.E (ind_35)	-0.077	-0.080	-0.078	-0.049	-0.100	**-0.017**	-0.113	-0.060	-0.094	**0.000**	**0.084**	-0.157
Transportation equipment (ind_37)	**3.051**	**3.001**	**3.040**	-0.086	-0.080	-0.091	-0.562	-0.100	-0.665	0.139	0.030	0.142
Electronic and other electrical equipment and C. (ind_36)	**4.010**	**3.020**	**5.020**	1.179	0.100	0.561	0.137	-0.030	0.047	-0.999	0.002	-0.149

Table 1.10: GPR, CR, and SOR values across models and industries (Continued)

Industry	sDOC			PCA			RF			NN3		
	GPR	*CR*	*SOR*	*GPR*	*CR*	*SOR*	*GPR*	*CR*	*SOR*	*GPR*	*CR*	*SOR*
Miscellaneous manufacturing industries (ind_39)	**3.644**	**3.300**	**3.758**	-0.062	-0.06	-0.074	-0.039	-0.028	-0.050	-0.082	-0.050	-0.103
General merchandise stores (ind_53)	**3.224**	**3.070**	**3.259**	-0.317	-0.20	-0.526	-0.292	-0.130	-0.446	-0.324	-0.160	-0.562
Apparel and accessory stores (ind_56)	0.053	0.060	0.080	-0.250	-0.160	-0.274	-0.488	-0.160	-0.311	**0.270**	**0.090**	**0.394**
Wholesale trade-durable goods (ind_50)	0.099	-0.100	-0.010	0.053	-0.002	0.028	-0.093	-0.060	-0.066	**3.500**	**1.510**	**2.550**
Holding and other investment offices (ind_67)	**0.653**	0.000	**0.254**	-0.434	-0.056	-0.159	0.405	**0.090**	0.295	0.189	0.057	0.233
Home furniture, furnishings, and E.S (ind_57)	**3.107**	**3.930**	**3.732**	-0.154	-0.070	-0.250	-0.132	-0.060	-0.152	-0.540	-0.170	-0.990
Security and commodity Brokers, D., E., and S. (ind_62)	**-0.066**	-0.050	**-0.008**	-0.100	-0.100	-0.179	-0.108	-0.060	-0.152	-0.540	-0.070	-0.162
Communications (ind_48)	-0.096	-0.038	-0.062	-0.335	-0.200	-0.212	**0.228**	**0.035**	**0.193**	-1.00	-0.085	-0.179
Electric, gas, and S.S. (ind_49)	0.068	0.070	0.030	0.230	0.0087	0.231	0.186	**0.090**	0.169	**0.268**	0.002	**0.302**
Health services (ind_80)	**3.183**	**3.077**	**2.344**	-0.181	-0.080	-0.230	-0.404	-0.130	-0.467	-0.116	-0.060	-0.147
Depository institutions (ind_60)	-0.042	-0.070	-0.048	0.356	0.110	0.401	0.128	0.010	0.087	**0.410**	**0.210**	**0.567**
Non-depository credit institutions (ind_61)	**1.844**	**1.870**	**1.766**	-0.417	-0.200	-0.508	-0.428	-0.084	-0.417	-0.380	-0.160	-0.476
Eating and drinking places (ind_58)	**3.701**	**3.090**	**3.736**	-0.361	-0.120	-0.216	0.990	0.160	0.589	0.074	0.006	0.062
Insurance carriers (ind_63)	-0.020	-0.030	-0.106	**0.510**	**0.200**	**0.465**	0.036	0.007	0.044	0.389	**0.200**	0.455
Motion pictures (ind_78)	**2.503**	**0.858**	-0.010	0.126	0.020	0.111	0.432	0.001	0.376	0.268	0.077	**0.260**
Miscellaneous retail (ind_59)	**2.760**	**2.600**	**2.325**	-0.361	-0.170	-0.726	-0.278	-0.120	-0.270	-0.375	-0.130	-0.722
Business services (ind_73)	-0.007	-0.005	-0.013	-0.157	-0.087	-0.190	-0.341	-0.120	-0.447	-0.102	-0.035	-0.099
Educational services (ind_82)	0.168	0.130	0.064	1.260	0.081	0.439	-0.102	-0.035	-0.097	**2.510**	**1.800**	**3.500**
Engineering, accounting, research, M., and R.S (ind_87)	**3.800**	**3.610**	**3.870**	-0.607	-0.100	-0.375	-0.493	-0.078	-0.307	-0.050	0.002	-0.210
Nonclassification establishment (ind_99)	0.036	0.033	0.060	-0.009	-0.110	-0.0001	**0.445**	**0.230**	**0.424**	-0.102	-0.100	-0.172
Metal mining (ind_10)	**3.003**	**3.020**	**3.152**	0.212	0.002	0.192	0.045	0.003	0.048	2.900	1.880	3.860

Table 1.10: GPR, CR, and SOR values across models and industries (Continued)

Industry	sDOC			PCA			RF			NN3		
	GPR	*CR*	*SOR*	*GPR*	*CR*	*SOR*	*GPR*	*CR*	*SOR*	*GPR*	*CR*	*SOR*
Bituminous coal and L.M (ind_12)	0.096	**0.050**	0.176	-0.074	-0.060	0.103	-0.058	-0.040	-0.060	**0.143**	-0.061	**0.188**
Mining and quarrying of nonmetallic minerals, except fuels (ind_14)	**3.076**	**3.010**	**3.103**	-0.311	-0.200	-0.513	-0.400	-0.140	-0.597	-0.272	-0.140	-0.444
Building construction general contractors and O.B (ind_15)	**2.134**	**2.043**	**2.222**	-0.271	-0.200	-0.428	-0.195	-0.095	-0.241	-0.273	-0.150	-0.433
Construction special trade contractors (ind_17)	**1.859**	**1.161**	**1.757**	-0.194	-0.096	-0.312	-0.620	-0.170	-0.961	-0.250	-0.111	-0.356
Tobacco products (ind_21)	**0.101**	**0.110**	**0.086**	-0.052	-0.098	-0.076	-0.313	-0.120	-0.391	0.052	0.003	0.083
Textile mill products (ind_22)	**2.914**	**2.900**	**2.594**	0.298	0.086	0.445	0.266	0.094	0.280	-0.247	-0.170	-0.598
Lumber and wood products, except furniture (ind_24)	**2.526**	**2.180**	**2.847**	0.268	0.065	0.380	0.568	0.140	0.317	-0.137	-0.050	-0.170
Furniture and fixtures (ind_25)	-0.062	-0.010	-0.027	-0.011	-0.057	-0.013	**0.267**	**0.160**	**0.237**	-0.031	-0.046	-0.038
Paper and allied products (ind_26)	0.066	**0.110**	0.160	**0.272**	0.098	0.317	-0.065	-0.037	-0.113	0.226	**0.110**	**0.289**
Leather and leather products (ind_31)	**3.881**	**3.840**	**3.583**	-0.078	-0.081	-0.126	-0.164	-0.076	-0.218	-0.521	-0.160	-0.894
Railroad transportation (ind_40)	**3.271**	**3.130**	**3.471**	-0.171	-0.100	-0.238	-0.463	-0.150	-0.572	-0.016	-0.023	-0.025
Motor freight transportation, and W. (ind_42)	0.018	0.090	0.134	**0.237**	**0.100**	**0.345**	-0.137	-0.060	-0.263	0.159	0.077	0.275
Water transportation (ind_44)	**1.809**	**1.717**	**1.647**	-0.442	-0.20	-0.784	0.050	0.003	0.051	-0.476	-0.180	-0.698
Transportation services (ind_47)	0.271	0.200	0.121	**0.881**	**0.400**	**0.939**	-0.005	-0.020	-0.007	0.423	0.200	0.606
Building materials, hardware, garden supply, and M.H.D (ind_52)	0.200	**0.070**	**0.155**	**0.116**	-0.031	0.144	-0.156	-0.076	-0.234	0.067	-0.008	0.116
Food stores (ind_54)	0.007	0.080	0.134	-0.084	-0.098	-0.134	**0.372**	**0.150**	**0.528**	0.025	-0.030	0.043
Automative dealers and gasoline S.S (ind_55)	**-0.011**	-0.050	**-0.025**	-0.096	-0.110	-0.144	-0.354	-0.130	-0.310	-0.049	-0.050	-0.079
Insurance agents, brokers and service (ind_64)	0.325	0.190	**0.828**	**0.687**	0.270	0.640	0.075	0.021	0.135	0.574	**0.280**	0.630
Real estate (ind_65)	**2.737**	**2.432**	**2.896**	-0.064	-0.048	-0.079	0.639	0.280	0.623	0.076	0.006	0.114
Hotels, rooming houses, camps and other lodging places (ind_70)	**2.117**	-0.010	**2.262**	-0.197	-0.180	-0.367	0.158	**0.015**	0.206	-0.158	-0.160	-0.251
Personal services (ind_72)	**1.025**	-0.120	-0.079	0.081	-0.081	0.098	0.021	-0.009	0.025	0.163	**0.025**	**0.187**
Automobile repairs, services, and parking (ind_75)	1.064	**1.003**	1.171	-0.294	-0.160	-0.378	-0.049	-0.038	-0.085	-0.204	-0.130	-0.239
Social services (ind_83)	-0.123	-0.090	-0.251	-0.438	-0.180	-0.596	**0.126**	**0.040**	**0.146**	-0.106	-0.130	-0.179
Miscellaneous services (ind_89)	-0.032	-1.300	-0.087	-0.591	-0.190	-0.963	-0.045	-0.044	-0.051	-0.231	-0.130	-0.410

Table 1.11: GPR, CR, and SOR values across industries for GARCH(1,1) (Part 1)

Industry	CR	SOR	GPR
Metal mining (ind_10)	-0.002	0.131	0.185
Bituminous coal and L.M (ind_12)	0.049	-0.418	-0.217
Oil and gas extraction (ind_13)	0.000	-0.020	-0.014
Mining and quarrying of nonmetallic minerals (ind_14)	-0.001	0.164	0.119
Building construction and other contracting (ind_15)	0.028	-0.343	-0.184
Heavy construction other than B.C.C (ind_16)	0.036	-0.303	-0.265
Construction special trade contractors (ind_17)	0.001	-0.358	-0.513
Food and kinred products (ind_20)	0.002	-0.124	-0.218
Tobacco products (ind_21)	0.012	-0.320	-0.280
Textile mill products (ind_22)	0.012	0.102	0.098
Apparel and other finished products (ind_23)	0.026	-0.639	-0.456
Lumber and wood products, except furniture (ind_24)	0.005	-0.272	-0.265
Furniture and fixtures (ind_25)	0.027	-0.387	-0.311
Paper and allied products (ind_26)	0.006	-0.126	-0.068
Priting, publishing, and allied industries (ind_27)	0.018	-0.100	-0.079
Chemicals and allied products (ind_28)	0.005	-0.128	-0.499
Petroleum refining and related industries (ind_29)	0.003	-0.227	-0.193
Rubber and miscellaneous plastic products (ind_30)	0.002	-0.102	-0.142
Leather and leather products (ind_31)	0.003	-0.358	-0.571
Stone, clay, glass, and concrete products (ind_32)	0.014	-0.635	-0.411
Primary metal industries (ind_33)	-0.033	0.573	0.380
Fabricatel metal products, except machinery and T.E (ind_34)	0.000	**0.120**	**0.074**
Industrial and commercial machinery and C.E (ind_35)	0.001	-0.292	-0.201
Electronic and other electrical equipment and C. (ind_36)	-0.002	0.213	0.188
Transportation equipment (ind_37)	0.004	0.081	0.068
Measuring, analyzing, and controlling instruments (ind_38)	-0.003	0.186	0.141
Miscellaneous manufacturing industries (ind_39)	0.019	-0.503	-0.416
Railroad transportation (ind_40)	0.000	0.258	0.231
Motor freight transportation and W. (ind_42)	0.003	-0.019	-0.019
Water transportation (ind_44)	0.000	-0.334	-0.336

Table 1.11: GPR, CR, and SOR values across industries for GARCH(1,1) (continued)

Industry	CR	SOR	GPR
Transportation by air (ind_45)	-0.004	**0.239**	**0.210**
Transportation services (ind_47)	0.007	-0.438	-0.363
Communications (ind_48)	0.013	-0.232	-0.123
Electric, gas, and sanitary services (ind_49)	0.006	0.053	0.033
Wholesale trade-durable goods (ind_50)	0.013	-0.166	-0.103
Wholesale trade-nondurable good (ind_51)	0.002	-0.126	-0.136
Building materials, hardware, garden supply, and M.H.D (ind_52)	0.018	-0.963	-0.528
General merchandise stores (ind_53)	-0.001	0.101	0.121
Food stores (ind_54)	0.037	-0.741	-0.395
Automotive dealers and gasoline S.S (ind_55)	**0.001**	-0.149	-0.119
Apparel and accessory stores (ind_56)	0.001	-0.149	-0.119
Home furniture, furnishings, and E.S (ind_57)	0.017	-0.266	-0.164
Eating and drinking places (ind_58)	-0.010	0.367	0.352
Miscellaneous retail (ind_59)	-0.038	0.700	0.702
Depository institutions (ind_60)	0.010	-0.097	-0.109
Non-depository credit institutions (ind_61)	0.010	-0.182	-0.144
Security and commodity brokers, D., E., and S. (ind_62)	**0.003**	-0.131	-0.211
Insurance carriers (ind_63)	0.011	0.039	0.033
Insurance agents, brokers and service (ind_64)	-0.007	0.199	0.186
Real estate (ind_65)	0.001	-0.143	-0.137
Holding and other investment offices (ind_67)	0.029	-0.305	-0.193
Hotels, rooming houses, camps, and other lodging places (ind_70)	0.000	-0.367	-0.268
Personal services (ind_72)	0.011	0.059	0.062
Business services (ind_73)	**0.000**	**0.021**	**0.017**
Automobile repairs, services, and parking (ind_75)	-0.029	**2.150**	**2.604**
Motion pictures (ind_78)	0.004	-0.379	-0.347
Amusement and recreation services (ind_79)	0.015	-0.350	-0.230
Health services (ind_80)	0.000	-0.097	-0.117
Educational services (ind_82)	0.009	-0.192	-0.136
Social services (ind_83)	0.020	-0.113	-0.073
Engineering, accounting, research, M., and R.S (ind_87)	-0.020	0.404	0.329
Miscellaneous services (ind_89)	**0.007**	**0.008**	**0.006**
Nonclassifiable establishment (ind_99)	0.071	-0.332	-0.271

1.5.7 Rolling R^2_{OoS} (in %) performance across models and industries

This section presents R^2_{OoS} over time across model and industries. A model is considered most successful if it consistently achieves the highest rolling R^2_{OoS} compared to other models, particularly during periods of low and high volatility, indicating superior and robust predictive performance.

Figure 1.1: Apparel & other finished products (ind_23)

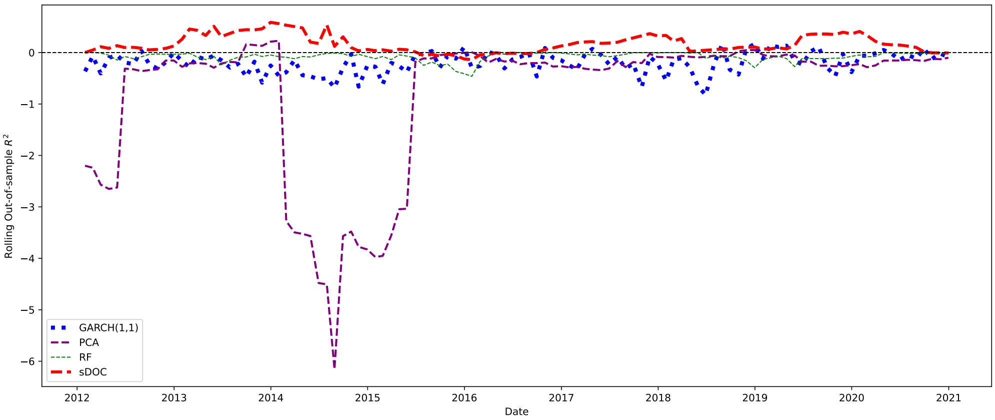


Figure 1.2: Wholesale non-durable goods (ind_51)

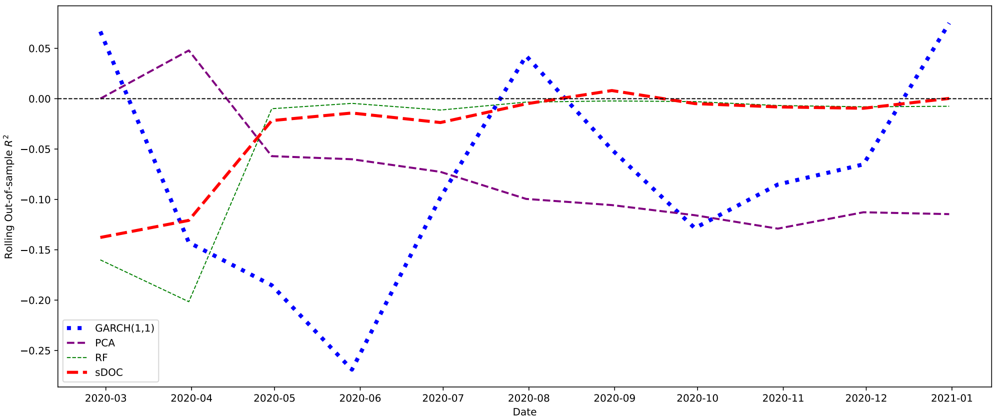


Figure 1.3: Chemicals and allied products (ind_28)

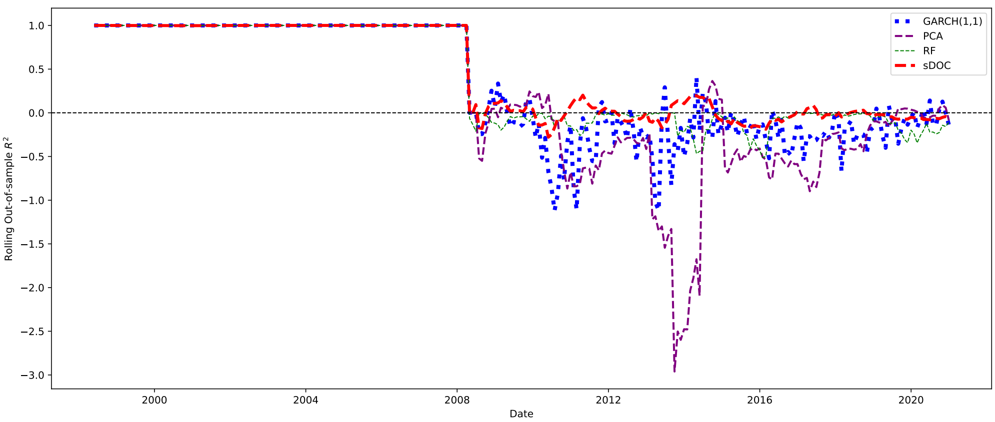


Figure 1.4: Rubber & miscellaneous plastic products (ind_30)

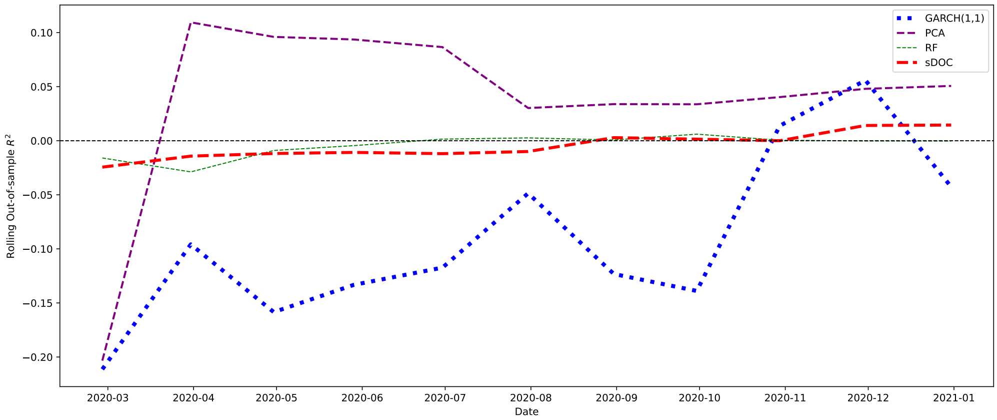


Figure 1.5: Oil & gas extraction (ind_13)

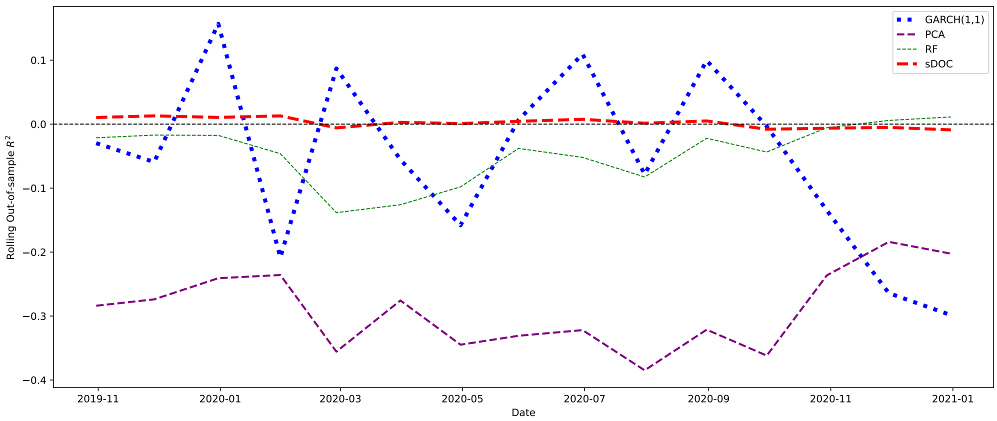


Figure 1.6: Petroleum refining & related industries (ind_29)

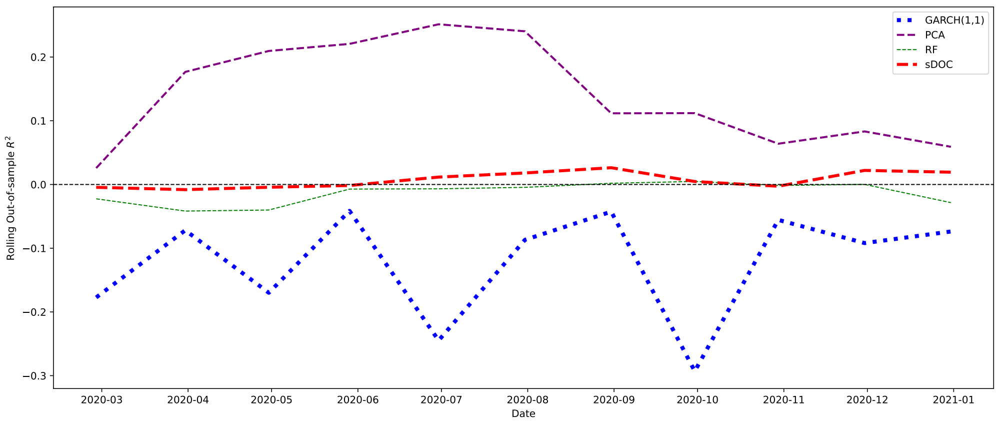


Figure 1.7: Food & kinred products (ind_20)

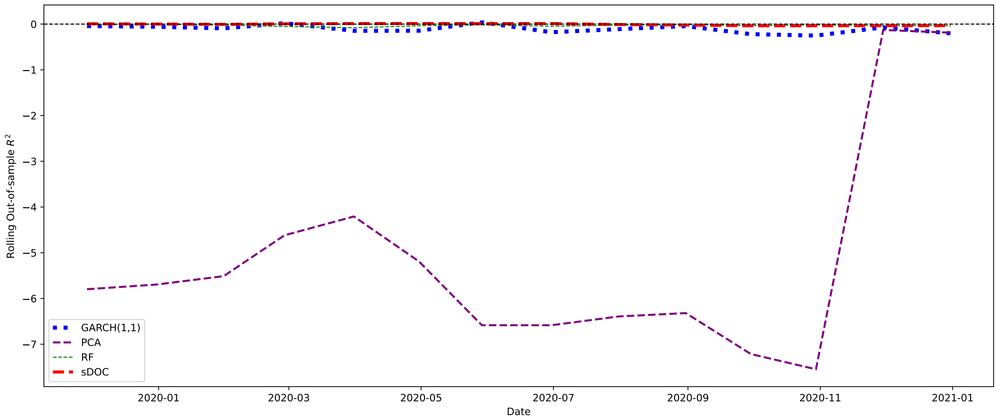


Figure 1.8: Heavy construction & other B.C.C (ind_16)

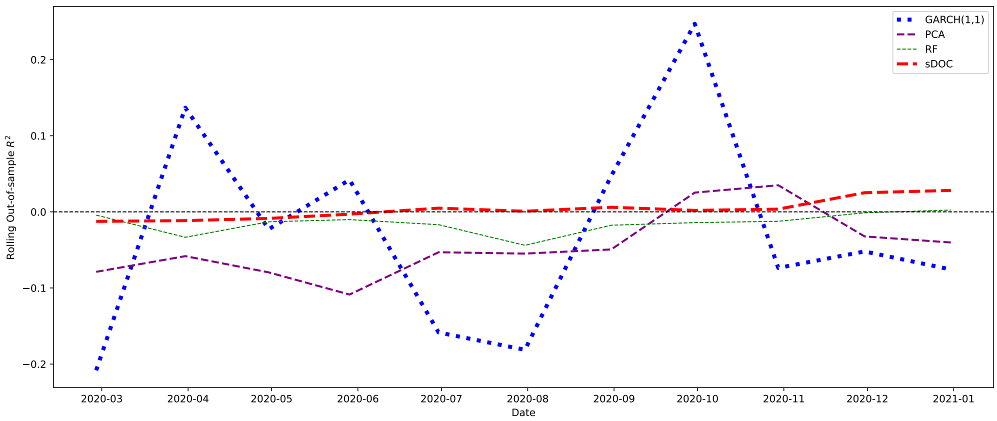


Figure 1.9: Fabricatel metal products, except machinery and T.P (ind_34)

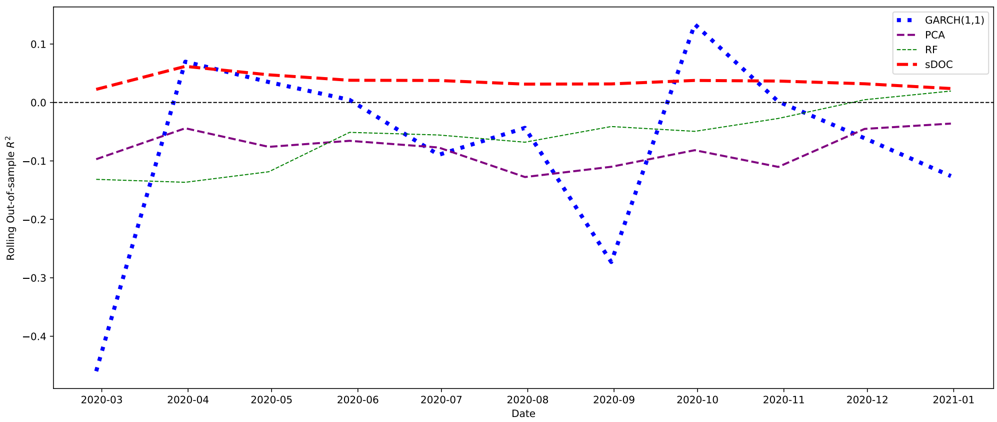


Figure 1.10: Amusement and recreation services (ind_79)

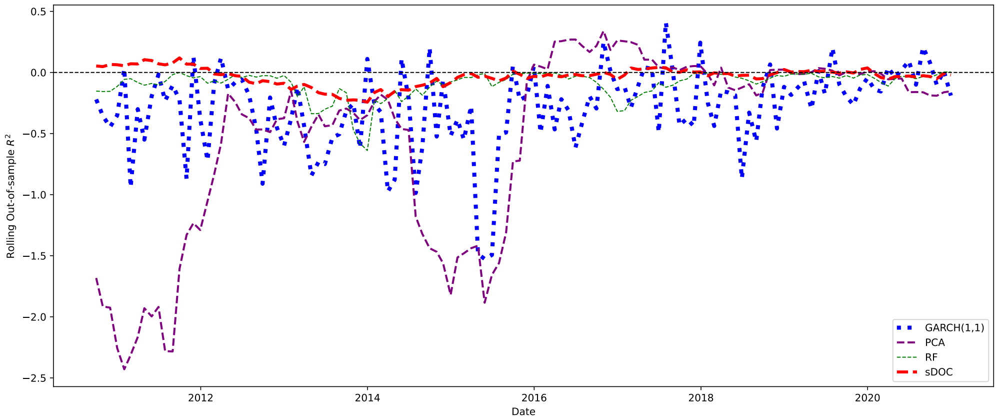


Figure 1.11: Transportation by air (ind_45)

Figure 1.12: Measuring, analyzing, & controlling instruments & other B.C.C (ind_38)

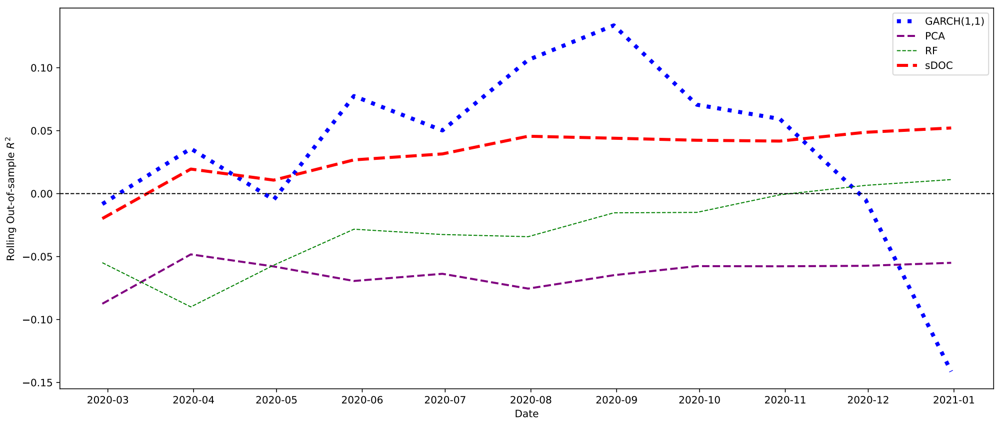


Figure 1.13: Stone, clay, glass, & concrete products (ind_32)

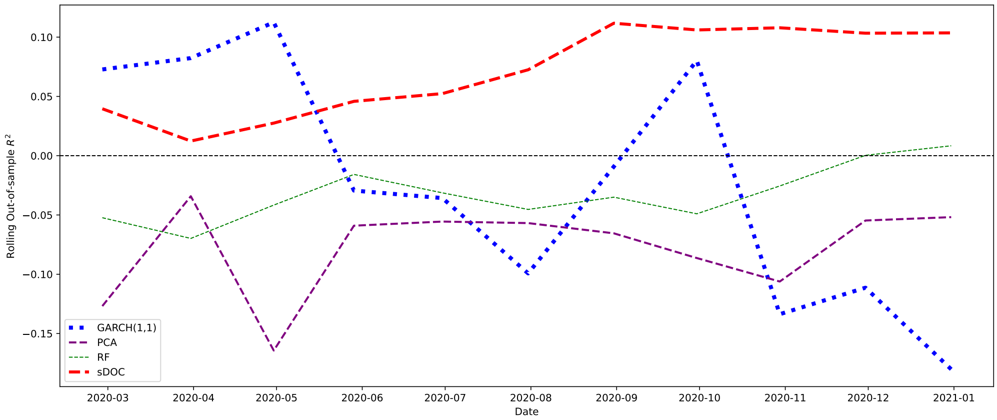


Figure 1.14: Primary metal industries (ind_33)

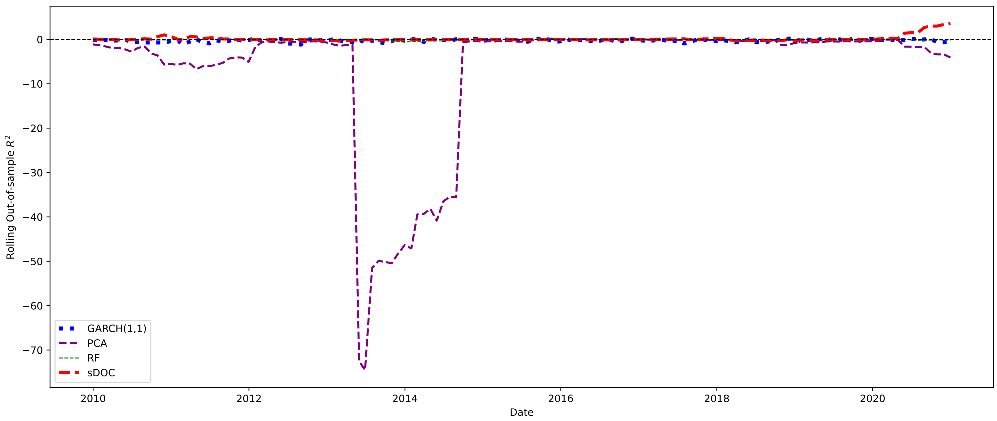


Figure 1.15: Industrial & commercial machinery & C.E (ind_35)

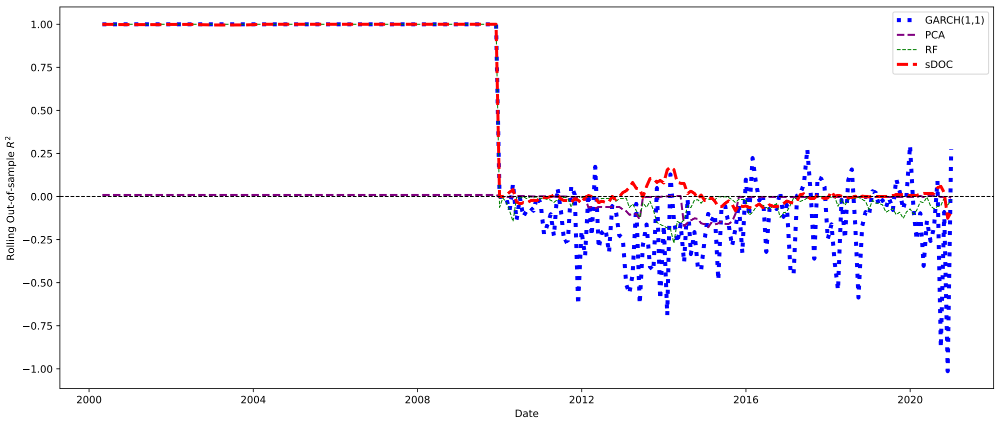


Figure 1.16: Transportation equipment (ind_37)

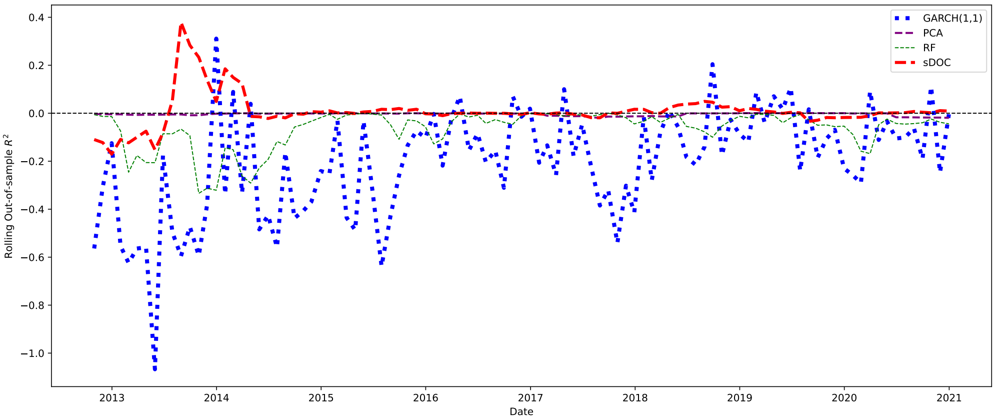


Figure 1.17: Electronic & other electrical equipment & C. (ind_36)

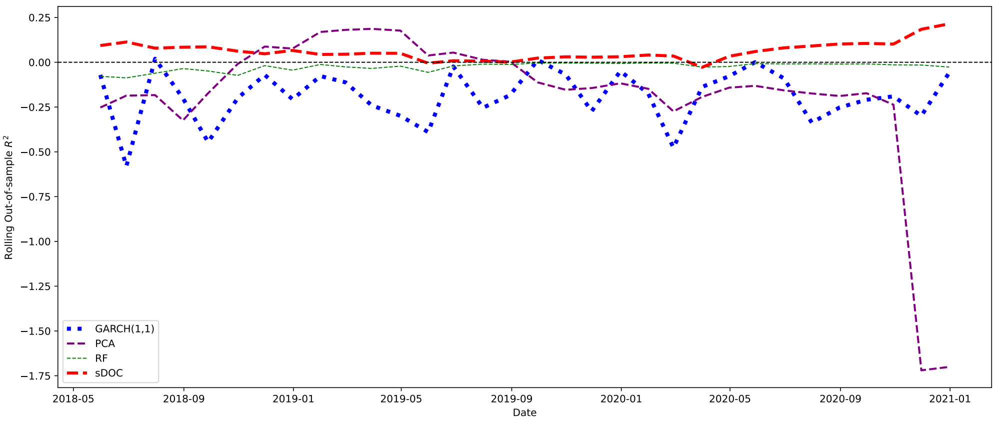


Figure 1.18: Miscellaneous manufacturing industries (ind_39)

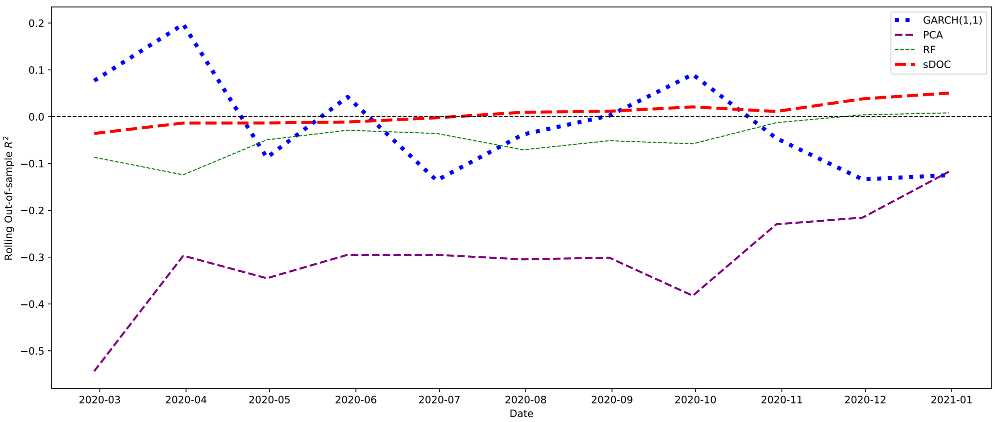


Figure 1.19: General merchandise stores (ind_53)

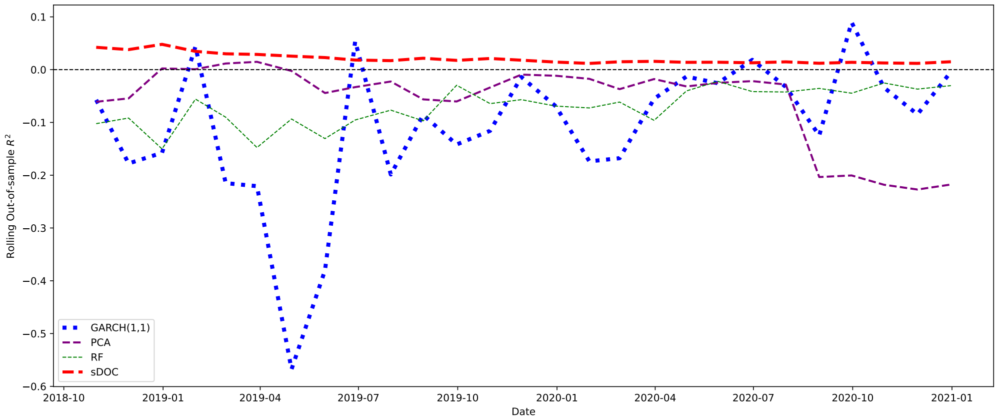


Figure 1.20: Apparel and accessory stores (ind_56)

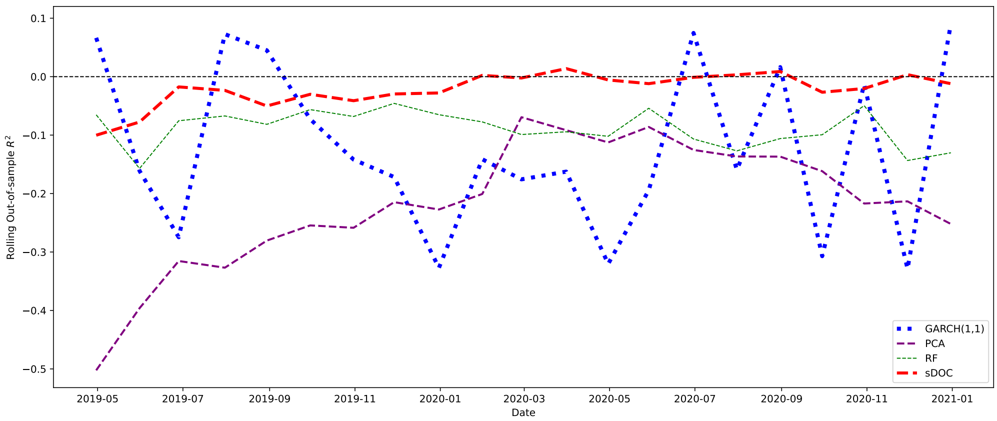


Figure 1.21: Wholesale trade-durable goods (ind_50)

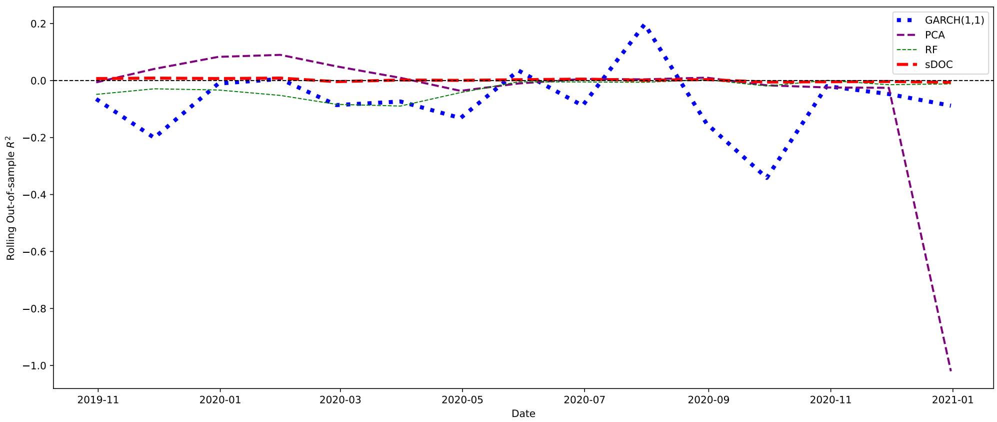


Figure 1.22: Holding & other investment offices (ind_67)

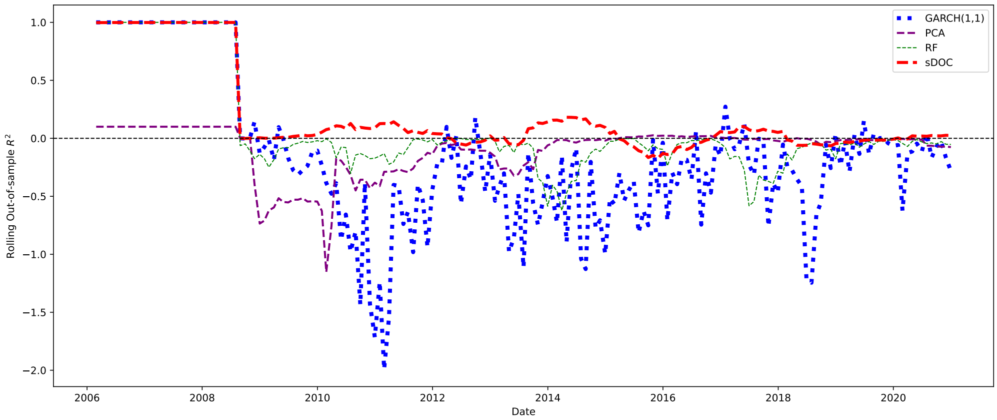


Figure 1.23: Home furniture, finishings, & E.S (ind_57)

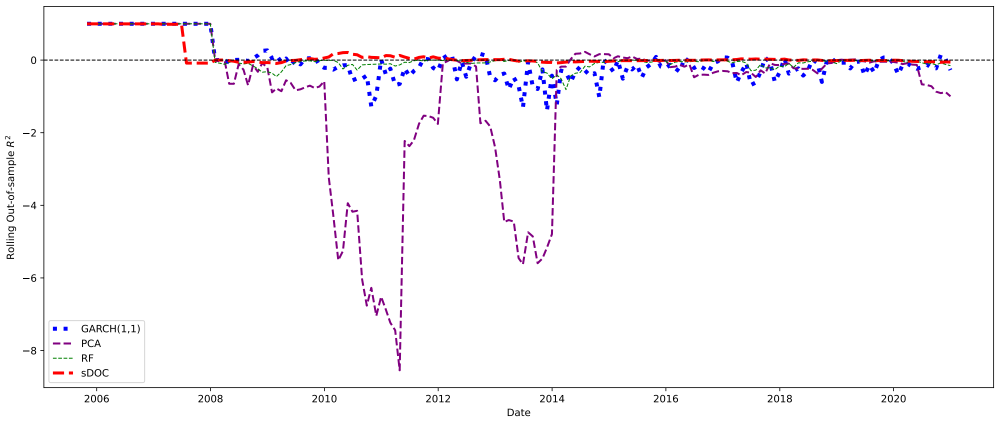


Figure 1.24: Security & commodity brokers, D., E., & S. (ind_62)

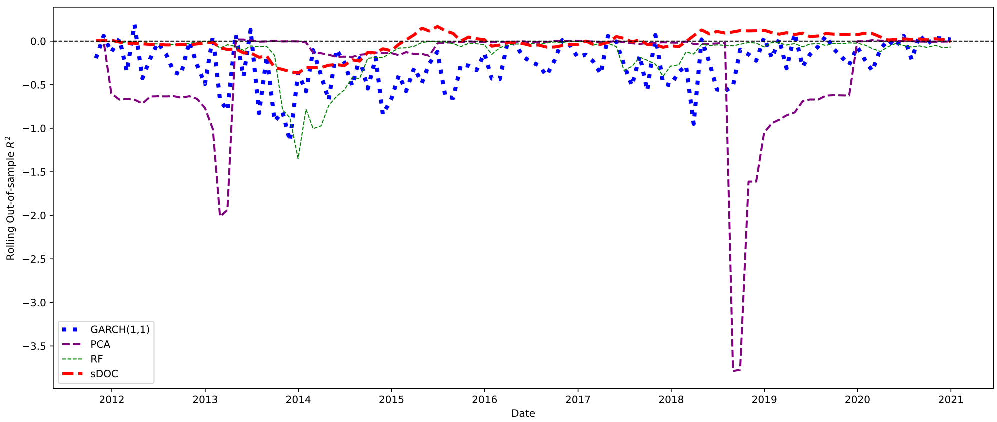


Figure 1.25: Communications (ind_48)

