ESSAYS ON MODELLING AND FORECASTING STOCK MARKETS WITH INVESTOR SENTIMENT

CHANG SUN

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

Department of Economics

Birmingham Business School

College of Social Sciences

University of Birmingham

October 2020

Abstract

This thesis provides an analysis of the predictability of stock returns in ten European countries. First, we discuss the relationship between investor sentiment and stock returns in panel and individual levels in chapter 1. As investor sentiment is an subjective variable and is not easily to be observed directly, we compare all measures of investor sentiment employed in past literature to find the suitable proxy for it. It seems that only the consumer confidence index (CCI) is standardized across all European countries. Also we use macroeconomic factors as control variables since they can improve predictions of returns. We show that investor sentiment can positively affect the stock return in all of these ten European countries. In chapter 2, we use both the univariate models (ARMA, ARMAX) and the multivariate forecasting models (VAR, BAR) to examine the predictability of stock returns. Among these models, the ARMAX performs better in out-of-sample prediction. Although VAR models are usually better forecasters than the ARMA and ARMAX, in this case they disappointed the expectations. Also, the Bayesian VAR improves the standard VAR and performs better in general. In chapter 3, we use three types of model averaging methods (SMA, BMA, AMA) to combine the predictions of individual models into a composite model to improve the predictive performance for stock returns. According to the out-of-sample results for averaged models, we can conclude that the forecasting performance of stock returns has been improved significantly by averaging individual models

with different weights. Among these three types of model averaging methods, the Bayesian Model Averaging performs better than the other two methods in most countries.

Acknowledgements

First and foremost, I would like to express my sincere appreciation to my first supervisor, Dr. Marco Barassi, who keeps guiding me and encouraging my research a lot during the process of my PhD. Throughout these four years, I received constructive and valuable advice of my thesis from each meeting we had. He is patient to me from the very beginning of my PhD and inspire me to think critically. I benefit immensely from the suggestion he gave me both in research and life. I would also thank my second supervisor, Dr. Joanne Ercolani, for her professional comments and effective advice to my research. She provides a tremendous help with my chapters and the presentations I gave.

Secondly I would like to thank both my external and internal examiners, Professor Fabio Spagnolo and Professor David Dickinson. Thank you for your precious suggestions on my thesis during the viva. And thank you Dr. Ioannis Karavias for attending my viva.

I truly want to thank my parents. Without their supports and encouragement, I would not even be a PhD student. It is their selfless love that teaches me always keeping positive attitude to life. They are always the closest family I care about, no matter where I am. I feel really guilty that I cannot accompany with them every day, but I am aware that they support every decision I made and would be very proud of me.

Last, I would like to give a huge thank you to all of my friends and colleagues. Particularly I should say thank you to Dr. Yuqian Zhao, Dr. Weiqing Tang, Dr. Shixuan Wang and Dr. Weifeng Zhou for the help with my work. Also, I owe a thank you to my best friends, Nan Xu, Chen Hou and Yanbo Yang for your patience of my complaints and tolerance of my whines about my research. I really enjoyed the happy time with you. Specially, I thank my partner, Ray. Thank you for all your help with my research and the care of my life.

Table of contents

List of Figures

List of Tables

Chapter 1

Introduction

In the past several decades, a large volume of literature have focused on modelling and predicting stock returns.

Before that, the efficient market hypothesis (EMH) constituted the main theoretical framework within which stock markets were analysed. According to the EMH, agents in the market are fully informed and rational. Empirically, stock prices were pure random walks and as such returns were unpredictable white noise processes. In simpler terms, this risk-adjustment hypothesis assumed that the stock market could always reach to the equilibrium level and market information were fully reflected by the stock price (Malkiel and Fama, 1970; Basu, 1977; Malkiel, 1989). Under the EMH, no matter the rational or irrational investors behaviours should follow the intrinsic market pattern, which leaves no role of investor feelings in stock markets.

However, as a series of financial crises occurred since 1970s, researchers began to realize that the EMH had failed to completely explain the behaviour of stock markets. Investor sentiment is also one of the key factors that had to be considered. De Long et al. (1990) defined the 'noise traders' as the unpredictable part in the stock market. They also pointed out that these noise traders would affect the expected stock returns by acting emotionally. Similarly, Campbell and Kyle (1993) used the noise trader theory to explain the anomalies in the US stock market. Since then, more and more studies began to investigate whether investor sentiment has significant effects on stock returns (Fisher and Statman, 2003; Brown and Cliff, 2004, 2005; Baker and Wurgler, 2007; Schmeling, 2009).

Different from financial factors or macroeconomic factors, investor sentiment is completely subjective and cannot be observed easily. Thus although important, it is not easy to find proper proxies for it. According to previous studies, there are abundant ways to measure investor sentiment which are usually divided into three types, namely direct, indirect and composite. Direct measurements refer to surveys generated by specific institutions by means of questionnaires that can reflect the attitude of investors towards the current and future state of the stock market. There are several popular surveys that has been used in past literature, such as the U.S. Michigan Consumer Sentiment (Lemmon and Portniaguina, 2006), the Investors Intelligence (II) (Fisher and Statman, 2000), the American Association of Individual Investors (AAII) (Fisher and Statman, 2003), the Consumer Confidence Index (CCI) (Schmeling, 2009). Except for the CCI, all of the other surveys are only focused on the US stock market.

A few past papers also adopted the financial indicators as the indirect measurements of stock returns. There are six of them used widely in empirical studies: the closed-end fund discount (CEFD) (Qiu and Welch, 2004), the dividend premium (Fama and French, 2001), the stock trading volume (Jones, 2002), the equity share in new issues (Baker and Stein, 2004) , the initial public offering (IPO) first day returns and its trading volume (Ljungqvist et al., 2006) respectively. Compared with direct measurements, different financial estimators reflect the investor sentiment in different aspects. It does not seem too convincing to use any single financial indicator as the proxy for investor sentiment. In fact, Baker and Wurgler (2006) generated a composite index (BW index) to measure investor sentiment by combining those financial indicators with various proportions. After comparing all different types of measurements, it seems that the direct CCI is the most suitable proxy for investor sentiment in each European stock market for the reason that only the CCI is collected in a consistent way across time and countries.

The thesis is organised as follow: in chapter 2, we investigate whether stock returns are affected by investor sentiment significantly in ten European countries. As mentioned above, we select the CCI as the unique proxy for investor sentiment to estimate and forecast stock returns for ten European countries. Also, we adopt several macroeconomic factors as other control variables. Alongside the consumer confidence index (CCI), we include the CPI inflation rate, a three months treasury interest rate (R) and the industrial production index (CPI) as our control variables (Rapach et al., 2005; Schmeling, 2009). We then merge the data of these ten European countries into a panel and we estimate the model both as a panel and for the individual countries to examine how the effects of investor sentiment on stock returns changes on the specific characteristics of each country and also produce forecasts in the two cases.

In chapter 3, we continue to investigate the predictability of stock returns by univariate and multivariate time series models. Following the steps of past literature (Henry, 2002; Anaghi and Norouzi, 2012), the first model we use to forecast stock returns is the Autoregressive Moving Average (ARMA) model which combines the auto-regression with the moving average process. The ARMA model can prove whether stock returns are self-predictive. However, as mentioned above, it is likely that stock returns are not only affected by their own past, but are also correlated with investor sentiment and macroeconomic factors (Baker and Wurgler, 2000; Rapach et al., 2005). Considering this, we employ the ARMAX model, which can add exogenous variables to the original ARMA so to improve its forecasting performance. As multivariate models, we use the vector-autoregression (VAR) model as it is widely adopted by past literature (Binswanger, 2004; Schmeling, 2009). Different from the ARMA, the VAR model can be considered as a whole system as it combines all of the variables into a matrix. Thus we can forecast all variables together instead of only focusing on the predictable values of stock returns, which can help us to observe the relationship between different variables. we also use a Bayesian VAR (BVAR), which compared with the traditional VAR has the property that all of the parameters can be regarded as random and estimated by the Bayesian methods. The coefficients for the long-lag terms would approach to zero, thereby efficiently avoid any over-fitting issues, typical of VAR models. Selecting proper lag lengths is also an important issue when we forecast stock returns with different classes of models. It can directly determine the quality of forecasting models. Following previous studies, we use several information criteria to select the lag length. In this way, the model with smaller information criteria is selected as optimal. To examine the quality of prediction, we then generate out-of-sample forecasts so that we can calculate the difference between forecasting values and the actual values out-of-sample.

It turns out that, the ARMA and ARMAX models generate better predictions than VAR and BVAR models which are usually though as better performing. This results, shows that it can still be controversial which model has the absolutely superior forecasting performance when comparing with others. An alternative way has been proposed to improve the accuracy of forecasting performance, namely the model averaging method (Wasserman et al., 2000; Burnham and Anderson, 2002; Montgomery and Nyhan, 2010; Symonds and Moussalli, 2011). This method refers to the process of combining the optimal models selected from each class by giving calculated weights to them. The optimal models with fitted lag length are selected by information criteria. In chapter 4, we merge these candidate models selected from each class of models introduced in chapter 3 into an averaged model. We use three types of model averaging methods to forecast stock returns and make comparison between the forecasting performance of individual and average models. We first use simple model averaging (SMA) where each of the candidate models would be weighted equally. The simple average model can be very efficient if all of the candidate models are well-specified. Bayesian model averaging (BMA), however, is more widely adopted in empirical studies (Wasserman et al., 2000; Posada and Buckley, 2004). The weight of BMA is calculated based on the Bayesian rule by estimating the log-likelihood function of the parameters. We also use the Akaike model averaging (AMA) to forecast stock returns. Similar with BMA, the weight of AMA is calculated based on the Akaike information criteria.

The thesis also contributes to the literature on the analysis of modelling and forecasting stock returns in several respects. To be specific, first, we observe the relationship between investor sentiment and stock returns in ten European countries instead of only focusing on a single stock market as in past literature. Also, we estimate predict the stock returns both in panel level and individually, which provides better comparisons of the specific effect of investor sentiment on stock returns among European countries. We also include macroeconomic factors alongside proxies for investor sentiment and organize them into the estimation model as control variables to avoid misspecification of the model. Furthermore, we contribute to the literature by comparing four different forecasting models (ARMA, ARMAX, VAR and BVAR) showing that stock returns are significantly correlated with not only their own past, but also with investor sentiment as well as the macroeconomic factors thereby adding evidence that the EMH is insufficient to explain the functioning of the stock market. We also compare the out-of-sample forecasting results for these models. Finally, we use model averaging methods to improve the accuracy of the predictive performance of stock returns by distributing different weights to each of the optimal models in each class of models and generate improved forecast.

Chapter 2

Investor Sentiment and Stock Market Returns:

Some Evidence for European Countries

2.1 Introduction

The efficient market hypothesis (EMH, hereafter), as the core content of classical financial theories, left no role for investor sentiment for several decades. This hypothesis is based on two assumptions. One is that asset prices can fully incorporate and reflect market information. The other is that the stock market can always reach an equilibrium level where the stock prices are equal to the rationally discounted value (Malkiel and Fama, 1970). Even though few investors are irrational, the market internal pattern can control the price fluctuation so that their irrational behaviour cannot affect market price. However, as several financial crashes have occurred in the past, a number of researchers have argued about the validity of the EMH and began to focus on the expectation of investors.

There were various definitions for investor sentiment. Zweig (1973) pointed out that investors usually tend to overestimate or underestimate the market situation when they make decisions, and the main cause for such bias was the existence of sentiment even in the most developed markets. Later in 1990s, De Long et al. (1990) presented the 'noise trader' theory, which defined the behaviours of irrational investors as 'the unpredictable part of unsophisticated opinion of investors'. Furthermore, Baker and Wurgler (2007) documented that investor sentiment was the expectation of investors about future stock market conditions and investment risks that were unable to be observed or estimated directly. Despite differences in word, all definitions involves the expectations of investors. With respect to our research, we define investor sentiment as the variable that represents the expectations of investors about the future of stock market. Specifically, a bullish investor expects higher future returns compared with the current returns while a bearish expects lower future returns relatively to current returns.

From the generalized definition, it is clear that investor sentiment is substantially unobserved and subjective. According to this characteristic of it, researchers attempted to find suitable proxies for investor sentiment. Unfortunately, they have not found a widely accepted measurement in recent studies. Generally speaking, all of the proxies that have been employed in past papers can be divided into three types: direct measures, indirect measures and composite indexes respectively. Prior works used surveys composed by different institutions as the direct measures. Compared with it, the indirect measures referred to financial indicators. For this study, we list measurements for sentiment and compare these measurements in some detail. Finally, considering the data periods and the standardized issue across countries, we decide to use consumer confidence index (CCI, hereafter) as our proxy for investor sentiment in modelling and forecasting stock returns.

Normally, investor sentiment can affect the stock returns in two possible ways, which are time-series effect and cross-sectional effect. The time-series effect can better reflect the fluctuation of variables based on time changing, while under the cross-sectional effect, the asset price and arbitrage constraints could be vary across different markets. Namely, it is easier to compare characteristics among different countries at the same time. Regarding these two effects, we first organize the data into a panel as a system and test the forecasting model jointly in order to figure out the relationship between investor sentiment and aggregate stock market returns on average across countries. Then we examine the forecasting model separately so that we can find the sentiment-return relation in any individual country. We can compare time-series effect and the cross-sectional effect in this way.

In recent years, more and more studies have began to focus on the the impact of investor sentiment on stock market. However, most of these works were concentrating on the American stock market(De Long et al., 1990; Brown and Cliff, 2004; Baker and Wurgler, 2006). These researches employed both direct and indirect proxies for sentiment to interpret one common conclusion: investor sentiment is significantly related to the American stock market returns. Our research contributes to the existing literature on spreading the research target to ten European countries, which are Austria (OE), Belgium (BG), Finland (FN), Germany (BD), Greece (GR), Italy (IT), Netherlands (NL), Portugal (PT), Spain (ES) and the United Kingdom (UK) respectively. We find that the effect of investor sentiment on stock market is still significant both jointly and individually. However, the correlation between sentiment and returns is positive which is contradictive with most of the conclusions in researches based on the US stock market. This result proves that it is not possible to simply transfer evidence obtained from the US stock market to other stock markets.

In addition, using CCI as the proxy for investor sentiment, this study constructed macroeconomic factors as control variables to avoid perfect collinearity. To be specific, we use the monthly change consumer price index (CPI), monthly treasury interest rate (IR) and monthly percentage change of industrial production index (IP) in all ten European countries since these macroeconomic factors are definitely not correlated with each other and can stay fixed in the regression model.

The rest of the chapter is organized as follows. Section 2 is an overview of existing literature and in the third section we propose the testable hypothesis. Section 4 introduces the measurement for investor sentiment and compare them in details. Section 5 provides a general data description and section 6 describes the model for forecasting regressions. Section 7 contains the empirical result based on the model in section 6 and section 8 is the conclusion part.

2.2 Theoretical Background

For many years, the efficient market hypothesis (EMH) has taken the leading role in classic financial theories which states that stock prices can be fully predictable via changes of discount rate and stock market information. Although EMH allowed investors to be irrational and behave accordingly, it strongly required the market to be orderly and well-organized. In other words, stocks can always trade on its fair value, which gives no chance to arbitrage in an efficient market. Apparently, the sentiment-return relation has no role in this framework based on the assumption that a stock market is always self-regulated. This theory seemed quite convincing until a series of events happened in the history of several stock markets, such as the Kennedy Slide of 1962 (also know as 'the Flashing Crash of 1962'), Brazillian Markets Crash of 1971, the dramatic rise in oil price in UK of 1973 and the Black Monday of 1987, showing that it was possible to 'break the market'. These crashes proved that EMH failed to explain abnormal stock prices. As a consequence, researchers began to focus on irrational behaviour of investors, namely what we can call now investor sentiment.

Zweig (1973) showed that stock prices move as a random walk with closed-end fund premiums which is the measurement of investor expectations even in the most popular stock market, such as the US one. Specifically, they documented that the price changing is caused by arbitrageurs instead of the rational investors and explained that biased expectation would affect stock value and stock returns systematically. This theory was at odds with the standard EMH by considering the irrational behaviour of investors. However, since their theory was only based on the price changing caused by arbitrageurs, the conclusion was not a definite one. Since then, a number of attempts have been made to explore the effect of individual investor sentiment on stock returns, especially on the most advanced stock market in the world: the US stock market. Typically, De Long et al. (1990) shed light on the theory of noise traders with an overlapping generation model, which suggested that some irrational investors trade on a 'noisy signal', adding a risk to diverge stock price from their intrinsic value. Lee et al. (1991) and Campbell and Kyle (1993) also explained the stock market anomalies using the theory of noise traders. They documented that closed-end funds, as the proxy for investor sentiment, had a significant relationship with stock returns in an imperfect market. After that, a number of researchers attempted to measure the effect of investor sentiment on stock returns quantitatively.

2.2.1 Definition of Investor Sentiment

Investor sentiment, as one of the key contents of behavioural finance theories nowadays, was first suggested by Keynes (1936). He stated that investors can both overreact and underreact to the stock market when they receive the same market signals. In other words, it is the behaviour of irrational investors to produce the likelihood that a stock price varies from its fair value. Subsequently, Zweig (1973) also documented that impulsive investors could still confound the stock market even based on the perfect sharing stock market information assumption. Later in 1990s, several stock market crashes motivated researchers to shift their attention to investor sentiment rather than the EMH. At that time, the definition of investor sentiment was closely connected with the 'noise trader' theory (De Long et al., 1990; Elton et al., 1998; Neal and Wheatley, 1998). However, the noise trader theory has been proved to be controversial in further researches. More recently, the concept of investor sentiment has been further expanded and is now more accurate (Fisher and Statman, 2000; Doukas and Milonas, 2004; Baker and Wurgler, 2006; Huang et al., 2015). Normally in behavioural finance, the definition of investor sentiment was the emotional opinion about both of a current and future stock market. To be specific, a bullish trader expects stock returns to be above the average level while a bearish trader expects lower returns instead. Based on the discussion about the definition of investor sentiment, it is clear that sentiment directly leads to market imperfection that are at odds with the EMH. In the meanwhile, it is not hard to see that investor sentiment is a relatively unobserved and subjective variable compared to stock market and economic fundamentals.

2.2.2 Different effect of sentiment

In recent years, a substantial number of studies have tried to understand the effect and the predictive ability of investor sentiment on stock market, both for any individual country and the global stock markets as a whole. Normally, we can classify the approach in two possible ways, namely a time series effect and a cross section effect. The time series effect focuses on a certain individual stock market in different time periods. The cross section effect concentrates more on the part of returns that can be affected by the cross section symmetric risks. Additionally, the cross section effect allows the asset price and arbitrage constraint to vary across different stock markets.

The research on the time series effect of sentiment can be dated back to 1970s. Kahneman and Tversky (1979) presented the prospect theory with descriptive model of asset risk. They also state that certainty contributes to avoid risky assets whereas the isolation effect results in the failure of taking dividends components into consideration. Hence, firms' performances will be inconsistent even facing the same portfolio choices. Moreover, firms may use decision weights instead of probabilities to describe the main reason of mispricing, that is the unpredictable risk created by irrational investors. Although they make brilliant contributions to explaining the prospect theory with risk conditions, their findings are mostly qualitative. In fact, they fail to evaluate the specific effect of irrational traders on stock assets and stock system risk in a quantitative way. Conversely, De Long et al. (1990) filled this gap with a new definition, called the noise trader. They generated a overlapping generation model to illustrate the noise trader theory which proves that irrational behaviour of noise traders is also the main source of system risk in stock markets. Their research is based on two assumptions: one is that the behaviour of irrational investors is completely unpredictable and the other is that arbitrageurs undertake the risk of mispricing caused by noise traders. Except for confuting the efficient market hypothesis, they also explain some abnormal situations in the stock markets, such as the equity premium puzzle, the closed-end fund discount puzzle, by focusing on the irrational investors models. Unfortunately, their research proved to be unconvincing in later papers in this field for endogeneity issues. Later, Shefrin and Statman (1994) created a Behavioural Capital Asset Model (BAPM) by adding investor sentiment into the traditional Capital Asset Pricing Model (CAPM). They stated that decisions of rational arbitragers and noise traders co-determined stock prices under disposition effect, which describes that investors always sell stock shares with stock increasing prices and buy stock shares with decreasing prices.

From the beginning of 2000s, researchers were not satisfied with simply discussing the sentiment-return relation, but began being concerned about how to measure the effect of investor sentiment in practical applications. The US stock market, as one of the most developed and advanced stock market in the world, was always the research subject over series of researches in this field. Fisher and Statman (2000) investigate the effects of sentiment on the US stock market. Specifically, they divided the investor sentiment into three groups: named individual, newsletter writers and Wall Street strategists respectively. They pointed out that sentiment was negatively correlated with stock returns and the relationship was weak only for newsletter writers but statistically significant for the other two types of investors. While Brown and Cliff (2004) used VAR models to combine all of estimators as a whole system and held a slightly different view: although stock market returns were strongly correlated with investor sentiment, sentiment played no role in predicting subsequent near-term stock market fluctuation. Baker and Wurgler (2007) constructed a component sentiment index (BW index, hereafter), which combined six indirect measures as proxies for sentiment via the first principal component methodology. They documented that future stock returns were conditional on the component sentiment index. Specifically, when sentiment was high, stocks earned relatively low subsequent returns under the cross-section effect. Recently, Huang et al. (2015) developed a new aligned investor sentiment index using the same indirect measures as proxies of BW index. Compared to the BW index, they eliminated the common noise component via the partial least-squares (PLS) method and stated that the modified sentiment index was more powerful in predicting the stock market as compared with the BW index.

Meanwhile, a number of pioneering studies focused on other developed stock market to study the common pattern of the sentiment-return relation. Doukas and Milonas (2004) examined whether the effect of investor sentiment, measured by closed-ends funds, was significant for the Greek stock market. The result showed that the closed-end funds had less effect on Greek stock returns compared to the US stock market. The convincing explanation for this was that closed-end funds could not fully symbolize sentiment and the Greek stock market was still an emerging market compared to other well-developed ones. Afterwards, Schmeling (2009) utilized the consumer confidence index as the proxy for investor sentiment to examine whether the future stock returns could be affected by sentiment in eighteen industrialized countries. To be specific, they divided the stock market into three groups, namely aggregate stock market, value stock market and growth stock market respectively. Their results were consistent with earlier findings for the US stock market, namely that more positive sentiment would lead to a lower future stock returns. Moreover, they used cross-sectional analysis to show that the sentiment effect would be stronger in markets lacking of integration and collectivism. Chang et al. (2012) investigated the effect of sentiment on twenty three different countries equity markets, covering both the developed and developing market spectrum. They found that there is a significantly negative correlation between the local sentiment and the future stock returns across stock markets in different countries. Also, it was notable that a stock market with higher information symmetry had stronger sentiment-return relation. Baker et al. (2012) constructed sentiment indices for six major stock markets based on the first principle component method. Furthermore, they found that global investor sentiment are negatively correlated with time-series of cross-sectional returns across different stock markets.

2.3 Our Hypothesis

Based on the preceding theoretical considerations, we propose the following hypothesis:

H1. Current investor sentiment can significantly affect the future stock returns in different European countries.

There is abundant evidence showing that the effect of investor sentiment on stock returns is significant. De Long et al. (1990) used the overlapping model to explain that irrational traders, also called the noise traders, were the main causes of mispricing. Since then, Brown and Cliff (2004) employed the VAR model to interpret the significant impact of sentiment on stock return while Baker and Wurgler (2007) constructed a composite index combining six proxies for sentiment, providing further evidence of the forecasting power of investor sentiment.

H2. The sentiment-return relation is negative.

Theoretically, regarding the sentiment-return relation, it is still controversial in the direction of it. Brown and Cliff (2004) pointed out that although sentiment had been proved to be correlated with stock returns, the predictive performance is weaker in near-term than in long run periods both in individual and multiple stock markets. Instead, Baker and Wurgler (2007), Schmeling (2009) and Huang et al. (2015) argued that sentiment was significantly negative correlated with stock returns both for individual country stock market and for international ones. Following (Schmeling, 2009), we will use data from 10 European countries and employ consumer confidence index as the standardized proxy for sentiment.

H3. As forecast horizon becomes longer, the effect of sentiment on average stock returns decrease.

Supposing there are only two types of investors in the market, which are irrational investors and arbitrageurs. As the market resources are limited, the expectation is that irrational investor will misprice the current price and the near-term stock price as well, but not prices in the long term (De Long et al., 1990). For this study, we will test the power of sentiment at different forecasting horizon jointly across different countries. The null hypothesis for the test is that the coefficient value of all proxies for sentiment equals to zero. Such that sentiment has known predictive power for all current and future returns.

2.4 Measurements of investor sentiment

As investor sentiment is relatively subjective and is hard to observe directly, it is essential to measure it properly. Prior work has suggested a variety of measurements for investor sentiment which can be roughly divided into three types: direct, indirect and composite measurement respectively.

2.4.1 Direct Measures

As previously noted, there are numbers of surveys with different questionnaires that can directly reflect investor sentiment. The most representative survey is published monthly by University of Michigan since 1952. It contains fifty questions covering the continental United States over five hundred telephone interviews per month used by Charoenrook (2005) and Lemmon and Portniaguina (2006) to examine the forecasting power of consumer confidence on economic fundamentals. Another survey is the Investors Intelligence (II) sentiment index conducted by Chartcraft in 1947. It provides a perspective on the bull-bear stock market via categorizing approximately 150 newsletters since 1960s. Fisher and Statman (2000) used II data as the proxy for investor sentiment and found the negative but weak correlation between sentiment and US stock returns. Later, Brown and Cliff (2005) documented that the II sentiment index, as the direct survey measure of sentiment, was one of the main causes of mispricing and was powerful to predict future stock returns. Similarly, there are also other surveys, such as the American Association of Individual Investors (AAII) sentiment surveys used by Fisher and Statman (2003) and the Gallup

Index of Investor Optimism used by Qiu and Welch (2004), which are widely considered as the direct measures of sentiment in many researches (Fisher and Statman, 2003; Qiu and Welch, 2004).

As introduced in former paragraph, these direct measurements are only suitable for the US stock market for the reason that questions in the surveys are based on the economic factors and condition of the US, which is difficult to be standardized and widely applied to other countries. When attempting to spread the research target to different countries, it is necessary to find a consistent way to measure investor sentiment across multiple stock markets. According to recent detailed analysis of investor sentiment, it is reasonable to use the consumer confidence index as direct measure of sentiment across countries for several reasons (Charoenrook, 2005; Lemmon and Portniaguina, 2006; Schmeling, 2009). Firstly, the consumer confidence index (CCI) is available for several individual countries compared to other direct measures. It is an indicator conducted to reflect the expectation of investors both in current and future stock market. To collect this index, this direct survey contains a standardized question pooled across different countries. Specifically, the content of questionnaires includes the expectation of future financial conditions, the expectation of future general economic state, the employment expectation and the expectation on savings for a future year. Additionally, a number of questions related to the situations for the past year are added into the survey aiming at locating the current financial situation of household. Secondly, this index is correlated with other sentiment proxies and hardly affected by the contents of news given its monthly frequency. As above, the CCI is the only suitable direct measure of sentiment in different countries. Since our research targets are European countries instead of single stock market, we also use the CCI as a proxy for investor sentiment to investigate the effect of it on European stock market returns.

2.4.2 Indirect Measurements

Compared with direct surveys on expectation of investors, indirect measures are financial indicators which are always taken as market weather vanes by stock market commentators. Considering past papers, there are numbers of indicators that have been used as proxies for investor sentiment. We will introduce six widely used indirect measures, which are the closed-end fund discount, stock trading volume, the initial public offering (IPO, hereafter) first day returns, the IPO volume, the equity share in new issues and the dividend premium respectively.

The closed-end funds discount (CEFD) refers to the average level of difference between the net asset value (NAV) of closed-end fund shares and their market price. It is marked as 'discount' if NAV is higher than market price and defined as 'premium' otherwise. The CEFD puzzle is a famous empirical finding that closed-end funds shares are normally sold at a discount compared to its NAV. To solve this puzzle, Zweig (1973) clarified that CEFD reflects the expectation of individual investors and De Long et al. (1990) employed a model to prove that the existence of noise traders may cause the unpredictable price changing which causes fluctuations on the CEFD as well. Meanwhile, Lee et al. (1991) and Chen et al. (1993) provided empirical evidence that discounts are caused by expectation of irrational investors in stock markets. Later, Neal and Wheatley (1998), Qiu and Welch (2004) and Ben-Rephael et al. (2012) discussed whether CEFD can be considered as an indirect measure of investor sentiment.

The stock trading volume refers to the number of stock shares traded in a certain market during a given period of time. Practically speaking, researchers always take the share turnover (TURN), which is the ratio of volume to average market shares, as an indirect measure of sentiment for the reason that investors prefer to trade more frequently in a high expectation of market. To be specific, Jones (2002) stated that bid-ask spreads are cyclical under high liquidity and turnover is negatively correlated with stock returns. While Baker and Stein (2004) employed market liquidity as a sentiment indicator and found that noise traders always overestimate stock returns under high liquidity in the short-term stock market rather than long-term ones.

IPO indicates the shares of a private company is offered to the public for the first time. According to past literature, the number of IPOs (NIPO) is cited as an indirect measure of sentiment. To be specific, higher NIPO reflects optimistic expectation about the stock market (Ritter, 1984). Furthermore, the IPO market is considered to be sensitive to sentiment mainly for the reason that first-day IPOs are always under-priced based on irrational investor sentiment consideration (Ljungqvist et al., 2006). Baker and Wurgler (2007) subsequently showed the significant impact of first-day IPO returns (RIPO) on investor sentiment.

The equity share in new issues is another common financial indicator that can be regarded as an indirect measure of sentiment. The definition of it is the gross equity issuance divided by the total of gross equity and debt issuance. Baker and Wurgler (2000) used such indicator to measure sentiment, and document that the effect of values of equity share on stock market returns is negative.

The dividend premium (PD) is the log difference between the average book-to-market ratios for payers and non-payers. Fama and French (2001) explained the declining incidence of dividends and found that enterprises prefer to pay dividends under a premium situation rather than a discount one. This phenomenon reflected that dividend premium measures sentiment well especially when enterprises make dividends decisions. Additionally, Baker and Wurgler (2004) developed the catering theory to show that managers were willing to pay dividends when investors put premium asset price on payers and not to pay when they prefered non-payers, which also proved that PD is useful to measure sentiment.

2.4.3 Composite Measurement

Although prior work found substantial proxies for sentiment, unfortunately, it was still controversial whether a specific direct survey or indirect indicator can measure sentiment properly. Therefore Baker and Wurgler (2006) constructed a new composite sentiment index (BW index, hereafter) that combined six indirect measures for sentiment which were shown above. They used the first principle component to obtain a new index and pointed out that there was a strongly negative effect of investor sentiment on returns in the cross-sectional American stock market. Huang et al. (2015) employed a new sentiment index (Aligned index, hereafter) using the same six proxies as Baker and Wurgler (2006) but in a more efficient way. Econometrically, they used the partial least squares (PLS) method to extract the most relevant common component from the proxies and split out noise or error information. This new index provides more evidence of the strong predictive power of investor sentiment for stock market and improves the widely accepted BW index as well.

2.5 Data

2.5.1 Descriptive statistics

In implementing our analysis, we use not seasonally adjusted, monthly stock return indexes as our dependent variable. The stock returns are from the electronic version of Morgan Stanley's Capital International Perspectives (MSCI). The most outstanding advantage of the MSCI data compared with other international databases is that they can eliminate the so called survivor bias comparatively (Fama and French, 1998). To be specific, the MSCI database includes not only the currently traded firms, but the historical data for disappearing firms as well. Our monthly returns, which can be collected from Datastream in a consistent manner for different countries, ranges from May 1985 through December

2015 wherever possible.

As previously noted, there are numbers of proxies for investor sentiment, both direct and indirect. However, most of these measurements are only available for the US stock market. For this study, we require to measure investor sentiment for different European countries, which are Austria (OE), Belgium (BG), Finland (FN), Germany (BD), Greece (GR), Italy (IT), Netherlands (NL), Portugal (PT), Spain (ES) and the United Kingdom (UK) respectively. Doing this in a consistent way calls for the measurement to be comparable across countries and acquirable for specified time duration. Following pioneering studies that also focused on multiple stock markets (Lemmon and Portniaguina, 2006; Schmeling, 2009; Chang et al., 2012), we also use the consumer confidence index as the proxy for investor sentiment for the reason that it seems to be the only suitable direct measurement of sentiment available for all these European countries. The data on consumer confidence index come from the European Commission website, which professionally conducts surveys and provides financial indicators for European countries. Moreover, the content of the questionnaires for consumer confidence index can be acquired as well. The data range of consumer confidence index is consistent with the dependent variable, the stock returns. However, the data limitation issue enforce some-what shorter period for some specific countries as can be seen in Table 2.1.

Additionally, to obtain a better specified model, we also employ monthly macroeconomic factors which are consumer price index (CPI), three months treasury bill interest rate (IR), industrial production index (IP) respectively. As these factors may contribute to explain and forecast stock returns. Also, they can be correlated with investor sentiment although not perfectly.

Table 2.1 shows the summary statistics for all of the variable in the dataset for the ten European countries and essential features are displayed in the summary data below. We also summarized the statistics for all countries in differences in the following table 2.2. From the table, we present important features of return based on stock prices in table 2.1 and the differentiated macroeconomics factors. The time period goes from 1985 May through December 2015 at monthly frequency, giving a total of 368 observations for each time-series. However, there are some variables that haven't been collected within the full sample. In this situation, the analysis of these countries is conducted slightly for shorter time periods. The available time period for each country is also provided on the table as an independent column.

Table 2.1: Summary statistics for all countries in levels

Country	Label	Start	Price mean	Price Std	CCI mean	CCI Std	CPI mean	CPI Std	IR mean	IR Std	IP mean	IP Std
Austria	OE	1995M10	572.3066	273.4436	-2.1173	8.1489	92.8844	10.5192	3.9457	1.4520	89.8037	16.2232
Belgium	BG	1985M05	714.3055	318.6210	-7.3866	8.6671	83.3935	15.1176	5.6141	2.4647	77.0425	17.4289
Finland	FN	1995M11	621.5077	337.5869	12.8847	5.7939	94.1392	9.2747	3.8857	1.5940	94.6776	12.0163
Germany	BD	1985M05	512.5415	238.0039	-7.2943	9.1908	85.6557	13.8313	4.8345	2.1387	89.0997	13.4326
Greece	GR	1997M06	906.1619	574.2423	-42.3318	18.3608	87.8418	12.8862	7.6197	4.8486	109.7693	13.5515
Italy	IT	1991M03	863.5654	344.5075	-16.4252	9.1395	86.7450	14.4772	6.1747	3.2728	106.6753	9.3766
Netherlands	NL	1985M05	763.9952	375.5253	0.5364	11.4571	83.6066	11.2678	4.9832	2.0860	85.6289	11.2678
Portugal	PT	1993M07	146.6662	50.8170	-25.0641	12.9149	88.7302	13.8250	5.9842	2.7055	110.5620	12.2642
Spain	ES	1986M06	649.3495	350.6960	-13.5285	10.6200	78.0971	20.3862	7.0170	3.7570	102.3986	12.5280
UK	UK	1985M05	1309.875	500.0483	-8.7902	8.7799	81.8147	17.5049	6.1734	2.8516	101.9236	6.6103

Table 2.2: Summary statistics for all countries in differences

Country	Label	Start	Return mean	Return Std	CCI mean	CCI Std	CPI mean	CPI Std	IR mean	IR Std	IP mean	IP Std
Austria	OE	1995M10	0.0034	0.0686	0.0002	0.0284	0.0015	0.0032	-0.0040	0.1168	0.0028	0.0152
Belgium	BG	1985M05	0.0074	0.0599	0.0013	0.0348	0.0016	0.0027	-0.0047	0.0752	0.0020	0.0225
Finland	FN	1995M11	0.0107	0.0985	-0.0001	0.0203	0.0013	0.0030	-0.0042	0.1253	0.0028	0.0152
Germany	BD	1985M05	0.0072	0.0661	0.0004	0.0269	0.0014	0.0032	0.0011	0.2024	0.0014	0.0015
Greece	GR	1997M06	-0.0052	0.1095	0.0019	0.0955	0.0020	0.0116	0.0030	0.0769	-0.0001	0.0291
Italy	IT	1991M03	0.0045	0.0687	0.0010	0.0362	0.0020	0.0020	-0.0059	0.0524	-0.0002	0.0130
Netherlands	NL	1985M05	0.0065	0.0571	0.0008	0.3686	0.0015	0.0044	-0.0033	0.0910	0.0013	0.0266
Portugal	PT	1993M07	0.0024	0.0632	0.0022	0.0404	0.0020	0.0045	-0.0037	0.0587	0.0002	0.0257
Spain	ES	1986M06	0.0072	0.0694	0.0013	0.0390	0.0027	0.0048	-0.0040	0.0521	0.0005	0.0171
UK	UK	1985M05	0.0055	0.0495	0.0012	0.3384	0.0022	0.0042	-0.0040	0.0451	0.0003	0.0010

2.5.2 Result for unit root tests

Since our variables in levels in table 2.1 are likely to be non-stationary, it is necessary to test whether the variables in first differences are stationary or not. According to past papers, we use Levin, Lin & Chu test (Levin et al., 2002) as a test for a panel unit root test with the null hypothesis being there is a common unit root in the panel processes. We also employ the Im, Pesaran and Shin (IPS) test, ADF-Fisher test and the PP-Fisher unit root tests (Choi, 2001; Schmeling, 2009) to examine the individual unit root process. The lag length selection is based on the modified Akaike information criterion (AIC) and the test assumes individual intercepts. The results in table 2.3 show that the variables we use are stationary in first differences.

2.6 Methodology

According to a number of pioneering studies in this field (Baker and Wurgler, 2006; Schmeling, 2009), we can start with a simple predictive regression which describes the relationship between the future stock returns and the current investor sentiment as follows:

$$r_{t+1} = \alpha + \beta \times sent_t + \eta_t \tag{2.1}$$

where r_{t+1} is the future stock market return at time $t+1$ and $sentiment_t$ is a proxy for current investor sentiment. Based on this model, a widely accepted finding for the US stock market is that investor sentiment is negatively correlated with future stock returns.

Table 2.3: Unit root tests result

Return	test statistic	p-value	Obs
Levin, Lin & Chu	-5.9753	***(0.00)	3497
Im, Pesaran and Shin	-16.4890	***(0.00)	3497
ADF-Fisher	313.554	***(0.00)	3497
PP-Fisher	313.554	***(0.00)	3596
Sent	**test statistic**	**p-value**	**Obs**
Levin, Lin & Chu	-26.4441	***(0.00)	3327
Im, Pesaran and Shin	-22.9106	***(0.00)	3327
ADF-Fisher	499.280	***(0.00)	3327
PP-Fisher	1442.27	***(0.00)	3383
CPI	**test statistic**	**p-value**	**Obs**
Levin, Lin & Chu	-46.3591	***(0.00)	3627
Im, Pesaran and Shin	-32.0199	***(0.00)	3627
ADF-Fisher	763.93	***(0.00)	3627
PP-Fisher	1589.17	***(0.00)	3660
IR	**test statistic**	**p-value**	**Obs**
Levin, Lin & Chu	-17.8026	***(0.00)	3219
Im, Pesaran and Shin	-20.3648	***(0.00)	3219
ADF-Fisher	426.812	***(0.00)	3219
PP-Fisher	1211.66	***(0.00)	3259
IPI	**test statistic**	**p-value**	**Obs**
Levin, Lin & Chu	-80.2006	***(0.00)	3621
Im, Pesaran and Shin	-50.8973	***(0.00)	3621
ADF-Fisher	729.811	***(0.00)	3621
PP-Fisher	1474.53	***(0.00)	3660

Asterisks refer to the level of significance where: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

In other words, in near-term aggregate stock market, the higher the expectation of investors is, the lower the returns will be in the future (Brown and Cliff, 2004).

