TESTING THE PREDICTIVE POWER OF TERM SPREAD IN THE EURO AREA

Jyväskylä University
School of Business and Economics

Master's Thesis

2021

Author: Sachindra Rimal
Economics Banking and International Finance
Supervisor: Juha-Pekka Junttila

JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

ABSTRACT

Author Sachindra Rimal	
Title Testing the Predictive Power of Term Spread In the Euro Area	
Subject Finance	Type of work Thesis
Date 25.1.2021	Number of pages 59

This thesis tests the predictive power of term spread in predicting the Euro area's real economic activities. The objectives of this study are to test the predictive power of term spread in the negative interest rate period in the Euro area, to examine the joint predictive power of term spread and EPU, and to reveal the Granger causality of the variables.

Term spread and GDP growth rate are significant variables; however, the term spread model is augmented with EPU. Term spread is derived from the three-month interest rate and triple 'A' rated ten-year government bond. The sample of this thesis ranges from 1999Q1 to 2019Q4. The in-sample model fit is tested with the full sample data, and the out-of-sample prediction is tested using the data before the negative interest rate period in the Euro area. The vector autoregression method is used in this study; furthermore, a linear model is estimated using some dummy variables such as the financial crisis 2008-9, high uncertainty period, and negative interest rate period.

The following are the five most significant findings of this thesis. First, the predictive power of term spread is low, but it has slightly increased during the negative interest rate period. Although term spread's predictive power is increasing, the estimate coefficients of term spread are not statistically significant yet. Such a low predictive power of the term spread is found in Germany, Italy, Spain, Belgium, Ireland, and Finland. Only in France, term spread has significant predictive power. Second, the relatively low predictive power of term spread is observed particularly during the recession caused by the European sovereign debt crisis and during the high uncertainty period. Third, the lags of GDP growth rate have better predicting power than the term spread has. The model's adjusted R^2 decreases by only 0.01 when term spread is removed from the independent variables, but the adjusted R^2 drops from 0.93 to 0.61 as the lags of GDP are removed from the independent variables, indicating that the real economic activities in the Euro area can be better predicted by GDP growth rate's lags than by term spread. Fourth, the estimate coefficients for EPU are almost zero and it cannot increase the model's predictive power either. Last, term spread Granger causes GDP growth in lower lags, optimally at lag two. A fragile form of bidirectional Granger causality between term spread and GDP growth rate is observed, while EPU does not Granger cause the GDP growth rate at all.

Keywords: Term Spread, Yield Curve, Forecasting, Predicting, Real Economic Activity, VAR model

Place of storage
Jyväskylä University Library

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

1 INTRODUCTION

Term spread, also known as the slope of the yield curve, is one of the most popular indicator variables in the macro-financial analysis. Indeed, it has been extensively used to forecast future real economic activities in several empirical studies (Harvey, 1988; Estrella and Hardouvelis, 1991; Estrella and Mishkin, 1997; Kozicki, 1997; Pena et al., 2006; Papadamou, 2009; Schunk, 2011; Dar et al., 2014; Hyozdenska, 2015a). Many studies have confirmed that term spread has been handy in predicting real economic activity. The predictive relationship between term spread and real economic activity is so well established that it is considered a stylized fact in financial economics. However, the research findings examining the predictive power of term spread have been time-varying at different times (Bismans and Majetti, 2011; Jardet, 2004; Morrel, 2018; Dong and Park, 2018; Kuosmanen, Rahko, and Vataja, 2019; Karlsson & Osterholm, 2020). If the predictive power of term spread can differ from time to time, then what about its predictive power for now?

The exact reasons for the time-varying predictive power of term spread are unknown, but some evidence suggests that its predictive power can depend on the variables added to the forecasting model based on term spread and the method employed in the study. For example, the results of Chionis and Gogas (2010), Gogas and Pragidis (2011), Kuosmanen and Vataja (2017), and Chen, Valadkhani, and Grant (2016) suggest that the predictive power increases as the term spread model is augmented with other variables. Also, the results from some studies (Paya et al., 2004; Pena and Rodriguez, 2006; Evgenidis and Siripoulos, 2014; Gogas et al., 2015; Gogas et al., 2015b; Gupta et al., 2020; Evgenidis, Papadamou, and Siripoulos, 2020) suggest that the methods employed in the study can impact the predictive power of term spread.

In the last two decades, the euro area economies have experienced several significant economic shocks from a series of economic events, such as the financial crisis 2008-9, the recession caused by the European sovereign debt crisis, and the negative interest rate period. Also, the level of economic uncertainty has risen sharply. As the level of economic uncertainty rises, the economic agents become more desperate to know about the economy's future. At this point, economic forecasting gets substantial attention, so does the predictive power of term spread. An accurate prediction of changes in future economic activities can undoubtedly be useful for economic agents and policymakers to make their economic decisions efficiently. For example, policymakers can be well prepared for anticipated changes that will happen in the economy, households can make plans to smooth their consumption when an uncertain future is anticipated, and businesses can plan for investment timings to avoid uncertainties in expected cash flows. In this regard, the predictive power of term spread to predict the future of the economy gets significant attention and importance. All of the above mentioned three significant economic periods have affected the Euro-area economy, and these could also have altered the predictive power of term spread.

