IMPROVING ACCURACY OF INFLATION FORECASTS USING MACROECONOMIC INFORMATION

ARPAD KOCSIS
2064 / i6093399

07/01/2016

Abstract

This thesis examines the effects of macroeconomic factors on inflation level and volatility in the Euro Area to improve the accuracy of inflation forecasts with econometric modelling. Inflation aggregates for the EU as well as inflation levels of selected countries are analysed, and the difference between these inflation estimates and forecasts are documented. The research proposes alternative models depending on the focus and the scope of inflation forecasts. I find that models with a Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) in mean process have better explanatory power for inflation variance compared to the regular GARCH models. The significant coefficients are different in EU countries in comparison to the aggregate EU-wide forecast of inflation. The presence of more pronounced GARCH components in certain countries with more stressed economies indicates that inflation volatility in these countries are likely to occur as a result of the stressed economy. In addition, other economies in the Euro Area are found to exhibit a relatively stable variance of inflation over time. Therefore, when analysing EU inflation one have to take into consideration the large differences on country level and focus on those one by one.

Chapter 1: General Introduction

Chapter One introduces general ideas, such as personal and academic motivation. Based on these personal interests, and academic arguments I phrase several research questions.

Introduction and motivation

Personal Motivation

The intention behind the choice of the topic "Improving accuracy of inflation forecasts using macroeconomic information" has been my personal motivation to contribute to the debate about inflation forecasts and their accuracy. The main point of preference for a contribution to this field is my personal interest in econometric modelling and in using statistical tools to forecast future economic movements. A practical application of this amplitude would give me the benefit of a more detailed knowledge of these tools, which have been initially triggered during one of the quantitative courses I took during my studies.

Econometrics deals with complex multivariate relationships between economic data, and econometric applications are mostly based on non-experimental or "field" data that are influenced by many factors. No matter whether we talk about a simple linear regression model or more advanced regression models, most of these classical econometric analyses use some basic assumptions which may be difficult to validate (McFadden in *Robust Methods in Econometrics*, 1999). Hence, the interpretation of the results, but also of the models, has to take into account these difficulties. This is what strikes me the most in the field of econometrics, and this is why I chose a closely related topic.

Furthermore, there are some special issues to be considered when estimating relationships between time series data. Time itself gives a natural flow of the observations which are absent when one uses cross-sectional dataset, and it ultimately leads to dynamic features of the observations. For example, autocorrelation, which is the cross-correlation of a signal with itself at different points in time, or heteroskedasticity especially in the case of autocorrelation raise my interest for different tests, methods, and applications to put in practice.

Putting these econometric methods in practice allows us to estimate different time series models. I find particularly challenging to forecast HICP for the next period in

time, and evidently putting the variance of inflation into this context. Besides the application of the different methodologies, the practical contribution of improving the inflation forecasting accuracy can be interesting combined with policy recommendations and it could provide a base for further research in this topic.

Academic Motivation

Inflation itself captures the rate of change of prices, which means the price level today reflects the patterns of past inflation. The rate of inflation over the past year is determined by:

$$Inflation = \frac{(P - P_{-1})}{P_{-1}} \tag{1}$$

Very high inflation is typically associated with very poor economic performance, with investment, consumption and output being depressed. When inflation is high, it tends to be more volatile and costly simply because it creates an uncertain environment and undermines the informational role of price levels (Carlin, W., & Soskice, D., 2005). On the other hand, if inflation is negative, meaning prices and wages will be lower in a year's time than they are now, the difficulty of cutting nominal wages raises alongside with the threat of a deflation trap levels (Carlin, W., & Soskice, D., 2005). Consequently, volatile inflation is distortive and masks the relevant changes in the economy causing misallocation of resources.

These arguments about high or negative inflation levels and the uncertainty around them lead to the assumption that if the expectations of the market are meant to be fulfilled, inflation level should be at a close to constant level (Carlin, W., & Soskice, D. (2005).

In the model developed by Carlin, W., & Soskice, D. (2005), we can observe that at higher inflation levels, people want to hold lower money balances, because they wish to economize on their holdings of money, so to restore the equilibrium in the money market, the real money supply must be lower than in the initial low inflation equilibrium (Carlin, W., & Soskice, D. (2005). This leads to additional costs, such as shoe-leather costs (people's trips to the bank or the cash machines more frequently), and menu costs (the time and effort involved changing price lists more frequently), hence making a higher, more volatile inflation environment considerably more costly (Carlin, W., & Soskice, D. (2005).

The book further argues, that policy makers should focus on a nominal anchor for the economy keeping the price level changes (inflation) low and stable. However, we still observe times with high inflation levels and also volatility. Those inflationary problems usually require painful and politically very unpopular solutions, resulting in unresolved social and political conflicts (Carlin, W., & Soskice, D., 2005).

Hence, it is of no surprise that over the past decades, policymakers, central banks, investors and private firms tried to predict the future levels of inflation. They did so because determining how the general price level will change is the key to an informed decision making process (Faust and Wright, 2013). In addition, policy makers use monetary and fiscal policy tools that have effects on inflation. Therefore forecasting HICP is also important to minimize distortions in inflation levels in the economy. However, economic theory often gives a steady state inflation equilibrium, but it is silent about the variation in inflation. Hence it is hard to choose a theory-based model for analyzing inflation variation, this is the motivation to use GARCH models allowing for variance changes over time. The case study by J. Faust and H. Wright examines the different approaches used to forecast HICP in the United States. However, so far only little attention has been dedicated to the actual volatility of the inflation that is the source of uncertainty.

Since the mid-1980s the variability of quarterly inflation has declined by about two thirds, the period known as Great Moderation (Bernanke, 2004). Faust and Wright argue that the presumption that this period will last forever is rather shaky (Faust and Wright, 2013). However, Bernanke remarks that lower volatility of inflation has advantages, such as improved market functioning, easier economic planning, and reduced hedging risks (Bernanke, 2004).

Especially after the recent financial crisis, it is clear that not only the level of inflation, but also its volatility, are of great importance. Determining the main drivers of this inflation volatility and forecasting in this setup would help decision makers to more accurately link future expectations to real levels of inflation. Particularly financial decision makers are interested in inflation and inflation uncertainty since discounts on (expected) future process of assets are directly linked to inflation (and interest rate) expectations.
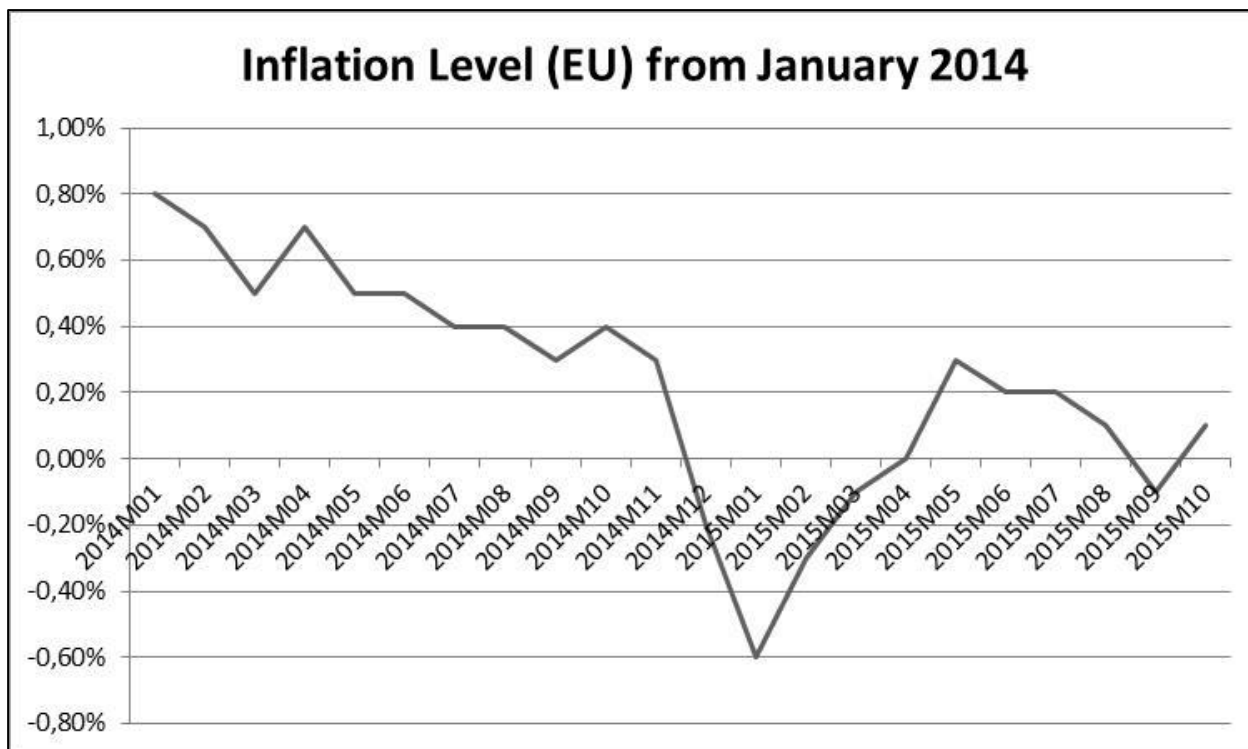
Graph 1: 2014 – 2015YTD EU Inflation Levels (monthly)

The graph above (Graph 1), however, clearly shows that the Euro Area experiences very low levels of inflation since January 2014. The expectation of the European Central bank is that low inflation will be prolonged but gradually return to 2% (Draghi, 2014). The speech by Mario Draghi, President of the ECB, "Monetary policy in a prolonged period of low inflation" at the ECB Forum on Central Banking (Sintra, 26 May 2014) sums up the current developments about low level inflation in the EU:

"...Low inflation is not particular to the Euro Area. Inflation is low across advanced economies, mainly due to the diminishing effect of oil prices on consumer prices. But looking at the anatomy of inflation, there are two factors specific to the Euro Area that contribute to especially low inflation here..."

The first is a common factor: the rise in the euro exchange rate and its effect on the price of internationally traded commodities. The second is a local factor: the process of relative price adjustment in certain Euro Area countries that pulls down aggregate inflation.

Falling commodity prices explain the lion's share of the disinflation the Euro Area has experienced since the end of 2011. Brent crude oil prices were down by around 7% in euro terms in the first quarter of this year, compared with a year earlier. Food prices were sharply down as well. In fact, these two components have together accounted for

around 80% of the decline in Euro Area HICP (Harmonised Index of Consumer Prices) inflation since late 2011.

To add to this, aggregate inflation has been dragged down by local factors linked to the sovereign debt crisis and the process of relative price adjustment in stressed countries. Several Euro Area countries are currently undergoing internal devaluation to regain price competitiveness, both internationally and within the currency union. The crucial adjustments vis-à-vis other Euro Area countries have to take place irrespective of changes in the external value of the euro.

So to sum up: falling energy and food prices, coupled with the effects of relative price adjustment in stressed countries, explain almost fully the disinflation we have seen in the Euro Area…" (Draghi, "Monetary policy in a prolonged period of low inflation" at the ECB Forum on Central Banking, 26 May 2014)

The European Union and the Euro Area consist of several very diverse economies in different stages. Therefore it is viable to anticipate that the main changes in inflation variance stems from those stressed countries, while other economies as part of the Euro Area (for example, Germany) exhibit a relatively stable variance of inflation over time.

According to these arguments, there is a need of disaggregation to country level analysis and observe how the different country specific trends and macroeconomic variables contribute to the overall inflation levels in the European Union. As Draghi argues, low inflation is not particular to the Euro Area because inflation tends to be low across advanced economies, mainly due to the diminishing effect of oil prices on consumer prices (Draghi, 2014). In those advanced economies, where inflation is low, the stability of the price level changes brings also a more predictable and relatively constant variance over time. So the main question remains, where does the variance change on aggregate level come from and if those more uncertain economies, in fact, are the initial source of inflation variance, how can one incorporate these characteristics into the inflation forecasting model?

Research Questions

The research idea focuses on the above-discussed factors and their effects on the inflation levels and inflation volatility in the Euro Area. The central question of the proposed thesis will be:

"Which macroeconomic factors affect the inflation level and its volatility and how?"

More precisely, this research attempts to find an alternative model to determine the levels of inflation in the Euro Area through focusing on its volatility and on the changes of this volatility over time. This leads to several sub questions:

- What are the main drivers of volatility changes?
- To what extent can these models forecast HICP accurately?