However, this model is only suitable for the short-term stock market with one period ahead. Another issue with this model failed to consider the fact that except for investor index, many other factors could also affect stock returns, which would cause the endogeneity issues. To improve this model, we estimate a long-horizon model for stock returns and convert the short-term sentiment-return model into the following form[1]:

$$\frac{1}{k}\sum_{j=1}^{k} r_{t+j}^{i} = \delta_0^{i.(k)} + \delta_1^{i.(k)} sent_t^i + \Psi^{i.(k)}\gamma^{i.(k)} + \xi_{t+j}^{i.(k)} \tag{2.2}$$

where the left hand side is the moving average for k periods of stock returns for each country as dependent variable and the right hand side is the investor sentiment as the independent variable with other control variables and error terms. For stock returns, k is the periods chosen for moving average and i is the symbol for different countries. Similarly to eq(1), $sent_t^i$ is the consumer confidence index as proxy for sentiment in individual countries. Furthermore, we define a matrix, $\Psi^{i.(k)}$, which is composed by macroeconomic factors as explanatory variables in the forecasting model. The reason for adding this new term is to control for the potential effect of other variables beside that of the sentiment index. Thus, the condition for this new term, either as control variable or explanatory variable, is that it should be uncorrelated with fundamental risk factors. In line with previous

[1]The model (2.2) is referred to the forecasting model generated by Schmeling (2009)

literature (Baker and Wurgler, 2006; Lemmon and Portniaguina, 2006; Schmeling, 2009; Huang et al., 2015), we use the monthly percentage change of industrial production index (IPI), the monthly percentage of consumer price index (CPI) and the monthly percentage change in treasury bill interest rate (IR) for these ten European countries into the matrix, Ψ, to net out the effect of commonly used covariates with investor sentiment.

In light of Eq.(2.2), we estimate a panel regression with fixed-effects. Doing this in a consistent way provides a chance to test the effect of sentiment on stock returns across different European countries jointly. Actually, we are not the first one trying to employ this in forecasting regression models. Ang and Bekaert (2006) investigated the predictive performance of dividend yield to stock returns, cash flows and the interest rate. Since they focused on four stock markets in four different countries instead of the individual one, they employ panel techniques for both short term and long term forecasting. Following their paper, we consider 10 European countries as a whole system and jointly estimate and forecast the Eq.(2.2) for different forecast horizons, which are 1, 3, 6 and 12 months respectively. Namely, we test whether there is a jointly significant effect of sentiment on the stock returns at entire European level, and use this results for forecasting purposes.

To check the predictability of stock returns, we use both in-sample and out-of-sample forecasts. Based on the model mentioned above, we first combine all these ten European countries into a panel and discuss the findings for in-sample panel regression so that we can figure out the effect of investor sentiment on stock returns. Then, we use out-of-sample prediction and access the model's forecast performance.

2.7 Empirical Results

In this section, we interpret the results at panel level, as well as the results for individual countries to compare the common points and differences among ten European countries.

2.7.1 Results for panel regressions

2.7.1.1 Results for in-sample and out-of-sample panel regressions

Normally, to test the performance of model's predictability, we split the whole data set into two sub-samples, which are in-sample and out-of-sample portions respectively. There are numbers of researches using this method to access the power of model forecasting. Rapach and Wohar (2006) used both in-sample and out-of-sample analysis to predict stock returns. Similarly, Molodtsova and Papell (2009) used real-time data to find the performance of out-of-sample exchange rate forecasting ability considering the Taylor rule [2] fundamentals. Following the pace of prior works, we separate our panel into two parts. Our whole data set includes 3650 with a time span from May 1985 to December 2015. Due to the data limitation and model restrictions, we take the observations starting from May 1985 until December 2010 as our in-sample portion and observations ranging from January 2011 to

[2] The Taylor rule is an econometric model that describes the relationship between Federal Reserve operating targets and the rates of inflation and gross domestic product growth. The original formula of Taylor rule is: $\mathrm{i}_t = \pi_t + r_t^* + a_\pi(\pi_t - \pi_t^*) + a_y(y_t - \bar{y}_t)$

December 2015 as the out-of-sample portion.

Panel A in table 2.4 presents the results for in-sample panel fixed effects regressions with the aggregate stock market returns as dependent variables. To be specific, in panel A, we provide the estimated coefficients of consumer confidence index as the proxy for investor sentiment and other macroeconomic factors at 1, 3, 6 and 12 months forecast horizons respectively. Also, the table shows the corresponding robust standard errors in parentheses and significance at usual levels. By examining forecasting regressions, when we net out macroeconomic factor from sentiment, it is clear that the CCI as the proxy for investor sentiment, has significantly effects on aggregate stock market returns in panel. This finding proves our hypothesis 1.

To be specific, a rise in sentiment proxy increases the stock returns by 22.15% with one month ahead as a start, while the positive effect of sentiment on returns shrink as the forecast horizon becomes longer and finally decreases to approximately 3.6% with 12 months ahead. However, the correlation between sentiment and stock returns is positive at all forecast horizons which is contradictive of prior works (Brown and Cliff, 2005; Baker and Wurgler, 2006; Huang et al., 2015). The possible reasons for this result are as follows. Firstly, most of the past papers concentrate on the US stock market while our research is based on the European stock markets instead of any individual country. The difference in research target leads to the diversity of our data, especially for countries not performing well. Also, as previously noted, some variables are available for shorter periods for the data limitation issue. As we regress the forecasting model as a panel, the missing data may also affect the coefficient value of independent variables. Secondly, compared to prior works that use annual or quarterly data for a short time period in individual countries, we focus on the monthly data for over 30 years period which is relatively long term for the stock market. The higher frequency data determines that observations fluctuate more. Another possible reason for it is about the macroeconomic factors. In this chapter, we employed a matrix composed by macro factors which are marginally significant to stock returns and reinforced but not perfectly correlated with each others, where some earlier researchers failed to consider the effect of macroeconomic factors on stock returns (Zweig, 1973; De Long et al., 1990).

In addition, it is interesting to find that as forecast horizon becomes longer, the effect of sentiment on average stock returns decreases, which also proves our hypothesis. We present the reason for this finding in economic implications. Overall, the declining marginal effect of sentiment on stock returns provides evidence that the expectation of noise traders can run out with longer periods of time. In other words, both the effect of noise traders and arbitrageurs are limited by the restricted market resources to a certain time period. As a result, in longer forecast horizon, the limits to arbitrage will become weaker and the temporary shock will be absorbed and the stock market would return to its fundamental price.

Panel B in table 2.4 shows the result for out-of-sample forecasting Root Mean Square Deviation (RMSD), also known as the prediction errors in out-of-sample forecast, for stock returns at different forecast horizons. To be specific, the RMSD is a widely used measure to the performance of predictive ability obtained via taking the differences between the forecasting value and the real value. According to the panel, we can find that RMSDs for returns remain at a low-value level all along as the forecast horizons expand from 1 month ahead to 12 months ahead. This reveals that the consumer confidence index as the proxy for consumer confidence, has a significantly positive effect on stock returns. Moreover, we observe that the RMSD decreases from 0.7263 under 1 month period ahead forecast horizon to 0.0186 under 12 months periods ahead forecast horizon progressively. In other words, the longer the forecast horizon is, the better the rolling window forecast the model is.

As can be noted from Panel A and Panel B in table 2.4, investor sentiment is positively correlated with stock market returns, which contradicts with our hypothesis 2 and past papers. The possible reasons for the difference between our results and past papers could be various. The first possible reason is the difference in the range of the observations. We estimate the stock returns in 10 European countries while most of the previous studies only concentrate on the US stock market. Baker and Wurgler (2006) constructed a composite index combined six indirect measures for investor sentiment and explored that investor sentiment has a negative effect on stock market returns for the reason that the US stock market is speculative and hard to arbitrage. The CCI is more suitable for analysis with multiple countries where BW index is better for a well-established individual country, for example, the US stock market.

Table 2.4: Panel fixed-effect regressions with different forecast horizons

Panel A: In-sample panel fixed effect regression, 2493 observations

	Forecast horizon			
	(1 month)	(3 months)	(6 months)	(12 months)
$\Delta Sent$	0.2215***	0.0968***	0.0614***	0.0364*
	(0.0409)	(0.0247)	(0.0180)	(0.0132)
ΔCPI	0.1391	-0.1030	-0.5028**	-0.4026***
	(0.3311)	(0.2001)	(0.1457)	(0.1073)
ΔR	-0.1937 ***	-0.1127***	-0.0933***	-0.0669***
	(0.0395)	(0.0239)	(0.0173)	(0.0128)
ΔIPI	0.0479	0.1128 *	0.0560*	-0.0006
	(0.0670)	(0.0423)	(0.0071)	(0.0005)
In sample observations	2,493	2,493	2,493	2,493

Panel B: Out of sample RMSD, 600 observations

	Forecast horizon			
	(1 month)	(3 months)	(6 months)	(12 months)
Return	0.0726	0.0428	0.0975	0.0186

Panel C: Comparison panel regression, 1481 observations

	Forecast horizon			
	(1 month)	(3 months)	(6 months)	(12 months)
$\Delta Sent^{SCH}$	-0.0933	-0.0625**	-0.0559***	-0.0182***
	(0.0607)	(0.0215)	(0.0105)	(0.0059)

The table shows the results for predictive panel regression results with future stock returns as dependent variable and CCI as independent variables. Macroeconomic factors are as control variables.
Asterisks refer to the level of significance where: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.5: OAT analysis based on Panel regression

Panel A: $\Delta Sent, 2668$ *observations*	*Forecast horizon*			
	(1 month)	(3 months)	(6 months)	(12 months)
$\Delta Sent$	0.2162***	0.0956***	0.0609***	0.0355***
	(0.0399)	(0.0240)	(0.0175)	(0.0129)
Panel B: $\Delta Sent$ $\Delta CPI, 2668$ *observations*	*Forecast horizon*			
	(1 month)	(3 months)	(6 months)	(12 months)
$\Delta Sent$	0.2161***	0.0952***	0.0591***	0.0341***
	(0.0399)	(0.0240)	(0.0175)	(0.0129)
ΔCPI	-0.0048	-0.1030	-0.5352***	-0.4498***
	(0.3311)	(0.1879)	(0.1369)	(0.1006)
Panel C: $\Delta Sent$ $\Delta R, 2668$ *observations*	*Forecast horizon*			
	(1 month)	(3 months)	(6 months)	(12 months)
$\Delta Sent$	0.2216***	0.0985***	0.0635***	0.0371***
	(0.0397)	(0.0239)	(0.0175)	(0.0129)
ΔR	-0.1485***	-0.0780***	-0.0712***	-0.0437***
	(0.0308)	(0.0186)	(0.0135)	(0.0100)
Panel D: $\Delta Sent$ $\Delta IPI, 2668$ *observations*	*Forecast horizon*			
	(1 month)	(3 months)	(6 months)	(12 months)
$\Delta Sent$	0.2154***	0.0935***	0.0598***	0.0356***
	(0.0399)	(0.0240)	(0.0175)	(0.0129)
ΔIPI	0.0416	0.1140***	0.0577*	-0.0032
	(0.0679)	(0.0408)	(0.0298)	(0.0219)
Panel E: $\Delta Sent$ ΔCPI $\Delta R, 2668$ *observations*	*Forecast horizon*			
	(1 month)	(3 months)	(6 months)	(12 months)
$\Delta Sent$	0.2219***	0.0983***	0.0618***	0.0357***
	(0.0398)	(0.0240)	(0.0174)	(0.0129)
ΔCPI	0.0989	-0.0765	-0.4880***	-0.4214***
	(0.3118)	(0.1878)	(0.1366)	(0.1005)
ΔR	-0.1492***	-0.0785***	-0.0679***	-0.0408***
	(0.0309)	(0.0186)	(0.0135)	(0.0100)
Panel F: $\Delta Sent$ ΔCPI $\Delta IPI, 2668$ *observations*	*Forecast horizon*			
	(1 month)	(3 months)	(6 months)	(12 months)
$\Delta Sent$	0.2154***	0.0930***	0.0580***	0.0341***
	(0.0399)	(0.0240)	(0.0175)	(0.0129)
ΔCPI	-0.0094	-0.1438	-0.5419***	-0.4497***
	(0.3125)	(0.1878)	(0.1369)	(0.1007)
ΔIPI	0.0416	0.1147***	0.0605**	-0.0009
	(0.0679)	(0.0408)	(0.0297)	(0.0219)
Panel G: $\Delta Sent$ ΔR $\Delta IPI, 2668$ *observations*	*Forecast horizon*			
	(1 month)	(3 months)	(6 months)	(12 months)
$\Delta Sent$	0.2206***	0.0964***	0.0623***	0.0371***
	(0.0398)	(0.0240)	(0.0175)	(0.0129)
ΔR	-0.1496***	-0.0765***	-0.0724***	-0.0437***
	(0.0308)	(0.1854)	(0.0135)	(0.0100)
ΔIPI	0.0555	0.1216***	0.0644**	0.0009
	(0.0677)	(0.0407)	(0.0006)	(0.0219)

The table shows the OAT results for predictive panel regression models. Asterisks refer to the level of significance where: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Schmeling (2009) found further international evidence for the negative relationship between investor sentiment and stock market in Panel C of Table 2.4. We thus replicate exact Schmeling's Work and get the same result in aggregate market, value stock market and growth stock market. Then we use the same time period, same variables and same data source as Shmeling used and merge data into our model to check the result. To be specific, we use the monthly stock market returns collected from Professor Kenneth French's website with time spanning from January 1985 to December 2005 which is exactly the same with Schmeling's paper. In Panel C of Table 2.4, we notice that investor sentiment is negatively correlated with stock market returns especially in long forecast horizons, which is the opposite of the result we got. Also, with longer forecast horizons, the effect of investor sentiment is better based on data used by (Schmeling, 2009). In other words, investor sentiment can affect stock market returns better in long-term rather than in short-term. This proves that the relationship between investor sentiment and stock market largely depends on the time span and the measurement of stock returns.

2.7.1.2 Sensitivity analysis

From a general perspective, sensitivity analysis is one of the key ingredients needed in building models and quality assurance. A number of researchers have focused on increasing in computing power of sensitivity techniques. Rabitz (1989) pointed out that the judicious application of sensitivity analysis can reflect the maximum capabilities of economical models. Recently, more and more research institutions begin to use sensitivity analysis as one of the best practices to improve the quality of scientific modelling, (see the Florida Commission on Hurricane Loss Projection Methodology(Iman et al., 2005)). Borgonovo and Plischke (2016) also found out that sensitivity analysis shed light on model behavior, model structure and the response of independent variables to changes in a model. Based on past papers, we use one of the most common and core approaches, One-at-a-time(OAT, hereafter), to assess our model. To be specific, the OAT method requires to add input variables one by one into the model to determine whether our model is robust and how the proxy for investor sentiment contribute to the model.

Table 2.5 shows the result for OAT analysis in our model. It is easy to see that consumer confidence index, as the proxy for investor sentiment, is our key variable. In Panel A, even in the primary model without any control variables, investor sentiment is still positively correlated with returns significantly on different forecast horizons. As we add the macroeconomic factors as the control variables into the model one by one, we can notice that the relationship between investor sentiment and stock market returns is still significantly positive, which means our model is robust and performs well.

2.7.2 Results for individual regressions

2.7.2.1 Results for in-sample forecasting regressions for individual countries

The coefficients reported in Table 2.6 directly show the effect of sentiment with different forecast horizons in individual European countries. To have a better comparison with panel forecasting results, we still divide the whole sample into two parts: the in-sample and the out-of-sample with the same time period for each subsample as we mentioned above. Hence we have 296 observations for in-sample estimation test and use 70 observations for out-of-sample forecasting ability test for each country.

Furthermore, we also consider the results for the US stock market for two reasons. Firstly, the US stock market is one of the most influential economic system in the world. Importing the result for US proves that our model can be widely used in well-developed stock markets in the world which makes our findings more convincing. Another reason is that considering most past papers focusing on the US stock market, adding the result for that US stock market follows the pioneers' findings and provides a better comparison with past papers. To simplify the table, we only present the coefficient of proxy for investor sentiment over forecast horizons of 1,3,6 and 12 months. The null hypothesis for this test is that the coefficients for all countries are zero. The p-value of the test is the probability of the hypothesis above. As shown in the table, a positive effect of consumer confidence index as a proxy for investor sentiment on stock returns is found for all these ten European countries at a 5%-level of significance in one month ahead forecast horizon. For the other three forecast horizons, the numbers of countries with significant sentiment-return relationship at a 5%-

level are 8, 7 and 6 respectively.

Overall, we can conclude that investor sentiment has a significantly positive effect on future stock returns at different forecast horizons (from 1 to 12 months) across European countries, which is in line with the results we get for our panel prediction. Considering the results for the US stock market, we find that the result follows the same logic as the European countries, which is in line with past papers. However, this result contradicts with our Hypothesis 2. The possible reason for this is that we use macroeconomic factors as control variables. Another potential reason is the different data frequency. Moreover, a quite interesting finding is that the effect of the investor sentiment on future stock returns diminishes across forecasting horizons, which is consistent with our Hypothesis 3. For example, a rise in proxy of sentiment can cause approximately 0.56% increasing in stock returns for 1 month forecast horizon in Austria. While the coefficient declines to 0.11% for 12 months forecasting horizon in the same country and the significance level also changes from 10%-level to less than 1%-level. In economic terms, this phenomenon can be explained by the noise trader theory (Schmeling, 2009). To be specific, with smaller arbitrage opportunity in longer time horizons, the effects of noise traders are washed out as time passes. The explanation seems quite convincing from the view point that stock market will always reach to an equilibrium in long time periods. Furthermore, although investor sentiment has a significant positive impact on future stock returns for most countries, we can still observe that such relationship seems not to hold for some specific European countries, especially in long forecast horizons. Thus, evidence cannot prove the relationship to the specific characteristics of a country.

2.7.2.2 Determining the performance of predictions

We examine the stock return out-of-sample predictive performance using correct sign prediction and the root mean square deviations (RMSDs). Table 2.7 shows the proportion of correct sign prediction while Table 2.8 and 2.9 make a comparison between the individual regression and panel coefficients of different forecast horizons in out-of-sample time periods.

The correct sign prediction, known as the possible indicator to measure the direction changes of the predictions, is widely used in many past literature(Pesaran and Timmermann, 1992; Gencay, 1999; Shahriari et al., 2016). Pesaran and Timmermann (1992) give the definition of the correct sign predictions which can be expressed as the following equation:

$$correct\ sign\ predictions = \frac{1}{T-(T_1-1))}\sum_{t=T_1}^{T} z_{z+s} \tag{2.3}$$

where

$$z_{z+s} = \begin{cases} 1 & \text{if } (y_{t+s}f_{t,s}) > 0 \\ 0 & \text{if } (y_{t+s}f_{t,s}) \leq 0 \end{cases}$$

Hence, it is possible to observe the proportion of correctly predicted signs and directional changes for the out-of-sample test time periods. After that, Gencay (1999) also used the

Table 2.6: Return predictability of individual countries across horizons

	Forecast horizon			
	1 month	3 months	6 months	12 months
OE	0.0059 ***	0.0045***	0.0029***	0.0011
	(0.0019)	(0.0012)	(0.0010)	(0.0007)
BG	0.0031 ***	0.0026 ***	0.0013**	0.0010**
	(0.0012)	(0.006)	(0.0005)	(0.0004)
FN	0.0042 **	0.0026	0.0027*	0.0022
	(0.0041)	(0.0022)	(0.0016)	(0.0013)
BD	0.0013***	0.0021 **	0.0013*	0.0006
	(0.0017)	(0.0009)	(0.0007)	(0.0005)
GR	0.0044**	0.0024**	0.0018**	0.0009*
	(0.0019)	(0.0010)	(0.0009)	(0.0007)
IT	0.0006***	0.0015*	0.0002	0.0001
	(0.0017)	(0.0008)	(0.0007)	(0.0005)
NL	0.0029***	0.0021***	0.0017***	0.0011***
	(0.0011)	(0.0005)	(0.0004)	(0.0003)
PT	0.0030***	0.0029***	0.0019***	0.0016***
	(0.0016)	(0.0009)	(0.0007)	(0.0006)
ES	0.0029***	0.0024***	0.0014***	0.0009 ***
	(0.0015)	(0.0007)	(0.0005)	(0.0004)
UK	0.0008***	0.0016***	0.0009***	0.0006***
	(0.0010)	(0.0005)	(0.0004)	(0.0003)
US	0.0548***	0.0391***	0.0182***	0.0168***
	(0.0137)	(0.0061)	(0.0051)	(0.0039)

The table shows results for predictive individual regression results with future stock returns as dependent variable and sentiment as well as macroeconomic factors as predictive variable.
Asterisks refer to the level of significance where: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.7: Correct Sign Prediction Result

	Forecast horizon			
	1 month	3 months	6 months	12 months
OE	0.49	0.53	0.57	0.51
BG	0.57	0.62	0.80	0.80
FN	0.60	0.65	0.54	0.64
BD	0.59	0.65	0.68	0.78
GR	0.54	0.66	0.71	0.61
IT	0.59	0.56	0.60	0.64
NL	0.46	0.66	0.68	0.69
PT	0.56	0.57	0.60	0.49
ES	0.47	0.54	0.62	0.64
UK	0.61	0.65	0.65	0.64
US	0.54	0.74	0.81	0.91

The table shows the proportion of the correct sign prediction across forecast horizons in out-of-sample time periods

sign predictions to examine the performance of out-of-sample linear and non-linear forecasting models. Furthermore, Barassi and Zhao (2017) used the sign prediction technique to forecast the energy demand in the UK.

Table 2.7 presents the percentage of the correct signs in the out-of-sample period. We can observe that most of the correction sign prediction reaches to a level of 50%. For example, the scale of sign predictions across different forecast horizons are 49%, 53%, 57% and 51% respectively in Austria. We also test the correct signs result in the US stock market which is above 50% overall. This indicates that our model can also reflect the environment of US stock market. In other words, the results increase the reliability of our forecasting model. However, there is no strong pattern of the proportion of sign prediction across forecast horizons.

Table 2.8: RMSDs of Individual Countries

	Forecast horizon			
	1 month	3 months	6 months	12 months
OE	0.0632	0.0342	0.0240	0.0164
BG	0.0480	0.0229	0.0171	0.0132
FN	0.0608	0.0311	0.0242	0.0165
BD	0.0557	0.0268	0.0188	0.0106
GR	0.2366	0.1260	0.0641	0.0592
IT	0.0606	0.0323	0.0210	0.0144
NL	0.0463	0.0228	0.0163	0.0118
PT	0.0699	0.0355	0.0272	0.0292
ES	0.0657	0.0306	0.0205	0.0169
UK	0.0445	0.0185	0.0115	0.0064
US	0.0365	0.0143	0.0116	0.0073

The table shows the RMSDs of individual countries across different forecast horizons in out-of-sample time periods

Table 2.8 presents the Root Mean Squared Deviation (RMSDs) of individual countries in out-of-sample time periods. As mentioned above, we have 60 out-of-sample observations for each of these ten countries. Normally, we use RMSDs to figure out the standard deviation of the unexplained variance. In other words, the smaller the RMSDs are, the better the model fits. Generally speaking, the RMSDs are quite small for all of the countries in table 2.8, which means we can predict the stock market returns with investor sentiment and our result is a relatively powerful one. Furthermore, we can find that RMSDs declines as time passes, which means the predictability of investor sentiment becomes better in longer forecast horizons.

To compare with the panel forecasting result, we substitute the coefficients in panel level

Table 2.9: RMSDs of Individual Countries Based on Panel Coefficients

	Forecast horizon			
	1 month	3 months	6 months	12 months
OE	0.0705	0.0360	0.0329	0.0167
BG	0.0748	0.0367	0.0339	0.0194
FN	0.0700	0.0378	0.0328	0.0181
BD	0.0681	0.0331	0.0324	0.0140
GR	0.1590	0.0788	0.0495	0.0507
IT	0.0749	0.0383	0.0340	0.0159
NL	0.0851	0.0434	0.0362	0.0198
PT	0.0905	0.0443	0.0373	0.0212
ES	0.0918	0.0415	0.0376	0.0203
UK	0.0720	0.0335	0.0333	0.0121

The table shows the RMSDs of individual countries based on the panel coefficient across different forecast horizons in out-of-sample time periods

into the individual countries and report the new RMSDs in Table 2.9. As we can observe, most of the RMSDs for individual countries based on panel coefficients are still less than 0.1, which means that the result for forecasting model is convincing. If we make a comparison between table 2.8 and 2.9, it is interesting to find that the RMSDs in individual levels are smaller than the RMSDs in panel level. This result proves the significant difference among these European countries.

2.8 Conclusion

This chapter has aimed to contribute to the empirical literature examining the effect of investor sentiment on aggregate stock market returns among European countries. As investor sentiment is an unobserved and subjective variable, it is crucial to find a suitable proxy for it. Especially in multiple stock market, it is hard to find a unified measurement for sentiment among different countries. After comparing various of measures for sentiment, we select a widely used proxy for investor sentiment, the consumer confidence index, which seems to be the only standardized direct measurement among different countries.

We find that there is a significant positive impact of investor sentiment on stock returns. Additionally, in order to investigate the sentiment-return relation in different countries, we employ both panel regression and individual regression in empirical analysis. Based on the forecasting model, we conclude that comparing with short-term or mid-term individual stock market (such as the US stock market), investor sentiment is significantly and positively correlated with future stock returns. In other words, future stock returns will rise with higher expectation of investors. After comparing the results for panel and individual regression, we find that there is no country-specific effect on the sentiment-stock relation in European stock market.

Overall, our research have three creative points. Firstly, we expand the research target to 10 European countries rather than individual country, which provides a chance to get a contradictive result compared to prior works that focus on the US Stock market. Furthermore, we introduce macroeconomic factors as control variables into the model thereby reducing the chance of misspecification of the model. Finally, we compare the panel regression results from individual regressions proving the significant effect of sentiment on stock market across countries.

Results suggest several avenues for future work. A better understanding of investor sentiment may shed light on patterns in aggregate stock market and prevent the market from sudden shock. Furthermore, the results suggest that descriptively accurate models of prices and expected returns need to incorporate a prominent role for investor sentiment.

Chapter 3

Forecasting Stock Market Returns with Univariate and Multivariate Models

3.1 Introduction

The stock market is always one the most studied markets in economics for the reason that it holds the highest amount capitals in trading. Also, the stock market is closely connected with the development of the economy, monetary policy, natural environment or even cultural differences (Schmeling, 2009; Kong et al., 2011; Phan et al., 2015). For several decades, numerous academic financial economists focused on proving that stock returns are predictable. Pioneering work supported the Efficient market hypothesis (EMH). This is one of the most popular theory in classic financial economics, which states that stock prices can be fully predictable via changes of discount rate and stock market information. The EMH is based on the perfect market assumptions. To be specific, there are no costs of capital distribution and information acquisition for investors under this assumption. In other words, prices of capitals can fully reflect stock market information and a stock is always traded around its fundamental value. Although the EMH allowed investors to be irrational and behave erratically, it strongly required the market to be orderly and well-organized, leaving no chance to arbitrage in efficient market. Apparently, the sentiment-return relation has no effect in this framework based on the assumption that stock market is always self-regulated. According to this theory, it is meaningless to trying predict any changes in future stock returns via historical pricing movements for the reason that prices could only be affected by the stock internal market pattern. However, this theory is strongly controversial. Believers argue it is not necessary to forecast future stock market and focus on the certain undervalued stocks which only account for as more in stock market. While detractors point to a series of crashes which happened in history of financial markets, such as the Kennedy Slide of 1962 (also know as 'the Flashing Crash of 1962'), the Brazillian Markets Crash of 1971, the dramatic rise in oil price in UK of 1973 and the Black Monday of 1987, and so on. These crashes proved that EMH failed to explain the abnormal stock price. As a consequence, more and more researchers began to realize that predicting the trend of stock market can shed light on abnormal phenomena and prevent stock market from getting crashes again. In early 1990s, De Long et al. (1990) broke the EMH theory and pointed out that stock returns were predictable via financial fundamentals in the U.S stock market. Later, Baker and Wurgler (2000); Baker and Stein (2004); Brown and Cliff (2004) also raised the idea that stock returns were predictable and positively correlated with macroeconomic variables and financial indicators. However, controversial opinions also appeared in the same period. Kothari and Shanken (1997) applied financial indicators and economic variables collected from small firms to prove that returns are not always predictable in specific period. Besides, Bossaerts and Hillion (1999) revealed that returns were not self-predictable via out-of-sample forecasts, which provided a contradictory result compared with past papers.

In this chapter, we aim to prove that stock returns are predictable in European countries. Taken on the basis of previous literature, we apply both univariate and multivariate models and examine their forecasting performance for stock returns. As univariate model, we use the Auto-regressive Moving Averaging model(ARMA, hereafter). As we know, ARMA is one of the most common method in modelling and forecasting, and has been widely used in numerous past papers (De Gooijer, 1989; Karanasos, 2001; Henry, 2002). There are also extension models of the basic ARMA model, for example, the ARIMA and ARMAX models, are also adopted extensively in many previous studies (Hannan et al., 1980; Johansen and FOSS, 1993; Lim et al., 2009). The main advantage of ARMA model is that moving average and autoregression are combined into the same process to improve the accuracy of the forecasting.

However, a lot of literature pointed out that the univariate model could only predict the stock market with its own past. They applied various multivariate model to make forecasting of stock returns. Schmeling (2009) used a panel VAR model to investigate the relationship between investor sentiment and stock market returns internationally. Pradhan et al. (2013) also applied a panel VAR model to observe the impact of stock market development on inflation and economic growth of sixteen Asian countries. They revealed that these variables are cointegrated, suggesting presence of a long-run equilibrium relationship among them. The main advantage of VAR model are that they can consider all of the variables as endogenous and merge them into a system to forecast. In other words, we can predict the dependant variable not only related with its own previous value, but also with lags of other variables. Based on the standard VAR model, the Bayesian VAR model sets the model parameters as random variables with assigning probabilities estimated by Bayesian method. The Bayesian VAR model is widely applied in macroeconomic variable forecasting. Carriero et al. (2009) forecasted the panel exchange rates with a large Bayesian VAR with driftless random walk prior. They proved that BVAR performed better than the univariate ARMA and ARMAX, especially in short horizons.

The first contribution of this chapter is that we forecast stock returns with ARMA, ARMAX, VAR and the BVAR. Although various of previous literature built these models to predict stock returns, they never applied both of them together. As usual, we generate in-sample model forecasts and examine the accuracy of the prediction via out-of-sample forecasting. Specifically we forecast the stock returns in ten European Countries instead of a single country. The aim of doing this is to find the common point and the different characteristics in different countries. As can be noticed, most of the literature only focused on the American stock market or a single developed stock market (De Long et al., 1990; Fama and French, 1998; Brown and Cliff, 2005; Ang and Bekaert, 2006; Huang et al., 2015; Zheng and Zhu, 2017). Thus enriching the investigation scope is also a contribution of this chapter.

The rest of this chapter is organized as follows: Section two provides the theoretical background and a general review of the literature. The data description and some primary test results are shown in the third section. In section four, we introduce the process of building each forecasting model and choosing the optimal model via information criteria. Section five contains the analysis and forecasting result. Also, we make comparison of the out-of-sample forecasting results to figure out the prediction performance and a summary concludes.

3.2 Theoretical Background

Investigation of the stock market is always an essential active part of the financial economic area mainly for the reason that it holds the greatest amount of capital compared with other types of financial markets. Furthermore, it plays the indispensable role in attracting and allocating the distributed savings and liquidity into the most profitable financial activities so that financial resources are adequately utilized in the national economy. In the stock market, both rational investors and speculators are required to determine the fundamental value of stocks and then compare it with the current market price so that they can trade in a better price. Thus, investigating the stock market is very important for both researchers and individual investors alike.

3.2.1 Predicting Stock Market Returns

The optimal allocation of resources is highly determined by the reaction of investors in the stock market. In other words, to understand the behaviour of stock market, it is necessary to predict the stock market returns by analyzing market information. Numerous studies have focused on predicting stock market returns over the past decades. Campbell (1987) showed that excess returns on bills, bonds and stocks are surprisingly predictable. Next year, Campbell and Shiller (1988) further proved that stock returns are predictable by forecasting the dividend price ratio one period ahead in the US stock market. They also proved that the stock market was predictable using real dividend growth, measured real discount rates, and unexplained factors, which yield the metric to judge the measured financial factors. Similarly, Fama and French (1988) pointed out that stock market was strongly influenced by dividends and explained the temporary effect of future stock returns on the stock market.

Later in 1990s, Fama and French (1992) predict the future stock returns with more financial factors, which are market β, size, leverage and book-to-market ratio respectively. Their work strongly proved that stock return is predictable in cross-section level. Kothari and Shanken (1997) and Pontiff and Schall (1998) proved that book-to-market (BM, hereafter) ratio and dividend yield could predict the stock market returns. However, Kothari and Shanken (1997) showed that stock market was not always efficient, especially in certain periods. The reason for it could be that dividend yields and BM were mostly influenced by smaller firms, so the characteristic of these financial variables determined the power of predictability. However, most of the literature in that period only found limited evidence of stock market return (US market) predictability via using in-sample estimation. So there even were some studies showing in-sample results could not be used for ou-of-sample(Bossaerts and Hillion, 1999).

In recent years, more and more researchers realize that it is more persuasive to use both in-sample and out-of-sample techniques to predict future stock returns. To be specific, Goyal and Welch (2003) suggested a simple, recursive residuals (out-of-sample) approach to evaluate the predictive ability of equity premium and stock returns in the U.S stock market. However, the out-of-sample results they got reflected that dividend ratios for U.S stock market were not predictable both in the short and long time horizons. To examine the out-of-sample forecasting performance of stock returns, they comprehensively reexamined the performance of the predicting models for stock returns by a variety of macroeconomic factors and financial variables. Unfortunately, they still found that the predictive regression models were not stable and were unable to beat the simple historical average benchmark forecast both in and out-of-sample. In historical average predictable regressions, one of the main assumptions was that the coefficients of the predictors should be equal to zero, in other words, the information from financial variables were not useful for predicting stock returns. They also attempted to organize all of the predictors into a single forecasting regression known as the "kitchen sink model". It is not surprising that the performance of this model was even worse than multiple forecasting regressions for the reason that in-sample is over-fitting in this model (Welch and Goyal, 2007). Later, these findings from (Welch and Goyal, 2007) have encouraged more attentions. Campbell and Thompson (2007) proved that predictability of macroeconomic factors were superior than the historical used fundamentals by generating an economically-motivated restricted model. According to their research, the weak restrictions are imposed on the signs of coefficients and the forecast returns must be non-negative. The restricted model improves the forecasting ability of predictors in out-of-sample and in some cases can beat the historical averaged benchmark. The more important breakthrough is that the out-of-sample tests are economically meaningful for mean-variance investors by imposing the restrictions of steady state valuation models in short horizons to reduce the volatility of stock returns. To improve the predicting ability of single forecasting regression models, Rapach et al. (2010) combined numerous variables into a new forecasting regression model to reduce the reduce the forecasting volatility to some extent. Their combination predictive model includes fifteen economic and financial variables, which can improve the accuracy of out-of-sample prediction and precede the historical averaged models in a consistent way over different time periods.

Additionally, several studies began to forecast the stock market in a more consistent way. Ang and Bekaert (2006) proved that stock returns were predictable not only in US stock market, but also in France, Germany, Japan and the UK. They used a forecasting regression with three instruments, which are short rate, dividend yield and earnings yield respectively and the result showed that only in short-run the out-of-sample tests are robust and stable. In long-horizon time periods, they proved that there was no evidence of predictability of stock returns. Generally speaking, the ability of predictor cross-countries were stronger compared with historical average predictors using local instruments (Ang and Bekaert, 2006). Also some researchers focused on certain industries and changed the data frequency. Phan et al. (2015) focused on the out-of-sample predictability of stock returns in oil industries. Then by using daily, monthly and quarterly data, Phan et al. (2015) tested whether different data frequency would affect the power of predictability of stock returns. They focused that stock returns were always predictable under different data frequencies and the predicting ability depended mainly on the characteristics of the explanatory variables.

In this chapter, we use in-sample forecasting regression to figure out the predictability of stock returns and use out-of-sample forecasts to evaluate our results. We use data on different countries rather than confining analysis to a single stock market to international stock market. In addition, we forecast the regression for each countries individually, thereby comparing the results for different countries.

3.2.2 Estimators for Stock Market Returns

In past papers, researchers used various explanatory variables to predict stock returns. These variables, which are highly interrelated factors in economics, can be divided into three categories, which are economic, financial specific variables and proxies for investor sentiment respectively. As financial variables are unlikely to be standardized across all of the ten European countries object our interest, we will apply both economic variables and proxies for investor sentiment to predict stock returns and make a full comparison of the results. In this part we will discuss these variables and explain the reason that we choose them.

3.2.2.1 Investor Sentiment

The failure of EMH made researchers to realize that irrational factors, such as investor sentiment, also plays an important role in financial market. As investor sentiment is a totally unobserved and subjective variable, the main difficulty of using investor sentiment as a variable is to determine a proxies for it appropriately. Based on prior work, measurements for investor sentiment can be divided into three categories, which are direct, indirect and composite measurement respectively [1]

.

Most of the direct sentiment measures are taken from a series of surveys generated by various associations through various questionnaires both for institutions and individual investors. As introduced in chapter 2, one of the most popular survey is created by the American Association of Individual Investors (AAII), which collects the answers from random individual participants each week since 1987. This survey aims to split the response from investors into bullish, neutral, bearish types. Another representative survey is Investor Intelligence (II) which categorizes over 100 newsletters to compile the bull-bear condition in the stock market each week. As the majority authors of these newsletters are financial professionals, researchers always regard this survey as the sentiment of institutions. These two direct measures seem to reflect the investor sentiment in a more straightforward way. However,they are limited since it can only be applied to the US market. To spread the research target to international stock market, it is necessary to find a more consistent way to measure investor sentiment.

Compared with direct measurement, the Consumer Confidence Index (CCI,hereafter), is reasonable to be taken as the proxy for investor sentiment when we focus on stock markets in different countries for the main reason that it can be standardized across countries compared with other direct measurements (Charoenrook, 2005; Lemmon and Portniaguina, 2006; Schmeling, 2009).

[1]Refer to 2.4 Measurements of investor sentiment in Chapter 2

To be specific, it is the only survey with standardized questionnaires which is distributed internationally. The question pool requires participants to give marks in all questions which contain four topics: the current financial conditions and expectation of the future financial conditions, the current general economic state and expectation of the future general economic state, the current employment and the employment status in the future, satisfaction of current savings and the expected savings in the future year. From these questions, we can clearly see that these indicators are conducted to reflect investor sentiment both in the current period and the future in a more comprehensive way. Additionally, CCI is clearly correlated with many economic variables and financial factors which improves its suitability as a proxy for investor sentiment.

The indirect sentiment measures are some financial indicators taken as the market signals, and could be more flexible and less strict compared with direct sentiment measurement. Previous work suggest a number of proxies for investor sentiment to use in time-series models. However, it is still controversial which financial factor can best represent an indirect measure of sentiment. Here we briefly introduce the most widely used indirect measures in past papers and compare them with the direct sentiment measures. The most common measurement that has been used is the closed-end funds discount (CEFD, hereafter), which refers to the average level of the difference between net asset value (NAV) of closed-end fund shares and their current market price (Zweig, 1973). It is marked as 'discount' if NAV is higher than market price and defined as 'premium' on the contrary (Zweig, 1973). Many researchers provide empirical evidence to explain that CEFD is caused by the expectations of irrational investors in stock market (Lee et al., 1991; Chen et al., 1993). Later, Neal and Wheatley (1998), Qiu and Welch (2004) and Ben-Rephael et al. (2012) discussed whether CEFD can be actually considered as an indirect measure of investor sentiment. Another widely used indirect measurement of investor sentiment is the trading volume. It is also known as the share turnover in practical research, refers to the number of shares traded in a certain market during a given period of time (Jones, 2002). To be specific, it based on the ratio of volume to average market shares, which is suitable to be used as a proxy for investor sentiment for the reason that it reflects the higher liquidity brought by irrational investors. In other words, higher liquidity in stock market represents the symptom of overvaluation of a stock price(Baker and Stein, 2004). Besides, there are numbers of other financial indicators which are taken as the reflection of investor sentiment to some extent, like the initial public offering (IPO), the Number of IPOs (NIPO) and the equity share and the dividend premium (PD) (Ritter, 1984; Ljungqvist et al., 2006; Baker and Wurgler, 2000; Fama and French, 2001; Baker and Wurgler, 2004, 2007). IPO refers to the stock of a private company is offered to the public for the first time. Prior work suggested that both IPO and NIPO can be cited as sentiment measurements. Ritter (1984) pointed out that higher IPO and NIPO could reflect the optimistic and pessimistic expectation of stock market. Later Ljungqvist et al. (2006) also demonstrated that IPO and NIPO are quite sensitive to irrational traders sentiment for the reason that first-day IPOs are always under-priced based on risk aversion investors consideration. Baker and Wurgler (2007) subsequently explored the impact of first-day IPO returns (RIPO) on investor sentiment. The equity share in new issues is another common financial indicator that can be regarded as an indirect measure of sentiment. The definition of it is the gross equity issuance divided by the total of gross equity and debt issuance. Baker and Wurgler (2000) used such indicator to measure sentiment and documented that the effect of values of equity share on stock market returns is negative. The dividend premium (PD) refers to the log difference between the average book-to-market ratios for payers and non-payers. Fama and French (2001) explained the declining incidence of dividends and found that enterprises prefer to pay dividend under a premium situation rather than a discount one. This phenomenon showed that PD is sensitive to sentiment especially when companies need to make important financial decisions.