First, the financial crisis 2008-9 originated from the United States following the bankruptcy of Lehman Brothers, and then it spread to Europe. The recession in

Eurozone turned severe from the fourth quarter of 2008 to the fourth quarter of 2009, resulting in five consecutive quarters with a negative GDP growth rate. Analyzing the quarterly GDP growth rate from OECD (2020) data for the Euro area, a closer picture of the severity of the recession can be explained clearly that the Eurozone quarterly GDP growth rate fell by 2.133% in the fourth quarter of 2008, and the growth rate further dropped by 5.647% in the first quarter of 2009 from the previous quarter. Further, the decline continued until the fourth quarter of 2009 by 5.36%, 4.49%, and 2.33%, respectively (adapted from OECD 2020). This financial crisis certainly gave a massive setback to the Eurozone economy. Second, the recession caused by the European sovereign debt crisis began after the collapse of Iceland's banking system in 2008. During this recession, some European economies such as Portugal, Italy, Ireland, Greece, and Spain experienced highly deepening risk positions of some financial institutions, steeply increasing government debt ratios, and rapid widening of government bond term spreads. The crisis led to a situation in which Greek, Portuguese, and Irish government bonds received a junk bond status from international credit rating agencies, making those countries harder to finance their budget deficits. There was a fear of financial contagion to other EU economies, which even led to the fear of the whole Euro system's collapse. As a solution to this crisis, the EU countries and the International Monetary Fund provided financial guarantees for the affected countries and controlled the crisis before it was too late. However, this crisis caused tax raises that created socio-political unrest in affected countries and stimulated investors' fears of European economies. Overall, this crisis definitely increased the economic uncertainty in the Euro area. Third, European Central Bank, including some central banks, implemented a very bold and controversial monetary policy tool of 'negative interest rates.' The implementation was mainly aimed at avoiding the deflationary spiral in the EU when annual inflation was minus 0.6% in 2015 and at stimulating the economy by demotivating the hoarding of money and cash balances in the banking sector. This policy introduced a strange situation in modern economic history by violating Irving Fisher's popular statement that states that if a commodity could be stored costlessly over time, the interest in units of that commodity could never fall below zero. Moreover, the policy is counterintuitive in the perspective of risk and return relationship in which a positive return must reward depositors who are taking the risk of default of their deposits. The policy has not been as successful as it was expected to be. The reason could be the problems in the transmission process where banks hesitate to issue new loans amid an uncertain economic environment in the economy. The negative interest rate policy is still in effect in the Eurozone, and the COVID 19 pandemic has made the aggregate economic circumstances even worse than ever before. All these events have hit the Eurozone hard, so they might well have altered the predictive power of term spread, too.

The abovementioned unique context developed over time in the EU area is what makes this economic area of particular interest in this study. In addition to the unique context, the EU area economy is one of the world's most influential economies because of its size, currency, and composition. It is the second-largest economic area in the world. Its currency, the euro, is the second-largest reserve currency in the world after the US dollar. The economy consists of different sizes of national economies with free movement of goods, services, capital, and labor. Upon choosing the EU area, it is also possible to test whether the results obtained from the EU area hold for the selected sets of national economies of the EU area. For comparison, this thesis includes Germany, France, Italy, and Spain as a set of core economies of the EU area, whereas Ireland, Belgium, and Finland as a set of small economies of the EU area. Thus, a re-examination of the predicting power of term spread to predict the real economic activities in the EU area is imminent and of high relevance, even though it is not a new topic and has been studied for decades.

This thesis's general aim is to re-examine whether term spread still can accurately predict the real economic activity in the EU area. To be more precise, there are three specific objectives set to achieve the abovementioned general aim. The objectives are to test the predictive power of term spread during the negative interest rates' era in the euro area. The additional objectives are to examine the joint forecasting power of term spread and economic policy uncertainty (EPU), see Baker et al. (2016), and reveal the prominent Granger causality relationships between the three mainly focused variables: term spread, EPU, and real economic activity. The following research question becomes imminent to address in this study to accomplish the stated objectives. How good is the predictive power of term spread in predicting the real economic activities in the Euro area? While addressing this research question, it is also essential to understand how term spread's predictive power has changed during the period of negative interest rates. Does the inclusion of EPU with term spread increase the predicting power of the model? What is the causal relationship between term spread, EPU, and real economic activity?