There are several potential macroeconomic factors, such as unemployment or other economic factors used in the USA, which can have an influence on the levels of inflation. Depending on data availability and the models to be used, the above mentioned questions will be addressed and the proposed model based on these variables will be used to answer another set of sub questions:

- What is the probability that the volatility of inflation will change in the next period?
- How severe will that change be?
- What will potentially cause inflation volatility to change?
- Are volatility changes accompanied by changes in expected inflation?
- What are the similarities and differences in inflation patterns across Euro Area countries?

In order to properly answer these proposed research questions, it is important to take a look at the underlying theories behind these topics.

Chapter 2: Theory

Chapter Two outlines the potential contribution of this papers, and introduces the theoretical background of inflation forecasting

Contribution

The academic contribution of this paper would be the focus on inflation volatility together with the levels of inflation. Depending on data availability, other econometric models can also be applied, for example, probit or logit models for predicting the probability of the change in inflation variance. Further, applying a robustness check will ensure that the results can be generalized. Additionally, a case study like the one by J. Faust and H. Wright has not been done in the Euro Area, therefore given good data availability a comprehensive review of used methods can be relevant to European private firms and policy makers. A further contribution is to address similarities and differences in inflation patterns across Euro Area countries.

Finally, most of the research about forecasting HICP has been limited to the USA. Given today's economy it is timely to look at the Euro Area or some representative countries and provide a study about the inflation volatility in those economies. For example, Rother finds evidence before the financial crisis that the negative consequences of inflation volatility are of particular concern, given the destabilising effects of discretionary fiscal policies (Rother, 2004). So given an even more uncertain time period it is of great importance to examine the main drivers of inflation volatility.

Theoretical background

There has been an extensive amount of research done about inflation and its properties, but the main focus in most studies was the levels of the inflation and the possibility to forecast the future levels of it. To mention a few, whether it is the Phillips-curve model, or market-based inflation hedging methods, the uncertainty component has not been fully encountered. In recent years, macroeconomic forecasting and its uncertainty have gotten more attention, especially after the recent financial crisis. It is also important to mention the probability of deflation (for USA it happened only in 2009 Q3). Without proper estimates of inflation and its volatility, this deflation is an "unexpected" event. This is also a reason for increased attention in the Euro Area to estimating and forecasting

volatility. Furthermore, Giordani and Söderlind argue that forecasters underestimate inflation uncertainty (Giordani and Söderlind, 2002).

Hence, forecasting HICP has become the main focus of many studies in the USA and many different approaches have been suggested. For example the model by Stock and Watson (2007), which allows the variances of the permanent and transitory disturbances evolve randomly over time, that is, an unobserved components model with stochastic volatility (UCSV) (Stock and Watson, 2007). Furthermore, Groen et al. (2013) claim that a model allowing for structural breaks in the error variance results in very accurate point and density forecasts (Groen et al., 2013).

Based on these academic articles and supporting theories, it is clear that nowadays many new methods are being used to improve the forecasts of inflation. However, J. Faust and H. Wright argue that a move to a fundamental re-think of the topic of inflation forecasting is called for (Faust and Wright, 2013). Their comprehensive comparison of methods shed light on the performance of these forecasting models in the USA, but on the other side a case study like theirs has not been done in the Euro Area. Although there are a lot of different aspects to consider when forecasting HICP, the attention to its volatility and the sources and main drivers have been avoided in most of these researches. The obvious question now is: "Why did these studies not look at the volatility component?" I plan to answer this question by collecting the most appropriate dataset.

Chapter 3: Methodology

Data

Source

The major source of data is the ECB's website, since the Euro Area inflation measures are published there. The exact span of the data is limited for some countries, hence selection of countries and time periods will be based on data availability. Explanatory variables will be macroeconomic factors (such as unemployment, marginal costs) and in addition economic factors from the USA. There are several other potential factors to be included, and these will be decided according to data availability and literature on inflation determinants. As a starting point, I collected the overall Harmonized Index of Consumer Prices (HICP) normalized monthly levels from 1996 (See Graph 2), as well as other potential explanatory variables such as unemployment or labour market productivity. The main source of datasets is OECD's database.

HICP - EU (2010=100)
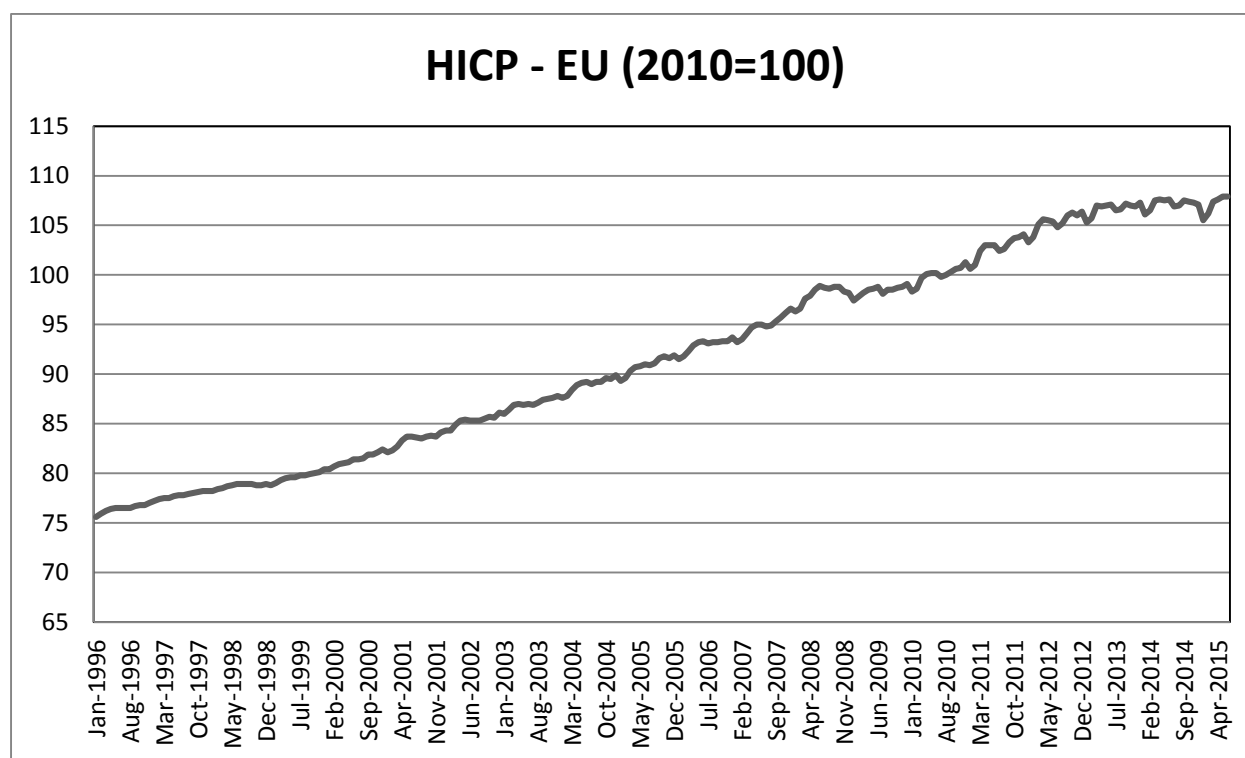
Graph 2: 1996 – 2015YTD EU HICP (monthly)

Finding a robust and reliable dataset for Europe is a difficult task. The composition of the European Union changed significantly over the past decades, additionally a recent ECB papers argues that:

"…a vast literature has emerged analysing the degree of inflation persistence. One of the central issues is whether inflation persistence follows a unit root. Recent empirical evidence not only suggests inflation has varied over time, but also that inflation is not an intrinsically persistent process…" (Lünnemann, P., and T. Mathä. "How Persistent is Disaggregated Inflation." *An Analysis* (2004)).

These issues result in some cumbersome dataset issues, which will be explained below after the data description. To correctly interpret the results, the clarification of definitions used by the OECD is needed in order to avoid misconceptions. Table 1 includes the definitions for the compositions referring the European Union:

Composition	Definition
Euro Area (19)	**Refers to the evolving composition of the Euro Area.** (Data prior to 2001 refer to EU11 (Belgium, Germany, Ireland, Spain, France, Italy, Luxembourg, the Netherlands, Austria, Portugal and Finland). Data from 2001 to 2006 refer to EU12 (EU11 plus Greece). Starting in January 2007, data refer to EU13 (EU12 plus Slovenia))
EU28	**Refers to the 28 countries of the European Union.** (Data prior to 1990 have been calculated by the OECD based on the first 12 Euro Area member states plus Denmark, Sweden and the United Kingdom.)

Table 1: OECD Definitions: Euro Area, EU28

All of the above used definitions, as well as variables inputs are common across all European countries to avoid misconceptions of definitions. Therefore, similar metrics have been used for the analysis. As I mentioned previously, the composition and time span availability of the data differ among variables and also among countries, hence following I list these features of each of the variables. All definitions are taken from the OECD's database and the wording is kept identical for reliability and validity reasons (See Table 2):

Data	Frequency	Availability	Measure	Definition by the OECD Database*
HICP	Monthly	For Euro Area and EU from January, 1996 to August, 2015	Index (2010 = 100)	The Harmonized Index of Consumer Prices (HICP) is the measure of prices used by the Governing Council for the purpose of assessing price stability. Within the European Union, Harmonized Indices of Consumer Prices (HICP) are consumer price indices compiled on the basis of a harmonized coverage and methodology. HICPs at the "all items" level are presented for 21 European countries
Unemployment	Monthly	For Euro Area from July, 1990 (estimated by ECB) and for EU from January, 2000 to September, 2015	Harmonized unemployment rate: all persons (index), seasonally adjusted	Harmonised unemployment rates define the unemployed as people of working age who are without work, are available for work, and have taken specific steps to find work. The uniform application of this definition results in estimates of unemployment rates that are more internationally comparable than estimates based on national definitions of unemployment
Unit Labour Cost	Quarterly	For Euro Area from Q1-1996 to Q2-2015	Index (2010 =100), seasonally adjusted	Unit labour costs (ULCs) measure the average cost of labour per unit of output. They are calculated as the ratio of total labour costs to real output. Quarterly ULCs can be decomposed into the components labour compensation per employee and output per person employed (employment-based labour productivity). Every effort has been made to ensure that data are comparable across countries.
M1	Monthly	For Euro Area from January 1990, estimated	Seasonally adjusted index based on 2010=100.	Narrow Money (M1) includes currency i.e. banknotes and coins, plus overnight deposits.
M3	Monthly	For Euro Area from January, 1990, estimated	Seasonally adjusted index based on 2010=100.	Broad money (M3) includes currency, deposits with an agreed maturity of up to two years, deposits redeemable at notice of up to three months and

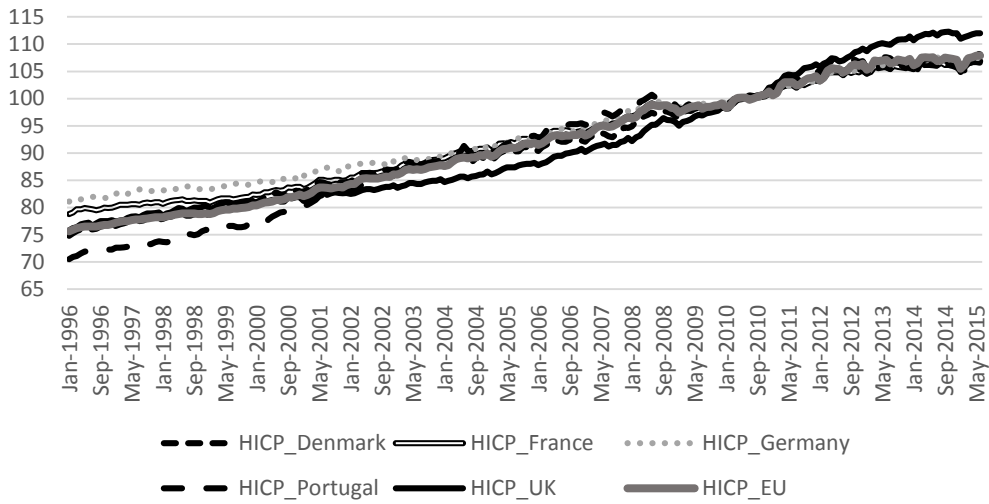|				repurchase agreements, money market fund shares/units and debt securities up to two years
Long term interest rates	Monthly	For Euro Area from January, 1990, estimated to September, 2015	Long-term interest rates, Per cent per annum	Long term (in most cases 10 year) government bonds are the instrument whose yield is used as the representative 'interest rate' for this area. Generally the yield is calculated at the pre-tax level and before deductions for brokerage costs and commissions and is derived from the relationship between the present market value of the bond and that at maturity, taking into account also interest payments paid through to maturity.
Short term interest rates	Monthly	For Euro Area from January, 1994 to September, 2015	Short-term interest rates, Per cent per annum	Short term rates are usually either the three month interbank offer rate attaching to loans given and taken amongst banks for any excess or shortage of liquidity over several months or the rate associated with Treasury bills, Certificates of Deposit or comparable instruments, each of three month maturity. For Euro Area countries the 3-month "European Interbank Offered Rate" is used from the date the country joined the euro.
GDP changes	Quarterly	For Euro Area from Q2-1995 to Q2-2015	Growth rate compared to previous quarter, seasonally adjusted	Gross domestic product (GDP) at market prices is the expenditure on final goods and services minus imports: final consumption expenditures, gross capital formation, and exports less imports. "Gross" signifies that no deduction has been made for the depreciation of machinery, buildings and other capital products used in production. "Domestic" means that it is production by the resident institutional units of the country.