From these considerations, we can see that each of these measures based on financial indicators is likely to include a sentiment component. The best way to explain these indicators at the same time is to combine them as a composite index. However, since our analysis focuses on multiple stock markets, it is quite difficult to standardize these indicators at a balance value for the reason that different countries have different measurement for financial indicators. Hence indirect and complex measurements are well suitable in single countries. In our research, we still use consumer confidence index as the proxy for investor sentiment since it seems that CCI is the only appropriate measures across countries.

3.2.2.2 Financial and Economic Fundamentals

We could also employ several financial fundamentals as predictors of stock returns, which are book to market ratio, earnings-price, cash earnings to price and dividend yield respectively. Several studies have proved that these financial ratios can predict both short-term and long-term stock market returns(Fama and French, 1988; Campbell, 1987; Goyal and Welch, 2003; Rapach and Wohar, 2006; Kong et al., 2011; Westerlund and Narayan, 2014; Phan et al., 2015). Book to market ratio (B/M) refers to the common shareholders equity divided by the market cap. This ratio is normally applied to find the fundamental value of an enterprise by comparing the book value with the market value. The book value is based on the general managing cost or the accounting value of an enterprise whereas the market value is determined by the marketing capitalization. Kong et al. (2011) find that BM can have positive effect on stock market returns. Additionally, they also prove that BM can predict stock returns better for the portfolio with lower market capitalisation and liquidity. Earnings-price (P/E) measures the current per-share earnings to the current share price, which is also known as the price multiple. P/E is widely used by analysts in determining the relative value of enterprise shares or in comparing the aggregate market. Normally, a higher P/E represents a higher participation of investors in stock market. Westerlund and Narayan (2014) pointed out that P/E can definitely predict stock returns and it is positively correlated with stock returns. Cash earnings to price (CE/P) refers to the ratio between the operating cash flow and the number of shares outstanding. This ratio is a profitability ratio which measures the financial performance of enterprises. Fama and French (1998); Campbell (1987) used the CE/P to measure the financial performance of enterprises. They take CE/P as a profitability ratio and proved that CE/P can also significantly affect the US stock returns in the short term. The dividend yield (D/P) refers to the ratio of the company's dividend over its share price. D/P, as one the the most fundamental financial indicators, have been used in large numbers of previous paper to prove the predictability of stock market returns (Goyal and Welch, 2003; Westerlund and Narayan, 2014; Phan et al., 2015). Compared with financial fundamentals, the economic fundamentals, especially the macroeconomic factors, are better to merge into the forecasting models. Many previous studies have shown that macroeconomic fundamentals could be used as predictors for stock returns (Bossaerts and Hillion, 1999; Welch and Goyal, 2007; Westerlund and Narayan, 2014). Bossaerts and Hillion (1999) used simple in-sample forecasting model to prove that economic indicators could predict the stock returns successfully. Welch and Goyal (2007) examined the predictive performance of the equity premium by using the out-of-sample forecasts and also proved that macroeconomic factors could forecast the stock efficiently.

In this chapter we use three macroeconomic fundamentals to predict stock returns, namely monthly consumer price index (CPI), monthly treasury interest rate (IR) and monthly percentage changing of industrial production index (IP).

3.3 Data

3.3.1 Descriptive statistics

As previously noted, we are interested in modelling and forecasting stock market returns across countries. Doing this in a consistent way , we use not-seasonally adjusted monthly stock return index as our dependent variable. The stock returns are from the electronic version of Morgan Stanley's Capital International Perspectives (MSCI). The most outstanding advantage for the MSCI data compared with other international database is that they can eliminate the survivor bias comparatively (Fama and French, 1998). To be specific, the MSCI database includes not only the currently trading firms, but the historical data for disappearing firms as well. Our monthly returns, which can be collected from Datastream in a consistent manner for different countries, ranges from May 1985 through December 2015 wherever possible.

In implementing our analysis, first we introduce the proxies for investor sentiment. Following pioneering works, there are numbers of proxies for investor sentiment, both direct and indirect ones. However, after comparing all these sentiment measurement, it seems that CCI is the only suitable proxy for investors sentiment that is common to different stock markets (Lemmon and Portniaguina, 2006; Schmeling, 2009; Chang et al., 2012). For this study, we require to measure investor sentiment in 10 different countries, which are Austria (OE), Belgium (BG), Finland (FN), Germany (BD), Greece(GR), Italy (IT), Netherlands (NL), Portugal(PT), Spain (ES) and the United Kingdom (UK) respectively.

Doing this in a consistent way calls for the measurement to be comparable across countries and acquirable for the specified time duration. The data on consumer confidence index comes from the European Commission website, which professionally conducts surveys and provides financial indicators for European countries. Moreover, the content of the questionnaires for consumer confidence index can be acquired as well. The data range of consumer confidence index is consistent with the dependent variable, stock returns. However, the data limitation issue enforces somewhat a shorter period for some countries that we will list in descriptive statistics table (table 3.1).

Additionally, to make a better comparison with different type of estimators, we use the consumer price index (CPI), three months treasury bill interest rate (IR), industrial production index (IP) respectively. Clearly, these factors may contribute to explain and forecast stock returns and numbers of prior work also proved this point (Bossaerts and Hillion, 1999; Welch and Goyal, 2007; Westerlund and Narayan, 2014). Also, they can be correlated with investor sentiment although not perfectly.

Panel A of the table 3.1 shows the uni-variate summary statistics for all of the data in ten countries in levels and essential features are displayed in the summary data below. We also summarized the statistics for all countries in differences in panel B of table 3.1. From the table, we present important features of return based on stock prices in table 3.1 and the differentiated macroeconomics factors. The time period of monthly data is from May 1985 through December 2015, giving a total of 368 observations for each time-series.

Table 3.1: Summary Statistics for All Countries

Panel A: Summary statistics for all countries in levels

Country	Label	Start	Price mean	Price Std	CCI mean	CCI Std	CPI mean	CPI Std	IR mean	IR Std	IP mean	IR Std
Austria	OE	1985M05	501.7500	238.5607	-2.0619	7.4498	84.0527	15.3937	4.7883	2.5732	76.7806	22.6077
Belgium	BG	1985M05	766.3796	332.9616	-6.9567	8.3107	83.3935	15.1176	5.6141	2.4647	77.0425	17.4289
Finland	FN	1995M11	465.6659	352.3154	-1.6467	3.8952	85.8350	14.6096	5.8165	3.7386	81.6005	20.7711
Germany	BD	1985M05	562.6675	265.0555	-8.0714	7.1449	85.6557	13.8313	4.8345	2.1387	89.0997	13.4326
Greece	GR	1988M01	635.9520	550.7181	-31.8666	19.2422	67.6783	28.8649	7.6197	5.3030	106.8248	11.3246
Italy	IT	1985M05	774.2981	331.2151	4.9042	8.7430	79.6281	19.7426	6.1747	3.8154	104.6948	9.7152
Netherlands	NL	1985M05	855.5219	439.6469	0.8123	19.2055	83.6066	11.2678	4.9832	2.0860	85.6289	11.2678
Portugal	PT	1988M01	127.8210	51.9672	-13.3928	12.1429	76.6693	23.9569	5.9842	3.5186	106.8605	13.1616
Spain	ES	1985M05	660.4650	347.4518	-11.8194	9.6582	76.7201	21.2798	7.2243	3.8501	101.6653	12.8917
UK	UK	1985M05	1395.1280	528.8090	-8.5750	7.5435	81.8147	17.5049	6.1734	2.8516	101.9236	6.6103

Panel B: Summary statistics for all countries in differences

Country	Label	Start	Return mean	Return Std	CCI mean	CCI Std	CPI mean	CPI Std	IR mean	IR Std	IP mean	IR Std
Austria	OE	1985M06	0.0024	0.0692	0.0002	0.0284	0.0015	0.0032	-0.0040	0.1168	0.0028	0.0152
Belgium	BG	1985M06	0.0045	0.0566	0.0013	0.0348	0.0016	0.0027	-0.0047	0.0752	0.0020	0.0225
Finland	FN	1988M02	0.0070	0.0833	-0.0001	0.0203	0.0013	0.0030	-0.0042	0.1253	0.0028	0.0152
Germany	BD	1985M06	0.0045	0.0583	0.0004	0.0269	0.0014	0.0032	0.0011	0.2024	0.0014	0.0015
Greece	GR	1988M02	-0.0013	0.1033	-36.5351	16.8112	67.6783	28.8649	7.6197	5.3030	106.8248	11.3246
Italy	IT	1985M06	0.0031	0.0667	0.0010	0.0362	0.0020	0.0020	-0.0059	0.0524	-0.0002	0.0130
Netherlands	NL	1985M06	0.0051	0.0515	0.0008	0.3686	0.0015	0.0044	-0.0033	0.0910	0.0013	0.0266
Spain	ES	1985M06	0.0055	0.0643	0.0013	0.0390	0.0027	0.0048	-0.0040	0.0521	0.0005	0.0171
Portugal	PT	1985M06	-0.0006	0.0602	-21.5276	13.7400	76.6693	23.9569	5.9842	3.5186	106.8605	13.1616
UK	UK	1985M06	0.0041	0.0451	0.0012	0.3384	0.0022	0.0042	-0.0040	0.0451	0.0003	0.0010

However, given some data limitation, some variables are not available for full sample. In this situation, the analysis of these countries is performed for slightly shorter time periods. The starting date for each country is provided on the table in the fourth column.

Table 3.2: Unit Root Test Results

	Im, Pesaran and Shin		*ADF Fisher Chi-sq*		*PP Fisher Chi-sq*	
	Statistic	Prob	Statistic	Prob	Statistic	Prob
OE	-15.668	0.0000***	255.211	0.0000***	655.230	0.0000***
BG	-32.186	0.0000***	548.710	0.0000***	659.702	0.0000***
FN	-28.924	0.0000***	452.481	0.0000***	678.733	0.0000***
BD	-18.834	0.0000***	301.476	0.0000***	716.332	0.0000***
GR	-16.357	0.0000***	257.495	0.0000***	569.502	0.0000***
IT	-22.592	0.0000***	346.826	0.0000***	697.158	0.0000***
NL	-24.528	0.0000***	420.703	0.0000***	583.324	0.0000***
PT	-24.573	0.0000***	399.343	0.0000***	590.213	0.0000***
ES	-23.684	0.0000***	368.345	0.0000***	681.750	0.0000***
UK	-21.566	0.0000***	336.612	0.0000***	708.114	0.0000***

This table represent the unit root test results in ten European countries countries individually. Asterisks refer to the level of significance where: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

3.3.2 Results for unit root tests

Since our variables in levels in table 3.1 are well known to be non-stationary, it is necessary to test whether their first differences are stationary or not. Following past papers, we employ the Im, Pesaran and Shin (IPS) test, ADF-Fisher test and the PP-Fisher unit root tests to examine the individual unit root process(Schmeling, 2009). The lag length selection is based on the modified Akaike information criterion (AIC) and the test model includes individual intercepts. The results are shown in table 3.2 which present that for both panel data and the individual time-series unit root test, the variables we use are stationary in first difference.

3.4 Methodology

In this section, we will outline the methodology that we use to obtain the forecasting values of stock market returns. To be specific, we will explain each time series model we use and then introduce the information criteria which will be applied for optimal model selection. Also, we adopt some quality measures for the chosen forecasting models.

3.4.1 ARMA

The first model we will apply is the Auto-Regressive Moving Average (ARMA, hereafter) model. It is one of the random time-series models developed by Peter Whittle in 1951 and publicized by (Box et al., 1970). It is also the most commonly used forecasting model in financial economics and usually constitutes a benchmark against which other forecasting techniques are compared. Henry (2002) used ARMA models to forecast stock market returns. Also, Anaghi and Norouzi (2012) employed the ARMA class of models to predict stock price. In this paper, we remove the non-stationary part in estimators and organise the following ARMA(p,q) parameterisation:

$$r_t^{(n)} = \eta^{(n)} + \sum_{i=1}^{p} \gamma_i^{(n)} r_{t-i}^{(n)} + \sum_{j=1}^{q} \theta_i^{(n)} \varepsilon_{t-i}^{(n)} + \varepsilon_t^{(n)} \tag{3.1}$$

In this model, r_t represents the stock market returns as dependent variable with the auto-regressive(AR) order of p and the moving average order(MA) of q. n represents different countries. ε_{t-i} are stationary white noises. The model is estimated by ordinary least squares and selected from the set of ARMA(p,q) with p and q = 0, 1, ..., 5 giving a total of 36 models estimated for each country. When q=0, it turns into an AR(p) process while p=0, it turns into an MA(q) process. To determine the optimal model, we use the information criteria and choose the model that minimizes them.

Once the optimal model has been chosen, we can use it to obtain the prediction values of stock returns in each country:

$$f_t^{(n)} = \hat{\eta}^{(n)} + \sum_{i=1}^{p} \hat{\gamma}_i^{(n)} r_{t-i}^{(n)} + \sum_{j=1}^{q} \hat{\theta}_i^{(n)} \varepsilon_{t-i}^{(n)} \tag{3.2}$$

3.4.2 ARMAX

The second model we use is the Auto-regressive Moving Average with Exogenous Terms (ARMAX, hereafter) model, which is based on the traditional ARMA model and then adding exogenous input terms into it. Compared with traditional ARMA, the ARMAX considers the exogenous factors into the model which is more persistent. In other words, the ARMAX (p, q, b) model contains the AR (p) and MA (q) models as the same form of ARMA model and a linear combination of b exogenous terms interpreted as d_t. The general express of ARMAX (p, q, b) should be written as the following equation:

$$r_t^{(n)} = \sum_{i=1}^{p} \gamma_i^{(n)} r_{t-i}^{(n)} + \sum_{j=1}^{q} \theta_i^{(n)} \varepsilon_{t-i}^{(n)} + \sum_{k=1}^{b} \eta_k^{(n)} d_{t-k}^{(n)} + \varepsilon_t^{(n)} \tag{3.3}$$

In this equation, the left hand side is still the stock market returns which is the same with the ARMA process. The difference of being the exogenous terms, is that d_t have been added to the right hand side. Similarly with ARMA model, we still use γ_i and θ_i as the parameters for autoregressive terms and moving average terms respectively. Furthermore,

η_k represents the parameter for the exogenous inputs term, d_t.

After determining the optimal model via information criteria, we can get the forecasting model of ARMAX as the following expression:

$$f_t^{(n)} = \sum_{i=1}^{p} \hat{\gamma}_i{}^{(n)} r_{t-i}^{(n)} + \sum_{j=1}^{q} \hat{\theta}_i{}^{(n)} \varepsilon_{t-i}^{(n)} + \sum_{k=1}^{b} \hat{\eta}_k{}^{(n)} d_{t-k}^{(n)} \tag{3.4}$$

3.4.3 VAR

Both ARMA and ARMAX models are self-forecast models, where the predicting value is mainly based on the serial correlation of historical data. The vector autoregression (VAR) model is more flexible and widely used. Brown and Cliff (2004) has used VAR model to consider investor sentiment and stock returns as a whole system. Following on this method, we consider stock market returns with all of the estimators as a whole system and organize them into the system Y_t, so that we can analyze the variables together. The system of Y_t composed by endogenous variables $Y_{1,t}$, $Y_{2,t}$, . . . , $Y_{n,t}$, where n represents for different countries where k starts from 1 to 9. Thus, the VAR model with lag length p can be formulated as the following equation:

$$Y_t = \mu + \sum_{i=1}^{p} \delta_i Y_{t-i} + \epsilon_t \tag{3.5}$$

where $Y = [R, \Delta Sent, \Delta CPI, \Delta IR, \Delta IP]$ for the monthly VAR. μ is a k×1 vector of constants which combining all estimators as a whole system, δ is the parameter for a time-invariant k×k matrix and ϵ_t is a k×1 vector of error terms.

We estimate each VAR by maximum likelihood (MLE). Then, we use the in-sample estimation results based on above model to iterate forward so that we can obtain the out-of-sample predictions as:

$$f_{t+1} = \hat{\mu} + \sum_{i=1}^{p} \hat{\delta}_i Y_{t-i} \tag{3.6}$$

Compared with other forecast models, the advantage of VAR models is that they provide predictions not only based on historical fitting of individual process but also by means of lags of other variables in the system.

3.4.4 BVAR

Although it is common to use VAR model to obtain forecasts, it has also been argued that VAR estimated by Bayesian methods (BVAR) would provide better forecast with more parsimonious models for the reason that standard VARs often incur in over-fitting problems (Carriero et al., 2009). Compared with standard estimation, the BVAR model treats model's parameters as random variables, and applies Bayesian estimation imposing restrictions on the dynamics of the parameters according to a specific type of prior. Based on this assumption, the coefficients on longer lagged variables are more likely to be near zeros, resulting in a more parsimonious estimation. To be specific, we start with the model showed in Equation (4), then adding posterior parameters. In the VAR system, there are K equations, and each one can be expressed as the following:

$$Y_{i,t} = \mu + \sum_{i=1}^{p} sum_{j=1}^{Q} \delta_{i,j} Y_{j,t-i} + \epsilon_t \tag{3.7}$$

In this case, the prior about coefficients are captured in the prior density function $\bar{\Omega} = \frac{p(Y_{i,t}|\delta_{i,j})p(Y_{i,t})}{p(Y_{i,t})}$ with $p(Y_{i,t}$ is the prior density function using to capture the prior about coefficient.

Hence, the predictions of $Y_{i,t}$ can be written as:

$$f_{i,t+1} = \mu + \sum_{i=0}^{p-1} sum_{j=1}^{Q} \hat{\delta_{i,j}} Y_{j,t-i} + \tag{3.8}$$

where we will choose the best model via information criteria.

3.4.5 Model Selection

For each model discussed above, we will consider different parameter settings and lag lengths, and assume that the best forecasting model exists among those considered here. After forecasting stock returns with the models above, we need to use information criteria to pick the model that has the best forecasting power. The Information criteria we will use are: Akaike Information Criterion (AIC), Schwarz's Bayesian Information Criterion (BIC) or Hannan-quinn Criterion (HQIC). These three types of information criteria are also widely adopted in many past papers (Wagenmakers and Farrell, 2004; Acquah, 2010; Dziak et al., 2020).

To be specific, the expressions for these three information criteria are as follows:

$$AIC = ln(\hat{\sigma}^2) + \frac{2K}{T} \tag{3.9}$$

$$BIC = ln(\hat{\sigma}^2) + \frac{K}{T} lnT \tag{3.10}$$

$$HQIC = ln(\hat{\sigma}^2) + \frac{2K}{T} lnlnT \tag{3.11}$$

where $\hat{\sigma}^2$ is the in-sample fitted error variance and it can be presented by the equation: $\hat{\sigma}^2(m) = \frac{\sum_{i=1}^{n} \hat{\varepsilon}^2(m)}{n}$). K is the number of parameters and T is the number of total in-sample observations. As we can notice in these expressions, the information criteria has two parts: the residual variance part and the penalty part. In practice, we often use the log of the likelihood function value rather than the residual sum of squares divided by the number of in-sample observations, T. In this case, the expressions of the information criteria would be given as:

$$AIC = \frac{-2l}{T} + \frac{2K}{T} \tag{3.12}$$

$$BIC = \frac{-2l}{T} + \frac{K}{T} lnT \tag{3.13}$$

$$HQIC = \frac{-2l}{T} + \frac{2K}{T} lnlnT \tag{3.14}$$

where l represents the averaged log likelihood function value adjusted by the penalty function, K and T are still the same with the former expressions.

After calculating the information criteria in different models, we select the model with the smallest information criteria as the optimal model with the most proper lag length. However, the models selected by different information criteria cannot always be the same. Dziak et al. (2020) discuss the model selection based on AIC and BIC, and compare these two information criteria carefully. They present that sometimes BIC performs better than AIC for the reason that BIC is more consistent. AIC has a probability to choose unnecessarily model when the number of observations is huge. These unnecessarily models will include more parameters and lag length than it should include. Meanwhile, they also admit that if we only rely on the BIC result, thus may cause an under-fitting of the model especially when the number of observations is modest.A great amount of literature also makes comparison among different information criteria but how to choose the optimal forecasting model is still a controversial issue. What we know is that BIC is more consistent compared with AIC, but AIC is more efficient than BIC. And HQIC result is somewhat between AIC and BIC. As a consequence, we can not simply decide which information criteria is better in all cases. It also depends on the number of parameters and observations, as well as the characteristics of the model.

3.4.6 Quality of forecasting models

After fitting the model in-sample it is important to evaluate the quality of the forecasts. According to past literature, there are several measures can be adopted to determine the accuracy of the forecasting value. Here, we use both of in-sample and out-of-sample forecasts.To be specific, we split the observations into in-sample data (1985-2010) and out-of-sample data (2011-2015). Firstly, we use in-sample data to estimate the forecasting model. Then we apply the model we build to generate the forecasting values out-of-sample. As we have the actual data for out-of-sample, we make comparison between the predictions and the actual value to figure out the power of the predicting models. Normally, the performance of forecasts can be measured by several statistical metrics, which are the root mean squared errors (RMSE), the mean absolute errors (MAE) and the mean absolute percentage errors (MAPE) respectively. The expressions for these measurements are:

$$RMSE = \sqrt{\sum_{t=S+1}^{S+m} (\hat{r_t} - r_t)^2/m}, m = 1, 2, \ldots, T - S \tag{3.15}$$

$$MAE = \sum_{t=S+1}^{S+m} \left| \hat{r_t} - r_t \right| /m, m = 1, 2, \ldots, T - S \tag{3.16}$$

$$MAPE = 100 \sum_{t=S+1}^{S+m} \left| \frac{\hat{r_t} - r_t}{r_t} \right| /m, m = 1, 2, \ldots, T - S \tag{3.17}$$

where S represents the in-sample time periods and m is the out-of-sample periods. Thus, $S + m$ is the total sample size, T. Also, r_t denotes the actual value in period t where $\hat{r_t}$ denotes the forecasting value in the same period with r_t. The model with the smaller values of these indicators predicts more accurate compared with others. If we take them individually, it is hard to describe the power of the model. However, if we compare these values between different countries, it shed good light on the performance of different predicting models.

3.5 Empirical Result

So far we have introduced two univariate models (ARMA and ARMAX) and two multivariate models (VAR and BVAR). In this section, we provide the empirical analysis of each model and the forecasting results of them. As mentioned in former part, we divided the full sample size into two parts. The in-sample part use monthly data from May 1985 until December 2010. The out-of-sample period is between January 2011 and December 2015. To start with, we use information criteria to select the optimal lag length and the best predicting performance in each type of the model.Then, we adopt both dynamic forecasting and static forecasting of out-of-sample size to make comparison among different models to determine which model can better predict the stock returns.

3.5.1 Results of ARMA model

The first step for estimating ARMA model is to identify the optimal ARMA parameterization from different combination of p and q. One of the typical methods is to use the auto-correlation function (ACF) and partial auto-correlation (PACF) (Rounaghi and Zadeh, 2016; Zhang et al., 2008). However, it is quite hard to interpret the ACF and PACF in practice for the reason that only very little real data can follow the pattern as the ideal plots of ACF and PACF. In most of the cases, using ACF and PACF would make very difficult to determine the model especially across different countries. Thus we use information criteria to figure out the optimal forecasting models. As Table 3.3 presents the optimal in-sample forecasting ARMA(p,q) model selected by the information criteria in ten European countries. To be specific, the model is estimated by the Ordinary Least Squares

Table 3.3: ARMA Information Criteria

Country	lag length	AIC	SC	HQ
OE	(2,3)	-2.454	-2.369	-2.420
BG	(2,3)	-2.690	-2.605	-2.656
FN	(1,1)	-1.835	-1.787	-1.816
BD	(1,1)	-2.485	-2.436	-2.466
GR	(1,1)	-1.554	-1.509	-1.536
IT	(1,1)	-2.474	-2.432	-2.457
NL	(1,1)	-2.751	-2.702	-2.731
PT	(3,1)	-2.704	-2.625	-2.672
ES	(1,1)	-2.414	-2.366	-2.395
UK	(2,3)	-3.045	-2.996	-3.025

(OLS) and chosen from a combination of $ARMA(p, q)$ where p and q start from 0 until 5, with a total of thirty six models with different combination of p and q in each country. The special case is that when $p = 0$, the process would be a $MA(q)$ process. Similarly, when $q = 0$, the process would be a $AR(p)$ process.

After determining the optimal combination of p and q, we estimate the model with in-sample data for all ten European countries. As shown in Appendix A, the sample models for most of the countries perform well in-sample, which strongly supports that stock market returns are predictable by its own past. To be specific, in Austria, Germany, Greece, Italy and UK, the stock returns in the past periods are positively correlated with the future stock returns. In other words, the higher stock returns in the current period would encourage investors to buy or sell more stocks in the future. Thus, the stock returns would be higher due to the optimistic investor behaviour and the increasing capital in stock market. However, in Belgium, Finland, Netherlands, Portugal, past stock returns have a negative effect on the future stock returns. As can be noticed, capitals in stock market of these countries are not as abundant as the countries with positive relationship between past and future returns. This means that most of the investors in these countries are not optimistic enough to the future stock markets even though they achieved high returns in current period, which is consistent with the noise trader theory (De Long et al., 1990; Brown and Cliff, 2005). The possible reason for this phenomena could be related to government policy and different cultural characteristics. Also we can find that there is no significant relationship between previous stock return and future returns in Spain, which continuously supports that individual country has its own characteristics (Schmeling, 2009; Baker et al., 2012).

Table 3.4 presents the ARMA out-of-sample forecasting result in these countries. We construct both dynamic and static methods to predict stock returns. Dynamic forecasting is a method of multi-steps prediction where the static method is only a sequence of one step ahead prediction. As a result, dynamic forecasts can rapidly converge on the expected mean value in long-run periods. From the table, we can find that the result of dynamic forecasts performs better than the results of static forecasts in most of the countries. Specifically, the predictive performance of stock returns performs better in the UK compared with other countries as it has the smallest error statistics. This finding is literally convincing as the UK stock market is widely accepted as a well-established and comprehensive market. While the out-of-sample forecasts for Greece is barely acceptable. The possible reason is the Greek government-debt crisis happened in 2009 can also strike the stock markets heavily.

Table 3.4: ARMA Out-of-sample Forecasting Result in Ten European Countries

Country	(p,q)	Dynamic RMSE	Dynamic MAE	Dynamic MAPE	Static RMSE	Static MAE	Static MAPE
OE	(2,3)	0.066	0.049	1.633	0.070	0.056	6.619
BG	(2.3)	0.050	0.037	1.596	0.053	0.040	5.167
FN	(1,1)	0.062	0.046	1.165	0.065	0.049	1.416
BD	(1,1)	0.061	0.045	1.310	0.061	0.046	1.310
GR	(1,1)	0.143	0.113	1.248	0.141	0.111	2.604
IT	(1,1)	0.065	0.052	1.082	0.066	0.053	1.181
NL	(1,1)	0.049	0.036	0.977	0.049	0.036	1.003
PT	(3,1)	0.064	0.049	1.068	0.070	0.055	1.722
ES	(1,1)	0.064	0.052	1.663	0.065	0.053	2.294
UK	(2,3)	0.039	0.030	1.509	0.040	0.030	1.768

3.5.2 Result of ARMAX model

The ARMA model as one of the most common used univariate models, is widely used in the research of stock returns. In recent years, many researchers have pointed out that introducing exogenous variables into univariate model can effectively increase the accuracy of forecasting (Ding and Chen, 2005; Zheng and Zhu, 2017). Others also pointed out that stock return is not only affected by its own past, but also affected by many other factors, including macroeconomic factors and investor feelings(De Long et al., 1990; Brown and Cliff, 2004; Schmeling, 2009). Since the investor feeling is a completely subjective and unobserved variable, it is essential to find proper proxy for it. We have explained why we chose the consumer confidence index as the proxy for investor feeling. For macroeconomic factors, we set consumer price inflation (CPI), three months treasury bill interest rate (R) and the industrial production function index (IP) as exogenous variables respectively.

Similar with the ARMA model, we also adopt information criteria to determine the optimal model as shown in the table 3.5. Compared with table 3.3, the lag length of ARMAX model increases in Greece and Italy. Also, ARMA (2,3) model changes to ARMAX(2,0) by

Table 3.5: ARMAX Information Criteria

Country	lag length	AIC	SC	HQ
OE	(2,3)	-2.464	-2.271	-2.386
BG	(2,3)	-2.751	-2.618	-2.698
FN	(1,1)	-1.588	-1.447	-1.531
BD	(1,1)	-2.507	-2.410	-2.468
GR	(2,2)	-1.870	-1.717	-1.808
IT	(2,3)	-2.615	-2.454	-2.550
NL	(1,1)	-2.765	-2.668	-2.726
PT	(3,1)	-2.749	-2.589	-2.684
ES	(1,1)	-2.514	-2.414	-2.474
UK	(2,0)	-3.105	-3.008	-3.066

adding exogenous variables in the UK. These changes prove that models of stock returns also depend on the characteristics of individual countries. The in-sample ARMAX model results are shown in Appendix B. Based on the result, we can find investor sentiment has significantly positive relationship with stock returns except for Austria and Netherlands. Table 3.6 also presents that dynamic forecasting process performs better than the simple static forecasting, which is consistent with the standard ARMA model. Except for Portugal and the UK, all of the other countries have lower values of RMSE, MAE and MAPE comparing with the ARMA model both in dynamic and static process. It is worth mentioning that the values of statistics drop a lot in the Greece stock market, which further proved our explanation to it in former part. Such comparisons prove that ARMAX performs better compared with traditional ARMA. In other words, adding exogenous variables into ARMA model can increase the accuracy of the predictions.

Table 3.6: ARMAX Out-of-sample Forecasting Results in Ten European Countries

Country	(p,q)	Dynamic RMSE	Dynamic MAE	Dynamic MAPE	Static RMSE	Static MAE	Static MAPE
OE	(2,3)	0.064	0.050	1.569	0.068	0.053	1.477
BG	(2,3)	0.049	0.039	1.439	0.049	0.039	1.460
FN	(1,1)	0.058	0.045	1.349	0.059	0.047	1.396
BD	(1,1)	0.056	0.043	1.546	0.056	0.043	1.561
GR	(2,2)	0.203	0.133	1.191	0.203	0.135	1.199
IT	(2,3)	0.059	0.045	1.243	0.059	0.045	1.212
NL	(1,1)	0.048	0.037	1.490	0.048	0.037	1.487
PT	(3,1)	0.068	0.052	1.492	0.077	0.060	1.510
ES	(1,1)	0.063	0.049	1.328	0.062	0.047	1.282
UK	(2,0)	0.041	0.033	1.586	0.041	0.032	1.510

3.5.3 VAR model

Comparing with self-forecasting models we discussed above, the VAR model can combine all of the variables into a whole system. To be specific, the forecasting values of the self-forecasting models, such as the standard ARMA, are mainly explained by the serial correlation of their own historical data. Although the ARMAX models can merge exogenous variables into them, the potential endogeneity issue may also affect the forecasting abilities. One of the advanced advantages for VAR model is that all of the variables are considered as endogenous and can be predicted together. In VAR models, the predicting values can be determined more than its own lags or the combinations of error terms, but also by the entire model system.

We have proved that all of the variables in the VAR models are stationary in differences. Also, we make Granger-Causality process to figure out to dependencies among variables. To be specific, the Granger-Causality test is used to check whether one variable in the past period can affect another variable in the current period significantly.

As can be inferred from table 3.7, investor sentiment Granger causes stock returns in most of the European countries (except for Finland and Germany), which means that the lagged sentiment values have significant effect on stock returns. It is easy to understand that the mood of investors would affect the stock price in certain ways.A bullish investor always expects the stock returns to be higher than the average level. As a result, if most of the investors prefer to buy more stocks at a certain period, the expected stock returns in next period would be high. On the contrary, a bearish investor often considers more about the high risks in stock market, which will affect their behavior of buying or selling stocks in the market. In that case, the expected stock returns would be lower in the next period compared with current period. Also, we can observe that in most of the countries, investor sentiment and macroeconomic variables Granger causes stock returns in general.

Moreover, in some other countries, such as Finland and Germany, we can find that the previous stock returns and macroeconomic variables can significantly influence current investor sentiment. Qiu and Welch (2004) and Schmeling (2009) have explained that investor sentiment could have relationship with returns or some macroeconomic variables. It is due to the fact that investors are not always rational when they make decisions. Especially meeting with a series of breaking news (either good or bad), changes in the stock market condition and macroeconomic development policies, investors would be exceedingly optimistic or pessimistic in a certain period.

The Granger-Causality test results further proved that in some conditions, previous sentiment and macroeconomic variables can affect current stock returns. And sometimes investor sentiment would also be driven by previous stock returns and macroeconomic variables. Therefore, it is more than reasonable that we organize these variables into a system to predict stock returns.

Table 3.7: Granger-Causality Test Result in 10 European Countries

OE:Dependent variable:RETURN				Dependent variable:CCI			
Excluded	Chi-sq	df	Prob.	Excluded	Chi-sq	df	Prob.
OECCI	5.344	1	0.021**	OERETURN	1.530	1	0.216
OECPI	1.233	1	0.267	OECPI	3.554	1	0.059**
OER	1.344	1	0.246	OER	0.003	1	0.954
OEIPI	4.806	1	0.028**	OEIPI	0.349	1	0.555
All	15.179	4	0.004***	All	4.878	4	0.300
BG:Dependent variable:RETURN				Dependent variable:CCI			
Excluded	Chi-sq	df	Prob.	Excluded	Chi-sq	df	Prob.
BGCCI	3.901	1	0.048**	BGRETURN	0.516	1	0.473
BGCPI	2.440	1	0.118	BGCPI	1.703	1	0.192
BGR	15.445	1	0.000***	BGR	0.000	1	0.995
BGIPI	1.042	1	0.307	BGIPI	1.479	1	0.224
All	23.520	4	0.000***	All	3.508	4	0.477
FN:Dependent variable:RETURN				Dependent variable:CCI			
Excluded	Chi-sq	df	Prob.	Excluded	Chi-sq	df	Prob.
FNCCI	0.644	1	0.422	FNRETURN	6.148	1	0.013**

FNCPI	0.213	1	0.644	FNCPI	1.860	1	0.173
FNIPI	0.402	1	0.526	FNIPI	0.085	1	0.771
FNR	0.441	1	0.507	FNR	8.569	1	0.003***
All	1.495	4	0.828	All	16.325	4	0.003***
BD:Dependent variable:RETURN				Dependent variable:CCI			
Excluded	Chi-sq	df	Prob.	Excluded	Chi-sq	df	Prob.
BDCCI	0.861	1	0.353	BDRETURN	0.608	1	0.436
BDCPI	0.559	1	0.455	BDCPI	2.154	1	0.142
BDR	0.212	1	0.645	BDR	0.115	1	0.734
BDIPI	1.975	1	0.160	BDIPI	8.103	1	0.004***
All	3.951	4	0.413	All	11.228	4	0.024**
GR:Dependent variable:RETURN				Dependent variable:CCI			
Excluded	Chi-sq	df	Prob.	Excluded	Chi-sq	df	Prob.
GRCCI	4.818	6	0.056*	GRRETURN	5.445	6	0.488
GRCPI	2.185	6	0.029**	GRCPI	5.010	6	0.543
GRR	6.432	6	0.377	GRR	6.342	6	0.386
GRIPI	4.051	6	0.670	GRIPI	6.908	6	0.329
All	23.431	24	0.049**	All	28.695	24	0.232
IT:Dependent variable:RETURN				Dependent variable:CCI			
Excluded	Chi-sq	df	Prob.	Excluded	Chi-sq	df	Prob.
ITCCI	0.374	2	0.830	ITRETURN	0.698	2	0.705
ITCPI	0.148	2	0.929	ITCPI	0.284	2	0.868
ITR	5.540	2	0.063	ITR	7.224	2	0.027**
ITIPI	8.162	2	0.017**	ITIPI	3.683	2	0.159

All	13.886	8	0.085*	All	10.745	8	0.217
NL:Dependent variable:RETURN				Dependent variable:CCI			
Excluded	Chi-sq	df	Prob.	Excluded	Chi-sq	df	Prob.
NLCCI	41.437	3	0.000***	NLRETURN	2.674	3	0.445
NLCPI	3.229	3	0.358	NLCPI	1.682	3	0.641
NLR	8.678	3	0.034**	NLR	3.129	3	0.372
NLIPI	0.989	3	0.804	NLIPI	2.329	3	0.507
All	48.989	12	0.000***	All	10.160	12	0.602
PT:Dependent variable:RETURN				Dependent variable:CCI			
Excluded	Chi-sq	df	Prob.	Excluded	Chi-sq	df	Prob.
PTCCI	1.086	2	0.581	PTRETURN	0.457	2	0.796
PTCPI	3.266	2	0.195	PTCPI	0.152	2	0.927
PTR	2.174	2	0.337	PTR	3.101	2	0.212
PTIPI	1.530	2	0.465	PTIPI	3.323	2	0.190
All	9.034	8	0.33	All	6.831	8	0.555
ES:Dependent variable:RETURN				Dependent variable:CCI			
Excluded	Chi-sq	df	Prob.	Excluded	Chi-sq	df	Prob.
ESCCI	6.687	6	0.351	ESRETURN	4.297	6	0.637
ESCPI	9.600	6	0.143	ESCPI	17.042	6	0.009***
ESR	12.976	6	0.043**	ESR	6.241	6	0.397
ESIPI	21.254	6	0.002***	ESIPI	3.899	6	0.690
All	47.520	24	0.003***	All	29.622	24	0.198
UK:Dependent variable:RETURN				Dependent variable:CCI			

Excluded	Chi-sq	df	Prob.	Excluded	Chi-sq	df	Prob.
UKCCI	0.867	1	0.035**	UKRETURN	0.263	1	0.608
UKCPI	0.019	1	0.891	UKCPI	0.640	1	0.424
UKR	9.616	1	0.002***	UKR	3.610	1	0.057
UKIPI	2.215	1	0.137	UKIPI	1.697	1	0.193
All	12.604	4	0.013**	All	6.026	4	0.197

Asterisks refer to the level of significance where: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

In modelling and forecasting VAR models, it is important to use stationary variables and determine the lag length of the model. Similar with ARMA and ARMAX model, we also use information criteria to determine lag length of VAR model. To be specific, we select the lag length from 0 to 8 in each country and table 3.8 represents the optimal lag length selected by the smallest information criteria.

After determining the optimal lag length, the VAR in-sample regression result is shown in Appendix C. We can observe that the previous stock returns have a significant positive effect on future returns in Austria, Belgium, Finland, and Spain when we consider all of these estimates into a whole system. We also prove that there is a positive relationship between the past investor sentiment and the future stock returns in Austria, Belgium, Netherlands and Spain. Besides, past returns can also positively affect investor decision in the future periods in Finland and Portugal. However, this results cannot explain for how long the effects would work through the VAR system. Hence, we need to introduce the impulse

Table 3.8: VAR Information Criteria

Country	lag	AIC	SC	HQ
OE	1	4.908983*	5.451518	5.129051
BG	1	5.411542*	5.782824	5.560147*
FN	1	6.774242*	7.379063	7.019971
BD	1	5.068960*	5.440242	5.217564*
GR	1	11.57794*	14.63461	12.81955
IT	2	4.821921*	5.646614	5.154622
NL	3	6.090337*	7.080422	6.486615*
PT	2	7.248404*	8.152293	7.614157
ES	6	6.908046*	8.88944	7.702248
UK	1	5.305122*	5.676404	5.453727*

responses in VAR models to help us analyze the relationship among variables inside the system.