This thesis uses the Euro area data for empirical analysis. The sample of the data ranges from 1999Q1 to 2019Q4. The starting point of the sample signifies the Euro area's establishment, and the ending point is 2019Q4, which is the quarter before the outbreak of the COVID 19 pandemic. The definition of the Euro area is slightly different from the definition of the European Union since the Euro area is a subset of the European Union. The European Union was established in the Maastricht treaty in 1992, while the Euro area was formed in 1999 as a monetary union of some European Union member states that decided to use the euro as their currency and sole legal tender. The expansion of the number of members in the European Union and the Euro area is still underway. Based on the most recent data, the Euro area includes 19 member countries: Germany, France, Italy, Spain, Austria, Belgium, Cyprus, Estonia, Finland, Greece, Ireland, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Portugal, Slovakia, and Slovenia. The Euro area has a significant contribution to the world economy. The data from the Eurostat (2020) provides that the Euro area has $13.3 trillion of the GDP in nominal terms, has a 342 million population, spans over 2.7 million square km, and makes $39 thousand GDP per capita, which is well above the global average.

Term spread, GDP growth rate as a proxy of real economic activity, and EPU are the variables used in this thesis's empirical analysis. The definition of term spread is simply the difference between short-term and long-term interest rates. The term spread used in this study is composed of the ten-year AAA-rated euro area central government bond yield minus the 3-month interest rate. The quarterly real GDP growth rate is used as a measure of real economic activity. Using the real GDP growth rate is more appropriate than using the nominal GDP rate for this study. The real GDP rate is a more accurate gauge of the change in production levels from one period to another, but the nominal GDP rate is a better gauge of consumer purchasing power. This study does not consider using the industrial production index since it covers only a part of the real GDP. The real GDP measures the price paid by the end-user, including value-added in the retail sector, which the industrial production index ignores. The European level EPU data are used as the Euro area EPU indicator in this study. Policy uncertainty refers to an economic risk in which the government policy's future path is uncertain, increasing risk premia and making businesses and individuals delay consumption and investment until the uncertainty has been decreased. The increase in EPU raises systemic risk and then raises the cost of capital in the economy. Consequently, a higher level of EPU can lower investments, mainly because of the irreversibility of investment. Higher EPU can have adverse effects on GDP growth and investment, with these effects estimated to be protracted through time (Caldara et al., 2019). 'Policyuncertainty.com' releases a monthly index of Global EPU that runs from January 1997 to the present. The Global EPU Index is a GDP-weighted average of national EPU indices for 21 countries (for more details, see Baker et al., 2016).

The vector autoregression (VAR) model is used as a method since it is a natural tool for time series forecasting. Also, the models are precise and straightforward to examine the predictive power of variables under study. The model is based on analyzing individual time series processes as a stochastic representation of the data and capturing the linear interdependencies among multiple time series. Each variable in the system has a regression equation explaining its evolution based on its own lagged values, the lagged values of the other model variables, and an error term. VAR modeling does not require as much knowledge about the forces influencing a variable as do structural models with simultaneous equations. The only prior knowledge required is a list of variables that can be hypothesized to affect each other intertemporally. The model developed in this thesis is fitted to forecast in-sample and out-of-sample. In addition to the VAR model, a linear regression model with dummy variables is also estimated to observe whether this model provides additional information that VAR models cannot provide.

This thesis begins with a brief introduction, followed by theoretical concepts and definitions in the second chapter. A literature review is presented in the third chapter. Data and methodology are detailed in the fourth chapter, empirical results are presented in the fifth chapter, and discussions and conclusions are presented in the sixth and seventh chapter, respectively.

2 THEORETICAL CONCEPTS AND DEFINITIONS

A deep theoretical understanding of the yield curve and the Taylor rule is essential to explain how interest rates and the real economy are linked with each other. In a typical economic environment, central banks attempt to stir the real economy with the help of changing short-term interest rates, which is one end of the yield curve.

2.1 The Yield Curve

A curve that connects different interest rates of identical bonds, except their difference in maturities, is known as the yield curve. The curve is also known as the term structure of interest rates. As it links the short-term interest rates to the real economy, it plays a central role in every economy. (Mishkin, 2011, 113-127)

Economists have observed three empirical facts of the yield curve. The first fact is that the interest rates on bonds of different maturities move together over time. The second fact is that when short-term interest rates are low, yield curves are more likely to have a shape of an upward slope, and as short-term interest rates are at a high level, yield curves are more likely to slope downward and be inverted. The third fact is that the yield curves almost always slope upward. A good theory must be able to explain the above mentioned three empirical facts of the yield curve.

Four theories have been put forward to explain the yield curve: the expectations theory, the segmented markets theory, the liquidity premium theory, and the preferred habitat theory. The expectations theory explains the first and second facts but fails to address the third fact. The segmented markets theory can explain only the third fact. The liquidity premium and preferred habitat theories can explain all three facts for the yield curve and hence are widely accepted views so far.

2.1.1 The Expectations Theory

The expectations theory states that the interest rate paid on a long-term bond equals an average of short-term interest rates that economic agents expect to occur over the maturity of the long-term bond. An assumption of this theory is that the holders of bonds do not have a preference on bonds of some maturity over bonds of other maturities, so holders do not hold a bond when the expected return from the bond is less than that of other bond having a different maturity. (Mishkin, 2011, 113-127)

If the expectations theory holds, then the following two strategies must have the same expected return. The first strategy is to purchase a one-year bond, and when it matures in one year, purchase another one-year bond. The second strategy is to purchase a two-year bond and hold it until maturity.