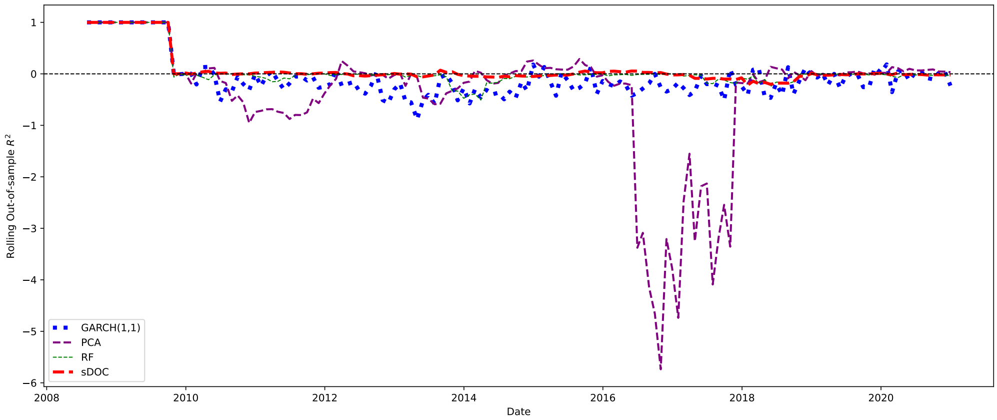


Figure 1.26: Electric, gas, & S.S. (ind_49)

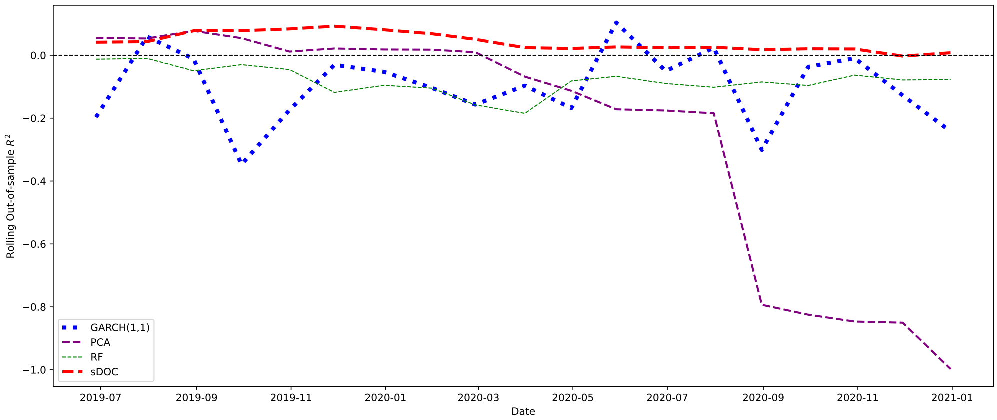


Figure 1.27: Health services (ind_80)

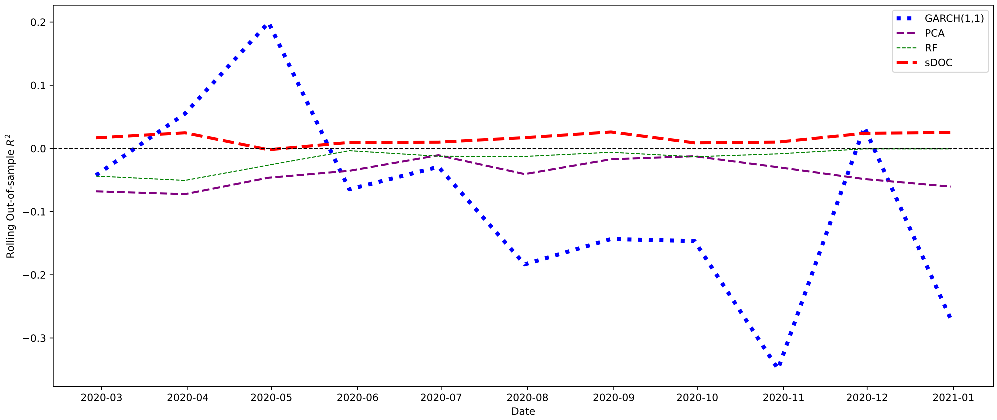


Figure 1.28: Depository institutions (ind_60)

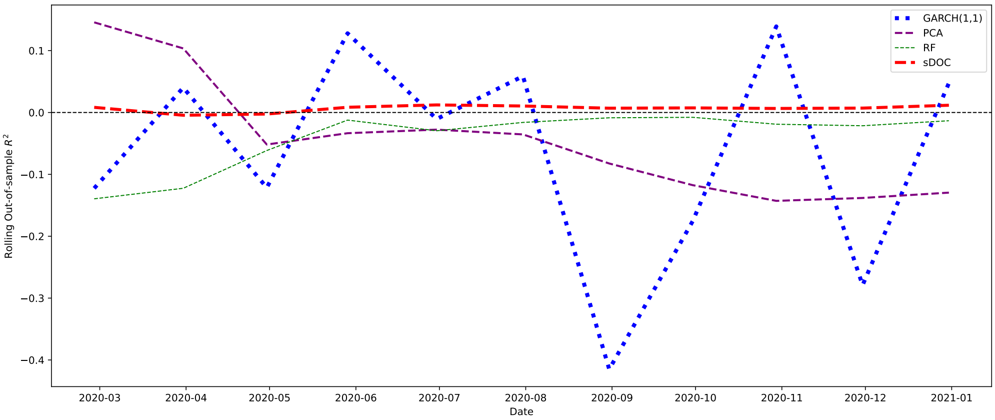


Figure 1.29: Non-depository credit institutions (ind_61)

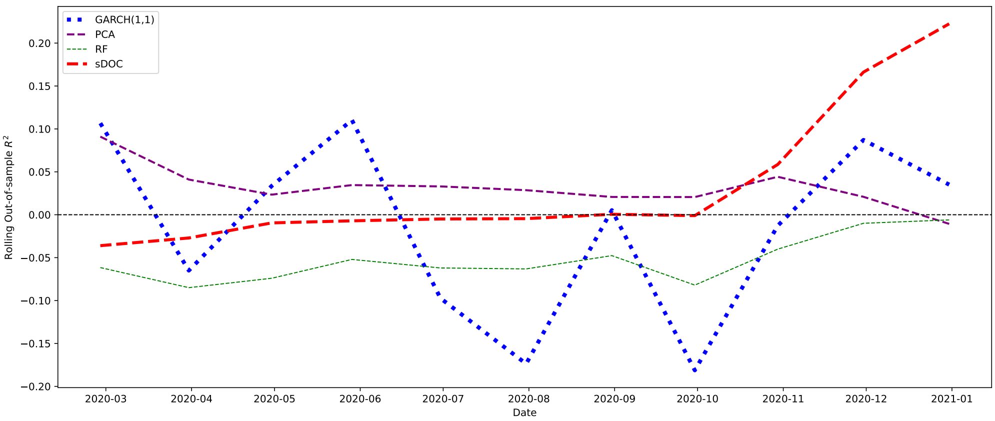


Figure 1.30: Eating & drinking places (ind_58)

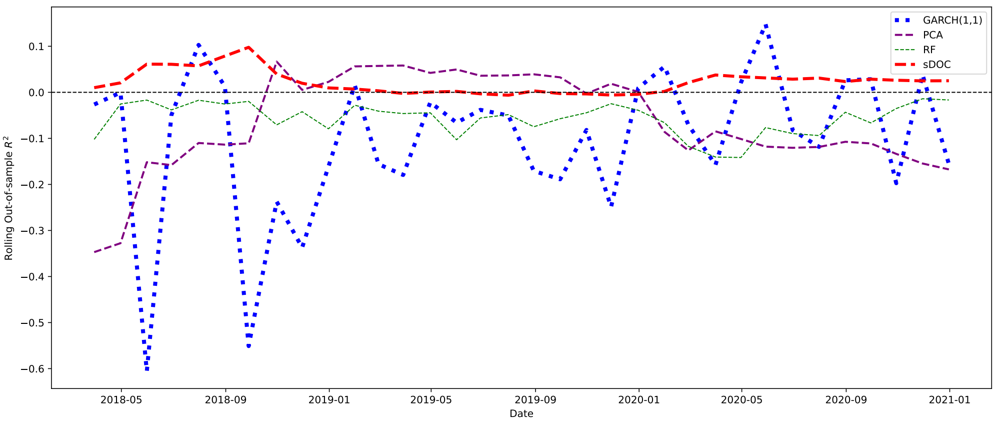


Figure 1.31: Insurance carriers (ind_63)

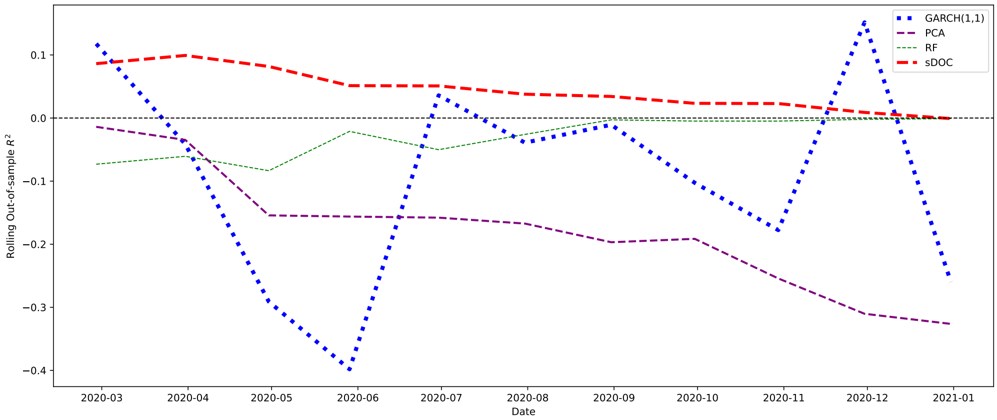


Figure 1.32: Motion pictures (ind_78)

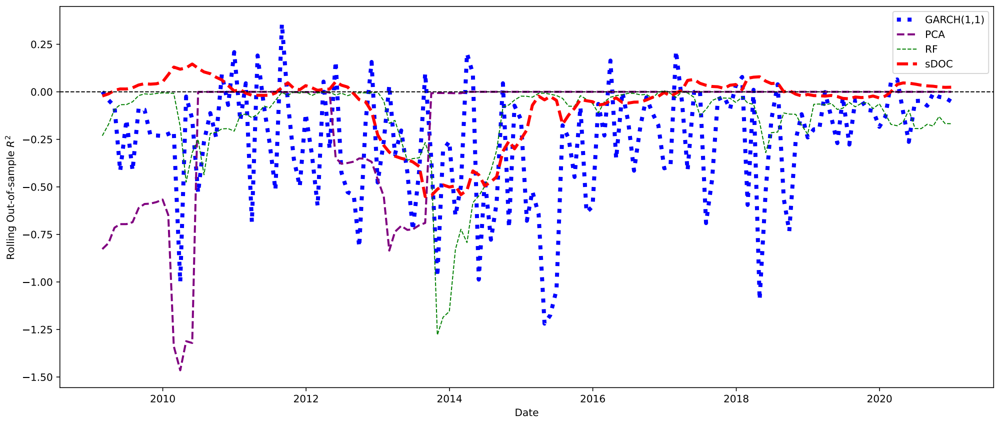


Figure 1.33: Miscellaneous retail (ind_59)

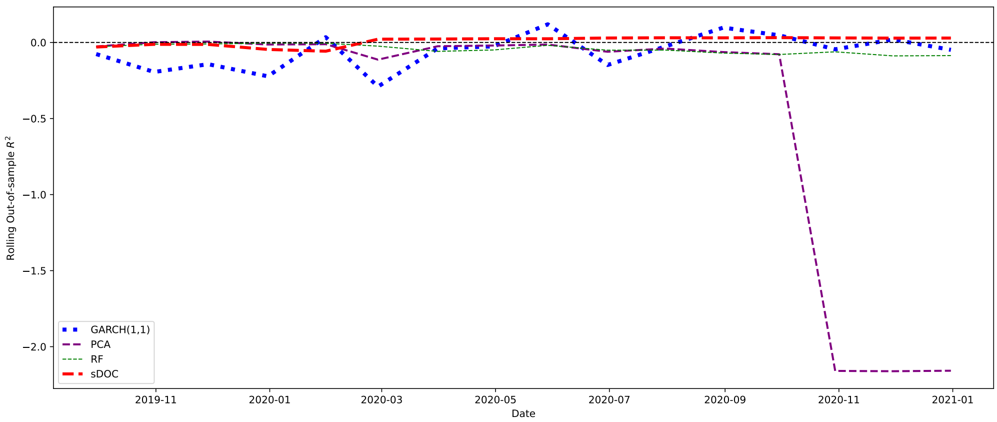


Figure 1.34: Business services (ind_73)

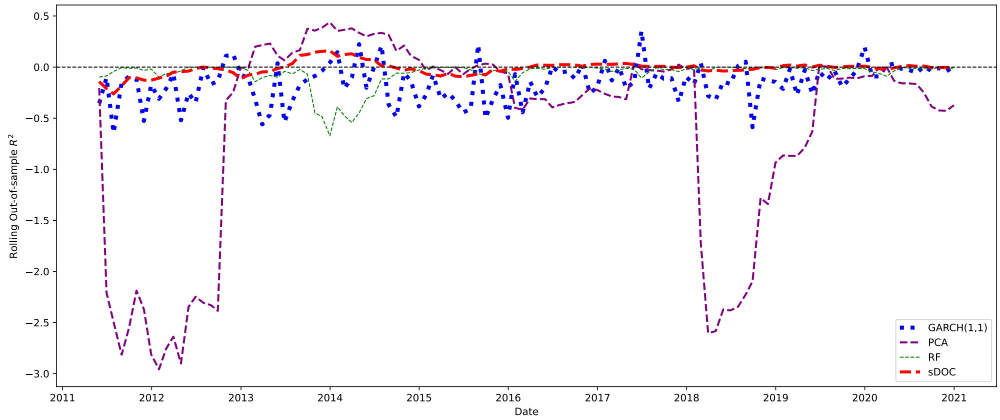


Figure 1.35: Educational services (ind_82)

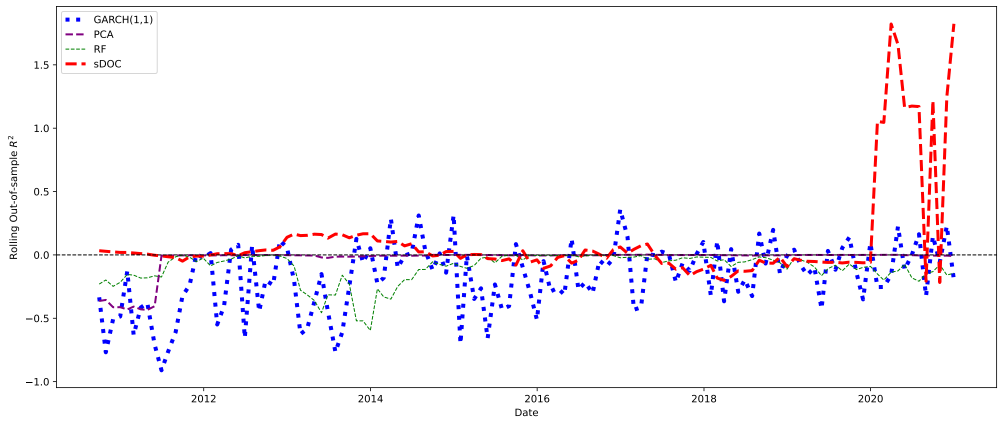


Figure 1.36: Engineering, accounting, research, M., & R.S (ind_87)

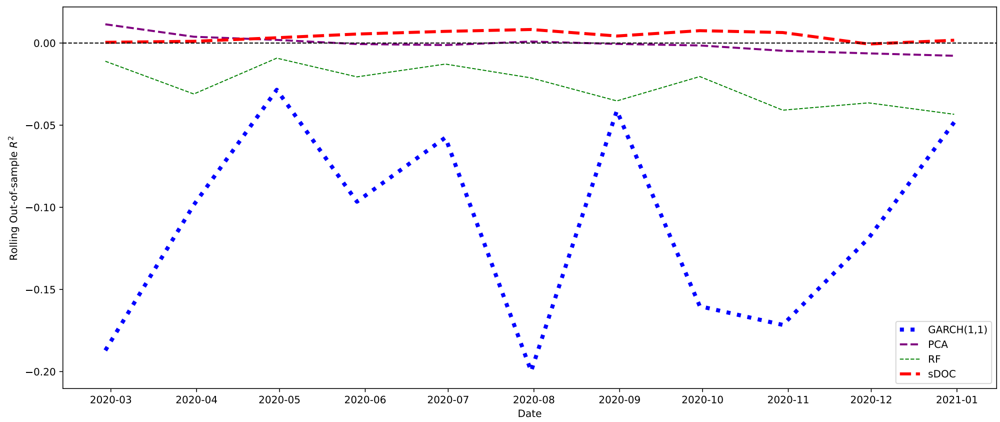


Figure 1.37: Nonclassifiable establishment (ind_99)

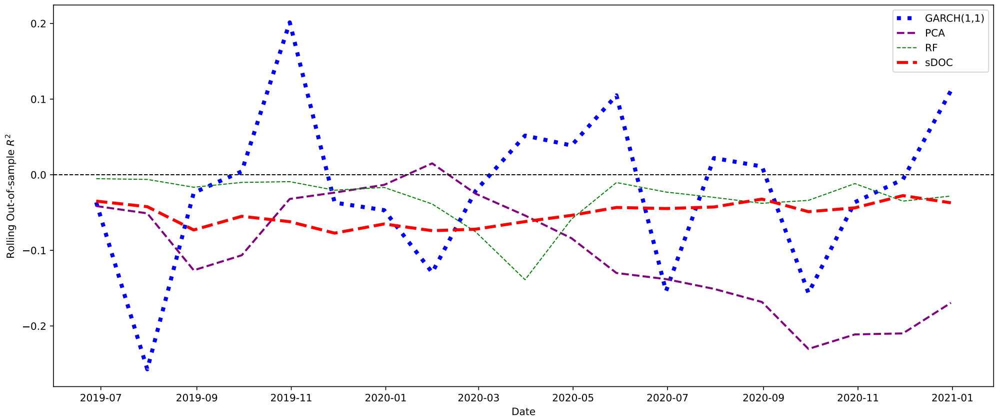


Figure 1.38: Metal mining (ind_10)

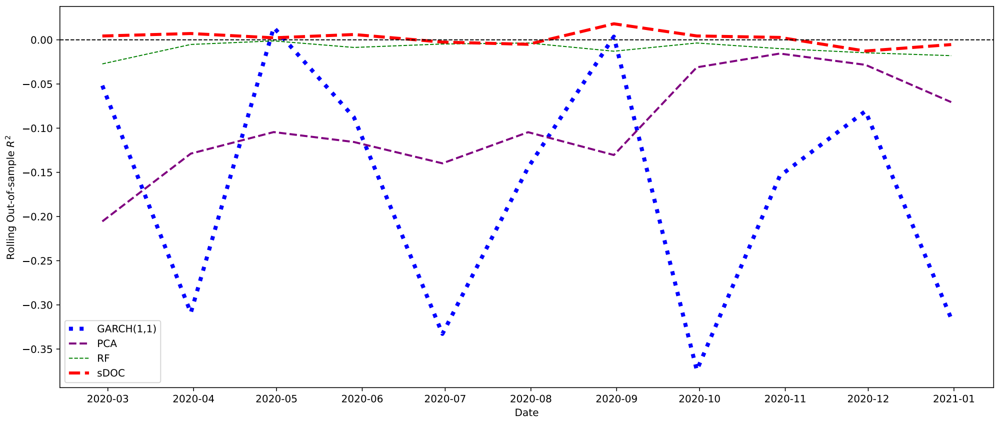


Figure 1.39: Bituminous coal & L.M (ind_12)

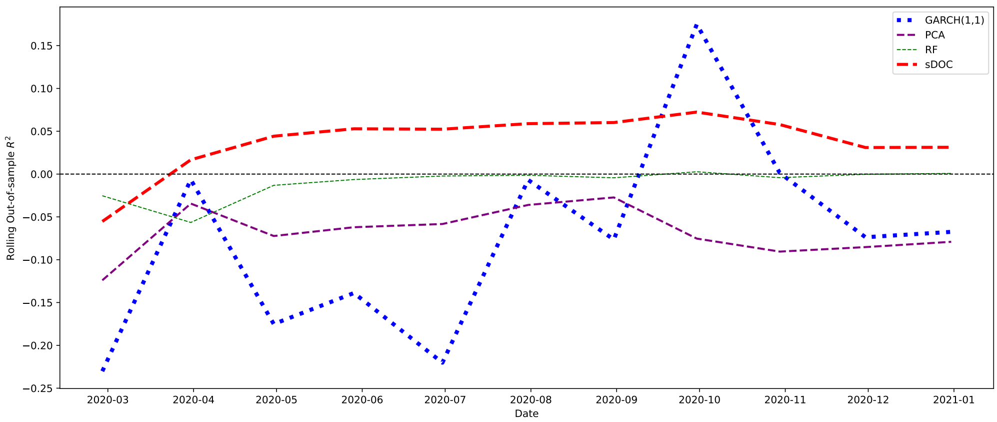


Figure 1.40: Mining & quarrying of nonmetallic minerals, except fuels (ind_14)

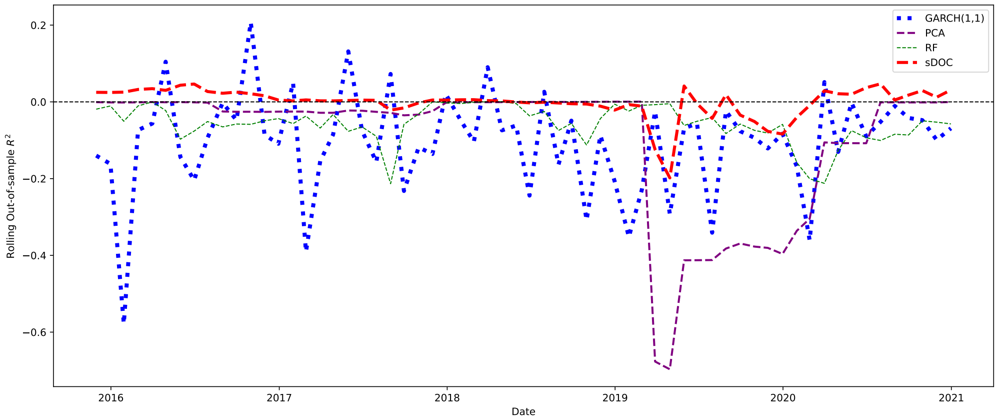


Figure 1.41: Building construction general contractors & O.B (ind_15)

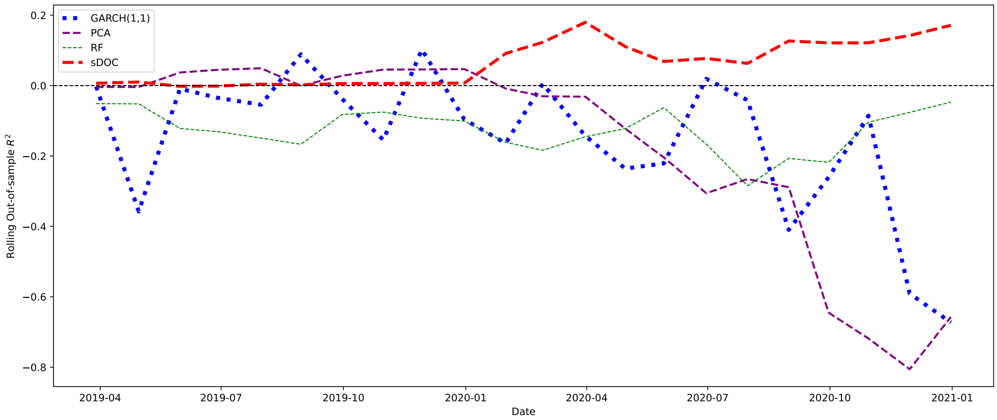


Figure 1.42: Construction special trade contractors (ind_17)

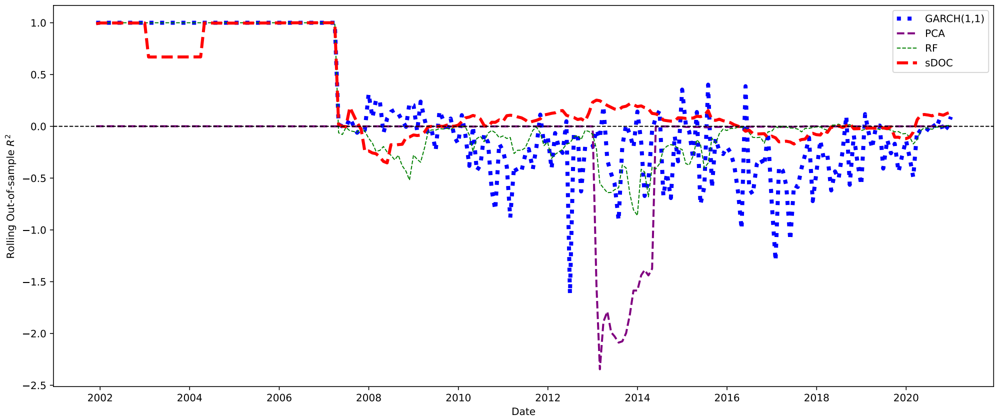


Figure 1.43: Tobacco products (ind_21)

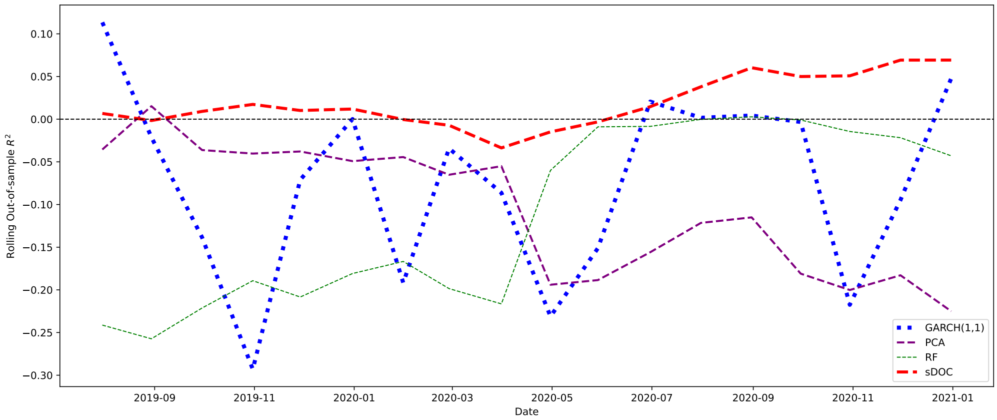


Figure 1.44: Textile mill products (ind_22)

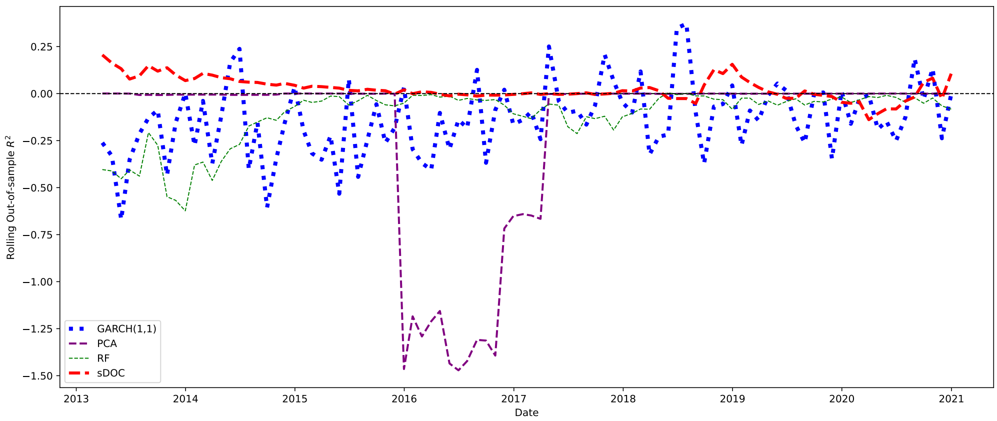


Figure 1.45: Lumber and wood products, except furniture (ind_24)

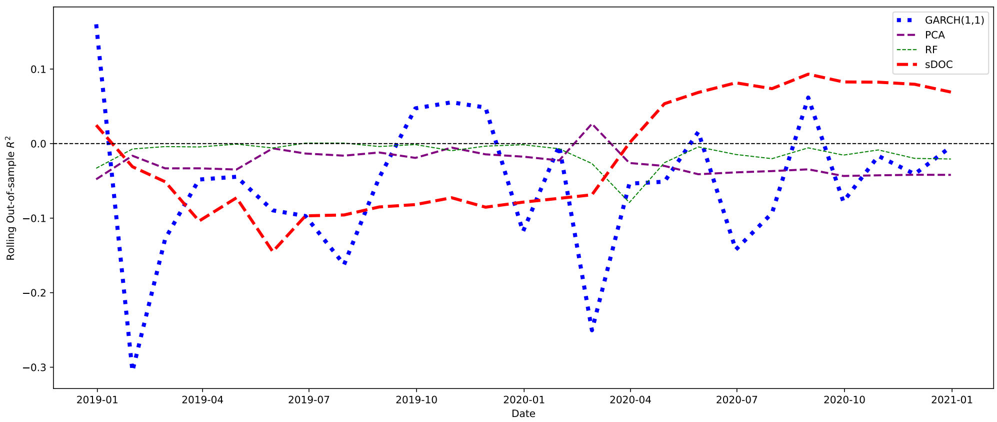


Figure 1.46: Furniture and fixtures (ind_25)

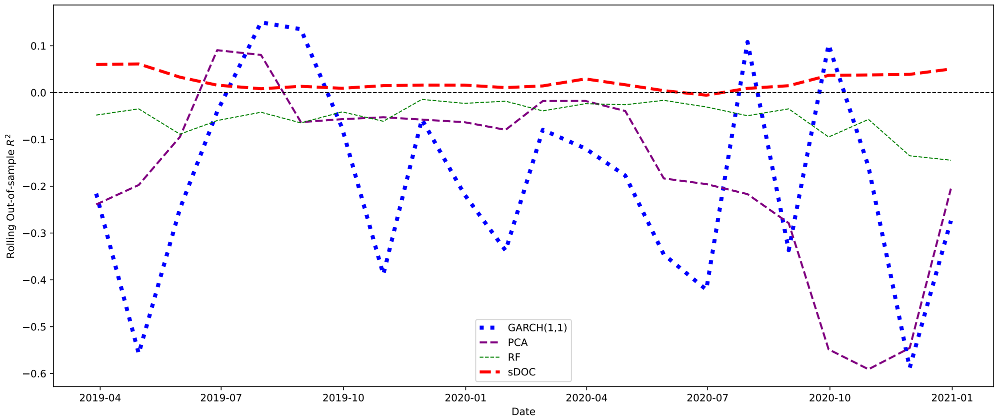


Figure 1.47: Paper and allied products (ind_26)

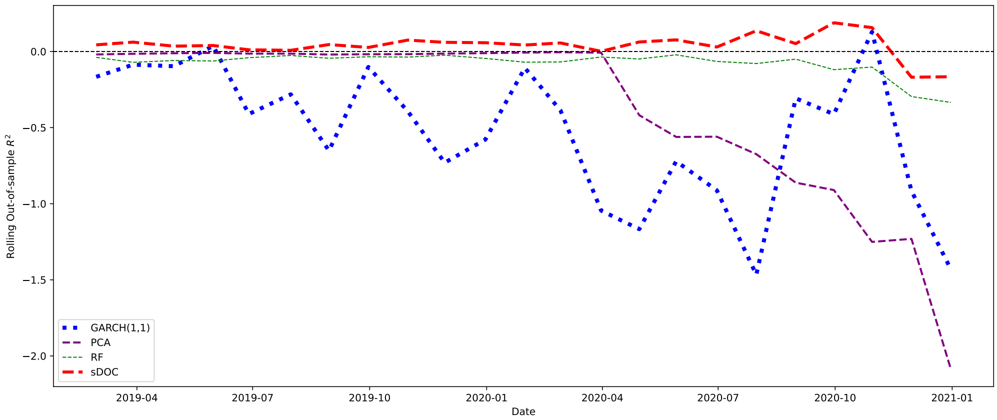


Figure 1.48: Leather and leather products (ind_31)

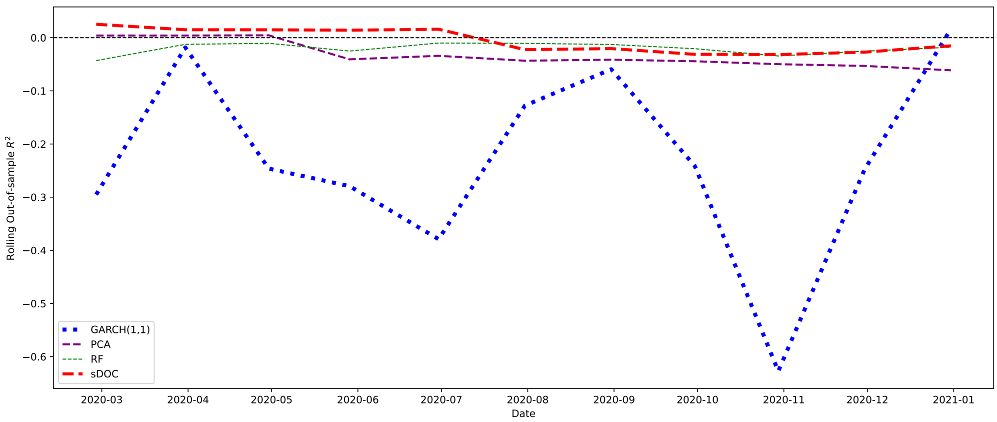


Figure 1.49: Railroad transportation (ind_40)

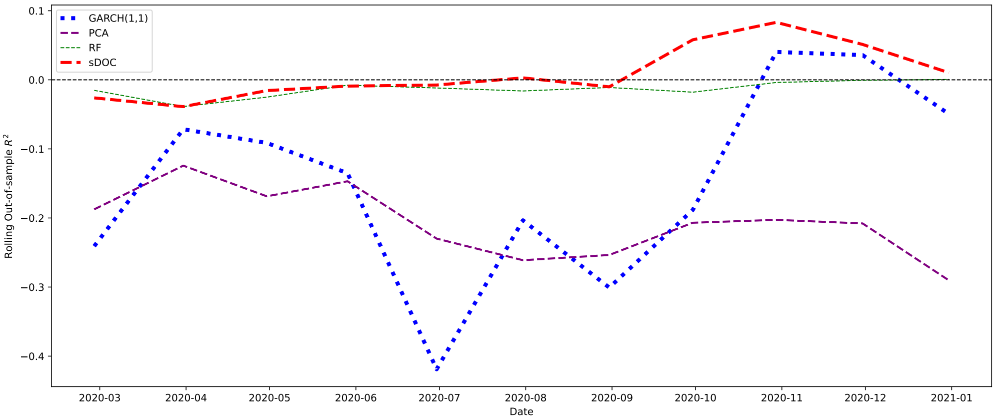


Figure 1.50: Motor freight transportation and W. (ind_51)

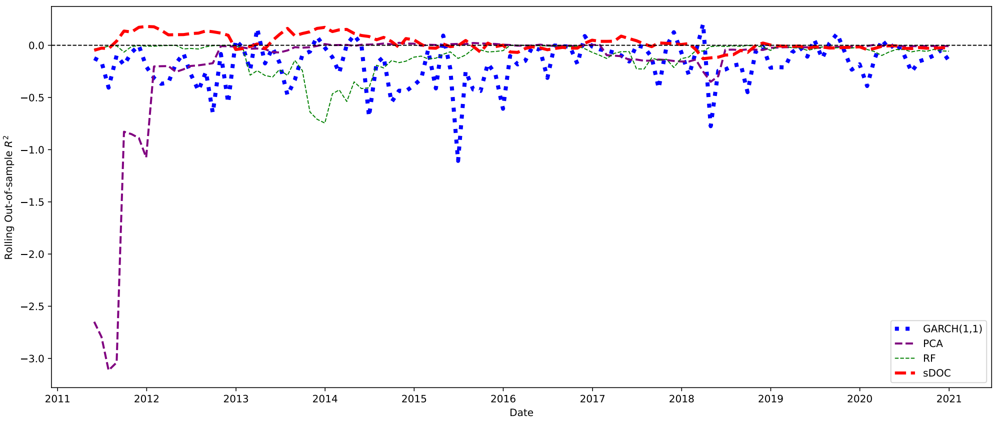


Figure 1.51: Water transportation (ind_52)

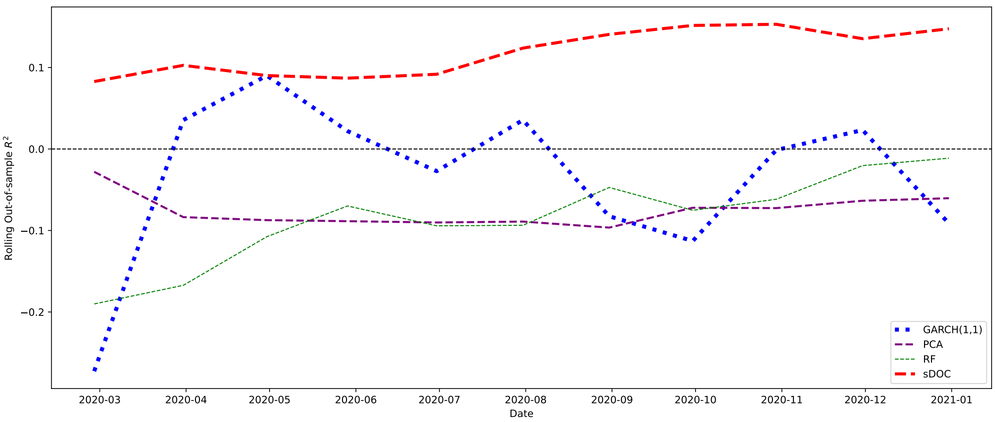


Figure 1.52: Transportation services (ind_47)

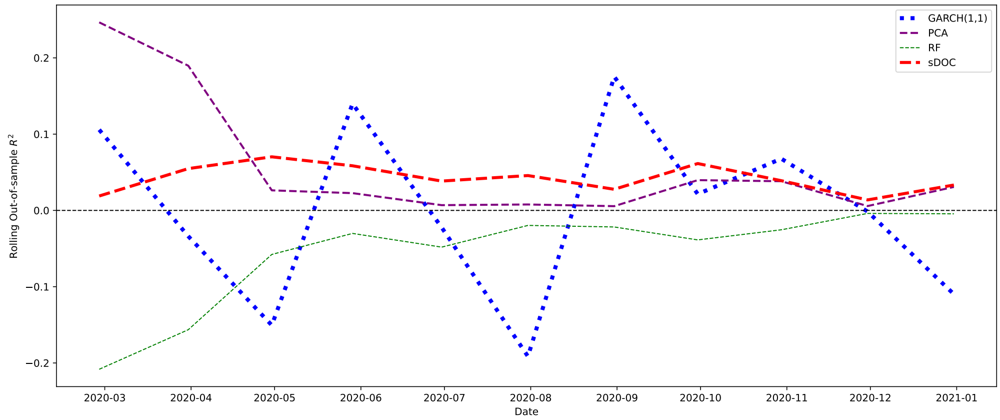


Figure 1.53: Building materials, hardware, garden supply, & M.H.D (ind_52)

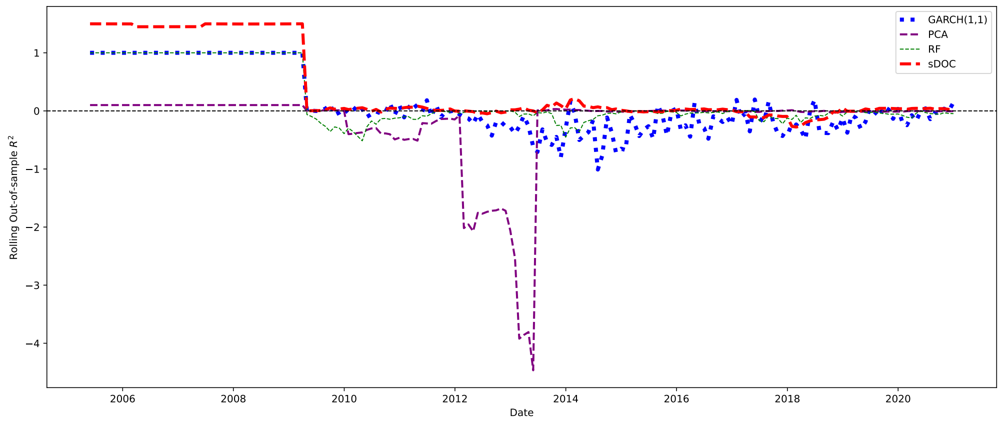


Figure 1.54: Food stores (ind_54)

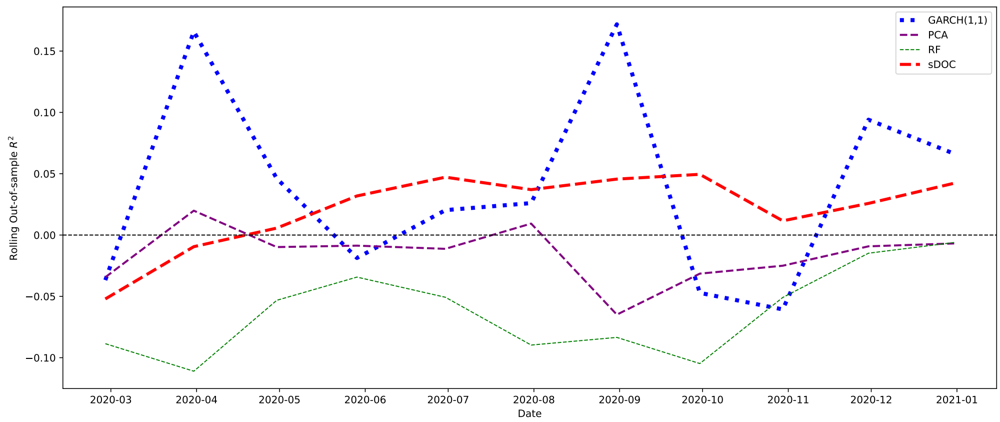


Figure 1.55: Automotive dealers & gasoline S.S (ind_55)

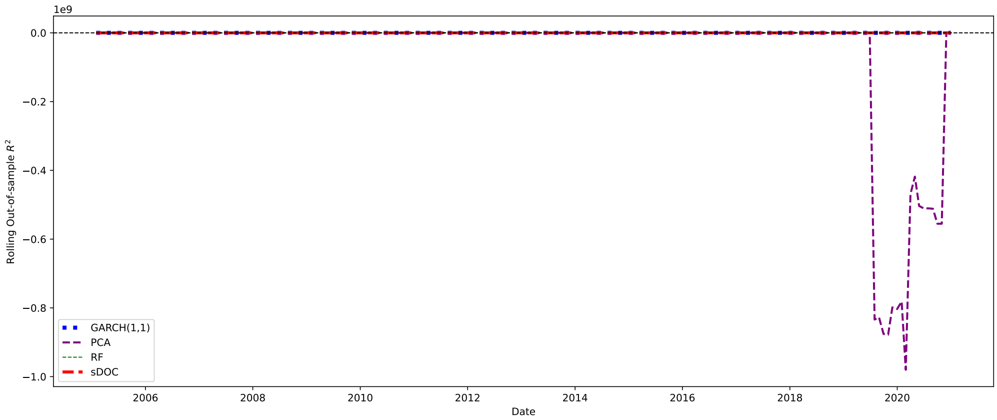


Figure 1.56: Insurance agents, brokers & service (ind_64)

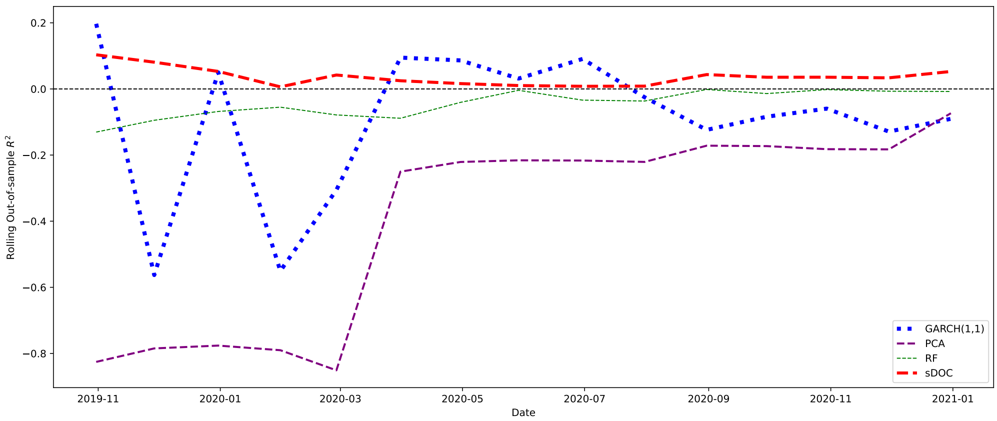


Figure 1.57: Real estate (ind_65)

Figure 1.58: Hotels, rooming houses, camps, & other lodging places (ind_70)

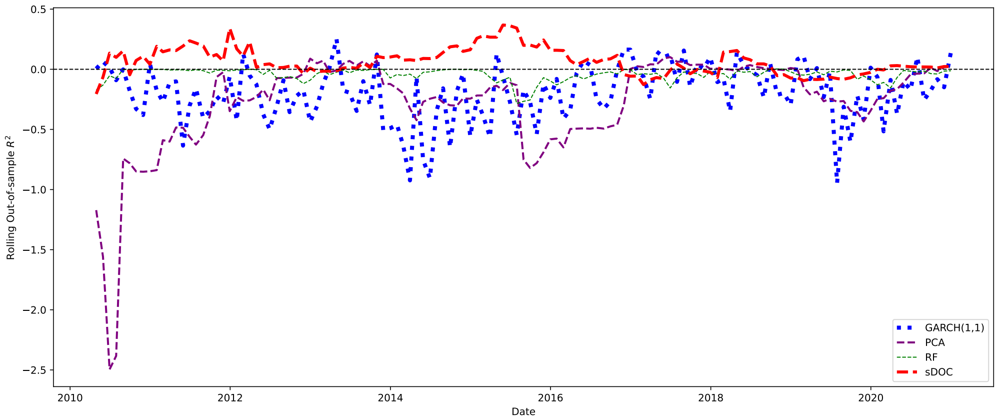


Figure 1.59: Personal services (ind_72)