The impulse responses explore the responsiveness of the dependant variables in the VAR model when a unit shock appears to the error in each separate model. To make it clear, we can observe the impact of unit shock on the current values, the tracing of the impact and when will the impact disappear in the future period. The pictures in appendix D display the combination responses of each variable to other variables in the system. Firstly, we can find that the linkages between different series are relatively weak due to the fact that in each pictures, the responses of the variable given shocks to itself are sharply down, where responses of other variables to the shock are quite small. This finding support the Granger causality test results.Also, we can notice that the responses of shocks will disappear after

Table 3.9: VAR Out-of-sample Forecasting results in European Countries

Country	Variable	Dynamic RMSE	Dynamic MAE	Dynamic MAPE	Static RMSE	Static MAE	Static MAPE
OE	OERETURN	0.201	0.116	1.113	0.087	0.067	1.096
	OECCI	2.497	1.908	1.441	2.488	0.190	2.215
	OECPI	0.486	0.360	1.403	0.488	0.040	3.454
	OER	0.179	0.141	0.804	0.177	0.014	6.513
	OEIPI	1.224	0.947	3.682	1.225	0.094	1.345
BG	BGRETURN	0.095	0.073	0.922	0.045	0.034	1.318
	BGCCI	3.067	2.408	1.766	3.068	2.408	2.032
	BGCPI	0.251	0.184	1.087	0.873	0.699	2.211
	BGR	0.203	0.155	0.752	0.194	0.137	1.757
	BGIPI	2.101	1.753	4.793	2.102	1.754	0.744
FN	FNRETURN	0.070	0.050	1.171	0.045	0.035	1.182
	FNCCI	2.620	2.114	1.729	2.618	2.118	2.619
	FNCPI	0.321	0.249	1.171	0.662	0.512	0.575
	FNR	0.163	0.126	0.651	0.161	0.127	0.947
	FNIPI	0.974	0.803	1.173	0.973	0.802	1.657
BD	BDRETURN	0.014	0.081	1.076	0.067	0.054	1.065
	BDCCI	1.930	1.350	1.098	1.928	1.347	1.303
	BDCPI	0.343	0.259	2.009	0.509	0.400	2.792
	BDR	0.160	0.124	0.987	0.156	0.121	1.689
	BDIPI	1.413	1.103	2.379	1.416	1.105	1.779
GR	GRRETURN	0.031	0.028	0.898	0.449	0.281	1.112
	GRCCI	5.042	3.754	1.615	5.054	3.782	5.183
	GRCPI	1.384	1.104	0.420	2.437	1.890	4.879
	GRR	2.121	1.337	0.785	2.115	1.287	1.349
	GRIPI	3.111	2.447	0.845	3.108	2.447	1.407
IT	ITRETURN	0.085	0.070	1.933	0.082	0.062	1.117
	ITCCI	3.267	2.445	1.353	3.289	2.458	5.630
	ITCPI	0.251	0.206	1.590	0.516	0.403	1.456
	ITR	0.317	0.235	0.685	0.332	0.258	0.997
	ITIPI	1.180	0.940	3.881	1.184	0.944	2.949
NL	NLRETURN	0.144	0.068	0.620	0.681	0.553	1.216
	NLCCI	4.233	3.419	4.221	4.257	3.462	2.261
	NLCPI	0.556	0.452	3.945	1.538	1.237	2.313
	NLR	0.162	0.125	1.244	0.152	0.124	4.487
	NLIPI	1.927	1.535	4.014	1.932	1.541	1.376
PT	PTRETURN	0.305	0.106	1.608	1.266	0.966	0.347
	PTCCI	2.695	2.160	1.198	2.757	2.196	1.835
	PTCPI	0.884	0.577	2.185	2.036	1.543	2.374
	PTR	0.618	0.490	3.289	0.594	0.464	5.941
	PTIPI	2.068	1.635	6.445	2.067	1.631	1.408
ES	ESRETURN	0.300	0.147	2.146	1.761	1.441	0.982
	ESCCI	3.713	3.014	6.391	3.705	2.941	2.415
	ESCPI	0.728	0.552	1.008	1.175	0.924	2.925
	ESR	0.296	0.229	4.564	0.316	0.235	3.613
	ESIPI	1.071	0.856	2.399	1.069	0.854	2.468
UK	UKRETURN	0.050	0.034	2.233	0.482	0.368	1.051
	UKCCI	2.817	2.138	1.245	2.806	2.131	1.407
	UKCPI	0.393	0.311	0.773	0.413	0.317	0.614
	UKR	0.164	0.130	0.571	0.153	0.118	0.494
	UKIPI	0.798	0.506	1.422	0.798	0.506	3.889

Table 3.10: BVAR Out-of-sample Forecasting Results in European Countries

Country	Variable	Dynamic RMSE	Dynamic MAE	Dynamic MAPE	Static RMSE	Static MAE	Static MAPE
OE	OERETURN	0.119	0.092	1.120	0.932	0.681	1.141
	OECCI	2.498	1.906	1.399	2.507	1.932	2.638
	OECPI	0.473	0.350	1.346	0.648	0.490	0.935
	OER	0.180	0.142	1.204	0.180	0.138	1.308
	OEIPI	1.224	0.947	3.785	1.226	0.948	1.020
BG	BGRETURN	0.115	0.078	0.905	0.654	0.518	1.165
	BGCCI	3.067	2.408	1.749	3.069	2.409	4.658
	BGCPI	0.224	0.170	1.041	0.286	0.223	2.124
	BGR	0.203	0.156	0.933	0.198	0.138	1.235
	BGIPI	2.101	1.753	5.266	2.100	1.754	2.311
FN	FNRETURN	0.308	0.096	2.946	1.082	0.858	1.015
	FNCCI	2.618	2.112	1.684	2.619	2.122	3.886
	FNCPI	0.319	0.254	1.435	0.506	0.425	1.466
	FNR	0.163	0.126	0.763	0.163	0.127	3.940
	FNIPI	0.974	0.803	1.097	0.974	0.805	1.318
BD	BDRETURN	0.097	0.074	1.059	0.745	0.539	1.209
	BDCCI	1.931	1.353	1.086	1.927	1.345	1.377
	BDCPI	0.354	0.262	1.950	0.583	0.433	2.043
	BDR	0.160	0.125	1.151	0.156	0.122	1.064
	BDIPI	1.413	1.103	2.452	1.417	1.105	2.347
GR	GRRETURN	0.558	0.399	0.97	3.530	2.408	0.999
	GRCCI	5.041	3.752	1.565	5.117	3.817	1.056
	GRCPI	1.385	1.106	0.488	2.295	1.745	2.229
	GRR	2.121	1.340	0.669	2.261	1.286	1.149
	GRIPI	3.112	2.447	0.857	3.113	2.446	0.963
IT	ITRETURN	0.073	0.059	1.884	0.786	0.545	1.086
	ITCCI	3.268	2.446	1.344	3.291	2.452	2.596
	ITCPI	0.249	0.201	1.761	0.508	0.396	0.767
	ITR	0.317	0.235	0.763	0.322	0.231	0.723
	ITIPI	1.180	0.940	4.078	1.180	0.941	1.967
NL	NLRETURN	0.258	0.080	2.428	0.938	0.737	1.029
	NLCCI	4.229	3.418	2.742	4.212	3.409	2.454
	NLCPI	0.524	0.441	0.685	0.632	0.536	0.822
	NLR	0.162	0.126	1.279	0.155	0.118	3.786
	NLIPI	1.927	1.535	3.746	1.925	1.534	2.504
PT	PTRETURN	0.218	0.085	1.884	1.279	0.993	1.093
	PTCCI	2.695	2.161	1.176	2.749	2.196	0.848
	PTCPI	0.700	0.488	0.823	0.915	0.719	0.701
	PTR	0.624	0.496	2.969	0.631	0.487	1.291
	PTIPI	2.067	1.634	0.742	2.068	1.633	2.183
ES	ESRETURN	0.145	0.090	1.537	0.809	0.631	1.045
	ESCCI	3.738	3.034	1.517	3.722	2.965	2.163
	ESCPI	0.648	0.484	1.428	0.760	0.564	0.671
	ESR	0.302	0.234	1.737	0.308	0.222	0.421
	ESIPI	1.071	0.855	1.623	1.069	0.855	1.257
UK	UKRETURN	0.040	0.030	1.098	0.497	0.369	1.279
	UKCCI	2.822	2.143	1.216	2.808	2.143	1.537
	UKCPI	0.395	0.313	0.828	0.397	0.319	2.185
	UKR	0.164	0.130	0.592	0.153	0.116	3.623
	UKIPI	0.798	0.506	1.159	0.799	0.507	4.652

3-4 periods in most of the countries except for Spain. In Spain, the responses of each variable diminishes in longer time horizon and it converges to zero but still fluctuates a bit in ten periods time. The reason for this peculiar case is that the model we choose for Spain includes 6 periods of lag length. Hence the responses would last longer compared with other countries.

table 3.9 shows the out-of-sample forecasting result of VAR model. Compared with ARMA and ARMAX, we can find that the values of RMSE, MAE and MAPE are bigger than ARMA and ARMAX model both in dynamic and static forecasting. As we know, the model with the smaller criteria performs better in forecasting. Thus ARMA and ARMAX models seem to have better forecasting performance compared with VAR.

3.5.4 Bayesian VAR model

BVAR model is the common VAR estimated by Bayesian methods. Compared with the standard VAR models, Bayesian VAR models take parameters as random variables and adopt prior probabilities to them. By providing the framework to reassign probability distributions for unobserved parameters, Bayesian VAR allows to absorb previous information about the parameters. The BVAR in-sample forecasting result is shown in Appendix E. As displayed in the table 3.10, the previous stock returns are positively correlated with the future stock returns in all of these ten European countries. Moreover, we can also observe that the previous value of investor sentiment has a positive effect on the future stock returns in Austria and Netherlands. Besides, the past values of macroeconomic variables would affect the forecasting value of returns among half of these European countries, which are Austria, Belgium, Netherlands, Spain and UK respectively.

Similar to the VAR model, we also display the impulse responses pictures in Appendix F. The effect of the unit shock fluctuates a lot in the first three periods, and then disappears in 4-6 periods, which is longer than the standard VAR. The reason for this longer affection is that the Bayesian VAR model reallocate the probability of distributions for unobserved parameters. As a result, when a shock added to a variable, all of the other variables would take more time periods to reach to the random distribution of the forecasting value.

table 3.10 displays the out-of-sample forecasting result of BVAR model. First, all of the dynamic forecasting estimators are smaller than the static estimators, which further proved that multi-steps forecasting result performs better compared with one-step ahead forecasting. Second, both of the dynamic RMSE and the Static RMSE decrease compared with the standard VAR model, which proves that BVAR provides more accurate prediction of stock returns than VAR.

3.6 Conclusion

In this chapter we produced both the univariate and multivariate models to predict stock returns. To be specific, we obtained the in-sample forecasts results from Auto-Regressive Moving-Average model, ARMA with exogenous terms, Vector Auto-regression model and the Bayesian VAR. Then we used AIC, BIC and HQIC information criteria to select the optimal model from each sets of models. To examine the quality of prediction, out-of-forecasting estimators (RMSE, MAE and MAPE) are generated to make comparison between different models.

Specifically, the ARMA and ARMAX models prove that stock returns are strongly predictable by their previous value in these countries. The past values of returns have always affect the future stock returns positively. As the out-of-sample forecasting results show, ARMAX performs better compared with ARMA. Since most of the previous studies with these models only focused on the US stock market, our results enriches the literature by forecasting several European stock markets. Moreover, as ARMAX model suggests that macroeconomic variables and the investor sentiment could be also factors that affecting the stock returns as well, we produce a standard VAR model to consider the stock returns, investor sentiment and macroeconomic variables into a whole system. Based on the result of VAR, we further obtain that stock returns are predictable by their own past. Also, investor sentiment in previous period has a positive effect on the stock returns in next period for the reason that emotional of investors is also an essential factor of the stock returns. Bullish investors expect the stock returns to be higher in future periods, which motivates them to invest or sell more in current period. As a result, the stock returns will increase as capitals are accumulated in stock market due to the confidence of investors. On the contrary, if investors are pessimistic about the future state in the stock market, they would decrease the investment in the future, which would diminish the liquid capitals in the market and cause returns to be lower in the futures. Besides, we also found that macroeconomic variables can affect stock return significantly in certain countries. The possible reason for this situation is that the development of stock market also relies on the government policies, the current economic condition and the specific characteristics of the country. Having used data for ten European stock markets, we notice that both there are common patterns for all of these ten countries, but the unique characteristics of each country also exists. As we consider the effect of macroeconomic variables on returns, we introduce the Bayesian VAR model to improve the performance of these variables by adding prior weight on the parameters. The results of BVAR suggested that BVAR has the smaller RMSE, MAE and MAPE than the standard VAR, which means BVAR performs better compared with the standard VAR when macroeconomic variables including in the model.

Generally speaking, the ARMA and ARMAX performs better than the VAR and BVAR models. Comparing these four models, we can find the advantage of the ARMA and ARMAX model is that they consider the moving average process into the model where the VAR and BVAR do not, which increases the accuracy of the predictability of stock returns. The disadvantage of these univariate models is that they can only reflect the power of prediction for the dependant variable on its own. And this is also the main advantage of VAR and BVAR models. They are more flexible and all of the variables can be taken as a system to forecast together. In that case, we can organise all variables as endogenous and it would be easy and informative to observe the relationship between them. However, univariate models could also be superior to multivariate models for the following reasons. First, since the VAR models include too many parameters, it may not suitable for small sample size as the degree of freedom would be rapidly used up, which may cause very large of standard errors. Also, the coefficient of VAR can only interpret the direction of the relationship but with no economic meanings. That is why we used the univariate and multivariate models at the same time. Given accurate predictions with low error such as those we obtained using various forecasting models, it is theoretically to observe stock market situations in advance. For regulators, forecasting stock returns accurately can shed light on abnormal financial activities and prevent the market from serious financial crisis. For individual investors, having a good overview of stock market can help them to make better choices in trading activities.

Chapter 4

Forecasting Stock Market Returns with Model Averaging Methods

4.1 Introduction

The stock market always plays an essential role in economics since it controls the largest amount of liquid capital. Stock price fluctuations can affect the future of financial stock market, thereby influencing behaviour in financial market and the development of the economy. For several decades, many researchers have been focusing on modelling and forecasting stock market returns to understand the inherent patterns of the stock market for the reason that both institutions and individuals wish to make rational decisions so that they can earn more in the stock market. Moreover, regulators can formulate policies to reduce the probability of unexpected financial crises happening in the future.

At the very beginning, the efficient market hypothesis (EMH) had the leading position within financial theories as it explained the market behaviour in great details. The main content of EMH is that market information is fully shared among all investors and stock prices will always reach a equilibrium since the trading activities follows the self-regulation patterns of stock markets (Basu, 1977; Malkiel, 1989). Later in 1990s, due to a series of financial crises, researchers realized that the EMH cannot explain stock activities completely. De Long et al. (1990) and Fisher and Statman (2000) pointed out the effect that noise traders may have on the stock markets. From then, a rich past literature developed attending to predict stock returns. Different types of methods have been adopted by researchers all over the world to model and forecast stock returns. Fama and French (1992) forecasted the stock returns in cross-sections with fundamental financial factors, which strongly proved that stock returns were predictable by means of other financial variables. After that, Pontiff and Schall (1998) and Fama and French (1998) also showed that stock market returns were correlated with financial factors. Moreover, they pointed out that stock returns were also correlated with their own past and could be predicted by macroeconomic factors. However, there was also some evidence on the lack of predictability of stock returns. Kothari and Shanken (1997) used financial variables collected by small firms to show that returns are not always predictable. Besides, Bossaerts and Hillion (1999) revealed that returns were not self-predictable out-of-sample forecasts, thereby providing a contradictory result compared with past papers.

In recent years, numerous of researchers began to examine the predictability of stock returns using different forecasting models. As we know, the most common univariate model is the Auto-Regressive Moving Average (ARMA) model, which has been widely used in numerous past papers (De Gooijer, 1989; Karanasos, 2001; Henry, 2002). There are also extensions of the basic ARMA model, for example, the ARIMA and ARMAX models, which are also widely adopted in previous studies (Hannan et al., 1980; Johansen and FOSS, 1993; Lim et al., 2009). However, as a lot of literature has pointed out that univariate model could only predict the stock markets using its own past, so later on, multivariate model s were also used to make prediction of stock returns. Specifically, Schmeling (2009) applied the vector autoregression (VAR) model to investigate the predictive power of investor sentiment in different types of stock markets internationally. He pointed out that the benefit of VAR models is that they can combine all variables into an entire system. He showed that stock returns were not only correlated with their own past, but also negatively affected by sentiment. Similarly, Carriero et al. (2009) built a large Bayesian VAR model to forecast exchange rates using the justification that exchange rates would co-move. They also discussed the advantages of multivariate model as compared to those of the simple random walk model. As we can see from above, there is still debate on which model to use when predicting stock returns. Actually, this is true for empirical analysis in general.

The issue is that it is unlikely to find a single model which performs better than any other models for the following reasons: Firstly, as any coin has two sides, due to the different properties among prediction methods, any single model has advantages and disadvantages. If we choose a regression model, usually it is because that we believe if we choose a particular model, it is believed that it has more benefits than disadvantages. Meanwhile, no one can deny that shortcomings of the model still exists. Secondly, most of the so-called optimal model is selected from a set of potential candidate models based on different and often subjective criteria. One should always believe that the optimal model among these candidate models really exists. But the truth is, we can only select the 'better' model compared with all of selected models but we can never find the 'best' one that would fit every problem since even selection statistics are affected by sample sizes and more broadly the characteristics of the data. Thirdly, the forecasting performance of certain models can potentially change in different periods of time. For example, the predictors of stock returns could be different in the past as compared to current ones. Furthermore, some variables may only be able to predict well in periods of expansion or recession. Since we need to find a common pattern among European countries, it would be necessary to standardize the variables to make better comparisons. For these reasons, a stream of literature has proposed an alternative way to combine different models for the purpose of forecasting rather than using an individual model: this is called model averaging. Model averaging method combines individual models by giving them different weights. The idea is that appropriate weighted average of the forecast obtained from individual models would improve the accuracy of prediction of the individual models used.

The main contribution of this paper is that we use model averaging to improve the forecasts of stock markets obtained by individual models. The analysis would be performed in two steps. The first step involves choosing the optimal model from each class to be combined in the averaging process. The second step is the choice of weight to attach to each nominated model. As in the previous studies, we are going to use both univariate models and multivariate models. To be specific, we are going to use ARMA and ARMAX models as univariate model while for multivariate model, we forecast VAR and BVAR models for ten European countries. As model averaging techniques, we are going to use the simple model averaging (SMA, hereafter), Bayesian model averaging (BMA, hereafter) and the smoothing Akaike model averaging (AMA, hereafter) and will make a comparison between these model averaging methods.

The rest of this chapter is organized as follows: Section two gives a brief literature review of individual model forecasting and model averaging methods. Section three describes the data and displays the primary test results. In section four, we introduce the models selected into model averaging and the model averaging methods we use. In section five, we analyze the forecasting results of individual models and averaged models, also making a comparison among the out-of-sample results. A summary concludes.

4.2 Theoretical Background

The stock market plays an indispensable role in attracting and allocating the distributed savings and liquidity into the most profitable financial activities so that the financial resources are adequately utilized in the national economy. In any stock market, both rational investors and speculators are required to determine the fundamental value of stocks and then compare it with the current market price so that they can trade in a better price. Thus, investigating stock market is very important for both researchers and individual investors alike.

4.2.1 Reviews of Stock Return Prediction

Behaviors of investors can determine the capital allocation in financial market. Thereby, to investigate the stock market, it is necessary to predict its returns by analyzing efficiently the available market information. Numerous studies have focused on predicting stock market returns over the past decades. Campbell (1987) showed that excess returns on bills, bonds and stocks are surprisingly predictable. Later, they continued their research and used the dividend price ratio as the a predictor of stock returns and forecast the stock market one period ahead (Campbell and Shiller, 1988). They also proved that the stock market was predictable using real dividend growth, measured real discount rates, and other unexplained factors, which yield the metric to judge the measured financial factors. Similarly, Fama and French (1988) pointed out that stock market was strongly influenced by dividends and explained the temporary effect of future stock returns on the stock market.

Later in 1990s, Fama and French (1992) detected the cross-sectional expected stock returns with more financial factors. To be specific, they used market β, size, leverage and book-to-market ratio to forecast stock returns and showed that the US stock returns are strongly predictable. Kothari and Shanken (1997) and Pontiff and Schall (1998) reported that book-to-market (BM, hereafter) ratio and dividend yield can predict stock market returns. However, Kothari and Shanken (1997) highlighted that the stock market was not always efficient, especially in certain periods. However, most of the literature in that period only found limited evidence of stock market return (US market) predictability in-sample. Furthermore, there are several papers proving that in-sample estimation often doesn produce good out-of-sample forecasts (Bossaerts and Hillion, 1999).

In recent years, more and more researchers have started using more efficiently both in-sample and out-of-sample analyses to predict future stock returns. To be specific, Goyal and Welch (2003) used a simple out-of-sample approach to examining the forecasting performance of the stock returns in the US market. However, the out-of-sample results they obtained reflected that dividend ratios for the American stock market were not predictable both in both short and long time horizons.

They comprehensively examined the performance of forecasting models for the stock returns with several macroeconomic factors and financial variables. However, the forecasting results showed that such predictions failed to beat the benchmark forecasts both in-sample and out-of-sample. In historical average predictive regressions, one of the main assumptions was that the coefficients of the predictors should be equal to zero, in other words, the information from financial variables were not useful for predicting stock returns. It is not surprising that the performance of this model was even worse than multiple forecasting regressions for the reason that the in-sample forecasting is often over-fitting the data (Welch and Goyal, 2007). Although their forecasting methods need to be improved, their findings have encouraged more efforts from other researchers. Campbell and Thompson (2007) proved that predictive regressions could beat the historical average return by generating an economically-motivated restricted model. According to their research, the weak restrictions are imposed on the signs of coefficients and the forecast returns must be non-negative. The restricted model improves the forecasting performance out-of-sample and can often beat the historical average benchmark. The more important point is that the out-of-sample predictions are meaningful for mean-variance investors as they impose the restrictions of steady state valuation models in short horizons to reduce the volatility of stock returns. To improve the predictive ability of single forecasting regression models, Rapach et al. (2010) combined numerous variables into a new forecasting regression model to reduce the forecast error variance. Their predictive model includes fifteen economic and financial variables, which help improving the accuracy of out-of-sample forecasting and beat the historical average model in a consistent way over different time periods.

On the other hand, several studies have begun to forecast the stock market in a more consistent way. Ang and Bekaert (2006) proved that stock returns were predictable not only for the US stock market, but also in France, Germany, Japan and the UK. They introduced a forecasting regression with three instruments, namely a short rate, dividend yield and earnings yield respectively and showed that only in short-run the out-of-sample forecasts are relatively accurate and stable. In long-horizon time periods, they proved that there was no evidence of predictability of stock returns. Generally speaking, the ability of predictors across all the countries were stronger compared with historical average predictors using local instruments (Ang and Bekaert, 2006). Other researchers have focused on certain industries and the effect that data frequency may have on forecasts' accuracy. Phan et al. (2015) focused on the out-of-sample predictability of stock returns in oil-related industries. Then using daily, monthly and quarterly data Phan et al. (2015) examined whether different data frequency would affect the power of predictability of stock returns. Their results proved that stock returns were always predictable under different data frequencies and the predictability depended mainly on the characteristics of the predictors.

In this chapter, we use in-sample forecasting regressions to study the predictability of stock returns. Then, among all the classes of models, we use the information criteria to determine the optimal model within each class. Then, we obtain and evaluate the out-of-sample forecasting for each of the optimal models we have selected.

4.2.2 Reviews of Model Averaging

Although many researchers have used a variety of forecasting models to predict stock returns, it is still controversial which class of models performs better as each modelling strategy has its own strengths and weaknesses. As mentioned above, even though one specific forecasting model produced more accurate predictions compared with others, it only means that this model performed better than others in some particular circumstance, in certain time period, for some specific data-set. Thus, unfortunately, it seems unrealistic to find a superior model that can systematically beat all of the other models. A way to find a "superior" forecasting model is to use combine the predictions obtained by several potential optimal candidate models by means of suitable averaging methods. To be specific, first we select the optimal model within each class, then obtain forecasts from each model and finally combine these forecasts into an average model by suitably weighting each of them. In model averaging, the model with higher weight would contribute to the forecasting performance more while the model with lower weight would have lesser role in forecasting the target variable. Thus, there are two tasks for us when using the model averaging, one is to find proper candidate model for each class and the other one is to obtain the weight for each of the candidate models.

The first study of model averaging approach can be dated back to 1960s. Bates and Granger (1969) used the airline passenger data to generate two forecasting models. Then they combined two separate sets of predictions into a new composite one and proved that the mean squared error of the new composite forecasts is lower compared with either of those produced by the two original models. Over time, model averaging has been widely applied in various research areas. Granger and Ramanathan (1984) forecasted the quarterly hog prices using both in-sample and out-of-sample individual predictions and proposed model averaging to improve the accuracy of forecasts. He also proved that model averaging is superior to individual forecasting models. There are now various methods to determine the weights of individual forecasts in model averaging. The most commonly used is the Bayesian model averaging. Raftery et al. (1997) described two procedures that can improve the individual model forecasting performance based on the Bayesian model averaging method. Similarly, Montgomery and Nyhan (2010) used the Bayesian model averaging to reduce the specification uncertainty of individual optimal models. Furthermore, they check the robustness of the model averaging result compared with individual model specifications and proved that model averaging could strongly strengthen the robustness of forecasting results in political science. Later, Koop and Korobilis (2012) forecasted quarterly US inflations using dynamic Bayesian model averaging. They suggested that dynamic model averaging could lead to substantial forecasting improvements over the benchmark regressions as well as over more sophisticated approaches, such as time varying coefficient models.

The Akaike model averaging is also a widely applied method to determine the weighting of the forecasts of the selected models. Posada and Buckley (2004) discussed the model selection and model averaging in genetics based on Akaike information criterion. They presented a general review of model selection according to the likelihood ratio test and the Akaike information criteria. They suggested that the Akaike information criteria showed great advantages compare with the hierarchical likelihood test in phylogenetics. They also compared the performance of Akaike model averaging and Bayesian model averaging in predicting the phylogeny. Cade (2015) proved that the Akaike model averaging can reduce the model specification uncertainty effectively. Further, he pointed out that the AIC weight could also be regarded as the assessment of selected optimal individual models. According to previous studies, we use the simple model averaging (SMA), Bayesian model averaging (BMA) and the Akaike model averaging (AMA) methods to determine the weights of candidate models in model averaging processes and compare them with individual forecasting models in European stock markets.

4.2.3 Predictors of Stock Market Returns

As we know, the stock return is not only correlated with its own past, but it would also be affected by many other factors. In previous studies, researchers used a variety of predictors to model stock returns. These can be divided into three categories: economic variables, financial factors and proxies for investor sentiment. As financial variables are unlikely to be standardized across all of the ten European countries, in previous chapters we have focused [1], on economic variables and proxies for investor sentiment to predict stock returns. In what follows, we briefly review the literature using these factors to determine which are more applicable to our research.

The Efficient Market Hypothesis (EMH) has been the dominant approach to model stock markets for several decades. As several financial crisis happened in 1980s, the failure of EMH to explain these abnormal phenomena, forced researchers to consider the cases that also irrational trading behaviour could affect the stock market and the information from the stock market was often asymmetric. In other words, it was highly unlikely that the market system would have self-adjustment mechanisms as mentioned in the EMH. Investor sentiment thus plays an important role in financial stock market. However, since investor sentiment is hard to observe and is largely dependent on the expectation of investors, it

[1] Refer to theoretical background part in chapter three

is crucial to find proper proxies for it. Based on prior works, measurements for investor sentiment can be divided into three types: direct, indirect and composite.

Most of the direct sentiment measures are taken as a series of surveys usually generated through various questionnaires both for institutions and individual investors. One of the most popular survey was created by the American Association of Individual Investors (AAII). It collected the answers from randomly selected participants each week since 1987. Another representative survey is Investor Intelligence (II). The limitation of these two surveys is that it can only be adopted in the US market. In this chapter we still focus on ten European countries. Hence, we continue using the Consumer Confidence Index (CCI,hereafter) as the proxy for investor sentiment. CCI is a reasonable candidate as the proxy for investor sentiment when we focus on stock markets in different countries for the mainly reason that it is usually standardized across countries compared with other direct measurements. (Charoenrook, 2005; Lemmon and Portniaguina, 2006; Schmeling, 2009).

.

Indirect sentiment measures, comprise some financial indicators often taken as the market signals, which can sometimes be more flexible and easily obtainable as compared with direct sentiment measures. Prior work suggested a numbers of proxies for investor sentiment to use in time-series models. The most common measurement that has been used is the closed-end funds discount (CEFD, hereafter), which refers to the average level of the difference between net asset value (NAV) of closed-end fund shares and their current market price (Zweig, 1973). CEFD is one of the most popular financial factors applied in past papers (De Long et al., 1990; Lee et al., 1991; Chen et al., 1993). However, Neal and Wheatley (1998), Qiu and Welch (2004) and Ben-Rephael et al. (2012) debated whether CEFD can be considered as an indirect measure of investor sentiment. The second widely used measurement is the stock trading volume, which is also known as the share turnover in practical research, refers to the number of stock shares traded in a certain market during a given period of time (Jones, 2002; Baker and Stein, 2004). Besides, there are numbers of other financial indicators which are taken as the reflection of investor sentiment, which are the initial public offering (IPO), the number of IPOs (NIPO), the equity share and the dividend premium (PD) (Ritter, 1984; Ljungqvist et al., 2006; Baker and Wurgler, 2000; Fama and French, 2001; Baker and Wurgler, 2004, 2007). Referring to the detailed discussion in chapter 3, we can see that each of these indirect measures may account for a specific component of investors sentiment. Indeed, the best thing to do, would be to combine them as a composite index. However, since our analysis focus on multiple stock markets, it is quite difficult to consistently obtain these indicators as different countries may have different way of measuring these financial indicators. Hence indirect and composite measurements are well suitable only for individual countries. In our research, we use the consumer confidence index as the unique proxy for investor sentiment since it seems that the CCI, being collected across the EU countries using the same questionnaire is the only appropriate candidate.

Apart from the measurements for investor introduced above, many researchers have proved that macroeconomic factors could also play an important role in forecasting stock returns. According to previously research, the most common used macroeconomic factors are the consumer price index (CPI), three-months treasury interest rates (R) and the index of industrial production (IPI) (Bossaerts and Hillion, 1999; Welch and Goyal, 2007; Westerlund and Narayan, 2014). To be specific, Bossaerts and Hillion (1999) used simple in-sample forecasting model to prove that economic indicators can predict stock returns successfully. Later, Welch and Goyal (2007) evaluated the predictability of the equity premium by various macroeconomic factors. Considering the effect of macroeconomic factors, we use the monthly change on CPI, monthly change in interest rate (R) and the monthly change of the IPI together with the proxy for investor sentiment to examine the predictability of stock returns in European stock markets.

4.3 Data Description and Unit Root Test Results

As previously noted, we are interested in modelling and forecasting stock market returns across countries. We use not seasonally adjusted, monthly stock return index as our dependent variable. Considering our requirement to international database, the stock returns are from the electronic version of Morgan Stanley's Capital International Perspectives (MSCI). The advantage of the MSCI data compared with other international database is that they can eliminate the survivor bias as the database includes not only currently traded firms, but disappearing firms too(Fama and French, 1998). The monthly returns, which can be collected from Datastream ranges from May 1985 through December 2015 wherever possible.

4.3.1 Data Discription

In implementing our analysis, we also use a proxy for investor sentiment. After comparing several sentiment measurement, it seems that CCI is the only suitable proxy for investors sentiment in different stock markets (Lemmon and Portniaguina, 2006; Schmeling, 2009; Chang et al., 2012). For this study, since we require to measure investor sentiment in 10 different countries, which are Austria (OE), Belgium (BG), Finland (FN), Germany (BD), Greece(GR), Italy (IT), Netherlands (NL), Portugal(PT), Spain (ES) and the United Kingdom (UK) respectively. we obtain the consumer confidence index from the European Commission. The European Commission, provides a database comprising professionally conducted surveys as well as economic and financial indicators for the EU countries. The data range of the consumer confidence index is consistent with the dependent variable, stock returns, starting from May 1985 until December 2015. However, for a few countries the sample period is shorter as displayed in the descriptive statistics table (Table 4.1).

We also include the macroeconomic factors as mentioned above in the attempt to improve the predictions of stock returns and avoid possible mis-specification of the model. To be specific, we use the consumer price index (CPI), a three months treasury bill interest rate (IR), industrial production index (IP) respectively. The choice of variables is suggested by a numbers of previous works that have shown this that these factors may contribute to explain and forecast stock returns. (Bossaerts and Hillion, 1999; Welch and Goyal, 2007;

Westerlund and Narayan, 2014).

Table 4.1 shows the summary statistics for all of the data for the ten European countries both in levels (panel A) and in first differences [2]. The time period of monthly data is from 1985 May through December 2015, giving a total of 368 observations for each time-series. However, this time span is not available for some variables. For these countries, the analysis will be performed for slightly shorter time periods. The available time period for each country is also provided on the table in the third column of table 4.1. As usual, We split the whole time period into in-sample and out-of-sample parts, where the in-sample data is used to obtain the optimal candidate forecasting models and the out-of-sample is adopted to examine the accuracy of prediction both for individual models and averaged model.

[2]Referring to the data description part in chapter 3 at page 68

Table 4.1: Summary statistics for all countries

Panel A: Summary statistics for all countries in levels

Country	Label	Start	Price mean	Price Std	CCI mean	CCI Std	CPI mean	CPI Std	IR mean	IR Std	IP mean	IR Std
Austria	OE	1985M05	501.7500	238.5607	-2.0619	7.4498	84.0527	15.3937	4.7883	2.5732	76.7806	22.6077
Belgium	BG	1985M05	766.3796	332.9616	-6.9567	8.3107	83.3935	15.1176	5.6141	2.4647	77.0425	17.4289
Finland	FN	1995M11	465.6659	352.3154	-1.6467	3.8952	85.8350	14.6096	5.8165	3.7386	81.6005	20.7711
Germany	BD	1985M05	562.6675	265.0555	-8.0714	7.1449	85.6557	13.8313	4.8345	2.1387	89.0997	13.4326
Greece	GR	1988M01	635.9520	550.7181	-31.8666	19.2422	67.6783	28.8649	7.6197	5.3030	106.8248	11.3246
Italy	IT	1985M05	774.2981	331.2151	4.9042	8.7430	79.6281	19.7426	6.1747	3.8154	104.6948	9.7152
Netherlands	NL	1985M05	855.5219	439.6469	0.8123	19.2055	83.6066	11.2678	4.9832	2.0860	85.6289	11.2678
Portugal	PT	1988M01	127.8210	51.9672	-13.3928	12.1429	76.6693	23.9569	5.9842	3.5186	106.8605	13.1616
Spain	ES	1985M05	660.4650	347.4518	-11.8194	9.6582	76.7201	21.2798	7.2243	3.8501	101.6653	12.8917
UK	UK	1985M05	1395.1280	528.8090	-8.5750	7.5435	81.8147	17.5049	6.1734	2.8516	101.9236	6.6103

Panel B: Summary statistics for all countries in differences

Country	Label	Start	Return mean	Return Std	CCI mean	CCI Std	CPI mean	CPI Std	IR mean	IR Std	IP mean	IR Std
Austria	OE	1985M06	0.0024	0.0692	0.0002	0.0284	0.0015	0.0032	-0.0040	0.1168	0.0028	0.0152
Belgium	BG	1985M06	0.0045	0.0566	0.0013	0.0348	0.0016	0.0027	-0.0047	0.0752	0.0020	0.0225
Finland	FN	1988M02	0.0070	0.0833	-0.0001	0.0203	0.0013	0.0030	-0.0042	0.1253	0.0028	0.0152
Germany	BD	1985M06	0.0045	0.0583	0.0004	0.0269	0.0014	0.0032	0.0011	0.2024	0.0014	0.0015
Greece	GR	1988M02	-0.0013	0.1033	-36.5351	16.8112	67.6783	28.8649	7.6197	5.3030	106.8248	11.3246
Italy	IT	1985M06	0.0031	0.0667	0.0010	0.0362	0.0020	0.0020	-0.0059	0.0524	-0.0002	0.0130
Netherlands	NL	1985M06	0.0051	0.0515	0.0008	0.3686	0.0015	0.0044	-0.0033	0.0910	0.0013	0.0266
Spain	ES	1985M06	0.0055	0.0643	0.0013	0.0390	0.0027	0.0048	-0.0040	0.0521	0.0005	0.0171
Portugal	PT	1985M06	-0.0006	0.0602	-21.5276	13.7400	76.6693	23.9569	5.9842	3.5186	106.8605	13.1616
UK	UK	1985M06	0.0041	0.0451	0.0012	0.3384	0.0022	0.0042	-0.0040	0.0451	0.0003	0.0010

Table 4.2: Unit Root test result

	Im, Pesaran and Shin		*ADF Fisher Chi-sq*		*PP Fisher Chi-sq*	
	Statistic	Prob	Statistic	Prob	Statistic	Prob
OE	-15.668	0.0000***	255.211	0.0000***	655.230	0.0000***
BG	-32.186	0.0000***	548.710	0.0000***	659.702	0.0000***
FN	-28.924	0.0000***	452.481	0.0000***	678.733	0.0000***
BD	-18.834	0.0000***	301.476	0.0000***	716.332	0.0000***
GR	-16.357	0.0000***	257.495	0.0000***	569.502	0.0000***
IT	-22.592	0.0000***	346.826	0.0000***	697.158	0.0000***
NL	-24.528	0.0000***	420.703	0.0000***	583.324	0.0000***
PT	-24.573	0.0000***	399.343	0.0000***	590.213	0.0000***
ES	-23.684	0.0000***	368.345	0.0000***	681.750	0.0000***
UK	-21.566	0.0000***	336.612	0.0000***	708.114	0.0000***

The table shows the unit root results for each country individually
Asterisks refer to the level of significance where: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

4.3.2 Result for Unit Root Tests

As the forecasting models we use are only suitable for stationary processes, the starting point is to make sure all the series used are stationary. Since the data in levels are non-stationary, it is necessary to test whether the series in first differences are stationary or not. To be specific, we employ the Im, Pesaran and Shin (IPS) test, ADF-Fisher test and the PP-Fisher unit root tests to examine the stationary of the first difference processes. The null hypothesis is that there is a unit root. The optimal lag length is selected by the modified Akaike information criterion (AIC). As shown in table 4.2, we have 0 probability to accept the null hypothesis, which means we have over the percentage of over 99.99 to reject the null hypothesis. All of series are therefore assumed to be stationary.

4.4 Methodology

As previously mentioned, different methods have their advantages and disadvantages. Selecting the optimal individual forecasting model for each class is the first issue to be solved before using model averaging. In this section we first introduce the individual forecasting models that we will merge into model averaging, and then we outline the different model averaging methods.

4.4.1 Selecting Individual Models

In this subsection, we briefly explain the individual forecasting models we use to predict stock returns, namely the ARMA, ARMAX, VAR and BVAR respectively. To be specific, we use ARMA models to examine whether the stock returns are predictable by its own past. Also, we apply the ARMAX, VAR and BVAR models to consider the case where the stock returns could be affected by investor sentiment and some other macroeconomic factors.

4.4.1.1 ARMA

The ARMA model, as one of the most commonly accepted univariate models, was first proposed by (Whittle, 1951) and made widely popular by (Box and Jenkins, 1968). It describes a weekly stationary process as the sum of an auto-regressive (AR, hereafter) and moving average (MA, hereafter) polynomials and has been widely adopted in many previous studies. Specifically, Karanasos (2001) applied the ARMA model to forecast the conditional mean and variance of stock returns. They derived the formula of the conditional variance and selected the expressions for the optimal predictors among all candidate models. Later, Henry (2002) discussed the long horizon predictability in stock returns internationally and proved that return predictability came from time variation in a large proportion rather than from long memory. Rounaghi and Zadeh (2016) used ARMA models to model and predict stock returns in SP 500 and London Stock Exchange in medium and long time horizons.