The expected return from the two periods by investing € 1 in the two-period bond and holding it for the two periods is calculated as

$$(1 + i_{2t})(1 + i_{2t}) - 1 = 1 + 2i_{2t} + (i_{2t})^2 - 1 = 2i_{2t} + (i_{2t})^2$$

$$= 2i_{2t}\ (When\ (i_{2t})^2\ is\ reasonably\ small\ )$$

Here, i_t refers to the interest rate on a one-period bond for time *t*, and i_{2t} refers to the interest rate on the two-period bond in time *t.*

With another strategy, in which one-period bonds are bought, the expected return on the €1 investment over the two periods is

$$(1 + i_t)(1 + i^e_{t+1}) - 1 = 1 + i_t + i^e_{t+1} + i_t(i^e_{t+1}) - 1 = i_t + i^e_{t+1} + i_t(i^e_{t+1})$$

$$= i_t + i^e_{t+1}\ \ (When\ i_t(i^e_{t+1})\ is\ reasonably\ small)$$

Here, i^e_{t+1} refers to the interest rate on a one-period bond expected for time *t+1.*

Both of these bonds will be held if these expected returns are equal, that is, when

$$2i_{2t} = i_t + i^e_{t+1}$$

$$or,\ \ i_{2t} = \frac{i_t + i^e_{t+1}}{2}$$

In this way, the n-period interest rate must equal the average of the one-period interest rates expected to repeat over the n-period life of the bond. The following equation shows the mathematical expression of the expectation's theory. (Mishkin, 2011, 113-127).

$$i_{nt} = \frac{i_t + i^e_{t+1} + i^e_{t+2} + \cdots + i^e_{t+(n-1)}}{n} \qquad (1)$$

If the short-term interest rate increases today, it tends to rise in the future. Thus, an increase in short-term rates increases people's expectations of increasing short-term rates in the future. Since long-term rates are the average of expected future short-term rates, an increase in short-term rates will also increase long-term rates, causing short- and long-term rates to move together. (Mishkin, 2011, 113-127).

When short-term rates are at a low level, people generally expect them to rise to some standard level in the near future, and the average rate of future expected short-term rates is higher compared to the present short-term interest rate. Therefore, long-term interest rates will be substantially above current short-term rates, and the yield curve would then have an upward slope. In contrast, if short-term rates are high, people usually expect them to come down. Long-term rates would then drop below short-term rates because the average of expected future short-term rates would be below current short-term rates, and the yield curve would slope downward and become inverted. (Mishkin, 2011, 113-127).

This theory cannot explain the fact that yield curves usually slope upward. A typical upward slope of yield curves implies that the short-term interest rate is typically expected to increase in the future. Practically, short-term interest rates are just as likely to decrease as they are to increase, and hence the expectations theory states that the typical yield curve should take a flat shape rather than an upward-sloping shape.

2.1.2 The Segmented Markets Theory

Segmented markets theory argues that markets for different-maturity bonds are segmented and not influenced by other segments. The theory strongly rejects the idea that bonds of different maturities serve as a substitute for each other. The theory assumes that investors have a particular holding period in their minds for their investment. Bonds that have shorter maturities have less interest rate risk than those with longer maturities. Investors tend to prefer short-term bonds over long-term bonds keeping the curve usually upward sloping. One of the reasons for preferring short-term bonds over long-term bonds could be the lower interest rate risk associated with shorter maturities than that of longer maturities. The demand and supply for a particular bond are responsible for the differing patterns of the curve. This theory cannot answer why interest rates on bonds of different maturities tend to move together and why the curve appears to be inverted when short term interest rates are high. Since expectations theory and segmented markets theory explain empirical facts that others cannot, the combination of these two theories becomes a logical step to follow. Combining these theories lead us to the liquidity premium and preferred habitat theories. (Mishkin, 2011, 113-127).

2.1.3 The Liquidity Premium Theory

This theory states that the interest rate on a long-term bond equals an average of short-term interest rates expected to occur until the maturity of the long-term bond and a liquidity premium. An assumption of this theory is that bonds of different maturities are partial substitutes. Investors are interested in short-term bonds because it is low exposure to the interest rate risk. A positive premium is required to induce investors to hold long-term bonds over short-term bonds.

$$i_{nt} = \frac{i_t + i^e_{t+1} + i^e_{t+2} + \cdots + i^e_{t+(n-1)}}{n} + l_{nt} \qquad (2)$$

Where l_{nt} refers to the liquidity premium of the n-period bond at time t, that is positive, and it increases with the term to maturity of the bond, n. (Mishkin, 2011, 113-127).

The liquidity premium theory and the preferred habitat theory conclude to the same point that investors usually prefer short-term bonds, and in order to prefer long-term bonds over short-term bonds, they need higher expected returns.