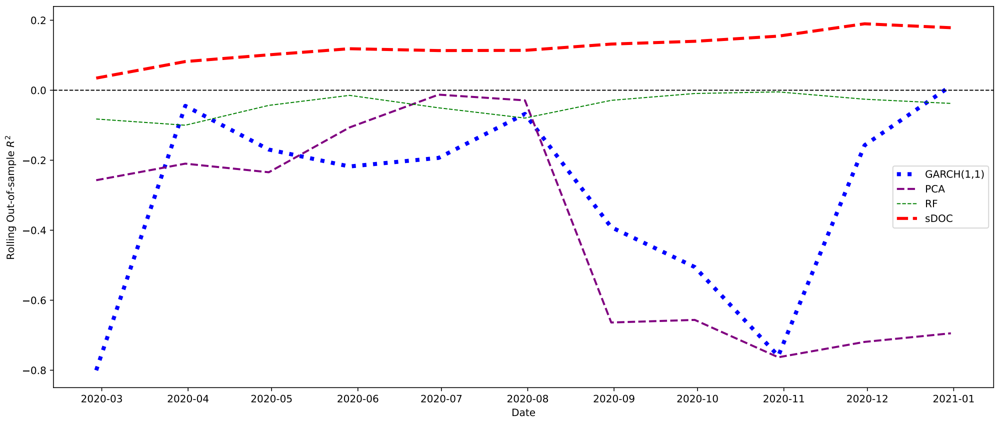


Figure 1.60: Automobile repairs, services, & parking (ind_75)

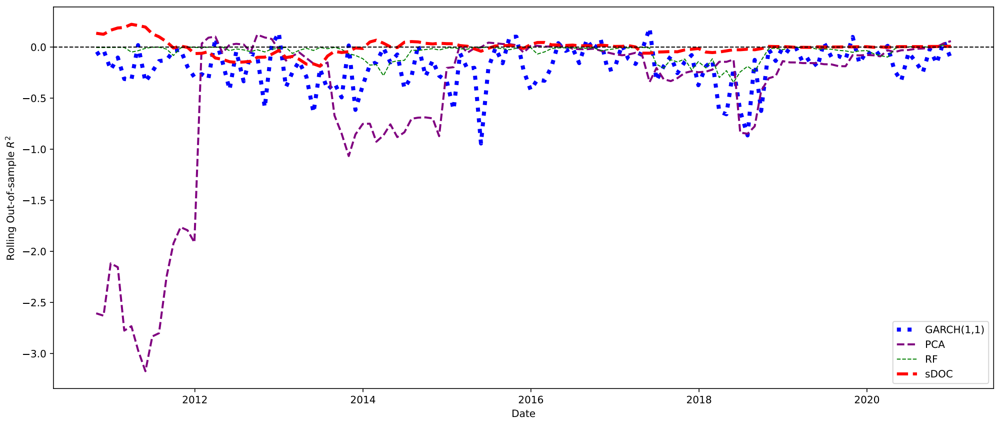


Figure 1.61: Social services (ind_83)

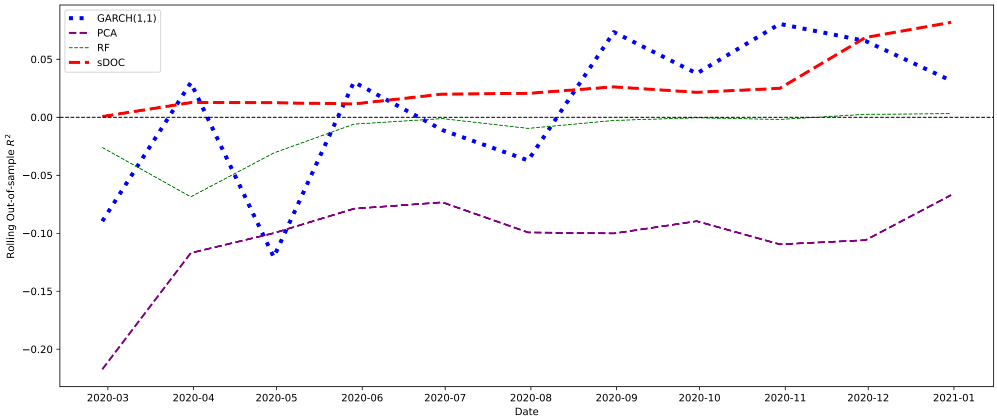


Figure 1.62: Miscellaneous services (ind_89)

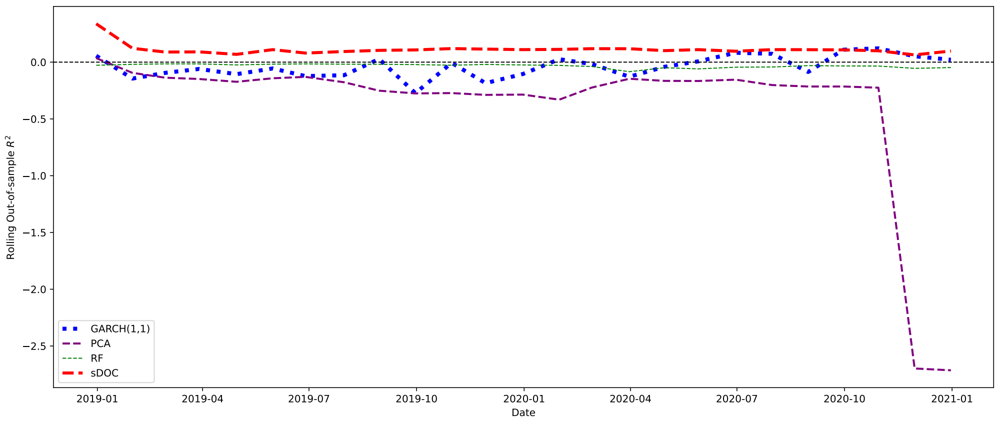
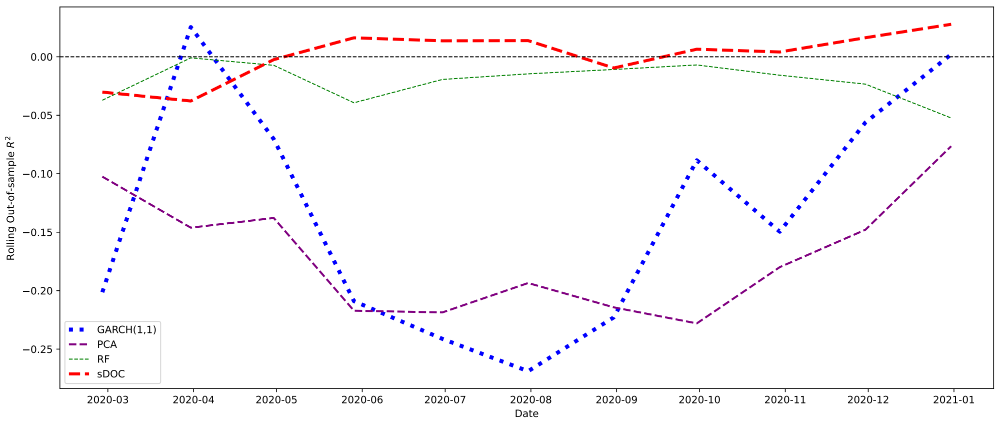


Figure 1.63: Printings, Publishing & allied industries (ind_27)



1.5.8 Description of firm characteristics and industry classification (SIC)

Table 1.12: Description of firm characteristics (Part 1)

Variables	Description
turn	Share turnover. It represents the average monthly trading volume for most recent 3 months scaled by number of shares outstanding in current month.
baspread	Monthly average of daily bid-ask spread divided by average of daily spread.
beta	Estimated market beta from weekly returns and equal weighted market returns for 3 years ending month t-1 with at least 52 weeks of returns.
idiovol	The monthly idiosyncratic volatility of stock. It represents the standard deviation of residuals from the regression using monthly returns over the prior 21 trading days
chmom	Cumulative returns from months t-6 to t-1 minus months t-12 to t-7.
mom1m	1-month cumulative return.
indmom	Equal weighted average industry 12-month returns.
aeavol	Average daily trading volume (vol) for 3 days around earnings announcement minus average daily volume for 1-month ending 2 weeks before earnings announcement divided by 1-month average daily volume.
bm	Book-to-market. It represents the book value of equity divided by end of fiscal-year-end market capitalization.
cash	Cash and cash equivalents divided by average total assets.
rd_mve	R&D expense divided by end-of-fiscal-year market capitalization.
cashdebt	Earnings before depreciation and extraordinary items divided by average total liabilities.
cashpr	Fiscal year end market capitalization plus long term debt minus total assets divided by cash and equivalents.
chatoia	2-digit SIC-fiscal-year mean adjusted change in sales divided by average total assets.
chinv	Change in inventory scaled by average total assets.
depr	Depreciation divided by PP&E.
ep	Annual income before extraordinary items divided by end of fiscal year market cap.
gma	Revenues minus cost of goods sold divided by lagged total assets.
lev	Total liabilities divided by fiscal year end market capitalization.
lgr	Annual percent change in total liabilities.
quick	(current assets - inventory)/current liabilities.
pchquick	Percent change in quick.
sp	Annual revenue divided divided by fiscal-year-end market capitalization.
stdacc	Standard deviation for 16 quarters of accruals scales by sales.
stdcf	Standard deviation for 16 quarters of cash flows divided by sales.

Table 1.12: Description of firm characteristics (continued-1)

chtx	Percent change in total taxes from t-4 to t.
pctacc	Same as acc except that the numerators is divided by the absolute value of ib.
invest	Annual change in gross property, plant, and equipment+annual change in inventories all scaled by lagged total assets.
currat	Current assets/current liabilities.
dy	Total dividends divided by market capitalization at fiscal year-end.
egr	Annual percent change in book value of equity.
operprof	Revenue minus cost of goods sold - SG&A expense - interest expense divided by lagged common shareholders'equity.
pchcurrat	Percent change in currat.
pchdepr	Percent change in depr.
pchgm_pchsale	Percent change in gross margin minus percent change in sales.
rd_mve	R&D expense divided by end-of-fiscal-year market capitalization.
pchsale_pchinvt	Annual percent change in sales minus annual percent change in inventory.
pchsale_pchrect	Annual percent change in sales minus annual percent change in receivables.
pchsale_pchxsga	Annual percent change in sales minus annual percent change in SG&A.
saleinv	Annual sales divided by total inventory.
pchsaleinv	Percent change in saleinv.
rd_sale	R&D expense divided by sales.
roaq	Income before extraordinary items divided by one quarter lagged total assets.
roavol	Standard deviation for 16 quarters of income before extraordinary items divided by average total assets.
roeq	Earnings before extraordinary items divided by lagged common shareholders'equity.
rsup	Sales from quarter t minus sales from quarter t-4 divided by fiscal-quarter-end market capitalization.
salecash	Annual sales divided by cash and cash equivalents.
salerec	Annual sales divided by accounts receivable.
tang	Cash holdings + 0.715 ×receivables + 0.547×inventory + 0.535×PPE/total assets.
sue	Unexpected quarterly earnings divided by fiscal-quarter-end market cap. Unexpected earnings is I/B/E/S actual earnings minus median forecasted earnings if available, else it is the seasonally differenced quarterly earnings before extraordinary items from compustat quarterly file.
chfeps	Mean analyst forecast in month prior to fiscal period end date from I/B/E/S summary file minus same mean forecast for prior fiscal period using annual earnings forecasts.
sfe	Analysts mean annual earnings forecast for nearest upcoming fiscal year from most recent month available prior to month of portfolio formation from I/B/E/S summary files scaled by price per share at fiscal quarter end.
chnanalyst	Change in nanalyst (number of analyst forecasts available I/B/E/S summary files in month prior to month of portfolio formation) from month t-3 to month t.
disp	Standard deviation of analyst forecasts in month prior to fiscal period end date divided by the absolute value of the mean forecast. Forecast data from I/B/E/S summary files.

Table 1.12: Description of firm characteristics (continued-2)

rev3	The past three-month return reversal.
rev4	The past four-month return reversal.
oa	Operating assets.
ol	Long term operations.
noa	Net operating assets.
noa	Net operating assets.
ibq_adj	Income before extraordinary items adjusted.
O-score	Operating scores.
grnoa	Growth in net operating assets.
illiq	Illiquidity. It represents the average monthly ratio of the absolute stock return to the dollar trading volume within a five-month period
rf	Risk free return.
Log_Mkt-rf	Log market return minus risk free return.
Log_Mkt-rf^2	Log market return minus squared risk free return.
Log_Mkt-rf^4	Log market return minus risk free return in power four.
Log_Mkt-rf_lag1	Log market return minus lag one risk free return.
Log_Mkt-rf_lag2	Log market return minus lag two risk free return.
Log_Mkt-rf_lag3	Log market return minus lag three risk free return.
r_sq	Squared return.
excess_log_return	Log return minus risk free rate.
coskew	Conditional skewness.
mom	Momentum. It is the cumulative return of a stock.
chsaleq	Change in sales quantity.
chatq	Change of asynchronous thread queue (ATQ).
size	The firm's size is computed as the product of the price per share and the number of shares outstanding (in million dollars).
gammaps	Measure for price impact based on price reversals for the equity market.
max	The maximum daily return. It represents the firm's extreme positive return defined as the average of five highest daily returns over the past 21 days.
acc	Annual income before extraordinary items (ib) minus operating cash flows divided by average total assets.
cfp	Operating cash flows divided by fiscal-year-end market capitalization.
agr	Annual percent change in total assets.
rev1	The past one-month return reversal.
rev2	The past two-month return reversal.

Table 1.13: Industry Classification (SIC) and Number of Firms by Industry (Part 1)

Last two digits of SIC (“a")	**Industries (“ind_a")**	**Number of firms**
10	Metal Mining	**14**
12	Bituminous Coal and Lignite Mining	**1**
13	Oil and Gas Extraction	**52**
14	Mining and Quarrying of Nonmetallic Minerals, except Fuels	**3**
15	Building Construction General Contractors and Operative Builders	**13**
16	Heavy Construction other than Building Construction Contractors	**12**
17	Construction Special Trade Contractors	**5**
20	Food and Kinred Products	**41**
21	Tobacco Products	**2**
22	Textile Mill Products	**10**
23	Apparel and other Finished Products Made from Fabrics and Similar Materials	**11**
24	Lumber and Wood Products, except furniture	**5**
25	Furniture and Fixtures	**16**
26	Paper and Allied Products	**16**
27	Printing, Publishing, and Allied Industries	**23**
28	Chemicals and Allied Products	**189**
29	Petroleum Refining and Related Industries	**11**
30	Rubber and Miscellaneous Plastics Products	**14**
31	Leather and Leather Products	**8**
32	Stone, Clay, Glass, and Concrete Products	**9**
33	Primary Metal Industries	**24**
34	Fabricatel Metal Products, except Machinery and Transportation Equipment	**32**
35	Industrial and Commercial Machinery and Computer Equipment	**104**
36	Electronic and other Electrical Equipment and Components, except Computer Equipment	**135**
37	Transportation Equipment	**55**
38	Measuring, Analyzing, and Controlling Instruments; Photografic, Medical, and Optical Goods; Watches and Clocks	**117**
39	Miscellaneous Manufacturing Industries	**17**
40	Railroad Transportation	**4**
41	Local and Suburban Transit and Interurban Highway Passenger Transportation	**0**
42	Motor Freight Transportation and Warehousing	**15**
43	United States Postal Service	**0**
44	Water Transportation	**4**
45	Transportation by Air	**13**
46	Pipelines, except Natural Gas	**0**
47	Transportation Services	**5**
48	Communications	**48**
49	Electric, Gas, and Sanitary Services	**74**
50	Wholesale Trade-Durable Goods	**54**
51	Wholesale Trade-Nondurable Good	**26**
52	Building Materials, Hardware, Garden Supply, and Mobile Home Dealers	**6**
53	General Merchandise Stores	**12**

Table 1.13: Industry Classification (SIC) and Number of Firms by Industry (continued)

Last two digits of SIC ("a")	Industries ("ind_a")	Number of firms
54	Food Stores	**5**
55	Automotive Dealers and Gasoline Service Stations	**15**
56	Apparel and Accessory Stores	**25**
57	Home Furniture, Furnishings, and Equipment Stores	**10**
58	Eating and Drinking Places	**26**
59	Miscellaneous Retail	**27**
60	Depository Institutions	**204**
61	Non-Depository Credit Institutions	**22**
62	Security and Commodity Brokers, Dealers, Exchanges, and Services	**32**
63	Insurance Carriers	**63**
64	Insurance Agents, Brokers and Service	**6**
65	Real Estate	**7**
67	Holding and other Investment Offices	**50**
70	Hotels, Rooming Houses, Camps, and other Lodging Places	**11**
72	Personal Services	**4**
73	Business Services	**178**
75	Automobile Repairs, Services, and Parking	**3**
76	Miscellaneous Repair Services	**0**
78	Motion Pictures	**5**
79	Amusement and Recreation Services	**16**
80	Health Services	**29**
81	Legal Services	**0**
82	Educational Services	**11**
83	Social Services	**3**
86	Membership Organizations	**0**
87	Engineering, Accounting, Research, Management, and Related Services	**54**
89	Miscellaneous Services	**1**
91	Executive, Legislative, and General Government, except Finance	**0**
92	Justice, Public Order, and Safety	**0**
93	Public Finance, Taxation, and Monetary Policy	**0**
94	Administration of Human Resource Programs	**0**
95	Administration of Environmental Quality and Housing Programs	**0**
96	Administration of Economic Programs	**0**
97	National Security and International Affairs	**0**
99	Nonclassifiable Establishment	**19**
TOTAL	-	**2026**

Chapter 2

A Regime-Switching Approach for Detecting Shock Transmission Types across Asset Markets, with an Application to Sovereign Bond Markets

2.1 Introduction

The manner and the extent to which shocks are transmitted across international asset markets has been, and continues to be, of much interest to researchers, financial investors, and policy-makers. It is specially relevant during periods of crises, when transmitted negative shocks are much larger than usual, and where the economic and financial damages can be substantial, both, in the country where the shock originates, and also in the countries where the shock propagates to. The subprime market collapse that occurred in the U.S. and that resulted in the Great Financial Recession of 2007-2008 is a recent example. The Covid-19 pandemic and its impacts worldwide is another.

In the academic literature related to this topic, one strand focuses on the theoretical channels through which shocks may propagate, such as bilateral trade, inefficiencies in financial markets, or herding behavior by investors (see, for instance, Claessens and Kose (2018)), while another strand concentrates on empirically differentiating between the types of shock transmission, which are interdependence, contagion, and decoupling (Dungey et al. (2004) provide a survey). In particular, studies try to distinguish between interdependence, where during a crisis, shocks, regardless of their size, transit from one market to the other via existing (normal) channels, and contagion or decoupling, where the interaction structures between the markets change during the transmission process. Although there is disagreement on the precise definition of the concepts, contagion is generally deemed to occur when the shock transmission results in additional (or, excessive) co-movement in the concerned asset markets, and decoupling occurs when this co-movement is less than what would have been observed in an interdependence situation (for example, see Dungey et al. (2020)).

To distinguish between interdependence and contagion, tests have been proposed.[1] While several approaches have been advanced in this regard, recent work has primarily adopted factor-based empirical frameworks for modeling the "normal" interactions, and has looked for breaks in this structure to signal the presence of contagion. For instance, Bekaert and Harvey (2003) and Bekaert et al. (2014) both model normal interactions between asset market returns via observable common factors in the means and a given variance-covariance structure for their residuals. Then, during a crisis period, contagion is deemed to occur, for the former, upon finding a significantly higher cross-correlation in the residual terms, and for the latter, when changes occur in either, the residual cross-correlations, or, in the factor exposures of the returns. A somewhat similar strategy to Bekaert et al. (2014) is also followed in Dungey and Renault (2018), although the estimation approach is very different.

In all of the above cases, break tests are applied to crisis periods where the start and end dates are designated *ex ante*. However, any lack of concordance between the chosen and actual dates for the aforementioned crisis episodes may be problematic for the results obtained (see Gravelle et al. (2006)). Another issue is that, since break tests are applied, they are likely to pick up only quick and abrupt contagion episodes, not more gradual ones. Campos-Martins and Amado (2022) demonstrate the existence of short-term and long-term contagion effects, and stress the importance of also distinguishing between them.

In this paper, we propose a regime-switching factor-based framework to test for either contagion, or decoupling, against the alternative of interdependence. We model asset market returns pairwise, with means that include observable asset-specific and global factors, and with a variance-covariance structure that allows for the presence of a common (unobservable) shock in addition to idiosyncratic shocks. The common shock series captures global and regional uncertainty elements beyond the systematic features already present in the means (for example, unexpected changes in world commodity markets, geopolitical and health conditions, and in U.S. monetary policy). The three shocks are assumed to evolve according to independent Markov regime-switching processes, with the low-variance regimes representing

[1] Historically, contagion has been studied more than decoupling.

normal times and the high-variance regimes as crisis times. Our crisis and normal periods are thus data-determined.

We then propose bootstrap-based Student's t-tests to examine whether high-variance idiosyncratic shocks originating in one of the markets significantly alter the interactions between the two markets on average, via changes in selected factor loadings in the means of these market returns. The test can be designed to detect either contagion or decoupling. Moreover, our estimated Markov probabilities can shed light on whether particular crises episodes were short term or long term events.

Our paper is related to Gravelle et al. (2006) who propose a regime-switching structure to test for shift contagion, but there are important differences between this study and theirs. The works of Bekaert et al. (2014), Dungey and Renault (2018), and others have shown the importance of including all of the sources affecting the mean processes of the financial asset returns, notably by taking into account global drivers which have proven important in the evolution of financial assets as globalization has increased (for a recent reference, see Miranda-Agrippino and Rey (2021)). Whereas the Gravelle et al. (2006) model mean lacks this dimension, our study includes it. Another feature stressed by the factor-based literature above is the need to pinpoint the specific market where the contagion shock originated and all the markets where the shock was transmitted to. Our study also follows this strategy. This is different from the proposed approach of Gravelle et al. (2006) where it is not possible to distinguish between contagion shock origin and destination. More precisely, our contagion-producing impetus originates as an idiosyncratic shock in one of the markets, and it gets transmitted to the other market via the mean structure of the assets. As such, our model empirically separates these contagion-producing shocks from the common shocks in the model, the latter assumed to be different in nature and having a different impact on the assets. That is, we think of common shocks as most likely originating outside of the two markets considered, and we allow them to only influence the variance-covariance structure of the assets (that is, their uncertainty). In contrast, Gravelle et al. (2006) assume that contagion-generating shocks originate from amongst the common shocks.

An example showing the above distinctions is the following: consider the case of two equity markets (say, of Brazil and Argentina), which both are influenced by a large common shock (say the Covid-19 pandemic shock). Such a shock is unlikely to alter the interactions at the mean level between the equity markets of the two countries, but it is likely to affect their volatilities. Contrast this to an important idiosyncratic shock originating in one of the markets (say, Brazil), and which may then "contage" to the other (Argentina) equity market via real, financial, and political channels that may be present between the two markets as it may have happened, for example, when a political scandal occurred in May 2017 in Brazil that strongly impacted the Brazilian equity markets and also spilled over to the Argentinian equity market via cross-listed firms – see Hillier and Loncan (2019). Thus, while our modeling strategy would distinguish between these different source shocks, Gravelle et al. (2006) does not.

We demonstrate our approach by testing for contagion and decoupling in the sovereign bond market of six Latin American country pairs. One reason for focusing on these markets is that the development of international financial markets in the 1980s expanded greatly the use of such debt instruments in these countries. Moreover, these markets experienced considerable volatility, and that appear to co-move increasingly during these periods (Lin et al., 2008). Another reason is that these cases were originally considered by Gravelle et al. (2006), and it is useful to compare results from our model to theirs.

Using monthly data, we estimate our model and the Gravelle et al. (2006) model over the sample period of May 1993 to September 2023.[2] A number of diagnostic checks then confirm that our factor-based modelling structure is better specified. We then apply our proposed tests to examine whether contagion or decoupling occurred, on average, over our full sample.[3] We find evidence, at the 5% level, of decoupling having occurred in two of our six country pair sovereign bond market cases (namely, Brazil/Mexico, where Brazil is the source of crisis and Mexico is the target country, and Argentina/Mexico, where Argentina is the source of crisis and Mexico is the target country). Below we discuss how this is likely due to Mexico adopting policies that shielded the country from these types of impacts. Interestingly, we find no evidence of contagion on average over our sample in any of our considered cases, implying only interdependence among these bond markets. Finally, we show that our model has good forecasting performance out of sample, notably outperforming the model proposed in Gravelle et al. (2006).

[2]The data sample in the Gravelle et al. (2006) study ended in 2001. Note also that while the original study used weekly data frequency, limitations in data availability have forced us to work with monthly data instead.

[3]Note that while we test for contagion/decoupling on average over our sample, with higher frequency data we could equally have applied our tests to specific crisis periods.

The remainder of this paper is organized as follows. Section 2.2 includes: (1) the description of the global and local PCA factors, (2) the definition of the Markov-switching factor-augmented model, (3) the description of the data, (4) the bootstrap-based Student's t-test algorithm, and (5) the forecasting evaluation. Section 2.3 provides the empirical results. Section 2.4 concludes.

2.2 Methodology

As shown in Miranda-Agrippino and Rey (2021), financial assets seem to be driven in part by a few common global financial features. An asset pair is also likely to be influenced by common conditions local to both, stemming notably from trade relations between their corresponding countries. In our modeling strategy, we will thus assume that the mean process of each of our asset pair returns is composed of three parts: one representing its common global component, another that captures its common local component (that is, specific only to the two markets), and a third part, which is an unpredictable error term.

We obtain two large sets of variables, the first containing observable series that capture global features shown to simultaneously affect many types of financial returns in the world, and the second containing trade-related variables and other variables specific to each market pair considered. We then apply Principal Component Analysis (PCA) to each of the two groups of series to yield our global and different country pair local factors. The asset returns of each market pair are thus assumed to load onto these two types of factors. In the following, we describe the construction of these principal components.

2.2.1 Construction of global PCA and local PCA features: PCA analysis

Let X_t be the matrix $T \times l$ of global relevant characteristics (common to every pair of financial assets), and P_t be the matrix $T \times (l_i + l_j)$ of (other) characteristics that are only relevant to the pair of considered assets i and j. So, $X_t = (x_t^{(1)}, ..., x_t^{(l)})$ and $P_t = (p_{i,t}^{(1)}, ..., p_{i,t}^{(l_i)}, p_{j,t}^{(1)}, ..., p_{j,t}^{(l_j)})$.

Let $M = (m_1, ..., m_v)$ and $N = (n_1, ..., n_{v_{ij}})$ be the $l \times v$ matrix of weights associated with the global characteristics and the $(l_i + l_j) \times v_{ij}$ matrix of weights associated with the local characteristics of the market pair i and j, respectively. $v < l$, $v_{ij} < (l_i + l_j)$, m_v represents the vector of linear combination of weights used to construct the v^{th} uncorrelated process, and $m_{v_{ij}}$ represents the vector of linear combination of weights used to construct the v_{ij}^{th}

uncorrelated process.

The $(T \times v)$ matrix of global PCA factors is given by $f_t^{(G)} = X_t M$ where the vector of linear combination of weights is defined as:

$\underset{M}{\text{Argmax}}\,\text{Var}(X_t M)$, s.t $M^\top M = 1$, $cov(X_t^\top m_v, X_t^\top m_l) = 0$, $l = 1, 2, .., (v-1)$, and $f_t^{(G)} = (f_{1,t}^{(G)}, ..., f_{v,t}^{(G)})$.

The first principal component (PC), $l = 1$, captures the most variance. The second PC, $l = 2$, is the solution to $\underset{m_2}{\text{argmax}}\,\text{Var}(X_t^\top m_2)$, s.t $\|m_2\|_2 = 1$, and $Cov(X_t^\top m_1, X_t^\top m_2) = 0$.

The $(T \times v_{ij})$ matrix of local PCA factors is given by $f_t^{(ij)} = P_t N$ where the vector of linear combination of weights is defined as:

$\underset{N}{\text{Argmax}}\,\text{Var}(P_t N)$, s.t $N^\top N = 1$, $cov(P_t^\top n_{v_{ij}}, P_t^\top n_{(l_i + l_j)}) = 0$, $(l_i + l_j) = 1, 2, .., (v_{ij} - 1)$, and $f_t^{(ij)} = (f_{1,t}^{(ij)}, ..., f_{v_{ij},t}^{(ij)})$.

Intuitively, while the global PCA factors represent v uncorrelated components that capture the maximum variance of the global features, the local PCA factors represent v_{ij} uncorrelated components that capture the maximum variance of local features of a country pair. The series on which the above PCA analysis is applied are given in the data section below.

2.2.2 The Markov Switching Factor Augmented (MS-FACTOR) model

To illustrate our methodology, and as in Gravelle et al. (2006), we consider sovereign bond yield spreads for pairs of countries, measured relative to comparable U.S. Treasury bonds.[4] In international financial markets, the spread of sovereign bonds is used to measure the riskiness of a country (that is, its risk of sovereign default), and it has been shown to be predictable (Bianchi et al., 2020; Cochrane & Piazzesi, 2005; Sarno et al., 2016; Thornton & Valente, 2012). We construct the bond return of a given country by taking the natural log of its bond spread, first differencing it, and multiplying this by 100. In terms of our notation, $r_{i,t}$ and $r_{j,t}$

[4]In this study, we use the J.P. Morgan EMBI+ Index to represent emerging sovereign bond yield spreads. This index includes only U.S. dollar–denominated sovereign bonds with a minimum maturity of 2.5 years, measured relative to comparable U.S. Treasury bonds.

thus represent these returns for countries i and j at time t.

From the earlier step, we retain the largest principal component of each of the global and local PCA factors.[5] We denote the global first factor as $f_{1,t}^{(G)}$, and the local first factor as $f_{1,t}^{(ij)}$. Thus, while the former captures the largest fraction of the common variation in the Global Financial Cycle (GFC) and the Global Trade and Commodity Cycle (GTCC; see Miranda-Agrippino and Rey (2021) for evidence on the existence of these cycles), the latter captures the largest fraction of the common variation in trade, production, and financial interaction variables directly relevant to the two markets considered. The returns of the two asset markets i and j are thus influenced by: (i) the lag of the observed local factor, $f_{1,t-1}^{(ij)}$; (ii) the lag of the observed global factor, $f_{1,t-1}^{(G)}$; (iii) a common shock, $z_{c,t}$; and (iv) idiosyncratic shock of each market, $z_{h,t}$, where $h = i, j$.

$$r_{i,t} = \theta_{1,t} f_{1,t-1}^{(ij)} + \lambda_1 f_{1,t-1}^{(G)} + u_{i,t}, \tag{2.1}$$

$$r_{j,t} = \theta_{2,t} f_{1,t-1}^{(ij)} + \lambda_2 f_{1,t-1}^{(G)} + u_{j,t}, \tag{2.2}$$

$$u_{i,t} = \sigma_{c,i,t} z_{c,t} + \sigma_{i,t} z_{i,t}, \tag{2.3}$$

$$u_{j,t} = \sigma_{c,j,t} z_{c,t} + \sigma_{j,t} z_{j,t}, \tag{2.4}$$

where, $\theta_{1,t}$ and $\theta_{2,t}$ are the time-varying loading parameters associated with $f_{1,t-1}^{(ij)}$ for the asset markets i, and j, respectively, while λ_1 and λ_2 are the loading parameters associated with the lag global factor affecting asset markets i, and j, respectively. The $u_{h,t}$ ($h = i, j$) are zero mean forecast errors and they are assumed to be uncorrelated across time (i.e, $\mathbf{E}_t[u_{h,t+k}|\mathcal{F}_{t-1}] = 0$ for all $k > 0$ and $\mathbf{E}[u_{h,t} u_{h,t'}|\mathcal{F}_{t-1}] = 0$, for $t \neq t'$, $\mathcal{F}_{t-1} = \{f_{1,t-1}^{(ij)}, f_{1,t-1}^{(G)}\}$). As in Gravelle et al. (2006), the forecast errors of the two asset returns are assumed to be contemporaneously correlated (i.e, $\mathbf{E}[u_{j,t} u_{i,t}|\mathcal{F}_{t-1}] \neq 0$).

The common shock $z_{c,t}$ represents unpredictable common global and regional disturbances, whereas the idiosyncratic shocks $z_{i,t}$ and $z_{j,t}$ denote unexpected shocks specific to individual markets. We assume that ($z_{c,t}$, $z_{i,t}$, $z_{j,t}$) are conditionally uncorrelated—both over time and with each other—and that they have zero conditional mean ($\mathbf{E}_t[z_{c,t+k}|\mathcal{F}_{t-1}] = 0$, $\mathbf{E}_t[z_{h,t+k}|\mathcal{F}_{t-1}] = 0$ for all $k > 0$, $\mathbf{E}[z_{c,t} z_{c,t'}|\mathcal{F}_{t-1}] = 0$, $\mathbf{E}[z_{h,t} z_{h,t'}|\mathcal{F}_{t-1}] = 0$ for $t \neq t'$ and

[5] Below we discuss how much of the existing common variability the largest principal components are able to capture. In the Appendix, we provide some sensitivity results to the use of additional factors.

$\mathbf{E}[z_{v,t} z_{v',t} | \mathcal{F}_{t-1}] = 0$ for $v \neq v'$). We also normalize the variances of all of these structural shocks at one, so that the corresponding σ impact coefficients can be interpreted as the standard deviations of the respective structural shocks.

The shock $z_{c,t}$ can be interpreted as a (latent) common uncertainty factor that influences both asset returns through their error structures (the subscript c denotes attributes related to the common shock). In this respect, the standard deviations, $\sigma_{c,i,t}$ and $\sigma_{c,j,t}$ capture the impact of this common latent factor on the variances of asset returns i and j, respectively. We can thus see that, in this model, co-movements between the two asset returns exist via the presence of the local and global factors found in the conditional means, and also, via the presence of the common uncertainty shocks found in their error structures.

Following Gravelle et al. (2006), and using much of their notation, we assume that each of the three shocks in the model ($z_{c,t}$, $z_{i,t}$, and $z_{j,t}$) switches between low (L)- and high (H)-volatility regimes according to distinct Markov switching processes, and where the underlying state variables are $A_{r,t}$, $r = c, i, j$. The impact coefficients of these shocks thus evolve according to:

$$\sigma_{c,h,t} = \sigma_{c,L,h}(1 - A_{c,t}) + \sigma_{c,H,h} A_{c,t}, \qquad h = i, j \tag{2.5}$$

$$\sigma_{h,t} = \sigma_{L,h}(1 - A_{h,t}) + \sigma_{H,h} A_{h,t}, \qquad h = i, j, \tag{2.6}$$

where $A_{r,t} = \{0, 1\}$, for $r = c, i, j$, and the values 0 and 1 denote low-, and high-volatility regimes, respectively. $\sigma_{c,L,h}$ are the impact coefficients of the common shock affecting asset h ($h = i, j$) in the low-volatility regime, while $\sigma_{c,H,h}$ are the corresponding high-volatility regime parameters. Similarly, $\sigma_{L,h}$ are the impact coefficients of the idiosyncratic shocks of asset h ($h = i, j$), in the low-volatility regime, whereas $\sigma_{H,h}$ denote the corresponding high-volatility parameters. The variances and the covariances of the forecast errors in the high volatility regime (i.e., for $A_{r,t} = 1$) are:

$$var(u_{i,t} | A_{c,t} = 1, \mathcal{F}_{t-1}) = \sigma^2_{c,H,i} + \sigma^2_{L,i}, \tag{2.7}$$

$$var(u_{j,t} | A_{c,t} = 1, \mathcal{F}_{t-1}) = \sigma^2_{c,H,j} + \sigma^2_{L,j}, \tag{2.8}$$

$$cov(u_{i,t}, u_{j,t} | A_{c,t} = 1, \mathcal{F}_{t-1}) = \sigma_{c,H,i} \times \sigma_{c,H,j}, \tag{2.9}$$

$$var(u_{i,t}|A_{i,t}=1,\mathcal{F}_{t-1})=\sigma^2_{c,L,i}+\sigma^2_{H,i}, \tag{2.10}$$

$$var(u_{j,t}|A_{j,t}=1,\mathcal{F}_{t-1})=\sigma^2_{c,L,j}+\sigma^2_{H,j}. \tag{2.11}$$

The transition probabilities to stay in the low- and high-volatility regime are:

$$Pr[A_{r,t}=0|A_{r,t-1}=0]=q_r, \tag{2.12}$$

$$Pr[A_{r,t}=1|A_{r,t-1}=1]=p_r, \tag{2.13}$$

where $r=c,i,j$. q_r and p_r are the probabilities of transitioning from a low-volatility regime to a low-volatility regime, and from a high-volatility regime to a high-volatility regime, respectively.

In the above modeling context, we assume that contagion or decoupling occurs conditional on the idiosyncratic shock of the origin market i being in the high-variance regime. However, this is only a necessary condition, and not a sufficient one. To establish that either contagion or decoupling has occurred, we require significant changes in the loadings of the local PCA factors entering the means of the asset returns. We formalize this idea by modeling the loadings on the local factors $\theta_{s,t}$, with $s=1,2$, as being dependent on the state of the asset i's volatility regime. This is presented in the equation below:

$$\theta_{s,t}=\theta_{s,L}(1-A_{i,t})+\theta_{s,H}A_{i,t}, \qquad s=1,2. \tag{2.14}$$

Finally, in the interest of maintaining model simplicity, we assume that the loadings of the lag global factor remain constant.

2.2.3 Data description

We use monthly emerging market bond yield spreads, which are spread-yields on the Emerging Market Bond Index Plus (EMBI+) constructed by J.P. Morgan and they are U.S.-dollar-dominated. The J.P. Morgan EMBI+ is a widely used benchmark index for measuring the performance of sovereign bonds issued by emerging market countries. EMBI+ spread refers to the difference between the yield on the bonds in the EMBI+ index and the yield on comparable U.S. Treasury securities (considered risk-free). We focus on three Latin American countries, namely, Brazil, Mexico, and Argentina.

To construct global PCA factors, we consider variables (considered relevant for capturing the GFC) that denote fluctuations in financial activity on a global scale (Rey, 2015), and also variables (considered as being part of the GTCC) that denote fluctuations in output, commodity market, and trade market on a global scale (Miranda-Agrippino & Rey, 2021).

The GFC variables include the volatility of the U.S. stock market (VIX index), world private liquidity, world cross-border capital flows, U.S. federal funds rate, U.S. real effective exchange rate, term, spread, and S.P500. The GTCC variables include spot price of WTI crude oil, world trade, and world industrial production as a proxy for global production. With the expansion of financial globalization, studies show that the latter drive many financial asset returns, including stock returns and sovereign bond spreads (Gilchrist et al., 2022; Miranda-Agrippino & Rey, 2021).

We also construct local PCA factors, specific to each sovereign bond market pair under consideration. A set of country-specific and pairwise trade-related variables is retained for each asset pair case. For the Mexico-Brazil pair, we include the real effective exchange rate, industrial production, inflation, trade, and the stock market index for both countries. For the Mexico-Argentina pair, we include the real effective exchange rate, industrial production, stock market index, inflation, and trade for both countries. For the Brazil-Argentina pair, we include the real effective exchange rate, industrial production, inflation, and trade for both countries.[6]

We test the stationarity of all these series using Dickey-Fuller tests and KPSS tests, and, where needed, we take first differences to make the series stationary. We then use the data sets described above to generate a single set of global principal components, as well as three sets of local principal components via the Principal Component Analysis method. For our base case, we make use of only the largest principal component of each of these. These are relatively informative as they represent, on average, 60% of the total variance of the relevant features representing the global factors, and approximately 65% of the total variance of the relevant "local" features for each market pair.

Our data span the period from May 1993 to September 2023. Emerging market bond spreads are collected from the World Bank website, and data for the other variables are collected from the Federal Reserve Economic Data (FRED), the World Bank, and Bloomberg.

[6] The variables used reflect the data that is available for each market pair.

2.2.4 A bootstrap-based Student's t-test for detecting contagion or decoupling

Our objective is to assess whether, during a crisis period, an idiosyncratic shock originating from one market and spilling over to the other can be interpreted as *interdependence*, *contagion*, or *decoupling*. We characterize *contagion* or *decoupling* as a significant change in the loadings of the local factor within the mean processes of the asset markets, given that the idiosyncratic shock of the origin market is in the high-variance regime. *Contagion* is identified when these factor loadings shift in a way that amplifies the impact of the original shock on the mean processes, whereas *decoupling* is said to occur if the impact is instead dampened.

After estimating the model (2.1)–(2.14) via maximum likelihood, and recalling that we are using the subscript L to designate a low-variance regime, and H, a high-variance regime, we consider the following hypotheses:

Or

$$\left\{\begin{array}{lll} H_0: & \left|\frac{\theta_{2,H}}{\theta_{1,H}}\right| = \left|\frac{\theta_{2,L}}{\theta_{1,L}}\right| & \Rightarrow \text{Interdependence} \\ H_{1a}: & \left|\frac{\theta_{2,H}}{\theta_{1,H}}\right| > \left|\frac{\theta_{2,L}}{\theta_{1,L}}\right| & \Rightarrow \text{Contagion} \end{array}\right\} \left\{\begin{array}{lll} H_0: & \left|\frac{\theta_{2,H}}{\theta_{1,H}}\right| = \left|\frac{\theta_{2,L}}{\theta_{1,L}}\right| & \Rightarrow \text{Interdependence} \\ H_{1b}: & \left|\frac{\theta_{2,H}}{\theta_{1,H}}\right| < \left|\frac{\theta_{2,L}}{\theta_{1,L}}\right| & \Rightarrow \text{Decoupling} \end{array}\right\}.$$

Given that the idiosyncratic shock originates in market i, the first hypothesis (H_0) implies that there is no statistical difference in the (absolute value of the) relative local factor loadings of the two markets across high and low variance regimes. In other words, we have a situation of interdependence. The second hypothesis (H_{1a}) designates a situation of contagion, since the factor loading of the target market relative to that of the origin market is greater in the high (crisis) regime versus in the low (normal) regime. That is, the impact of the origin shock has been amplified in the target market. As for the third hypothesis (H_{1b}), it describes a situation of decoupling, with the impact of the origin shock having become dampened in the target market.

To distinguish between these hypotheses, we propose the following bootstrap-based Student's t-test procedure:

Step 1: Estimate the model (2.1)–(2.14) with the observed data without imposing the equality constraint H_0, and obtain the estimated transition probabilities and estimated factor loadings. Then, compute the absolute ratio of the local factor loadings, in the high regime ($\widehat{\tau}_1$)

and in the low regime ($\widehat{\tau}_2$). Thus, we have $\widehat{\tau}_1 = \left|\frac{\widehat{\theta}_{2,H}}{\widehat{\theta}_{1,H}}\right|$, while $\widehat{\tau}_2 = \left|\frac{\widehat{\theta}_{2,L}}{\widehat{\theta}_{1,L}}\right|$. Next, compute the value of the statistic $tstat = \sqrt{T}(\widehat{\tau}_1 - \widehat{\tau}_2)$. Note that a value of zero designates a situation of interdependence under the null hypothesis.

Step 2: Now, impose the null hypothesis, and estimate the model to get the estimated matrix of transition probabilities. This problem is basically a *constrained* Maximum Likelihood Estimation (CMLE). Calculate the ratios $\widehat{\tau}_{null,1} = \left|\frac{\widehat{\theta}_{null,2,H}}{\widehat{\theta}_{null,1,H}}\right|$ and $\widehat{\tau}_{null,2} = \left|\frac{\widehat{\theta}_{null,2,L}}{\widehat{\theta}_{null,1,L}}\right|$.

Step 3: Generate the state variables $A_{c,t}$, $A_{i,t}$ and $A_{j,t}$ for $t = 0, ..., T$ as Markov chains using the estimated transition probabilities estimated in Step 2.

Step 4: Draw three $(T \times 1)$ sequences of standard Gaussian variables, to represent $z_{c,t}$, $z_{i,t}$ and $z_{j,t}$, $t = 0, ..., T$, and where all are independent of the Markov chains generated in Step 3.

Step 5: Given the samples $z_{c,t}$, $z_{i,t}$ and $z_{j,t}$, $t = 0, ..., T$ obtained in Step 4, the estimated parameters of the model (2.1)–(2.14) obtained in Step 2, and the state variables obtained in Step 3, draw two bootstrap samples $r^*_{i,t}$ and $r^*_{j,t}$, $t = 0, ..., T$, for a market pair (i, j).

Step 6: Fit the samples $r^*_{i,t}$ and $r^*_{j,t}$ to the model (2.1)–(2.14) without imposing any constraint. Next, compute the absolute ratio of the local factor loadings in the high regime ($\widehat{\tau}^*_1$) and in the low regime ($\widehat{\tau}^*_2$). Then, compute the *normalized* bootstrap statistic:

$tstat^* = \sqrt{T}((\widehat{\tau}^*_1 - \widehat{\tau}^*_2) - (\widehat{\tau}_{null,1} - \widehat{\tau}_{null,2}))$, where $(\widehat{\tau}_{null,1} - \widehat{\tau}_{null,2}) = 0$ by construction.

Step 7: Loop through Steps 3 to 6 $B = 499$ times, and obtain B bootstrap statistics: $tstat^*_1,...,tstat^*_B$.

Assuming that the bootstrap statistics are independent realizations, there are no ties between these normalized bootstrap statistics and the *tstat*. Now, count the number of times (say, $\overline{B}$) when the sample statistic *tstat* is smaller than or equal to the simulated statistics $tstat^*_b$, $b = 1, ..., B$. This gives us the right-sided bootstrap p–value:

$$p\text{-value}^* = \frac{\overline{B} + 1}{B + 1}.$$

If p-value* is a small value (or $\leq \alpha$, where α is a given level of significance), so that *tstat* is highly likely to be positive, then we have *contagion*. Otherwise, we do not have *contagion*.

To test for *decoupling*, we calculate the left-sided bootstrap p-value by counting the number of time $(\overline{B})$ when the sample statistic *tstat* is greater than or equal to the bootstrap statistics $tstat_b^*$, $b = 1, ..., B$. In this case, if p-value* is small (or $\leq \alpha$) so that *tstat* is highly likely to be negative, we have *decoupling*. Otherwise, we do not have *decoupling*.

2.2.4.1 Justification of the validity of the Bootstrap-based Student's t-test

While proving the theoretical validity of the proposed procedure is quite challenging and beyond the scope of this paper, we nevertheless advance some arguments as to the likely validity of the testing procedure that we have proposed.

We start by noting that our approach follows along the lines of Kasahara and Shimotsu (2018), who develop a rigorous framework to justify bootstrap-based inference in Markov regime-switching models, specifically for testing the number of regimes. In their work, the authors establish the validity of the bootstrap likelihood ratio test statistic (LRTS) for determining whether a model with M regimes provides a significantly better fit than one with (M-1) regimes. This is accomplished by proving that, under a set of regularity conditions—such as compactness of the parameter space, regime separability, identifiability, and smoothness of the likelihood function—the bootstrap LRTS converges in distribution to the correct limiting distribution, even in models with regime-dependent heteroskedasticity and nuisance parameters. Inspired by this theoretical architecture, we propose that our bootstrap-based Student's t-test—used to test for contagion or decoupling versus interdependence in a regime-switching factor model—can be justified analogously. Specifically, we construct a normalized test statistic that captures the difference in relative factor loadings between high- and low-variance regimes, and evaluate it using bootstrap samples drawn under the null-imposed Markov-switching structure.