According to former papers, the ARMA(p,q) model in our chapter is shown as follows:

$$r_t^{(n)} = \eta^{(n)} + \sum_{i=1}^{p} \gamma_i^{(n)} r_{t-i}^{(n)} + \sum_{j=1}^{q} \theta_i^{(n)} \varepsilon_{t-i}^{(n)} + \varepsilon_t^{(n)} \tag{4.1}$$

On the left hand side, r_t is the stock returns with AR process with p lag length and MA process with q lag length. n represents ten European countries. On the right hand side, ε_{t-i} is the stationary white noise. This expression is estimated by ordinary least squares (OLS) method. We set the maximum values of p and q are 5 which gives a total of 36 ARMA(p, q) models in each country. There are two types of special cases among these models: one is that q=0, the model would become an AR(p) process while p=0, it would be an MA(q) process. We use information criteria to select the optimal values of p and q. After determining the optimal model, we can use the following forecasting expression to calculate the predictions out-of-sample.

$$f_t^{(n)} = \hat{\eta}^{(n)} + \sum_{i=1}^{p} \hat{\gamma}_i{}^{(n)} r_{t-i}^{(n)} + \sum_{j=1}^{q} \hat{\theta}_i{}^{(n)} \varepsilon_{t-i}^{(n)} \tag{4.2}$$

Comparing the predicted value with the actual value provides information on the forecasting performance of the model.

4.4.1.2 ARMAX

Although the ARMA model can predict the stock returns using its historical features, some studies point out that returns may also correlated with other factors such as investor sentiment and macroeconomic factors (De Long et al., 1990; Baker and Wurgler, 2000; Rapach et al., 2005). To account for this, researchers add more exogenous variables to the standard ARMA to obtain the ARMAX model, which has been proved to improve the forecasting performance effectively (Zadrozny, 1988; Akal, 2004; Zheng and Zhu, 2017). In this chapter, we also set up an ARMAX model to include, as exogenous variables, investor sentiment and macroeconomic factors. The expression of the ARMAX (p, q, b) is written as follows:

$$r_t^{(n)} = \sum_{i=1}^{p} \gamma_i^{(n)} r_{t-i}^{(n)} + \sum_{j=1}^{q} \theta_i^{(n)} \varepsilon_{t-i}^{(n)} + \sum_{k=1}^{b} \eta_k^{(n)} d_{t-k}^{(n)} + \varepsilon_t^{(n)} \tag{4.3}$$

where r_t are again the stock market returns for each European country, p and q are still the lag length of ARMA components. The main difference is that the exogenous inputs terms are expressed as d_t and η_k is the parameter of them.

Similar with ARMA model, we also use the usual information criteria to select the optimal model among the sets of models of each country, we can obtain the following expression to calculate the forecasting values:

$$f_t^{(n)} = \sum_{i=1}^{p} \hat{\gamma}_i^{(n)} r_{t-i}^{(n)} + \sum_{j=1}^{q} \hat{\theta}_i^{(n)} \varepsilon_{t-i}^{(n)} + \sum_{k=1}^{b} \hat{\eta_k}^{(n)} d_{t-k}^{(n)} \tag{4.4}$$

4.4.1.3 VAR

As the models mentioned above are all univariate, we also use the VAR model. Compared with the univariate parameterisations, the VAR combines all of the variables into a whole system. As a result, it is more flexible and can improve the accuracy of the predictions. Binswanger (2004) modelled a bivariate structural VAR to investigate the relationship between stock price and the growth rates of industrial production. Similarly in another study of stock returns, Schmeling (2009) generated a panel VAR and found that the stock returns was significantly affected by investor sentiment. In this chapter, we consider the stock returns, investor sentiment and macroeconomic factors as a matrix Y_t into the VAR model so that them as a whole system. Then the expression for the VAR model is shown as below:

$$Y_t = \mu + \sum_{i=1}^{p} \delta_i Y_{t-i} + \epsilon_t \tag{4.5}$$

In this expression, $Y = [R, \Delta Sent, \Delta CPI, \Delta IR, \Delta IP]$ represents the monthly variables matrix where n indicates different countries. represents the lag length and δ is the parameter for the lagged term of Y, which is a time-invariant $kkmatrix$. μ and ϵ_t are the $k1$ vectors, namely the constant term and the error term respectively. Each VAR model is estimated by maximum likelihood (MLE) method.

We set the maximum lag length equals to 8 and use the standard information criteria to select the optimal VAR model among these sets of models. The forecasting value can be calculated by the following expression:

$$f_{t+1} = \hat{\mu} + \sum_{i=1}^{p} \hat{\delta}_i Y_{t-i} \tag{4.6}$$

The difference between VAR and former models is that the forecasting value we calculate is still a matrix including all of the variables. In other words, the predicting values of stock returns are determined by its own historical values and the past values of other variables as well.

4.4.1.4 BVAR

As we know, one of the advantages of VAR models is that they usually produce better forecasts than univariate models and they also do not need to make any assumption on the exogeneity of the variables modelled. However, given that the VAR are symmetric systems they are likely to be overparameterised and thus over-fit the data (Carriero et al., 2009). Based on the standard VAR, the Bayesian VAR used Bayesian methods to estimate the parameters. Specifically, all of the parameters are taken as randomly evolving and are assigned prior probabilities. In this way, coefficients of the variables with longer lag length would be close to 0 and the uncertainty of the parameters would be reduced, namely improving the accuracy of the prediction.

In this chapter, we use the BVAR model alongside the standard VAR. The only difference is the way of obtaining the parameter estimators. The added posterior parameters are shown in the following expression:

$$Y_{i,t} = \mu + \sum_{i=1}^{p} sum_{j=1}^{Q} \delta_{i,j} Y_{j,t-i} + \epsilon_t \tag{4.7}$$

where the prior probabilities for coefficients are based on the prior density function $\bar{\Omega} = \frac{p(Y_{i,t}|\delta_{i,j})p(Y_{i,t})}{p(Y_{i,t})}$, where $p(Y_{i,t}$ is the prior probability density function.

Similar to VAR model, the forecasting values can be calculated by the following equation:

$$f_{i,t+1} = \mu + \sum_{i=0}^{p-1} sum_{j=1}^{Q} \hat{\delta_{i,j}} Y_{j,t-i} + \tag{4.8}$$

4.4.2 Information Criteria Model Selection

At this point, we need to select the optimal lag lengths for each class of models discussed above. The basic assumption is that the best forecasting model could always exist within each class of models. Following previous studies, we again use the information criteria to determine the optimal candidate model in this chapter. The model with the smallest information criteria is suggested to be the optimal model of its class.

The Information criteria we will adopt for model selections are: the Akaike Information Criterion (AIC), the Schwarz's Bayesian Information Criterion (SBIC) and the Hannan-quinn Criterion (HQIC). These three types of information criteria are also widely adopted in many past papers (Wagenmakers and Farrell, 2004; Acquah, 2010; Dziak et al., 2020). These three information criteria can be expressed as follows:

$$AIC = ln(\hat{\sigma}^2) + \frac{2K}{T} \tag{4.9}$$

$$SBIC = ln(\hat{\sigma}^2) + \frac{K}{T} lnT \tag{4.10}$$

$$HQIC = ln(\hat{\sigma}^2) + \frac{2K}{T} lnlnT \tag{4.11}$$

where $\hat{\sigma}^2$ represents the in-sample fitted error $\hat{\epsilon_t}(m) = y\hat{y}(m)$, K is the number of parameters and T is the number of total in-sample observations. As we can see, the information criteria is a function of the residual variance part and a penalty component for over-parameterisation. In practice, we often use the log of the likelihood function value rather than the residual sum of squares divided by the number of in-sample observations, T.

4.4.3 Model Averaging

After selecting the optimal individual model in each class, the next issue is to determine the weight that each optimal model should have in producing forecasts. In this subsection, we will outline the three methods of model averaging that we will use to improve the performance of prediction of the individual optimal models: Simple Model Averaging, Bayesian Model Averaging and the Akaike Model Averaging respectively.

4.4.3.1 Simple Model Averaging

The simple model averaging (SMA, hereafter) method attach equal weights to each of the optimal models we obtain from the different sets of models estimated in-sample. The weights for SMA can be expressed as:

$$w_s = \frac{1}{M} \tag{4.12}$$

where w_s is the weight of each selected model and M is the number of the candidates forecasting models.

Thus the forecasting expression of SMA can be written as follows:

$$f_t = \sum_{s=1}^{M} \frac{1}{M} f_t(s) \tag{4.13}$$

As we know, the SMA can improve the predictive performance of the individual models under the assumption that all of the candidates forecasting models are well specified. If the selected models are not well-specified, the accuracy of the prediction may decrease significantly.

4.4.3.2 Bayesian Model Averaging

Bayesian model averaging (BMA, hereafter) is also known as one of the most common used averaging methods. The basic assumption for BMA is that we can always find one model that performs better than all of the other candidate models. This means that with better performing, the optimal model will have more weight than the other optimal ones which would have slightly less weight attached to them. Given some priors, the probabilities given to the optimal model in each class are called the Bayesian posterior probabilities. To be specific, researchers use the likelihood function to determine the fit of each model (Wasserman et al., 2000; Posada and Buckley, 2004; Montgomery and Nyhan, 2010). To start with, we define the likelihood function (L) as:

$$L = P(Y \mid M, \beta, \zeta, \eta) \tag{4.14}$$

Y represents the data and M comes from the prior probability distribution M$\sim P(M).\beta$ is the vector of model parameter and ζ is the tree topology. η is the vector of branch lengths. They are however nuisance parameters, which should be removed from the model from the inference by assigning prior probabilities to obtain the marginal probabilities of the model. We organize the model likelihood expressions as following:

$$P(Y \mid M) = (Y \mid M, \beta, \zeta, \eta) P(\beta, \zeta, \eta \mid M) d\beta d\zeta d\eta$$

(4.15)

Using the Bayesian Rule, we can calculate the Bayes factors:

$$B_{is} = \frac{P(Y \mid M_i)}{P(Y \mid M_s)} \tag{4.16}$$

A Bayesian solution is used to select the optimal model with the highest posterior probability through multiple models. And the posterior probability of the model is expressed as:

$$P(M_s \mid Y) = \frac{P(Y \mid M_s) P(M_s)}{\sum_{s=1}^{R} P(Y \mid M_s) P(M_s)} \tag{4.17}$$

To make the model likelihoods easier to compute, we often impose the Bayesian Information Criteria (BIC, hereafter) (Schwarz et al., 1978) to calculate the weights of BMA for the reason that the BIC is developed as the approximation value of the log likelihood function.

In other words, the difference between BIC values can be used to estimate the approximated values of the log of the Bayes factors (Wasserman et al., 2000; Posada and Buckley, 2004). Here we use the BIC in log likelihood model expression as shown below:

$$BIC = \frac{-2l}{T} + \frac{K}{T} lnT \tag{4.18}$$

where l represents the averaged log likelihood function value adjusted by a penalty function, K is the number of estimated parameters and T is the size in-sample. In this way, the weights of BMA can be approximated as the following:

$$w_s^b = \frac{exp(-\frac{1}{2}(BIC_s))}{\sum_{p=1}^{M} exp(-\frac{1}{2}(BIC_m))} \tag{4.19}$$

Hence, the forecasting value can be calculated by:

$$f_t = \sum_{i=1}^{M} w_i^b f_t(i) \tag{4.20}$$

where $w_i^b = w_m^b$ and M represents the number of models we select.

4.4.3.3 Akaike Model Averaging

The Akaike Model Averaging (AMA, hereafter) can also be used to both select optimal models. Actually, there is evidence that AMA may also work more effectively than BMA to produce better forecasts (Posada and Buckley, 2004; Symonds and Moussalli, 2011). Similarly, we can also use the log likelihood version of Akaike information criteria (AIC) to determine the weight function of AMA in the same way as we do for BMA.

$$AIC = \frac{-2l}{T} + \frac{2K}{T} \tag{4.21}$$

where l is still the log likelihood function. We can calculate the Akaike weight that should be given to each of the predictive models that we have selected in by the following equation:

$$w_s^a = \frac{exp(-\frac{1}{2}(AIC_s))}{\sum_{p=1}^{M} exp(-\frac{1}{2}(AIC_m))} \tag{4.22}$$

And the forecasting model with AMA is given as:

$$f_t = \sum_{i=1}^{M} w_i^a f_t(i) \tag{4.23}$$

where w_i^a equals to w_m^a and M represents the number of selective models.

4.5 Analysis of Forecasting Results

In this section, we discuss the forecasting results and the evaluation of them obtained by different model averaging methods. To be specific, we generate both dynamic and static out-of-sample forecasts. Also, to measure the quality of model averaging, we use three different statistics to make comparison among these models, namely the root mean squared error (RMSE, hereafter), the mean absolute error (MAE, hereafter) and the mean absolute percentage error (MAPE) respectively.

Apart from comparing different model averaging methods, we also make comparison between the results predicted by model averaging methods and individual forecasting model to examine whether model averaging actually improve the accuracy of individual optimal forecasting models.

4.5.1 Comparisons Among Individual Forecasting Models

We forecast the stock returns for ten European countries using ARMA, ARMAX, VAR and BVAR models. Table 4.3 presents the dynamic and static RMSE, MAE and MAPE respectively. Generally speaking, it is not suprisingly to find that the dynamic prediction performs better than static ones since the RMSE, MAE and MAPE in dynamics are smaller than they are in statics in all of these ten European countries. The main difference between the dynamic and static forecasting is that dynamics are the multi-steps ahead forecasts, while the statics are one-step-ahead. As a result, the dynamic forecasts can converge to the long-term unconditional mean value more rapidly as the forecasting horizon increases. However, the static forecasting results cannot perform the same with dynamics for the reason that they are a fixed one-step-ahead rolling forecast.

Another interesting and perhaps surprising finding is that the ARMA and ARMAX model performs better compared to VAR and BVAR for all of the ten countries. The logical thinking is that if we combine all the relevant variables into a system model and evaluate them together, the forecasting performance should be better Avramov (2002); Binswanger (2004); Cochrane (2008). However, there are also many researchers that found that the forecasting performance of ARMA model is in some cases still superior (Swider and Weber, 2007; Kambouroudis et al., 2016). The possible reasons of such results could be as follows: firstly, since we selected the optimal model in each class by information criteria, these forecasting models include different lag length in different countries. As the forecasting values of the variables are correlated with their own past, there would be a chance that model with the longer lag length would be more precise. However, models with too many lags may also cause over-fitting. The second possible reason is that ARMA models includes the MA process where the VAR model doesn't. If the estimation process has the existing MA patterns by extra auto-regressives lags, ARMA or ARMAX model may perform better since they include the MA process in the model directly.

In greater details, for most of the countries, the ARMAX models obtain a smaller RMSE, MAE, MAPE than the ARMA models except for Greece, Portugal and the UK. This result seems convincing for the reason that ARMAX follows the same process with ARMA but also adding more exogenous variables, namely investor sentiment and macroeconomic factors, which are likely to improve the accuracy of predictions. The reason that ARMA performs better for UK is because the auto-regressive lags decrease from 3 to 0 for the ARMAX, and this may clealy affect the accuracy of forecasts for the model. For Portugal, when macroeconomic factors are added into the model, the $AR(2)$ and $AR(3)$ lags are no longer significant for stock returns. Another special case is in Greece, where the lags of the MA process increase from 1 to 2 when adding investor sentiment and macroeconomic factors into the model. This change could perhaps cause some over-fitting of the model and decrease the accuracy of forecasting performance.

Finally, we make comparisons between VAR and BVAR models. The BVAR model performs better than the standard VAR in over half of these ten countries, which proves that giving prior probabilities to parameters can improve the forecasting performance to some extent.

However, the BVAR is performing always better than the VAR model as shown for Belgium, Finland, Greece and UK. The possible reason for this perhaps surprising occurrence, is that the BVAR model is more suitable for when we include macroeconomic variables due to its fundamental properties. In fact, if macroeconomic factors perform poorly in-sample, Bayesian methods would not include them in the forecasting model.

4.5.2 Comparison Among Forecasting with Model Averaging Methods

Looking at the forecasts for individual countries, the ARMAX model seems to perform better than all other models. However, other models may also have some forecasting power that cannot be neglected. To improve accuracy of stock return forecasts, a better way is to merge them in a "composite" forecasting model, which contains all of these selected optimal models from each class suitably weighted. This process is the model averaging. We use the SMA, BMA and AMA to determine the weight of selected models and the forecasting results are shown in table 4.4.

Generally speaking, there are no huge differences among the statistics calculated for different model averaging methods. This implies that the weights for each selected model are actually relatively close. In other words, each of the individual methods we use to forecast is fairly well-specified and there are no obvious shortcomings of any single optimal model. Specifically, we can find the BMA method performs better in Belgium, Finland, Italy, Netherland and Spain, where the SMA predicts more accurately for Austria and Greece by using the dynamic RMSE. By dynamic MAE, we can observe that BMA has the best forecasting performance for Italy, Netherlands and Spain. While looking at the values of dynamic MAPE, it suggests that the AMA performs better in Austria, Finland, Greece, Portugal and the UK. Also we can notice that in Germany and Netherlands, the BMA performs better compared with other model averaging methods. Finally, SMA performs best when we observe the RMSE values in the Austria, Greece or consider the MAPE values for the Germany and Italy.

Similarly with the results in individual countries, we can find the RMSE, MAE and MAPE for dynamic forecasting models are smaller than the static ones (except for Italy and Netherlands).

4.5.3 Comparisons between Individual Forecasting Model and Model Averaging

Although many researchers state model averaging can improve the forecasting performance compared with individual models, little have made comparison between individual forecasting models and the model averaging method. By comparing table 4.3 and table 4.4, we can find some similarities across these countries. The first finding is that all of the evaluation statistics of model averaging methods, no matter dynamic or static, report lower values than those for any individual model. In other words, model averaging methods have significantly superior forecasting performance compared with the selected optimal models in each class. Secondly, the values of RMSE and the MAE in model averaging reduced over half compared with those calculated by any selected single forecasting models. Thirdly, as we can notice that the values of MAPE, as the percentage error, were over 1 in most of the selected optimal models. With model averaging methods, the values of MAPE are all below 0.6. All of these three findings strongly support that the forecasting performances of stock returns can be significantly improved by model averaging.

In the meanwhile, we can notice that the RMSE and MAE in Greece take the maximum value compared with other countries both in multiple models and model averaging forecasts. To be specific, the RMSEs of Greece in different models are all above 0.1 in table 4.3, where the values drop to 0.059 in table 4.4. Although it still follows the pattern that model averaging is superior to any single model, the forecasting performance is relatively weak compared with other countries. One possible reason is that the Greek government-debt crisis happened between late 2009 and 2010, which strikes the investor confidence a lot and also the financial market fluctuates a lot during these years. Also, the dynamic RMSEs of Greece are larger than the static ones in single models. In other words, static forecasting performs better than dynamic one, which means that returns in Greece stock market are more predictable in short-term period rather than long-term periods. However, dynamic forecasting shows better performance than static one by model averaging. This also supports that model averaging can improve the forecasting accuracy in long periods of time.

4.6 Conclusion

As forecasting the stock returns is a very important topic in financial economics both for researchers and practitioners, it is necessary to find better methods to predict them more accurately. academic research have focused on this topic for several decades already(Baker and Wurgler, 2000; Swider and Weber, 2007; Schmeling, 2009; Carriero et al., 2009; Zheng

Table 4.3: Individual Forecasting results in European Countries

Country	Model	Dynamic RMSE	Dynamic MAE	Dynamic MAPE	Static RMSE	Static MAE	Static MAPE
OE	ARMA	0.066	0.049	1.633	0.070	0.056	6.619
	ARMAX	0.064	0.050	1.569	0.068	0.053	1.477
	VAR	0.201	0.116	1.113	0.865	0.676	1.096
	BVAR	0.119	0.092	1.120	0.932	0.681	1.141
BG	ARMA	0.050	0.037	1.596	0.053	0.040	5.167
	ARMAX	0.049	0.039	1.439	0.049	0.039	1.460
	VAR	0.095	0.073	0.922	0.448	0.348	1.318
	BVAR	0.115	0.078	0.905	0.654	0.518	1.165
FN	ARMA	0.062	0.046	1.165	0.065	0.049	1.416
	ARMAX	0.058	0.045	1.349	0.059	0.047	1.396
	VAR	0.070	0.050	1.711	0.453	0.356	1.182
	BVAR	0.308	0.096	2.946	1.082	0.858	1.015
BD	ARMA	0.061	0.045	1.310	0.061	0.046	1.310
	ARMAX	0.056	0.043	1.546	0.056	0.043	1.561
	VAR	0.147	0.081	1.076	0.675	0.548	1.065
	BVAR	0.097	0.074	1.059	0.745	0.539	1.209
GR	ARMA	0.143	0.113	1.248	0.141	0.111	2.604
	ARMAX	0.203	0.133	1.191	0.203	0.135	1.199
	VAR	0.318	0.287	0.898	4.449	2.813	1.112
	BVAR	0.558	0.399	0.97	3.530	2.408	0.999
IT	ARMA	0.065	0.052	1.082	0.066	0.053	1.181
	ARMAX	0.059	0.045	1.243	0.059	0.045	1.212
	VAR	0.085	0.070	1.933	0.829	0.623	1.117
	BVAR	0.073	0.059	1.884	0.786	0.545	1.086
NL	ARMA	0.049	0.036	0.977	0.049	0.036	1.003
	ARMAX	0.048	0.037	1.490	0.048	0.037	1.487
	VAR	0.144	0.068	0.620	0.68	0.55	1.216
	BVAR	0.258	0.080	2.428	0.938	0.737	1.029
PT	ARMA	0.064	0.049	1.068	0.070	0.055	1.722
	ARMAX	0.068	0.052	1.492	0.077	0.060	1.510
	VAR	0.305	0.106	1.608	1.266	0.966	0.347
	BVAR	0.218	0.085	1.884	1.279	0.993	1.093
ES	ARMA	0.064	0.052	1.663	0.065	0.053	2.294
	ARMAX	0.063	0.049	1.328	0.062	0.047	1.282
	VAR	0.300	0.147	2.146	1.761	1.441	0.982
	BVAR	0.145	0.090	1.537	0.809	0.631	1.045
UK	ARMA	0.039	0.030	1.509	0.040	0.030	1.768
	ARMAX	0.041	0.033	1.586	0.041	0.032	1.510
	VAR	0.050	0.034	2.233	0.482	0.368	1.051
	BVAR	0.040	0.030	1.098	0.497	0.369	1.279

Table 4.4: Forecasting Performances in Different Model Averaging Methods

Country	Model	Dynamic RMSE	Dynamic MAE	Dynamic MAPE	Static RMSE	Static MAE	Static MAPE
	SMA	0.02612	0.00798	0.28115	0.02787	0.00879	0.48680
	BMA	0.02613	0.00798	0.27992	0.02787	0.00879	0.49130
OE	AMA	0.02614	0.00798	0.27282	0.02789	0.00879	0.49082
	SMA	0.02003	0.00618	0.30666	0.02039	0.00630	0.30298
	BMA	0.02002	0.00618	0.30633	0.02039	0.00630	0.30183
BG	AMA	0.02003	0.00618	0.30650	0.02040	0.00630	0.30110
	SMA	0.02435	0.00741	0.18656	0.02545	0.00795	0.25716
	BMA	0.02438	0.00742	0.18643	0.02546	0.00796	0.25536
FN	AMA	0.02440	0.00742	0.18624	0.02548	0.00796	0.25534
	SMA	0.02403	0.00735	0.22505	0.02465	0.00751	0.23183
	BMA	0.02403	0.00766	0.23056	0.02507	0.00761	0.23617
BD	AMA	0.02403	0.00735	0.22512	0.02465	0.00752	0.23188
	SMA	0.05927	0.01885	0.53382	0.06254	0.02047	1.33757
	BMA	0.05942	0.01885	0.56128	0.06285	0.02055	1.37338
GR	AMA	0.05936	0.01885	0.55206	0.06276	0.02053	1.36353
	SMA	0.02504	0.00802	0.16802	0.02475	0.00798	0.17709
	BMA	0.02499	0.00799	0.16828	0.02471	0.00796	0.17630
IT	AMA	0.02500	0.00800	0.16821	0.02471	0.00796	0.17637
	SMA	0.01939	0.00580	0.16651	0.01884	0.00574	0.19729
	BMA	0.01939	0.00580	0.16646	0.01884	0.00574	0.19797
NL	AMA	0.01939	0.00581	0.16662	0.01884	0.00574	0.19657
	SMA	0.02720	0.00878	0.20790	0.02843	0.00922	0.27147
	BMA	0.02720	0.00878	0.20871	0.02846	0.00924	0.27291
PT	AMA	0.02720	0.00878	0.20863	0.02847	0.00925	0.27326
	SMA	0.02510	0.00812	0.25024	0.02558	0.00821	0.49470
	BMA	0.02509	0.00811	0.25069	0.02555	0.00820	0.49250
ES	AMA	0.02510	0.00812	0.25024	0.02547	0.00818	0.48046
	SMA	0.01592	0.00490	0.23842	0.01619	0.00493	0.25932
	BMA	0.01592	0.00490	0.23853	0.01619	0.00493	0.25936
UK	AMA	0.01592	0.00490	0.23857	0.01619	0.00493	0.25934

and Zhu, 2017). In this chapter, we used model averaging methods to improve the accuracy of predictions for the stock returns in ten European countries. We firstly selected optimal forecasting models from each class for them to be subsequently averaged. To be specific, we considered both univariate and multivariate models, namely ARMA, ARMAX, VAR and BVAR, to predict stock markets in ten European countries. Then we used information criteria to select the optimal model among all of these classes of models. Although a variety of models have been adopted in analyzing stock markets, researchers still argued that it was hard to find a model that could be strongly superior than all of the other models. To solve this issue, we used model averaging, which combined the optimal models from each class using weights based on the usual information criteria. The main advantage of model averaging is that it can keep the benefits and weaken the shortcoming of a single forecasting model by given different weight to each model. We adopted three model averaging methods, that were the simple model averaging, Bayesian model averaging and the Akaike model averaging.

We made comprehensive comparisons among different types of models and different model averaging methods. The first conclusion we get is that dynamic forecasts always performs better than the static forecasts, both for individual forecasting models and using model averaging methods. Secondly, since the optimal models we selected are well-specified, the weight of each model under BMA and AMA are not too different. Overall, the BMA performs best across these three model averaging methods while there are also cases where AMA or SMA performs (equally) better. Last but least is that, no matter which model averaging method is used, as highlighted in previous studies, the forecasting performance of it is always superior to that of any individual forecasting model we selected (Hoeting et al., 1999; Posada and Buckley, 2004; Claeskens et al., 2008).

In brief, we contributed to the literature on forecasting stock return in three ways as, firstly we successfully used model averaging methods to improve the accuracy of forecasting in ten European countries. Secondly, we made a thorough comparison between the selected optimal forecasting models and average models. In previous studies of this area, most of the researchers only concentrated on generating predictions generated by single optimal forecasting models (Baker and Wurgler, 2006; Swider and Weber, 2007; Rounaghi and Zadeh, 2016). Although there has been research concentrating on model averaging methods (Montgomery and Nyhan, 2010; Liao et al., 2019), very few papers offered thorough comparison between individual models and model averaging, especially in empirical analysis. Finally, it is important to highlight that improving the forecasting performance of stock returns can help understanding the stock market better for regulators as well as being beneficial for investors to make rational decision.

Chapter 5

Conclusion

Forecasting the stock returns is always a very important issue in financial applications. The failure of the EMH to explain the market anomalies brought researchers' attention to the effect of investor sentiment on stock returns. A bullish investor may often overvalue the stock return while a bearish investor may "bet" on the contrary. In chapter 1, we aim to investigate the relationship between investor sentiment and stock returns. Since investor sentiment seems unlikely to be observed directly, we have to choose suitable proxy for it. After comparing all types of measurements for investor sentiment, we use the CCI, as the proxy for investor sentiment, to estimate and forecast stock returns in ten European countries. Both the panel fixed-effect and the individual model regression show that the investor sentiment can positively affect stock returns in all of these ten European countries. In other words, the future stock returns will increase with higher expectation of investors and will decrease with lower expectation of investors. By comparing the results of estimations between panel and individual countries, we cannot find strong country-specific differences in terms of significance of investors sentiment. In detail, in some countries the effect of investors sentiment on stock returns may be stronger in magnitude but it is overall significant and has the same direction for all the countries thereby providing a direction for future works on predicting stock returns.

In chapter 2, we also used both univariate models (ARMA, ARMAX) and multivariate models (VAR and BVAR) to observe the in-sample forecasts for stock returns. We employ information criteria to determine the optimal forecasting model from each class of models. We then compare the performance of the chosen model out-of-sample. In general, we observe that the stock returns are predictable mostly by their own past, and partly by investor sentiment and macroeconomic factors in those European countries. In details, the ARMAX performs better than the standard ARMA model and BVAR provides more accurate forecasts of stock returns compared with the traditional VAR model. Another interesting finding is that in some countries, the ARMAX provides the best forecasting performance among these four models. However, there are also some special cases that BVAR performs better than any other forecasting models.

In chapter 3, we use model averaging method to obtain a composite average model containing all of the optimal models selected in chapter 2 within each class, to improve the quality of forecasting of stock returns. Specifically, model averaging involves distribute different weight to each of the optimal models and then obtain a weighted average of the individual forecasts. We attributed the weights to these candidate optimal models by the simple model averaging, Bayesian model averaging and Akaike model averaging methods. Then we make general comparisons of out-of-sample forecasting errors among each of the individual models and the average models. Firstly the out-of-sample results show that the dynamic forecasts have more accuracy than the static ones, both in single models and averaged models. Since the results for model averaging methods do not have huge differences, we can conclude that the optimal models we select within each class, are well-specified. However, even with little differences among them can prove that BMA performs better than the other two model averaging methods. Finally, no matter which averaging method we consider, it always forecasts more accurately compared with any of the individual models, thereby proves that model averaging can definitely improve the accuracy of forecasting performance compared with the optimal single models.

To conclude, our analysis provides evidence of the predictability of stock returns in European countries from different aspects. The first is that the relationship between investor sentiment and stock returns can help the stock market regulators by observing the intrinsic patterns in stock markets. Moreover, forecasting stock returns in a precise way would allow to prevent the stock market from sudden breaks more effectively. With higher accuracy of forecasting performance of stock returns, professionals can manage stock market better and investors are encouraged to gain more in stock market.

Appendix A

ARMA In-sample Model Results

Table A.1: ARMA In-sample Forecasting Model Result

OE	Variable	Coef.	Std.Error	Prob
	C	0.004	0.006	0.538
	AR(1)	0.948	0.107	0.000***
	AR(2)	-0.643	0.107	0.000***
	MA(1)	-0.756	0.107	0.000***
	MA(2)	0.418	0.112	0.000***
	MA(3)	0.309	0.050	0.000***
BG	Variable	Coef.	Std.Error	Prob
	C	0.004	0.005	0.383
	AR(1)	-1.176	0.051	0.000***
	AR(2)	-0.863	0.050	0.000***
	MA(1)	1.344	0.073	0.000***
	MA(2)	1.139	0.084	0.000***

	MA(3)	0.143	0.058	0.014***
FN	Variable	Coef.	Std.Error	Prob
	C	0.009	0.006	0.152
	AR(1)	-0.448	0.220	0.042**
	MA(1)	0.593	0.195	0.003***
BD	Variable	Coef.	Std.Error	Prob
	C	0.005	0.005	0.360
	AR(1)	0.801	0.443	0.072*
	MA(1)	-0.773	0.471	0.102
GR	Variable	Coef.	Std.Error	Prob
	C	0.005	0.009	0.528
	AR(1)	0.883	0.133	0.000***
	MA(1)	-0.839	0.154	0.000***
IT	Variable	Coef.	Std.Error	Prob
	C	0.004	0.006	0.427
	AR(1)	0.871	0.149	0.000***
	MA(1)	-0.822	0.175	0.000***
NL	Variable	Coef.	Std.Error	Prob
	C	0.004	0.004	0.286
	AR(1)	-0.975	0.052	0.000***
	MA(1)	0.957	0.066	0.000***
PT	Variable	Coef.	Std.Error	Prob
	C	0.001	0.005	0.832
	AR(1)	-0.738	0.086	0.000***

	AR(2)	0.118	0.058	0.042**
	AR(3)	0.144	0.068	0.035**
ES	Variable	Coef.	Std.Error	Prob
	C	0.007	0.005	0.113
	AR(1)	-0.314	0.396	0.429
	MA(1)	0.411	0.384	0.285
UK	Variable	Coef.	Std.Error	Prob
	C	0.005	0.003	0.159
	AR(1)	0.888	0.312	0.005***
	AR(2)	-0.677	0.256	0.009***
	MA(1)	-0.918	0.322	0.005***
	MA(2)	-0.625	0.276	0.024**
	MA(3)	-0.056	0.053	0.299

Appendix B

ARMAX In-sample Model Result

Table B.1: ARMAX In-sample Forecasting Results

OE	Variable	Coef.	Std.Error	Prob
	C	-0.007	0.0099	0.9405
	OECCI	-0.005	0.0019	0.7987
	OECPI	0.0342	0.0223	0.1268
	OEIPI	-0.0033	0.0032	0.2963
	OER	0.0433	0.0356	0.2256
	AR(1)	0.9061	0.1447	0.0000***
	AR(2)	-0.6023	0.1348	0.0000***
	MA(1)	-0.7459	0.1606	0.0000***
	MA(2)	0.5094	0.1504	0.0009***
	MA(3)	0.2793	0.0843	0.0011***
BG	Variable	Coef.	Std.Error	Prob
	C	-0.0005	0.0049	0.9251

	BGCCI	0.0063	0.0012	0.0000***
	BGCPI	0.0233	0.0143	0.1028
	BGIPI	0.0025	0.0022	0.2491
	BGR	-0.0306	0.0208	0.1423
	AR(1)	-0.4206	0.0329	0.0000***
	AR(2)	-0.9514	0.0411	0.0000***
	MA(1)	0.5102	1.3117	0.6976
	MA(2)	1.0200	5.6976	0.8580
	MA(3)	0.0428	0.2405	0.8590
FN	Variable	Coef.	Std.Error	Prob
	C	0.0051	0.0093	0.5865
	FNCCI	0.0093	0.0041	0.0256**
	FNCPI	0.0036	0.0253	0.8867
	FNIPI	0.0061	0.0040	0.1268
	FNR	0.0134	0.0437	0.7593
	AR(1)	-0.5899	0.3146	0.0625*
	MA(1)	0.6926	0.2855	0.0163**
BD	Variable	Coef.	Std.Error	Prob
	C	0.0046	0.0052	0.3676
	BDCCI	0.0047	0.0016	0.0036***
	FNCPI	0.0031	0.0158	0.8468
	BDIPI	0.0013	0.0024	0.5825
	BDR	0.0443	0.0219	0.0434**
	AR(1)	0.4748	2.2118	0.8302

	MA(1)	-0.4931	2.1841	0.8215
GR	Variable	Coef.	Std.Error	Prob
	C	0.000	0.008	0.969
	GRCCI	0.006	0.002	0.001***
	GRCPI	0.003	0.008	0.694
	GRIPI	0.000	0.003	0.990
	GRR	-0.082	0.016	0.000***
	AR(1)	0.436	0.049	0.000***
	AR(2)	-0.981	0.055	0.000***
	MA(1)	-0.486	0.069	0.000***
	MA(2)	0.970	0.086	0.000***
IT	Variable	Coef.	Std.Error	Prob
	C	-0.0065	0.0063	0.3027
	ITCCI	0.0061	0.0015	0.0001***
	ITCPI	0.0378	0.0311	0.2243
	ITIPI	0.0078	0.0026	0.0033***
	ITR	-0.0593	0.0185	0.0016***
	AR(1)	-0.3237	0.0729	0.0000***
	AR(2)	-0.9591	0.0746	0.0000***
	MA(1)	0.2627	0.0923	0.0048***
	MA(2)	0.8918	0.0949	0.0000***
	MA(3)	-0.0880	0.0646	0.1745
NL	Variable	Coef.	Std.Error	Prob
	C	0.0056	0.0041	0.1716

	NLCCI	0.0015	0.0010	0.1332
	NLCPI	-0.0061	0.0093	0.5151
	NLIPI	0.0018	0.0016	0.2518
	NLR	0.0523	0.0180	0.0039***
	AR(1)	-0.9686	0.0820	0.0000***
	MA(1)	0.9520	0.0973	0.0000***
PT	Variable	Coef.	Std.Error	Prob
	C	0.0004	0.0052	0.9320
	PTCCI	0.0062	0.0013	0.0000***
	PTCPI	0.0141	0.0118	0.2332
	PTIPI	0.0001	0.0015	0.9677
	PTR	-0.0332	0.0196	0.0930*
	AR(1)	-0.8510	0.0917	0.0000***
	AR(2)	0.0034	0.0706	0.9613
	AR(3)	0.0879	0.0817	0.2829
	MA(1)	0.8985	0.0786	0.0000***
ES	Variable	Coef.	Std.Error	Prob
	C	0.0026	0.0039	0.5176
	ESCCI	0.0068	0.0014	0.0000***
	ESCPI	0.0116	0.0121	0.3386
	ESIPI	0.0027	0.0020	0.1684
	ESR	-0.0378	0.0132	0.0046***
	AR(1)	0.6001	0.2123	0.0050***
	MA(1)	-0.7056	0.1904	0.0003***

UK	Variable	Coef.	Std.Error	Prob
	C	0.0041	0.0033	0.2249
	UKCCI	0.0048	0.0011	0.0000***
	UKCPI	0.0038	0.0102	0.7065
	UKIPI	0.0056	0.0027	0.0410**
	ESR	0.0029	0.0092	0.7542
	AR(1)	-0.0667	0.0474	0.1599
	AR(2)	-0.1565	0.0483	0.0013***

Appendix C

VAR In-sample Model Result

Table C.1: VAR In-smple Forecasting Results in Ten European Countries

OE	Variable	OERETURN	OECCI	OECPI	OER	OEIPI
	OERETURN(-1)	0.233***	3.672	0.603**	0.407**	0.963
		(0.072)	(2.969)	(0.241)	(0.161)	(1.404)
	OECCI(-1)	0.004**	-0.092	-0.015**	0.003	0.024
		(0.002)	(0.075)	(0.006)	(0.004)	(0.036)
	OECPI(-1)	0.024	-1.722	0.048	0.080	0.097
		(0.022)	(0.914)	(0.074)	(0.050)	(0.432)
	OER(-1)	-0.038	-0.077	0.043	0.222***	0.544
		(0.032)	(1.344)	(0.109)	(0.073)	(0.635)
	OEIPI(-1)	0.008**	0.093	0.011	-0.012	-0.153**
		(0.004)	(0.158)	(0.013)	(0.009)	(0.075)
	C	-0.004	0.288	0.124***	-0.023*	0.281**

		(0.006)	(0.244)	(0.020)	(0.013)	(0.116)
BG	Variable	BGRETURN	BGCCI	BGCPI	BGR	BGIPI
	BGRETURN(-1)	0.087*	2.147	0.047	0.279*	0.898
		(0.052)	(2.988)	(0.214)	(0.170)	(1.489)
	BGCCI(-1)	0.002**	-0.089	0.007*	0.001	0.033
		(0.001)	(0.061)	(0.004)	(0.004)	(0.030)
	BGCPI(-1)	-0.025	-1.049	0.156***	0.117**	0.481
		(0.016)	(0.804)	(0.058)	(0.048)	(0.401)
	BGR(-1)	-0.071***	-0.006	0.073	0.290***	0.513
		(0.018)	(0.918)	(0.659)	(0.055)	(0.457)
	BGIPI(-1)	0.002	0.132	-0.006	-0.006	-0.366***
		(0.002)	(0.108)	(0.008)	(0.006)	(0.054)
	C	0.005	0.196	0.117***	-0.030**	0.160
		(0.004)	(0.212)	(0.015)	(0.013)	(0.106)
FN	Variable	FNRETURN	FNCCI	FNCPI	FNR	FNIPI
	FNRETURN(-1)	0.134*	3.674**	-0.326	-0.038	0.164
		(0.081)	(1.482)	(0.216)	(1.613)	(0.131)
	FNCCI(-1)	0.004	-0.052	0.004	0.224***	0.004
		(0.004)	(0.079)	(0.012)	(0.086)	(0.007)
	FNCPI(-1)	-0.015	-0.775	0.117	0.430	0.040
		(0.032)	(0.568)	(0.083)	(0.618)	(0.050)
	FNR(-1)	0.034	2.662***	0.110	2.438**	0.208***
		(0.051)	(0.909)	(0.133)	(0.990)	(0.080)
	FNIPI(-1)	-0.003	0.025	0.000	-0.257**	-0.013*