2.1.4 The Preferred Habitat Theory

The preferred habitat theory states that investors prefer bonds of one maturity over another, and they will be interested in buying bonds of different maturities only if they earn a higher expected return. This means the interest rate on a long-term bond has to be equal to an average rate of short-term interest rates that are expected to occur over the life of the long-term bond plus a liquidity premium. This theory also suggests that bonds of different maturities are partial substitutes. It is so because bonds of different maturities cannot be a substitute to each other, and each bond's interest rate with a different maturity is determined by the demand for and supply of that bond. (Mishkin, 2011, 113-127).

2.2 The Yield Curve Models

Understanding the yield curve models that are applied to estimate the yield curves indeed deepens the knowledge on the yield curve. Several models can be applied for empirical yield curve estimations considering the goodness of fit of the curve. European central bank releases daily euro area yield curves based on the Svensson model, see Nymand-Andersen (2018), who conducts a detailed study on the European central bank's yield curve estimation models. His findings report that the Deutsche Bundesbank, the Banco de Espana, the Banca d'Italia, and the Banque de France have used parametric models such as the Svensson model and the Nelson and Siegel model. The Nelson and Siegel model contains the slope, level, and curvature factors (Nelson & Seigel, 1987), whereas the Svensson model adds the hump factor to the Nelson and Siegel model (Svensson, 1994). The Svensson model seems to be a broader model than the Nelson and Siegel model. However, either model is not free from possible problems due to collinearity.

The Nelson and Siegel model specify a functional form for the instantaneous forward rate, f(t), as follows:

$$f(t) = [B_0 \quad B_1 \quad B_2] \begin{bmatrix} 1 \\ exp^{-\frac{t}{\tau}} \\ (t/\tau)exp^{-\frac{t}{\tau}} \end{bmatrix},$$

which can be expressed as follows:

$$f(t) = B_0 + B_1 \exp\left(-\frac{t}{\tau}\right) + B_2 \frac{t}{\tau} \exp\left(-\frac{t}{\tau}\right) \ldots\ldots (3)$$

Here B_0, B_1, B_2, t, τ are vector parameters, and B_0 and τ must be positive.

In the previous equation (3), Svensson (1994) adds a fourth term, with B_3 and τ_2 additional parameters, which refers to the hump-shape. The necessary condition is that t_2 must be positive.

The Svensson model presents a functional form for the instantaneous forward rate, f(t), as follows:

$$f(t) = B_0 + B_1 exp\left(-\frac{t}{\tau_1}\right) + B_2 \frac{t}{\tau_1} exp\left(-\frac{t}{\tau_1}\right) + B_3 \frac{t}{\tau_2} exp\left(-\frac{t}{\tau_2}\right) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4)$$

However, the Bank of England, the Federal Reserve Bank of New York, the Bank of Japan, and the Bank of Canada have used spline-based models such as the Waggoner cubic spline method with a three-tiered step-wise linear penalty function and the variable roughness penalty method or Waggoner model with a smooth penalty function (Waggoner, 1997)

An optimized linear combination of the basis spline generated with the De Boor algorithm designs the cubic spline model (De Boor et al., 1978). First, the augmented set of knot points are created as follows:

$\{dk\}_{k=1}^{k+6}$ Where, $d_1=d_2=d_3=d_4=s_1$, $d_{k+4} = d_{k+5} = d_{k+6} = s_k$, $d_{k+3} = s_k$ $\forall$ k in [1; K]

A cubic spline is a vector of h=p+2 cubic B-splines presented over the domain.

A B-spline is presented by the following recursion, where r = 4 for a cubic B-spline and
$1 \leq k \leq p$:

$$\theta_k^r(m) = \frac{\theta_k^{r-1}(m)\, x\, (m-d_k)}{d_{k+r-1} - d_k} - \frac{\theta_{k+1}^{r-1}(m)\, x\, (m-d_{k+r})}{d_{k+r} - d_{k+1}} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (5)$$

For $m \in [0, M]$, with $\theta_k^r(m) = \begin{cases} 1, \forall m \in [d_k;\ d_{k+1}[\\ 0, \quad else \end{cases}$

Thus, ultimately the vector
$\theta^r(m) = (\theta_1^r(m), \ldots., \theta_p^r(m)) = \theta_1(m), \ldots., \theta_p(m)$ ……………………………(6)
is achieved.

Then, the linear combination of this basis can construct any spline

$$\beta = (\beta_1, \ldots., \beta_p)^r$$

For a given maturity interval $[m_{min;}\ m_{max}]$, Waggoner (1997) explained a step-wise penalty function that is unchanged in three maturity breakdowns at three different levels, that all are to be fixed in advance.

$$\lambda(m) \begin{cases} a, \forall\, m \in [m_{min}: m_1[\\ b, \forall m \in [m_1; m_2[\\ c, \forall m \in [m_2; m_{max}[\end{cases}$$

The Federal Reserve Bank has used the step-wise penalty functions to estimate the yield curves using the values for a = 0.1, b = 100, c = 100,000, S = 0.1, μ = 1. This model can be modified with a continuous penalty function instead of the step-wise penalty functions as used by the Bank of England. (Nymand-Anderson, 2018).