We further rely on similar principles of conditional bootstrap consistency and weak convergence under Markov-switching structures. By simulating bootstrap samples based on the constrained maximum likelihood estimates under the null hypothesis of interdependence (where the relative local factor loadings remain constant across regimes), and using the associated estimated transition probabilities, we generate bootstrap statistics that preserve the regime structure and parameter uncertainty inherent in the original model. As in Kasahara and Shimotsu (2018), convergence of the bootstrap distribution can be established using a continuous mapping theorem and stochastic equicontinuity of the objective function, provided certain regularity conditions hold. These include the consistency and asymptotic normality of the estimators, the existence of a true regime structure, and boundedness of the parameter space. Thus, even though we do not present a formal proof, our procedure is grounded in the same principles of asymptotic theory and weak convergence used in bootstrap inference for regime-switching models, and satisfies similar regularity conditions to ensure its theoretical soundness.

2.3 Empirical results

2.3.1 Econometric validation

For completion, we first re-visit the original Gravelle et al. (2006) study using our monthly data. We estimate this model over the full sample (May 1993 to September 2023). We then report the obtained estimates in Table 2.6 of the Appendix. Before drawing any conclusions about how the dynamics of shocks may have changed across regimes and country pairs, we verify whether the model fits the data well.

Table 2.1: Diagnostics Checks; Verifying Fit of Gravelle et al. (2006) Model

	Market pairs					
	Mexico	Argentina	Brazil	Argentina	Mexico	Brazil
$Q(1)$	0.012	0.039	0.003	0.044	0.032	0.012
$Q(4)$	0.032	0.010	0.025	0.075	0.012	0.034
JB	0.420	0.432	0.070	0.139	0.033	0.204
$ARCH(1)$	0.626	0.511	0.303	0.567	0.909	0.372
$ARCH(4)$	0.980	0.855	0.519	0.801	0.998	0.562

This table reports the p-values of several diagnostic tests for each market pair. $Q(k)$ refers to the Ljung-Box test for the null of no autocorrelation up to lag k. JB represents the Jarque-Bera test for the null of normality. $ARCH(k)$is the Lagrange Multiplier test for the null of no $ARCH$ effects of order k.

Table 2.1 presents the results of autocorrelation, normality, and volatility tests applied to the standardized residuals of the bond returns for each country pair. We observe that while the null hypotheses of normality and no autoregressive conditional heteroskedasticity (ARCH) are broadly not rejected (p-values > 5%), the null hypothesis of no autocorrelation is strongly rejected overall for all country pairs (p-values < 5%). This remaining autocorrelation in the residuals likely indicates that the interaction at the mean level between the two bond returns has not been well-captured in this model.

Table 2.2: Diagnostics Checks; Verifying Fit of Our Proposed Model (MS-FACTOR)

	Market pairs					
	Brazil	Mexico	Brazil	Argentina	Mexico	Brazil
$Q(1)$	0.250	0.056	0.059	0.137	0.211	0.359
$Q(4)$	0.586	0.200	0.054	0.556	0.427	0.598
JB	0.345	0.355	0.059	0.235	0.840	0.001
$ARCH(1)$	0.949	0.490	0.463	0.784	0.717	0.679
$ARCH(4)$	0.987	0.929	0.567	0.887	0.853	0.685
	Mexico	Argentina	Argentina	Mexico	Argentina	Brazil
$Q(1)$	0.077	0.182	0.058	0.108	0.199	0.455
$Q(4)$	0.110	0.085	0.060	0.082	0.554	0.549
JB	0.078	0.477	0.003	0.088	0.840	0.075
$ARCH(1)$	0.681	0.753	0.910	0.640	0.085	0.137
$ARCH(4)$	0.845	0.807	0.922	0.640	0.085	0.137

This table reports the p-values of several diagnostic tests for each market pair. $Q(k)$ refers to the Ljung-Box test for the null of no autocorrelation up to lag k. JB represents the Jarque-Bera test for the null of normality. $ARCH(k)$ is the Lagrange Multiplier test for the null of no $ARCH$ effects of order k.

We next examine the fit of our proposed specification (denoted hereafter the MS-FACTOR model), integrating our desired features into the means and the variance-covariance structures of the returns.[7] We now observe from Table 2.2 that the null hypothesis of no autocorrelation is not rejected for all country pairs (p-values >5%). Additionally, with the exception of two cases, the null hypothesis of normality is also not rejected. This suggests that the lagged local factor and lagged global risk factor capture fairly adequately the autocorrelations in the standardized residuals. Therefore, our proposed model specification appears to be well-suited for conducting our contagion and decoupling analyses.

2.3.2 Estimation results of the MS-FACTOR model

Before testing formally for contagion or decoupling, we examine the estimated impacts of idiosyncratic and common shocks on bond spread returns across regimes. Next, we assess the estimated global factor loadings. Then, in the following section, we analyze the transmission of idiosyncratic shocks between the bond markets.

[7]Moreover, we remind that contagion or decoupling shocks are assumed not to emanate from the common shocks in this model setup.

2.3.2.1 Impacts of the common and idiosyncratic shocks in low and high regimes

Table 2.3: Estimated Parameters of MS-FACTOR Model

Parameters	Brazil/Mexico	Brazil/Argentina	Mexico/Brazil	Mexico/Argentina	Argentina/Mexico	Argentina/Brazil
Variance parameters						
$\sigma_{c,L,1}$	**6.754**	**5.474**	**5.973**	**5.342**	**6.939**	**5.754**
	(1.105)	(0.527)	(0.880)	(2.064)	(1.500)	(0.635)
$\sigma_{c,L,2}$	**6.874**	**8.966**	**6.340**	**5.129**	**2.935**	**6.191**
	(1.074)	(0.435)	(0.838)	(1.975)	(0.926)	(0.529)
$\sigma_{c,H,1}$	**21.489**	**7.209**	**21.528**	**19.814**	**16.305**	**21.147**
	(4.581)	(2.991)	(3.622)	(3.125)	(2.104)	(3.492)
$\sigma_{c,H,2}$	**21.870**	**49.845**	**21.082**	**19.013**	**15.595**	**21.107**
	(4.346)	(7.714)	(3.533)	(3.124)	(1.827)	(3.340)
$\sigma_{L,1}$	3.392	**4.939**	**6.207**	**5.154**	**4.893**	**6.226**
	(1.869)	(0.400)	(0.767)	(2.108)	(1.969)	(0.383)
$\sigma_{L,2}$	3.094	0.027	1.573	**6.867**	1.934	0.500
	(2.107)	(0.276)	(2.926)	(1.494)	(1.179)	(0.939)
$\sigma_{H,1}$	**8.924**	**14.726**	**8.220**	**6.079**	**65.067**	**53.889**
	(1.542)	(1.230)	(0.980)	(2.457)	(15.588)	(10.298)
$\sigma_{H,2}$	**13.742**	**12.239**	**9.593**	**66.802**	**7.210**	**9.640**
	(2.234)	(3.127)	(1.070)	(15.984)	(0.626)	(0.948)
Mean parameters						
$\theta_{1,L}$	-0.493	0.770	-0.042	-1.171	-0.027	0.314
	(0.478)	(0.567)	(0.469)	(0.651)	(0.108)	(0.489)
$\theta_{2,L}$	**-0.808**	0.016	-0.796	-0.374	-0.531	**0.996**
	(0.370)	(0.258)	(0.568)	(0.775)	(0.353)	(0.415)
$\theta_{1,H}$	**-3.651**	1.230	-16.574	4.406	-19.266	**-17.962**
	(1.180)	(4.477)	(16.365)	(13.723)	(23.540)	(8.071)
$\theta_{2,H}$	**-1.624**	0.111	-4.178	4.790	-4.806	**-4.891**
	(0.441)	(1.017)	(2.470)	(17.470)	(2.976)	(2.112)
λ_1	**0.917**	**1.171**	**1.627**	0.657	**1.523**	**1.674**
	(0.373)	(0.281)	(0.409)	(0.391)	(0.395)	(0.412)
λ_2	**0.730**	**1.662**	**1.051**	**1.378**	**0.883**	**1.313**
	(0.370)	(0.344)	(0.341)	(0.407)	(0.403)	(0.307)

Country 1 and Country 2 are listed to the left and right of the slash (/), respectively, and the shock is assumed to originate in Country 1. L and H denote the low- and high-volatility regimes. Standard deviations are reported in parentheses. Estimated parameters in **bold** are statistically significant using a two-sided 5% Student's t-test ($|t|$ >1.96).

Table 2.3 reports estimates of model parameters related to common shocks, idiosyncratic shocks, and common factors. The estimated impact coefficients of the idiosyncratic shocks are overall significant in both low- and high-volatility regimes. In the low-volatility regime, these coefficients ($\sigma_{L,1}$ and $\sigma_{L,2}$) range from 0.500 to 4.939 for Brazil, from 0.027 to 6.867 for Argentina, and from 1.934 to 6.207 for Mexico. They increase significantly in the high-volatility regime ($\sigma_{H,1}$ and $\sigma_{H,2}$) ranging from 8.924 to 14.726 for Brazil, from 12.239 to 66.802 for Argentina, and from 6.079 to 13.742 for Mexico. The estimated standard deviations indicate that, on average, Argentina's idiosyncratic volatility is significantly higher than that of other countries during crisis periods. This can be explained by the numerous political, fiscal, and debt crises that Argentina has faced, making its economy risky and increasing the volatility of bond spreads. Mexico's idiosyncratic volatility, on the other hand, appears to be lowest on average during periods of crisis. This can be explained by the fact that, in the aftermath of the peso crisis of 1994, the country implemented major structural reforms across all economic sectors, which contributed importantly to reducing uncertainties and risks in economic activity and financial markets.

The estimated impact coefficients of the common uncertainty shocks are similar across countries in the low-volatility regime ($\sigma_{c,L,1}$ and $\sigma_{c,L,2}$). In addition, while these estimates do increase significantly during crises ($\sigma_{c,H,1}$ and $\sigma_{c,H,2}$), their amplitudes remain similar across countries. This suggests that on average, the common volatility of bond spread returns is similar across countries in both normal and crisis periods over our study sample. Thus, our model appears able to adequately disentangle idiosyncratic shocks from common shocks.

2.3.2.2 Impacts of a change in the global factor on bond spread returns

Table 2.3 also shows that the estimated global factor loadings (λ_1 and λ_2) are generally positive and significant, with similar amplitudes across countries. This indicates that a change in the global factor significantly raises the spread of emerging sovereign bonds. It also suggests a strong interaction between sovereign bond risk and the global factor. Since the sovereign bonds of emerging countries are primarily issued in international financial markets, the cost of borrowing faced by these countries is highly dependent on changes in global financial conditions, world production, and world trade. Our model shows that the global factor significantly influences the value of emerging sovereign bonds. This result is related to two strands of the financial literature.

First, in terms of empirical research on the determinants of sovereign spreads, Uribe and Yue (2006), Akıncı (2013), and Gilchrist et al. (2019) examine the effect of changes in global factors, particularly U.S. interest rates and monetary policy, on sovereign spreads. Longstaff et al. (2011) and Ang and Longstaff (2013) decompose sovereign CDS (credit default swap) spreads into local and global components, highlighting the importance of global factors. Our results thus confirm the influence of global conditions, showing that this global factor accounts for a substantial proportion of the variation in sovereign bond spreads.

Second, our result is related to recent research that emphasizes the importance of the Global Financial Cycle, such as Rey (2015), Miranda-Agrippino and Rey (2020), and Gilchrist et al. (2022), as well as the combined influences of the Global Financial Cycle and the Global Trade-Commodity Cycle, as discussed in Miranda-Agrippino and Rey (2021). In particular, Rey (2015) links the VIX to a common factor in international asset prices, as the VIX is widely considered a measure of uncertainty and risk aversion in financial markets. Gilchrist et al. (2022) incorporate both the VIX and the Global Financial Cycle to examine their impact on sovereign default risk, beyond the effects of other global factors, including U.S. monetary shocks. Miranda-Agrippino and Rey (2021) show that the Global Financial Cycle and the Global Trade-Commodity Cycle co-move and influence international asset prices.

2.3.2.3 Idiosyncratic shock transmissions based on the MS-FACTOR model

When the idiosyncratic shock of the market generating the crisis significantly increases, we see that, except for two country pairs: Brazil/Mexico and Argentina/Brazil (where Brazil and Argentina are the source countries of the crises, respectively), the estimated loadings on the lagged local factors generally do not change significantly. This suggests that the interactions between the other country pairs over our sample duration are, on average, only interdependences. In other words, in periods of crises, large shocks originating from the source country market spill over to the target country market via the usual channels, similar to the way smaller shocks propagate.

While the above results are informative, we cannot draw definitive conclusions about the degree of contagion or decoupling across markets simply by comparing the changes in each local factor loading across volatility regimes. We formalize the analysis in the next section, where test outcomes are reported.

2.3.3 The bootstrap-based Student's t-test results

Table 2.4 reports the Student test p-values based on the bootstrap-based distribution. Evidence, at the 5% level, for the presence of decoupling is obtained for two of our country pairs, namely, Argentina/Mexico and Brazil/Mexico; their p-value* is less than 5%. For the other country pairs, we conclude that there is only interdependence, given the non-rejection of the null hypothesis.[8]

Table 2.4: Bootstrap t-Test Results and Crisis Transmission Type

Country pairs	Source	Target	t-stat	p-value*	Crisis transmission type
Argentina/Mexico	Argentina	Mexico	-3.08	**0.002**	Decoupling
Brazil/Mexico	Brazil	Mexico	-2.86	**0.004**	Decoupling
Argentina/Brazil	Argentina	Brazil	-0.650	0.506	Interdependence
Brazil/Argentina	Brazil	Argentina	1.750	0.078	Interdependence
Mexico/Brazil	Mexico	Brazil	-1.850	0.064	Interdependence
Mexico/Argentina	Mexico	Argentina	-1.860	0.060	Interdependence

This table reports the origin and destination countries of the crisis, the corresponding sample t-statistics, p-values, and the identified nature of crisis transmission. Statistically significant p-values are shown in bold.

To interpret these findings in light of the various crises that have occurred over our sample period, it is helpful to examine, for each country pair, the filtered probabilities of the high-volatility regimes for the idiosyncratic and the common shocks. Figures 2.1-2.3 and 2.4-2.6 below show these filtered probabilities for the Brazil/Mexico and Argentina/Mexico cases, respectively.

When examining the filtered probabilities of the high-volatility idiosyncratic shocks, we observe that many more high probability shocks have occurred in Argentina and Brazil compared to Mexico. For Argentina (see Figure 2.2), there are notable peaks at the end of 2001, early 2002, between late 2005 and early 2006, in 2013, at the end of 2019, and in 2020.

These peaks likely correspond to the following turbulent events: Argentina's sovereign debt default (end 2001), the country's currency crisis (between late 2005 and early 2006,

[8]The lack of evidence of contagion in the remaining four cases may also be explained by the possibility that the data contain more than two underlying regimes, which would limit the ability of a two-state framework to capture more complex dynamics.

reflecting economic challenges such as rising inflation, energy sector issues, and social unrest), the political crisis during Cristina Fernandez de Kirchner's presidency (2013), and Argentina's financial crisis in 2019, marked by sharp currency depreciation, soaring inflation, high interest rates, and severe economic contraction. Lastly, the 2020 peak likely corresponds to the impact of the Argentina defaulting on its debt in May 2020, exacerbating the country's ongoing economic challenges, specially after the Covid-19 lockdown in March of the same year.

For Brazil (see Figure 2.5), we observe probability peaks in the mid-1990s (early 1993 and late 1994), between 1995 and 1996, mid-1998, in 1999, between 2002 and 2006, in 2015, and between 2018 and 2019. These peaks likely correspond to the following turbulent events: the banking crisis (mid-1990s) triggered by the implementation of the Real Plan in 1994, which stabilized hyperinflation but disrupted the banking sector,[9] the impact of the Mexican peso on Brazil's currency (1995-1996), the effects of the Asian financial crisis (1997) and the Russian financial crisis (in mid-1998), Brazil's currency crisis (1999), and the economic and political challenges (2002-2006), including the energy crisis, financial instability, and political uncertainty. Additionally, Brazil's recession in 2015 is significant. Lastly, the 2018-2019 peaks likely correspond to socio-economic crises, including political uncertainty, high unemployment, and significant fiscal challenges affecting the country's growth and stability.

Finally, in the case of Mexico, we see high volatility shock probability peaks only between the years 1994 and 1995. These probability peaks likely correspond to the Mexico peso crisis (the Tequila crisis), which resulted from a sudden devaluation of the Mexican peso but which seems to have had persistent effects over the two years. All other estimated shocks for Mexico are low-volatility ones.

Figures 2.2-2.3 and 2.5-2.6 also confirm these results, decoupling for the bond market pairs of Argentina/Mexico and Brazil/Mexico respectively. We can see that, on average, not only has Mexico not generated major crises since the peso crisis of 1994, but it also has not been significantly affected by crises originating in neighboring countries. This is not surprising given that Mexico implemented several reforms to strengthen its economic and financial sectors after the 1994 crisis. Below, we outline Mexico's economic landscape since the mid-1990s.

[9]Prior to the plan, banks profited from high inflation through practices like "float", where they earned returns by holding deposits before processing payments. The abrupt decline in inflation post-implementation led to liquidity shortages, reduced profitability, and instability in the sector.

A) Estimated filtered probabilities of high volatility shocks; Argentina–Mexico

Figure 2.1: Common Shock

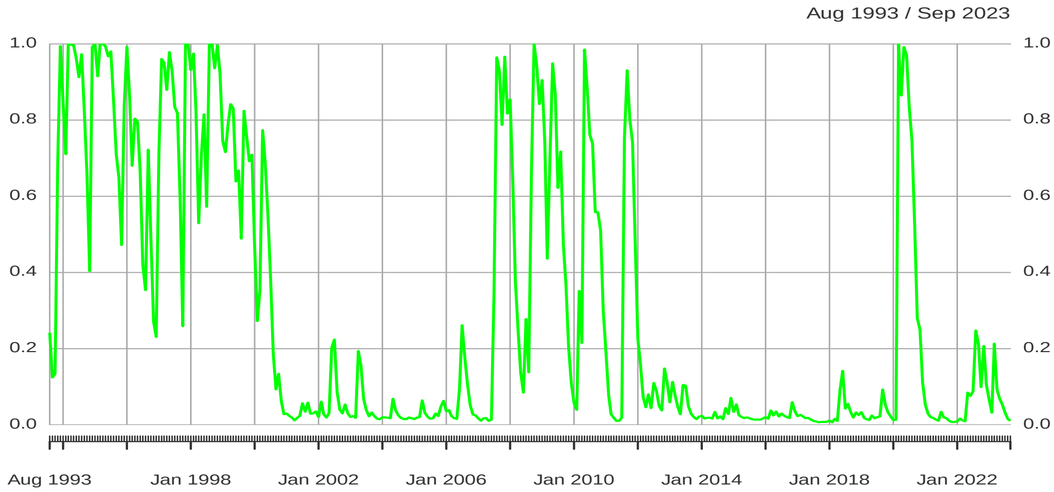


Figure 2.2: Idiosyncratic Shock, Argentina (Shock origin)

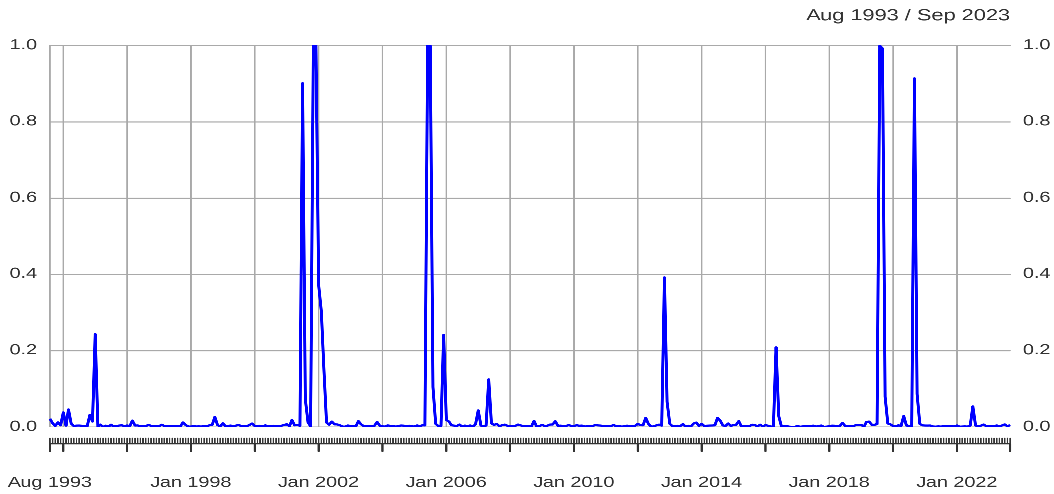
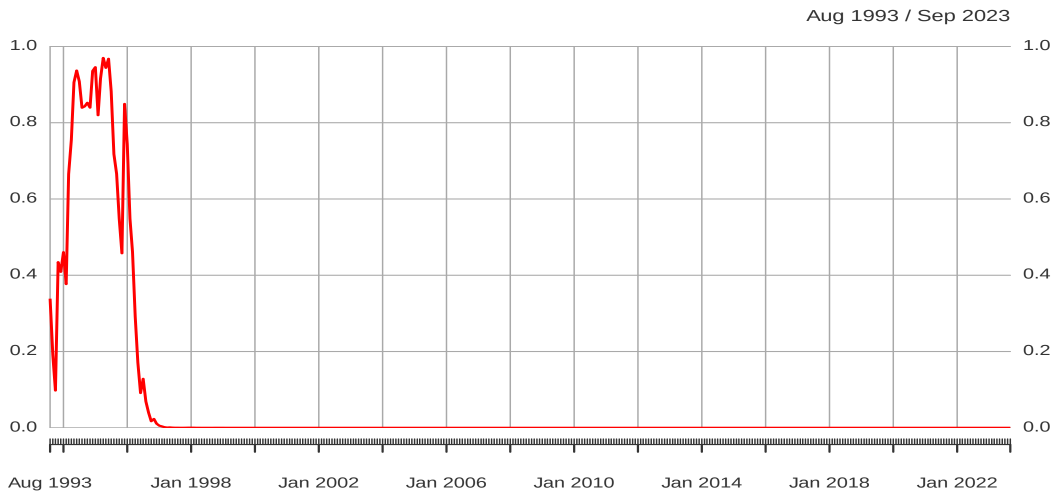


Figure 2.3: Idiosyncratic Shock, Mexico (Shock target)



B) Estimated filtered probabilities of high volatility shocks; Brazil–Mexico

Figure 2.4: Common Shock

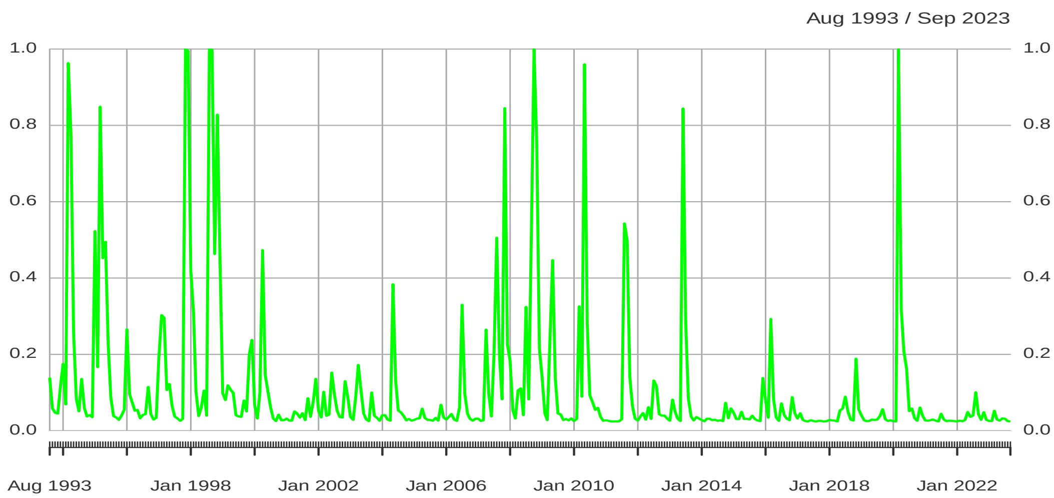


Figure 2.5: Idiosyncratic Shock, Brazil (Shock origin)

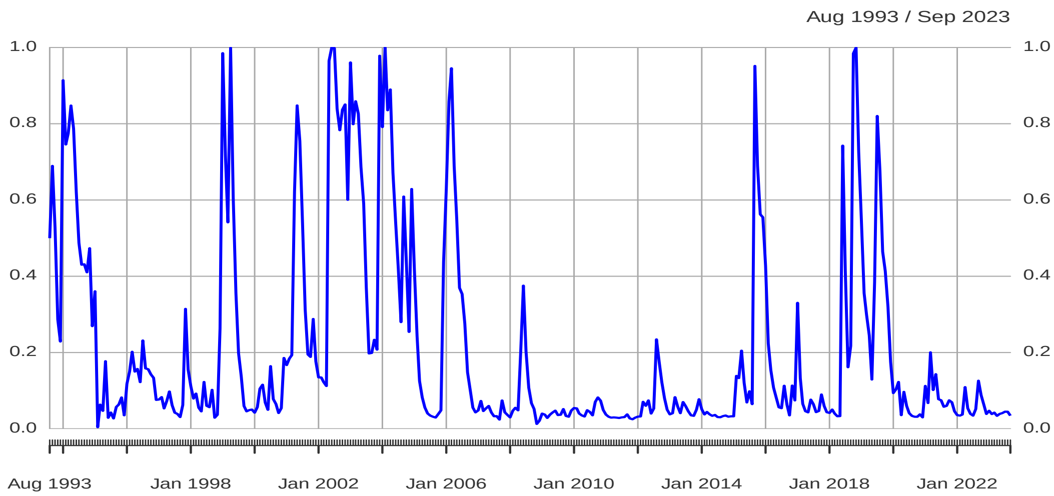
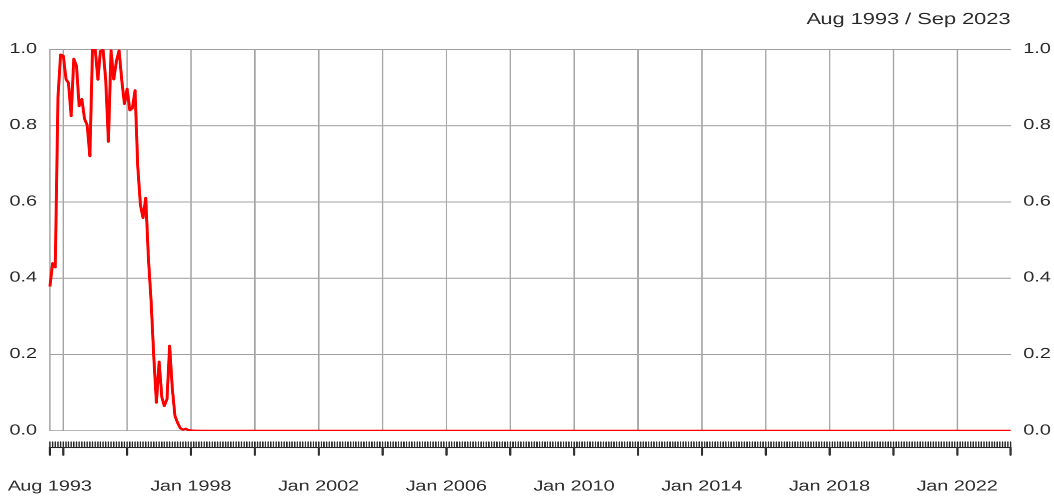


Figure 2.6: Idiosyncratic Shock, Mexico (Shock target)



First, Mexico's fiscal policy has become more robust. Currently, and despite a fiscal deficit and the potential economic risks associated with declining oil prices, the country is better prepared to manage shocks with tools such as an oil revenue stabilization fund. Moreover, Mexico holds a strong reputation in financial markets, supported by its low public debt level. Additionally, inflation remains moderate due to a credible monetary policy regime and the Mexican Central Bank's readiness to lower interest rates if necessary. Second, Mexico has been implementing a flexible exchange rate regime, which acts as a shock absorber by allowing currency depreciation to support exports during crises. Third, Mexico has greatly benefited from a diversified trade portfolio and maintains particularly strong trade relations with the United States and Canada compared to the rest of Latin America, including Brazil and Argentina.

When examining the filtered probabilities of high-volatility common shocks (Figures 2.1 and 2.4), we observe distinct spikes that align with well-documented episodes of global and regional financial crises. The most prominent common global crises appear during the periods of 1997, 1998, 2007–2009, 2011–2013, and February 2020, whereas the common regional shocks are concentrated around 1993–1995 and 1998–1999. These periods reflect major systemic disruptions that transmitted volatility across countries and markets, particularly affecting emerging economies through capital flows and risk premia adjustments.

Focusing first on global shocks, the spike in 1997 coincides with the Asian financial crisis, which began in Thailand after the collapse of the baht due to speculative attacks and unsustainable foreign debt. This turmoil rapidly spread across Asia and reached Latin America through global investor sentiment, causing sovereign bond spreads in countries like Brazil, Mexico, and Argentina to rise significantly (Kaminsky & Reinhart, 2000; Radelet et al., 1998). The 1998 spike aligns with the Russian financial crisis, where a sovereign default and currency collapse undermined confidence in all emerging markets, raising bond spreads despite sound fundamentals in Latin America (Odling-Smee, 2006). The 2007–2009 Global Financial Crisis, triggered by the collapse of Lehman Brothers, severely impacted Latin America—especially Mexico due to its economic integration with the U.S. (Claessens et al., 2010). The 2011–2013 period reflects the European Debt Crisis, where unsustainable debt levels in Greece, Portugal, and Spain created contagion effects that spilled over to Latin America via capital outflows and elevated sovereign risk premiums (Mink & De Haan, 2013). Finally, the spike in early 2020 corresponds to the onset of the Covid-19 pandemic, which abruptly halted global trade and capital flows, unleashing severe financial volatility (IMF, 2020; Zhang et al., 2020).

Turning to regional shocks, the first major peak around 1993–1995 corresponds to the Mexican peso crisis (Tequila Crisis), which was driven by large fiscal deficits, declining reserves, and political instability. The subsequent devaluation spilled over across Latin America through capital flight, currency depreciations, and increased bond spreads—particularly in Argentina and Brazil (Han et al., 2003; Sachs et al., 1996). The second peak, between 1998–1999, is associated with the Brazilian currency crisis, triggered by loss of investor confidence and unsustainable fiscal imbalances, leading to a sharp devaluation of the real. This event propagated regionally through bond markets as investors reassessed risk in other emerging markets with similar vulnerabilities (Kaminsky & Reinhart, 2000).

In summary, Figures 2.1 and 2.4 show broadly consistent outcomes with regard to the identification of the common shocks. While the exact estimated timing and magnitude of the peaks may vary across country pairs—due to differences in exposure, transmission mechanisms, and market-specific dynamics—the overall patterns validate the model's ability to capture systemic episodes of volatility that influence Latin American sovereign bond spreads.

2.3.4 Discussion

Our results are related to the literature on financial contagion in Latin American bond markets. Using weekly data over the sample period from January 1991, to September 2001, Gravelle et al. (2006) find evidence of shift-contagion having occurred (on average, over the sample period) in the case of the country pairs Brazil/Mexico and Venezuela/Mexico. However, as discussed above, this model misses the dimension of global effects impacting the mean processes of the returns – a feature that is by now strongly established in the empirical literature. In addition, the model attributes all possible contagion events between the asset pair returns to the shocks that are common to these markets, regardless of whether these originate within these markets or outside it, and without distinguishing between origin/destination of the spillover shock.

In a different study, Han et al. (2003) propose a methodology similar to that of Karafiath et al. (1991) to examine the spillover effects of the Mexican peso crisis to emerging debt markets. In this model, the start and end dates of the crisis are assumed to be known in advance, allowing the data to be separated into "normal" behavior during the tranquil period and "anomalous" behavior during the crisis. They find that the Latin American region was more exposed to the Mexican crisis than emerging markets from other regions. They also show that direct trade linkages were not a transmission channel of the crisis. This result is consistent with findings by Van Rijckeghem and Weder (2001) and Kaminsky and Reinhart (2000), who find that bilateral trade between emerging markets typically represents only a small share of total exports. Rather, the transmission channel of the Mexican crisis was through trade competition in third markets. According to Glick and Rose (1999), countries lose competitiveness when their trading partners devalue, leading to trade deficits and a gradual decline in the international reserves of their central banks, eventually, making these countries themselves targets of speculative attacks. Our results are also in line with these conclusions since we find that, on average, the local factor (which includes trade with Mexico) does not appear to be a significant transmission channel.

With higher frequency data, other contagion-factor models (with different modeling frameworks) have found contagion effects in other types of markets. Bekaert et al. (2014) employ a three-factor model to examine the contagion effects of the 2007-2008 Global Financial Crisis on country-industry portfolios. They find evidence of contagion through both the global financial factor and the domestic factor. Dungey and Renault (2018) extend this model by incorporating factors into the conditional mean that capture the joint conditional variance of returns. They also find evidence of contagion in various Asian currency markets during 1997-1998, in U.S. sectoral equity indices during 2007-2009, and in the CDS market during the European Debt Crisis of 2010-2013.

As for studies having focused on decoupling in financial markets, Dooley and Hutchison (2009) examine the transmission of the U.S. subprime financial crisis to emerging markets. They find that asset markets in emerging markets were able to decouple for several months from this shock due to several actions. In particular, these countries had hedged some of their currency positions, monitored and controlled foreign currency debt by non-financial firms and they had accumulated sufficient reserves.

Furthermore, Dungey et al. (2020) analyze the transmission of shocks between global banking, domestic banking, and the non-financial sector for eleven Eurozone countries. They propose a Markov-switching Factor Augmented VAR model to distinguish between contagion, interdependence, and decoupling. They find that global banking shocks created a contagion situation in the banking sectors of smaller states, exacerbating their crises, while non-financial sectors largely decoupled, as they secured alternative financing.

2.3.5 Out-of-sample forecasting evaluation

2.3.5.1 Rolling window forecasting

In this section, we compare the out-of-sample one-month-ahead forecasting ability of our model to that of the Gravelle et al. (2006) model. We proceed as in Chang et al. (2023) to apply this forecasting exercise.

In the model by Gravelle et al. (2006), the mean of each sovereign bond return is equal to its time-varying expected return, which depends on the state of the common shock. In contrast, in our model, the mean of each sovereign bond return is a function of both the global factor and the local factor, with a constant global factor loading and a time-varying local factor loading that depends on the state of the country-specific shock generating the crisis.

Let T_e be a fixed window size. Data from T_0 up to the time the forecast is made is used to generate one-month ahead forecasts as follows: if the first observation is at time $\ell_e = T_0$, then the first forecast uses the data from the time interval $[\ell_e, \ell_e + (T_e - 1)]$ to obtain an out-of-sample forecast return at time $\ell_e + T_e$ as follow:

In Gravelle et al. (2006):

$$\widehat{r}_{i,\ell_e+T_e} = \mathbf{E}[\widehat{\mu}_{i,\ell_e+T_e}], \quad i = 1, 2, \tag{2.15}$$

$$\mathbf{E}[\widehat{\mu}_{i,\ell_e+T_e}] = \widehat{\mu}_i^* P(A_{c,\ell_e+T_e} = 1) + \widehat{\mu}_i(1 - P(A_{c,\ell_e+T_e} = 1)), \tag{2.16}$$

where $P(A_{c,\ell_e+T_e} = 1)$ are the filtered probabilities of the high volatility regime of the common shock for the country pair.

In our model:

$$\widehat{r}_{i,\ell_e+T_e} = \mathbf{E}[\widehat{\theta}_{1,\ell_e+T_e}|\mathcal{F}_{\ell_e+T_e-1}] f^{(ij)}_{\ell_e+T_e-1} + \widehat{\lambda}_1 f^{(G)}_{\ell_e+T_e-1}, \tag{2.17}$$

$$\widehat{r}_{j,\ell_e+T_e} = \mathbf{E}[\widehat{\theta}_{2,\ell_e+T_e}|\mathcal{F}_{\ell_e+T_e-1}] f^{(ij)}_{\ell_e+T_e-1} + \widehat{\lambda}_2 f^{(G)}_{\ell_e+T_e-1}, \tag{2.18}$$

$$\mathbf{E}[\widehat{\theta}_{l,\ell_e+T_e}|\mathcal{F}_{\ell_e+T_e-1}] = \widehat{\theta}_{l,H}P(A_{i,\ell_e+T_e} = 1|\mathcal{F}_{\ell_e+T_e-1}) + \widehat{\theta}_{l,L}(1 - P(A_{i,\ell_e+T_e} = 1|\mathcal{F}_{\ell_e+T_e-1})), \quad (2.19)$$

where $l \in \{1,2\}$, and $P(A_{i,\ell_e+T_e} = 1|\mathcal{F}_{\ell_e+T_e-1})$ are the filtered probabilities of the high volatility regime of the country i shock generating the crisis. The local and global factor are estimated using data in the time interval $[\ell_e, \ell_e + (T_e - 1)]$. The second forecast uses the data from the time interval $[\ell_e + 1, \ell_e + T_e]$ to produce a forecast return at time $\ell_e + T_e + 1$ and so on. Finally, we compute the root mean squared errors (RMSE) of the two models.

2.3.5.2 Forecast results

We perform one-month ahead rolling-window forecasts, setting the rolling-window sample sizes to 344 months.[10]

Table 2.5 reports the out-of-sample root-mean-squared errors (RMSE) for each country for the two models. We can see that, except in one case, the MS-FACTOR model outperforms the Gravelle et al. (2006) model handily. This demonstrates the importance of not ignoring interactions at the mean level among the bond returns and validates our proposed extensions to the original model.

Table 2.5: RMSE in % Across Models

MS-FACTOR model					
Mexico	Brazil	Mexico	Argentina	Brazil	Argentina
2.750	**2.580**	**3.550**	**21.000**	**2.210**	22.970
Gravelle et al. (2006) model					
Mexico	Brazil	Mexico	Argentina	Brazil	Argentina
5.970	5.140	5.000	22.660	4.080	**22.670**

This table reports the root mean squared errors (RMSE) for each country pair. For the MS-FACTOR model, the reported results correspond to cases where the shock originates in: (1) Mexico, for the Mexico/Brazil and Mexico/Argentina pairs; and (2) Brazil, for the Brazil/Argentina pair. Results remain qualitatively similar when the shock originates in the other market of each pair. The winner RMSE are in bold.

[10] This (approximately) 26 year window size minimized numerical/convergence problems when estimating the model over each window.

2.4 Conclusion

We propose a modeling strategy for investigating the nature of the transmission of shocks between two asset markets during periods of crises. Our regime-switching model builds on the specification of Gravelle et al. (2006), and which we modify in several dimensions. First, we allow for co-movements in the mean processes of the asset returns, complementing co-movements in the volatility structure of these returns. The variables included in the means are PCA factors that capture both global and market-specific (local) impacts on these returns. Second, while we also separate high-volatility common shocks from high-volatility idiosyncratic shocks, we assume that it is only the latter that can generate contagion/decoupling, and that these occur when the interdependence structure of the means is altered significantly. We propose a bootstrap Student's t-test to test for contagion or decoupling. Applications to three pairs of Latin American countries reveal evidence, at the 5% level, of decoupling having occurred (on average, over our sample) in the bond market pairs of Brazil/Mexico and Argentina/Mexico. We conclude that in the rest of the bond market pairs, there is on average, only interdependence.

For future research, possible extensions include: (1) allowing for time-variation also in the loadings of the global factors of the mean processes of the returns, and possibly letting these switch with the state of the common shocks; (2) considering multiple asset markets, rather than focusing on a pair of markets each time; (3) exploiting higher frequency data to test for contagion or decoupling over individual periods of crises.

2.5 Appendix: Estimates of Gravelle et al. (2006) model and robustness check

2.5.1 Use of AI in research

In this study, OpenAI ChatGPT and Grammarly were used to check the grammar, catch typos, and enhance the writing to improve the clarity of the text.

2.5.2 Estimates of Gravelle et al. (2006) model

Table 2.6 reports estimates from the Gravelle et al. (2006) model. σ_{c1} and σ_{c2} represent the impact coefficients of the common shock in the low-volatility regime for countries 1 and 2, respectively, while σ^*_{c1} and σ^*_{c2} denote the corresponding coefficients in the high-volatility regime. Similarly, σ_1 and σ_2 capture the impact coefficients of idiosyncratic shocks 1 and 2 in the low-volatility regime, with σ_1^* and σ_2^* representing these coefficients in the high-volatility regime. Finally, μ_1 and μ_2 (μ_1^* and μ_2^*) refer to the estimated means for countries 1 and 2 during the low-volatility (high-volatility) regime of the common shock respectively.

Table 2.6: Estimated Parameters of Gravelle et al. (2006) Model

Parameters	Brazil/Argentina	Mexico/Brazil	**Mexico/Argentina**
Variance parameters			
σ_{c1}	**5.543** (0.484)	**5.751** (0.579)	**4.844** (0.475)
σ_{c2}	**9.572** (0.460)	**6.906** (0.482)	**6.813** (0.762)
σ_{c1}^*	**5.543** (0.484)	**17.538** (2.722)	**21.623** (2.904)
σ_{c2}^*	**58.231** (9.410)	**16.493** (2.910)	**21.915** (3.195)
σ_1	**4.788** (0.342)	**4.208** (0.350)	0.155 (0.518)
σ_2	**0.003** (0.001)	**1.012** (0.143)	**6.057** (0.554)
σ_1^*	**14.113** (1.168)	**13.455** (2.185)	**7.078** (0.527)
σ_2^*	**9.836** (3.140)	**7.855** (1.713)	**67.847** (15.955)
Mean parameters			
μ_1	**-1.552** (0.464)	**-2.041** (0.484)	**-1.135** (0.423)
μ_2	-0.459 (0.518)	**-2.595** (0.522)	-0.518 (0.597)
μ_1^*	10.199 (13.542)	**13.004** (5.048)	6.314 (4.560)
μ_2^*	**6.867** (2.828)	**11.796** (5.311)	6.365 (4.838)

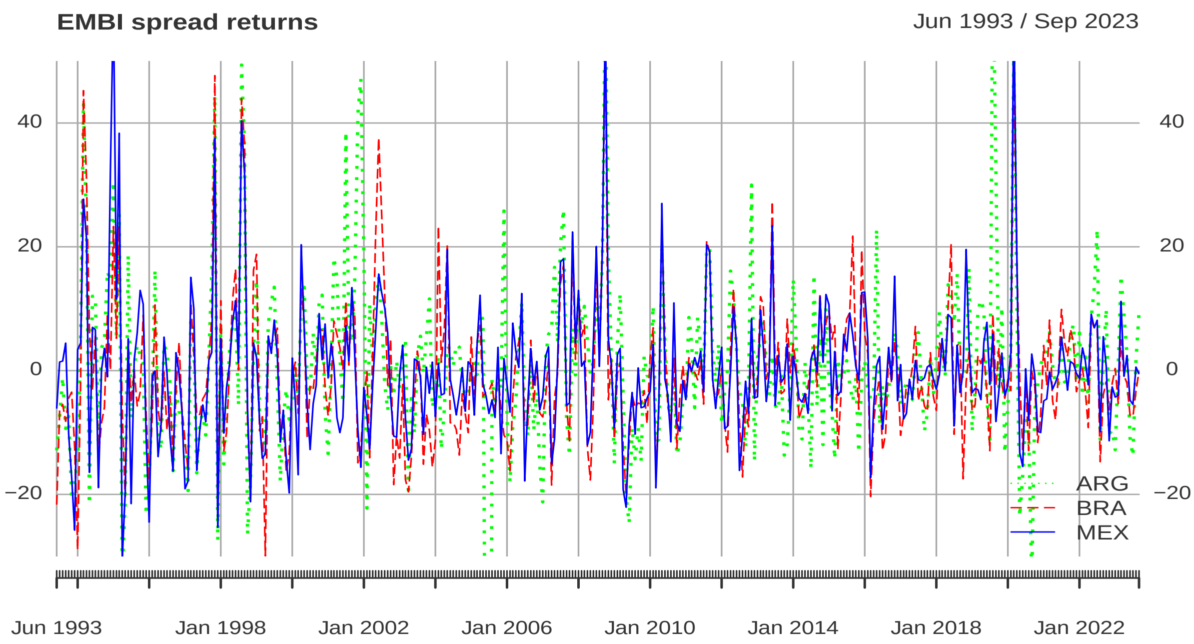
Subscripts 1 and 2 refer to country 1 and country 2, respectively. σ_{c1} and σ_{c2} are the impacts of the common shock in the low-volatility regime, while σ_{c1}^* and σ_{c2}^* represent the impacts in the high-volatility regime. Likewise, σ_1 and σ_2 measure the impacts of idiosyncratic shocks in the low volatility, and σ_1^* and σ_2^* in the high volatility. μ_1 and μ_2 denote the mean returns in the low-volatility regime, with μ_1^* and μ_2^* representing the corresponding means in the high-volatility regime. Standard errors are in parentheses. Estimated parameters in **bold** are statistically significant using a two-sided 5% Student's t-test ($|t| > 1.96$).

When comparing the results across country pairs, we find that the estimated impact coefficients of both common and idiosyncratic shocks increase significantly in the high-volatility regime for all cases. In addition, we observe notable differences in the estimated mean values. Compared to Table 2.3, we find that in the low-volatility regime, the estimated amplitudes of both common and idiosyncratic shocks in Gravelle et al. (2006) are similar to those produced by our model. In contrast, in the high-volatility regime, the idiosyncratic shock amplitudes reported by Gravelle et al. (2006) are higher than those in our results, whereas the estimated amplitudes of the common shocks in our model are generally larger than theirs. This discrepancy may be explained by the inclusion of dynamic factors in the mean equation, which helps to more accurately capture the underlying error structures.

2.5.3 Stationarity of sovereign bond spread returns

The Figure 2.7 illustrates the spread EMBI+ returns of Argentina, Brazil, and Mexico between June 1993 and September 2023 and we can observe considerable co-movement in these returns.

Figure 2.7: Stationary sovereign bond spread returns



2.5.4 Robustness check

In this section, we analyze the robustness of our model to the inclusion of the second largest global factor in the means of the returns. Together the first and second factors capture 85% of the total variance of the global features.[11]

[11]Note that increasing the number of local factors, or the number of the global and local factors simultaneously, causes convergence problems. Indeed, convergence is numerically infeasible in several of these cases.