		(0.005)	(0.085)	(0.012)	(0.092)	(0.008)
	C	0.005	0.196	0.117***	-0.030	0.160
		(0.004)	(0.212)	(0.015)	(0.013)	(0.106)
BD	Variable	BDRETURN	BDCCI	BDCPI	BDR	BDIPI
	BDRETURN(-1)	0.014	1.641	0.396*	0.303**	1.951*
		(0.059)	(2.104)	(0.222)	(0.146)	(1.141)
	BDCCI(-1)	0.002	0.088	-0.004	-0.001	0.093***
		(0.002)	(0.058)	(0.006)	(0.004)	(0.032)
	BDCPI(-1)	0.011	0.801	-0.124**	0.067*	0.649**
		(0.015)	(0.546)	(0.058)	(0.038)	(0.296)
	BDR(-1)	-0.011	0.273	0.057	0.265***	0.784
		(0.022)	(0.803)	(0.085)	(0.056)	(0.435)
	BDIPI(-1)	0.004	0.297***	0.001	0.004	-0.147***
		(0.003)	(0.104)	(0.011)	(0.007)	(0.057)
	C	0.002	0.086	0.138***	-0.020*	0.056
		(0.004)	(0.158)	(0.017)	(0.011)	(0.086)
GR	Variable	GRRETURN	GRCCI	GRCPI	GRR	GRIPI
	GRRETURN(-1)	-0.023	-0.086	0.232	0.225	1.311
		(0.085)	(3.398)	(0.760)	(0.361)	(2.439)
	GRCCI(-1)	0.005	-0.128	0.002	-0.005	0.037
		(0.002)	(0.081)	(0.0018)	(0.009)	(0.058)
	GRCPI(-1)	0.001	-0.165	0.013	0.023	0.439
		(0.009)	(0.357)	(0.080)	(0.038)	(0.256)
	GRR(-1)	-0.046	-2.114	-0.098	0.138	0.700

		(0.020)	(0.801)	(0.179)	(0.085)	(0.575)
	GRIPI(-1)	0.002	0.115	-0.013	0.016	-0.422
		(0.003)	(0.101)	(0.023)	(0.011)	(0.072)
	C	0.002	0.115	-0.013	0.016	-0.422
		(0.008)	(0.319)	(0.071)	(0.034)	(0.229)
IT	Variable	ITRETURN	ITCCI	ITCPI	ITR	ITIPI
	ITRETURN(-1)	-0.062	1.104	0.049	0.574**	3.525**
		(0.072)	(2.773)	(0.134)	(0.280)	(1.455)
	ITRETURN(-2)	-0.078	-1.978	-0.203	0.194	2.203
		(0.073)	(2.803)	(0.136)	(0.283)	(1.470)
	ITCCI(-1)	0.001	-0.143**	0.002	0.006	-0.001
		(0.002)	(0.070)	(0.003)	(0.007)	(0.037)
	ITCCI(-2)	0.001	-0.055	-0.002	0.001***	-0.025
		(0.002)	(0.069)	(0.003)	(0.004)	(0.036)
	ITCPI(-1)	-0.006	0.612	0.219***	0.156	1.028
		(0.035)	(1.344)	(0.065)	(0.136)	(0.705)
	ITCPI(-2)	0.013	0.168	0.213***	-0.330**	0.490
		(0.035)	(1.352)	(0.066)	(0.136)	(0.709)
	ITR(-1)	-0.038**	1.358**	0.022	0.297***	0.414
		(0.017)	(0.653)	(0.032)	(0.066)	(0.342)
	ITR(-2)	-0.003	-1.478**	0.014	0.038	-0.018
		(0.017)	(0.673)	(0.033)	(0.068)	(0.353)
	ITIPI(-1)	0.001	-0.008	0.007	-0.001	-0.121
		(0.003)	(0.128)	(0.006)	(0.013)	(0.067)

	ITIPI(-2)	0.010***	0.239*	0.013**	0.003	0.162**
		(0.003)	(0.126)	(0.006)	(0.013)	(0.066)
	C	-0.002	-0.189	0.103***	0.008	-0.261
		(0.009)	(0.335)	(0.016)	(0.034)	(0.176)
NL	Variable	NLRETURN	NLCCI	NLCPI	NLR	NLIPI
	NLRETURN(-1)	-0.078	-0.237	0.306	0.234	2.823
		(0.060)	(3.415)	(0.270)	(0.173)	(2.018)
	NLRETURN(-2)	0.013	3.730	0.010	0.155	4.626**
		(0.059)	(3.370)	(0.266)	(0.170)	(1.991)
	NLRETURN(-3)	-0.001	4.005	0.406	0.042	2.752
		(0.057)	(3.254)	(0.257)	(0.164)	(1.923)
	NLCCI(-1)	0.006***	-0.083	-0.003	0.008**	-0.011
		(0.001)	(0.059)	(0.005)	(0.003)	(0.035)
	NLCCI(-2)	0.004***	-0.028	-0.008	0.000	0.039
		(0.001)	(0.063)	(0.005)	(0.003)	(0.037)
	NLCCI(-3)	0.001	0.071	-0.007	-0.002	0.049
		(0.001)	(0.063)	(0.005)	(0.003)	(0.037)
	NLCPI(-1)	0.004	-0.733*	0.239***	0.035	0.230
		(0.011)	(0.611)	(0.048)	(0.031)	(0.361)
	NLCPI(-2)	-0.013	0.053	-0.128**	0.043	-0.019
		(0.011)	(0.629)	(0.050)	(0.032)	(0.372)
	NLCPI(-3)	0.018*	-0.274	-0.559***	0.032	0.492
		(0.011)	(0.617)	(0.049)	(0.031)	(0.365)
	NLR(-1)	-0.036*	2.053*	0.108	0.302***	0.163

		(0.021)	(1.169)	(0.092)	(0.059)	(0.691)
	NLR(-2)	-0.026	-0.629	-0.080	-0.119*	0.219
		(0.022)	(1.228)	(0.097)	(0.062)	(0.725)
	NLR(-3)	-0.021	-0.137	0.117	0.162***	0.467
		(0.020)	(1.159)	(0.092)	(0.059)	(0.685)
	NLIPI(-1)	-0.001	-0.055	-0.011	0.004	-0.577***
		(0.002)	(0.098)	(0.008)	(0.005)	(0.058)
	NLIPI(-2)	0.000	-0.017	-0.010	0.001	-0.381***
		(0.002)	(0.106)	(0.008)	(0.005)	(0.063)
	NLIPI(-3)	0.001	0.109	0.003	0.005	-0.150**
		(0.002)	(0.098)	(0.008)	(0.005)	(0.058)
	C	0.003	0.102	0.182***	-0.024	0.125
		(0.004)	(0.224)	(0.018)	(0.011)	(0.132)
PT	Variable	PTRETURN	PTCCI	PTCPI	PTR	PTIPI
	PTRETURN(-1)	-0.003	0.012	-0.083	0.355	0.908
		(0.074)	(3.351)	(0.339)	(0.264)	(3.099)
	PTRETURN(-2)	0.024	0.272*	0.613	-0.211	3.038
		(0.074)	(0.160)	(0.340)	(0.265)	(3.107)
	PTCCI(-1)	0.001	0.107*	0.005	-0.009	0.047
		(0.002)	(0.056)	(0.008)	(0.006)	(0.007)
	PTCCI(-2)	0.001	-0.092	-0.018**	0.001	0.030
		(0.002)	(0.075)	(0.008)	(0.006)	(0.069)
	PTCPI(-1)	0.020	-0.201	0.421***	0.031	1.267**
		(0.015)	(0.658)	(0.067)	(0.052)	(0.608)

	Variable					
	PTCPI(-2)	-0.022	0.207	-0.371***	-0.002	-0.230
		(0.015)	(0.663)	(0.067)	(0.052)	(0.613)
	PTR(-1)	-0.027	-1.253	0.068	0.313***	0.985
		(0.020)	(0.921)	(0.093)	(0.073)	(0.852)
	PTR(-2)	-0.001	1.491*	-0.038	0.101	-1.407*
		(0.021)	(0.953)	(0.096)	(0.075)	(0.881)
	PTIPI(-1)	0.001	0.075	0.004	-0.004	-0.583***
		(0.002)	(0.075)	(0.008)	(0.006)	(0.069)
	PTIPI(-2)	0.002	-0.065	0.006	0.004	-0.290***
		(0.002)	(0.075)	(0.008)	(0.006)	(0.070)
	C	0.002	-0.036	0.172***	-0.014	-0.199
		(0.005)	(0.247)	(0.025)***	(0.019)	(0.228)
ES	Variable	ESRETURN	ESCCI	ESCPI	ESR	ESIPI
	ESRETURN(-1)	-0.099*	1.964	-0.102	0.279	1.266
		(0.065)	(2.636)	(0.235)	(0.269)	(1.518)
	ESRETURN(-2)	-0.179***	-3.569	0.068	0.070	1.292
		(0.065)	(2.637)	(0.235)	(0.267)	(1.519)
	ESRETURN(-3)	-0.111*	0.335	-0.267	0.033	0.528
		(0.065)	(2.650)	(0.236)	(0.270)	(1.527)
	ESRETURN(-4)	-0.068	-2.666	0.240	0.681**	0.221
		(0.064)	(2.613)	(0.233)	(0.266)	(1.505)
	ESRETURN(-5)	-0.058	1.976	-0.035	-0.556**	0.177
		(0.064)	(2.605)	(0.232)	(0.265)	(1.501)
	ESRETURN(-6)	-0.021	0.118	-0.282	-0.415*	-0.368

	(0.064)	(2.589)	(0.231)	(0.264)	(1.492)
ESCCI(-1)	0.003*	-0.167***	-0.006	-0.005	0.062*
	(0.002)	(0.063)	(0.006)	(0.006)	(0.036)
ESCCI(-2)	0.003*	0.026	-0.005	0.011*	0.045
	(0.002)	(0.065)	(0.006)	(0.007)	(0.037)
ESCCI(-3)	0.001	-0.022	0.004	0.006	0.029
	(0.002)	(0.065)	(0.006)	(0.007)	(0.038)
ESCCI(-4)	0.001	0.104	0.006	-0.010	0.022
	(0.001)	(0.065)	(0.006)	(0.007)	(0.038)
ESCCI(-5)	-0.001	-0.046	0.002	0.001	0.056
	(0.002)	(0.063)	(0.006)	(0.007)	(0.038)
ESCCI(-6)	-0.002	-0.225***	0.001	0.003	0.095***
	(0.002)	(0.064)	(0.006)	(0.007)	(0.037)
ESCPI(-1)	0.011	-0.828	0.271***	0.070	-0.224
	(0.016)	(0.643)	(0.057)	(0.065)	(0.370)
ESCPI(-2)	-0.029*	-0.677	-0.109*	-0.018	-0.193
	(0.016)	(0.663)	(0.059)	(0.068)	(0.382)
ESCPI(-3)	-0.029*	-1.298*	-0.385***	-0.027	-0.148
	(0.016)	(0.656)	(0.059)	(0.067)	(0.382)
ESCPI(-4)	0.001	-1.077	0.208***	-0.062	-0.025
	(0.016)	(0.660)	(0.059)	(0.067)	(0.380)
ESCPI(-5)	-0.014	-0.850	-0.144**	-0.025	-0.125
	(0.016)	(0.670)	(0.060)	(0.068)	(0.386)
ESCPI(-6)	-0.026	-1.868***	0.413***	0.033	-0.273

	(0.016)	(0.649)	(0.058)	(0.066)	(0.374)
ESR(-1)	-0.040***	-1.054*	0.054	0.246***	0.134
	(0.015)	(0.616)	(0.055)	(0.063)	(0.355)
ESR(-2)	-0.008	0.317	-0.040	0.177***	0.307
	(0.016)	(0.636)	(0.057)	(0.065)	(0.366)
ESR(-3)	-0.014	0.604	-0.023	0.144**	-0.069
	(0.015)	(0.620)	(0.055)	(0.063)	(0.357)
ESR(-4)	-0.007	-0.131	-0.022	0.065	-0.134
	(0.015)	(0.604)	(0.054)	(0.061)	(0.348)
ESR(-5)	-0.006	0.372	0.029	-0.021	0.033
	(0.015)	(0.598)	(0.053)	(0.061)	(0.345)
ESR(-6)	-0.005	-0.992*	-0.046	-0.152**	-0.297
	(0.014)	(0.579)	(0.052)	(0.059)	(0.333)
ESIPI(-1)	0.001	0.142	0.020**	-0.008	-0.504***
	(0.003)	(0.107)	(0.010)	(0.011)	(0.061)
ESIPI(-2)	0.004	0.130	0.009	0.014	-0.075
	(0.003)	(0.118)	(0.011)	(0.012)	(0.068)
ESIPI(-3)	0.011***	0.127	0.014	0.023*	0.150**
	(0.003)	(0.118)	(0.010)	(0.012)	(0.068)
ESIPI(-4)	0.009***	0.114	0.014	-0.018	0.148**
	(0.003)	(0.122)	(0.010)	(0.012)	(0.070)
ESIPI(-5)	0.001	0.076	0.008	-0.003	0.218***
	(0.003)	(0.122)	(0.011)	(0.012)	(0.070)
ESIPI(-6)	0.003	0.016	-0.001	-0.007	0.145**

		(0.003)	(0.111)	(0.010)	(0.011)	(0.064)
	C	0.023**	1.260***	0.153***	-0.003	0.243
		(0.010)	(0.381)	(0.034)	(0.039)	(0.220)
UK	Variable	UKRETURN	UKCCI	UKCPI	UKR	UKIPI
	UKRETURN(-1)	-0.052	1.677	0.492	0.301	-0.651
		(0.059)	(3.272)	(0.360)	(0.282)	(1.099)
	UKCCI(-1)	0.001	-0.137**	0.002	0.003	0.010
		(0.001)	(0.059)	(0.006)	(0.005)	(0.020)
	UKCPI(-1)	-0.001	0.421	0.018	0.029	-0.166
		(0.009)	(0.526)	(0.058)	(0.045)	(0.177)
	UKR(-1)	-0.035***	-1.213*	0.167**	0.314***	0.234
		(0.011)	(0.638)	(0.070)	(0.055)	(0.214)
	UKIPI(-1)	0.004	0.213	0.007	-0.008	-0.296***
		(0.003)	(0.164)	(0.018)	(0.014)	(0.055)
	C	0.004	-0.114	0.168***	-0.022	0.088
		(0.003)	(0.189)	(0.021)	(0.016)	(0.063)