Table 2: OECD Definitions: Macroeconomic Variables

*All definitions are set by the OECD Database. Online at: < https://data.oecd.org/>
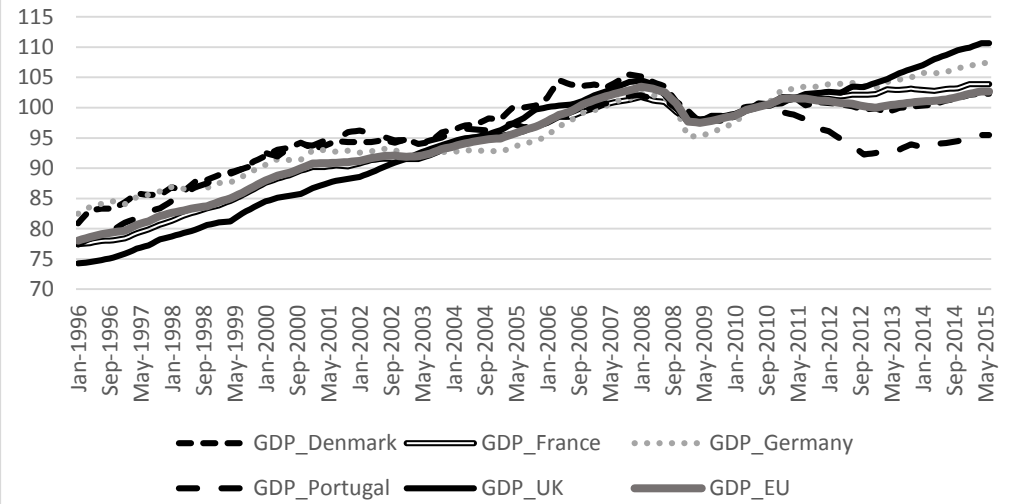
Based on data availability, and to allow the best fit possible, the data used will be referring the Euro Area from January, 1996 to June, 2015 for all variables. That leaves us with 234 observations for each of the above variables. The small sample size can be a drawback of the analysis. Furthermore, as one can notice, not only the availability and the composition are different, but also the frequency of the datasets differ from each other, therefore I needed to interpolate the GDP changes indicator and the ULC dataset to monthly. This will be a drawback of the analysis further on. Most of the dataset becomes stationary after the first difference, however, the HICP indicator only becomes stationary after the second difference. Possible explanation is seasonality or the existence of change in mean series over time. Standard tests for unit roots then fail and see higher AR coefficients (in the extreme this can lead to a unit root), Maddala and In-Moo Kim point out that although often used, the ADF test is "useless" in practice, because there is no uniformly powerful test for the unit root hypothesis (Maddala, In-Moo Kim, 1999). The aforementioned ECB paper examines this issue more closely and finds out that allowing for a structural break in the series avoids spuriously high inflation persistence estimates. An exogenous break at the start of EMU stage III and the effects of an important modification to the HICP data collection methodology, i.e. the inclusion of sales prices lead to significant mean differences in the date over the analyzed period (Lünnemann, P., and T. Mathä, 2004).

I will follow their methodology to find these structural breaks and include an explanatory variable (such as a dummy variables) to take care of mean changes, then the HICP conditional on this variable will become stationary after the first difference. To formally test, I will run a Chow test for the following breakpoints: 2001Q1 for the inclusion of sales and 1999Q1 for the EMU stage III. To visualize the different types of datasets, and their characteristics, all datasets have been normalized to the 2010 levels, according to the OECD standards. The quarterly data (ULC and GDP) have been interpolated by attributing the quarterly changes equally to each respective month within the quarter (See Graphs 3-10):

HICP (2010=100)

Legend: HICP_Denmark, HICP_France, HICP_Germany, HICP_Portugal, HICP_UK, HICP_EU

GDP (2010=100)

Legend: GDP_Denmark, GDP_France, GDP_Germany, GDP_Portugal, GDP_UK, GDP_EU

Short Term Interest Rates (2010=100)

Legend: ST_Denmark, ST_France, ST_Germany, ST_Portugal, ST_UK, ST_EU

Long Term Interest Rates (2010=100)

Legend: LT_Denmark, LT_France, LT_Germany, LT_Portugal, LT_UK, LT_EU

14

M1 (2010=100)

Legend: M1_Denmark, M1_UK, M1_EU, M1_Germany, M1_France, M1_Portugal
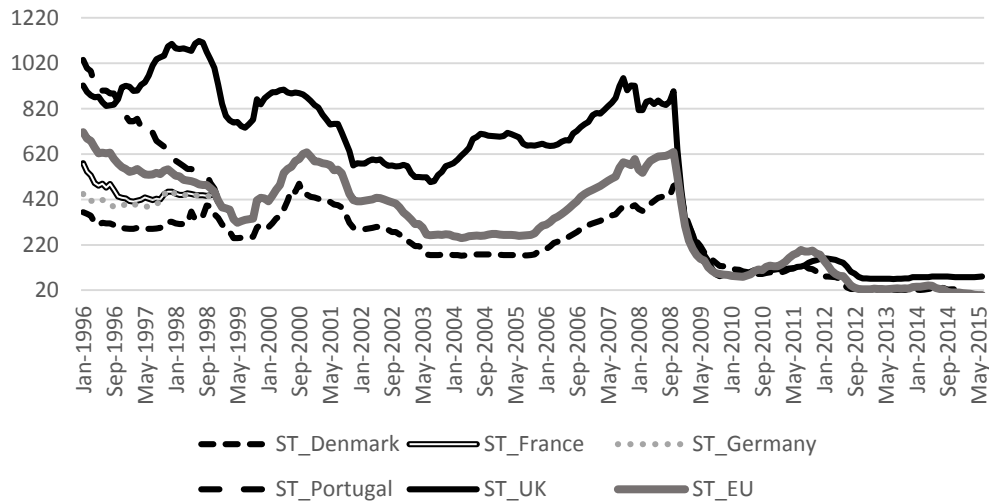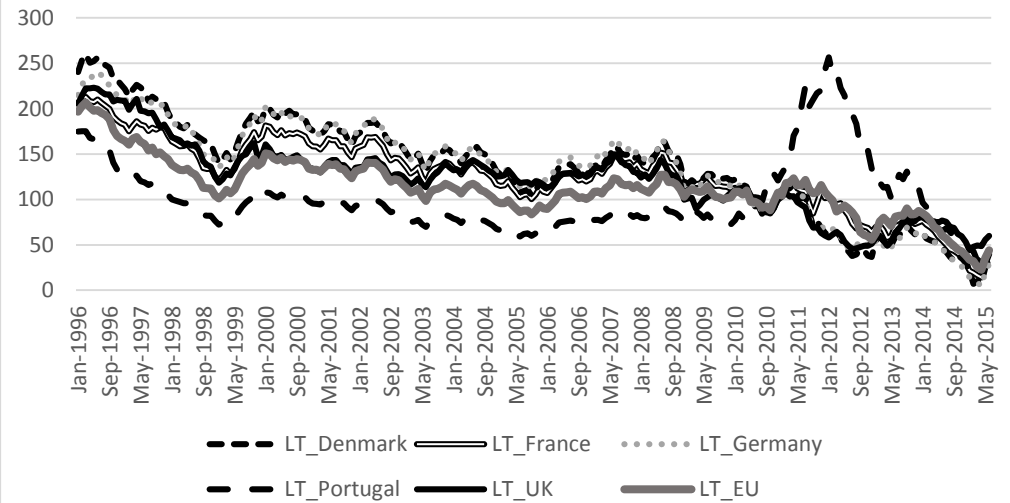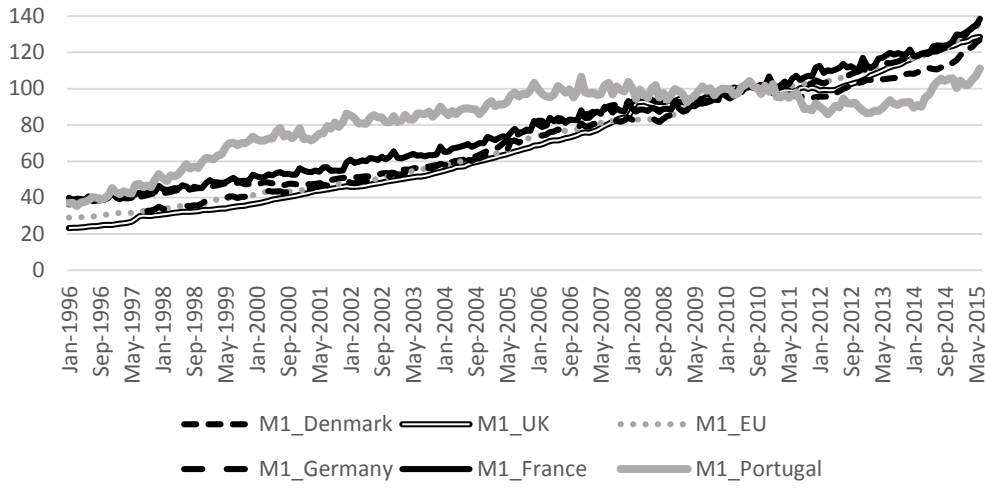
Harmonized Unemployment Rate (2010=100)

Legend: HUR_Denmakr, HUR_France, HUR_Germany, HUR_Portugal, HUR_UK, HUR_EU

M3 (2010=100)

Legend: M3_Denmark, M3_UK, M3_EU, M3_Germany, M3_France, M3_Portugal

ULC (2010=100)