2.5.4.1 Estimated MS-FACTOR model with additional global factor

Table 2.7: Estimated parameters of MS-FACTOR Model with Additional Global Factor

Parameters	Brazil/Mexico	Brazil/Argentina	Mexico/Brazil	Mexico/Argentina	Argentina/Mexico	Argentina/Brazil
Variance parameters						
$\sigma_{c,L,1}$	**6.682** (1.290)	**4.237** (0.518)	5.757 (7.360)	**5.929** (0.704)	**7.003** (1.347)	**5.685** (0.650)
$\sigma_{c,L,2}$	**6.944** (1.313)	**8.726** (0.479)	6.610 (8.713)	**3.785** (0.741)	**2.915** (0.807)	**6.065** (0.599)
$\sigma_{c,H,1}$	**21.740** (4.551)	**21.208** (3.450)	**20.107** (2.956)	**16.536** (1.697)	**16.202** (1.182)	**20.808** (3.407)
$\sigma_{c,H,2}$	**22.300** (4.395)	**22.254** (3.378)	**17.210** (5.568)	**15.213** (1.969)	**15.419** (1.561)	**20.842** (3.304)
$\sigma_{L,1}$	3.679 (2.153)	**4.512** (0.345)	3.879 (11.802)	0.071 (1.069)	**4.956** (1.759)	**6.205** (0.412)
$\sigma_{L,2}$	2.879 (2.921)	0.045 (0.094)	2.563 (21.492)	**7.539** (0.385)	1.670 (1.124)	0.536 (2.218)
$\sigma_{H,1}$	**9.358** (1.775)	**10.581** (0.789)	3.879 (11.801)	**8.345** (1.856)	**65.996** (16.218)	**53.570** (10.199)
$\sigma_{H,2}$	**13.620** (2.273)	**57.947** (12.157)	**10.408** (3.988)	**66.880** (16.044)	**7.125** (0.566)	**9.700** (0.968)
Mean parameters						
$\theta_{1,L}$	-0.450 (0.462)	0.439 (0.436)	-0.874 (1.794)	**-1.537** (0.715)	-0.205 (0.374)	0.302 (1.092)
$\theta_{2,L}$	-0.653 (0.486)	-0.689 (0.620)	-1.198 (1.383)	-0.998 (0.754)	**-0.879** (0.384)	0.995 (0.651)
$\theta_{1,H}$	-4.109 (2.565)	0.934 (0.779)	0.984 (3.230)	0.940 (2.787)	-30.322 (31.252)	-18.091 (16.711)
$\theta_{2,H}$	-1.547 (1.505)	0.755 (0.634)	3.370 (3.253)	2.175 (2.796)	-4.914 (3.501)	**-5.197** (2.613)
λ_1	**0.905** (0.350)	**1.205** (0.280)	**0.806** (0.305)	0.636 (0.382)	**1.472** (0.416)	**1.678** (0.418)
λ_2	0.232 (0.332)	**-0.182** (0.260)	**0.655** (0.276)	0.645 (0.508)	-0.383 (0.371)	-0.282 (0.627)
λ_3	**0.731** (0.344)	**1.452** (0.356)	**1.192** (0.421)	**1.316** (0.421)	0.778 (0.398)	**1.313** (0.311)
λ_4	**0.646** (0.327)	-0.038 (0.088)	0.023 (0.213)	-0.193 (0.671)	0.372 (0.356)	-0.328 (0.452)

Country 1 and Country 2 are listed to the left and right of the slash (/), respectively, and the shock is assumed to originate in Country 1. Standard deviations appear in parentheses. Estimated parameters in **bold** are statistically significant using a two-sided 5% Student's t-test ($|t|$ >1.96). The coefficients λ_1 and λ_2 denote the loadings of global factors 1 and 2 for Country 1, while λ_3 and λ_4 denote the corresponding loadings for Country 2.

Table 2.7 presents the model estimates. It shows that when an additional global factor is included in the conditional mean of each return, the impact coefficients of idiosyncratic shocks for Brazil and Mexico in the low-volatility regime become insignificant, while those for Argentina remain significant overall. In the high-volatility regime, however, the corresponding estimates ($\sigma_{H,1}$ and $\sigma_{H,2}$) increase significantly. Interestingly, when compared to our proposed model, their amplitudes decrease for Brazil and increase for Argentina. For example, $\sigma_{H,1}$ and $\sigma_{H,2}$ range from 9.700 to 10.581 for Brazil and from 53.570 to 66.880 for Argentina. For Mexico, the amplitudes of the high-volatility idiosyncratic standard deviations remain similar to those in our model.

In addition, the impact coefficients of the unexpected common shocks are significant in the low-volatility regime ($\sigma_{c,L,1}$ and $\sigma_{c,L,2}$) across countries and they are similar to those obtained from our proposed model in the same regime. During crises, the corresponding estimates increase significantly ($\sigma_{c,H,1}$ and $\sigma_{c,H,2}$), with amplitudes that remain similar across countries and in line with those from our model. Moreover, the estimated local factors remain insignificant, and the estimated global factor loadings also become insignificant.

2.5.4.2 Bootstrap t-test based on MS-FACTOR model with additional global factor

Table 2.8: Bootstrap t-Test Results and Crisis Transmission Type based on MS-FACTOR Model with Additional Global Factor

Country pairs	Source	Target	t-stat	p-value*	Crisis transmission type
Argentina/Mexico	Argentina	Mexico	-2.83	**0.004**	Decoupling
Brazil/Mexico	Brazil	Mexico	-1.93	0.054	Interdependence
Argentina/Brazil	Argentina	Brazil	-0.522	0.600	Interdependence
Brazil/Argentina	Brazil	Argentina	1.809	0.070	Interdependence
Mexico/Brazil	Mexico	Brazil	-1.792	0.072	Interdependence
Mexico/Argentina	Mexico	Argentina	-1.835	0.066	Interdependence

This table reports the origin and destination countries of the crisis, the corresponding sample t-statistics, p-values, and the identified nature of crisis transmission. Statistically significant p-values are shown in bold.

C) Estimated filtered probabilities of high volatility shocks; Argentina–Mexico

Figure 2.8: Common Shock

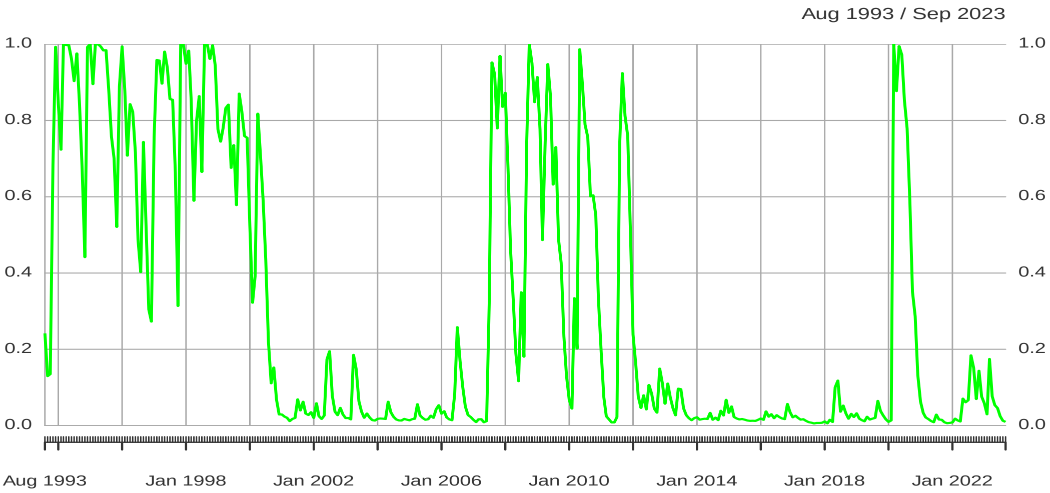


Figure 2.9: Idiosyncratic Shock, Argentina (Shock origin)

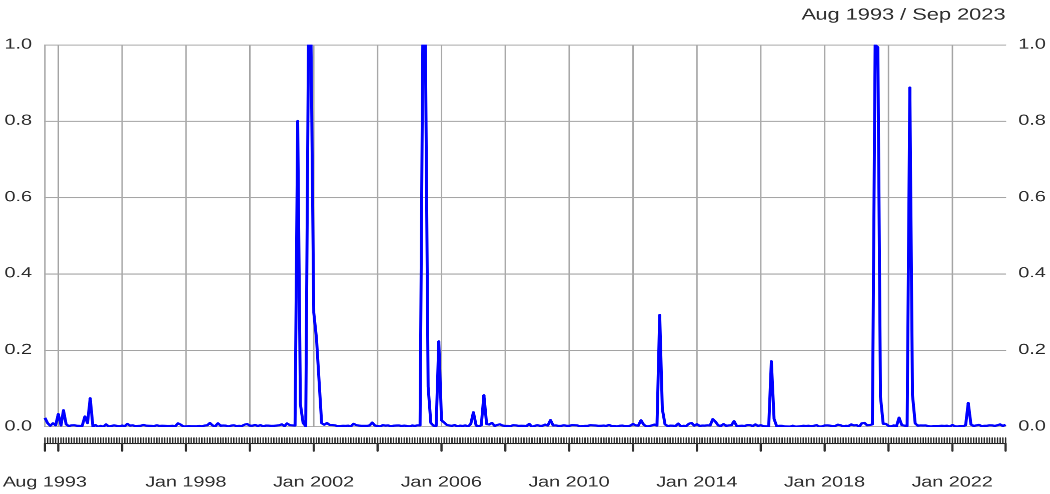
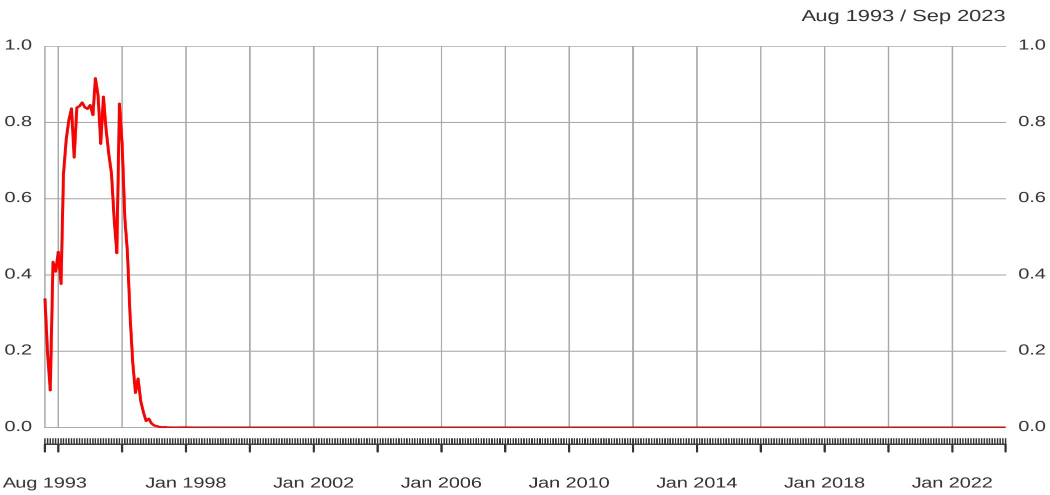


Figure 2.10: Idiosyncratic Shock, Mexico (Shock target)



D) Estimated filtered probabilities of high volatility shocks; Brazil–Mexico

Figure 2.11: Common Shock

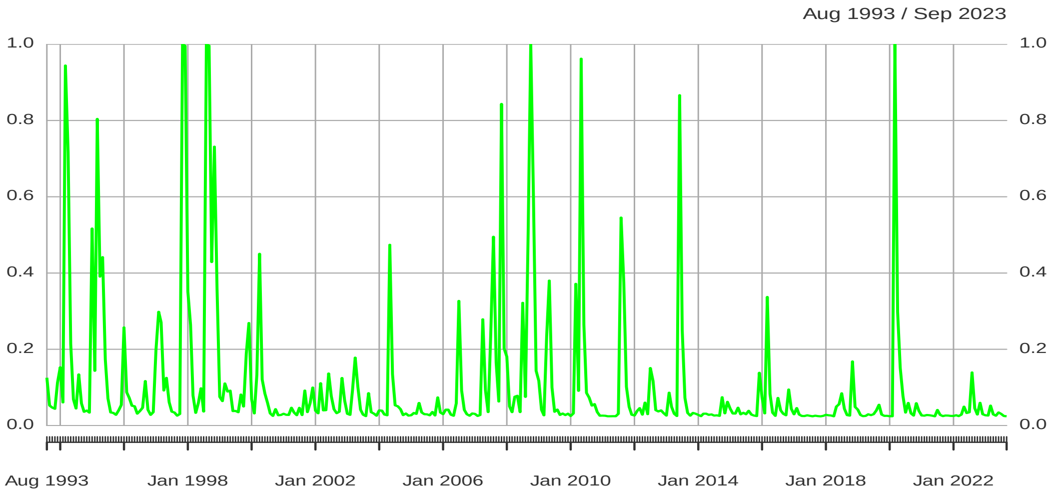


Figure 2.12: Idiosyncratic Shock, Brazil (Shock origin)

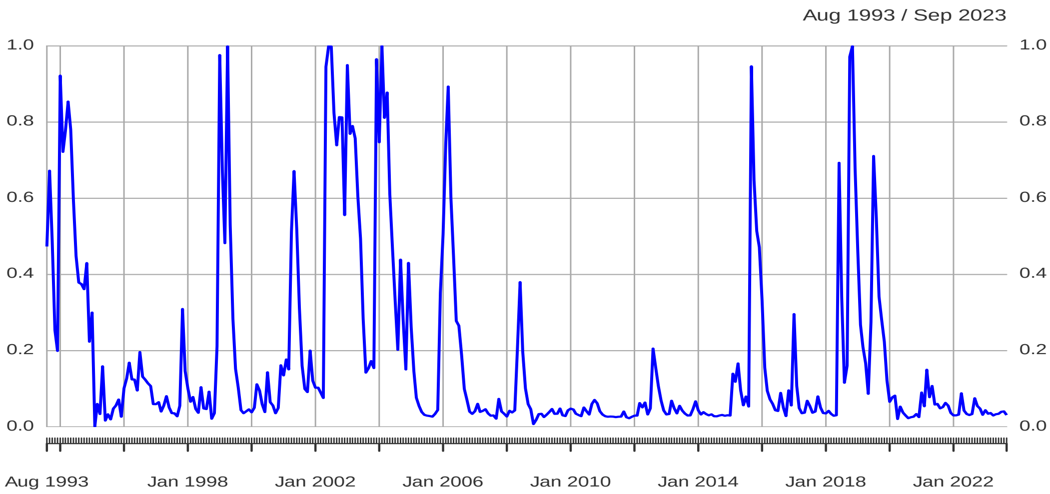


Figure 2.13: Idiosyncratic Shock, Mexico (Shock target)

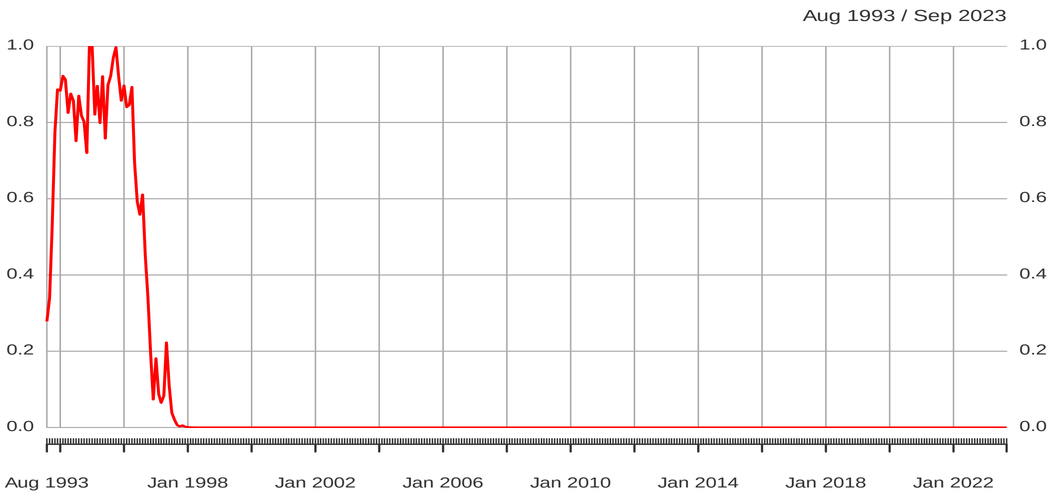


Table 2.8 reports the p-values from the Student's t-test based on the bootstrap distribution. At the 5% significance level, there is evidence of decoupling only for the Argentina/Mexico pair, where Argentina is the crisis-originating country. For the other country pairs, the results indicate interdependence on average. The estimated filtered shock probabilities in the high-volatility regime—both for common and idiosyncratic shocks—for the Argentina/Mexico and Brazil/Mexico pairs (Figures 2.8–2.13) show no substantial changes when an additional global factor is included, compared to the baseline model outcomes (Figures 2.1-2.6). On the whole, these results confirm the conclusions we drew based on our baseline model specifications.

Chapter 3

Commodity–Equity Linkages: A Markov-Switching Contagion Framework

3.1 Introduction

Despite the expanding literature on financial spillovers across asset markets, the link between commodity prices and equity market returns remains underexplored, especially the connection between a country's stock market and the world price of its main export commodity (Pinto-Ávalos et al., 2024). Most existing studies adopt a global or panel-based perspective, analyzing average dynamics across countries or asset classes (Basher et al., 2012; Ready, 2018). Such aggregation can mask important country-specific transmission channels, particularly in commodity-dependent economies whose macroeconomic performance and corporate valuations are tightly coupled to movements in their key export commodities. This paper examines country-specific such spillovers, aiming to fill this gap.

Focusing on individual countries offers practical, policy-related, and theoretical benefits. For investors, a granular view of how shocks to a country's key export commodity feed into that country's equity market may aid with achieving a better portfolio allocation, informing hedging decisions, and sharpening the measurement of country-specific risk premia. For policymakers, tracing this link can clarify how swings in global commodity prices can tighten or loosen domestic financial conditions (notably, credit growth and capital flows). Such information is crucial when designing stabilization policies, including the provision of precautionary funds and the setting of counter-cyclical fiscal rules or macro-prudential measures that shield the economy from commodity-driven volatility. On the theoretical side, examining how export-price shocks transmit to equity returns can deepen our understanding of the channels—trade, income, and cross-border finance—through which global disturbances affect open economies (Arezki et al., 2017).

Commodity sectors are also likely to embed forward-looking information about the domestic business cycle. Raw materials such as crude oil, copper, and iron-ore are both, critical inputs, and major revenue sources for listed firms, and their price fluctuations thus affect corporate cash flows, consumer demand, and, ultimately, stock valuations (Liu et al., 2019). Rising oil prices, for example, will boost revenues (and share prices) for upstream producers while squeezing margins for energy-intensive manufacturers. Empirically, metals and energy benchmarks have been shown to anticipate shifts in economic activity (Cheng & Xiong, 2014). Moreover, oil-demand surprises influence business sentiment (Hamilton, 2009) and correlate positively with equity returns (Kilian & Park, 2009).

Concrete country cases underline these mechanisms. Among developed exporters, Canada's S&P/TSX rises and falls with world crude-oil prices, mirroring the energy sector's large share in gross domestic product (GDP) and market capitalization; Norway's Oslo Børs shows a similar oil sensitivity, while iron-ore booms and busts ripple through Australia's mining-heavy ASX. In the emerging-market sphere, Brazil's B3 is tightly linked to iron-ore and soybean prices, and Chile's IPSA tracks copper cycles that also affect fiscal revenues. These examples demonstrate how commodity shocks can simultaneously reshape national income, external balances, and stock-market valuations (Connolly & Orsmond, 2011; Hasan & Mahbobi, 2013; Iturrieta & Federowski, 2017; Samora & McGeever, 2021).

Finally, the commodity–equity nexus is justified by the observed heterogeneity in pricing behavior across different international commodity markets: different commodities display idiosyncratic price dynamics across time and market segments. Daskalaki et al. (2014) show that many commodity markets are segmented, with pricing patterns largely independent of one another, while Erb and Harvey (2006) and Kat and Oomen (2006) document strong temporal variation in individual commodity behavior. A country's primary export commodity may therefore exert a unique influence on its equity market. Unpacking these commodity-specific linkages is essential for determining whether observed co-movements represent genuine contagion, shared exposure to global risk factors, or regime-dependent interactions—an analytical frontier that this study seeks to advance.

3.1.1 Related literature

Previous empirical studies have reported co-movements between commodity and equity returns, as well as significant spillover effects—often characterized as contagion—from commodity price shocks to equity market performance. Some of these notably employ a Dynamic Conditional Correlation Generalized Autoregressive Conditional Heteroskedasticity (DCC-GARCH) approach. This method captures the co-movement between commodity and equity returns using the estimated time-varying conditional correlations between asset returns while allowing for conditional heteroskedasticity or volatility.

For developed economies, Creti et al. (2013) analyze data from 25 commodity markets and S&P 500 returns from 2001 to 2011 and find significant co-movement between equity and commodity returns which tends to increase during the 2007-2009 Global Financial Crisis (GFC). Similar results have been observed in emerging markets as well. Sadorsky (2014) investigates the relationship between an aggregate emerging market equity index and a set of commodity markets including oil, copper, and wheat, from 2000 to 2012. He finds that the co-movement between stock and commodity market returns increases significantly during the GFC. Additionally, Roy and Roy (2017) examine the co-movement between commodity future price index, the Indian foreign exchange rate, and the Indian equity market. They find that the correlation between commodity future price index and equity market return increases significantly during the GFC, which they conclude to be evidence of contagion between these markets.

Another family of empirical research uses a Vector Autoregressive (VAR) approach, first introduced by Diebold and Yilmaz (2009) and later revised in Diebold and Yilmaz (2012), to examine the spillovers between commodity market prices and equity returns.[1] For developed countries, Bouri et al. (2021) analyze spillovers between S&P 500 returns and oil and gold markets using high-frequency data. They report evidence of contagion between these markets, as they find an intensification of the relationships between them during periods of financial distress. Similar results are observed in emerging markets. Ahmed and Huo (2021) investigate the dynamic relationship between Chinese stock returns and a range of commodity markets (oil, gold, industrial metals, and agricultural commodities), and they report finding evidence of

[1] Diebold and Yilmaz (2012) extend their earlier framework by introducing a generalized spillover index, which quantifies how much volatility in one asset or market is attributable to shocks from others, using a forecast error variance decomposition that is robust to variable ordering.

contagion among these markets. Dey and Sampath (2020) explore spillovers between the gold price and the real estate and banking sectors in India, also concluding that the relationship between these markets strengthens during financial distress.

However, although the two families of studies described above conclude that the effects found constitute contagion, they often neglect to control for global risk factors that, as explained below, can likewise drive the co-movement of almost all financial asset returns (see Miranda-Agrippino and Rey (2021), for a recent discussion). In contrast to the findings of earlier studies, more recent research incorporating the influence of global risk factors presents mixed results. For example, Silvennoinen and Thorp (2013) estimate the correlation of a set of commodity markets with United States (U.S.) stock returns using a smooth transition DCC-GARCH model, with VIX included as a threshold variable to capture the effect of global risk aversion. They find that during the global financial crisis, while the correlation between stock returns and oil returns significantly increases, that between metal prices and stock returns increases modestly, whereas no correlation is observed between food prices and stock returns. Zhang and Hamori (2021), using the Diebold and Yilmaz (2012) approach, find that commodity markets, particularly oil, played a less prominent role in transmitting shocks to U.S., Japanese, and German stock returns during the Covid-19 pandemic. They argue that spillovers between these stock markets are mainly explained by financial shocks, with a systemic factor related to global equity market volatility, driven by infectious diseases, playing a role.

Pinto-Ávalos et al. (2024) revisit commodity–equity linkages using both a multivariate DCC-GARCH model and the Diebold and Yilmaz (2012) spillover index. Their sample spans 2000–2023 and covers nine major commodity-exporting countries—advanced and emerging alike—allowing them to track dynamics across three key crisis episodes: the global financial crisis, the European Sovereign Debt Crisis (ESDC), and the Covid-19 pandemic. They document sharp increases in time-varying conditional correlations during each of these periods. Yet, once global risk proxies such as the VIX or investor-sentiment measures are included, the apparent surges in co-movement no longer signal commodity-driven contagion.

3.1.2 Motivation

The existing approaches that control for global risk factors when examining contagion from commodity returns to equity returns suffer from important limitations. The multivariate

DCC-GARCH and its smooth-transition variant can track time-varying correlations between two asset returns, yet they cannot pinpoint where the shocks driving those correlations originate. As discussed in Pinto-Ávalos et al. (2024) who analyze commodity–equity market co-movements for several exporting countries, these models can inform on the presence of excess co-movement, but not on where the spark for it originated.

Additionally, the Diebold and Yilmaz (2012) approach only captures linear effects between market returns. It summarizes average return interactions and therefore overlooks the possibility that financial markets oscillate between distinct regimes—bull versus bear phases, boom versus recession, or high- versus low-sentiment periods—each with its own patterns of returns, volatility, and investor behavior.

Moreover, in the multivariate DCC-GARCH and Diebold and Yilmaz (2012) frameworks, break tests are typically imposed around *ex ante* start- and end-dates for the crisis episodes of interest. If those pre-selected windows do not align with the crises' true timing, the resulting inferences can be misleading (Gravelle et al. (2006)). Do et al. (2018), for instance, demonstrate that pinpointing the onset of the 2007–2009 GFC is challenging because stress accumulated gradually rather than erupting on a single, clear-cut date. Likewise, Lane (2012) shows that the ESDC emerged from a mix of pre-existing fiscal vulnerabilities and external shocks, complicating efforts to specify an exact start.

Finally, since break tests are applied, the multivariate DCC-GARCH model is likely to pick up only quick and abrupt contagion episodes, not more gradual ones, while the smooth transition DCC-GARCH model is likely to pick up only gradual contagion, not more quick and abrupt ones. Campos-Martins and Amado (2022) demonstrate the simultaneous existence of short-term and long-term contagion effects and stress the importance of being able to distinguish between these types.

3.1.3 Contributions

This study re-examines how country-specific commodity shocks influence domestic equity returns in major commodity-exporting economies, while explicitly conditioning on global risk. To capture this transmission mechanism, we build a two-regime Markov-switching model that overcomes some of the important limitations of earlier empirical approaches. Commodity and equity returns for each country are estimated jointly so that co-movements are reflected not only in expected returns but also in the full variance–covariance matrix. Three independent Markov chains underlie that matrix: one for common shocks, the latter proxying unobserved episodes of global turbulence (for example, surprise geopolitical events or abrupt shifts in U.S. monetary policy), and two others for the idiosyncratic shocks that are specific to each country's commodity–equity pair. These chains switch between low-volatility (normal) and high-volatility (crisis) regimes, and the regime probabilities are inferred from the data rather than imposed *ex ante*. The framework can therefore detect both sudden spikes and slow-moving build-ups in stress.

Equity returns are allowed to depend on their own one-period lag, but the strength of this dependence—captured by the autoregressive coefficient—is allowed to vary with the volatility regime of the idiosyncratic commodity shock. Specifically, when the commodity market enters a high-volatility state, the autoregressive coefficient shifts from its low-regime value to a high-regime value. This state-dependent structure enables the model to identify spillovers by testing whether the dynamics of equity returns, particularly their persistence, changes in response to stress originating in the commodity market. If the autoregressive coefficients in the low- and high-volatility regimes differ, it indicates that the equity market's internal adjustment mechanism is sensitive to external commodity-specific turbulence, providing direct evidence of regime-contingent spillovers from commodities to equities. We define contagion as a situation in which an idiosyncratic commodity shock moves into its high-volatility regime and, simultaneously, the loading of the lagged equity return shifts in a way that amplifies the impact of this shock on the mean process. A likelihood-ratio test provides a formal gauge of whether this loading shift is statistically significant.[2]

We apply the model to daily data from six major commodity-exporting countries—Canada, Norway, the U.S., Mexico, Chile, and Russia—covering the period 2006 to 2024.[3] The sample

[2]Our definition of contagion—an amplification in the *mean* equation that occurs when an idiosyncratic commodity shock switches to its high-volatility state—fits the objective of this paper, which is to trace how first-moment spillovers alter expected equity returns. Portfolio managers, however, also worry about contagion operating through higher-order co-moments such as co-skewness, co-kurtosis, and co-volatility. Tests that target those channels already exist (see, e.g., Fry-McKibbin and Hsiao (2018); Fry et al. (2010); Bae et al. (2003)), and can be embedded in a Markov-switching framework by allowing the state variable to govern not only the idiosyncratic volatility parameters but also regime-dependent higher-moment loadings (e.g., a state-contingent co-skewness coefficient or a co-kurtosis term that rises in the high regime). Extending the model in that direction would permit formal "higher-moment contagion" tests while preserving the intuitive interpretation of regimes that underpins the current specification. However, implementing such an extension in a full multivariate Markov-switching framework remains computationally demanding, especially when estimating joint state-dependent dynamics in higher moments alongside recursive likelihood evaluation. For this reason, we focus on first-moment contagion in this paper, leaving higher-moment extensions for future research.

[3]In future versions of this paper, we will consider a larger set of countries for our analysis.

is fairly representative as it includes high-income countries (Canada, U.S., Norway) as well as middle-income and emerging economies (Mexico, Chile, Russia), which allows us to analyze structural differences in policy capacity, market depth, and vulnerability. Data availability and reliability are also strong in our chosen set, as these countries provide relatively long, high-frequency time series for both commodity prices and equity indices, enabling robust econometric analysis. In terms of policy and systemic variation, the sample includes countries with differences along these dimensions, for example reliances on sovereign wealth funds (Norway, Russia), hedging strategies (Mexico), or fiscal rules with stabilization funds (Chile). Historically, all six countries were materially affected during the Global Financial Crisis, the European Sovereign Debt Crisis, and the Covid-19 pandemic, with documented commodity price shocks and equity market volatility. Their experiences are widely studied, ensuring access to published research for support and citation. Finally, the manageable scope of this sample allows for a more granular examination of transmission channels, policy responses, and empirical evidence. This approach aligns with academic standards for comparative studies.

Estimations are conducted separately over three crisis periods: the Global Financial Crisis, the European Sovereign Debt Crisis, and the Covid-19 pandemic. The rationale for this approach is twofold. First, the three crises had distinct underlying causes and affected commodity prices in different ways. For example, the Global Financial Crisis was triggered by the collapse of U.S. mortgage-backed securities, which sharply reduced both U.S. demand and global demand. This collapse in global demand led to a steep decline in commodity prices, representing a commodity demand shock. The European Sovereign Debt Crisis, by contrast, was driven by fiscal imbalances and excessive public debt in several European countries. In response, governments contracted fiscal policy, which sharply reduced European demand and likewise depressed commodity prices, again reflecting a demand shock. The Covid-19 pandemic was a global health crisis that resulted in widespread lockdowns. These lockdowns caused sharp declines in both global supply and global demand, generating simultaneous commodity demand and supply shocks. Second, the shock spillovers differ across these crisis periods. During traditional financial crises such as the GFC and ESDC, transmission channels tend to amplify shocks primarily through heightened uncertainty and systemic vulnerabilities in credit, liquidity, confidence, and risk premia. By contrast, the Covid-19 pandemic introduced unique transmission dynamics, combining financial stress with real-economy disruptions.

Diagnostic checks confirm that our proposed specifications are well behaved, with standardized residuals generally exhibiting no significant serial dependence or regime-specific heteroskedasticity. Likelihood-ratio tests provide statistical evidence of contagion at the 5% significance level in Norway (when the model employs the VIX index as the global risk) and in the United States (when it uses the put-call ratio (PCR) as the global risk) during the Global Financial Crisis.

Finally, we evaluate the predictive performance of our proposed Markov Switching Augmented Factor model (which we will hereafter refer to as MS-FACTOR). Forecasts of commodity and equity returns generated by MS-FACTOR are benchmarked against three widely used alternatives: a VAR(p), a random walk, and a DCC-GARCH(1,1), across the three major crisis episodes—the Global Financial Crisis, the European Sovereign Debt Crisis, and the Covid-19 pandemic. The PCR-based MS-FACTOR often delivers the lowest root-mean-squared errors (RMSEs) for equity returns across both short-term (1-day ahead) and medium-term (20-day ahead) horizons, followed by its VIX-based variant. In contrast, the ranking for commodity return forecasts differs: DCC-GARCH(1,1), and occasionally VAR, outperform our proposed model at short-term horizons, although MS-FACTOR exhibits notable improvements mainly during high-volatility periods, such as the Covid-19 crisis. These patterns persist across medium-term horizons, albeit with less significant differences. The random walk model yields the weakest performance across asset types and horizons, reinforcing the value of regime-sensitive structures.

The remainder of this paper is organized as follows. Section 3.2 presents our Markov Switching with Observed Factor model, our contagion tests, and the description of the data. Section 3.3 provides the empirical results. Section 3.4 concludes.

3.2 Methodology

Miranda-Agrippino and Rey (2021) provide evidence of the existence of a global risk factor that drives almost all risky asset prices (commodities and equities). They also show that one of the main proxies of the global risk factor frequently used is the VIX index, which captures risk aversion. Other studies such as Pinto-Ávalos et al. (2024), Zhang and Hamori (2021), and Silvennoinen and Thorp (2013) highlight that commodity and equity returns may also co-move under the influence of alternative global risks, such as investor sentiment, which can act as an important determinant of asset prices (Baker & Wurgler, 2006; Statman

& Fisher, 2002). Studies such as Baek et al. (2005) suggest that investor sentiment may provide a more powerful explanation for short-term asset price fluctuations than traditional fundamental indicators. A widely used proxy for investor sentiment is the put–call ratio (PCR), which tracks actual investor hedging behavior. This ratio can amplify feedback loops between equities and commodities and often shows coordination with the VIX in contagion studies (Oyster, 1997; Pinto-Ávalos et al., 2024). For robustness checks, we also consider the U.S. economic policy uncertainty (EPU) index proposed by Baker et al. (2016), which reflects macroeconomic policy uncertainty and is crucial for explaining disruptions in global trade cycles and commodity demand.

Although each indicator points to contagion mechanisms, we consider VIX, PCR, and U.S. EPU separately for the following reasons.[4] (1) Parsimony: given the high correlation among them, Bekaert et al. (2014) show that including multiple measures does not significantly improve the explanatory power of contagion models. Similarly, Miranda-Agrippino and Rey (2021) emphasize parsimony in global financial cycle proxies to avoid redundancy; (2) Indicator-specific interpretability: VIX is forward-looking, PCR is behavioral, and EPU is structural. Using one indicator at a time allows clearer interpretation of which risk dimension drives spillovers (Baker et al., 2016); (3) Empirical evidence of sufficiency: Rey (2013) identifies VIX as the dominant proxy for Global Financial Cycle shocks; (4) Robustness and comparability: following many benchmark studies (Bekaert et al., 2014; Kilian & Park, 2009), we use one proxy at a time to ensure comparability across samples and crises.

We exclude lagged commodity returns from the mean equation in line with theoretical models and empirical findings showing that, once current fundamentals are controlled for, past commodity price movements offer no additional predictive value. From a theoretical perspective, competitive-storage and cost-of-carry frameworks suggest that spot and futures prices adjust immediately to reflect determinants such as inventory levels, convenience yields, hedging pressure, and prevailing market sentiment. This adjustment leads returns to follow a martingale process, rendering past shocks uninformative for current expectations (LeRoy, 1989; Samuelson, 2016). Any residual serial dependence would be eliminated through arbitrage

[4]Other global risk measures could also be considered, such as the international illiquidity risk metric proposed by (Amihud, 2019; Amihud, 2002), defined as the average absolute return per dollar of trading volume. Higher values indicate lower market liquidity because small trades generate large price movements, whereas lower values reflect deeper markets with minimal price impact. This measure has been shown to predict international stock returns, particularly during crisis periods, as illiquidity carries a risk premium. However, Racicot et al. (2025) document that this global illiquidity risk measure and the VIX index highly co-move, implying that VIX already captures much of the information contained in international liquidity conditions.

strategies such as cash-and-carry or calendar spreads (Brennan, 1976; Working, 1949). Empirically, studies on individual contracts Fama and French (2016), diversified commodity portfolios Erb and Harvey (2006), and factor-sorted panels Bakshi et al. (2019) consistently find near-zero autocorrelation in commodity returns once fundamental variables are included.[5]

Thus, our equity-return equation includes its own one-period lag, capturing short-run persistence, while we do not include any lagged dependent terms in our commodity equations. Crucially, the loading on the lagged return of the equity equation is allowed to vary depending on the volatility regime of the corresponding commodity-specific market shock. This choice is motivated by Rossi (2012), who documents a statistically significant directional influence running from commodities to equities: higher international commodity prices today tend to be followed by stronger equity performance tomorrow.[6]

3.2.1 Markov Switching Augmented Factor (MS-FACTOR) model

To illustrate the methodology across multiple countries, we pair—within each country—its principal export commodity with its domestic equity index. For every country in the sample, we compute commodity and equity returns by taking the natural logarithm of their prices, first-differencing them, and multiplying by 100. In our notation, $r_{1,t}$ represents the commodity market return, while $r_{2,t}$ represents the equity market return. The return $r_{1,t}$ is influenced by (i) an intercept term c_1, (ii) the observable global risk factor f_t, and (iii) an unexpected error term $\varepsilon_{1,t}$. As for $r_{2,t}$, it is influenced by: (i) an intercept term, c_2, (ii) its own one-period lag, $r_{2,t-1}$, (iii) the observable global risk factor, f_t, and (iv) an unexpected error, $\varepsilon_{2,t}$:

$$r_{1,t} = c_1 + \theta_1 f_t + \varepsilon_{1,t}, \tag{3.1}$$

$$r_{2,t} = c_2 + \beta_t r_{2,t-1} + \theta_2 f_t + \varepsilon_{2,t}, \tag{3.2}$$

where θ_1 and θ_2 denote the impacts of the global risk factor f_t on the means of $r_{1,t}$ and $r_{2,t}$, respectively, and β_t is a time-varying loading parameter associated with the lagged equity return $r_{2,t-1}$. The error terms $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ are assumed to be correlated.

[5]We try empirical estimations including the first lagged commodity return in the commodity equation and find that the model does not converge, confirming both the theoretical predictions and previous empirical studies.

[6]Rossi (2012) reports a positive and statistically significant relationship between current commodity prices and lagged equity returns, indicating that shocks originating in commodity markets propagate forward to affect subsequent equity-market performance.

We consider that the global risk factor follows an autoregressive process (Junttila & Vataja, 2018; Romo, 2012):

$$f_t = \alpha f_{t-1} + \omega_t, \tag{3.3}$$

where ω_t represents the unpredictable component of the global risk factor. The error terms $\varepsilon_{i,t}$, for $i = 1, 2$, are specified as:

$$\varepsilon_{i,t} = \delta_i \xi_t + \eta_{i,t}, \tag{3.4}$$

where ξ_t represents other unpredictable shocks common to both asset markets, δ_i quantifies the sensitivity of the error term $\varepsilon_{i,t}$ to these influences, and where $\eta_{i,t}$ denotes the market-specific disturbance; that is $\eta_{1,t}$ and $\eta_{2,t}$ are uncorrelated.

Substituting equation (3.3), (3.4) into (3.1) and (3.2) yields:

$$r_{1,t} = c_1 + \lambda_1 f_{t-1} + \theta_1 w_t + \delta_1 \xi_t + \eta_{1,t}, \tag{3.5}$$

$$r_{2,t} = c_2 + \beta_t r_{2,t-1} + \lambda_2 f_{t-1} + \theta_2 w_t + \delta_2 \xi_t + \eta_{2,t}, \tag{3.6}$$

where λ_1 and λ_2 are the impact coefficients of the lagged global risk factor on market returns $r_{1,t}$ and $r_{2,t}$, respectively.

Let $\gamma_{i,t}$ denote the total unpredictable common disturbance affecting the market i:

$$\gamma_{i,t} = \theta_i \omega_t + \delta_i \xi_t, \tag{3.7}$$

where, $\gamma_{i,t}$ and $\eta_{i,t}$, $(i = 1, 2)$ are assumed to be uncorrelated.

We consider that each disturbance $\eta_{i,t}$ is driven by an idiosyncratic shock $z_{i,t}$, representing financial and economic conditions specific to that particular market:

$$\eta_{i,t} = \sigma_{i,t} z_{i,t}. \tag{3.8}$$

Also, let the common disturbance $\gamma_{i,t}$ be driven by a common shock $z_{c,t}$, which captures all kinds of global risks:

$$\gamma_{i,t} = \sigma_{c,i,t} z_{c,t}. \tag{3.9}$$

Now assume that, conditioning on $\mathcal{F}_{t-1} = \{r_{2,t-1}, f_{t-1}\}$,the three shocks $(z_{1,t},\ z_{2,t},\ z_{c,t})$ are standardized normally distributed variables with mean 0 and variance 1, and are uncrorrelated two by two. The parameters $\sigma_{i,t}$ and $\sigma_{c,i,t}$ thus represent the magnitudes of the idiosyncratic and common shocks affecting market i, respectively.

Substituting equations (3.8) and (3.9) into (3.5) and (3.6) yields:

$$r_{1,t} = c_1 + \lambda_1 f_{t-1} + u_{1,t}, \tag{3.10}$$

$$r_{2,t} = c_2 + \beta_t r_{2,t-1} + \lambda_2 f_{t-1} + u_{2,t}, \tag{3.11}$$

$$u_{1,t} = \sigma_{c,1,t} z_{c,t} + \sigma_{1,t} z_{1,t}, \tag{3.12}$$

$$u_{2,t} = \sigma_{c,2,t} z_{c,t} + \sigma_{2,t} z_{2,t}. \tag{3.13}$$

In addition, we assume that the impact of the common and idiosyncratic shocks switch between low- and high-volatility regime according to distinct Markov-switching processes, and where the underlying state variables are $A_{n,t}$, $n = c, 1, 2$. The state variable $A_{n,t}$ takes values in $\{0, 1\}$, where 0 corresponds to the low-volatility regime (L) and 1 corresponds to the high-volatility regime (H). The transition probability matrix for each shock is

$$\mathbf{P}_n = \begin{pmatrix} p_{HH} & p_{HL} \\ p_{LH} & p_{LL} \end{pmatrix}, \tag{3.14}$$

where $p_{HH} = Pr[A_{n,t} = 1 | A_{n,t-1} = 1]$, $p_{LL} = Pr[A_{n,t} = 0 | A_{L,t-1} = 0]$, $p_{HL} = 1 - p_{HH}$, and $p_{LH} = 1 - p_{LL}$, with $n = c, 1, 2$. Therefore, $\sigma_{c,i,t}$ and $\sigma_{i,t}$ evolve as follows:

$$\sigma_{c,i,t} = \sigma_{c,i,L}(1 - A_{c,t}) + \sigma_{c,i,H} A_{c,t}, \tag{3.15}$$

$$\sigma_{i,t} = \sigma_{i,L}(1 - A_{i,t}) + \sigma_{i,H} A_{i,t}, \quad (i = 1, 2), \tag{3.16}$$

where $\sigma_{c,i,L}$ represents the impact of the common shock $z_{c,t}$ on the market i in the low-volatility regime, while $\sigma_{c,i,H}$ captures the corresponding impact in the high-volatility regime. Similarly, $\sigma_{i,L}$ measures the impact of the idiosyncratic shock $z_{i,t}$ on the market i in the low-volatility regime, while $\sigma_{i,H}$ measures the corresponding impact in the high-volatility regime $(i = 1, 2)$.

Finally, we assume that the loading of the lagged equity return β_t changes according to the volatility regime of the idiosyncratic commodity shock $z_{1,t}$:

$$\beta_t = \beta_L(1 - A_{1,t}) + \beta_H A_{1,t}, \tag{3.17}$$

where, β_L is the impact of the lagged equity return when the idiosyncratic commodity shock is in its low-volatility regime, while β_H is the corresponding coefficient when that shock is in its high-volatility regime.

Note that, in the model above, we assume that the intercept terms c_1 and c_2 remain constant, in line with standard empirical implementations of the CAPM model. This approach is supported by studies such as McDonald et al. (2009) and Zarifhonarvar (2023), which employ fixed intercepts. While alternative approaches—e.g., Lettau and Ludvigson (2001b) and Brennan et al. (2004)—permit time-varying intercepts, we favor the static formulation to reduce the number of estimated parameters and enhance computational efficiency. Diagnostic checks (see Tables 3.16-3.19 in the Appendix) confirm that our chosen specification adequately fits the data. For similar reasons of parsimony, we also maintain constant loadings on the lagged global risk factor.

3.2.2 Contagion test

To assess whether there is evidence of contagion, we test whether the loading on the lagged equity return varies in a manner that amplifies the influence of commodity-specific shocks on the equity mean equation. Specifically, we compare the regime-dependent coefficients β_L (low-volatility regime) and β_H (high-volatility regime), conducting a likelihood ratio test to determine whether the difference between these loadings is statistically significant.

H_0: $\beta_L = \beta_H$ vs H_1: $\beta_L \neq \beta_H$,

where the null hypothesis (H_0) implies that there is no statistical difference in the loadings of the lagged return of the equity market in the high and low volatility regime. We refer to it as a situation of no contagion. The alternative hypothesis (H_1) implies a statistically significant difference between the two loadings. Under (H_0), we estimate the restricted model and obtain its log-likelihood $L_{restricted}$; under (H_1), we estimate the unrestricted model and obtain $L_{unrestricted}$. The likelihood-ratio statistic is then

$$LR = -2(L_{restricted} - L_{unrestricted}), \tag{3.18}$$

which under H_0, asymptotically follows a χ^2 distribution with one degree of freedom ($\chi^2(1)$). If we reject the null hypothesis of equality and find that the absolute value of the estimated β_H is greater than the absolute value of the estimated β_L, then we conclude that contagion has occurred.

3.2.3 Data

The dataset consists of daily observations of country-specific equity indices and corresponding 1-year commodity futures prices for six commodity-exporting economies. For Canada, the equity series is the S&P/TSX Composite Index traded on the Toronto Stock Exchange, and the commodity is Western Canadian Select crude (heavy, sour) traded on NYMEX (New York Mercantile Exchange), representing about 12.7% of GDP over the period 2006–2022. For Mexico, the equity series is the IPC (Índice de Precios y Cotizaciones) traded on the Bolsa Mexicana de Valores market, and the commodity is Maya crude (heavy, sour), traded in U.S. and international markets, accounting for roughly 7% of GDP over 2006–2022. For Russia, the equity series is the MOEX Russia Index traded on the Moscow Exchange (MOEX), while the commodity is Urals crude (a heavy, sour Brent proxy) traded on ICE (Intercontinental Exchange), contributing around 19% of GDP during 2006–2022. For the U.S., the equity series is the S&P 500 Index traded on NYSE/NASDAQ, and the commodity is WTI crude oil (light, sweet) traded on NYMEX, representing only about 0.2% of GDP over 2006–2022. For Norway, the equity series is the OBX Index, a traded on the Oslo Børs (Oslo Stock Exchange), paired with Brent crude oil (light, sweet) traded on ICE, which accounts for approximately 25.7% of GDP across 2006–2022. For Chile, the equity series is the IPSA (Índice de Precios Selectivo de Acciones) traded on the Santiago Stock Exchange, and the commodity is high-grade copper traded on COMEX (Commodity Exchange, New York), representing about 26.3% of GDP over the 2006–2022 period.[7]

We use commodity futures prices instead of spot prices because futures markets are more dynamic and respond faster to price adjustments. The commodity prices are denominated in U.S. dollars, while the equity market prices are in domestic currency. This allows us to avoid

[7] Oil futures are priced differently in different countries; that is, the price depends on many factors such as the oil's grade (e.g., heavy, light, sour, sweet), where it is produced, where it is sold, transportation costs, etc. We use information from https://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=pet&s=mcrexus1&f=a and https://unctadstat.unctad.org/EN/Index.html, FED and datastream to compute the percentage share of GDP.

introducing confounding factors associated with fluctuations in the value of the U.S. dollar.[8]

To control for global risk that contributes to the co-movement between commodity and equity market returns, we incorporate separately into the mean of each return the lagged VIX index and the lagged put–call ratio (PCR). The PCR is defined as the ratio of the trading volume of put options to that of call options on the Chicago Board Options Exchange (CBOE). During periods of financial turmoil, risk-averse investors anticipating a decline in asset values tend to purchase more put options and fewer call options to hedge against potential losses. As a result, the PCR typically increases during episodes of elevated financial stress.