Appendix D

Impulse Responses of VAR Models

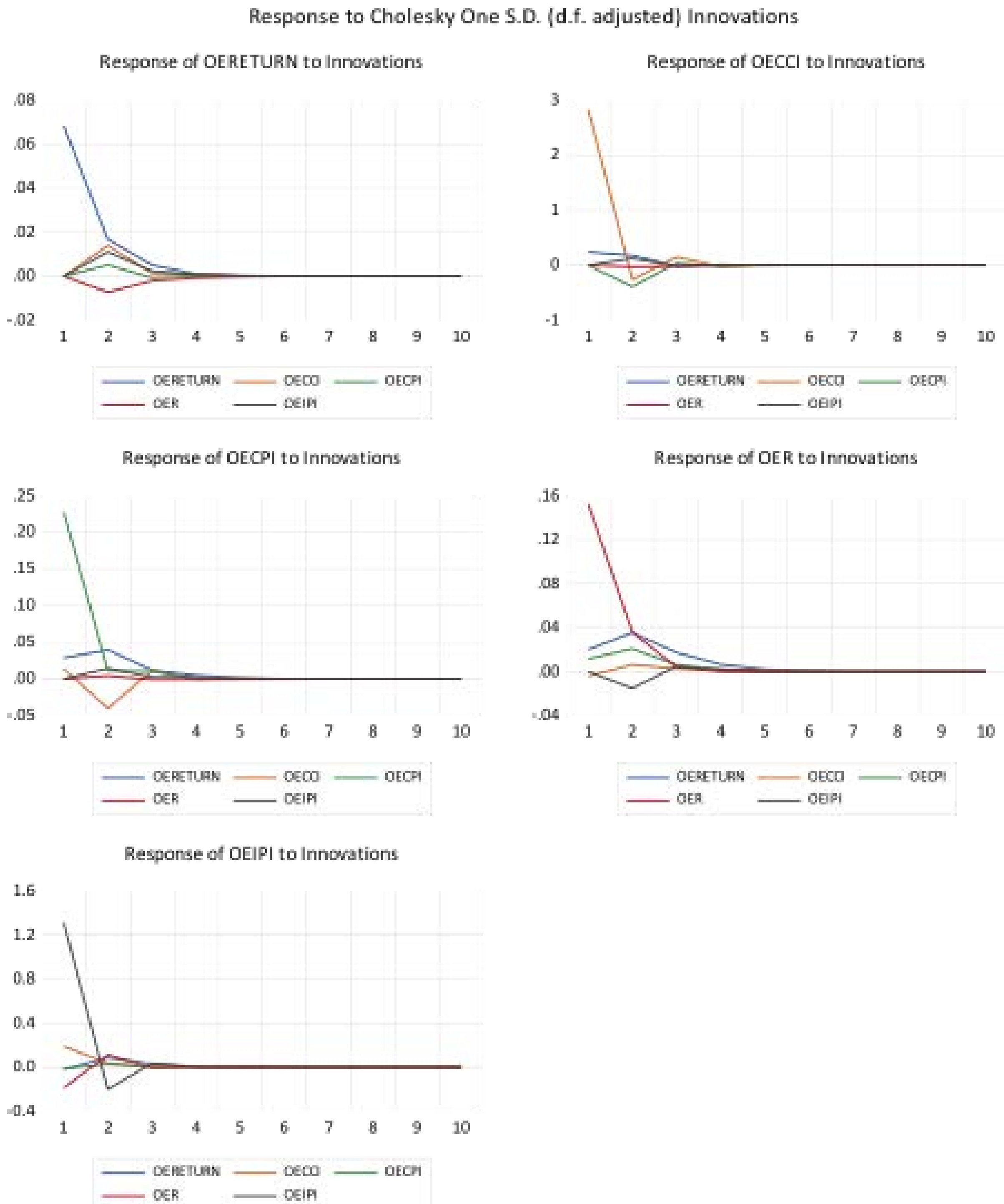

Figure D.1: Impulse Responses of VAR in Austria

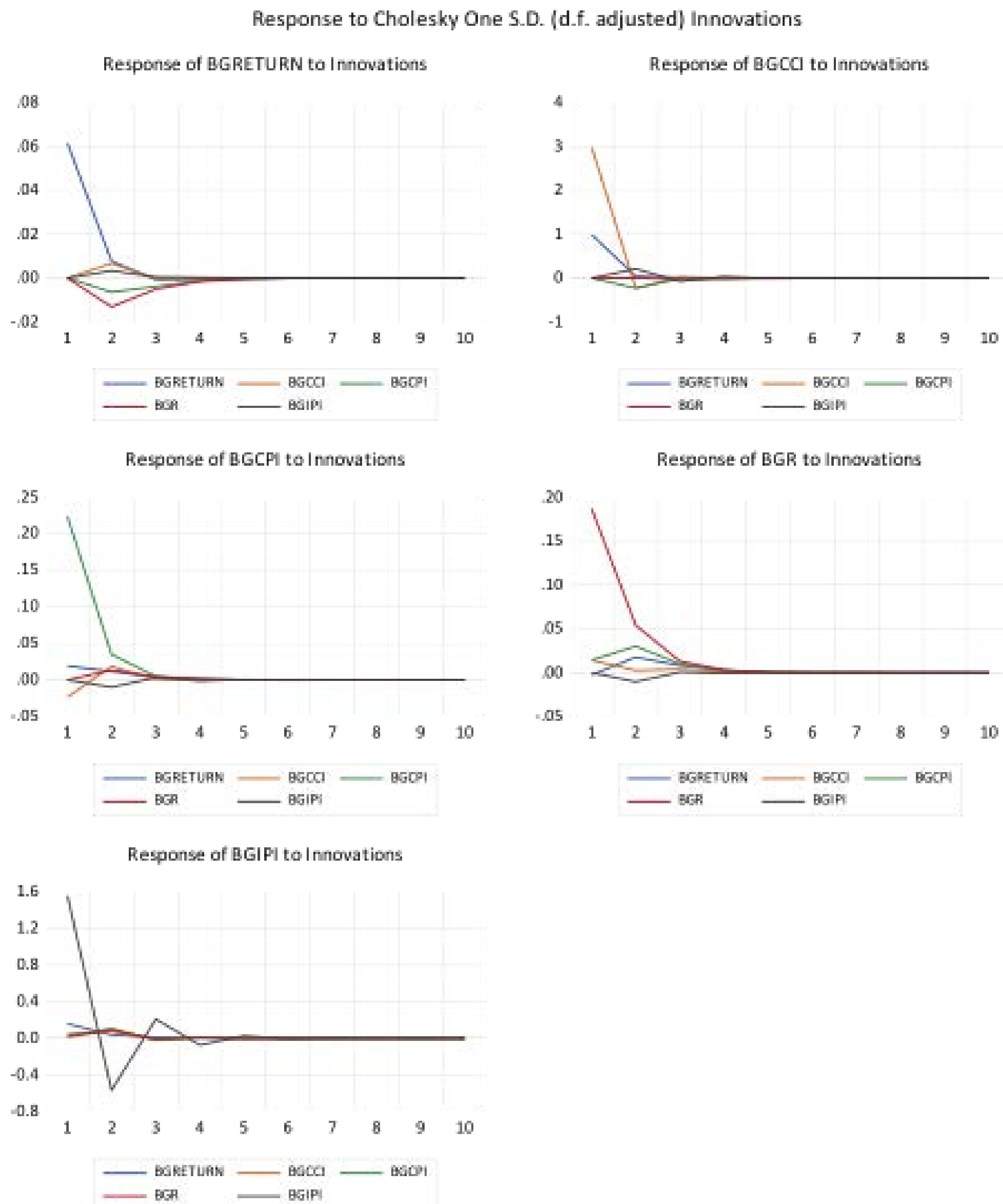

Figure D.2: Impulse Responses of VAR in Belgium

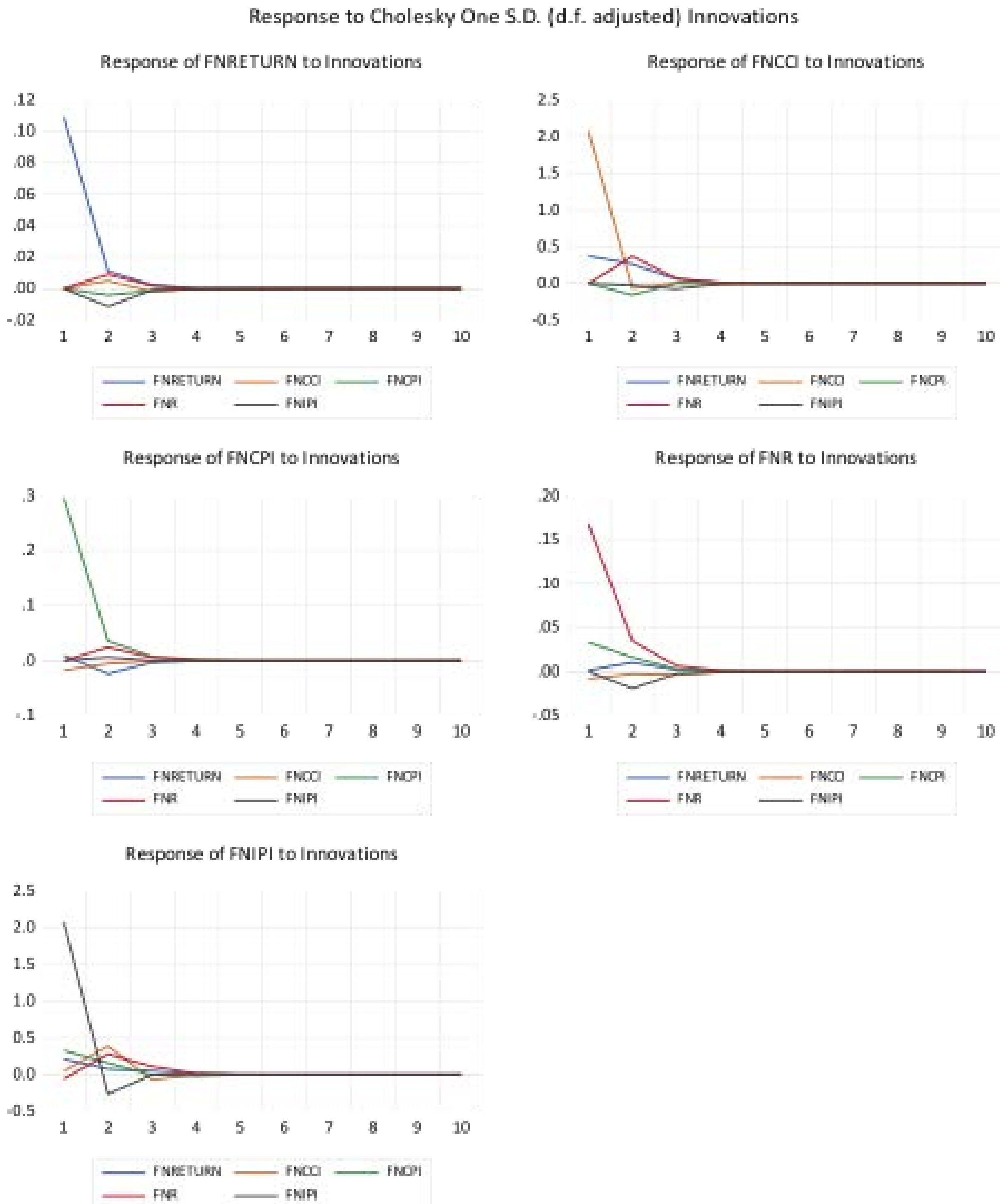

Figure D.3: Impulse Responses of VAR in Finland

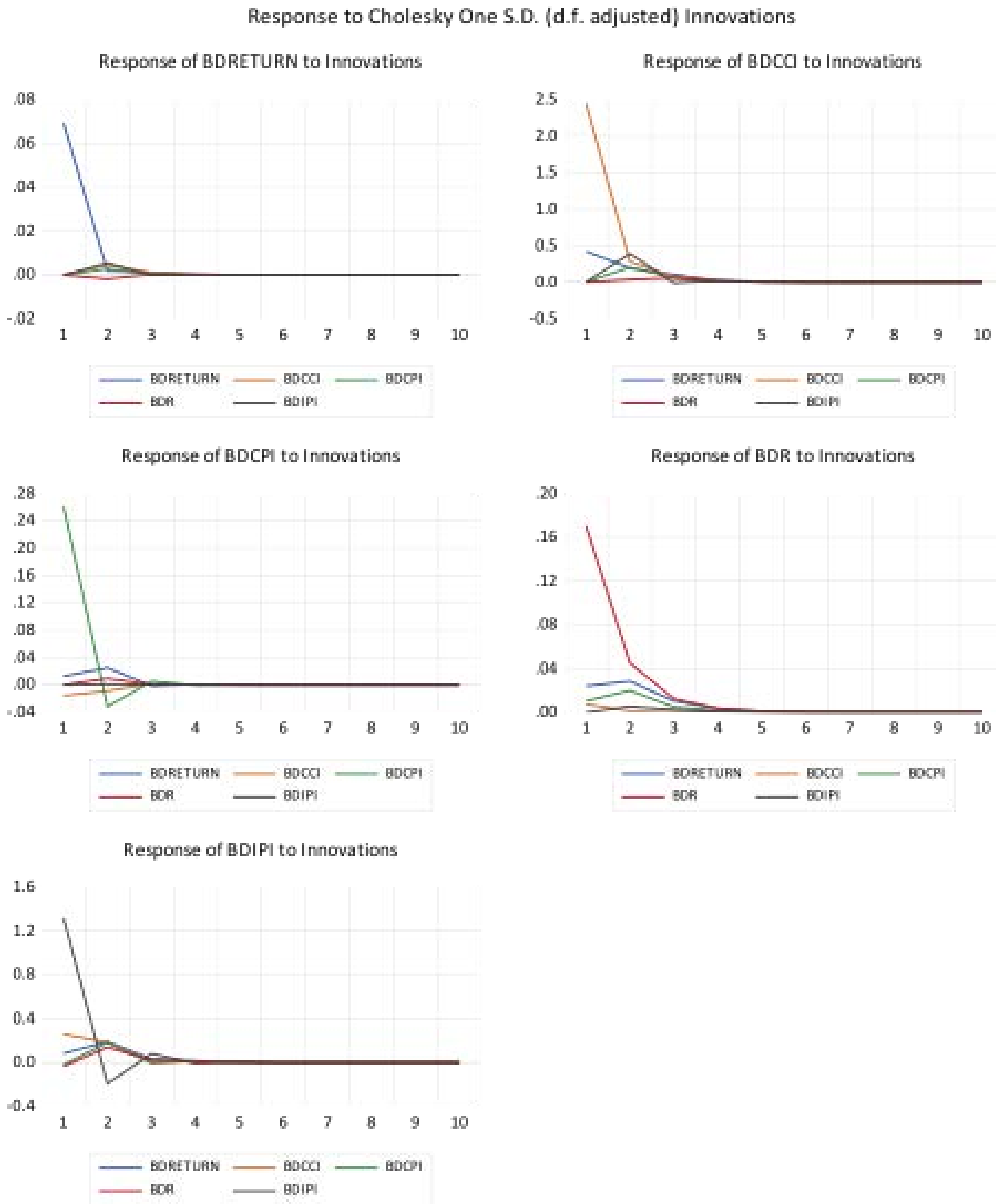

Figure D.4: Impulse Responses of VAR in Germany

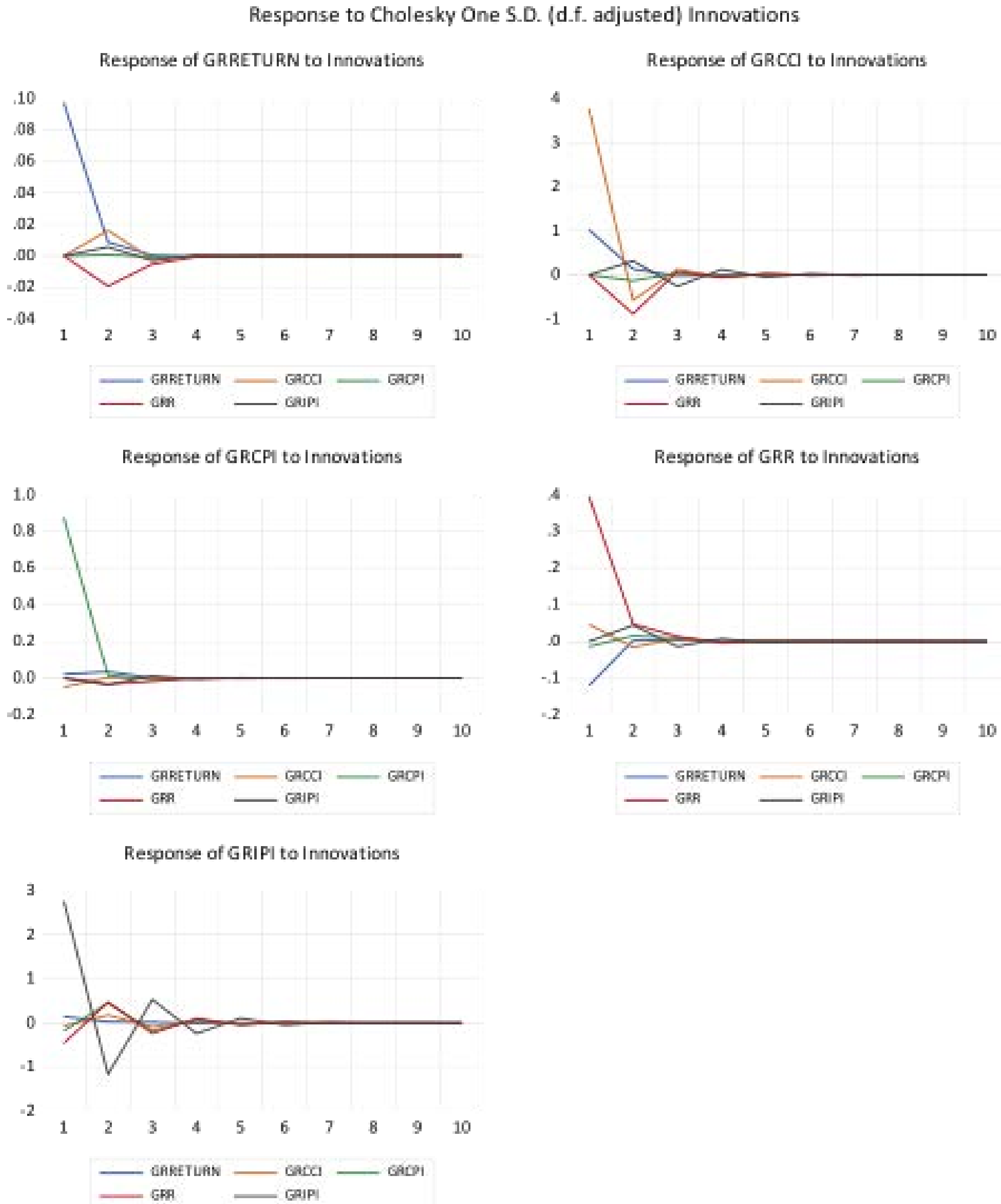

Figure D.5: Impulse Responses of VAR in Greece

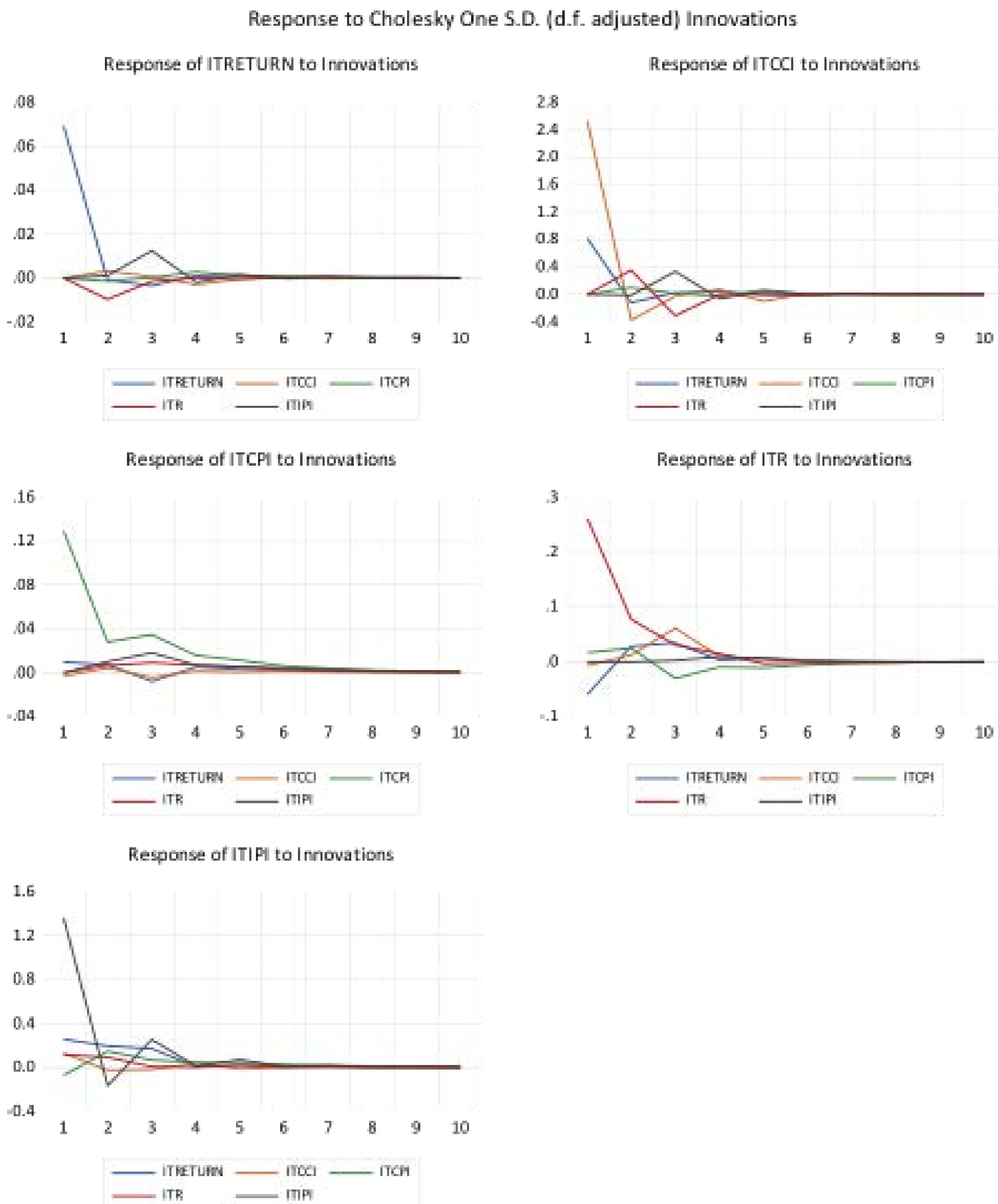

Figure D.6: Impulse Responses of VAR in Italy

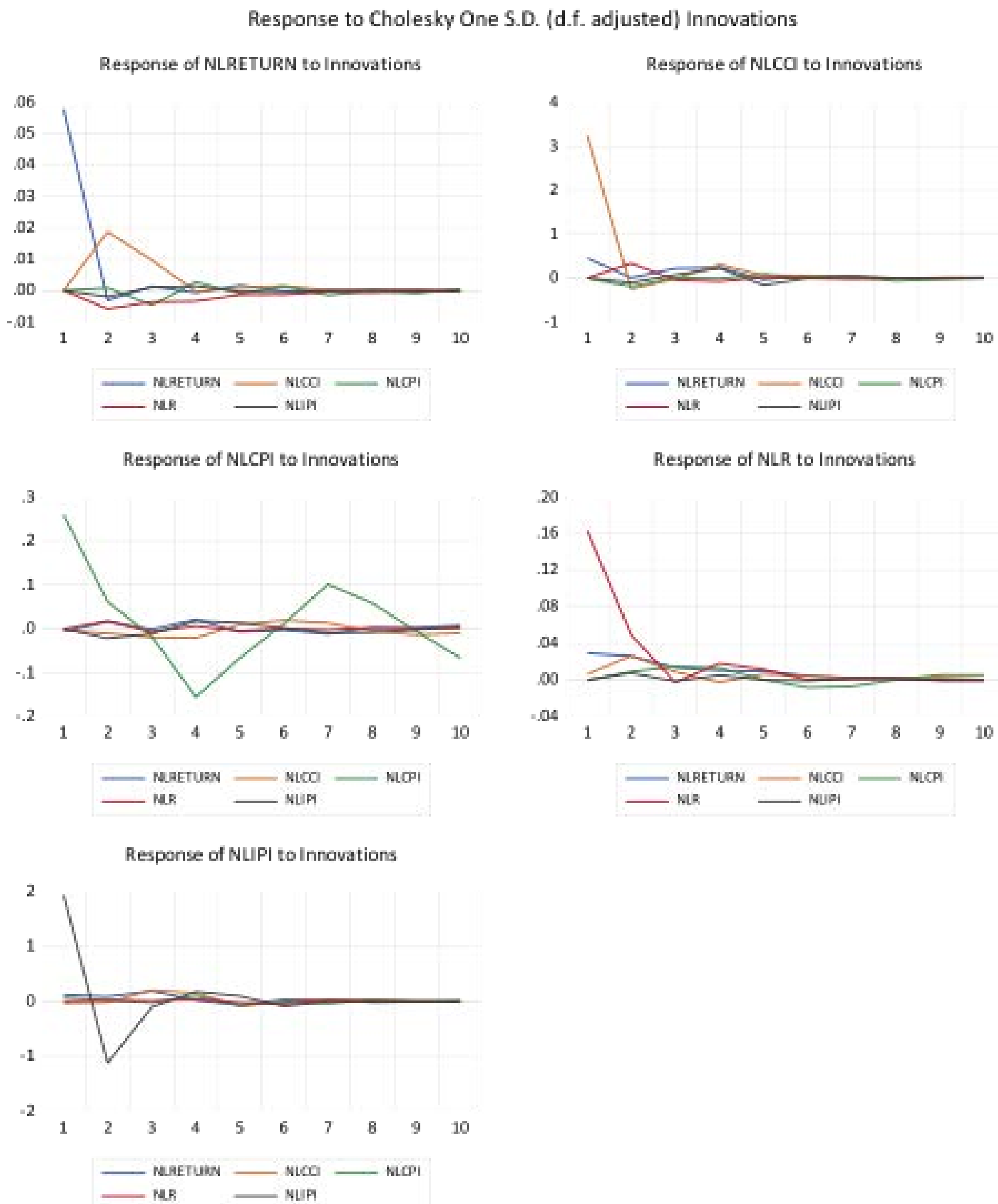

Figure D.7: Impulse Responses of VAR in Netherlands

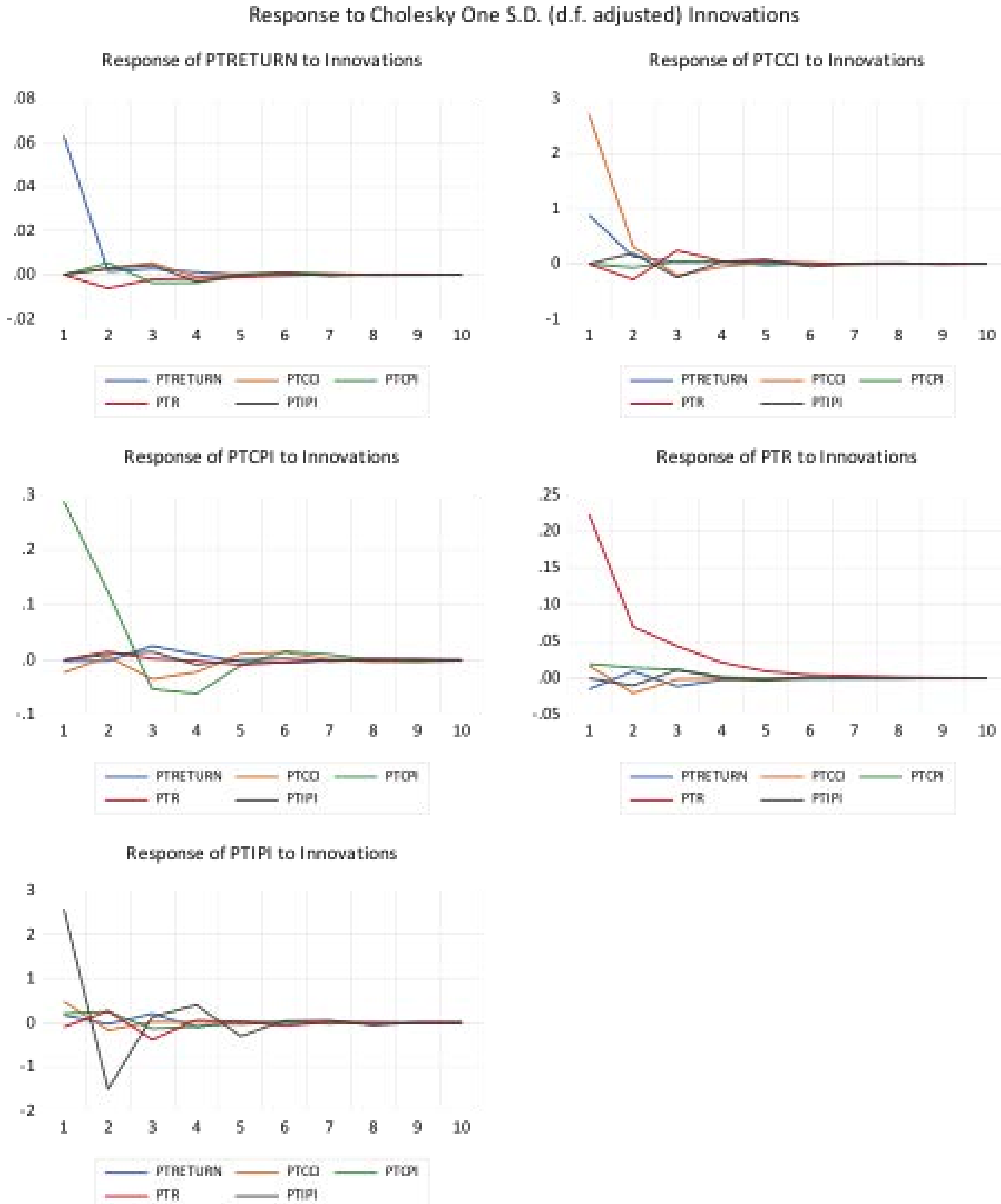

Figure D.8: Impulse Responses of VAR in Potugal

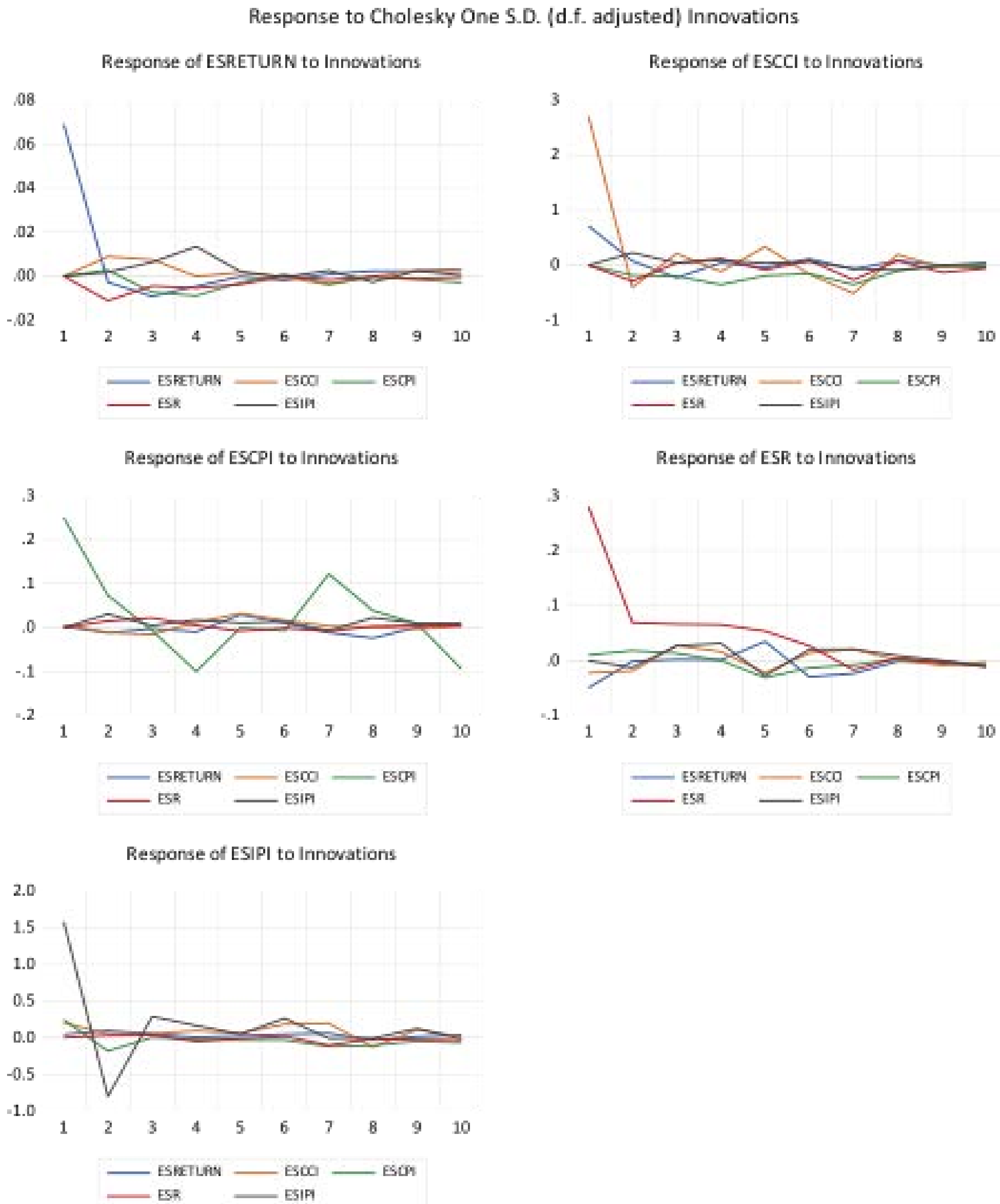

Figure D.9: Impulse Responses of VAR in Spain

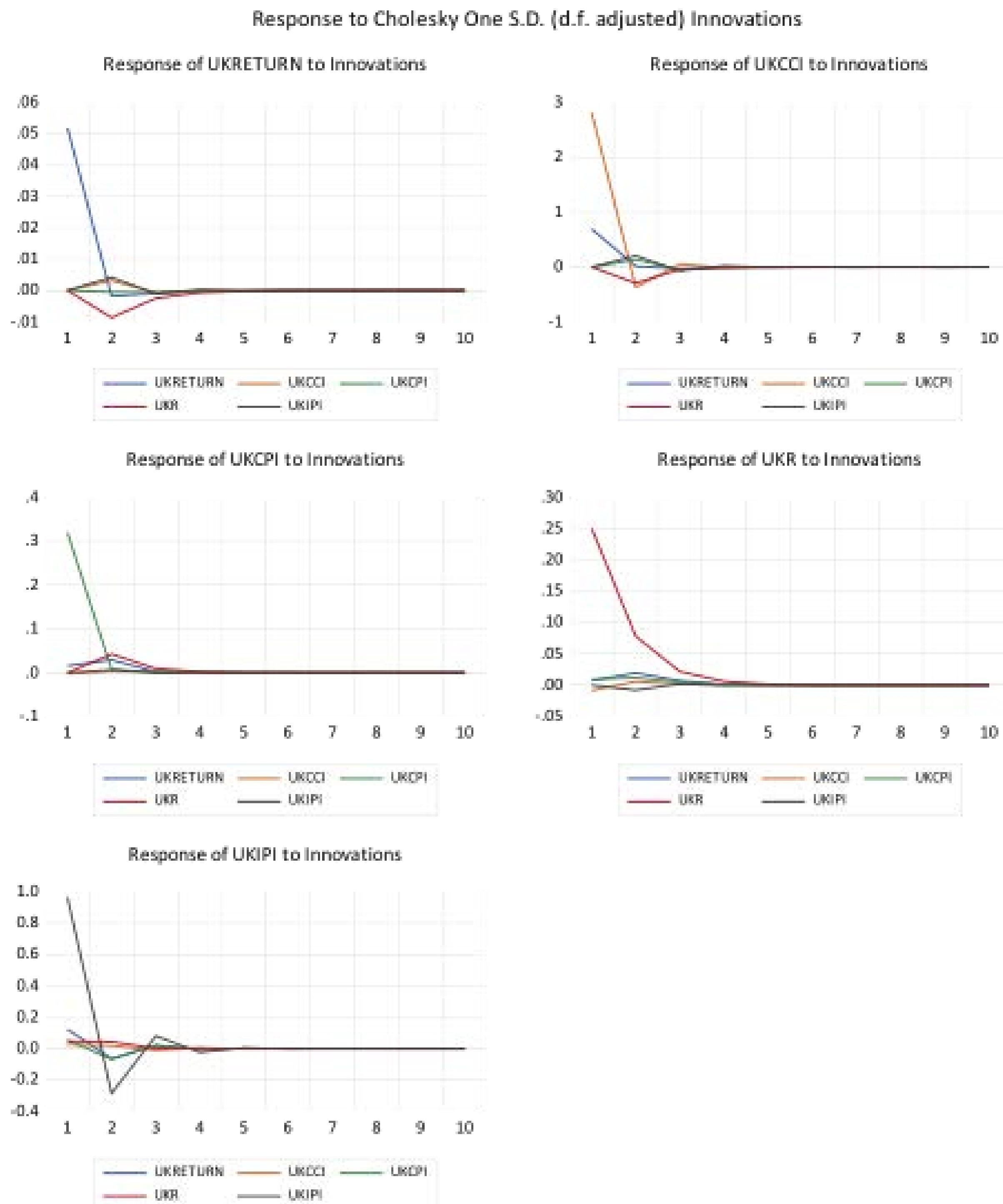

Figure D.10: Impulse Responses of VAR in UK

Appendix E

BVAR In-sample Model Result

Table E.1: BVAR Results in European Countries

OE	Variable	OERETURN	OECCI	OECPI	OER	OEIPI
	OERETURN(-1)	0.507***	1.463	0.282	0.212*	0.759
		(0.060)	(2.390)	(0.198)	(0.133)	(1.121)
	OECCI(-1)	0.003*	0.307***	-0.010**	0.002	0.001
		(0.001)	(0.060)	(0.005)	(0.003)	(0.028)
	OECPI(-1)	0.009	-1.241*	0.403	0.045	0.059
		(0.018)	(0.730)	(0.061)	(0.040)	(0.342)
	OER(-1)	-0.037	0.147	0.012	0.512***	0.527
		(0.027)	(1.079)	(0.089)	(0.060)	(0.506)
	OEIPI(-1)	0.006*	-0.008	0.005	-0.006	0.261***
		(0.003)	(0.127)	(0.011)	(0.007)	(0.060)
	C	-0.002	0.225	0.080***	-0.013	0.186*

		(0.006)	(0.235)	(0.020)	(0.013)	(0.110)
BG	Variable	BGRETURN	BGCCI	BGCPI	BGR	BGIPI
	BGRETURN(-1)	0.338***	-1.771	-0.002	0.250*	0.368
		(0.051)	(2.531)	(0.182)	(0.153)	(1.264)
	BGCCI(-1)	0.001	0.204***	0.007*	0.000	0.024
		(0.001)	(0.052)	(0.004)	(0.003)	(0.026)
	BGCPI(-1)	-0.025*	-0.433	0.364*	0.075*	0.276
		(0.014)	(0.694)	(0.050)	(0.042)	(0.347)
	BGR(-1)	-0.047*	-0.314	0.035	0.458*	0.277
		(0.016)	(0.801)	(0.058)	(0.049)	(0.400)
	BGIPI(-1)	0.001	0.098	-0.006	-0.006	-0.052
		(0.002)	(0.095)	(0.007)	(0.006)	(0.048)
	C	0.005	0.104	0.088*	-0.020*	0.133
		(0.004)	(0.205)	(0.015)	(0.012)	(0.102)
FN	Variable	FNRETURN	FNCCI	FNCPI	FNR	FNIPI
	FNRETURN(-1)	0.437***	0.661	-0.152	0.052	-0.299
		(0.061)	(1.193)	(0.164)	(0.095)	(1.203)
	FNCCI(-1)	-0.001	0.370***	0.001	0.001	0.110*
		(0.003)	(0.062)	(0.008)	(0.005)	(0.062)
	FNCPI(-1)	-0.010	-0.331	0.441***	0.002	0.042
		(0.022)	(0.442)	(0.061)	(0.035)	(0.446)
	FNR(-1)	0.027	1.317*	0.017	0.491***	0.996
		(0.038)	(0.748)	(0.103)	(0.060)	(0.753)
	FNIPI(-1)	-0.005	-0.018	-0.003	-0.006	0.293***

		(0.003)	(0.060)	(0.008)	(0.005)	(0.061)
	C	0.007	0.074	0.071***	-0.009	0.155
		(0.009)	(0.171)	(0.024)	(0.014)	(0.173)
BD	Variable	BDRETURN	BDCCI	BDCPI	BDR	BDIPI
	BDRETURN(-1)	0.269***	0.195	0.257	0.186	1.469
		(0.051)	(1.833)	(0.191)	(0.127)	(1.014)
	BDCCI(-1)	0.000	0.331***	-0.002	-0.001	0.047*
		(0.001)	(0.051)	(0.005)	(0.004)	(0.028)
	BDCPI(-1)	0.006	0.682	0.159***	0.045	0.517*
		(0.013)	(0.479)	(0.050)	(0.033)	(0.265)
	BDR(-1)	-0.016	0.044	0.026	0.446***	0.598
		(0.020)	(0.709)	(0.074)	(0.049)	(0.392)
	BDIPI(-1)	0.003	0.159*	0.002	0.003	0.152***
		(0.003)	(0.091)	(0.009)	(0.006)	(0.050)
	C	0.002	-0.065	0.103***	-0.014	0.039
		(0.004)	(0.157)	(0.016)	(0.011)	(0.087)
GR	Variable	GRRETURN	GRCCI	GRCPI	GRR	GRIPI
	GRRETURN(-1)	0.431***	-2.483	0.102	0.390	0.758
		(0.064)	(2.573)	(0.559)	(0.267)	(1.815)
	GRCCI(-1)	0.001	0.320***	0.005	-0.005	0.047
		(0.002)	(0.063)	(0.014)	(0.007)	(0.045)
	GRCPI(-1)	-0.001	-0.016	0.399***	0.020	0.306
		(0.007)	(0.286)	(0.062)	(0.030)	(0.202)
	GRR(-1)	-0.008	-1.460**	-0.028	0.482***	0.786*

		(0.015)	(0.614)	(0.133)	(0.064)	(0.433)
	GRIPI(-1)	0.002	0.125	-0.005	0.016*	0.063
		(0.002)	(0.083)	(0.018)	(0.009)	(0.059)
	C	-0.001	-0.213	0.147**	0.005	-0.092
		(0.008)	(0.325)	(0.071)	(0.034)	(0.229)
IT	Variable	ITRETURN	ITCCI	ITCPI	ITR	ITIPI
	ITRETURN(-1)	0.309***	-1.913	-0.004	0.562**	1.347
		(0.058)	(2.219)	(0.107)	(0.231)	(1.172)
	ITRETURN(-2)	-0.015	-0.124	-0.062	0.069	0.494
		(0.041)	(1.547)	(0.075)	(0.161)	(0.817)
	ITCCI(-1)	-0.001	0.219***	0.002	0.004	-0.011
		(0.001)	(0.057)	(0.003)	(0.006)	(0.030)
	ITCCI(-2)	0.001	0.005	-0.001	0.008*	-0.003
		(0.001)	(0.040)	(0.002)	(0.004)	(0.021)
	ITCPI(-1)	-0.012	0.473	0.480***	0.078	0.930
		(0.029)	(1.112)	(0.054)	(0.116)	(0.587)
	ITCPI(-2)	0.006	0.015	0.052	-0.123	0.038
		(0.021)	(0.817)	(0.040)	(0.085)	(0.431)
	ITR(-1)	-0.013	0.822	0.016	0.513***	0.165
		(0.014)	(0.532)	(0.026)	(0.056)	(0.281)
	ITR(-2)	0.004	-0.585	0.008	-0.013	-0.053
		(0.010)	(0.384)	(0.019)	(0.040)	(0.203)
	ITIPI(-1)	-0.002	-0.083	0.006	-0.002	0.240***
		(0.003)	(0.105)	(0.005)	(0.011)	(0.056)

	ITIPI(-2)	0.003	0.078	0.003	0.002	0.074*
		(0.002)	(0.075)	(0.004)	(0.008)	(0.040)
	C	0.002	-0.099	0.085***	-0.009	-0.167
	(0.008)	(0.289)	(0.014)	(0.030)	(0.153)	
NL	Variable	NLRETURN	NLCCI	NLCPI	NLR	NLIPI
	NLRETURN(-1)	0.253***	-0.806	0.306	0.099	1.749
		(0.052)	(2.788)	(0.221)	(0.143)	(1.679)
	NLRETURN(-2)	0.018	1.639	-0.074	0.001	1.416
		(0.039)	(2.050)	(0.163)	(0.105)	(1.235)
	NLRETURN(-3)	-0.003	0.744	0.124	0.018	0.066
		(0.029)	(1.546)	(0.123)	(0.079)	(0.931)
	NLCCI(-1)	0.004***	0.202***	-0.001	0.005**	-0.004
		(0.001)	(0.051)	(0.004)	(0.003)	(0.031)
	NLCCI(-2)	0.001	0.001	-0.001	0.001	0.019
		(0.001)	(0.039)	(0.003)	(0.002)	(0.023)
	NLCCI(-3)	0.001	0.022	-0.001	0.001	0.011
		(0.001)	(0.029)	(0.002)	(0.001)	(0.017)
	NLCPI(-1)	0.002	-0.570	0.440	0.029	0.152
		(0.010)	(0.519)	(0.041)	(0.027)	(0.313)
	NLCPI(-2)	-0.005	0.079	-0.175***	0.023	0.004
		(0.008)	(0.428)	(0.034)	(0.022)	(0.258)
	NLCPI(-3)	0.006	-0.070	-0.184***	0.007	0.105
		(0.006)	(0.339)	(0.027)	(0.017)	(0.204)
	NLR(-1)	-0.035*	1.291	-0.010	0.475***	0.279

		(0.018)	(0.979)	(0.078)	(0.050)	(0.589)
	NLR(-2)	-0.004	-0.240	-0.030	-0.055	0.151
		(0.014)	(0.744)	(0.059)	(0.038)	(0.448)
	NLR(-3)	-0.003	-0.048	0.037	0.033	0.043
		(0.010)	(0.559)	(0.044)	(0.029)	(0.337)
	NLIPI(-1)	-0.001	-0.039	-0.008	0.003	-0.118**
		(0.001)	(0.079)	(0.006)	(0.004)	(0.048)
	NLIPI(-2)	0.001	-0.007	-0.004	0.001	-0.063*
		(0.001)	(0.062)	(0.005)	(0.003)	(0.038)
	NLIPI(-3)	0.001	0.036	0.002	0.001	0.005
		(0.001)	(0.047)	(0.004)	(0.002)	(0.028)
	C	0.003	0.085	0.114***	-0.014	0.100
		(0.004)	(0.212)	(0.017)	(0.011)	(0.128)
PT	Variable	PTRETURN	PTCCI	PTCPI	PTR	PTIPI
	PTRETURN(-1)	0.356***	-2.861	-0.143	0.253	0.731
		(0.059)	(2.667)	(0.271)	(0.210)	(2.488)
	PTRETURN(-2)	0.007	0.237	0.135	-0.084	0.893
		(0.041)	(1.841)	(0.187)	(0.145)	(1.717)
	PTCCI(-1)	-0.001	0.414***	0.005	-0.006	-0.029
		(0.001)	(0.060)	(0.006)	(0.005)	(0.055)
	PTCCI(-2)	0.001	-0.033	-0.004	0.001	0.013
		(0.001)	(0.041)	(0.004)	(0.003)	(0.038)
	PTCPI(-1)	0.008	0.049	0.543***	0.016	0.504
		(0.012)	(0.538)	(0.055)	(0.042)	(0.502)

	PTCPI(-2)	-0.009	0.123	-0.146***	-0.002	-0.223
		(0.009)	(0.391)	(0.040)	(0.031)	(0.364)
	PTR(-1)	-0.012	-0.680	0.025	0.544***	0.588
		(0.016)	(0.717)	(0.073)	(0.057)	(0.669)
	PTR(-2)	0.001	0.444	-0.028	0.004	-0.542
		(0.012)	(0.520)	(0.053)	(0.041)	(0.485)
	PTIPI(-1)	0.001	0.045	-0.001	-0.003	-0.049
		(0.001)	(0.058)	(0.006)	(0.005)	(0.054)
	PTIPI(-2)	0.001	-0.045	0.001	0.002	-0.016
		(0.001)	(0.043)	(0.004)	(0.003)	(0.040)
	C	0.001	-0.046	0.111***	-0.010	-0.049
		(0.005)	(0.232)	(0.024)	(0.018)	(0.216)
ES	Variable	ESRETURN	ESCCI	ESCPI	ESR	ESIPI
	ESRETURN(-1)	0.293***	-0.439	0.033	0.157	1.317
		(0.055)	(2.151)	(0.190)	(0.224)	(1.229)
	ESRETURN(-2)	-0.043	-1.111	-0.021	0.009	0.762
		(0.039)	(1.543)	(0.136)	(0.161)	(0.882)
	ESRETURN(-3)	-0.003	0.398	-0.039	0.022	0.206
		(0.030)	(1.155)	(0.102)	(0.120)	(0.660)
	ESRETURN(-4)	0.001	-0.232	0.030	0.077	0.151
		(0.023)	(0.908)	(0.080)	(0.095)	(0.519)
	ESRETURN(-5)	-0.002	0.118	0.020	-0.051	0.126
		(0.019)	(0.745)	(0.066)	(0.078)	(0.425)
	ESRETURN(-6)	0.002	-0.073	-0.008	-0.011	0.065

	(0.016)	(0.629)	(0.055)	(0.066)	(0.359)
ESCCI(-1)	0.001	0.183***	-0.005	-0.001	0.034
	(0.001)	(0.053)	(0.005)	(0.006)	(0.030)
ESCCI(-2)	0.001	0.039	-0.002	0.005	0.013
	(0.001)	(0.039)	(0.003)	(0.004)	(0.022)
ESCCI(-3)	0.001	-0.007	0.001	0.001	0.005
	(0.001)	(0.029)	(0.003)	(0.003)	(0.017)
ESCCI(-4)	0.001	0.016	0.001	-0.001	0.004
	(0.001)	(0.023)	(0.002)	(0.002)	(0.013)
ESCCI(-5)	0.001	-0.002	0.001	0.001	0.005
	(0.001)	(0.019)	(0.002)	(0.002)	(0.011)
ESCCI(-6)	0.001	-0.013	0.001	0.001	0.006
	(0.001)	(0.016)	(0.001)	(0.002)	(0.009)
ESCPI(-1)	0.007	-0.257	0.360***	0.062	-0.208
	(0.012)	(0.484)	(0.043)	(0.050)	(0.276)
ESCPI(-2)	-0.014	-0.229	-0.075***	0.005	0.070
	(0.010)	(0.388)	(0.034)	(0.040)	(0.222)
ESCPI(-3)	-0.003	-0.112	-0.184***	-0.012	0.007
	(0.008)	(0.307)	(0.027)	(0.032)	(0.175)
ESCPI(-4)	0.001	-0.118	0.037	-0.015	-0.013
	(0.006)	(0.255)	(0.023)	(0.027)	(0.146)
ESCPI(-5)	0.001	-0.107	0.004	0.002	-0.010
	(0.005)	(0.209)	(0.019)	(0.022)	(0.119)
ESCPI(-6)	-0.001	-0.110	0.056	0.001	-0.025

	(0.005)	(0.177)	(0.016)	(0.018)	(0.101)
ESR(-1)	-0.019*	-0.390	0.046	0.478***	0.044
	(0.013)	(0.500)	(0.044)	(0.052)	(0.286)
ESR(-2)	0.001	0.270	-0.005	0.033	0.091
	(0.009)	(0.373)	(0.033)	(0.039)	(0.213)
ESR(-3)	-0.002	-0.011	-0.009	0.028	0.015
	(0.007)	(0.278)	(0.025)	(0.029)	(0.159)
ESR(-4)	-0.001	-0.090	-0.002	0.006	-0.026
	(0.006)	(0.220)	(0.019)	(0.023)	(0.126)
ESR(-5)	0.001	0.024	0.006	-0.002	-0.006
	(0.005)	(0.180)	(0.016)	(0.019)	(0.103)
ESR(-6)	0.001	-0.053	-0.001	-0.008	-0.022
	(0.004)	(0.152)	(0.013)	(0.016)	(0.087)
ESIPI(-1)	0.001	0.042	0.012*	-0.009	-0.052
	(0.002)	(0.087)	(0.008)	(0.009)	(0.050)
ESIPI(-2)	0.001	0.007	0.003	0.005	0.054
	(0.002)	(0.067)	(0.006)	(0.007)	(0.039)
ESIPI(-3)	0.002	0.006	0.001	0.006	0.036
	(0.001)	(0.050)	(0.004)	(0.005)	(0.029)
ESIPI(-4)	0.001	0.001	0.002	-0.005	-0.001
	(0.001)	(0.040)	(0.004)	(0.004)	(0.023)
ESIPI(-5)	0.001	0.002	0.001	0.001	0.017
	(0.001)	(0.033)	(0.003)	(0.003)	(0.019)
ESIPI(-6)	0.001	-0.004	0.001	0.001	0.003

		(0.001)	(0.028)	(0.002)	(0.003)	(0.016)
	C	0.006	0.165	0.166***	-0.017	0.074
		(0.006)	(0.241)	(0.021)	(0.025)	(0.138)
UK	Variable	UKRETURN	UKCCI	UKCPI	UKR	UKIPI
	UKRETURN(-1)	0.226***	-1.604	0.332	0.243	-0.858
		(0.051)	(2.810)	(0.310)	(0.241)	(0.938)
	UKCCI(-1)	0.001	0.157***	0.002	0.003	0.007
		(0.001)	(0.051)	(0.006)	(0.004)	(0.017)
	UKCPI(-1)	-0.002	0.327	0.271***	0.015	-0.156
		(0.008)	(0.457)	(0.051)	(0.039)	(0.153)
	UKR(-1)	-0.026***	-0.779	0.110	0.474***	0.121
		(0.010)	(0.562)	(0.062)	(0.048)	(0.188)
	UKIPI(-1)	0.002	0.160	0.003	-0.008	0.005
		(0.003)	(0.144)	(0.016)	(0.012)	(0.048)
	C	0.003	-0.069	0.126***	-0.015	0.071
		(0.003)	(0.185)	(0.020)	(0.016)	(0.062)