The model defined by Anderson and Sleath (2001) for penalty function f(m) of m and three fixed parameters L, S, and U, which satisfy the following relationship:

$$log_{10}\, \lambda(m) = L - (L - S)\, x\, exp\, (-\frac{m}{\mu}) \ldots\ldots\ldots\ldots(7)$$

The values used for L=100,000, S=0.1, μ = 1.

Central banks can also use hybrid models if the model is parsimoniously reflecting a smooth curve and flexible enough to capture movements of the curve.

2.3 The Taylor Rule

The Taylor (1993) rule is a targeting monetary policy rule that acts as a reaction function used by central banks. The rule is designed to stabilize the economic activity by setting up an optimal level for the Fed Funds rate based on the inflation gap between the targeted inflation rate and actual inflation rate, and the output gap between the real and the potential output. The following equation mathematically explains the rule:

$$i = r^* + \pi + a_\pi(\pi - \pi^*) + a_y(y - y^*) \ldots\ldots\ldots\ldots(8)$$

where; i = nominal Fed Funds rate, r* = real Federal Funds rate, π = rate of inflation, $\pi *$ = target inflation rate, y = logarithm of real output, y* = logarithm of potential output. As the rule of thumb proposed by John Taylor (1993), the coefficients a_π and a_y should be set to 0.5.

The intuition behind the rule is that the monetary authorities should raise nominal interest rates more than the increase in the inflation rate. If the authorities do not raise the nominal interest rates more than the rise in the inflation rate, then the real interest rates fall as inflation rises. The rise in inflation causes monetary easing leading to a further rise in future inflation, which creates serious instability of the economy.

An explanation for the role of the output gap in the Taylor rule can be done using the concept of the Phillips curve. The Phillips curve states that inflation and unemployment have a stable and inverse relationship and claims that inflation comes with economic growth, which in turn leads to less unemployment. It is likely that changes in inflation are induced by the state of the economy with respect to its productive capacity, which is the proxy for potential GDP.

Thus, Taylor's rule is a useful tool for monetary authorities; however, putting the monetary policy on autopilot with a Taylor rule with fixed coefficients would be a bad idea.

2.4 Interest Rates and Real Economic Activity

Understanding the effects of expansionary monetary policy helps to understand the dynamics between short-term interest rates and real economic activities. The following schematic statement demonstrates how the increase in short-term interest rates impacts real economic activities.

$$Expansionary\ monetary\ policy\ \Rightarrow\ i_r \downarrow \Rightarrow WACC \downarrow \Rightarrow I \uparrow \Rightarrow AD \uparrow \Rightarrow Y \uparrow \text{……..(9)}$$

An expansionary monetary policy refers to a fall in real interest rates ($i_r \downarrow$). When the real interest rate falls, the cost of debt for corporations decreases and hence also lowers the weighted average cost of capital (WACC). The degree of lowering the WACC depends on the capital structure of a corporation. The WACC decreases significantly for corporations having a significantly higher debt-to-equity ratio. For example, the banking sector has a significantly high debt-to-equity ratio; thus, this sector is more sensitive to the changes in interest rates than other sectors. The WACC is considered as one of the tools to make investment decisions; however, the decision taken based on the WACC can be misleading due to the mixing up of the project's value with the tax shield. The decreased weighted average cost of capital increases the net present value of corporate projects, which in turn increases the probability of the acceptance of the projects. In favorable prospects, corporations increase investments creating more employment and higher demand for goods. Aggregate demand of the economy increases because of the increased employment level and investments. The output increases because of the increment in the aggregate demand. (Mishkin, 2011, 651-55).

Changes in interest rates not only affect corporations but also impact on decision making of households. The decrease in interbank rates or policy rates leads to a decrease in bank loan rates and deposit rates. As the deposit rates fall, households prefer spending or investing over saving because relatively low-interest rates discourage people from depositing in banks. Households increase their consumption demand, and hence, aggregate demand in the economy rises, and as a result, the output increases. (Mishkin, 2011, 651-655).

The change in short-term interest rates initially affects all short-term money market interest rates, and then the effect extends to the whole spectrum of interest rates in the economy. The effect even hits the long-term interest rates that are tied up with corporate investments. How efficiently the effects transmit from the changes in short-term interest rates to the real economy largely depends on the quality of the financial markets and the banking sector. (Mishkin, 2011, 651-655).

The changes in interest rate also affect the level of asset prices, for example, the prices of bonds, equity, and real estate. The decreased short-term interest rate boosts the supply of bonds, increases the equity prices, and increases the prices of real estate. Bond issuers find the decreased interest rate as the relatively cheaper mode of financing; therefore, bond supply increases. When the short-term interest rate decreases, investors tend to prefer equity over bonds; thus, the equity prices go up. Increased equity prices can raise households' and corporations' real demand due to their strengthened net worth position. As the interest rates decrease, the bank loan rates also decrease, which makes investing in real estate attractive; hence the real estate demand and prices can also increase. As a result, employment and real economic activities boost. Changed demand levels of bonds, equities, and real estate due to the decrease in the short-term interest rate can impact the aggregate demand of the economy. As equity prices rise, the market valuation of corporations increases, enhancing the replacement of debt capital to equity capital. The replacement can lead to a lower cost of capital of corporations, enhancing investment spending. (Mishkin, 2011, 651-655).