To assess the robustness of our findings, we consider lagged U.S. economic policy uncertainty (EPU) index, developed by Baker et al. (2016). Robustness checks using the U.S. EPU are reported in Tables 3.12–3.15 of the Appendix. The U.S. EPU Index quantifies economic uncertainty by analyzing the frequency and context of news articles related to fiscal policy, trade developments, and geopolitical risks, as well as measuring disagreement among economic forecasts. Elevated U.S. EPU levels signal increased uncertainty, which is commonly associated with reduced investment activity, slower economic growth, greater market volatility, and lower asset valuations.

All three global factors are sourced from Bloomberg. Crude oil and equity market prices are obtained from Datastream, while copper prices for Chile and its equity market data are retrieved from the CEIC Asia database (formerly known as the China Economic Information Center). The sample period spans from January 10, 2006, to June 8, 2024.

3.3 Empirical results

3.3.1 Econometric validation

Before analyzing the dynamic relationship between commodity and equity returns in each country during the three major crises (the Global Financial Crisis, the European Sovereign Debt Crisis, and the Covid-19 pandemic), we first check whether our model provides an adequate statistical fit over each examined period. Tables 3.16–3.19 in the Appendix present

[8] If we convert equity prices into U.S. dollars, every equity return would mix two things at once: (1) the genuine move in that country's stock market, plus (2) the daily change in the U.S./local-currency exchange rate. By leaving equities in local currency, we measure the pure stock-market movement. Commodity prices remain in U.S. dollars because that is their natural quote, and any variation in the U.S. dollar affects all equally.

diagnostic tests on the standardized residuals, including autocorrelation (Ljung-Box), normality (Jarque-Bera), and conditional volatility (ARCH LM) tests, performed separately for commodity and equity returns during all three crises. Specifically, Tables 3.16 and 3.19 display results when the VIX is employed as the global risk factor, while Tables 3.17 and 3.18 present results using the put–call ratio instead.

3.3.1.1 Serial correlation in standardized residual (Ljung-Box tests)

Tables 3.16–3.19 report the Ljung-Box Q-statistics for the standardized residuals of each country's commodity and equity returns.[9] Across our 36 cases (six countries, three crisis periods, and two different options for our global financial measures), we find that the null hypothesis of no autocorrelation is overwhelmingly not rejected at the 5% significance level (p-values > 5%). This suggests that our model effectively captures the predictable dynamics of the return series. Figures 3.10-3.15 in the Appendix visually reinforce this conclusion, as they show no discernible patterns in the plotted standardized residuals of the asset returns.

3.3.1.2 Normality test (Jarque-Bera tests)

Tables 3.16 and 3.17 in the Appendix also present the results of Jarque–Bera (JB) normality tests, applied to the above-described standardized residuals. Again, we find that the null hypothesis of normality is majoritarily not rejected at the 5% significance level. In the context of our Markov-switching model, this outcome further confirms that our model appropriately captures regime-dependent dynamics in joint asset returns, effectively adjusting for structural shifts and changing market conditions. Moreover, this result holds regardless of which global risk factor is included: the VIX (reported in Table 3.16) or the put–call ratio (reported in Table 3.17).

3.3.1.3 Heteroskedasticity test (ARCH tests)

Finally, the above-mentioned tables also report the results of the ARCH-LM (Autoregressive Conditional Heteroskedasticity Lagrange Multiplier) tests, applied at lag 4 to our standardized residuals. We see that, again, for all six countries and for the three crisis periods, and regardless of the global risk proxy used, the null hypothesis of no ARCH effects of order four is largely

[9] We examine the hypothesis of no-autocorrelation over several possible lag lengths (specifically, 1,4, 30, and 90).

not rejected at the 5% significance level. This suggests that our model is able to adequately capture volatility dynamics up to at least four lags.[10]

3.3.2 Estimation results of the MS-FACTOR model

Before investigating statistically for the presence of contagion amongst our return pairs, we first examine the estimated impacts of the various model shocks on the variances of the commodity and equity returns across different regimes and crisis periods. This initial analysis helps us to understand how volatility behaves under various market conditions. Next, we assess the estimated influence of the global risk factor in shaping the mean returns of each market, highlighting its contribution to directional movements. Finally, we explore the transmission of commodity-specific shocks to equity returns, offering insight into the mechanisms driving cross-market interactions and potential spillover effects. Tables 3.1–3.3 present the estimated model parameters over each of the three major crisis periods, respectively: the Global Financial Crisis, the European Sovereign Debt Crisis, and the Covid-19 pandemic.

3.3.2.1 Variances of the different shocks

When the VIX is used as a global risk proxy, we find that idiosyncratic shock volatilities exhibit marked regime dependence. In low-volatility regimes, the impact coefficients—$\sigma_{1,L}$ for commodities and $\sigma_{2,L}$ for equities—across countries range from 0.001 to 1.927 and from 0.083 to 1.167, respectively, prior to the GFC; from 0.000 to 1.298 and from 0.005 to 1.031, respectively, prior to the ESDC; and from 0.000 to 1.625 and from 0.001 to 0.924, respectively, prior to the Covid-19 pandemic. These coefficients significantly increase in the high-volatility regimes (denoted $\sigma_{1,H}$ and $\sigma_{2,H}$). We now see that they range from 1.815 to 4.563 and from 0.800 to 5.117, respectively, during the GFC; from 1.492 to 5.690 and from 0.828 to 2.949, respectively, during the ESDC; and from 1.051 to 6.958 and from 0.686 to 4.610, respectively, during the Covid-19 pandemic.

[10]Choosing lag 4 for daily data is a pragmatic compromise: it checks for conditional variance patterns over the last four trading days (roughly equivalent to a week, excluding weekends), which is long enough to capture short-term persistence in volatility without overfitting or introducing noise from more distant lags.

Table 3.1: Estimated Parameters of the MS-FACTOR Model during the Global Financial Crisis

	Norway		**Mexico**		**U.S.**		**Canada**		**Russia**		**Chile**	
Parameter	VIX	PCR	VIX	PCR	VIX	PCR	VIX	PCR	VIX	PCR	VIX	PCR
Variance parameters												
$\sigma_{c,1,L}$	**1.310**	**0.678**	0.005	**0.340**	**1.042**	**0.157**	**0.969**	**0.763**	**0.600**	**0.660**	0.406	**0.486**
	(0.249)	(0.084)	(0.041)	(0.115)	(0.338)	(0.000)	(0.070)	(0.105)	(0.120)	(0.121)	(0.240)	(0.102)
$\sigma_{c,2,L}$	**0.628**	**1.404**	**0.949**	**1.008**	**0.357**	0.467	**0.804**	**1.030**	**1.119**	**1.128**	**0.302**	**0.846**
	(0.154)	(0.044)	(0.057)	(0.016)	(0.137)	(0.645)	(0.063)	(0.051)	(0.243)	(0.087)	(0.050)	(0.038)
$\sigma_{c,1,H}$	**2.425**	**2.482**	**2.300**	**2.718**	**5.260**	**0.499**	**2.251**	**2.258**	**3.523**	**1.849**	**1.628**	**1.259**
	(0.279)	(0.241)	(0.249)	(0.293)	(0.380)	(0.145)	(0.245)	(0.313)	(0.231)	(0.204)	(0.216)	(0.194)
$\sigma_{c,2,H}$	**2.718**	**3.666**	**1.759**	**2.340**	**1.800**	**1.270**	**2.818**	**2.931**	**3.575**	**3.135**	**1.216**	**2.180**
	(0.217)	(0.183)	(0.167)	(0.176)	(0.235)	(0.242)	(0.158)	(0.163)	(0.271)	(0.189)	(0.086)	(0.139)
$\sigma_{1,L}$	**1.422**	**1.786**	**1.927**	**1.969**	**1.807**	**2.042**	**1.675**	**1.766**	0.001	**1.847**	**1.888**	**1.993**
	(0.217)	(0.063)	(0.056)	(0.062)	(0.274)	(0.066)	(0.056)	(0.050)	(0.007)	(0.063)	(0.086)	(0.093)
$\sigma_{2,L}$	**1.167**	0.015	0.234	0.077	**1.060**	0.234	0.083	0.021	**1.050**	**0.974**	**0.689**	0.001
	(0.093)	(0.027)	(0.152)	(0.382)	(0.062)	(1.267)	(0.317)	(0.062)	(0.239)	(0.054)	(0.052)	(0.021)
$\sigma_{1,H}$	**3.857**	**3.856**	**4.563**	**4.514**	**1.815**	**5.656**	**3.758**	**3.830**	**1.976**	**4.589**	**3.852**	**3.913**
	(0.297)	(0.271)	(0.355)	(0.309)	(0.203)	(0.328)	(0.240)	(0.246)	(0.086)	(0.328)	(0.328)	(0.316)
$\sigma_{2,H}$	**3.573**	0.088	**1.895**	**1.576**	**3.256**	**3.202**	**0.800**	0.140	**5.117**	**7.253**	**2.339**	0.003
	(0.534)	(0.135)	(0.206)	(0.193)	(0.209)	(0.240)	(0.074)	(2.321)	(0.454)	(1.003)	(0.291)	(0.047)
Mean parameters												
λ_1	**-0.118**	**-0.650**	-0.001	**-1.610**	**-0.085**	**-0.960**	**-0.090**	0.991	**-0.168**	**-0.219**	**-0.088**	**-0.219**
	(0.044)	(0.132)	(0.024)	(0.218)	(0.040)	(0.288)	(0.044)	(0.771)	(0.044)	(0.005)	(0.016)	(0.005)
λ_2	**-0.298**	**-1.225**	0.012	0.358	**-0.092**	**-0.179**	**-0.093**	**-0.319**	**-0.253**	**-1.332**	**-0.083**	**-0.536**
	(0.036)	(0.550)	(0.020)	(0.519)	(0.042)	(0.047)	(0.032)	(0.094)	(0.045)	(0.631)	(0.022)	(0.049)
β_L	**-0.107**	-0.061	0.054	0.053	-1.174	**-0.122**	**-0.120**	0.073	**-0.288**	**-0.079**	**0.140**	**0.138**
	(0.036)	(0.038)	(0.038)	(0.045)	(1.880)	(0.041)	(0.046)	(0.044)	(0.113)	(0.037)	(0.042)	(0.046)
β_H	**-0.373**	0.028	**-0.140**	**-0.147**	-0.099	**-0.154**	**-0.271**	**-0.211**	-0.089	-0.055	**0.196**	**0.167**
	(0.079)	(0.056)	(0.072)	(0.071)	(0.057)	(0.081)	(0.072)	(0.079)	(0.034)	(0.073)	(0.096)	(0.074)

For each country, subscripts "1" and "2" denote the commodity and equity markets, respectively. The shock originates in the commodity market. Standard deviations are reported in parentheses. VIX and PCR are the global risk factors included in the model. Parameters in bold are statistically significant using a two-sided 5% test ($|t|$ >1.96).

Table 3.2: Estimated Parameters of the MS-FACTOR Model during the European Sovereign Debt Crisis

	Norway		Mexico		U.S.		Canada		Russia		Chile	
Parameter	VIX	PCR	VIX	PCR	VIX	PCR	VIX	PCR	VIX	PCR	VIX	PCR
Variance parameters												
$\sigma_{c,1,L}$	0.007	**0.809**	**0.822**	**0.821**	**0.446**	**0.581**	**0.697**	**0.884**	**1.537**	**1.536**	**0.711**	**0.717**
	(0.037)	(0.042)	(0.173)	(0.174)	(0.158)	(0.108)	(0.110)	(0.127)	(0.057)	(0.059)	(0.026)	(0.113)
$\sigma_{c,2,L}$	0.025	**1.168**	**0.492**	**0.491**	**0.342**	**0.441**	**0.454**	**0.509**	**0.853**	**0.853**	**0.541**	**0.583**
	(0.096)	(0.034)	(0.105)	(0.103)	(0.121)	(0.205)	(0.075)	(0.073)	(0.058)	(0.058)	(0.026)	(0.044)
$\sigma_{c,1,H}$	**1.520**	**1.610**	**2.401**	**2.399**	**1.442**	**1.505**	**2.397**	**2.848**	**5.061**	**5.115**	**1.225**	**1.208**
	(0.092)	(0.088)	(0.183)	(0.188)	(0.115)	(0.170)	(0.234)	(0.210)	(0.437)	(0.431)	(0.000)	(0.291)
$\sigma_{c,2,H}$	**1.392**	**2.512**	**1.436**	**1.435**	**1.107**	**1.143**	**1.563**	**1.640**	**1.859**	**1.996**	**1.683**	**1.682**
	(0.072)	(0.108)	(0.120)	(0.114)	(0.063)	(0.098)	(0.114)	(0.120)	(0.427)	(0.364)	(0.152)	(0.163)
$\sigma_{1,L}$	**1.254**	1.441	**1.256**	**1.257**	**1.298**	**1.291**	**0.672**	0.443	0.000	0.004	**1.295**	**1.296**
	(0.051)	(0.034)	(0.125)	(0.120)	(0.065)	(0.110)	(0.152)	(0.263)	(0.000)	(0.030)	(0.050)	(0.052)
$\sigma_{2,L}$	**0.760**	0.004	**0.505**	**0.503**	**0.324**	0.345	**0.595**	**0.632**	**1.031**	**1.040**	0.005	0.058
	(0.045)	(0.039)	(0.105)	(0.103)	(0.115)	(0.243)	(0.063)	(0.062)	(0.038)	(0.042)	(0.032)	(0.280)
$\sigma_{1,H}$	**4.360**	**4.151**	**5.690**	**5.700**	**4.916**	**4.926**	**2.516**	**2.157**	**1.492**	**1.501**	**2.829**	**2.836**
	(0.343)	(0.306)	(0.462)	(0.461)	(0.334)	(0.343)	(0.267)	(0.213)	(0.133)	(0.135)	(0.129)	(0.135)
$\sigma_{2,H}$	**2.697**	**3.565**	**1.250**	**1.249**	**2.048**	**2.019**	**1.889**	**2.117**	**2.949**	**2.941**	**0.828**	**0.829**
	(0.204)	(0.747)	(0.094)	(0.095)	(0.132)	(0.153)	(0.162)	(0.216)	(0.146)	(0.156)	(0.065)	(0.075)
Mean parameters												
λ_1	**-0.018**	0.095	-0.003	**-0.193**	-0.006	**0.333**	**-0.087**	**-0.206**	**-0.078**	**-0.785**	**-0.021**	**-0.329**
	(0.006)	(0.215)	(0.032)	(0.022)	(0.008)	(0.015)	(0.029)	(0.073)	(0.028)	(0.109)	(0.005)	(0.070)
λ_2	**-0.200**	**-1.369**	-0.009	**-0.263**	0.020	**-0.409**	**-0.062**	**-1.307**	**-0.149**	**-0.673**	**-0.021**	-0.077
	(0.031)	(0.396)	(0.015)	(0.028)	(0.023)	(0.097)	(0.022)	(0.289)	(0.031)	(0.044)	(0.006)	(0.352)
β_L	**-0.160**	**-0.062**	0.036	0.031	0.003	-0.051	-0.052	**-0.089**	-0.036	-0.036	**0.196**	**0.191**
	(0.031)	(0.028)	(0.052)	(0.028)	(0.090)	(0.063)	(0.044)	(0.044)	(0.042)	(0.042)	(0.038)	(0.038)
β_H	**-0.181**	-0.092	0.062	**-0.159**	**-0.120**	-0.016	-0.060	**-0.082**	**-0.120**	**0.126**	**0.129**	-0.037
	(0.084)	(0.085)	(0.104)	(0.068)	(0.074)	(0.035)	(0.071)	(0.038)	(0.042)	(0.057)	(0.060)	(0.175)

For each country, subscripts "1" and "2" denote the commodity and equity markets, respectively. The shock originates in the commodity market. Standard deviations are reported in parentheses. VIX and PCR are the global risk factors included in the model. Parameters in bold are statistically significant using a two-sided 5% test ($|t| > 1.96$).

Table 3.3: Estimated Parameters of the MS-FACTOR Model during the Covid-19 Pandemic

	Norway		Mexico		U.S.		Canada		Russia		Chile	
Parameter	VIX	PCR	VIX	PCR	VIX	PCR	VIX	PCR	VIX	PCR	VIX	PCR
Variance parameters												
$\sigma_{c,1,L}$	**0.883**	**0.789**	**1.598**	**2.112**	**1.072**	**0.507**	**0.482**	0.002	**0.876**	**1.446**	0.137	**0.377**
	(0.184)	(0.215)	(0.110)	(0.121)	(0.425)	(0.135)	(0.073)	(0.003)	(0.000)	(0.101)	(0.178)	(0.053)
$\sigma_{c,2,L}$	**0.510**	**0.622**	**0.287**	**0.304**	**0.235**	**0.571**	**0.518**	**0.652**	**1.074**	**1.179**	**1.207**	**1.351**
	(0.106)	(0.179)	(0.058)	(0.050)	(0.116)	(0.040)	(0.028)	(0.030)	(0.024)	(0.067)	(0.113)	(0.100)
$\sigma_{c,1,H}$	**2.955**	**3.212**	**19.225**	**15.230**	**4.922**	**6.062**	**0.921**	**0.660**	**1.670**	**2.378**	**1.051**	**1.653**
	(0.472)	(0.656)	(3.574)	(2.455)	(1.413)	(2.122)	(0.210)	(0.040)	(0.057)	(0.241)	(0.564)	(0.451)
$\sigma_{c,2,H}$	**2.536**	**2.535**	**0.487**	**0.303**	**1.078**	**1.220**	**5.848**	**3.888**	**2.397**	**3.650**	**1.855**	**5.317**
	(0.208)	(0.222)	(0.058)	(0.050)	(0.327)	(0.408)	(0.934)	(0.427)	(0.057)	(0.245)	(0.557)	(0.936)
$\sigma_{1,L}$	**1.525**	**1.575**	0.000	0.109	**1.296**	**1.602**	**1.157**	**1.103**	**1.625**	**1.386**	**0.913**	**0.892**
	(0.111)	(0.131)	(0.001)	(1.583)	(0.384)	(0.086)	(0.076)	(0.178)	(0.055)	(0.070)	(0.135)	(0.064)
$\sigma_{2,L}$	0.001	0.469	**0.924**	**0.920**	**0.503**	0.000	0.016	0.002	0.003	0.000	0.001	0.248
	(0.011)	(0.241)	(0.036)	(0.033)	(0.062)	(0.000)	(0.319)	(0.035)	(0.002)	(0.003)	(0.023)	(0.000)
$\sigma_{1,H}$	**6.958**	**7.117**	**2.951**	**7.996**	**5.749**	**3.671**	**4.806**	**5.149**	**0.660**	**4.834**	**1.051**	**1.796**
	(0.639)	(0.681)	(0.317)	(2.693)	(0.753)	(0.774)	(0.379)	(0.919)	(0.112)	(0.368)	(0.564)	(0.140)
$\sigma_{2,H}$	**0.686**	**1.196**	**2.301**	**2.284**	**3.347**	**1.052**	**1.104**	0.004	**2.776**	**1.513**	**1.855**	**4.610**
	(0.084)	(0.367)	(0.223)	(0.221)	(0.995)	(0.105)	(0.092)	(0.008)	(0.278)	(0.100)	(0.557)	(0.143)
Mean parameters												
λ_1	**-0.060**	**-2.523**	**-0.118**	**-2.407**	**-0.081**	**-3.100**	**-0.029**	**-0.973**	**-0.062**	**-0.535**	**-0.020**	**-0.374**
	(0.005)	(1.074)	(0.048)	(1.161)	(0.003)	(1.354)	(0.008)	(0.078)	(0.005)	(0.121)	(0.002)	(0.069)
λ_2	**-0.108**	**-1.832**	**-0.032**	**-0.988**	**-0.059**	**0.514**	**-0.049**	**-0.434**	**-0.169**	**-0.211**	**-0.099**	**-1.205**
	(0.022)	(0.530)	(0.006)	(0.153)	(0.007)	(0.646)	(0.024)	(0.198)	(0.044)	(0.096)	(0.030)	(0.188)
β_L	**-0.175**	**-0.133**	-0.033	-0.010	**-0.152**	-0.035	**-0.106**	-0.038	**-0.085**	-0.033	**-0.149**	**0.174**
	(0.043)	(0.043)	(0.112)	(0.062)	(0.091)	(0.062)	(0.062)	(0.066)	(0.040)	(0.037)	(0.095)	(0.079)
β_H	**-0.248**	0.066	0.012	0.426	0.383	**-0.660**	-0.267	-0.147	**-0.198**	**-0.139**	**-0.135**	**-0.064**
	(0.115)	(0.081)	(0.073)	(0.230)	(0.012)	(0.012)	(0.150)	(0.182)	(0.081)	(0.079)	(0.041)	(0.028)

For each country, subscripts “1” and “2” denote the commodity and equity markets, respectively. The shock originates in the commodity market. Standard deviations are reported in parentheses. VIX and PCR are the global risk factors included in the model. Parameters in bold are statistically significant using a two-sided 5% test ($|t| > 1.96$).

Importantly, we note that across all three crises, the idiosyncratic volatilities of oil and copper—key export commodities for the examined countries—are, on average, significantly higher than that observed in their corresponding equity markets. This disparity stems from several underlying economic and financial conditions unique to each crisis episode.

During the Global Financial Crisis, a sharp collapse in global demand, extensive supply-chain disruptions, and a surge in speculative trading in commodity futures markets intensified price swings. As equity markets plummeted, investors turned to commodities as alternative assets, amplifying volatility through feedback loops that pushed oil and metal prices into extreme fluctuations.[11] In the European Sovereign Debt Crisis, increased risk aversion and heightened safe-haven flows contributed to erratic movements in commodity markets, especially for energy-exporting economies such as Russia. Sporadic trade interruptions further destabilized prices, making commodities more reactive than equities to geopolitical and financial stressors. Finally, the Covid-19 pandemic brought about an unprecedented freeze in international trade and industrial activity, triggering severe supply-demand imbalances. This was most dramatically illustrated by the brief collapse of WTI oil futures to below zero, reflecting a breakdown in pricing mechanisms and storage constraints.[12] Although equity markets experienced elevated volatility during the pandemic, the scale and nature of shocks to commodity markets produced even greater idiosyncratic variance, underscoring their heightened sensitivity to real-economy disruptions.

When we instead incorporate the put–call ratio as the global risk factor, similar regime-sensitive behavior is observed, though commodity volatilities in the high-volatility regime tend to be slightly more pronounced than under VIX. This confirms that, regardless of the global proxy used, commodity markets remain more sensitive to the examined crisis (GFC, ESDC, Covid-19) compared to equity markets.

Our results also reveal that, before each crisis period—i.e. in the low-volatility regime—the impact coefficients of the common shocks are fairly similar across markets and countries. During crises, these coefficients ($\sigma_{c,1,L}$ and $\sigma_{c,2,L}$) rise significantly but their magnitudes remain largely uniform across both countries and asset classes. This pattern suggests that, on average,

[11] Commodities can act as safe-haven assets during periods of financial turmoil because investors often rebalance their portfolios toward alternative assets when traditional markets are under stress. Evidence shows that oil displays safe-haven properties during crises such as the GFC, with correlations to other asset classes becoming less positive, reflecting investor rotation (Tronzano, 2020). Further support comes from recent findings that both oil and metals can serve as safe-haven alternatives during crisis periods, even if return correlations diverge (Naeem et al., 2022; Sokhombela, 2024)

[12] On April 20, 2020, the May contract for WTI settled at minus $37.63 a barrel.

common shocks affect commodity and equity returns similarly across regimes and markets in our sample. It also indicates that our model appears able to successfully disentangle idiosyncratic volatility from common sources of volatility, capturing regime dependence while preserving cross-market coherence.

3.3.2.2 Impacts of the global risk factor

Tables 3.1–3.3 show that during crisis periods (the GFC, ESDC, and Covid-19 pandemic), the estimated lagged global factor loadings (λ_1 and λ_2) are generally negative and statistically significant across both asset classes and all examined countries. Increases in the global risk proxies—VIX or the PCR—are therefore associated with next period declines in oil and copper returns as well as domestic equity returns. This points to a strong co-movement between investor-sentiment indicators and market performance: when risk aversion rises, returns in both commodities and equities fall. The evidence underscores the central role of global risk factors in shaping asset prices, particularly under financial stress, and connect the two strands of the financial literature.

Thus, our findings align with recent research that highlights the role of the Global Financial Cycle in shaping international asset prices, as shown by Passari and Rey (2015), Coudert and Gex (2008), and Cheng et al. (2015). Passari and Rey (2015) demonstrate that the VIX index significantly depresses stock returns during crisis events such as the global financial crisis (2007–2009) and the European sovereign debt crisis. Coudert and Gex (2008) corroborate this view by emphasizing that shifts in investor risk aversion, as reflected in the VIX index, serve as predictors for episodes of equity market distress. Cheng et al. (2015) further support this by identifying a pronounced negative correlation between declines in commodity futures prices and elevated levels of the VIX.

Second, various studies examine the market trading behavior of options as a measure of investor sentiment and its linkage to future stock returns. For instance, Pan and Poteshman (2006) show that trading volumes in the options market contain valuable information for forecasting U.S. equity returns. Bathia and Bredin (2016) highlight the significance of investor sentiment—proxied by the put-call ratio—in explaining stock market performance across G7 countries.

3.3.2.3 Transmissions of idiosyncratic shocks

During the Global Financial Crisis, an increase in idiosyncratic volatility in the oil market—from the low regime to the high regime—coincides with a significant amplification in the impact of lagged equity returns on the conditional mean of equity returns in several countries. Specifically, when the model incorporates the VIX as the global risk, this amplification is observed in Norway and Canada; when using the PCR, it is evident in the United States. Estimated loadings on lagged equity returns become more negative in the high-volatility regime: -0.107 to -0.373 in Norway, -0.120 to -0.271 in Canada and -0.122 to -0.154 in the U.S. A more negative coefficient possibly reflects more pronounced risk-off behavior in investor sentiment under stress.[13] Conversely, in Chile—where copper is a key export—the results show a different pattern. As idiosyncratic volatility in the copper market intensifies, the estimated loading of lagged equity returns increases: from 0.140 to 0.196 when VIX is used, and from 0.138 to 0.167 with PCR. A rising (positive) coefficient suggests that past equity returns have a stronger contemporaneous influence on equity return changes, possibly pointing to synchronized optimism or coordinated risk-taking.

During the European Sovereign Debt Crisis, similar volatility shifts in the oil market prompt a modest increase in the magnitude of the lagged equity return estimated coefficient in Norway (from –0.160 to –0.181, VIX-driven), indicating a stronger transmission of oil dynamics to equity returns under heightened uncertainty. In contrast, Canada shows relatively stable behavior when using PCR, with the loading barely changing (from –0.089 to –0.082). Chile again demonstrates asymmetry: increased idiosyncratic copper volatility leads to a decline in the lagged equity coefficient from 0.196 to 0.129 (VIX-based).

During the Covid-19 pandemic, when oil market volatility sharply increased, we see stronger negative impacts of lagged equity returns in Norway and Russia (from –0.175 to –0.248 and –0.085 to –0.198, respectively, under VIX). These changes again imply intensified influence of past equity market conditions on current equity returns. Chile's copper market shows mixed behavior: under the VIX-based model, the equity return coefficient shifts slightly from –0.149 to –0.135, whereas under the PCR model, the coefficient reverses from 0.174 to –0.064, indicating a regime-sensitive adjustment in cross-market response dynamics.

[13] Risk-off behavior describes the risk sentiment where financial market participants reduce their exposure to risk to focus on protecting their capital.

In other cases, changes in idiosyncratic oil market volatility do not correspond with significant changes in the estimated equity return loadings. This suggests that during crisis periods, even large fluctuations in oil markets transmit through familiar channels, and do not trigger structural breaks or anomalous reactions in equity-linked dynamics.

Given that some countries exhibit statistically significant amplification in the impact of lagged equity returns on current equity returns as idiosyncratic oil volatility increases, we now proceed to examine statistically whether these changes reflect a regime shift—a structural break—or mark the onset of contagion.

3.3.3 Likelihood ratio test results

Tables 3.4 and 3.5 report the p-values from likelihood ratio tests when either the VIX or the PCR is used as the global risk in the model, respectively. Given our results, we focus on the cases where the test is significant at the 5% level, and where we also find that, during periods when the idiosyncratic commodity shock is in the high-volatility regime, the estimated impact of the lagged equity returns increases unambiguously. The latter signals an amplification of the commodity shock in the equity market, which we interpret as contagion. We thus find evidence for contagion in two of our six country cases: Norway and the U.S., both during the GFC, the former when VIX is used as the global proxy in our model, and the latter, when the PCR is used instead.

Chile also exhibits an LR test p-value below 0.05 during the Covid-19 pandemic, both when VIX or PCR are used. However, in the two cases, we see that the absolute value of the estimated impact on the lagged equity return does not amplify when the idiosyncratic copper market volatility is in the high-volatility regime. Thus, while there is structural change, it cannot be interpreted as evidence of contagion.[14]

For the remaining countries, there is no statistically significant evidence of contagion across the crisis periods analyzed. This implies that increases in commodity-specific volatility do not uniformly translate into structural breaks or amplified equity-return responses across all cases, underscoring the heterogeneous nature of cross-market linkages.

[14] What we observe is, rather, decoupling, defined as a scenario in which an idiosyncratic commodity shock enters a high-volatility regime, while the loading on lagged equity returns shifts in a way that does not amplify its impact on the mean process (Dungey et al., 2020).

Table 3.4: Likelihood Ratio Test for Contagion: VIX as Global Factor

	Norway	**Mexico**	**U.S.**	**Canada**	**Russia**	**Chile**
Panel A: Global Financial Crisis						
LR	**23.000**	0.000	0.000	0.007	1.740	0.360
p-value	**0.000**	0.984	1.000	0.931	0.187	0.549
Panel B: European Sovereign Debt Crisis						
LR	0.005	0.000	2.660	0.011	0.260	1.060
p-value	0.945	1.000	0.103	1.000	0.610	0.303
Panel C: Covid-19 Pandemic						
LR	0.003	0.140	0.535	0.940	0.126	**37.100**
p-value	0.956	0.708	1.000	0.332	0.723	**0.000**

The likelihood-ratio statistic (LR) is for the null of no contagion against the alternative of contagion for each market pair across the crisis periods; the corresponding p-values are also reported. The underlined bold values are statistically significant.

Table 3.5: Likelihood Ratio Test for Contagion: PCR as Global Factor

	Norway	**Mexico**	**U.S.**	**Canada**	**Russia**	**Chile**
Panel A: Global Financial Crisis						
LR	0.125	0.007	**50.420**	0.007	0.080	0.010
p-value	0.218	0.932	**0.000**	0.932	0.777	0.919
Panel B: European Sovereign Debt Crisis						
LR	0.003	0.180	1.900	0.540	0.460	1.020
p-value	0.958	0.671	0.169	0.462	0.498	0.313
Panel C: Covid-19 Pandemic						
LR	0.000	0.023	0.116	0.340	0.125	**17.620**
p-value	1.000	0.878	0.369	0.560	0.724	**0.000**

The likelihood-ratio statistic (LR) is for the null of no contagion against the alternative of contagion for each market pair across the crisis periods; the corresponding p-values are also reported. The underlined bold values are statistically significant.

To interpret these findings in the context of each of the different crises periods considered, it is informative to examine, for each country-specific market pair, the filtered probabilities of the high-volatility regime associated with the common and idiosyncratic shocks. The figures related to the common shocks help to identify and differentiate these global crisis episodes, allowing us to assess how a particular crisis originating outside specific commodity and equity markets examined can simultaneously affect both. Meanwhile, the figures for the idiosyncratic shocks illustrate how fluctuations specific to a particular country's commodity market propagates to its corresponding equity market. Our focus is on Norway during the Global Financial Crisis, and when the VIX is used as the global risk factor, and on the U.S. during the same crisis period, when the PCR serves as the global risk proxy.

When examining the filtered probabilities of high-volatility common shocks during the GFC period (Figure 3.1 and Figure 3.4), we observe notable peaks in August 2007, around March 2008, mid-September 2008, and from late 2008 until mid-2009. These peaks likely correspond to the following turbulent events: the collapse of subprime mortgage lenders and early credit market tensions (August 2007), the collapse of Bear Stearns (around March 2008), and the collapse of Lehman Brothers (mid-September 2008), which triggered a financial meltdown. Lastly, the period from late 2008 to mid-2009 corresponds to the ongoing global economic policy uncertainty due to bailouts and recession fears. These global shocks had a simultaneous and pronounced impact on both commodity and equity markets.[15]

Since commodities are priced in U.S. dollars and strongly sensitive to global demand conditions, the sharp contraction in international economic activity during the GFC led to significant declines in commodity prices. Export-oriented commodity markets responded strongly to these demand-side shocks, generating disruptions in pricing structures. For instance, the filtered probabilities of high idiosyncratic volatility in oil markets for Norway and the U.S. (Figure 3.2 and Figure 3.5, respectively) show noticeable peaks between October 2008 and May 2009.[16] This period corresponds to the aftermath of Lehman Brothers' collapse

[15] While the model is estimated on data from two distinct countries (Norway and the U.S.), and using different proxies for the global financial measure (VIX and PCR), the fact that our obtained estimated high-regime filtered probabilities for the common shock evolve in a practically identical manner in Figures 3.1 and 3.4 reassures us on the validity of our proposed specification for the objectives of this paper.

[16] Figures 3.2 and 3.5 correspond to the idiosyncratic Brent oil shocks and idiosyncratic WTI crude oil shocks, respectively, during the GFC. We observe some differences because, although both WTI and Brent oil prices collapsed due to falling global demand, WTI was more severely affected by storage and transportation bottlenecks at Cushing, Oklahoma. These constraints created a localized oversupply and caused WTI to trade at a persistent discount to Brent. Brent oil, being seaborne, remained a better indicator of global fundamentals. This divergence reflected temporary market segmentation rather than distinct global shocks (Büyük şahin

in mid-September 2008, which triggered widespread global panic and a severe contraction in international demand. The collapse in global economic activity contributed to a dramatic decline in oil prices, underscoring how export-oriented commodity markets react to global demand shocks.

Between mid-2008 and the end of the year, crude oil prices plunged from approximately \$140 per barrel to below \$40, driven by collapsing global demand and widespread recession fears. More precisely, and according to historical data from the U.S. Energy Information Administration (EIA), West Texas Intermediate (WTI) prices peaked at around \$145 in July 2008 and dropped to under \$38 by December 2008 (EIA, 2008). In response to the crisis, governments introduced aggressive monetary and fiscal measures. The U.S. Federal Reserve slashed interest rates to near zero and launched quantitative easing programs (Kohn & Sack, 2018), while the Obama administration enacted the \$787 billion American Recovery and Reinvestment Act in February 2009 (Congress, 2009). Despite these interventions, financial markets remained volatile, with uncertainty lingering through the early phases of the recovery. Signs of stabilization began to surface between March and May 2009, exemplified by the coordinated global response at the April G20 summit in London (G20, 2009). Nonetheless, oil demand remained muted, and market volatility persisted throughout the year.

The contraction in global demand led to steep declines in oil prices, which in turn dampened earnings expectations, triggered equity sell-offs, and eroded overall stock market valuations. Norway's economy and equity market—both deeply tied to the oil sector, were particularly vulnerable. As oil prices collapsed, expected revenues from petroleum exports fell sharply, resulting in significant declines on the Oslo Stock Exchange (OSE), especially within the energy, shipping, and oil services sectors (see high-probability spikes between September 2008 and January 2009 in Figure 3.3). Similarly, the United States—being both a major oil consumer and producer—experienced severe repercussions. The plunge in oil prices inflicted heavy losses on energy firms, including producers, refiners, and oilfield service companies, which comprise a substantial portion of U.S. equity indices. This contributed to broad-based declines in market performance during the crisis period (see high-probability spikes between 2008 and 2009 in Figure 3.6).

et al., 2013). Mexico and Russia oil prices also collapsed as global demand fell. However, the fluctuations of these prices depend also on the type of oil, where it is produced and how it is transported.

Panel A: Estimated filtered probabilities of high volatility shocks — Oil–equity market (Norway) during the GFC; VIX used in the model

Figure 3.1: Common shock

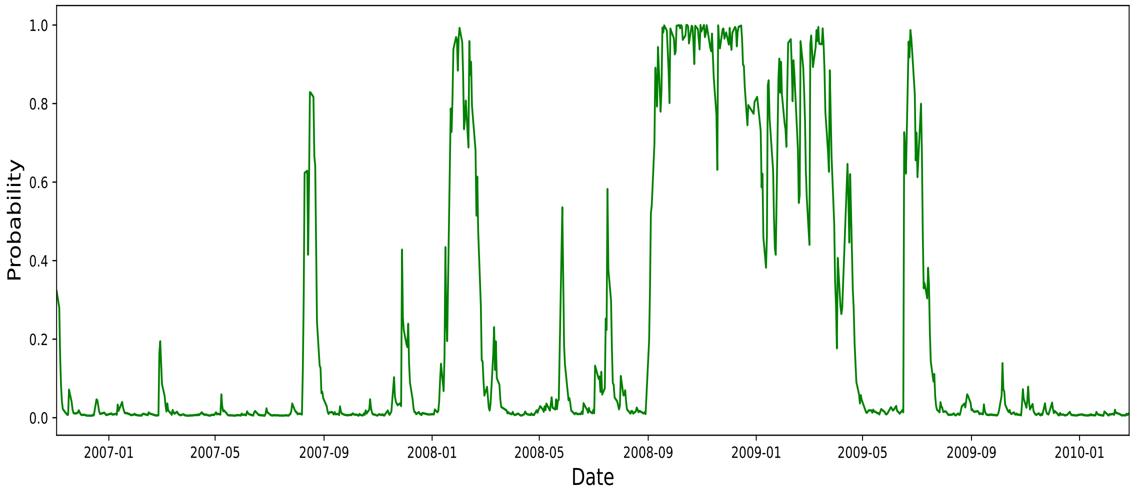


Figure 3.2: Idiosyncratic shock, Norway oil market (shock origin)

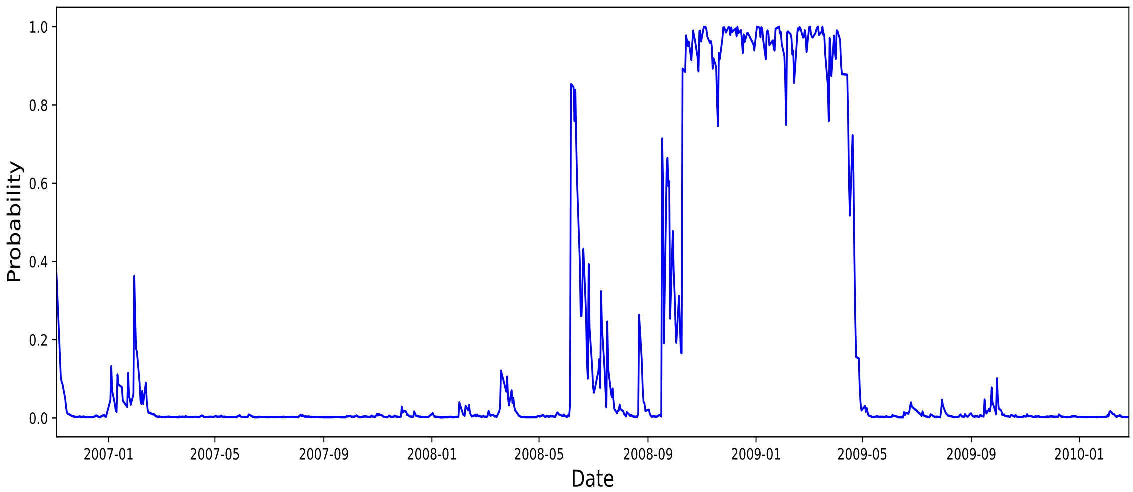
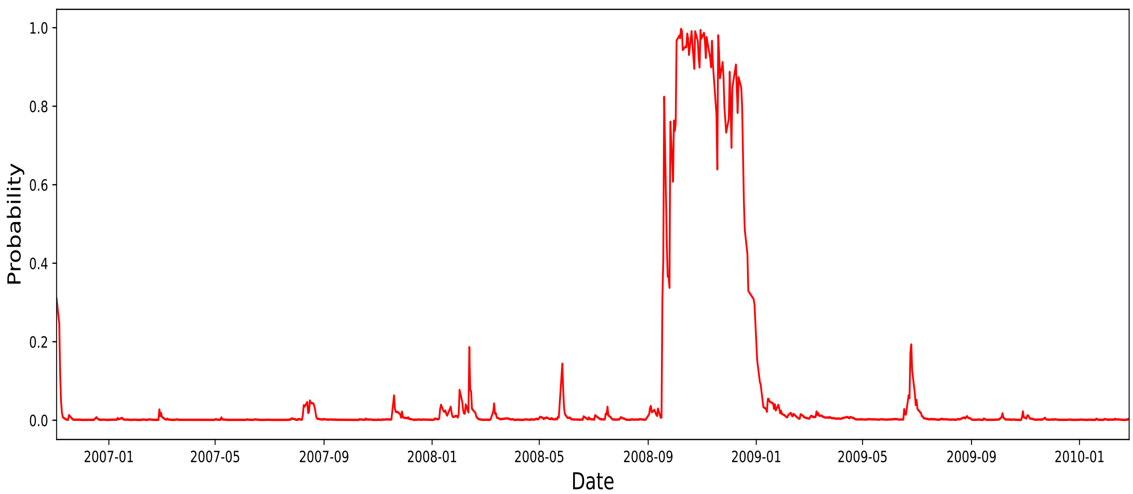


Figure 3.3: Idiosyncratic shock, Norway equity market (shock target)



Panel B: Estimated filtered probabilities of high volatility shocks — Oil-equity market (U.S.) during the GFC; PCR is used in the model

Figure 3.4: Common shock

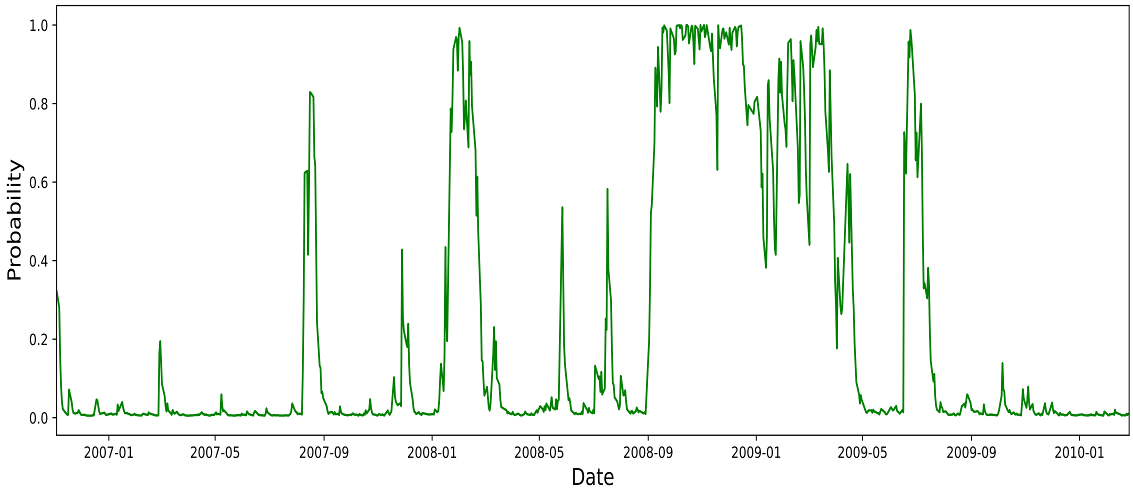


Figure 3.5: Idiosyncratic Shock, U.S. Oil market (Shock origin)

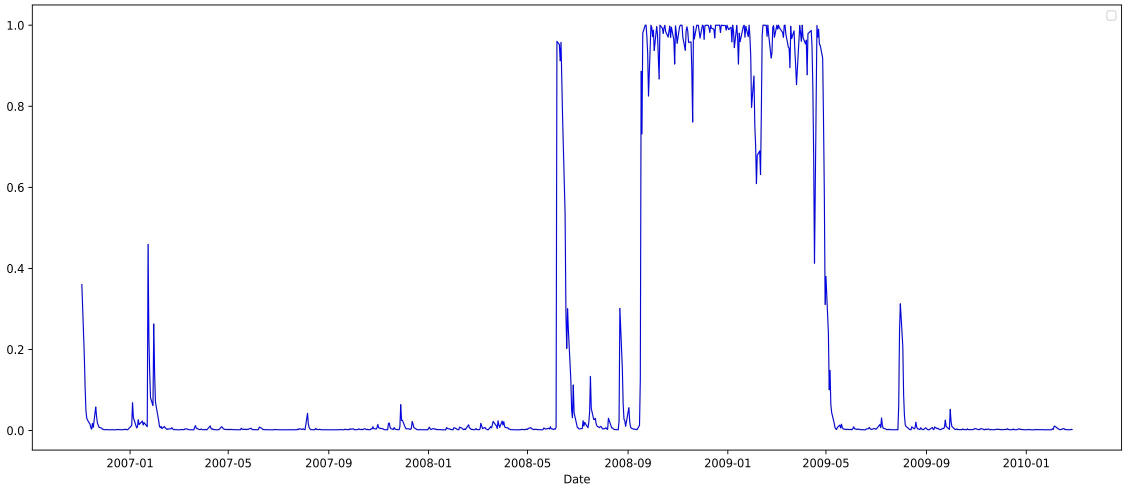
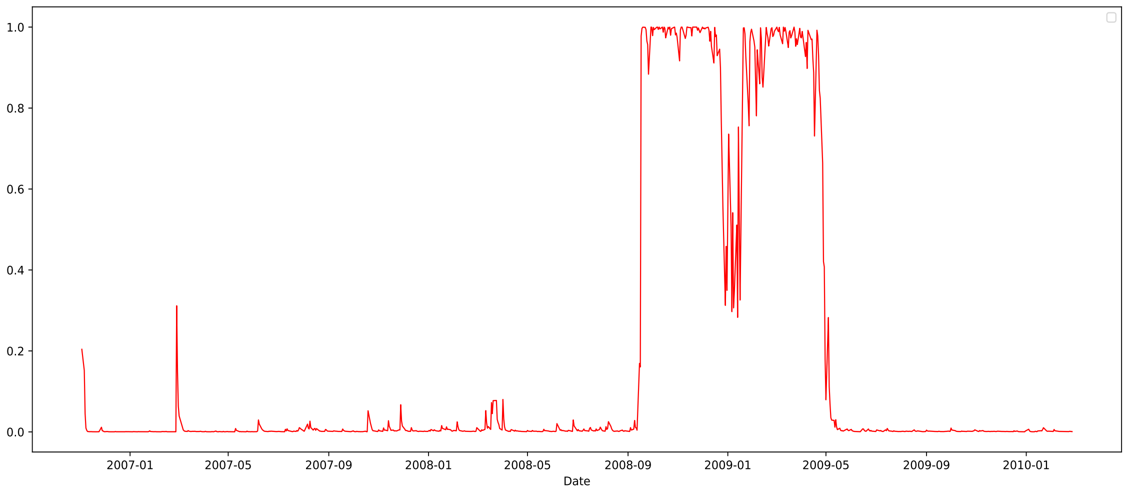


Figure 3.6: Idiosyncratic Shock, U.S. Equity market (Shock target)



3.3.4 Discussion

The findings of this study contribute to the growing literature on financial spillovers and contagion in commodity-exporting economies, particularly in the presence of global risk factors. Our analysis also reveals meaningful heterogeneities in spillover and contagion outcomes between high-income and emerging economies.