Legend: ULC_Denmark, ULC_France, ULC_Germany, ULC_Portugal, ULC_UK, ULC_EU
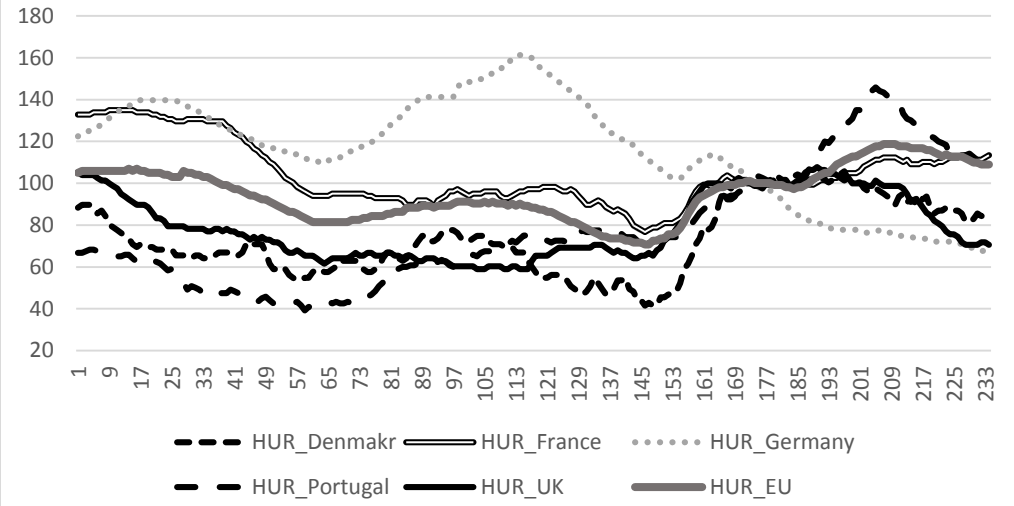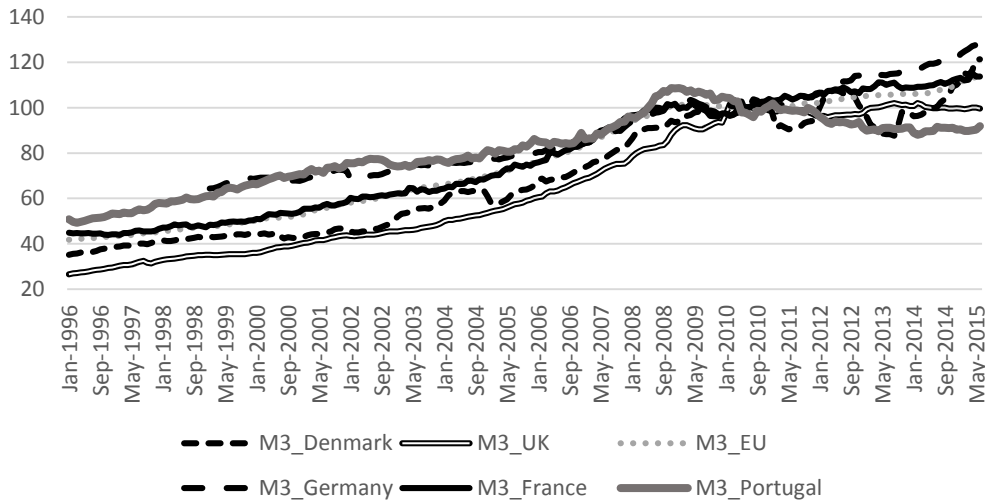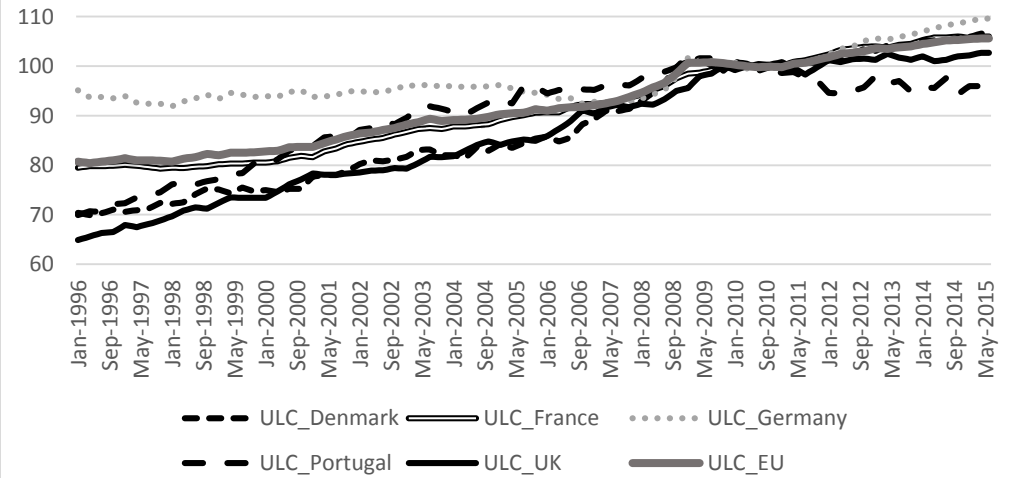
Graphs 3-10: Macroeconomic Variables

Region	Whole Sample		1996-1999		2000-2003		2004-2007		2008-2011		2012-2015 June	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Denmark	⇒1.96	⬇0.0470	⇒1.89	⬆0.0088	⇒1.89	⬇0.0112	⬇1.96	⬇0.0089	⬇2.00	⬆0.0114	⬇2.02	⬇0.0025
France	⬆1.96	⬇0.0425	⬆1.91	⬇0.0043	⬆1.91	⬇0.0098	⬆1.97	⬇0.0093	⬇2.00	⬇0.0108	⬇2.02	⇒0.0035
Germany	⬆1.97	⬆0.0388	⬆1.92	⬆0.0047	⬆1.92	0.0072	⬆1.97	⬆0.0100	⬇2.00	⬆0.0104	⬆2.03	⬆0.0047
Portugal	⬇1.96	⬆0.0580	⬇1.87	⬆0.0106	⬇1.87	0.0187	⬆1.97	⬆0.0131	⬇2.00	⬇0.0106	⬆2.03	⇒0.0032
United Kingdom	⬆1.96	⬆0.0527	⬆1.90	⬆0.0081	⬆1.90	0.0067	⬆1.95	⬆0.0113	⬇2.00	⬆0.0124	⬆2.04	⬆0.0076
All regions	⇒1.96	⇒0.0487	⇒1.89	⇒0.0069	⇒1.89	⬇0.0112	⇒1.96	⇒0.0111	⬆2.00	⬇0.0106	⬆2.03	⬆0.0043

Table 3: Regional Statistics per Period for HICP (after log transformation)

Graph 11: EU28 HICP Levels 1996 – 2015YTD
(All data have been normalized to 2010=100 points.)

HICP as dependent variable

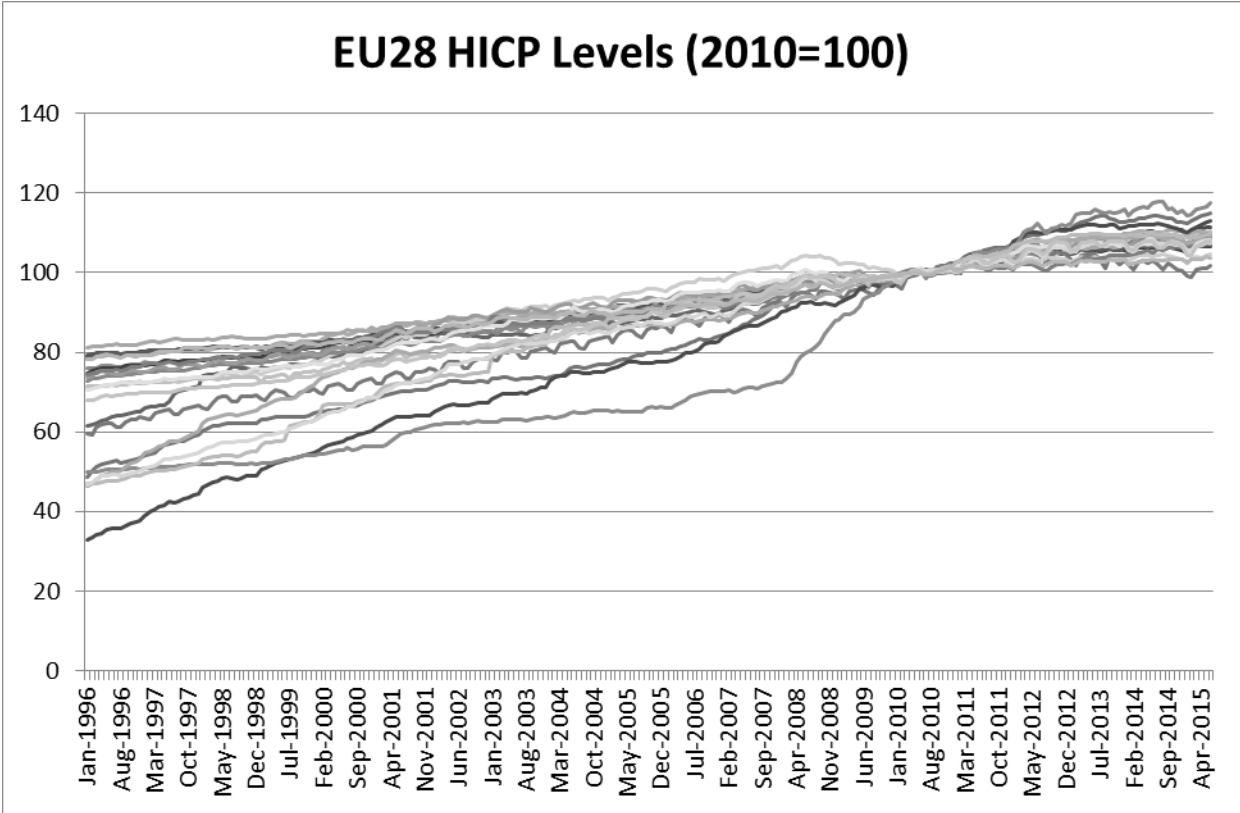Beck et al. (2009) claim that there exist large and long-lasting differences in inflation rates across European regions (see Table 3 and Graph 11), which are not related to business cycle or income growth dynamics (Beck, G. W., Hubrich, K., & Marcellino, M., 2009). They attribute these differences mainly to market distortions and other structural characteristics, which further implies welfare affecting behavior (Beck, G. W., Hubrich, K., & Marcellino, M., 2009). Thus there is a clear motivation to discuss this topic in further detail, from both a forecasting and volatility point of view.

When the focus is on inflation levels of a specific country, for example Germany, there are clear changes of dynamics of inflation over the past decade (See Graph 12). These changes in inflation dynamics over time, and across regions call for a closer examination of explanatory variables, not just from an economic point of view, but also from a pure statistical and econometric one. The set of graphs above illustrate the variety of behaviour observed with time series data. In the above section, I paid special attention to stationarity of the variables. Nonstationary variables display a "wandering behaviour", while a stationary data displays a more "fluctuating" behaviour. The graphs below clearly show that most of the macroeconomic variables I plan to incorporate in my modelling exhibit some sort of "trending" behaviour. The main reason why it is essential to make sure that all data is stationary, prior starting the modelling is that there is a big chance of obtaining significant results from unrelated data when nonstationary time series are used in regression analysis, these are called spurious regressions.

This ties back to my earlier reasoning at the beginning of Chapter 1 about interpretation of the results, since spurious regressions make no sense in reality, although purely on the results, they might well be very significant in terms of numbers and residuals would be highly persistent. Besides nonstationarity, I will also test for cointegration between the below variables. This cointegrating relation is a long run relation. That means that the one series and the other series have a unit root (a stochastic trend), but they don't stray away too far from each other as the error term is stationary. To test this long run equilibrium I will run an Engle-Granger procedure, because when predicting future levels of inflation, one could expect a long run relationship of some variables. Stock and Watson argue here that in general terms, inflation forecasts produced by the Phillips curve are more accurate than using methods

based on other macroeconomic variables. However, the forecasts can be improved significantly with measures of real aggregate activity, other than unemployment only (Stock and Watson, 1999). For example, the expectations hypothesis of the term structure of interest rates suggests that the spreads between interest rates of different maturities incorporate forecasts of inflation made by market participants. Similarly, the quantity theory of money predicts that, in the long run, the rate of inflation is determined by the long-run growth rate of monetary aggregates (Stock and Watson, 1999, p305 Forecasting HICP in Journal of Monetary Economics 44 (1999) 293- 335). Their reasoning gives a very good incentive for my choices of variables I plan to use, such as long term and short term interest rates, as well as M1 and M3 levels.

While describing the choice of specific variables I will use, one might ask why HICP is used, especially taking into account its technical difficulties. First, CPI, the other common variable measuring inflation, is not available for Europe, and HICP is more appropriate for the European market. Both the Consumer Price Index (CPI) and the Harmonised Indices of Consumer Prices (HICP) are designed to measure, in index form, the change in the average level of prices paid for consumer goods and services by all private households. Both the CPI and the HICP are used to measure consumer inflation (Central Statistics Office, 2012). There are common similarities in both measures. The purpose of both CPI and HICP is to measure the change in the average level of prices of a fixed basket of consumer goods and services. They both include all private households and foreign tourists on holiday, but exclude institutional households. CPI and HICP use the same price data collected from the same retails outlets/service providers. Compilation and aggregation of both measures are done with the same methodology developed by Eurostat.

However, next to these similarities, there are also a few differences. For example, CPI was first used in the United States already in March 1922, whereas HICP was launched in 1996. The current base reference period of the CPI is 2011, while in the case of HICP is 2010 (as used in the regression analysis later on). The coverage is also different in both cases. The Household Charge was added to the "Miscellaneous goods and services" item heading, however the following item heading are excluded from the HICP's scope:

- Mortgage interest
- Building materials
- Motor cycle/car taxes and insurance

- House insurance
- Union subscriptions

The HICPs enable international comparisons of inflation rates to be made between member states within the European Union. The standardized scope and coverage of the HICP measure makes it a better dependent variable, as those will be similar in all member states of the European Union.