In contrast, contractionary monetary policy actions where the central banks increase the short-term interest rates have opposite effects on the real economic activities. Thus, changes in the short-term interest rates can have an impact on real economic activities in the economy.

3 LITERATURE REVIEW ON PREVIOUS EMPIRICAL RESULTS

This chapter presents a brief literature review on the ability of term spread to forecast real economic activity. There is an extensive amount of literature on the nexus of term spread – real economic activity.

The starting point of the literature review for this thesis dates to the late '80s and early '90s; however, the yield curve has been considered as one of the leading economic indicators since the 1930s. Many empirical studies have confirmed the positive predictive relationship between term spread and real economic activities, establishing a new stylized fact in monetary economics, while a few empirical studies have doubted the predictive power of term spread. In such a context, some key questions become essential to address while reviewing the literature for the purposes of this study. Is term spread indeed useful in forecasting real economic activity? If it is useful, then how stable has the predictive relationship been in the last 30 years?

3.1 Usefulness of Term Spread in Forecasting

Already Harvey (1988) and Estrella and Mishkin (1997) have examined the relationship between term spread and subsequent real activity. Estrella and Mishkin (1997) conclude that the yield curve is a simple and accurate measure to help guide European monetary policy. This conclusion holds true for the US economy as well. Harvey (1988) focuses on the US economy, whereas Estrella and Mishkin (1996) focus primarily on a sample, from 1973 to 1995, of major European economies: France, Germany, Italy, and the United Kingdom. Harvey (1988) tests the consumption capital asset pricing model and provides evidence on predictability only up to 3 quarters into the future, confirming that term spread contains information about future consumption. Since consumption and real economic activity are highly correlated, logically, it implies that term spread captures the information about the future real activity. So, these two independent studies conducted on the United States and Europe arrive at similar conclusions implying that term spread is indeed useful in forecasting real economic activity.

There are more pieces of evidence covering a long period and several economies to support the findings from Harvey (1988) and Estrella and Mishkin (1997). For example, Estrella and Hardouvelis (1991), Kozicki (1997), Pena et al. (2006), Papadamou (2009), Schunk (2011), Dar et al. (2014), and Hyozdenska (2015a) find strong shreds of evidence for the positive predictive relationship between term spread and real economic activities. Estrella and Hardovelis (1991) use the yield curve as a predictor of real economic activity using the US data for 33 years, starting from 1955. In this study, real economic activity refers to non-durables, services, consumer durables, and investment. The study presents evidence that term spread can predict cumulative changes in real output for up to 4 years. Kozicki (1997) investigates the predictive power of term spread, derived from 10 years bond and three months bill, for real economic growth in Australia, Canada, France, Germany, Italy, Japan, Sweden, Switzerland, the UK, and the US. Using the data from 1970 to 1996, this study confirms that spread has the optimum predictive ability for real growth in the next year. In addition, this study notes that the spread matters most for predicting real growth, whereas the level of short rates matters most for predicting inflation. Term spread is not only useful in forecasting real economic activities in major economies around the globe but also in relatively small economies in Europe. Papadamou (2009) examines the role of term spread on real economic activity using the data from the Czech Republic, Poland, Hungary, and Slovakia. The study data are from 1995M1 to 2004M4, and the term spread is derived from the 10-year government bond rate and the 3-month money market rate. He finds that the interest rate spread has some predictive power over the future 24 months, and he notes that term spread is a better indicator in countries with low and stable inflation than in countries with high and volatile inflation. In the case of the Czech Republic, the spread explains 43% of the variation of the growth, providing strong evidence for the usefulness of term spread in forecasting.

After the great financial crisis of 2008-9, Schunk (2011) reformulated the study of Estrella and Mishkin (1998) by focusing on probability predictions of rising or falling real GDP growth and inflation. This study not only argues for the usefulness of the yield curve in forecasting real economic activity but also points out that knowing whether the yield curve is currently in the process of getting steeper or getting flatter would add to the useful information content of the yield curve. There is evidence that term spread has been useful in forecasting emerging economies like the Indian economy. Using the data from October 1996 to April 2011, Dar, Samantaraya, and Shah (2014) examine the predictive power of spreads for output growth within aggregate and time scale framework using wavelet methodology. They find that the predictive power holds only at lower frequencies for the spreads that are constructed at the shorter end and at the policy-relevant areas of the yield curve. However, spreads that are constructed at the longer end of the yield curve do not seem to have predictive information for output growth. They observe that the use of wavelet methodology is of better value than ordinary least squares in their context. Hyozdenska (2015a) examines the relationship between the term spread and the economic activity of the United Kingdom, Iceland, Switzerland, Norway, and Russia between the years 2000 and 2013. She divides the sample into two parts: before 2008 and after 2008. She observes the poor predictive power of the yield curve in the first part of the sample, and it increases after 2008 in Iceland, Russia, and Great Britain. The result shows that the best predictive lags are a lag of four and five quarters. In this way, evidence suggests that term spread has been remarkably useful in forecasting real economic activity in several economies in the last three decades.