Appendix F

Impulse Responses of BVAR Models

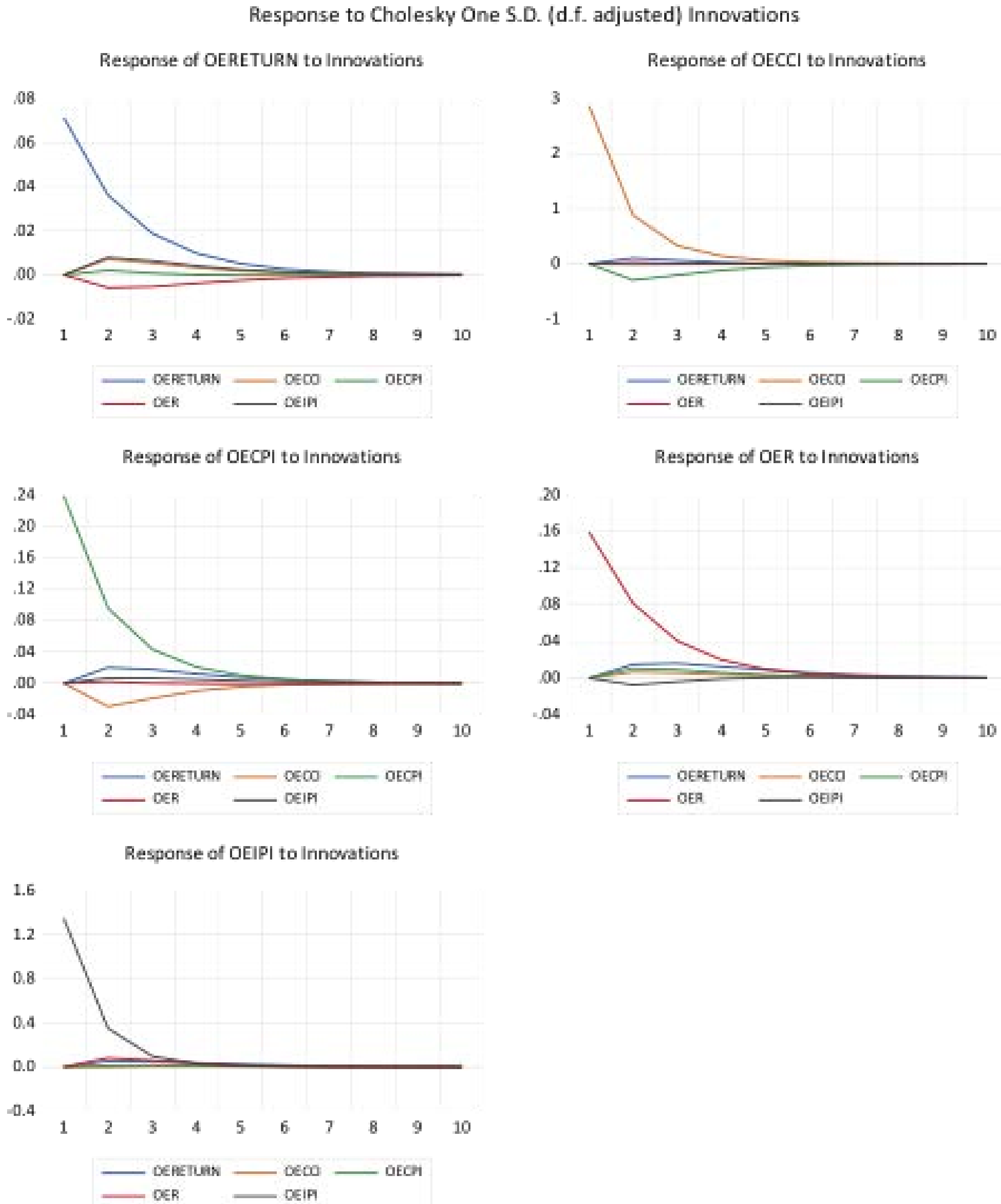

Figure F.1: Impulse Responses of BVAR in Austria

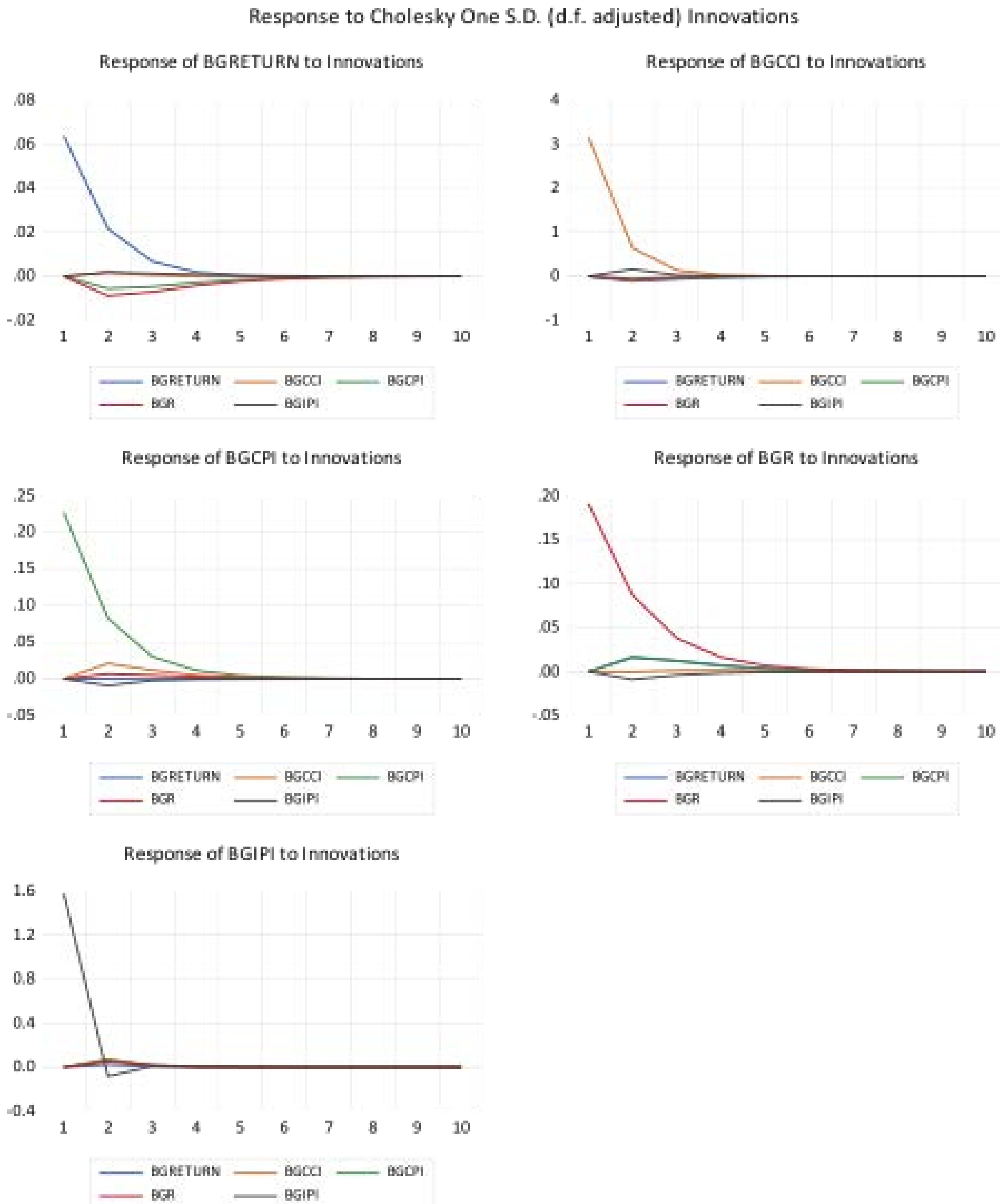

Figure F.2: Impulse Responses of BVAR in Belgium

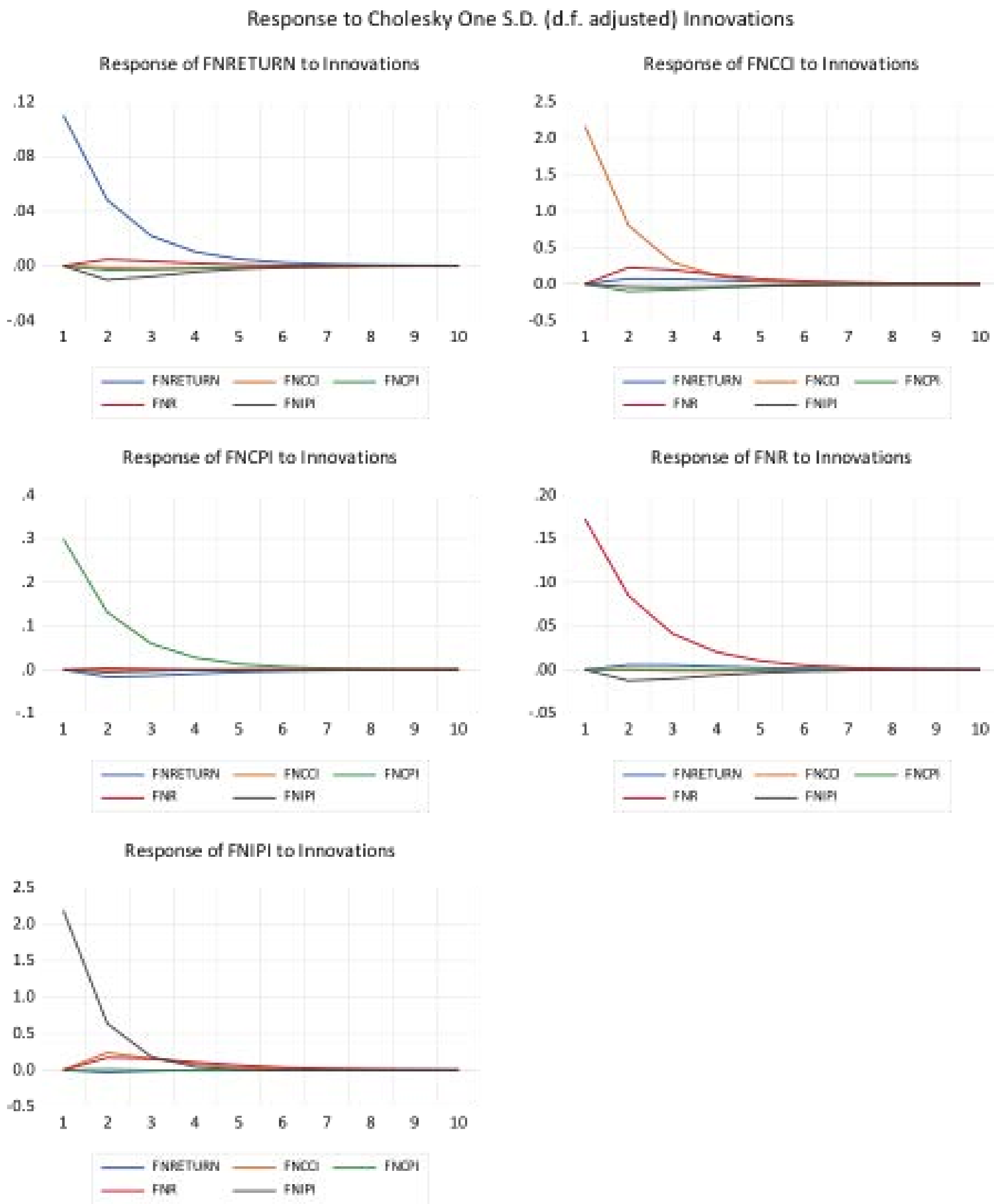

Figure F.3: Impulse Responses of BVAR in Finland

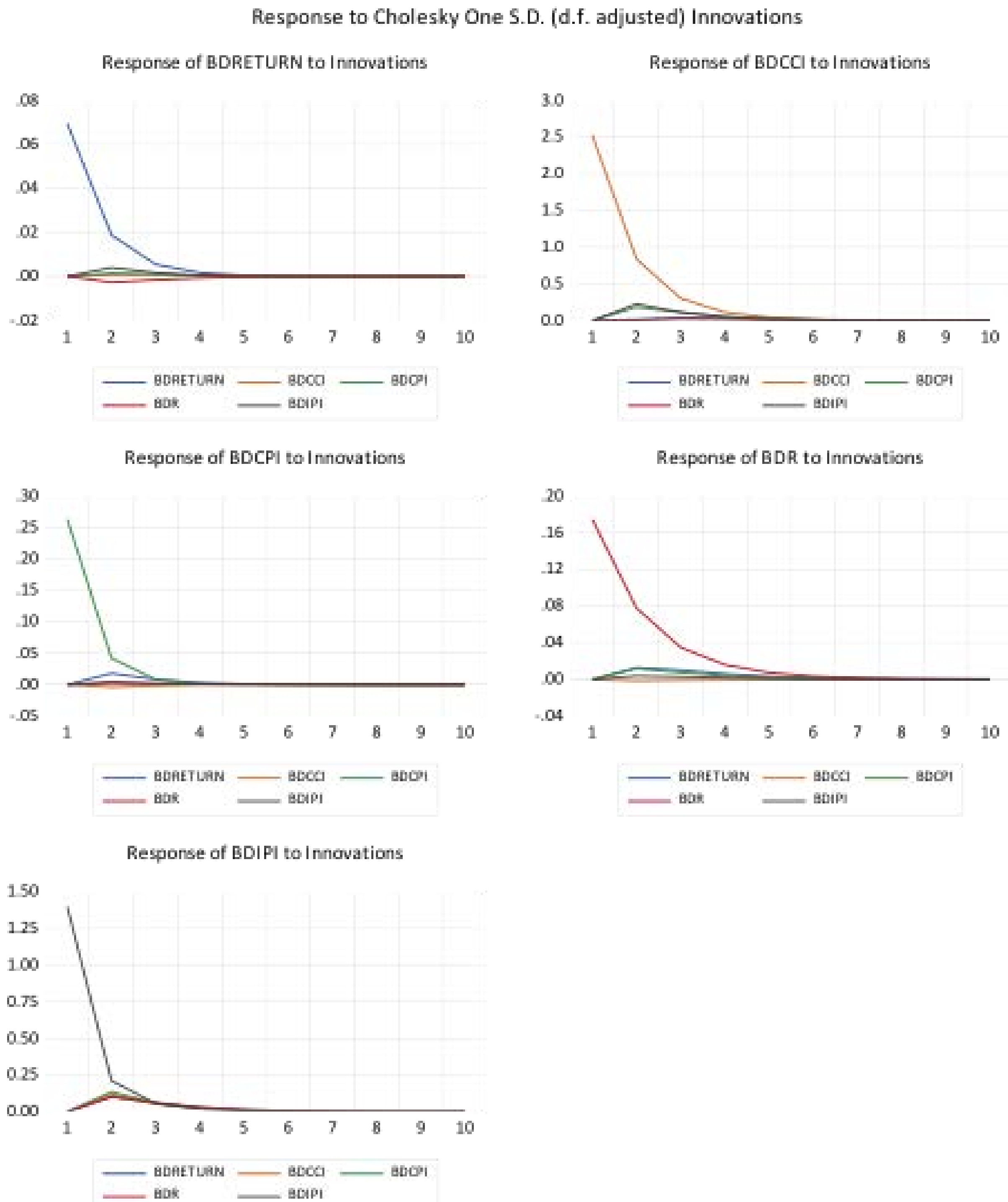

Figure F.4: Impulse Responses of BVAR in Germany

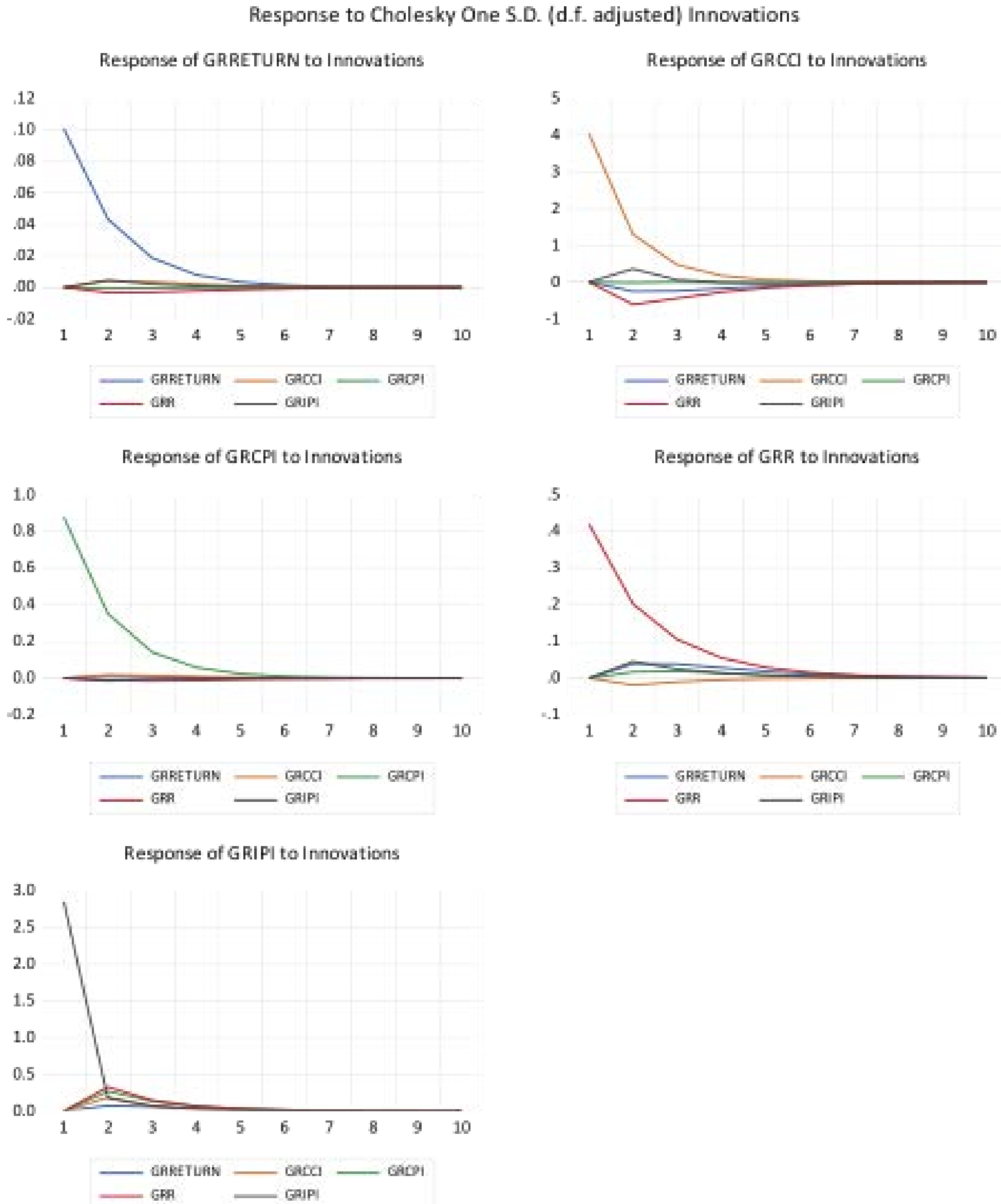

Figure F.5: Impulse Responses of BVAR in Greece

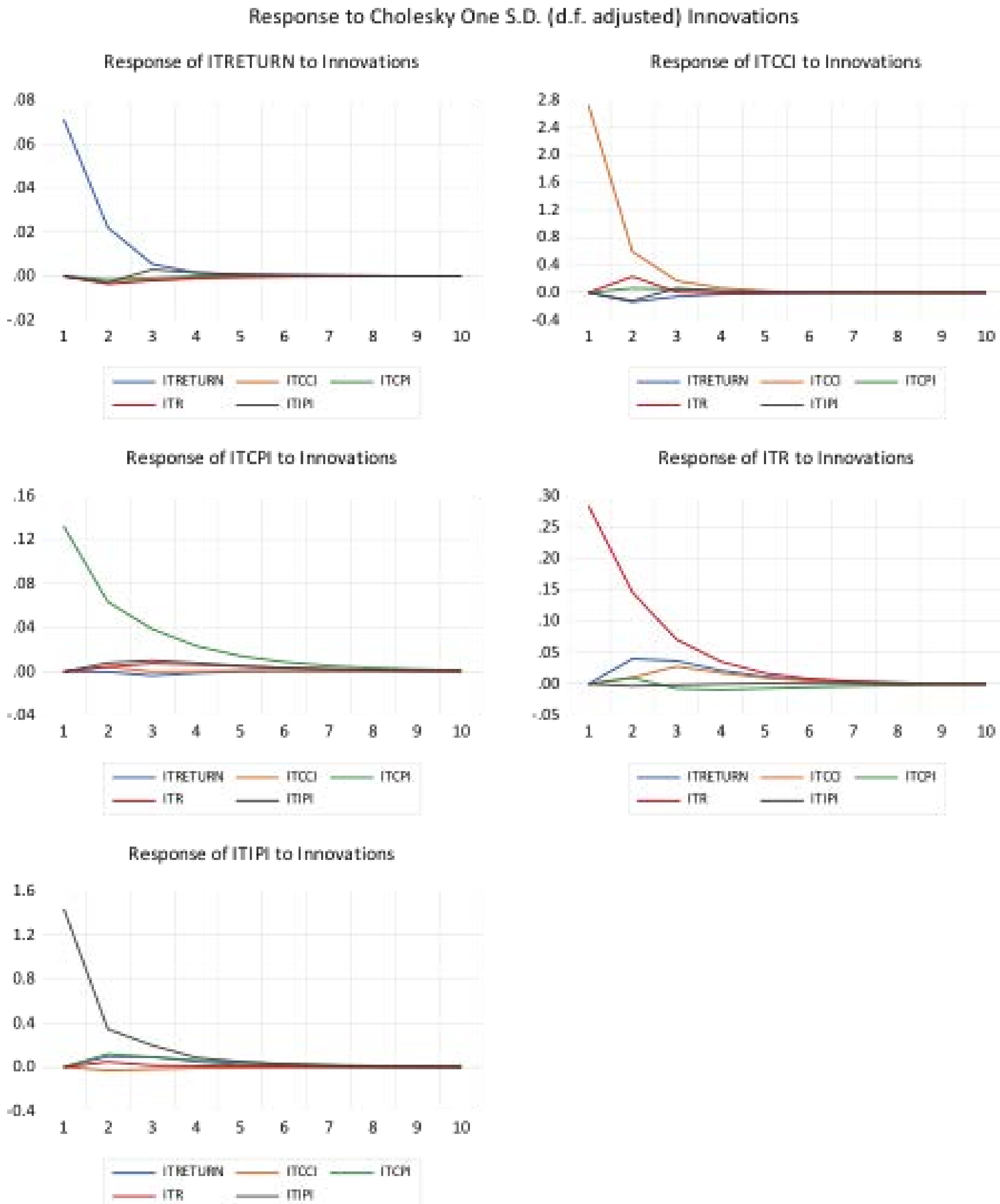

Figure F.6: Impulse Responses of BVAR in Italy

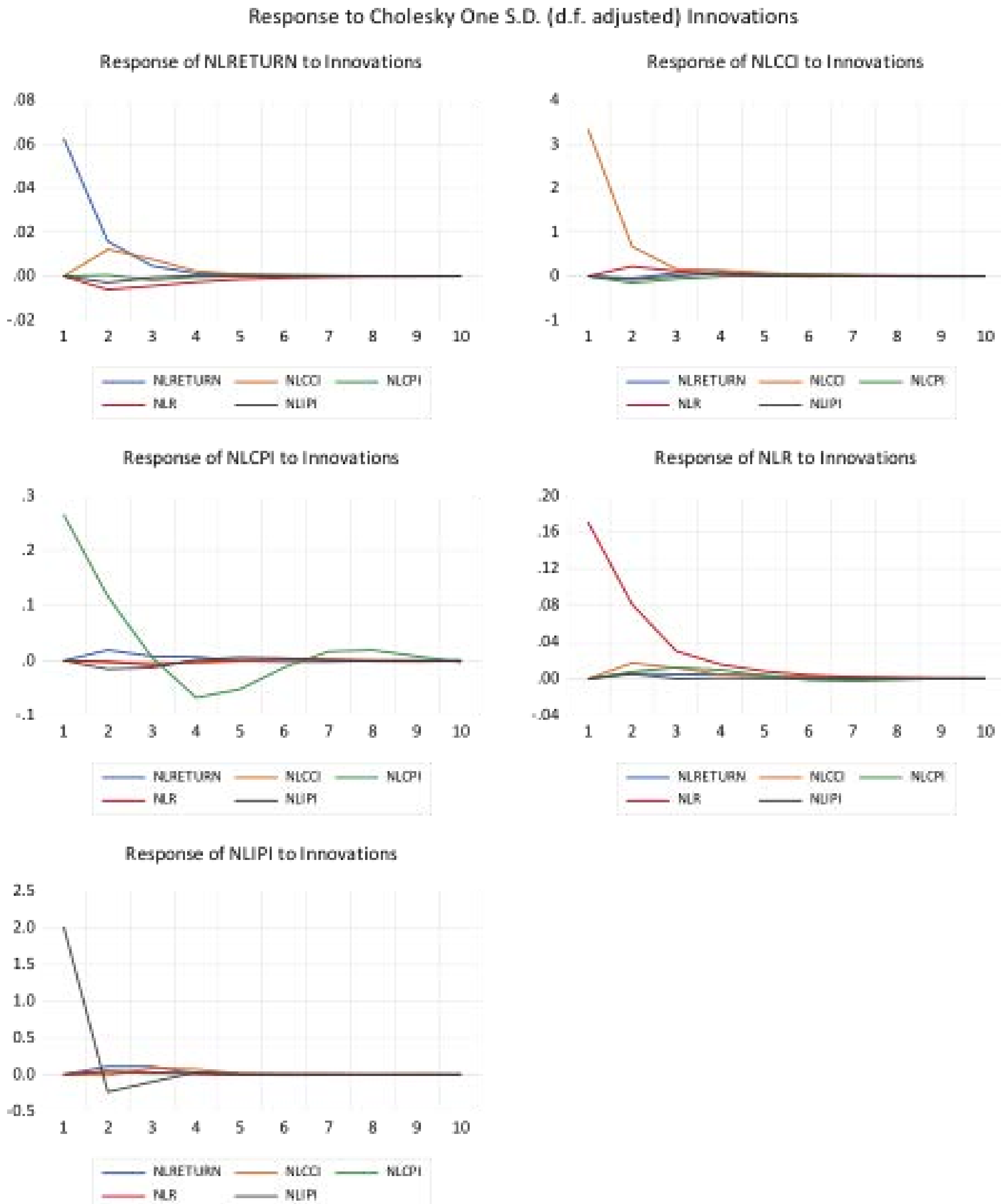

Figure F.7: Impulse Responses of BVAR in Netherlands

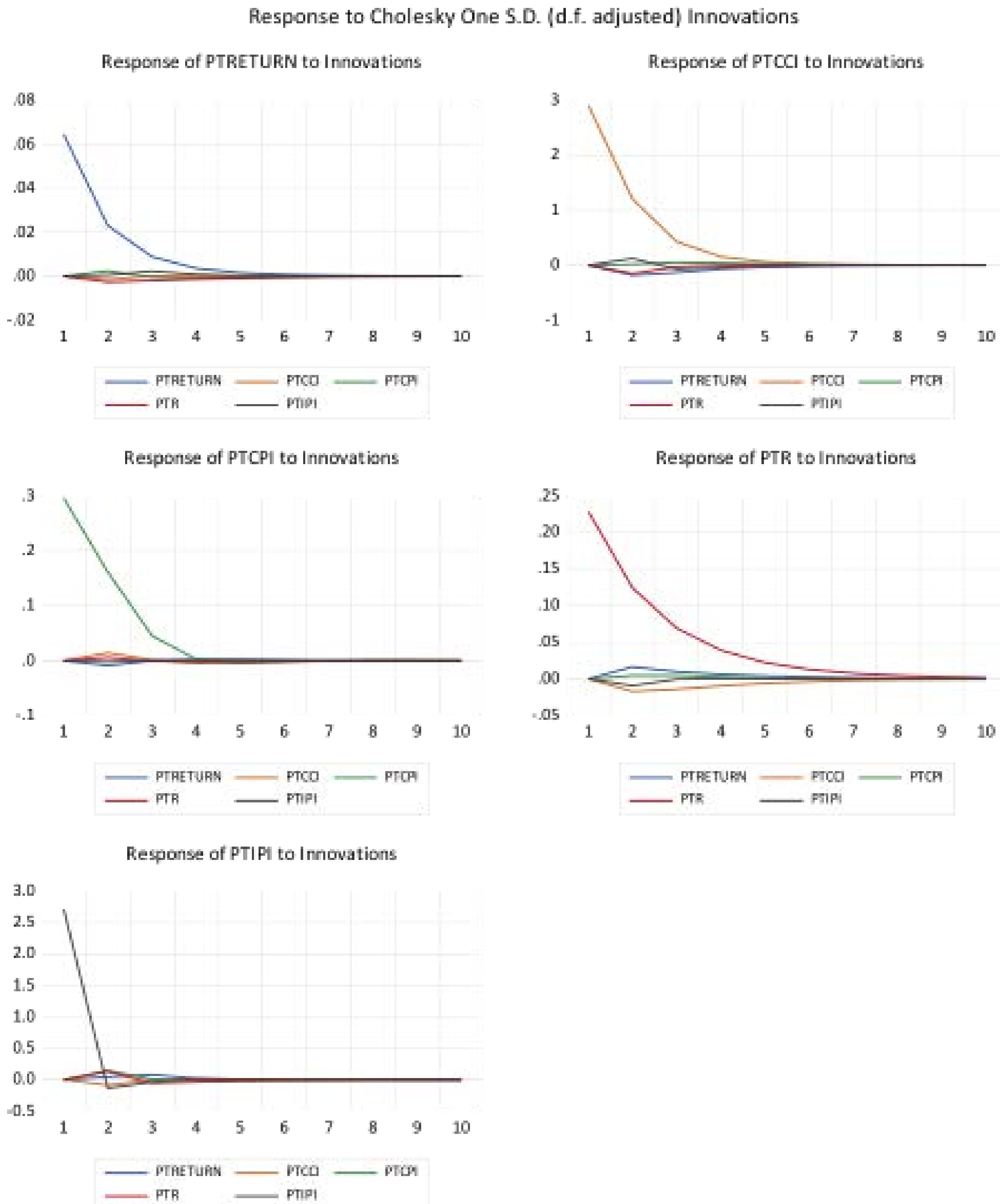

Figure F.8: Impulse Responses of BVAR in Portugal

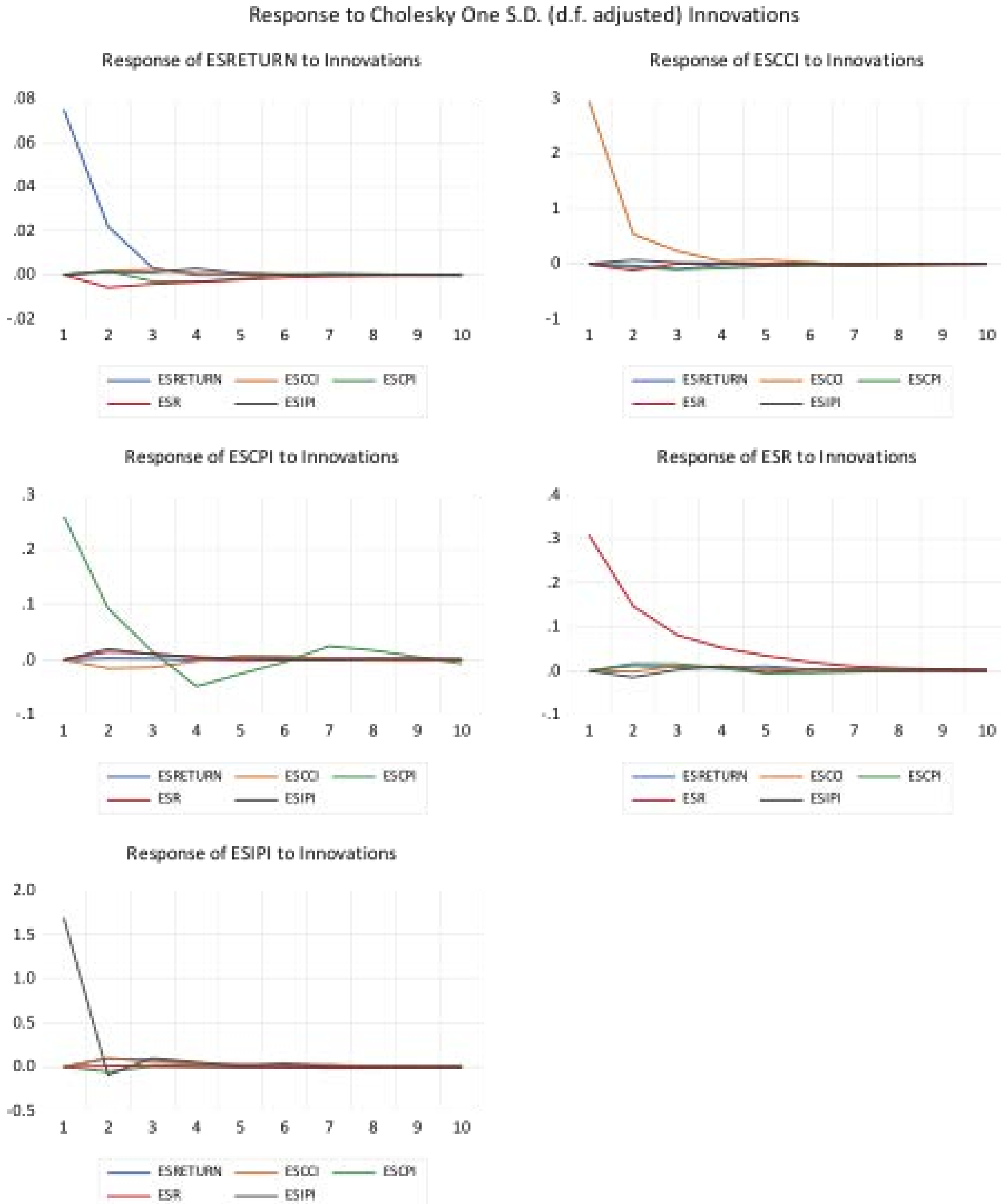

Figure F.9: Impulse Responses of BVAR in Spain

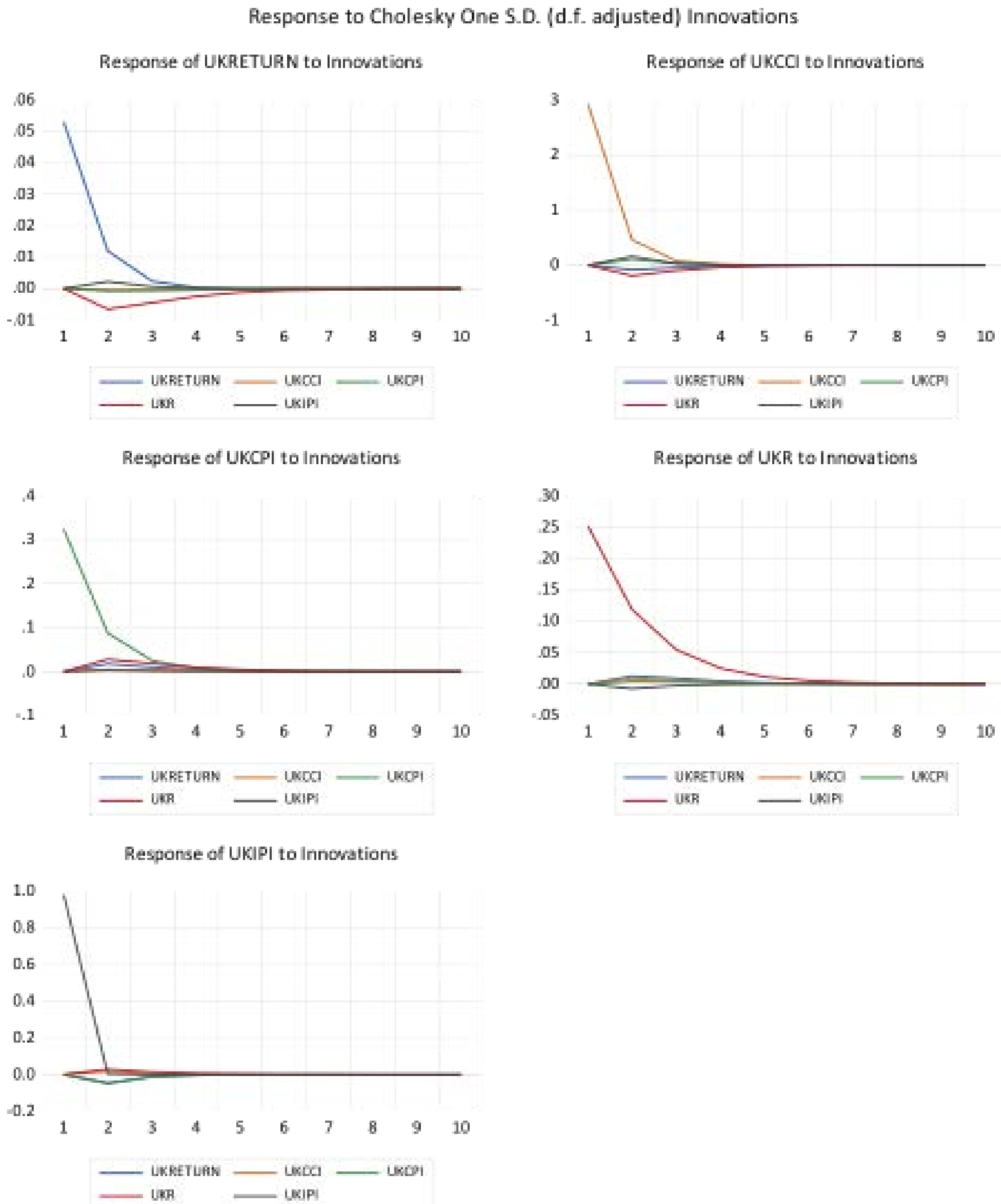

Figure F.10: Impulse Responses of BVAR in UK

List of References

Acquah, H. D.-G. (2010). Comparison of akaike information criterion (aic) and bayesian information criterion (bic) in selection of an asymmetric price relationship. *Journal of Development and Agricultural Economics*, 2(1):001–006.

Akal, M. (2004). Forecasting turkey's tourism revenues by armax model. *Tourism Management*, 25(5):565–580.

Anaghi, M. F. and Norouzi, Y. (2012). A model for stock price forecasting based on arma systems. In *2012 2nd International Conference on Advances in Computational Tools for Engineering Applications (ACTEA)*, pages 265–268. IEEE.

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

Avramov, D. (2002). Stock return predictability and model uncertainty. *Journal of Financial Economics*, 64(3):423–458.

Baker, M. and Stein, J. C. (2004). Market liquidity as a sentiment indicator. *Journal of Financial Markets*, 7(3):271–299.

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

Baker, M. and Wurgler, J. (2004). A catering theory of dividends. *The Journal of Finance*, 59(3):1125–1165.

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

Baker, M. and Wurgler, J. (2007). Investor sentiment in the stock market. *The Journal of Economic Perspectives*, 21(2):129–151.

Baker, M., Wurgler, J., and Yuan, Y. (2012). Global, local, and contagious investor sentiment. *Journal of Financial Economics*, 104(2):272–287.

Barassi, M. R. and Zhao, Y. (2017). Combination forecasting of energy demand in the uk. *Available at SSRN 2907296*.

Basu, S. (1977). Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis. *The journal of Finance*, 32(3):663–682.

Bates, J. M. and Granger, C. W. (1969). The combination of forecasts. *Journal of the Operational Research Society*, 20(4):451–468.

Ben-Rephael, A., Kandel, S., and Wohl, A. (2012). Measuring investor sentiment with mutual fund flows. *Journal of Financial Economics*, 104(2):363–382.

Binswanger, M. (2004). How important are fundamentals?—evidence from a structural var model for the stock markets in the us, japan and europe. *Journal of International Financial Markets, Institutions and Money*, 14(2):185–201.

Borgonovo, E. and Plischke, E. (2016). Sensitivity analysis: a review of recent advances. *European Journal of Operational Research*, 248(3):869–887.

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

Box, G. E. and Jenkins, G. M. (1968). Some recent advances in forecasting and control. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 17(2):91–109.

Box, G. E., Jenkins, G. M., and Reinsel, G. (1970). Time series analysis: forecasting and control holden-day san francisco. *BoxTime Series Analysis: Forecasting and Control Holden Day1970*.

Brown, G. W. and Cliff, M. T. (2004). Investor sentiment and the near-term stock market. *Journal of Empirical Finance*, 11(1):1–27.

Brown, G. W. and Cliff, M. T. (2005). Investor sentiment and asset valuation. *The Journal of Business*, 78(2):405–440.

Burnham, K. P. and Anderson, D. R. (2002). Model selection and.

Cade, B. S. (2015). Model averaging and muddled multimodel inferences. *Ecology*, 96(9):2370–2382.

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

Campbell, J. Y. and Kyle, A. S. (1993). Smart money, noise trading and stock price behaviour. *The Review of Economic Studies*, 60(1):1–34.

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

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

Carriero, A., Kapetanios, G., and Marcellino, M. (2009). Forecasting exchange rates with a large bayesian var. *International Journal of Forecasting*, 25(2):400–417.

Chang, C., Faff, R., and Hwang, C.-Y. (2012). Local and global sentiment effects, and the role of legal, information and trading environments. *February*, 28:2012.

Charoenrook, A. (2005). Does sentiment matter. *Unpublished working paper. Vanderbilt University.*

Chen, N.-F., Kan, R., and Miller, M. H. (1993). Are the discounts on closed-end funds a sentiment index? *The Journal of Finance*, 48(2):795–800.

Choi, I. (2001). Unit root tests for panel data. *Journal of international money and Finance*, 20(2):249–272.

Claeskens, G., Hjort, N. L., et al. (2008). Model selection and model averaging. *Cambridge Books*.

Cochrane, J. H. (2008). State-space vs. var models for stock returns. *Unpublished paper, Chicago GSB.*

De Gooijer, J. G. (1989). Testing non-linearities in world stock market prices. *Economics Letters*, 31(1):31–35.

De Long, J. B., Shleifer, A., Summers, L. H., and Waldmann, R. J. (1990). Noise trader risk in financial markets. *Journal of political Economy*, 98(4):703–738.

Ding, F. and Chen, T. (2005). Identification of hammerstein nonlinear armax systems. *Automatica*, 41(9):1479–1489.

Doukas, J. A. and Milonas, N. T. (2004). Investor sentiment and the closed-end fund puzzle: Out-of-sample evidence. *European Financial Management*, 10(2):235–266.

Dziak, J. J., Coffman, D. L., Lanza, S. T., Li, R., and Jermiin, L. S. (2020). Sensitivity and specificity of information criteria. *Briefings in bioinformatics*, 21(2):553–565.

Elton, E. J., Gruber, M. J., and Busse, J. A. (1998). Do investors care about sentiment? *The Journal of Business*, 71(4):477–500.

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

Fama, E. F. and French, K. R. (1992). The cross-section of expected stock returns. *the Journal of Finance*, 47(2):427–465.

Fama, E. F. and French, K. R. (1998). Value versus growth: The international evidence. *The journal of finance*, 53(6):1975–1999.

Fama, E. F. and French, K. R. (2001). Disappearing dividends: changing firm characteristics or lower propensity to pay? *Journal of Financial economics*, 60(1):3–43.

Fisher, K. L. and Statman, M. (2000). Investor sentiment and stock returns. *Financial Analysts Journal*, 56(2):16–23.

Fisher, K. L. and Statman, M. (2003). Consumer confidence and stock returns. *The Journal of Portfolio Management*, 30(1):115–127.

Gencay, R. (1999). Linear, non-linear and essential foreign exchange rate prediction with simple technical trading rules. *Journal of International Economics*, 47(1):91–107.

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

Granger, C. W. and Ramanathan, R. (1984). Improved methods of combining forecasts. *Journal of forecasting*, 3(2):197–204.

Hannan, E. J., Dunsmuir, W. T., and Deistler, M. (1980). Estimation of vector armax models. *Journal of Multivariate Analysis*, 10(3):275–295.

Henry, O. T. (2002). Long memory in stock returns: some international evidence. *Applied Financial Economics*, 12(10):725–729.

Hoeting, J. A., Madigan, D., Raftery, A. E., and Volinsky, C. T. (1999). Bayesian model averaging: a tutorial. *Statistical science*, pages 382–401.

Huang, D., Jiang, F., Tu, J., and Zhou, G. (2015). Investor sentiment aligned: A powerful predictor of stock returns. *The Review of Financial Studies*, 28(3):791–837.

Iman, R. L., Johnson, M. E., and Watson Jr, C. C. (2005). Uncertainty analysis for computer model projections of hurricane losses. *Risk Analysis: An International Journal*, 25(5):1299–1312.

Johansen, T. A. and FOSS, B. (1993). Constructing narmax models using armax models. *International journal of control*, 58(5):1125–1153.

Jones, C. M. (2002). A century of stock market liquidity and trading costs.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica (pre-1986)*, 47(2):263.

Kambouroudis, D. S., McMillan, D. G., and Tsakou, K. (2016). Forecasting stock return volatility: a comparison of garch, implied volatility, and realized volatility models. *Journal of Futures Markets*, 36(12):1127–1163.

Karanasos, M. (2001). Prediction in arma models with garch in mean effects. *Journal of Time Series Analysis*, 22(5):555–576.

Keynes, J. M. (1936). The general theory of employment, investment, and money. *London and New York, quoted from: https://sites. google. com/site/biblioeconomicus/KeynesJohnMaynard-TheGeneralTheoryOf EmploymentInterestAndMoney. pdf [Accessed 07 Jan 2014].*

Kong, A., Rapach, D. E., Strauss, J. K., and Zhou, G. (2011). Predicting market components out of sample: asset allocation implications. *The Journal of Portfolio Management*, 37(4):29–41.

Koop, G. and Korobilis, D. (2012). Forecasting inflation using dynamic model averaging. *International Economic Review*, 53(3):867–886.

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

Lee, C., Shleifer, A., and Thaler, R. H. (1991). Investor sentiment and the closed-end fund puzzle. *The Journal of Finance*, 46(1):75–109.

Lemmon, M. and Portniaguina, E. (2006). Consumer confidence and asset prices: Some empirical evidence. *The Review of Financial Studies*, 19(4):1499–1529.

Levin, A., Lin, C.-F., and Chu, C.-S. J. (2002). Unit root tests in panel data: asymptotic and finite-sample properties. *Journal of econometrics*, 108(1):1–24.

Liao, J., Zong, X., Zhang, X., and Zou, G. (2019). Model averaging based on leave-subject-out cross-validation for vector autoregressions. *Journal of Econometrics*, 209(1):35–60.

Lim, C., McAleer, M., and Min, J. C. (2009). Armax modelling of international tourism demand. *Mathematics and Computers in Simulation*, 79(9):2879–2888.

Ljungqvist, A., Nanda, V., and Singh, R. (2006). Hot markets, investor sentiment, and ipo pricing. *The Journal of Business*, 79(4):1667–1702.

Malkiel, B. G. (1989). Efficient market hypothesis. In *Finance*, pages 127–134. Springer.

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

Molodtsova, T. and Papell, D. H. (2009). Out-of-sample exchange rate predictability with taylor rule fundamentals. *Journal of international economics*, 77(2):167–180.

Montgomery, J. M. and Nyhan, B. (2010). Bayesian model averaging: Theoretical developments and practical applications. *Political Analysis*, pages 245–270.

Neal, R. and Wheatley, S. M. (1998). Do measures of investor sentiment predict returns? *Journal of Financial and Quantitative Analysis*, 33(4):523–547.

Pesaran, M. H. and Timmermann, A. (1992). A simple nonparametric test of predictive performance. *Journal of Business & Economic Statistics*, 10(4):461–465.

Phan, D. H. B., Sharma, S. S., and Narayan, P. K. (2015). Stock return forecasting: some new evidence. *International Review of Financial Analysis*, 40:38–51.

Pontiff, J. and Schall, L. D. (1998). Book-to-market ratios as predictors of market returns. *Journal of Financial Economics*, 49(2):141–160.

Posada, D. and Buckley, T. R. (2004). Model selection and model averaging in phylogenetics: advantages of the aic and bayesian approaches over likelihood ratio tests. In *Syst. Biol.* Citeseer.

Pradhan, R. P., Arvin, M. B., Samadhan, B., Taneja, S., et al. (2013). The impact of stock market development on inflation and economic growth of 16 asian countries: A panel var approach. *Applied Econometrics and International Development*, 13(1):203–220.

Qiu, L. and Welch, I. (2004). Investor sentiment measures. Technical report, National Bureau of Economic Research.

Rabitz, H. (1989). Systems analysis at the molecular scale. *Science*, 246(4927):221–226.

Raftery, A. E., Madigan, D., and Hoeting, J. A. (1997). Bayesian model averaging for linear regression models. *Journal of the American Statistical Association*, 92(437):179–191.

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

Rapach, D. E. and Wohar, M. E. (2006). In-sample vs. out-of-sample tests of stock return predictability in the context of data mining. *Journal of Empirical Finance*, 13(2):231–247.

Rapach, D. E., Wohar, M. E., and Rangvid, J. (2005). Macro variables and international stock return predictability. *International journal of forecasting*, 21(1):137–166.

Ritter, J. R. (1984). The" hot issue" market of 1980. *Journal of Business*, pages 215–240.

Rounaghi, M. M. and Zadeh, F. N. (2016). Investigation of market efficiency and financial stability between s&p 500 and london stock exchange: monthly and yearly forecasting of time series stock returns using arma model. *Physica A: Statistical Mechanics and its Applications*, 456:10–21.

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

Schwarz, G. et al. (1978). Estimating the dimension of a model. *The annals of statistics*, 6(2):461–464.

Shahriari, M., Sichani, O. A., Gharibshah, J., and Jalili, M. (2016). Sign prediction in social networks based on users reputation and optimism. *Social Network Analysis and Mining*, 6(1):91.

Shefrin, H. and Statman, M. (1994). Behavioral capital asset pricing theory. *Journal of financial and Quantitative Analysis*, 29(3):323–349.

Swider, D. J. and Weber, C. (2007). Extended arma models for estimating price developments on day-ahead electricity markets. *Electric Power Systems Research*, 77(5-6):583–593.

Symonds, M. R. and Moussalli, A. (2011). A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using akaike's information criterion. *Behavioral Ecology and Sociobiology*, 65(1):13–21.

Wagenmakers, E.-J. and Farrell, S. (2004). Aic model selection using akaike weights. *Psychonomic bulletin & review*, 11(1):192–196.

Wasserman, L. et al. (2000). Bayesian model selection and model averaging. *Journal of mathematical psychology*, 44(1):92–107.

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

Westerlund, J. and Narayan, P. (2014). Testing for predictability in conditionally heteroskedastic stock returns. *Journal of Financial Econometrics*, 13(2):342–375.

Whittle, P. (1951). *Hypothesis testing in time series analysis*, volume 4. Almqvist & Wiksells boktr.

Zadrozny, P. (1988). Gaussian likelihood of continuous-time armax models when data are stocks and flows at different frequencies. *Econometric Theory*, pages 108–124.

Zhang, D.-y., Song, H.-w., and Chen, P. (2008). Stock market forecasting model based on a hybrid arma and support vector machines. In *2008 International Conference on Management Science and Engineering 15th Annual Conference Proceedings*, pages 1312–1317. IEEE.

Zheng, C. and Zhu, J. (2017). Research on stock price forecast based on gray relational analysis and armax model. In *2017 International Conference on Grey Systems and Intelligent Services (GSIS)*, pages 145–148. IEEE.

Zweig, M. E. (1973). An investor expectations stock price predictive model using closed-end fund premiums. *The Journal of Finance*, 28(1):67–78.