Pinto-Ávalos et al. (2024) examine, among other cases, the same pairs of markets and episodes of financial distress as those considered in our analysis. Controlling for the same global risk proxies as the ones we have used here, and on the basis of a DCC-GARCH model and of a Diebold-Yilmaz-based analysis, they report no evidence of contagion in Norway. A possible reason for the difference between our conclusion and theirs is that, as discussed earlier, DCC-GARCH models do not distinguish the source and destination of volatility shocks. This lack of directional identification raises critical concerns regarding whether volatility transmission originates from external global factors or is driven by market-specific forces. In a similar vein, the Diebold–Yilmaz spillover index captures only linear transmission mechanisms and fails to accommodate potential nonlinear dynamics that typically become more relevant during periods of crisis. These methodological limitations may result in mischaracterized cross-market linkages and an underestimation of regime-dependent, asymmetric spillovers—particularly those that become more pronounced under high-volatility conditions.

In a separate study, Silvennoinen and Thorp (2013) shed light on how global risk aversion, measured via the VIX index, affects correlations between U.S. equity returns and various commodity markets. They observe that, during the Global Financial Crisis, the correlation between oil and equity returns rises significantly—interpreted as evidence of contagion—while food commodities show weaker interconnections. Although their use of a smooth transition DCC-GARCH model helps to highlight these sectoral differences, their framework lacks key structural elements that can clarify the nature and direction of spillovers. In contrast, the findings from our MS-FACTOR model provide a more nuanced view of financial contagion. Unlike Silvennoinen and Thorp (2013)'s model, our framework explicitly distinguishes between low- and high-volatility regimes and separates common global shocks from idiosyncratic country- and market-specific fluctuations. This enables us to identify not just the existence of contagion, but also the conditions under which it intensifies—particularly during episodes like the GFC, ESDC, and Covid-19 pandemic. Most importantly, our model is able to provide more information than their approach by pinpointing the origin and directionality of spillovers. We show that in countries such as Norway and U.S., idiosyncratic shocks from the oil market significantly influence equity returns during periods of heightened volatility. These domestic transmission channels, which are overlooked in Silvennoinen and Thorp (2013)'s setup, are crucial for understanding the full mechanics of cross-market linkages during crises.

Zhang and Hamori (2021) employ the Diebold and Yilmaz (2012) approach alongside the frequency dynamics method developed by Baruník and Křehlík (2018) to examine the impact of oil price shocks on equity returns during the Covid-19 pandemic, focusing on three major economies: the United States, Japan, and Germany. Their findings indicate that global risk—measured through a systemic factor capturing global equity market volatility associated with infectious disease—is the dominant transmission channel during this period. In contrast, oil spillover effects play only a marginal role in explaining equity return fluctuations once global risk is taken into account. While, considering a wider set of developed and emerging markets, our results affirm that global risk factors such as the VIX and the put–call ratio substantially influence commodity and equity market returns, in contrast to those authors, we also reveal significant country-level spillovers, particularly from commodity markets into domestic equities during high-volatility regimes. For example, during the Covid-19 crisis, our model captures a strong volatility transmission from Chile's copper market shocks to its equity returns—underscoring the relevance of idiosyncratic dynamics in commodity-exporting economies. Thus, whereas Zhang and Hamori (2021) emphasize the role of global channels, our findings demonstrate that domestic transmission mechanisms can be equally impactful—particularly under conditions of financial stress.

The transmission of idiosyncratic commodity shocks is more clearly observed in high-income countries (U.S., Canada, and Norway) and appears substantially weaker in emerging economies (Russia, Mexico, and Chile). Although the equity markets of high-income countries are more diversified, commodity price shocks primarily affect sectoral indices such as energy and mining. These sectors are largely privately driven, sizable in scale, and their production plays a crucial role in sustaining domestic economic growth. Consequently, commodity shocks can significantly influence the stock valuations of these markets. By contrast, in emerging markets, Mexico's oil exports represent only about 7% of GDP, Russia's oil sector is heavily state-controlled and not fully reflected in the MOEX index, and equity returns in both economies respond predominantly to global risk rather than domestic commodity fluctuations. Chile stands as a partial exception: although copper constitutes a large share of GDP, its strong fiscal rules and stabilization funds absorb much of the associated volatility, thereby limiting the transmission of idiosyncratic copper shocks to the equity market.

These structural differences also help explain why contagion is detected only in high-income economies (U.S. and Norway). In Norway, where oil exports account for roughly 26% of GDP, global oil price fluctuations naturally spill over into equity valuations, especially during periods of heightened volatility, making contagion more likely. In the U.S., the collapse in aggregate demand during the Global Financial Crisis sharply reduced commodity prices, and despite the diversification of U.S. equity markets, this shock was transmitted through valuation adjustments, higher risk premiums, and crisis-induced co-movement, revealing the limits of diversification during systemic downturns. In contrast, emerging markets show no evidence of contagion: Mexico's modest exposure to oil limits its vulnerability to commodity price collapses; Chile's credible macroeconomic framework buffers copper-driven volatility; and Russia's capital controls and geopolitical segmentation reduce financial integration with global markets. In Canada, a diversified economy and strong fiscal institutions dampen external spillovers despite meaningful oil dependence. Taken together, these factors explain why contagion emerges in the U.S. and Norway but not in the other economies examined.

3.3.5 Out-of-sample forecasting evaluations

3.3.5.1 Rolling window out-of-sample forecasting

In this section, we compare the out-of-sample forecasting performance of our model with three widely used benchmarks: the random-walk model, VAR(p), and DCC-GARCH(1,1). The evaluation covers three crisis periods—the Global Financial Crisis, the European Sovereign Debt Crisis, and the Covid-19 pandemic—on one-day and 20-day forecast horizons.[17] We follow the procedure of Chang et al. (2023) to carry out this forecast exercise.

Our forecast model is defined as:

$$r_{1,t+h} = c_1 + \lambda_1 f_t + u_{1,t+h}, \tag{3.19}$$

[17]The 1-day and 20-day windows are widely used in financial forecasting to capture meaningful trends while maintaining model reliability. Beyond the 20-day window, commodity and equity returns become difficult to forecast because these assets are extremely volatile.

$$r_{2,t+h} = c_2 + \beta_{t+h} r_{2,t} + \lambda_2 f_t + u_{2,t+h}. \tag{3.20}$$

Let T_e be a fixed window size. If the first observation is at time $\ell_e = T_0$, we obtain the first forecast as follows:

Step 1: Estimate the model using data from the time interval $[\ell_e, \tau]$ for the independent variables $r_{2,t}$ and f_t, and from $[\ell_e + h, \tau + h]$ for the dependent variables $r_{1,t+h}$ and $r_{2,t+h}$, where $\tau = \ell_e + (T_e - 1)$. Next, extract the estimated parameters: $\widehat{\lambda}_1$, $\widehat{\lambda}_2$, $\widehat{c}_1$, $\widehat{c}_2$, $\widehat{\beta}_H$, $\widehat{\beta}_L$. Additionally, compute the filtered probabilities at the final in-window observation, denoted $\alpha_{n,\tau} = P(A_{n,\tau} = 1|\mathcal{F}_\tau)$ ($n = c, 1, 2$ and $\mathcal{F}_\tau = \{r_{2,\tau}, f_\tau\}$). Finally, obtain the estimated transition matrix governing the regime-switching dynamics:

$$\mathbf{P} = \begin{pmatrix} p_{HH} & p_{HL} \\ p_{LH} & p_{LL} \end{pmatrix}.$$

Step 2: Compute the out-of-sample forecasts at time $\tau + h + 1$:

$$\widehat{r}_{1,\tau+h+1} = \widehat{c}_1 + \widehat{\lambda}_1 f_{\tau+1}, \tag{3.21}$$

$$\widehat{r}_{2,\tau+h+1} = \widehat{c}_2 + \mathbf{E}[\widehat{\beta}_{\tau+h+1}|\mathcal{F}_\tau] r_{2,\tau+1} + \widehat{\lambda}_2 f_{\tau+1}, \tag{3.22}$$

$$\mathbf{E}[\widehat{\beta}_{\tau+h+1}|\mathcal{F}_\tau] = \widehat{\beta}_H P(A_{1,\tau+h+1} = 1|\mathcal{F}_\tau) + \widehat{\beta}_L (1 - P(A_{1,\tau+h+1} = 1|\mathcal{F}_\tau)), \tag{3.23}$$

where

$$P\big(A_{1,\tau+h+1} = 1 \,\big|\, \mathbf{F}_\tau\big) = \big[\,\alpha_{1,\tau},\, 1 - \alpha_{1,\tau}\big]\,\mathbf{P} \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Then, slide the estimation window forward by one observation-from $[\ell_e, \tau]$ to $[\ell_e + 1, \tau + 1]$ for the independent variables, and from $[\ell_e + h, \tau + h]$ to $[\ell_e + h + 1, \tau + h + 1]$ for the dependent variables so that its length remains T_e. Re-estimate the model on this new sample (**Step 1**) and generate the next set of forecasts (**Step 2**), now targeting $\widehat{r}_{1,\tau+h+2}$ and $\widehat{r}_{2,\tau+h+2}$. Repeat the two steps, advancing the window by one observation each time, until the final origin of the forecast $\tau^* = T$, where T is the last date in the entire sample.

After completing all iterations, collect the rolling out-of-sample forecasts, denoted by $\widehat{r}_{1,t}$ and $\widehat{r}_{2,t}$, along with their realized counterparts $r_{1,t}$ and $r_{2,t}$. Next, compute the rolling forecast errors as $\widehat{u}_{i,t} = r_{i,t} - \widehat{r}_{i,t}$ for $i = 1, 2$, and calculate the root-mean-squared error (RMSE) for each return series $i = 1, 2$:

$$\text{RMSE}_{i,h} = \sqrt{\frac{1}{N_{i,h}} \sum_t \widehat{u}^2_{i,t}},$$

where $N_{i,h}$ is the total number of h-step forecasts produced over the rolling evaluation period.

3.3.5.2 Forecast results

Tables 3.6-3.11 report out-of-sample root mean squared errors (RMSE) for each country over three crisis periods, namely the Global Financial Crisis, the European Sovereign Debt Crisis, and the Covid-19 pandemic.[18]

When comparing forecasting methods across short-term (1-day ahead) and medium-term (20-day ahead) horizons, a distinct divergence emerges between equity and commodity returns. At the 1-day horizon, the equity returns are generally most accurately predicted by the MS-FACTOR model, especially when the global risk factor is the put–call ratio (PCR). The VIX-based variant often ranks second, while DCC-GARCH(1,1), VAR, and the random walk generally underperform. For commodities, the ranking reverses: DCC-GARCH(1,1)—and occasionally VAR—tend to yield the lowest short-term errors, while MS-FACTOR rarely leads outside of the Covid-19 crisis period. This pattern generally persists over a 20-day horizon, though with less sharpness. MS-FACTOR (PCR/VIX) generally retains its advantage for equities, the VAR slightly outperforming it in specific cases (e.g., Russia). The DCC-GARCH(1,1) and the VAR continue to lead commodity forecasts, while the random walk remains consistently the weakest benchmark across asset classes and forecast horizons.

These differences reflect the underlying dynamics of each asset class. Equity returns hinge on changes in risk premia and investor sentiment, so allowing the conditional mean to vary in response to a global sentiment proxy pays off; PCR, drawn from options trading, gives a better positioning signal than VIX (which is based on volatility). Commodity returns, by contrast, have little predictable content but they exhibit strong volatility clustering and short-memory in covariances. Models that focus on second-moment dynamics and short-lag interactions—DCC-GARCH and VAR—therefore fit them better than a state-dependent mean model.

[18] The figures of the rolling window forecast errors are ready and available upon request.

Table 3.6: Canada: Out-of-Sample Forecast Accuracy (RMSE, %)

Market returns	1-Day Ahead					20-Days Ahead				
	MS-FACTOR VIX	MS-FACTOR PCR	DCC GARCH(1,1)	VAR(p)	RW	MS-FACTOR VIX	MS-FACTOR PCR	DCC GARCH(1,1)	VAR(p)	RW
Panel A: Global Financial Crisis										
Commodity return	1.544	1.578	1.071	**1.061**	1.601	1.355	1.400	**1.191**	1.195	1.762
Equity return	**1.070**	**1.068**	2.079	2.106	2.915	**1.072**	**1.070**	2.125	2.130	3.137
Panel B: European Debt Crisis										
Commodity return	0.984	1.001	0.724	**0.716**	1.063	0.898	0.906	**0.723**	**0.723**	1.059
Equity return	**0.838**	**0.809**	1.480	1.509	2.108	**0.789**	1.005	0.978	0.805	2.193
Panel C: Covid-19 Pandemic										
Commodity return	1.002	0.992	**0.929**	0.934	1.016	0.822	0.850	0.931	**0.764**	1.005
Equity return	**1.336**	**1.333**	1.576	1.905	2.588	1.421	**0.895**	1.425	0.905	1.525

This table reports root-mean-squared errors (RMSE) in % across models and crisis periods. Lower values shown in **bold** identify the model with superior forecast accuracy for each return series. MS-FACTOR forecasts are computed with two alternative global risk factors (VIX and PCR). RW denotes the random-walk benchmark. p is the optimal VAR lag length, set to p=1 in Panels A and B and p=7 in Panel C.

Table 3.7: Russia: Out-of-Sample Forecast Accuracy (RMSE, %)

Market returns	1-Day Ahead					20-Days Ahead				
	MS-FACTOR VIX	MS-FACTOR PCR	DCC GARCH(1,1)	VAR(p)	RW	MS-FACTOR VIX	MS-FACTOR PCR	DCC GARCH(1,1)	VAR(p)	RW
Panel A: Global Financial Crisis										
Commodity return	2.335	2.338	**2.213**	2.333	2.630	**2.348**	**2.338**	2.531	2.541	3.562
Equity return	2.282	2.295	**2.276**	2.288	2.785	1.567	1.670	1.562	**1.556**	1.564
Panel B: European Debt Crisis										
Commodity return	1.496	1.507	**1.193**	1.213	1.710	1.192	1.201	**1.190**	1.194	1.728
Equity return	**0.971**	**0.974**	1.431	1.548	2.058	1.004	**0.335**	0.488	0.359	0.405
Panel C: Covid-19 Pandemic										
Commodity return	**1.521**	**1.524**	4.907	4.953	5.534	**2.531**	**2.514**	4.883	4.903	5.168
Equity return	2.556	2.523	**2.483**	2.580	3.538	**1.163**	**1.157**	2.472	2.437	3.521

This table reports root-mean-squared errors (RMSE) in % across models and crisis periods. Lower values shown in **bold** identify the model with superior forecast accuracy for each return series. MS-FACTOR forecasts are computed with two alternative global risk factors (VIX and PCR). RW denotes the random-walk benchmark. p is the optimal VAR lag length, set to p=1 in Panels A and C, and p=11 in Panel B.

Table 3.8: Norway: Out-of-Sample Forecast Accuracy (RMSE, %)

Market returns	1-Day Ahead					20-Days Ahead				
	MS-FACTOR VIX	MS-FACTOR PCR	DCC GARCH(1,1)	VAR(p)	RW	MS-FACTOR VIX	MS-FACTOR PCR	DCC GARCH(1,1)	VAR(p)	RW
Panel A: Global Financial Crisis										
Commodity return	1.295	1.328	1.471	**1.470**	2.001	1.655	1.660	**1.664**	1.671	2.395
Equity return	1.480	**1.447**	2.070	2.081	2.972	**1.501**	**1.461**	2.109	2.091	3.117
Panel B: European Debt Crisis										
Commodity return	1.089	1.095	1.098	**1.082**	1.611	1.099	1.106	1.096	**1.095**	1.709
Equity return	1.221	**0.708**	1.471	1.493	2.032	1.473	**0.807**	0.810	0.815	0.809
Panel C: Covid-19 Pandemic										
Commodity return	1.540	1.533	0.999	**0.965**	1.509	0.822	0.850	0.931	**0.764**	1.005
Equity return	**2.008**	**0.976**	2.145	2.210	3.095	1.417	**0.981**	1.140	1.127	1.333

This table reports root-mean-squared errors (RMSE) in % across models and crisis periods. Lower values shown in **bold** identify the model with superior forecast accuracy for each return series. MS-FACTOR forecasts are computed with two alternative global risk factors (VIX and PCR). RW denotes the random-walk benchmark. p is the optimal VAR lag length, set to p=1 in Panel A, p=12 in Panel B, and p=7 in Panel C.

Table 3.9: Mexico: Out-of-Sample Forecast Accuracy (RMSE, %)

Market returns	1-Day Ahead					20-Days Ahead				
	MS-FACTOR VIX	MS-FACTOR PCR	DCC GARCH(1,1)	VAR(p)	RW	MS-FACTOR VIX	MS-FACTOR PCR	DCC GARCH(1,1)	VAR(p)	RW
Panel A: Global Financial Crisis										
Commodity return	1.160	1.167	**1.135**	1.184	1.618	1.215	1.220	**1.214**	1.218	1.837
Equity return	**1.023**	**1.016**	2.157	2.146	3.143	**1.191**	**1.141**	2.197	2.191	3.344
Panel B: European Debt Crisis										
Commodity return	0.849	0.855	**0.699**	0.704	0.981	0.798	0.801	**0.698**	0.699	0.993
Equity return	**0.807**	**0.709**	1.471	1.493	2.032	1.405	**0.721**	1.400	0.765	2.306
Panel C: Covid-19 Pandemic										
Commodity return	1.002	1.005	**0.932**	0.935	1.419	0.917	0.923	**0.928**	0.931	1.255
Equity return	**0.934**	**0.935**	1.973	2.125	2.847	**0.937**	**0.935**	1.965	1.957	2.715

This table reports root-mean-squared errors (RMSE) in % across models and crisis periods. Lower values shown in **bold** identify the model with superior forecast accuracy for each return series. MS-FACTOR forecasts are computed with two alternative global risk factors (VIX and PCR). RW denotes the random-walk benchmark. p is the optimal VAR lag length, set to p=12 in Panel A, and p=1 in Panels B and C.

Table 3.10: Chile: Out-of-Sample Forecast Accuracy (RMSE, %)

Market returns	1-Day Ahead					20-Days Ahead				
	MS-FACTOR VIX	MS-FACTOR PCR	DCC GARCH(1,1)	VAR(p)	RW	MS-FACTOR VIX	MS-FACTOR PCR	DCC GARCH(1,1)	VAR(p)	RW
Panel A: Global Financial Crisis										
Commodity return	1.339	1.336	**0.784**	0.797	1.149	0.905	0.904	0.879	**0.854**	1.347
Equity return	1.956	**0.766**	2.002	2.020	2.857	**1.577**	**1.020**	1.014	2.127	3.283
Panel B: European Debt Crisis										
Commodity return	0.801	0.777	**0.606**	0.620	0.851	0.765	0.758	**0.605**	0.611	0.885
Equity return	**0.804**	**0.783**	1.251	1.267	1.793	1.704	1.711	**1.251**	1.255	1.991
Panel C: Covid-19 Pandemic										
Commodity return	**1.568**	**1.571**	1.861	1.946	2.583	**1.570**	**1.577**	1.852	1.861	2.368
Equity return	1.880	1.862	**1.587**	1.615	2.542	1.500	1.483	1.384	**1.229**	1.346

This table reports root-mean-squared errors (RMSE) in % across models and crisis periods. Lower values shown in **bold** identify the model with superior forecast accuracy for each return series. MS-FACTOR forecasts are computed with two alternative global risk factors (VIX and PCR). RW denotes the random-walk benchmark. p is the optimal VAR lag length, set to p=1 in Panels A and B, and p=2 in Panel C.

Table 3.11: U.S.: Out-of-Sample Forecast Accuracy (RMSE, %)

Market returns	1-Day Ahead					20-Days Ahead				
	MS-FACTOR VIX	MS-FACTOR PCR	DCC GARCH(1,1)	VAR(p)	RW	MS-FACTOR VIX	MS-FACTOR PCR	DCC GARCH(1,1)	VAR(p)	RW
Panel A: Global Financial Crisis										
Commodity return	1.033	1.044	**1.031**	1.036	1.442	1.114	1.110	**1.086**	1.096	1.566
Equity return	**1.012**	**1.016**	2.126	2.114	3.049	**1.219**	**1.215**	2.139	2.144	3.184
Panel B: European Debt Crisis										
Commodity return	0.839	0.842	**0.829**	0.844	1.158	0.922	0.927	0.828	**0.827**	1.162
Equity return	**0.859**	**0.861**	1.627	1.639	2.207	**0.675**	1.009	0.723	0.733	0.735
Panel C: Covid-19 Pandemic										
Commodity return	1.124	1.120	**0.947**	1.069	1.512	1.057	1.055	**0.948**	1.052	1.343
Equity return	2.333	**1.567**	2.236	2.408	3.423	2.258	**1.259**	2.229	2.316	2.237

This table reports root-mean-squared errors (RMSE) in % across models and crisis periods. Lower values shown in **bold** identify the model with superior forecast accuracy for each return series. MS-FACTOR forecasts are computed with two alternative global risk factors (VIX and PCR). RW denotes the random-walk benchmark. p is the optimal VAR lag length, set to p=3 in Panel A, p=12 in Panel B, p=1 in Panel C.

3.4 Conclusion

This paper revisits the dynamic relationship between commodity and equity markets by analyzing the influence of country-specific commodity shocks—particularly those related to a country's primary export commodities—on domestic equity returns, while controlling for global risk factors. To address limitations in previous empirical approaches, we implement a regime-switching framework that jointly models commodity and equity returns for each market pair; the mean equations incorporate a lagged global risk proxy, specifically the VIX and put–call ratio, while the variance–covariance structure reflects both common and idiosyncratic shocks, each governed by independent Markov-switching processes. Crucially, crisis and normal periods are identified endogenously through filtered regime probabilities rather than imposed ex-ante.

Using data from six major commodity-exporting economies between 2006 and 2023, we assess the existence of contagion across three episodes of financial distress: the Global Financial Crisis (GFC), the European Sovereign Debt Crisis, and the Covid-19 pandemic. Contagion is detected at the 5% significance level in two cases—Norway when the model incorporates the VIX, and in the U.S. when the PCR is used instead both, during the GFC. The MS-FACTOR model also demonstrates strong predictive performance. Compared to popular benchmark models, including VAR (p), DCC-GARCH(1,1), and the random walk, the PCR-based MS-FACTOR generally achieves the lowest root mean squared errors (RMSE) for equity returns at both short-term (1-day ahead) and the medium-term (20-day ahead) horizons, followed by its VIX-based variant. In contrast, forecasting performance for commodity returns displays a different ranking: DCC-GARCH(1,1) and occasionally VAR outperform in short-term horizons, while the MS-FACTOR shows relatively better forecast performances mainly during crisis episodes, such as during the Covid-19 period. Across medium-term horizons, these patterns persist with some mild variation. The random walk model yields the weakest performance across asset types and horizons, reinforcing the value of regime-sensitive structures.

Future research could extend these insights by examining the contagion between commodities and several more asset classes, such as fixed income, foreign exchange, and real estate, which would allow us to develop a broader understanding of the dynamic interaction between these financial assets. In addition, exploring the sector-level dynamics within equity markets, particularly in energy, mining, and manufacturing, would offer a more detailed view of how commodity exposure amplifies crisis spillovers.

3.5 Appendix: Robustness checks and diagnostic fit results of the MS-FACTOR model

3.5.1 Use of AI in research

In this study, OpenAI ChatGPT and Grammarly were used to check the grammar, catch typos, and enhance the writing to improve the clarity of the text.

3.5.2 Robustness check

In this section, we assess the robustness of our model by re-estimating it using the economic policy uncertainty (EPU) index as the global risk.

3.5.2.1 Estimated parameters model over the crisis periods

Tables 3.12–3.14 present the model estimates for the Global Financial Crisis, the European Sovereign Debt Crisis, and the Covid-19 pandemic, respectively. Across the three crisis periods, the impact coefficients of idiosyncratic shocks increase significantly under the high-volatility regime. Compared to our baseline specification, copper and oil still exhibit substantially higher idiosyncratic volatility than their corresponding equity markets in all examined countries. The impact coefficients of common shocks also increase significantly during crisis periods, with magnitudes broadly similar to those obtained in our proposed model.

The estimated loadings of the lagged economic policy uncertainty remain significantly negative overall, reinforcing its persistent influence across markets. Moreover, with the exception of a few countries, the estimated coefficients of the lagged equity return show limited variation in the high-volatility regime. This result is similar to that obtained when incorporating either VIX or PCR as global risk measure into the model.

Table 3.12: Estimated Parameters of the MS–FACTOR Model during the Global Financial Crisis

Parameter	Norway	Mexico	U.S.	Canada	Russia	Chile
Variance Parameters						
$\sigma_{c,1,L}$	**1.210**	**1.161**	0.185	**0.938**	1.435	**0.633**
	(0.098)	(0.599)	(0.319)	(0.143)	(1.337)	(0.110)
$\sigma_{c,2,L}$	**0.878**	0.291	0.434	**0.521**	1.254	**0.799**
	(0.014)	(0.207)	(0.586)	(0.083)	(1.140)	(0.036)
$\sigma_{c,1,H}$	**4.032**	**2.969**	**0.499**	**3.202**	**15.250**	**4.126**
	(0.437)	(0.324)	(0.160)	(0.210)	(2.110)	(0.393)
$\sigma_{c,2,H}$	**2.927**	**2.098**	**1.259**	**1.780**	**5.123**	**1.103**
	(0.269)	(0.241)	(0.209)	(0.149)	(1.259)	(0.039)
$\sigma_{1,L}$	**1.218**	**1.628**	**2.051**	0.189	1.612	**0.616**
	(0.205)	(0.423)	(0.067)	(0.243)	(1.180)	(0.160)
$\sigma_{2,L}$	**1.100**	**0.969**	0.295	**0.830**	1.358	**0.009**
	(0.086)	(0.066)	(0.832)	(0.049)	(1.051)	(0.000)
$\sigma_{1,H}$	**2.176**	**4.325**	**5.643**	**1.953**	**4.714**	**2.162**
	(0.298)	(0.347)	(0.332)	(0.101)	(0.527)	(0.092)
$\sigma_{2,H}$	**2.897**	**1.865**	**3.206**	**3.025**	**5.317**	**1.686**
	(0.246)	(0.267)	(0.232)	(0.257)	(0.455)	(0.121)
Mean Parameters						
λ_1	**-0.338**	**-0.121**	**-0.312**	**-0.026**	**-0.083**	**-0.132**
	(0.001)	(0.001)	(0.021)	(0.001)	(0.002)	(0.041)
λ_2	**-0.015**	**-0.037**	**-0.033**	0.001	**0.003**	**-0.045**
	(0.001)	(0.008)	(0.002)	(0.002)	(0.001)	(0.003)
β_L	**-0.035**	0.050	**-0.128**	-0.003	**-0.086**	**0.705**
	(0.036)	(0.037)	(0.039)	(0.027)	(0.000)	(0.112)
β_H	**-0.820**	**0.147**	**-0.154**	**-0.121**	-0.004	0.055
	(0.089)	(0.066)	(0.085)	(0.054)	(0.104)	(0.047)

For each country, subscripts "1" and "2" denote the commodity and equity markets, respectively. The shock originates in the commodity market. EPU is the global risk included in the mean of the model. Standard deviations are reported in parentheses. Parameters in bold are statistically significant using a two-sided 5% test ($|t|$ >1.96).

Table 3.13: Estimated Parameters of the MS–FACTOR Model during the European Sovereign Debt Crisis

Parameter	Norway	Mexico	U.S.	Canada	Russia	Chile
Variance Parameters						
$\sigma_{c,1,L}$	**1.385**	**0.822**	**1.125**	**0.996**	**1.861**	**0.715**
	(0.056)	(0.134)	(0.074)	(0.028)	(0.044)	(0.101)
$\sigma_{c,2,L}$	**0.800**	**0.488**	**0.683**	**0.460**	**0.802**	**0.538**
	(0.049)	(0.083)	(0.042)	(0.040)	(0.053)	(0.037)
$\sigma_{c,1,H}$	**4.581**	**2.411**	**2.313**	**2.870**	**5.311**	**1.218**
	(0.388)	(0.180)	(0.202)	(0.249)	(0.457)	(0.231)
$\sigma_{c,2,H}$	**2.320**	**1.431**	**2.392**	**1.673**	**1.825**	**1.684**
	(0.281)	(0.114)	(0.124)	(0.127)	(0.471)	(0.158)
$\sigma_{1,L}$	**0.094**	**1.252**	**1.256**	0.002	0.000	**1.295**
	(0.000)	(0.097)	(0.054)	(0.005)	(0.003)	(0.051)
$\sigma_{2,L}$	**0.931**	**0.510**	0.000	**0.698**	**1.116**	0.062
	(0.033)	(0.079)	(0.001)	(0.025)	(0.039)	(0.128)
$\sigma_{1,H}$	**2.176**	**1.745**	**5.698**	**4.993**	**2.226**	**0.086**
	(0.298)	(0.143)	(0.463)	(0.348)	(0.181)	(0.012)
$\sigma_{2,H}$	**4.581**	**1.250**	**0.808**	**1.990**	**2.958**	**0.829**
	(0.108)	(0.093)	(0.059)	(0.157)	(0.160)	(0.072)
Mean Parameters						
λ_1	**-0.084**	**-0.002**	**-0.060**	0.003	**-0.084**	**-0.035**
	(0.009)	(0.001)	(0.004)	(0.007)	(0.001)	(0.001)
λ_2	**-0.083**	**-0.002**	**-0.092**	**-0.065**	**-0.070**	0.000
	(0.002)	(0.001)	(0.005)	(0.001)	(0.001)	(0.000)
β_L	-0.062	0.043	-0.035	-0.038	-0.070	**0.191**
	(0.036)	(0.029)	(0.032)	(0.036)	(0.673)	(0.035)
β_H	0.005	0.077	**-0.169**	0.024	-0.053	**0.128**
	(0.006)	(0.086)	(0.067)	(0.067)	(0.107)	(0.056)

For each country, subscripts "1" and "2" denote the commodity and equity markets, respectively. The shock originates in the commodity market. EPU is the global risk included in the mean of the model. Standard deviations are reported in parentheses. Parameters in bold are statistically significant using a two-sided 5% test ($|t|$ >1.96).

Table 3.14: Estimated Parameters of the MS–FACTOR Model during the Covid-19 Pandemic

Parameter	Norway	Mexico	U.S.	Canada	Russia	Chile
Variance Parameters						
$\sigma_{c,1,L}$	**1.497**	**1.625**	**1.131**	**0.020**	**0.816**	**0.611**
	(0.379)	(0.089)	(0.475)	(0.000)	(0.062)	(0.062)
$\sigma_{c,2,L}$	0.038	**0.230**	**0.222**	0.069	**1.060**	**0.734**
	(0.090)	(0.038)	(0.107)	(0.158)	(0.060)	(0.112)
$\sigma_{c,1,H}$	**2.059**	**3.644**	**4.712**	**1.520**	**1.642**	**1.045**
	(0.230)	(0.939)	(1.015)	(0.053)	(0.119)	(0.109)
$\sigma_{c,2,H}$	**1.170**	**0.472**	**0.926**	**0.273**	**2.641**	**4.992**
	(0.280)	(0.080)	(0.328)	(0.034)	(0.140)	(0.722)
$\sigma_{1,L}$	0.776	0.005	**1.223**	**0.010**	**1.653**	**0.813**
	(0.619)	(0.071)	(0.112)	(0.000)	(0.054)	(0.092)
$\sigma_{2,L}$	**0.692**	**0.925**	**0.502**	**0.617**	0.0003	0.005
	(0.034)	(0.033)	(0.065)	(0.020)	(0.004)	(0.038)
$\sigma_{1,H}$	**2.836**	**7.459**	**19.829**	**4.817**	**4.841**	**1.833**
	(0.128)	(0.697)	(3.421)	(0.369)	(0.388)	(0.138)
$\sigma_{2,H}$	**2.503**	**2.237**	**1.180**	**5.816**	**3.395**	**1.367**
	(0.238)	(0.204)	(0.111)	(0.913)	(0.437)	(0.104)
Mean Parameters						
λ_1	0.000	**-0.046**	0.001	0.003	**-0.011**	0.000
	(0.001)	(0.001)	(0.002)	(0.004)	(0.002)	(0.001)
λ_2	**-0.052**	**-0.063**	**-0.093**	0.005	-0.001	-0.000
	(0.001)	(0.001)	(0.001)	(0.004)	(0.002)	(0.001)
β_L	**-0.097**	-0.000	-0.065	**-0.124**	-0.020	**0.168**
	(0.039)	(0.039)	(0.066)	(0.045)	(0.029)	(0.080)
β_H	-0.112	-0.011	**0.777**	**1.306**	**-0.138**	0.044
	(0.099)	(0.064)	(0.853)	(0.097)	(0.091)	(0.070)

For each country, subscripts "1" and "2" denote the commodity and equity markets, respectively. The shock originates in the commodity market. EPU is the global risk included in the mean of the model. Standard deviations are reported in parentheses. Parameters in bold are statistically significant using a two-sided 5% test ($|t|$ >1.96).

3.5.2.2 Likelihood ratio test

Table 3.15 reports the results of the likelihood ratio test. Evidence of contagion at the 5% significance level is found only for the oil–equity market pair in Norway during the GFC. For the other countries, no statistically significant signs of contagion are detected in the sample periods.

Table 3.15: Likelihood Ratio Test for Contagion: EPU as Global Risk Factor

	Norway	**Mexico**	**U.S.**	**Canada**	**Russia**	**Chile**
Panel A: Global Financial Crisis						
LR	**8.540**	0.008	0.169	0.005	0.001	0.016
p-value	**0.003**	0.799	0.459	0.946	0.974	0.678
Panel B: European Sovereign Debt Crisis						
LR	0.004	0.160	0.006	0.005	0.263	0.960
p-value	0.888	0.895	0.940	0.947	0.308	0.127
Panel C: Covid-19 Pandemic						
LR	0.004	0.005	0.233	0.252	0.113	0.118
p-value	0.888	0.945	0.328	0.316	0.600	0.598

The likelihood-ratio statistic (LR) tests the null of no contagion against the alternative of contagion for each market pair across the crisis periods; corresponding p-values are reported. Underlined bold values are statistically significant.

Panel C: Estimated filtered probabilities of high volatility shocks — Oil-equity market (Norway) during the GFC; U.S. EPU used in the model

Figure 3.7: Common shock

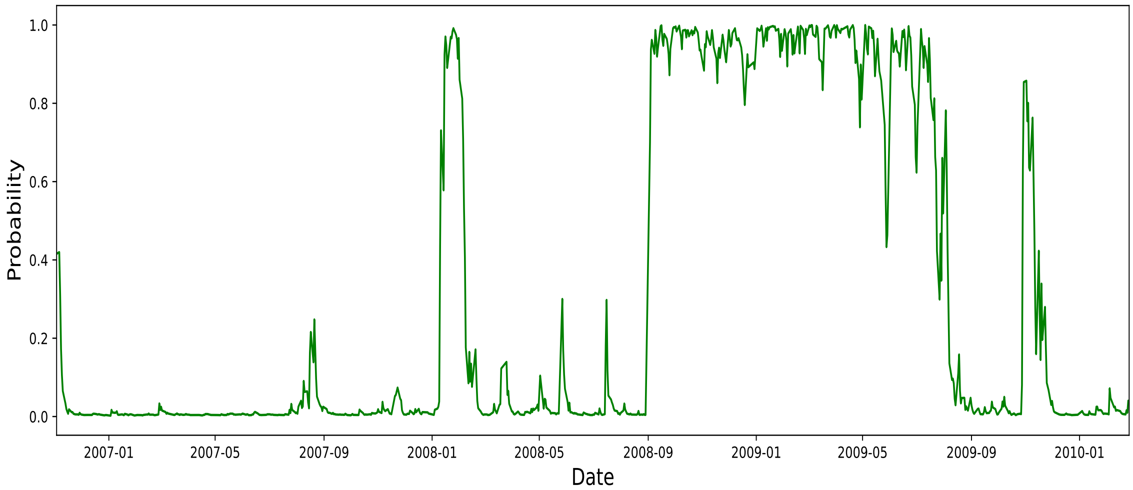


Figure 3.8: Idiosyncratic Shock, Norway Oil market (Shock origin)

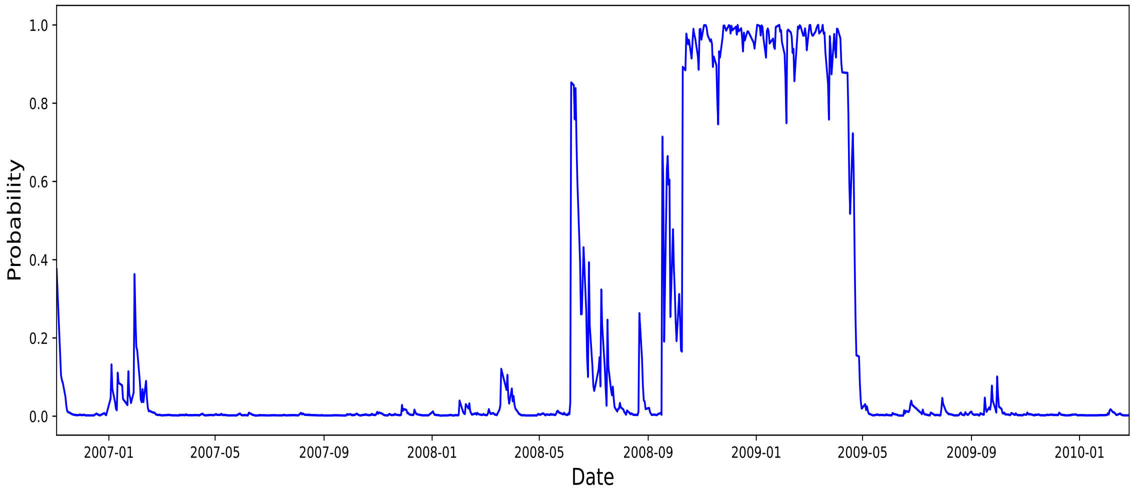
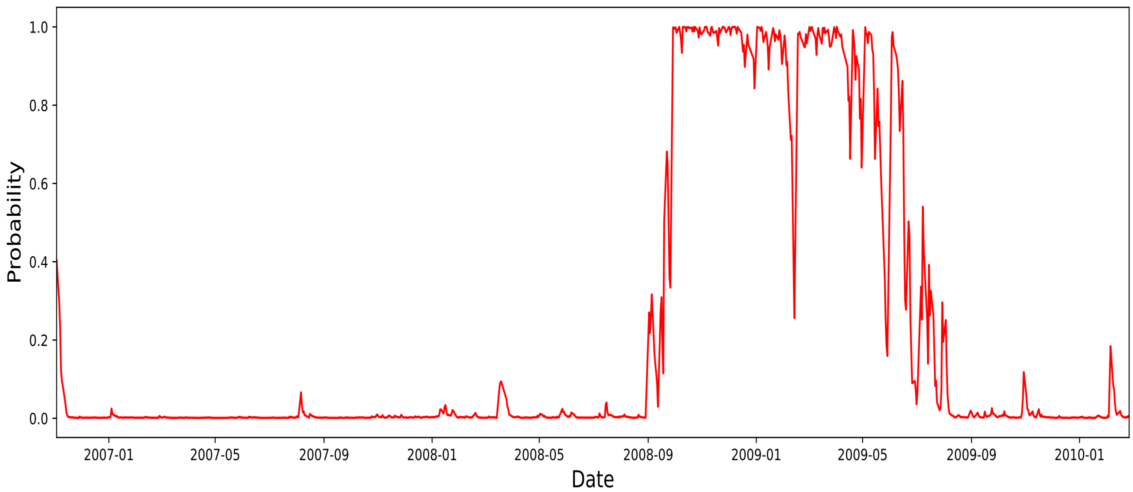


Figure 3.9: Idiosyncratic Shock, Norway Equity market (Shock target)



3.5.3 Diagnostic checks of the MS-FACTOR model

Table 3.16: Diagnostic Checks; Verifying Fit of MS-FACTOR Model (VIX)

Panel A1: Global Financial Crisis						
Test	Norway	Mexico	U.S.	Canada	Russia	Chile
$Q_1(1)$	0.066	0.699	0.174	0.064	0.619	0.065
$Q_1(4)$	0.106	0.748	0.291	0.196	0.759	0.060
$Q_2(1)$	0.438	0.645	0.849	0.444	0.357	0.858
$Q_2(4)$	0.603	0.363	0.792	0.903	0.821	0.368
JB_1	0.733	0.449	0.816	0.125	0.120	0.661
JB_2	0.065	0.709	0.077	0.069	**0.001**	0.393
$ARCH_1(4)$	0.541	0.417	0.187	0.537	0.059	0.819
$ARCH_2(4)$	0.770	0.649	**0.032**	0.065	0.054	0.098
Panel A2: European Sovereign Debt Crisis						
$Q_1(1)$	0.374	0.101	0.927	0.064	0.310	0.073
$Q_1(4)$	0.876	0.284	0.973	0.222	0.621	0.329
$Q_2(1)$	0.952	0.852	0.825	0.330	0.831	0.761
$Q_2(4)$	0.834	0.835	0.886	0.494	0.870	0.152
JB_1	0.721	0.063	0.661	0.061	0.269	0.680
JB_2	0.058	**0.004**	**0.020**	0.060	0.066	0.298
$ARCH_1(4)$	0.228	0.059	0.087	0.067	0.811	0.101
$ARCH_2(4)$	0.330	0.060	0.829	0.070	0.077	0.108
Panel A3: Covid-19 Crisis						
$Q_1(1)$	0.371	0.180	0.695	0.076	0.175	0.355
$Q_1(4)$	0.427	0.062	0.060	0.242	0.452	0.553
$Q_2(1)$	0.543	0.789	0.854	0.749	0.934	0.928
$Q_2(4)$	0.733	0.503	0.385	0.615	0.829	0.237
JB_1	0.068	**0.011**	0.090	0.080	0.068	0.167
JB_2	0.066	0.467	0.069	0.072	0.951	0.090
$ARCH_1(4)$	0.350	**0.045**	0.376	**0.031**	0.481	0.159
$ARCH_2(4)$	**0.018**	0.072	0.059	0.052	0.687	0.080

This table reports the p-values of several diagnostic tests for each market pair during the Global Financial Crisis, European Sovereign Debt Crisis, and the Covid-19 pandemic. For each country, subscripts "1" and "2" refer to commodity and equity markets, respectively. $Q_1(k)$ and $Q_2(k)$ denote the the Ljung-Box test for the null of no autocorrelation up to lag k. JB_1 and JB_2 represent the Jarque-Bera test for the null of normality. $ARCH_1(4)$ and $ARCH_2(4)$ are the Lagrange Multiplier test for the null of no ARCH effects of order 4. P-values $< 5\%$ are in bold.

Table 3.17: Diagnostic Checks; Verifying Fit of MS-FACTOR Model (PCR)

Panel A1: Global Financial Crisis						
Test	Norway	Mexico	U.S.	Canada	Russia	Chile
$Q_1(1)$	0.151	0.994	0.334	0.064	0.684	0.067
$Q_1(4)$	0.311	0.452	0.427	0.377	0.690	0.115
$Q_2(1)$	0.902	0.301	0.941	0.126	0.997	0.536
$Q_2(4)$	0.599	0.252	0.767	0.440	0.961	0.360
JB_1	0.831	0.663	0.579	0.117	0.084	0.583
JB_2	0.068	0.768	**0.017**	0.069	0.730	**0.011**
$ARCH_1(4)$	0.418	0.518	0.486	0.350	0.319	0.144
$ARCH_2(4)$	**0.004**	0.522	0.075	0.066	0.066	0.061
Panel A2: European Sovereign Debt Crisis						
$Q_1(1)$	0.824	0.128	0.886	0.402	0.770	0.074
$Q_1(4)$	0.947	0.339	0.960	0.842	0.788	0.361
$Q_2(1)$	0.828	0.794	0.919	0.060	0.529	0.716
$Q_2(4)$	0.494	0.867	0.951	0.193	0.913	0.142
JB_1	0.870	0.088	0.551	0.064	0.254	0.670
JB_2	0.412	0.062	0.066	0.074	0.077	0.472
$ARCH_1(4)$	0.088	0.100	0.088	0.080	0.777	0.122
$ARCH_2(4)$	0.078	**0.002**	0.132	0.088	0.055	0.077
Panel A3: Covid-19 Crisis						
$Q_1(1)$	0.330	0.563	0.700	0.067	0.225	0.480
$Q_1(4)$	0.390	0.062	0.072	0.147	0.560	0.539
$Q_2(1)$	0.451	0.584	0.817	0.944	0.911	0.586
$Q_2(4)$	0.719	0.583	0.222	0.855	0.660	0.146
JB_1	0.090	0.124	0.088	0.070	0.076	0.319
JB_2	**0.010**	0.316	0.065	0.069	0.104	0.072
$ARCH_1(4)$	0.210	0.104	0.616	0.377	0.634	0.075
$ARCH_2(4)$	0.075	0.055	0.701	0.094	0.234	0.088

This table reports the p-values of several diagnostic tests for each market pair during the Global Financial Crisis, European Sovereign Debt Crisis, and the Covid-19 pandemic. For each country, subscripts "1" and "2" refer to commodity and equity markets, respectively. Q_1(k) and Q_2(k) denote the the Ljung-Box test for the null of no autocorrelation up to lag k. JB_1 and JB_2 represent the Jarque-Bera test for the null of normality. $ARCH_1(4)$ and $ARCH_2(4)$ are the Lagrange Multiplier test for the null of no ARCH effects of order 4. P-values $< 5\%$ are in bold.