Based on the differences outlined above by the Central Statistics Office, we can conclude that HICP in fact is a better index for inflation due to market properties.
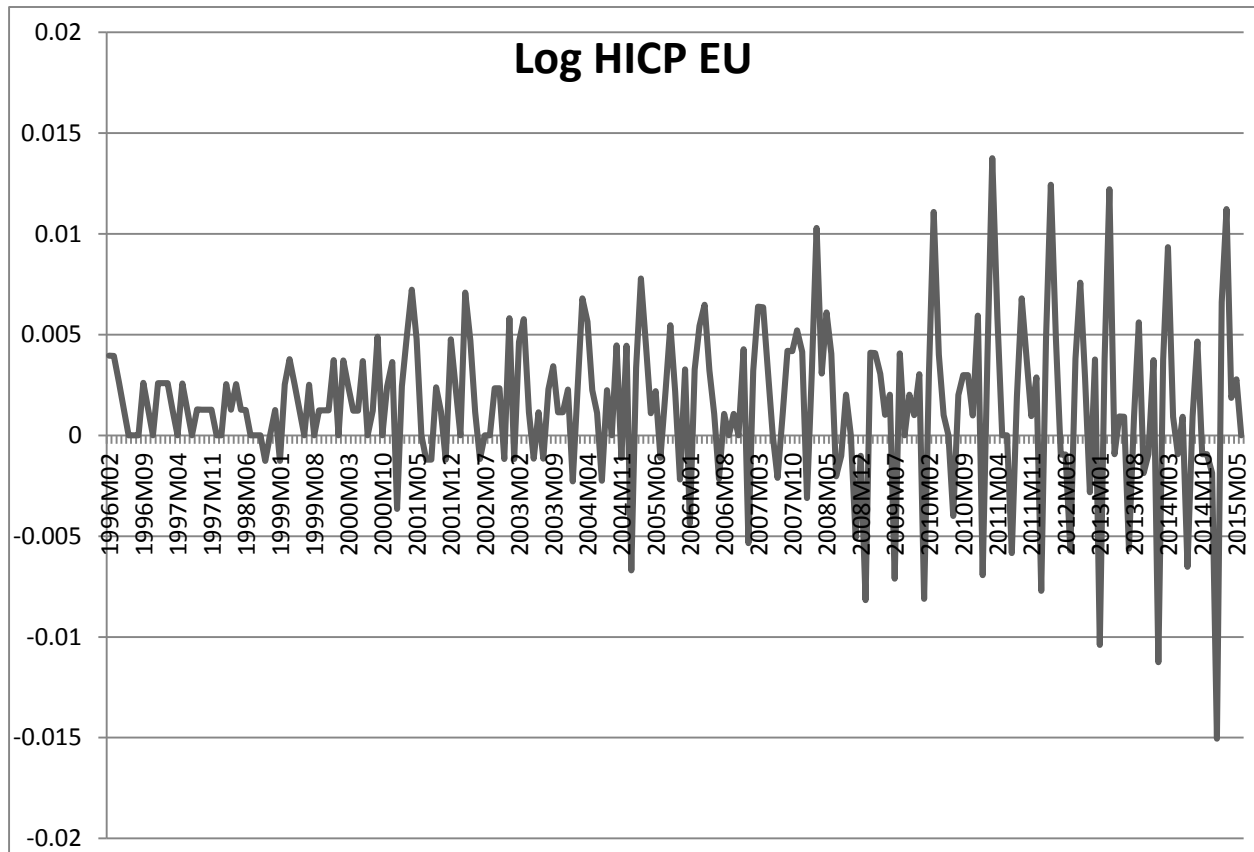
Graph 13: HICP EU: Log differences

Variance of Inflation

Furthermore, a good motivation for the proposed thesis is that after looking at log differences of the HICP data for Europe, there is a clear difference in the variance of the observed variable (See Graph 13).

To formally test whether there is a statistically significant difference in the variance of the observed variable, one can conduct a VR (Variance Ratio) test. By exploiting the fact that the variance of random walk increments is linear in all sampling intervals, the VR methodology consists of testing the random walk hypothesis against stationary alternatives. This means that the sample variance of k-period return is k times the sample variance of one-period (Charles, A., & Darné, O. (2009)).

The VR at lag k is then defined as the ratio between (1/k)th of the k-period to the variance of the one-period return. To test the hypothesis that a given time series is a collection of independent and identically distributed observations, the central idea of the variance ratio test is based on the observation that when returns are uncorrelated over time, we should have V(k) = 1 (Charles, A., & Darné, O. (2009)).

Null Hypothesis: LOGHICP is a martingale			
Sample: 1996M01 2015M06			
Included observations: 232 (after adjustments)			
Heteroskedasticity robust standard error estimates			
User-specified lags: 2 4 8 16			
Joint Tests		Value	df	Probability
Max |z| (at period 4)*	3.805.108	232	**0.0006**
Individual Tests			
Period	Var. Ratio	Std. Error	z-Statistic	Probability
2	0.641223	0.102827	-3.489.123	**0.0005**
4	0.326548	0.176986	-3.805.108	**0.0001**
8	0.161048	0.250334	-3.351.337	**0.0008**
16	0.085799	0.350285	-2.609.874	**0.0091**
Test Details (Mean = -1.7070694897e-05)			
Period	Variance	Var. Ratio	Obs.
1	2.8E-05	--	232
2	1.8E-05	0.64122	231
4	9.1E-06	0.32655	229
8	4.5E-06	0.16105	225
16	2.4E-06	0.08580	217

Table 3: VR Test

In the case of inflation, it is clear that for any chosen period (k=1,2,….,16) the variance of these subsequence of data is different, hence there is clear variance change over time (See Table 3 above).

Furthermore, this methodology can be further used to indicate high persistency, and consequently a unit root. As discussed in the previous section, high inflation persistence can lead to misinterpreted unit root tests. The measurement of the degree of this persistence is actually another way of testing for unit roots.

As previously discussed, the VR is the variance of the kth difference of a time series by k-times the variance of its first difference (Maddala, In-Moo Kim, 1999), where:

$$VR_k = \frac{V_k}{V_1} \qquad (2)$$

A large econometric literature actually finds that inflation is highly persistent, almost mimicking the behaviour of a random walk process (Altissimo, F., Bilke, L., Levin, A., Mathä, T., & Mojon, B. (2006)). Initially, the motivation behind using the variance ratio test was the paper testing the random walk hypothesis of the Euro exchange rates by Jeng-Hong Chen (2008). Variance ratio test is a powerful tool to test the random walk hypothesis. His work compares three different variance ratio tests and concludes that the Euro/U.S. Dollar exchange rate market is considered weak-form efficient. His results are less relevant to my

topic, however the methodology he used to interpret the variance ratio tests helps me to understand that there are in fact changes in inflation variance over time. The first graph above clearly shows that variance changes over time, as well as the result of the standard Eviews variance ratio test's null hypothesis cannot be rejected, hence there is statistical evidence of change in inflation variance over time. These support my initial decision to focus on the variance of inflation as well, and not just on predicting the future levels of inflation itself.

Methodology

For determining and examining the effects of macroeconomic factors on inflation and its volatility, I need to determine first which measure of inflation to use. For the decision I will rely on the relevant literature mentioned above to justify my choice between CPI, GDP deflator or other possible measures. To estimate inflation levels and volatility I will use standard inflation (time series) models allowing for mean changes in inflation and variance changes (such as the GARCH model). These models will be extended to include additional explanatory variables for inflation and inflation variance. To identify the changes in inflation and its volatility I will use standard tests, such as the Chow break point test. Since I have to take into account the mean changes I will include dummy variables for this purpose.

There is no ultimate model to forecast HICP, hence I will be modelling different scenarios with different sets of variables and compare these models with the standard GARCH and AR models to test with forecast performance tools which of these models perform better. Since it has been discussed in the previous section, there is a good indication of changes of variance of inflation over time. The ARCH models, where the time varying variance is a function of a constant term plus a term lagged once, the square of the error in the previous period, will be the main starting point of the analysis, which then will be expanded with the macroeconomic variables.

To point out an arguable caveat of inflation forecasting is the argument by Stock and Watson in their paper "Modelling Inflation after the Crisis" (2010). There is a considerable instability in inflation forecasting models, which leads forecasters to use multivariate models. With their extension of their initial time-varying unobserved components model (2007) they hypothesize different scenarios with different predictive variables and conclude that there are

several sources of uncertainty within the inflation forecasting models. First, if the volatility of the trend component within inflation varies little then inflation will revert to the trend, however, if the trend itself becomes more volatile then the trend tracks inflation (Stock and Watson, 2010). They also point out that there is a regional component that can change the dynamics of inflation, where the conventional monetary policy becomes ineffective, than most models fails to handle the forecasting.

These difficulties and arguments are compounded and make an even bigger influence in the given European situation, therefore upon review of the most commonly used models for inflation forecasting, I will try to incorporate the best of these methods to improve the accuracy of inflation forecasts.

To sum up this section, Stock and Watson make a very good observation in their other paper "Why has U.S. inflation become harder to forecast" (2007):

Depending on the perspective we take, the rate of price inflation has become both harder and easier to forecast. On the one hand, inflation (along with many other macroeconomic time series) is much less volatile than it was historically, and the root mean squared error of naïve inflation forecasts has declined sharply. In this sense, inflation has become easier to forecast: the risk of inflation forecasts, as measured by mean squared forecast errors (MSFE), has fallen (Stock and Watson, 2007)

On the other hand, if we take univariate models as a benchmark and measure the relative improvement of standard multivariate forecasting models, we find that those models are improving less and less over a naïve forecast of average twelve-month inflation by its average rate over the previous twelve months. In this sense, inflation has become harder to forecast, at least, it has become much more difficult for an inflation forecaster to provide value added beyond a univariate model (Stock and Watson, 2007). These characteristics and new dimensions changed irreversibly the properties of inflation forecasts.

Chapter 4: Results

Research Design

As discussed before, specific properties of inflation and other macroeconomic data need to be analysed prior to the modelling of the inflation forecasting equation. The dependent variable an all analyses of this thesis is the HICP measure (definition and description of these data are provided in Chapter 3). First, to test for stationarity, I employed an Augmented Dickey Fuller (ADF) test. It turns out that the dataset itself might not have a unit root, however, after including the trend in the test, it turned out to be significant and the variable became stationary on the 10% level. To be certain, I also have conducted a Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test also testing for unit root. I continued doing so with all variables on country levels. I have run ADF tests and if a variable showed a unit root I took the log difference and made all variables stationary, if applicable (Results of all stationarity tests are provided in Appendix A).

After transforming all variables to ensure stationarity, I employed standard time series econometric models for stationary data. These models include the Auto-Regressive (AR) model, Moving Average (MA) model and the combination of the former models, namely ARMA models. Since the dataset is on a monthly basis, I have included 12 lags at first to account for potential (annual) seasonality. Then I have removed the non-significant lags of inflation one by one according to t-tests, thereby improving both the R-squared, and the Akaike info criterion. I have done the same with the macroeconomic variables, using a significance level of 5% in all testss.

To formally test these models capability for forecasting, I have done an in-sample (both dynamic and static) forecasting and compared these models based on their RMSE value and MAE value. I continued doing this until in the given framework the measures of comparison (R-squared, Akaike criterion, MFSE) cannot be further improved. To reach that point, I will need to use more sophisticated models, allowing for mean and variance change over time, such as GARCH models.

Improving inflation forecasting using macroeconomic variables has two purposes. One being to actually forecast HICP, while the other being to judge how the variance of inflation will evolve over time. For the second purpose, including the log of the variance in the

equation in the GARCH environment gives a better indication whether the model is able to forecast HICP or the variance of inflation better.

After deciding on an aggregate model for the Euro Area, the following step is to break this framework down to country level analysis. There are certain macroeconomic indicators, and different economy components in each of the Euro Area countries which will help in the understanding of country level differences, as well as to uncover the source of inflation volatility on the Euro Area aggregate level.

Interestingly, on the country level analysis, there are clear differences not just in the significance level of macroeconomic variables, but also whether the GARCH component is significant or not on country level. This means that the overall change of inflation variance stems from country level differences, and those variance changes over time are not present in some countries (GARCH coefficient insignificant). This indicates the diversity of the Euro Area economy and hence makes it harder to forecast HICP properly.

Therefore I follow the steps of the initial Euro Area modelling, and compare the best model and estimation results for each country to the one of the Euro Area. For each country, as well as the Euro Area, the best performing econometric model is chosen according to in-sample fit. This leads us to two models on the Euro Area-wide aggregate level, and one model for each country.

After deciding on the model, I conduct a theoretical "out-of-sample" forecast to judge how well alternative models perform. For this, I divide the dataset into two subsamples. The first period runs from 1996M01 until 2008M06 and the second from 2008M07 until 2015M06. 2008 mid-year is chosen as a breaking point since the start of the financial crisis in the EU affected inflation levels substantially. If the model forecasts are accurate during the crisis period and observed inflation levels lie within the 95% confidence level for inflation forecasts, the model should be robust enough to work well in other periods too.

As institutions have a different focus and approach towards inflation level changes, they will either want to look at dynamic or static forecasting. I will conduct both methods within the same framework for each country.

As a last step, to evaluate the overall performance for the aggregate EU forecast, I also forecast in the model with the rolling window method one period ahead (12 months) and repeat the steps three times for a more complete overview.

In the following section, the results are shown with detailed explanation of the evolution of various model choices depending on country and purpose of the model. All unit root test results, regression outputs and forecasting outcome can be found in the Appendix.

Results

Measures of evaluation

As explained in the previous section, the final decision of the chosen model will be based on several criteria. Depending on the purpose of the model, and for what it will be used, there are several different ratios and information criteria to compare the models.

The first purpose of the analysis is to explain inflation, hence the models can be compared based on how well the regression fits. For this, the obvious measure is the adjusted R squared. R squared measures the success of the regression in predicting the values of the dependent variable within the sample and is between 0 and 1. The higher the measure the better it fits. However, as ARMA models use many explanatory variables and iterations, the adjusted R squared should be used. This is a close relative of R squared, where the formula corrects for sample size and number of predictors:

$$Adjusted\ R^2 = 1 - \frac{(1-R^2)(N-1)}{(N-p-1)} \tag{3}$$

where p is the number of predictors,
and N is the total sample size.

However, adding variables to a model always increases R squared and equivalently the likelihood, whereas, information criteria are likelihood based criteria to evaluate / compare model performances. There are two common information criteria: Akaike Information Criterion (AIC) and/or Schwarz Bayesian Criterion (SBC). Both of these measures are founded on information theory. They offer a relative estimate of the information lost when a given model is used to represent the process that generates the data. In doing so, it deals with the trade-off between the goodness of fit of the model and the complexity of the model:

$$AIC = -2\ln(L) + 2n \tag{4}$$
$$SBC = -2\ln(L) + \ln(T)\,n \tag{5}$$

where ln(L) is the maximized value of the log of the likelihood,

T is the number of observations

and n is the number of explanatory variables in the model.

For model selection, the model providing smallest information criteria is preferred. Both of these improve with the likelihood of the model and they penalize for too many parameters in the model (including irrelevant variables will be penalized through n). One advantage of the information criteria is that they are able to compare nested models unlike R squared. For example, AR(1) model is an AR(2) model with parameter restriction $\beta 2 = 0$:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \varepsilon_t \tag{6}$$

Although AIC/SBC only provide informal test for model comparison since there is no test statistic or critical value to evaluate the performance of the model. In cases, where AIC will indicate a different model as the better one, I will rely in SBC due to its asymptotic properties.

Another way of evaluating the models is to make a forecast for a historical period. There are two choices of the forecasting methods: dynamic and static forecasting. Dynamic forecasting calculates forecasts for periods after the first period in the sample by using the previously forecasted values of the lagged left-hand variable. These are also called n-step ahead forecasts. Whereas, the static forecasting uses actuals rather than forecasted values, but it can only be used when actual data are available. These are also called 1-step ahead or rolling forecasts.

Furthermore, there are two possible procedures for analyzing or comparing the quality of forecasts: Divide available time series into two subperiods:

- in-sample period for estimation of parameters where the period should be long enough to yield well-estimated coefficients;
- out-of-sample period (for investigating the quality of the forecasts) where the period should be long enough to have good power when analyzing or comparing the quality of the forecasts.

Then the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) are simply given by:

$$MAE = \frac{\Sigma |e_t|}{n} \tag{7}$$

$$RMSE = \sqrt{\Sigma \frac{e_t^2}{n}} \tag{8}$$

where et = Yt - ft = forecast errors
(Yt = actual values; ft = forecast values)
and n = number of forecasts.

They both measure how close forecasts or predictions are to the eventual outcomes, but the use of RMSE is more common because compared to the MAE, RMSE amplifies and severely punishes large errors.

The last measure to be used in the case of the GARCH model comparisons is the Theil Inequality Coefficient :

$$U = \frac{\sqrt{\frac{1}{n}\Sigma(Y_t - f_t)^2}}{\sqrt{\frac{1}{n}\Sigma f_t^2} + \sqrt{\frac{1}{n}\Sigma Y_t^2}} \tag{9}$$

If U = 0, Yt = ft for all forecasts and there is a perfect fit. The Theil Inequality Coefficient can be decomposed into three proportions of inequality - bias, variance and covariance. The sum of these three proportions is 1. The interpretation of these three proportions is as follows:

- Bias: Indication of systematic error. Whatever the value of U, we would hope that bias is close to 0. A large bias suggests a systematic over or under prediction.
- Variance: Indication of the ability of the forecasts to replicate the degree of variability in the forecasted variable.
- Covariance: This proportion measures unsystematic error and ideally, this should have the highest proportion of inequality.

From this measure, the change of bias and variance proportions between the simple GARCH and GARCH with log(Var) will be used to evaluate the forecasting power of those models for different purposes.

As a last step after evaluating the models based on the above-described measures, we are interested in the forecasting power of the models. One has to take into consideration that the used dataset does not have sufficiently enough number of observations to make a solid split between in-sample and out-of-sample. Therefore, the rolling window method will be used because it performs better due to the structural changes in the estimation period. One has to make the deliberation that the sample is not large enough to get a conclusive outcome only based on the rolling window method.

Regression Outputs

As explained before, the analysis of the forecasting models starts with using the EU aggregate data, and then the chosen model will be modified and/or changed to account for county level differences. The ratios and measures introduced in the previous section will be examined sequentially and the model will be evaluated based on those. In cases where a GARCH model will be used, the R squared will be avoided as those cannot be interpreted in that context.

For modelling purposes I needed to transform all data to stationary data. I did so with the EU HICP by taking the log of all observations and included a "@trend" term to de-trend the data, and by doing so I can reject the ADF test hypothesis (LOG1 (=log(HICP EU) has a unit root) on a 10% significance level. To ensure stationarity I also conducted a KPSS test where the hypothesis is that LOG1 is stationary which cannot be rejected (See Appendix A.1 for the test results). I have repeated the same steps for all macroeconomic variables for both EU and country levels.

After making the necessary transformations for both dependent and explanatory variables, I started with the EU models. The simplest model is an ARMA(1,1) model without any macroeconomic variables (See Appendix B.1). Since the dataset runs on a monthly basis, I have regressed 12 lags from both AR and MA terms separately. Both models (AR(12) and MA(12)) work better than the simple ARMA(1,1), however as the data is highly persistent the results of these models can be further improved by combining the models. By doing so, most of the AR and MA terms become insignificant. Systematically deleting the insignificant AR and MA terms one by one results in Combined ARMA model with AR(1), AR(6), AR(7), MA(6) and MA(12) terms staying significant (See Appendices B.2 – B.4). This represents the

seasonality characteristic of inflation and the fact that the inflation expectations usually change over a course of six months.

After deciding on the time series terms to include in the model, I have introduced the macroeconomic variables one by one and keeping the significant ones. This resulted in a model where short-term interest rate and M3 are significant, which is not a surprise since there is a long lasting relationship between those variables and inflation (See Appendix B.5).

So far, all measures and ratios used to evaluate the models have improved with the above models. Although these models have a relatively good forecasting power, we have not yet taken into consideration the variance change over time. Actually one of the underlying assumptions of ARMA models is that variance constant over time, which is clearly not the case (See Chapter 3 page 24). This can be validated through having a Simple GARCH model, where without any macroeconomic variables or further explanatory variables. The Simple GARCH model, where the GARCH term in the Variance Equation is highly significant, performs already better on the basis of AIC and SBC as well as the static RMSE and MAE measures (See Appendix B.6).

The Simple GARCH model can be further corrected with introducing the significant macroeconomic variables. This refines both information criteria measures and forecasting ratios (See Appendix B.7). The results so far are summarized in Table 4 below:

Model	Equation	Adjusted R-squared	AIC	SBC	RMSE (dynamic and static)	MAE (dynamic and static)
ARMA(1,1)	log(HICP) = c + @trend + AR(1) + MA(1)	0.998895	-8310474	-8236643	0.009546 / 0.003706	0.007222 / 0.002641
MA(12)	log(HICP = c + @trend + MA(1) + ... + MA(12)	0.999301	-8696461	-8474966	0.009400 / 0.002905	0.007078 / 0.002284
AR(12)	log(HICP = c + @trend + AR(1) + ... + AR(12)	0.999291	-8696937	-8475442	0.011295 / 0.002959	0.009022 / 0.002236
Combined ARMA	log(HICP) = c + @trend +AR(1) + AR(6) + AR(7) + MA(6) + MA(12)	0.999563	-91725	-905437	0.009007 / 0.002348	0.006284 / 0.001804
Combined ARMA with macroeconomic variables	log(HICP) = c + @trend +AR(1) + AR(6) + AR(7) + MA(1) + MA(6) + MA(11) + MA(12) + log(ST) + log(M3)	0.999577	-9207504	-9028126	0.008543 / 0.006309	0.006309 / 0.001762
Simple GARCH	log(HICP) = c + @trend + AR(1) + AR(6) + AR(7) + MA(6) + Ma(11) + MA(12)	N/A	-9267258	-9104828	0.011437 / 0.002343	0.008467 / 0.001819
GARCH with macroeconomic variables	log(HICP) = c + @trend + AR(1) + AR(6) + AR(7) + MA(1) + MA(6) + MA(11) + MA(12) + log(ST) + log(M3)	N/A	-9325351	-9131025	0.010159 / 0.002303	0.008081/ 0.001816
GARCH with log(Var)	log(HICP) = c + AR(1) + AR(6) + AR(7) + MA(1) + MA(6) + Ma(11) + MA(12) + log(GARCH)	N/A	-9239525	-9047563	0.010398 / 0.001849	0.00787 / 0.001849
GARCH in Mean with log(Var) wth macroeconomic variables	log(HICP) = c + AR(1) + AR(6) + AR(7) + MA(1) + MA(6) + MA(11) + MA(12) + log(GARCH) + log(ST) + log(M3)	N/A	-1567776	-1537535	0.027821*	0.024784*

Table 4: EU Regression Models
*log of non positive number prohibits the static forecasting.

Until now the focus was shifted towards explaining inflation and forecasting within those frameworks. Now we turn to explaining the variance of inflation. This can be done by including the GARCH component in the mean equation as explanatory variable. Since all variables have gone through a log transformation, the GARCH component will be introduced the same way as log(GARCH). This model does perform similarly as the GARCH with macroeconomic variables (See Appendix B.8.1), however, based on the performance of forecasting HICP I select the best performing model as the model without the log(GARCH) component, as it slightly worsens the forecasting power.

Although the coefficient of the log(GARCH) component is not significant over the whole sample, it becomes significant when I break the sample into two subsamples. This gives a motivation to further update the model. As a next step I introduced the same significant macroeconomic variables from the GARCH model. This creates a model where the information criteria are almost doubled, and the forecasting evaluation measures are lowered significantly (See Table 4). Still based on purely the forecasting measures the GARCH model with macroeconomic variables is the better one.

On the other hand, the reason of including the GARCH component in the mean equation was to explain the variance of inflation better. Therefore, we turn to the Theil Inequality Coefficient as evaluation measure. The overall coefficient stays approximately stable, but the distribution of the three proportions change when looking at the GARCH model and the GARCH in mean model (See Table 5).