Table 3.18: Diagnostic Checks (continued); Verifying Fit of MS-FACTOR Model (PCR)

Panel A1: Global Financial Crisis						
Test	Norway	Mexico	U.S.	Canada	Russia	Chile
$Q_1(30)$	0.666	0.781	0.895	0.927	0.843	0.336
$Q_1(90)$	0.727	0.551	0.301	0.777	0.925	0.284
$Q_2(30)$	0.825	0.945	0.654	0.441	0.963	0.572
$Q_2(90)$	0.212	0.247	0.788	0.338	0.634	0.278
Panel A2: European Sovereign Debt Crisis						
$Q_1(30)$	0.878	0.342	0.680	0.897	0.842	0.876
$Q_1(90)$	0.714	0.433	0.841	0.661	0.891	0.946
$Q_2(30)$	0.397	0.971	0.842	0.150	0.431	0.065
$Q_2(90)$	0.071	0.870	0.495	0.321	0.532	0.188
Panel A3: Covid-19 Crisis						
$Q_1(30)$	0.884	**0.012**	0.107	0.216	0.814	0.522
$Q_1(90)$	0.890	0.127	**0.014**	0.517	0.699	0.374
$Q_2(30)$	0.452	0.473	0.807	0.141	**0.042**	0.163
$Q_2(90)$	0.223	0.062	0.947	0.441	0.184	0.288

This table reports the p-values of autocorrelation tests for each market pair during the Global Financial Crisis, European Sovereign Debt Crisis, and the Covid-19 pandemic. For each country, subscripts "1" and "2" refer to commodity and equity markets, respectively. $Q_1(k)$ and $Q_2(k)$ denote the the Ljung-Box test for the null of no autocorrelation up to lag k=$\{30, 90\}$. P-values < 5% are in bold.

Table 3.19: Diagnostic Checks (continued); Verifying Fit of MS-FACTOR Model (VIX)

Panel A1: Global Financial Crisis						
Test	Norway	Mexico	U.S.	Canada	Russia	Chile
$Q_1(30)$	0.506	0.726	0.804	0.912	0.632	0.196
$Q_1(90)$	0.628	0.462	0.307	0.719	0.693	0.211
$Q_2(30)$	0.700	0.898	0.755	0.731	0.847	0.627
$Q_2(90)$	0.289	0.209	0.817	0.544	0.422	0.287
Panel A2: European Sovereign Debt Crisis						
$Q_1(30)$	0.910	0.308	0.643	0.718	0.842	0.869
$Q_1(90)$	0.813	0.418	0.811	0.277	0.897	0.949
$Q_2(30)$	0.575	0.970	0.843	0.254	0.562	0.062
$Q_2(90)$	**0.029**	0.876	0.446	0.600	0.404	0.206
Panel A3: Covid-19 Crisis						
$Q_1(30)$	0.889	**0.011**	0.086	0.418	0.662	0.631
$Q_1(90)$	0.949	0.296	0.069	0.717	0.779	0.435
$Q_2(30)$	0.247	0.527	0.870	**0.018**	0.415	0.191
$Q_2(90)$	**0.040**	0.137	0.989	0.103	0.636	0.357

This table reports the p-values of autocorrelation tests for each market pair during the Global Financial Crisis, European Sovereign Debt Crisis, and the Covid-19 pandemic. For each country, subscripts "1" and "2" refer to commodity and equity markets, respectively. $Q_1(k)$ and $Q_2(k)$ denote the the Ljung-Box test for the null of no autocorrelation up to lag k=$\{30, 90\}$. P-values < 5% are in bold.

Figure 3.10: Standardized Residuals; Model with VIX; GFC period

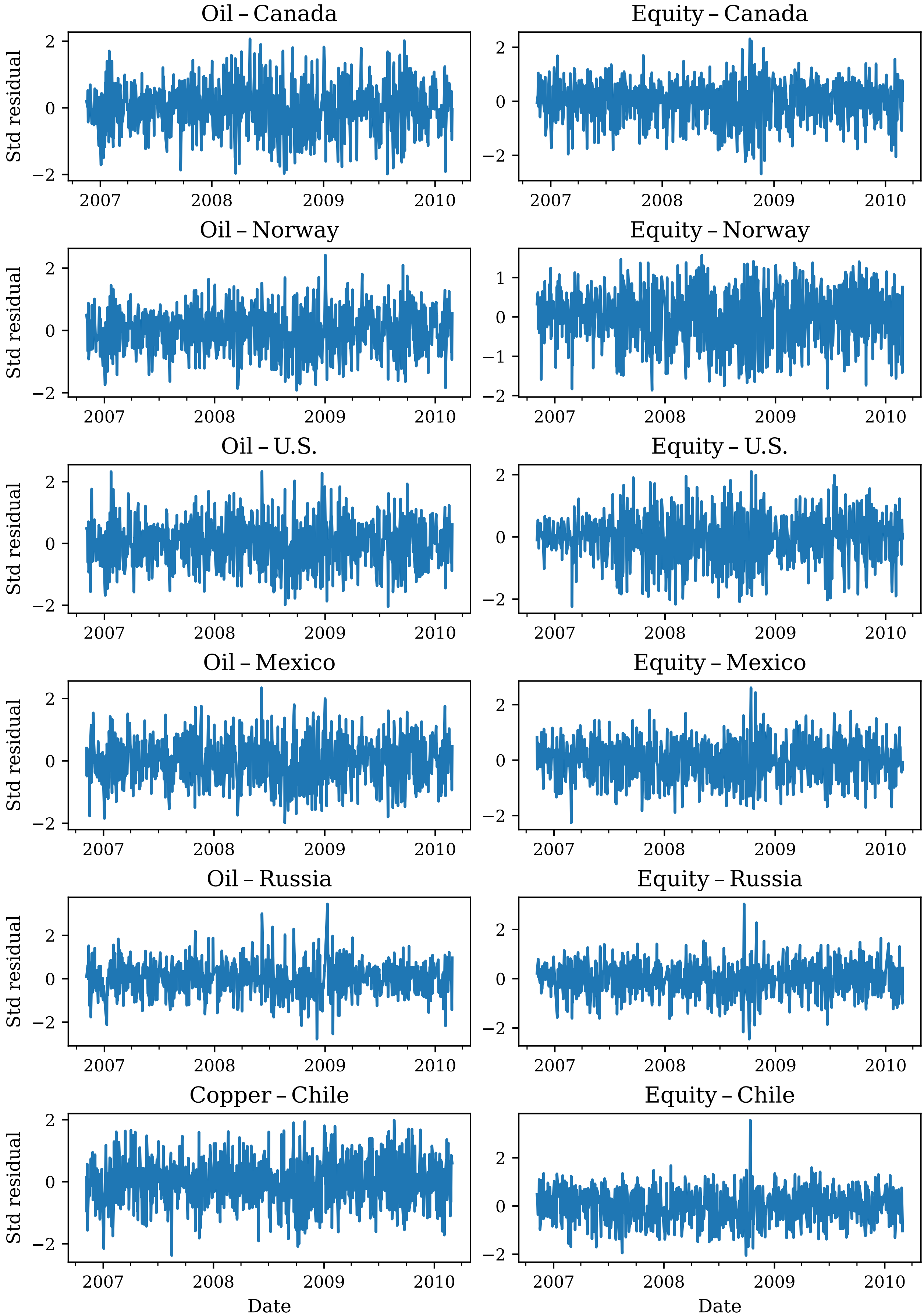


Figure 3.11: Standardized Residuals; Model with PCR; GFC period

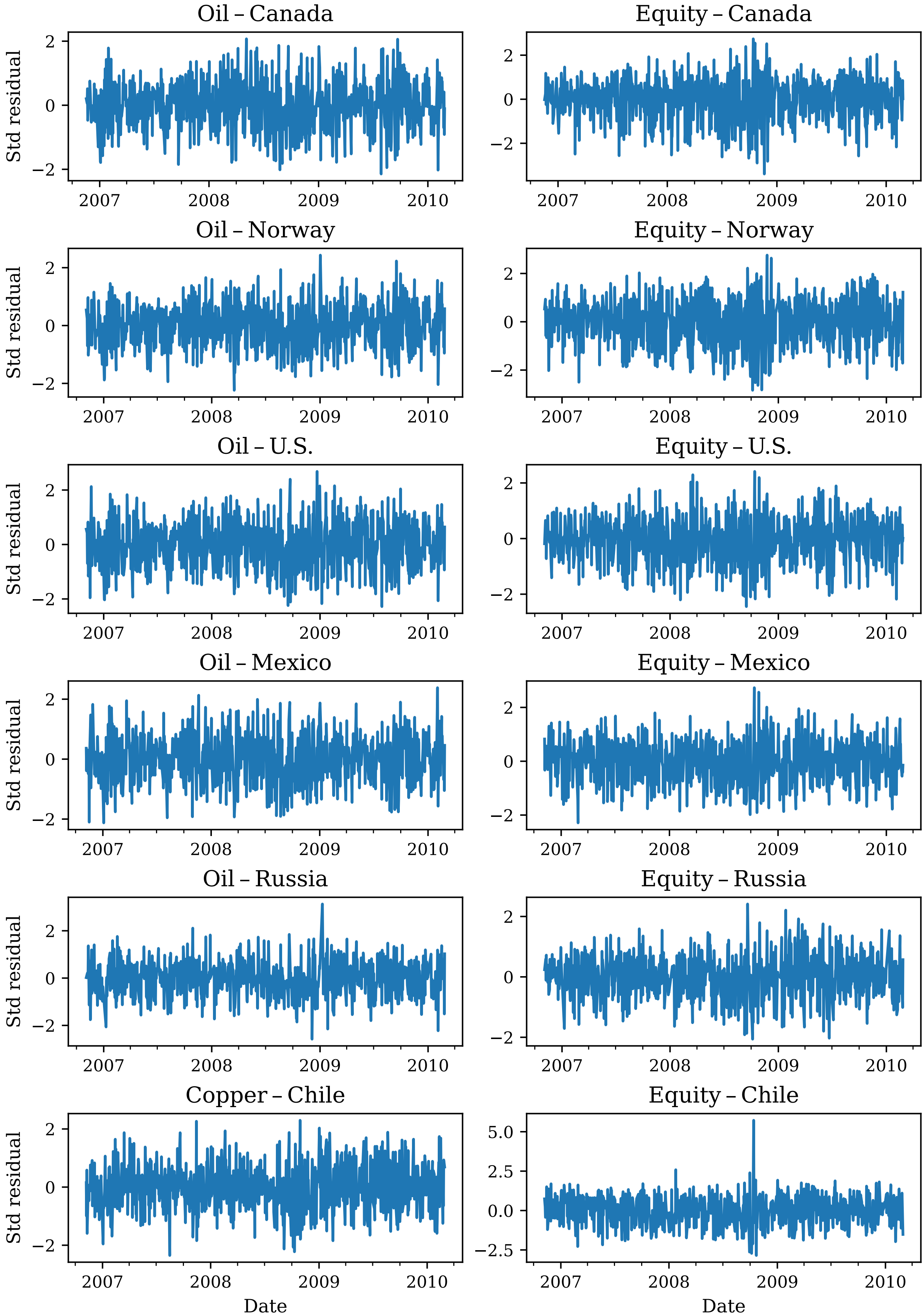


Figure 3.12: Standardized Residuals; Model with VIX; ESD period

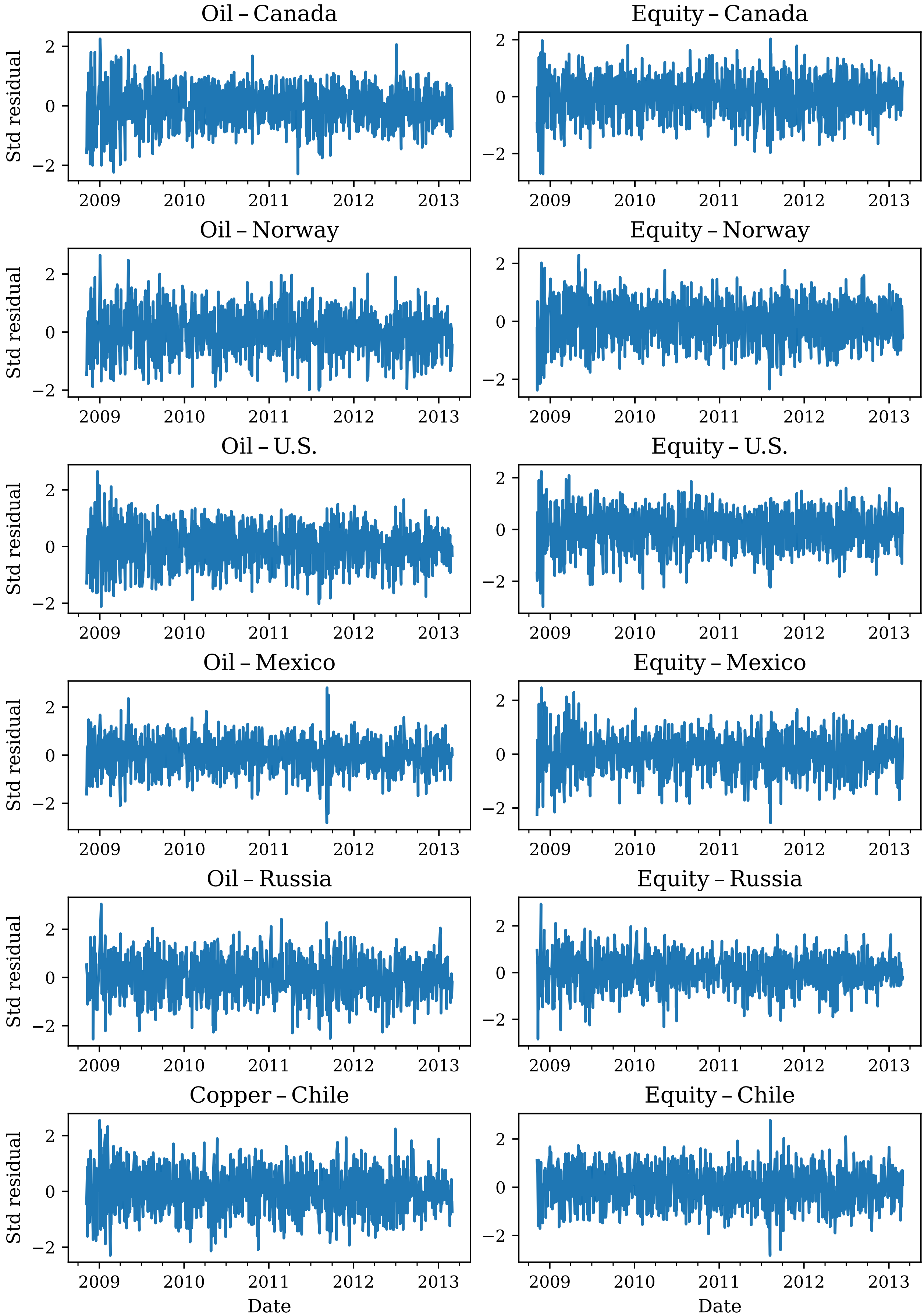


Figure 3.13: Standardized Residuals; Model with PCR; ESD period

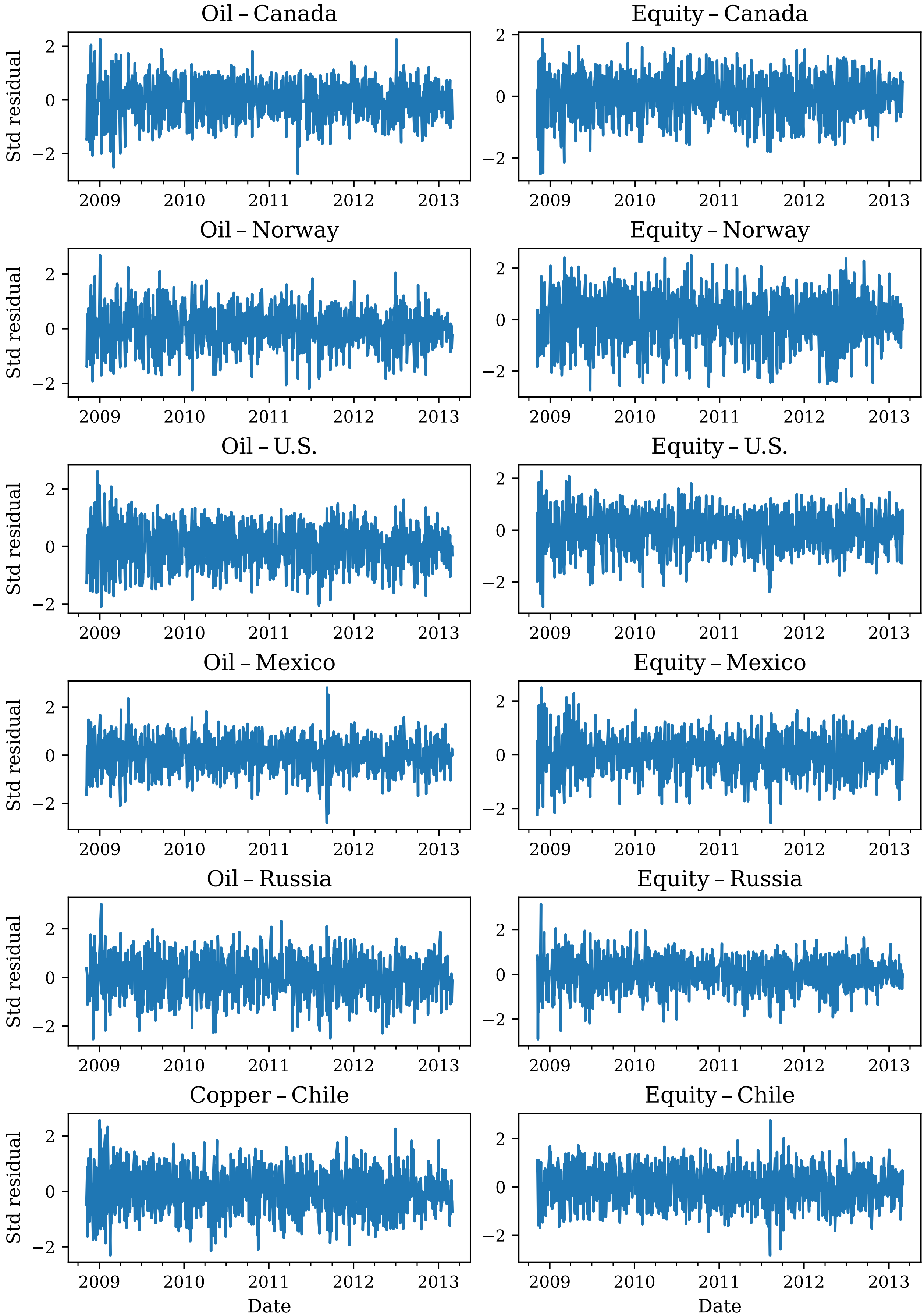


Figure 3.14: Standardized Residuals; Model with VIX;Covid19 pandemic

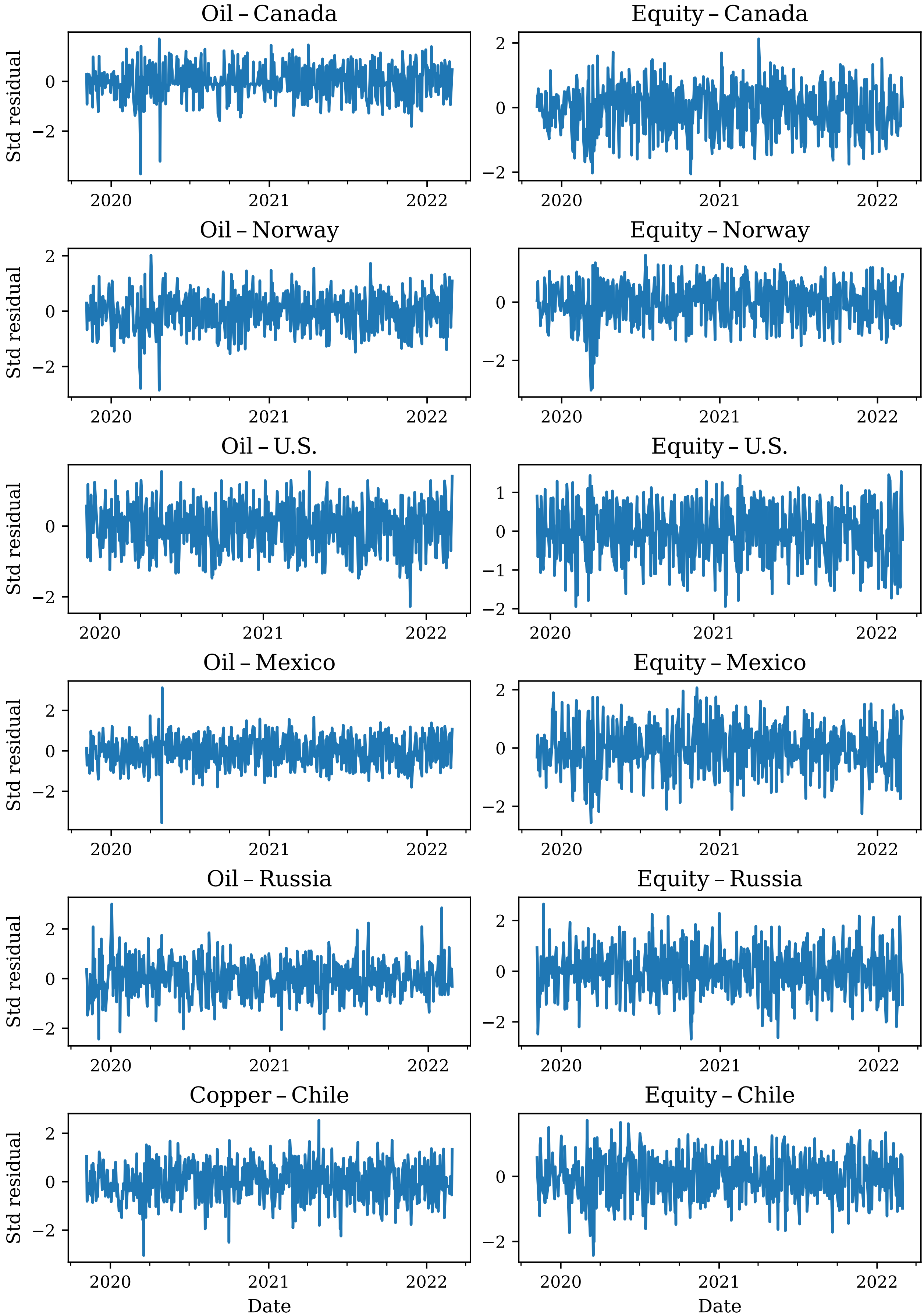
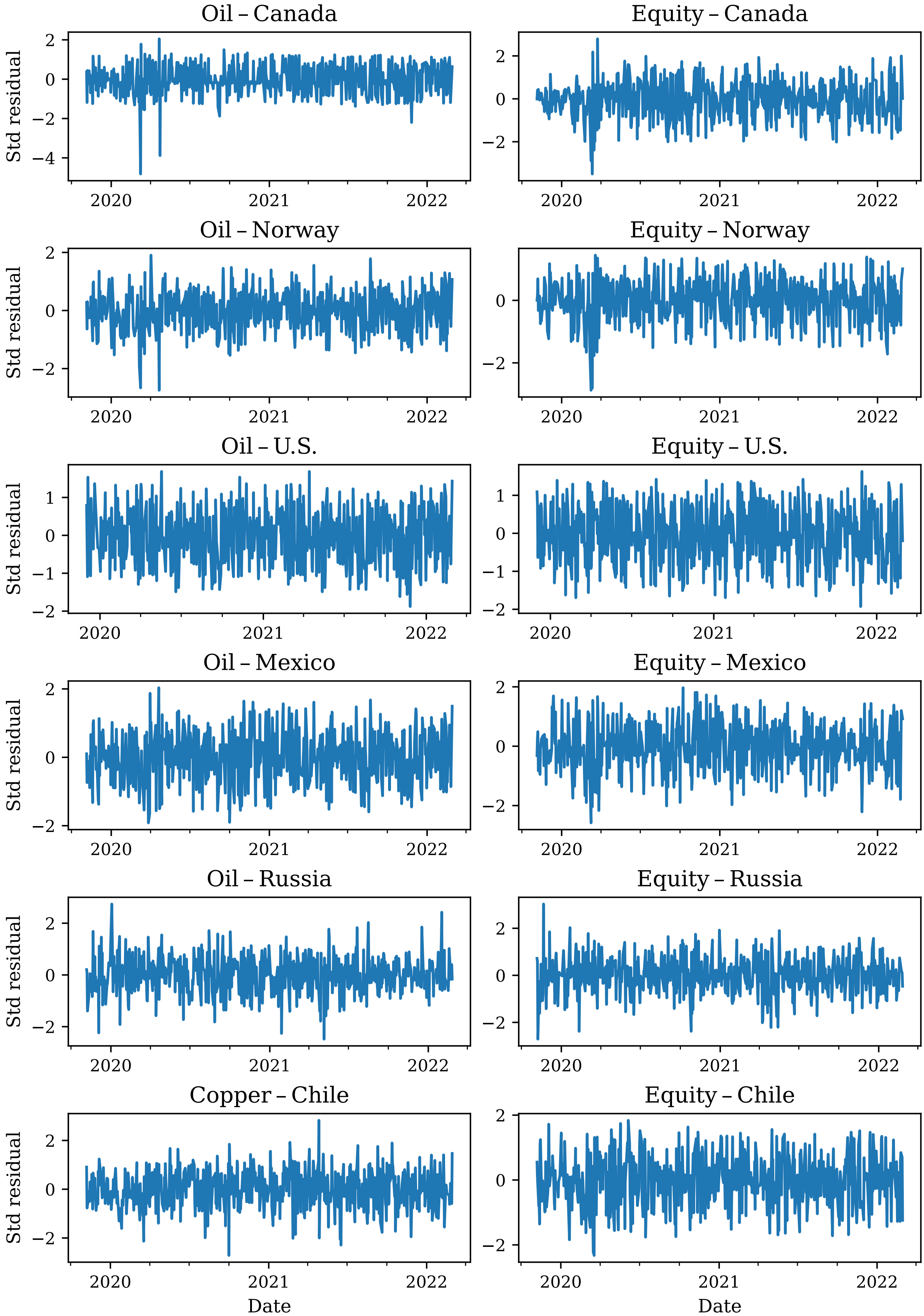


Figure 3.15: Standardized Residuals; Model with PCR;Covid19 pandemic



Bibliography

Ahmed, A. D., & Huo, R. (2021). Volatility transmissions across international oil market, commodity futures and stock markets: Empirical evidence from China. *Energy Economics*, *93*, 104741.

Akıncı, Ö. (2013). Global financial conditions, country spreads and macroeconomic fluctuations in emerging countries. *Journal of International Economics*, *91*(2), 358–371.

Amihud, Y. (2019). Illiquidity and stock returns: A revisit. critical finance review 8 (1–2): 203–221.

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of financial markets*, *5*(1), 31–56.

Ang, A., & Longstaff, F. A. (2013). Systemic sovereign credit risk: Lessons from the U.S. and Europe. *Journal of Monetary Economics*, *60*(5), 493–510.

Arezki, R., Ramey, V. A., & Sheng, L. (2017). News shocks in open economies: Evidence from giant oil discoveries. *The Quarterly Journal of Economics*, *132*(1), 103–155.

Bae, K.-H., Karolyi, G. A., & Stulz, R. M. (2003). A new approach to measuring financial contagion. *The Review of Financial Studies*, *16*(3), 717–763.

Baek, I.-M., Bandopadhyaya, A., & Du, C. (2005). Determinants of market-assessed sovereign risk: Economic fundamentals or market risk appetite? *Journal of International Money and Finance*, *24*(4), 533–548.

Baker & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *The Journal of Finance*, *61*(4), 1645–1680.

Baker, Bloom, N., & Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, *131*(4), 1593–1636.

Bakshi, G., Gao, X., & Rossi, A. G. (2019). Understanding the sources of risk underlying the cross section of commodity returns. *Management Science*, *65*(2), 619–641.

Baruník, J., & Křehlík, T. (2018). Measuring the frequency dynamics of financial connectedness and systemic risk. *Journal of Financial Econometrics*, *16*(2), 271–296.

Basher, S. A., Haug, A. A., & Sadorsky, P. (2012). Oil prices, exchange rates and emerging stock markets. *Energy Economics*, *34*(1), 227–240.

Bathia, D., & Bredin, D. (2016). An examination of investor sentiment effect on G7 stock market returns. In *Contemporary issues in financial institutions and markets* (pp. 99–128). Routledge.

Bekaert, G., & Harvey, C. R. (2003). *Market integration and contagion* (Working Paper No. 9510). National Bureau of Economic Research.

Bekaert, G., Ehrmann, M., Fratzscher, M., & Mehl, A. (2014). The global crisis and equity market contagion. *The Journal of Finance*, *69*(6), 2597–2649.

Bianchi, D., Büchner, M., & Tamoni, A. (2020). Bond risk premiums with machine learning. *The Review of Financial Studies*, *34*(2), 1046–1089.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327.

Bouri, E., Lei, X., Jalkh, N., Xu, Y., & Zhang, H. (2021). Spillovers in higher moments and jumps across us stock and strategic commodity markets. *Resources Policy*, *72*, 102060.

Brennan, M. J. (1976). The supply of storage. In *The economics of futures trading* (pp. 100–107). Springer.

Brennan, M. J., Wang, A. W., & Xia, Y. (2004). Estimation and test of a simple model of intertemporal capital asset pricing. *The Journal of Finance*, *59*(4), 1743–1776.

Butaru, F., Chen, Q., Clark, B., Das, S., Lo, A. W., & Siddique, A. (2016). Risk and risk management in the credit card industry. *Journal of Banking & Finance*, *72*, 218–239.

Büyük şahin, B., Lee, T. K., Moser, J. T., & Robe, M. A. (2013). Physical markets, paper markets and the wti-brent spread. *The Energy Journal*, *34*(3), 129–152.

Çakmaklı, C., & van Dijk, D. (2016). Getting the most out of macroeconomic information for predicting excess stock returns. *International Journal of Forecasting*, *32*(3), 650–668.

Campbell, J. Y., & Thompson, S. B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? *The Review of Financial Studies*, *21*(4), 1509–1531.

Campos-Martins, S., & Amado, C. (2022). Financial market linkages and the sovereign debt crisis. *Journal of International Money and Finance*, *123*, 102596.

Chang, Y., Herrera, A. M., & Pesavento, E. (2023). Oil prices uncertainty, endogenous regime switching, and inflation anchoring. *Journal of Applied Econometrics*, *38*(6), 820–839.

Cheng, I.-H., & Xiong, W. (2014). Financialization of commodity markets. *Annual Review of Financial Economics*, *6*(1), 419–441.

Cheng, I.-H., Kirilenko, A., & Xiong, W. (2015). Convective risk flows in commodity futures markets. *Review of Finance*, *19*(5), 1733–1781.

Claessens, S., & Kose, M. A. (2018). *Frontiers of macrofinancial linkages* (BIS Papers No. 95). Bank for International Settlements.

Claessens, S., Dell'Ariccia, G., Igan, D., & Laeven, L. (2010). Cross-country experiences and policy implications from the global financial crisis. *Economic Policy*, *25*(62), 267–293.

Cochrane, J. H., & Piazzesi, M. (2005). Bond risk premia. *The American Economic Review*, *95*(1), 138–160.

Congress, U. (2009). American recovery and reinvestment act of 2009.

Connolly, E., & Orsmond, D. (2011). The mining industry: From bust to boom | conference–2011 [Available from the Reserve Bank of Australia].

Coudert, V., & Gex, M. (2008). Does risk aversion drive financial crises? testing the predictive power of empirical indicators. *Journal of Empirical Finance*, *15*(2), 167–184.

Creti, A., Joëts, M., & Mignon, V. (2013). On the links between stock and commodity markets' volatility. *Energy Economics*, *37*, 16–28.

Dai, Z., Zhu, H., & Kang, J. (2021). New technical indicators and stock returns predictability. *International Review of Economics & Finance*, *71*, 127–142.

Daskalaki, C., Kostakis, A., & Skiadopoulos, G. (2014). Are there common factors in individual commodity futures returns? *Journal of Banking & Finance*, *40*, 346–363.

Dey, S., & Sampath, A. (2020). Returns, volatility and spillover–a paradigm shift in india? *The North American Journal of Economics and Finance*, *52*, 101110.

Diebold, F. X., & Yilmaz, K. (2009). Measuring financial asset return and volatility spillovers, with application to global equity markets. *The Economic Journal*, *119*(534), 158–171.

Diebold, F. X., & Yilmaz, K. (2012). Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of Forecasting*, *28*(1), 57–66.

Do, A., Powell, R., Singh, A., & Yong, J. (2018). When did the global financial crisis start and end? *Proceedings of 3rd Business Doctoral and Emerging Scholars Conference, Perth, Australia*.

Dooley, M., & Hutchison, M. (2009). Transmission of the *U.S.* subprime crisis to emerging markets: Evidence on the decoupling–recoupling hypothesis. *Journal of International Money and Finance*, *28*(8), 1331–1349.

Dungey, M., & Renault, E. (2018). Identifying contagion. *Journal of Applied Econometrics*, *33*(2), 227–250.

Dungey, M., Fry, R., & Martin, V. L. (2004). Currency market contagion in the ASIA-PACIFIC REGION. *Australian Economic Papers.*, *43*(4).

Dungey, M., Flavin, T. J., & Lagoa-Varela, D. (2020). Are banking shocks contagious? evidence from the eurozone. *Journal of Banking & Finance*, *112*, 105386.

EIA, U. (2008). Petroleum other liquids – WTI spot price.

Erb, C. B., & Harvey, C. R. (2006). The strategic and tactical value of commodity futures. *Financial Analysts Journal*, *62*(2), 69–97.

Fama, E. F., & French, K. R. (2008). Dissecting anomalies. *The Journal of Finance*, *63*(4), 1653–1678.

Fama, E. F., & French, K. R. (2016). Commodity futures prices: Some evidence on forecast power, premiums, and the theory of storage. In *The World Scientific Handbook of Futures Markets* (pp. 79–102). World Scientific.

Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, *81*(3), 607–636.

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

Forni, M., Hallin, M., Lippi, M., & Reichlin, L. (2005). The generalized dynamic factor model: One-sided estimation and forecasting. *Journal of the American Statistical Association*, *100*(471), 830–840.

Forni, M., Hallin, M., Lippi, M., & Zaffaroni, P. (2015). Dynamic factor models with infinite-dimensional factor spaces: One-sided representations. *Journal of Econometrics*, *185*(2), 359–371.

Forni, M., Hallin, M., Lippi, M., & Zaffaroni, P. (2017). Dynamic factor models with infinite-dimensional factor space: Asymptotic analysis. *Journal of Econometrics*, *199*(1), 74–92.

Freyberger, J., Neuhierl, A., & Weber, M. (2020). Dissecting characteristics nonparametrically. *The Review of Financial Studies*, *33*(5), 2326–2377.

Fry, R., Martin, V. L., & Tang, C. (2010). A new class of tests of contagion with applications. *Journal of Business & Economic Statistics*, *28*(3), 423–437.

Fry-McKibbin, R., & Hsiao, C. Y.-L. (2018). Extremal dependence tests for contagion. *Econometric Reviews*, *37*(6), 626–649.

G20. (2009). G20 london summit 2009.

Giglio, S., & Xiu, D. (2021). Asset pricing with omitted factors. *Journal of Political Economy*, *129*(7), 1947–1990.

Gilchrist, S., Yue, V., & Zakrajšek, E. (2019). U.S. monetary policy and international bond markets. *Journal of Money, Credit and Banking*, *51*(S1), 127–161.

Gilchrist, S., Wei, B., Yue, V. Z., & Zakrajšek, E. (2022). Sovereign risk and financial risk. *Journal of International Economics*, *136*, 103603.

Giovannelli, A., Massacci, D., & Soccorsi, S. (2021). Forecasting stock returns with large dimensional factor models. *Journal of Empirical Finance*, *63*, 252–269.

Glick, R., & Rose, A. K. (1999). Contagion and trade: Why are currency crises regional? *Journal of International Money and Finance*, *18*(4), 603–617.

Gravelle, T., Kichian, M., & Morley, J. (2006). Detecting shift-contagion in currency and bond markets. *Journal of International Economics*, *68*(2), 409–423.

Green, J., Hand, J. R., & Zhang, X. F. (2017). The characteristics that provide independent information about average us monthly stock returns. *The Review of Financial Studies*, *30*(12), 4389–4436.

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies*, *33*(5), 2223–2273.

Hamilton, J. D. (2009). *Causes and consequences of the oil shock of 2007-08* (tech. rep.). National Bureau of Economic Research.

Han, K. C., Lee, S. H., & Suk, D. Y. (2003). Mexican peso crisis and its spillover effects to emerging market debt. *Emerging Markets Review*, *4*(3), 310–326.

Hasan, S., & Mahbobi, M. (2013). The increasing influence of oil prices on the canadian stock market. *The International Journal of Business and Finance Research*, *7*(3), 27–39.

Hillier, D., & Loncan, T. (2019). Political uncertainty and stock returns: Evidence from the brazilian political crisis. *Pacific-Basin Finance Journal*, *54*, 1–12.

Hutchinson, J. M., Lo, A. W., & Poggio, T. (1994). A nonparametric approach to pricing and hedging derivative securities via learning networks. *The Journal of Finance*, *49*(3), 851–889.

IMF, A. (2020). World economic outlook: The great lockdown. *International Monetary Fund: Washington.*

Iturrieta, F., & Federowski, B. (2017). Emerging markets—Chile stocks see highest rise in two years on copper prices [Accessed July 13, 2025]. *Reuters.*

Junttila, J., & Vataja, J. (2018). Economic policy uncertainty effects for forecasting future real economic activity. *Economic Systems*, *42*(4), 569–583.

Kaminsky, G. L., & Reinhart, C. M. (2000). On crises, contagion, and confusion. *Journal of International Economics*, *51*(1), 145–168.

Karafiath, I., Mynatt, R., & Smith, K. L. (1991). The Brazilian default announcement and the contagion effect hypothesis. *Journal of Banking & Finance.*, *15*(3).

Kasahara, H., & Shimotsu, K. (2018). Testing the number of regimes in Markov regime switching models. *arXiv preprint arXiv*.

Kat, H. M., & Oomen, R. C. (2006). What every investor should know about commodities, part ii: Multivariate return analysis. *Alternative Investment Research Centre Working Paper*, (33).

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

Kelly, Pruitt, S., & Su, Y. (2019). Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics*, *134*(3), 501–524.

Kelly, Malamud, S., & Zhou, K. (2024). The virtue of complexity in return prediction. *The Journal of Finance*, *79*(1), 459–503.

Khandani, A. E., Kim, A. J., & Lo, A. W. (2010). Consumer credit-risk models via machine-learning algorithms. *Journal of Banking & Finance*, *34*(11), 2767–2787.

Kilian, L., & Park, C. (2009). The impact of oil price shocks on the us stock market. *International Economic Review*, *50*(4), 1267–1287.

Kohn, D. L., & Sack, B. (2018). Monetary policy during the financial crisis. *Terwin Money Management LLC; Bank of America Merrill Lynch*.

Kozak, S., Nagel, S., & Santosh, S. (2020). Shrinking the cross-section. *Journal of Financial Economics*, *135*(2), 271–292.

Kynigakis, I., & Panopoulou, E. (2022). Does model complexity add value to asset allocation? evidence from machine learning forecasting models. *Journal of Applied Econometrics*, *37*(3), 603–639.

Lane, P. R. (2012). The european sovereign debt crisis. *Journal of Economic Perspectives*, *26*(3), 49–68.

LeRoy, S. F. (1989). Efficient capital markets and martingales. *Journal of Economic Literature*, *27*(4), 1583–1621.

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

Lettau, M., & Ludvigson, S. (2001b). Resurrecting the CAPM: A cross-sectional test when risk premia are time-varying. *Journal of Political Economy*, *109*(6), 1238–1287.

Lettau, M., & Ludvigson, S. C. (2010). Measuring and modeling variation in the risk-return trade-off. *Handbook of Financial Econometrics: Tools and Techniques*, 617–690.

Lettau, M., & Pelger, M. (2020). Factors that fit the time series and cross-section of stock returns. *The Review of Financial Studies*, *33*(5), 2274–2325.

Lewellen, J. (2014). The cross section of expected stock returns. *Forthcoming in Critical Finance Review, Tuck School of Business Working Paper*, (2511246).

Light, N., Maslov, D., & Rytchkov, O. (2017). Aggregation of information about the cross section of stock returns: A latent variable approach. *The Review of Financial Studies*, *30*(4), 1339–1381.

Lin, C.-L., Wang, M.-C., & Gau, Y.-F. (2008). Emerging bond market volatility: Dynamic interdependency and volatility transmission between the us and emerging bond markets. *Annual Conference*.

Liu, Z., Ding, Z., Lv, T., Wu, J. S., & Qiang, W. (2019). Financial factors affecting oil price change and oil-stock interactions: A review and future perspectives. *Natural Hazards*, *95*, 207–225.

Longstaff, F. A., Pan, J., Pedersen, L. H., & Singleton, K. J. (2011). How sovereign is sovereign credit risk? *American Economic Journal: Macroeconomics*, *3*(2), 75–103.

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

Matteson, D. S., & Tsay, R. S. (2011). Dynamic orthogonal components for multivariate time series. *Journal of the American Statistical Association*, *106*(496), 1450–1463.

McDonald, J., Michelfelder, R. A., & Theodossiou, P. (2009). Robust regression estimation methods and intercept bias: A capital asset pricing model application. *Multinational Finance Journal*, *13*(3/4), 293–321.

Mincer, J. A., & Zarnowitz, V. (1969). The evaluation of economic forecasts. In *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance* (pp. 3–46). NBER.

Mink, M., & De Haan, J. (2013). Contagion during the greek sovereign debt crisis. *Journal of International Money and Finance*, *34*, 102–113.

Miranda-Agrippino, S., & Rey, H. (2021). *The global financial cycle* (Working Paper No. 29327). National Bureau of Economic Research.

Miranda-Agrippino, S., & Rey, H. (2020). U.S. monetary policy and the global financial cycle. *The Review of Economic Studies*, *87*(6), 2754–2776.

Moritz, B., & Zimmermann, T. (2016). Tree-based conditional portfolio sorts: The relation between past and future stock returns [Available at SSRN 2740751].

Naeem, M. A., Hasan, M., Arif, M., Suleman, M. T., & Kang, S. H. (2022). Oil and gold as a hedge and safe-haven for metals and agricultural commodities with portfolio implications. *Energy Economics*, *105*, 105758.

Odling-Smee, J. (2006). The IMF and Russia in the 1990s. *IMF Staff Papers*, *53*(1), 151–194.

Oyster, M. (1997). Progressive acrophobia and the put/call ratio cure. *Futures: News, Analysis & Strategies for Futures, Options & Derivatives Traders*, *26*(10).

Pan, J., & Poteshman, A. M. (2006). The information in option volume for future stock prices. *The Review of Financial Studies*, *19*(3), 871–908.

Passari, E., & Rey, H. (2015). Financial flows and the international monetary system. *The Economic Journal*, *125*(584), 675–698.

Pinto-Ávalos, F., Bowe, M., & Hyde, S. (2024). Revisiting the pricing impact of commodity market spillovers on equity markets. *Journal of Commodity Markets*, *33*, 100369.

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

Racicot, F.-E., Rentz, W. F., & Théoret, R. (2025). Is illiquidity priced in an international factor pricing model? a dynamic panel data application with robust iv. *International Journal of Finance & Economics*, *30*(1), 282–314.

Radelet, S., Sachs, J. D., Cooper, R. N., & Bosworth, B. P. (1998). The East Asian financial crisis: Diagnosis, remedies, prospects. *Brookings papers on Economic activity*, *1998*(1), 1–90.

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

Ready, R. C. (2018). Oil prices and the stock market. *Review of Finance*, *22*(1), 155–176.

Rey, H. (2013). Dilemma not trilemma: The global financial cycle and monetary policy independence. *Federal Reserve Bank of Kansas City Economic Policy Symposium.*

Rey, H. (2015, May). *Dilemma not trilemma: The global financial cycle and monetary policy independence* (Working Paper No. 21162). National Bureau of Economic Research.

Romo, J. M. (2012). Volatility regimes for the vix index. *Revista de Economía Aplicada*, *20*(59), 111–134.

Rossi, B. (2012). The changing relationship between commodity prices and equity prices in commodity exporting countries. *IMF Economic Review*, *60*(4), 533–569.

Roy, R. P., & Roy, S. S. (2017). Financial contagion and volatility spillover: An exploration into indian commodity derivative market. *Economic Modelling*, *67*, 368–380.

Sachs, J. D., Tornell, A., & Velasco, A. (1996). *Financial crises in emerging markets: The lessons from 1995* (tech. rep.). Brookings Institution. Washington, D.C.

Sadorsky, P. (2014). Modeling volatility and correlations between emerging market stock prices and the prices of copper, oil and wheat. *Energy Economics*, *43*, 72–81.

Samora, R., & McGeever, J. (2021). Iron ore to displace soybeans as brazil's main export earner. *Reuters*.

Samuelson, P. A. (2016). Proof that properly anticipated prices fluctuate randomly. *The World Scientific Handbook of Futures Markets*, 25–38.

Sarno, L., Schneider, P., & Wagner, C. (2016). The economic value of predicting bond risk premia. *Journal of Empirical Finance*, *37*, 247–267.

Schwager, J. D. (2012). *Hedge fund market wizards: How winning traders win.* John Wiley & Sons.

Silvennoinen, A., & Thorp, S. (2013). Financialization, crisis and commodity correlation dynamics. *Journal of International Financial Markets, Institutions and Money*, *24*, 42–65.

Sokhombela, A. L. L. (2024). *Assessing the performance of safe haven assets during major crises*. University of Johannesburg (South Africa).

Sortino, F. A., & Van Der Meer, R. (1991). Downside risk. *Journal of Portfolio Management*, *17*(4), 27–31.

Statman, M., & Fisher, K. L. (2002). Consumer confidence and stock returns. *Santa Clara University Dept. of Finance Working Paper*, (02-02).

Stock, J. H., & Watson, M. W. (2002). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, *97*(460), 1167–1179.

Thornton, D. L., & Valente, G. (2012). Out-of-sample predictions of bond excess returns and forward rates: An asset allocation perspective. *The Review of Financial Studies*, *25*(10), 3141–3168.

Tronzano, M. (2020). Safe-haven assets, financial crises, and macroeconomic variables: Evidence from the last two decades (2000–2018). *Journal of risk and financial management*, *13*(3), 40.

Uribe, M., & Yue, V. Z. (2006). Country spreads and emerging countries: Who drives whom? *Journal of International Economics*, *69*(1), 6–36.

Van Rijckeghem, C., & Weder, B. (2001). Sources of contagion: Is it finance or trade? *Journal of International Economics*, *54*(2), 293–308.

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

Working, H. (1949). The theory of price of storage. *The American Economic Review*, *39*(6), 1254–1262.

Xu, X., & Liu, W.-h. (2024). Forecasting the equity premium: Can machine learning beat the historical average? *Quantitative Finance*, *24*(10), 1445–1461.

Yao, J., Li, Y., & Tan, C. L. (2000). Option price forecasting using neural networks. *Omega*, *28*(4), 455–466.

Young, T. W. (1991). Calmar ratio: A smoother tool. *Futures*, *20*(1), 40–41.

Zarifhonarvar, A. (2023). The capital asset pricing model: A new empirical investigation.

Zhang & Hamori, S. (2021). Crude oil market and stock markets during the covid-19 pandemic: Evidence from the us, japan, and germany. *International Review of Financial Analysis*, *74*, 101702.

Zhang, Wei, Y., Ma, F., & Yi, Y. (2019). Economic constraints and stock return predictability: A new approach. *International Review of Financial Analysis*, *63*, 1–9.

Zhang, Hu, M., & Ji, Q. (2020). Financial markets under the global pandemic of covid-19. *Finance Research Letters*, *36*, 101528.