Measure	GARCH with macroeconomic variables (dynamic / static)	GARCH with log(Var) with macroeconomic variables (dynamic)	GARCH with log(Var) (dynamic / static)
Theil Inequality Coefficient	0.001124 / 0.000255	0.003070	0.001150 / 0.000260
Bias Proportion	0.000001 / 0.000004	0.793188	0.240211 / 0.018096
Variance Proportion	0.011423 / 0.002220	0.093699	0.000053 / 0.000475
Covariance Proportion	0.988576 / 0.997776	0.113113	0.759736 / 0.981439

Table 5: Mean in GARCH compared

Therefore, depending on the purpose of the forecast one need to judge which model to choose. Between the two GARCH in mean models, the one without macroeconomic variables have a more reliable distribution of the proportions, as the covariance proportions should be the highest of all three. The GARCH in mean model with macroeconomic variables has a very high Bias Proportion, and both RMSE and MAE are lower than the other models.

The results outlined in Figure 18 and 19 indicate that macroeconomic variables do have a powerful impact on forecasting accuracy. As the historical data and the tests and analysis in Chapter 3 showed that there is a clear change in inflation variance over time, the GARCH models are the appropriate models to choose and depending on the purpose of that forecast one has to decide between the GARCH or GARCH in mean equations.

After completing the evaluation of the models with information criteria and in-sample forecasting, the final step in the Euro Area-wide modelling is to evaluate the performance of the models with out-of-sample forecasts. Dividing the sample into two periods (For details see Chapter 4 Research Design section) and comparing the forecasted values to the actual ones gives a good indication how well the model performs. The purpose of this is that beyond the coverage of the data there is no way of evaluating the forecasting power of models, therefore treating the second sub-sample as out-of-sample and putting those next to the actual values can help deciding on the forecasting model.

The breaking point is 2008M06, and the period after is treated as out-of-sample. With dynamic forecasting in the GARCH in mean model, the actual inflation levels are within the forecasted 95% confidence interval. That indicates that even in a period where the financial crisis created a very uncertain environment a model allowing for variance change over time and having macroeconomic variables as explanatory variables in the equation the model performs well and its use is realistic (See Appendix C.1.1). We usually are interested in a one-step ahead forecast, as this gives a short-term future outlook for the upcoming period. This can be done with a rolling window forecast. I repeat the same steps as before, but instead of forecasting until 2015M06, I only forecast it until 2009M06 (See Appendix C.1.2). Then I repeat it twice with shifting the in-sample period one-year ahead and forecast with that period again one additional year (See Appendix C.1.3). This results in a more robust forecasting method, and all actual values are within the forecasted 95% interval for all three consecutive rolling window periods. For static forecasting I have followed the same methodology, but in the regular GARCH model, as the log of non positive numbers prevent the one year ahead static forecasting in the GARCH in mean model.

After evaluating and deciding which model to use for forecasting HICP in the Euro Area, we can turn to the country level models. As an initial step, I have regressed the AR(12) and MA(12) models and then combined them together in an ARMA model. Already these results show significant differences to the EU ARMA model, as well as between each country. This indicates that inflation expectations changes have a different pace and intensity over the course of a year in every country making the Euro Area-wide forecasting harder on an aggregate level.

Then I have introduced the macroeconomic variables in each equation separately. This step resulted in an interesting conclusion. Every country has different significant macroeconomic variables within the ARMA framework compared to each country level equation and also compared to the aggregate EU level equation. The differences in significance levels, and the diversity of significant macroeconomic variables mirror the complexity of the economy of the Euro Area. Each and every country has contrasting environments and dynamics within, therefore I have decided to model the country level equations separately and not by modifying the aggregate EU model.

Moreover when introducing the GARCH model, only in France and in Portugal are the coefficient significant. This means that in other countries, like Denmark, UK and Germany, where inflation levels are more stable, the variance change over time is not significant. Hence in those countries the model of choice is a Combined ARMA model with macroeconomic variables. In the case of France and Portugal, I continued the process by looking at the GARCH in mean models, however, just as in the case of the EU aggregate model, the log(GARCH) component worsens the performance of inflation forecasting, and delivers an improvement to the variance change forecasts. In the case of Denmark, the GARCH component turns out to be slightly significant but only on 10% level, and when dividing the sample into sub-periods the effect disappears.

The signs of all significant macroeconomic variables in each country coincide with the signs of the variables in the aggregate EU forecasting. This demonstrates that the aggregate level forecasting replicates the overall country level economic dynamics but on a different scale. The table below summarizes the significant variables in each equation and the significant AR, MA, GARCH and GARCH in mean coefficients, if applicable (Table 6):

Explanatory Variable Significance	Denmark	Germany	UK	France	Portugal	EU
@trend	+*	+*	+*	+*	+*	+*
HUR	0	0	0	0	0	0
ULC	0	0	0	+*	0	0
M1	0	-*	-*	0	0	0
M3	0	-*	0	0	0	-*
Long Term Interest Rate	0	0	0	-*	0	0
Short term Interest Rate	0	0	0	+*	0	+*
GDP	-*	0	0	0	0	0
AR lag terms	1,2,3,4,5,6,11,12	1,2,3,5,6,7,8,10,11	3,4,5,6,9,10,11,12	6	1,2,6,11	1,6,7
MA lag terms	1,2,3	2,5,6,7,10,12	1,2,3,4	1,2,3	-	6,11,12
Garch	+*	0	+*	+*	+*	+*
Log(var)	0	0	0	+*	0	+*

Table 6: Significance of explanatory variables

Comparison of the models between the countries are not necessarily possible in this setup, so I have summarized the evaluation measures and ratios in the table below. As we can see, all measures are in line with each other between the different country models. The difference comes in when we consider how many explanatory variables we need to get to those similar levels (See Table 7). In the cases, where the Combined ARMA models have been chosen, there are many AR and MA terms significant (See Appendix B.9 – B.11). In the case of France and Portugal, the GARCH model delivers a comparable set of measures, however in the Portuguese model there are no significant macroeconomic variables present, whereas in the French model, ULC, long- and short-term interest rates are significant. This sheds light on clear heterogeneity within the EU economies and calls for a more detailed explanation why the country level models differ from the EU aggregate model.

When we look at the results more carefully, it is clear that the French and Portuguese economies are in a more difficult position and are in a different stage of development than the other three countries. The significant GARCH coefficient indicates (See Appendix B.12 and B.13) that in stressed economies the uncertainty is more significant than in countries with a more stable inflation. This summarizes the difficulties around the EU inflation forecasting, because one has to take into account not just the different macroeconomic variables, but also the variance changes over time in certain countries.

	Denmark	Germany	UK	France	Portugal
Model of Choice	Combined ARMA model with macroeconomic variables	Combined ARMA model with macroeconomic variables	Combined ARMA model with macroeconomic variables	GARCH model with macroeconomic variables	GARCH model with macroeconomic variables
Adjusted R squared	0.999362	0.998834	0.999566	N/A	N/A
AIC	-8874923	-8714690	-8944609	-8692085	-8127236
SBC	-8652753	-8381370	-8707628	-8512707	-8009105
Dynamic RMSE	0.009566	0.007311	0.028985	0.007342	0.024009
Dynamic MAE	0.007462	0.005737	0.024993	0.0058888	0.018937
Static RMSE	0.002692	0.002793	0.002522	0.003149	0.004525
Static MAE	0.002190	0.002164	0.002018	0.002417	0.003202

Table 7: Country Level Models of Choices

The modelling process continued with out-of-sample forecasting (both dynamic and static) on country levels. In all cases the static forecasting performs better and follows the actual values more closely. As the breaking point of the sub-sample is in the middle of the financial crisis, the dynamic forecasting depicts those dynamics for the future. In Denmark, and in Germany where the economy was stable enough to tackle the uncertain environment faster, the dynamic forecasting still performs well for the entire forecasted period. In the case of United Kingdom, the actual inflation started to increase after the crisis, the dynamic model cannot depict the entire period correctly. While in France and in Portugal where the inflation levels dropped significantly due to the financial crisis, the actual inflation levels are above the upper bound of the forecasted ones (See Appendix C.2 – C.6). Since the ECB and also the countries have introduced measures to boost the economy, and amplify the recovery form the financial crisis, the static models perform more accurately.

Investors and institutions can look at inflation level announcements and use those to re-adjust the forecasting models hence for those market players the static forecasting can be a more appropriate one. Whereas central banks have to rely on dynamic forecasting, but they have the tools to influence the overall economy in a greater detail, so they are able to incorporate those changes and future expectations into the model.

Robustness Checks

A common exercise after conducting the modelling part of empirical studies is a "robustness check". It examines how certain regression coefficients behave when the regression specification is modified. Adding or removing certain variables can change the core coefficients signs and significance levels. However, if those core coefficients do not change much, it can be interpreted as plausible, reliable and robust results. If so, the model can be a basis for policy analysis and economic insight.

I have done so by adding three extra variables in each of the chosen models for both on aggregate EU level, and on country level. As the HICP data clearly exhibits a seasonal behaviour (See Graph 13), I have defined a binary dummy variable. The data has seasonal peaks in January and March therefore I have introduced two separate dummy variables. I set one variable with M01 as 1, another variable with M03 as 1, and all other months as 0 to take care of seasonality. The second extra variable is an oil price index, since an oil price increase/decrease have an immediate effect on both oil importer and exporter countries, and consequently on inflation. The last variable for robustness check is the exchange rate between the EUR and USD.

One drawback of this robustness check, that there was no available exchange rate from 1996, only from 1999, therefore there are only 197 observation points instead of the original 234. Still the results indicate that the inclusion of these variables do not lead to significant improvement nor change of the model. In some cases, those variables have significant coefficients, however, the measures used to judge the models do not improve significantly. The test results are summarized in Table 8 below (For more details See Appendix D)

Extra Variable	EU	Denmark	Germany	United Kingdom	France	Portugal
Seasonality Dummy (M1)	-*	-*	-*	-*	0	-*
Seasonality Dummy (M3)	0	0	0	0	+*	-*
Oil Price Index	0	0	+*	0	0	0
Exchange Rate	0	0	0	0	0	+*
Significant Improvement?	No	No	No	No	No	No
Change of Model?	No	No	No	No	No	No

Table 8: Robustness Check

None of the results showed significant improvement. The information criteria did not improve, nor did R-squared and in most cases including the extra variables for robustness check worsened the models. The signs of the core coefficients did not change, and as the inclusion of the four extra variables do not lead to a significantly better model in any of the cases (EU and country level), the models should not change.

After introducing the regression and forecasting outputs, we now can turn to answer the research questions of this paper. The last section in this Chapter will summarize the interpretation of the results in the context of the aggregate EU model and the country level models.

Interpretation of Results

"Which macroeconomic factors affect the inflation level and its volatility and how?" has been the main questions of this paper. To answer this question, we need to distinguish between the Euro Area aggregate forecasting and the country level forecasting. Overall, money supply (M1) and the short-term interest rate have a clear effect on the inflation levels. Furthermore, the GDP changes in Denmark, the M1 and M3 in Germany, the M1 in the United Kingdom, the ULC, and the long- and short-term interest rates in France are significant on country level. All models are able to forecast the future inflation levels within a set confidence interval even in a difficult period like in the middle of the financial crisis. As very little attention is dedicated to the volatility changes over time, it is important to look at country

level uncertainties, which turned out to be the focal point of the overall volatility change on the aggregate EU level.

The similarities and differences in inflation patterns across Euro countries indicate that the analysis needs to go beyond the common economic factors on the aggregate level in order to be able to potentially explain the future inflation levels on country level. One similarity is that all macroeconomic variables have the same sign of coefficients, whereas the differences stem from the fact that in each country the significance levels of macroeconomic variables and the absence of significant GARCH coefficient in certain countries differ.

These results after the robustness check still hold therefore can be interpreted as reliable in this framework. However, another angle of this analysis is the practical application of these models. As I used publicly available information, professional forecasters who have more access to proper data have the opportunity to improve the models further.

Chapter 5: Discussion

Implications

As we have seen in the previous section with the regression outputs, there are clear differences between the aggregate Euro-wide level model and the country-specific models. These differences stem from the fact that the economies of those countries are in different stages, especially after the recent financial crisis those differences are more apparent.

According to the most recent OECD Economic Survey about the Euro Area, the weak private sector balance sheets and impressive financial consolidation after the crisis resulted in a weakened effect of the supportive monetary policy steps. The financial fragmentation, high level debts and low confidence resulted in inflation rates falling further down. This subdued outlook is to be extended to the medium term with the ECB´s monetary policy staying accommodative.

The mentioned financial fragmentation can only be solved with the case of establishing of a banking union, so as the recovery advances, deflation risks can disappear. This, however, does not mean that other risks and difficulties will be solved on country level as well. Unemployment rates stand at very high levels in several countries, and inflation being very low in many countries.

example case of Denmark, where inflation has been low and stable for the past years due to sound policy decisions and institutions, or the United Kingdom where monetary accommodation was successfully implemented so inflation has fallen below 2% and wage pressures disappeared. Solid country level monetary decisions in these well-developed countries will guide the withdrawal of monetary support dependent on economic developments with maintaining inflation expectations anchored. In these countries there is no explanation of the aggregate Euro-wide concerns and difficulties which the ECB faces overall. The reasons should be looked in countries with specific challenges.

In France, where weak growth, unfavourable labour market conditions and recent decline in food and energy prices have resulted in a negative pressure on inflation, core inflation reaching very low levels. Furthermore, in countries like Portugal where the debt reduction burden is tight and the highly-indebted corporate sector being unable to cover their debt service expenses, the recovery will be even more difficult in a deflationary or very low inflation level period.

Also Germany faces a subdued inflation rate for a longer period with declining import prices and a negative output gap. This output gap might be closing as the OECD report does not include the recent impact of increased immigration.

Reading through these OECD Economic Surveys, it is clear that the challenges the Euro Area faces stem from country specific economic developments and risks. Therefore my findings are in line with research paper by Marcellino et al (February, 2002): That forecasts constructed by aggregating the country-specific models are more accurate than forecasts constructed using aggregate data (Marcellino et al, 2002).

Continuing fiscal consolidation with national fiscal framework are part of the expansionary monetary policy stance over an extended period. This is subject to the outlook for price developments over the medium term, hence forecasting HICP levels, and the price levels are key for ensuring a sustained recovery in the Euro Area. The ECB´s response to these developments will be the underlying step to the quantitative definition of price stability.

Further study

My main focus here is to improve the forecast accuracy of inflation, which is a difficult task, particularly on country level, due to data availability. Fostering economic recovery includes some forms of consolidation; hence a consolidation of country-wide databases is called for. This would allow analysts to dig deeper in the relationships between macroeconomic factors across borders and observing the impact of these on the Euro Area-wide aggregate level. The scope of this work stops here. Analysing the different macroeconomic variables and modelling inflation on both aggregate and some country level raises new questions which can be subject of further analysis and research.

All models used in this paper are restricted to univariate models. With better data availability the scope of the analysis could be expanded to all countries connecting the models together and analysing the overall effects to the Euro Area-wide level. This would also allow others to explore the possibility of comparing other models in use for inflation forecasting. In the introductory section, I have mentioned that there is a number or research done on inflation forecasting methods in the United States, which could be replicated in the Euro Area.

Another factor, which has not been included in the scope of this thesis, is the effect of exchange rate on inflation. We know that relatively higher inflation levels tend to reduce

the value of the currency. High inflation in a certain country means that goods increase in price quicker than other European goods. Therefore the country's goods become less competitive. Demand for exports fall, resulting in less demand in the currency. This depreciation in the exchange rate should cause inflation to increase. These dynamics work both ways, hence the foreign exchange rates can have a significant effect on country level inflation when examining cross-border influences.

This leads to another point for further study. The direct source of GARCH effects in the EU aggregate model has not been fully explored. Linking the country-level variance changes over time to the EU aggregate model would give more insights into the dynamics of the overall price level in the EU. This could lead to a broader cross-country influence analysis, which would give better insights into the push and pull effects of different macroeconomic variables across countries and on the overall EU inflation levels.

There is a shortcoming of the used model on the aggregate EU level. The log difference approach which on one takes care of stationarity, and the other hand normalizes all datasets, does not all to conduct a static forecasting. As this is a caveat of the methodology, future work could include different estimation techniques to compare all models with all possible forecasting methods.

To sum up, improvements to data availability, influences of country level macroeconomic variables on the aggregate EU level and cross country impacts could be further explored. Also, looking into the country level differences, finding other explanatory variables to improve the models and expanding the scope of the modelling to all European countries would give an opportunity to have a full overview of inflation dynamics in the EU. Another possibility is to compare the performance of multivariate models, and models predicting and assigning probabilities to volatility changes. These could help answer the questions around the impact of volatility changes on inflation expectations and help forecasters set a benchmark to measure how severe the change would be.

Chapter 6: Conclusion

After the recent financial crisis, not only the level of inflation, but also its volatility became important. Determining the main drivers of this inflation volatility and forecasting can help decision makers to more accurately link future expectations to real levels of inflation. The European Union and the Euro Area consist of several very diverse economies in different stages. Therefore, there is a need of disaggregation to country level analysis and observation how the different country specific trends and macroeconomic variables contribute to the overall inflation levels in the European Union.

The research idea focused on the determinants of inflation levels and inflation volatility in the Euro Area. The attempt was to find an alternative model to determine the levels of inflation in the Euro Area through focusing on its volatility and on the changes of this volatility over time. The central question of the thesis was: *"Which macroeconomic factors affect the inflation level and its volatility and how?"*

To answer this question, we needed to distinguish between the Euro Area aggregate forecasting and the country level forecasting. Overall, money supply (M1) and the short-term interest rate have a clear effect on the inflation levels. Furthermore, the GDP changes in Denmark, the M1 and M3 in Germany, the M1 in the United Kingdom, the ULC, and the long- and short-term interest rates in France are significant on country level.

Another decision was about static or dynamic forecasting. Investors and institutions can look at inflation level announcements and use those to re-adjust the forecasting models hence for those market players the static forecasting can be a more appropriate one. Whereas, central banks have to rely on dynamic forecasting.

However, all models were robust and able to forecast the future inflation levels within the set 95% confidence interval even in a difficult period like in the middle of the financial crisis. Very little attention was dedicated to the volatility changes over time, so the paper looked at country level uncertainties, which turned out to be the focal point of the overall volatility change on the aggregate EU level.

One similarity was across all models within the examined countries and the aggregate EU level that all macroeconomic variables have the same sign of coefficients.

Whereas, the differences stem from the fact, that in each country the significance levels of macroeconomic variables and the absence of significant GARCH coefficient in certain countries differ.

To sum up, there are clear differences between the aggregate Euro-wide level model and the country-specific models. These differences stem from the fact that the economies of those countries are in different stages, especially after the recent financial crisis those differences are more apparent, therefore the forecasting should incorporate those differences and take into consideration when looking at Euro Area-wide forecasting on an aggregate level.

References

Altissimo, F., Bilke, L., Levin, A., Mathä, T., & Mojon, B. (2006). Sectoral and aggregate inflation dynamics in the Euro Area. Journal of the European Economic Association, 4(2- 3), 585-593.

Andersen, T. G., & Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 885-905.

Ang, A., Bekaert, G., & Wei, M. (2007). Do macro variables, asset markets, or surveys forecast HICP better?. *Journal of Monetary Economics*, *54*(4), 1163-1212.

Angeloni, I., & Ehrmann, M. (2007). Euro Area inflation differentials. *The BE Journal of Macroeconomics*, *7*(1).

Basturk N., 2014: EBC4097 – Quantitative Methods for financial economics Part 2, class notes.

Basturk, N., Cakmakli, C., Ceyhan, P., & van Dijk, H. K. (2013). *Posterior-Predictive Evidence on US Inflation using Phillips Curve Models with Non-Filtered Time Series* (No. 13-011/III). Tinbergen Institute Discussion Paper.

Beck, G. W., Hubrich, K., & Marcellino, M. (2009). Regional inflation dynamics within and across Euro Area countries and a comparison with the United States. Economic Policy, 24(57), 142-184.

Bernanke, B. S. (2004). *The great moderation.* Remarks by Governor Ben S. Bernanke at the meetings of the Eastern Economic Association, Washington, DC February 20, 2004. Eastern Economic Association, Washington, DC, 20.

Bernanke, B. S., & Woodford, M. (1997). *Inflation forecasts and monetary policy* (No. w6157). National Bureau of Economic Research.

Bos, C. S., Franses, P. H., & Ooms, M. (2002). Inflation, forecast intervals and long memory regression models. *International Journal of Forecasting*, *18*(2), 243-264.

Bowerman and O'Connell, 1993: Forecasting and Time Series: An Applied Approach, Third Edition, The Duxbury Advanced Series in Statistics and Decision Sciences, Duxbury Press, Belmont, CA. ISBN: 0-534-93251-7

Carlin, W., & Soskice, D. (2005). Macroeconomics: imperfections, institutions, and policies. OUP Catalogue.

Charles, A., & Darné, O. (2009). Variance ratio tests of random walk: an overview. Journal of Economic Surveys, 23(3), 503-527.

Chen, J. H. (2011). Variance ratio tests of random walk hypothesis of the Euro exchange rate. International Business & Economics Research Journal (IBER),7(12).

Cristadoro, R., Venditti, F., & Saporito, G. (2008). *Forecasting HICP and tracking monetary policy in the Euro Area: does national information help?*. ECB Working Paper Series No 900 / May 2008, online at: <http://eulib.com/documents/ecbwp900.pdf>

Faust J, Wright JH. 2013. Forecasting HICP. In Handbook of Economic Forecasting, Elliott G, Timmermann A (eds), Vol. 2. Elsevier: Amsterdam; 2–56.

Fischer, S. (1993). The role of macroeconomic factors in growth. *Journal of Monetary Economics*, *32*(3), 485-512.

Forni, M., Hallin, M., Lippi, M., & Reichlin, L. (2003). Do financial variables help forecasting HICP and real activity in the Euro Area?. *Journal of Monetary Economics*, *50*(6), 1243-1255.

Giordani, P., & Söderlind, P. (2003). Inflation forecast uncertainty. *European Economic Review*, *47*(6), 1037-1059.

Groen, J. J., Paap, R., & Ravazzolo, F. (2013). Real-time inflation forecasting in a changing world. *Journal of Business & Economic Statistics*, *31*(1), 29-44.

Hall, J., & Jääskelä, J. P. (2011). Inflation volatility and forecast accuracy. *Australian Economic Review*, *44*(4), 404-417.

Koop, G., & Korobilis, D. (2012). Forecasting HICP Using Dynamic Model Averaging . International Economic Review, 53(3), 867-886.

Lu, X., & White, H. (2014). Robustness checks and robustness tests in applied economics. Journal of Econometrics, 178, 194-206.

Lünnemann, P., & Mathä, T. Y. (2004). How Persistent is Disaggregate Inflation?: An Analysis Across EU 15 Countries and HICP Sub-indices. Europ. Central Bank.

Lünnemann, P., and T. Mathä. "How Persistent is Disaggregated Inflation." An Analysis (2004).

Maddala, G. S. (1998). In-Moo Kim (1998) Unit Roots, Cointegration, and Structural Change.

Marcellino, M., Stock, J. H., & Watson, M. W. (2003). Macroeconomic forecasting in the Euro Area: Country specific versus area-wide information. *European Economic Review*, *47*(1), 1-18.

Mario Draghi, President of the ECB, *"Monetary policy in a prolonged period of low inflation"* at the ECB Forum on Central Banking (Sintra, 26 May 2014)

McFadden, D. (1999). Second Half Reader. Econometrics 240B: Chapter 7. *University of California Berkeley.*

Özkan, H., & Yazgan, M. E. (2014). Is forecasting HICP easier under inflation targeting?. *Empirical Economics* (January 2014), 1-18.

Rother, P. (2004). *Fiscal policy and inflation volatility*. ECB Working Paper Series NO. 317 / MARCH 2004, online at:
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.165.2953&rep=rep1&type=pdf>

Stock, J. H., & Watson, M. W. (1999). Forecasting HICP. *Journal of Monetary Economics*, *44*(2), 293-335.

Stock, J. H., & Watson, M. W. (2007). Why has US inflation become harder to forecast?. *Journal of Money, Credit and banking*, *39*(s1), 3-33.

Stock, J. H., & Watson, M. W. (2010). Modeling inflation after the crisis (No. w16488). National Bureau of Economic Research.

Wu, J. L., Chen, S. L., & Lee, H. Y. (2003). Sources of inflation uncertainty and real economic activity. *Journal of Macroeconomics*, *25*(3), 397-409.

Žďárek, V., & Aldasoro, J. I. (2009). *Inflation differentials in the Euro Area and their* determinants: an empirical vies (No. 450). Kiel advanced studies working papers.