Acknowledging the stylized fact that the yield curve is useful in forecasting, some studies focus on decomposition of the curve to examine which component of the curve contains more information for real future activity. The level, curvature, and slope of the yield curve can be examined separately to get a deep understanding of the usefulness of the yield curve in forecasting. For example, Argyropoulos and Tzavalis (2016) provide evidence that the slope and curvature factors of the yield curve contain more information about future changes in economic activity than term spread itself.

They also argue that these two factors reflect different information about future economic activity, which is smoothed out by term spread. They find that the slope factor has predictive power on economic activity over longer horizons ahead, and the curvature factor has predictive power on shorter movements of future economic activity. This study's limitation is that the results hold only for developed economies. Hannikainen (2017) analyzes the predictive content of the level, slope, and curvature of the yield curve for US real activity in a data-rich environment. He finds that the slope contains predictive power but not the level and curvature. The predictive power of the yield curve factors fluctuates over time. The economic conditions matter for the predictive ability of the slope. Inflation persistence long emerges as a key variable that affects the predictive power of the spread. The spread tends to forecast the output growth better when inflation is highly persistent.

Recession and real economic activity are closely related to each other since a recession refers to a significant decline in real economic activity. In general, a recession refers to at least two consecutive quarters of negative growth in real GDP. For the United States, the NBER provides the most widely accepted definition of a recession. In this regard, this section touches on a short review of the literature on the use of term spread in recession forecasting. Indeed, the literature on term spread forecasting recession moves parallelly to the literature on term spread forecasting real economic activity. For example, Estrella and Mishkin (1998), Hasegawa (2009), Moersch and Pohl (2011), Stuart (2020) find the spread useful in forecasting recessions. Estrella and Mishkin (1998) find that term spread outperforms other indicators for generating parsimonious predictions of the probability of a recession, especially at horizons of three and greater the three quarters. Hasegawa (2009) examines, in the Japanese economy from January 1979 to March 2004, if term spread contains information on the future economic recessions' likelihood applying a probit model considering the stability of the relationship between the spread and the future recessions. He finds that a structural change in the relationship between term spread and future recessions occur at the end of 1996. He also finds that the Japanese term spread contains more accurate information on future recessions than the stock returns and nominal money supply before the structural break. Moersch and Pohl (2011) examine the ability of term spread to predict recessions for seven countries. The data sample for the United States and France is from 1970 to 2008, and the sample for Japan is from 1980 to 2008. The result indicates that the predictive power of term spread is best for Canada, Germany, the United States, and the United Kingdom. The short-term interest rate predicts a recession better than the term spread in France and Australia. They also note that monetary policy action is not the only factor that influence term spread. Stuart (2020) examines the ability of term spread to predict a recession in Switzerland, using monthly data during the period 1974 to 2017. She composes a term spread by using 10-year government bond yields and 3-month interbank rates or Swiss Libor rate. She makes four crucial findings from the study. First, she finds that term spread contains useful information for predicting recessions for horizons up to 19 months. Second, she finds that the state of the economy has a role in forecasting recessions. The result shows that the present state of the South African economy stays in its current state for a short forecast horizon, but in a longer forecast horizon, the economy is likely to change state. Third, results from the structural breaks test at several different plausible points show that the relationships between term spread and real economic activity are stable over the entire 43-year sample. Fourth, the inclusion of the KOF business course indicator and M1 growth variables in the model enhances the overall fit of the model at prediction horizons of 4 to 18 months in in-sample and out-of-sample testing. Thus, many studies confirm that term spread is not only a useful variable in forecasting real economic activity but also it is useful in forecasting recessions.

The shape of the yield curve of one economy can be useful to forecast a recession in other economies' which are closely connected. This can be possible due to several reasons, such as growing interdependence among economies in production processes, increasing capital flows among economies, and increasing the flow of resources around the globe. So, a recession in one economy can have an impact on other closely linked economies. In recent studies, Fullerton et al. (2017) examine the predictive capacity of term spread for the United States metropolitan economies situated along the border with Mexico. The results suggest that the flattening of the yield curve for either country tends to increase the probability of recessions in border economies.

3.2 Time-Varying Predictive Power of Term Spread

Despite the past evidence for the predictive power of term spread, many studies find that the stability in the predictive power of term spread has been inconsistent over time. Bismans and Majetti (2011) compare the ability of term spread with the euro-US dollar exchange rate in predicting French recessions over the period 1979 – 2010. They also compare static probit models with dynamic probit models to produce the recession probabilities. They find that the dynamic specification performs better than the static specification, and they argue that the exchange rate has higher predictive power than yield curve spread, and their out-of-sample results confirm the predominant role assigned to the exchange rate in predicting the latest recession occurred in 2008-9. Hvozdenska (2015b) analyzed the relationship between term spread and the economic activities of selected countries between 2000 and 2013. The result shows that prediction ability before and after the 2008 crisis is different. There can be several possible reasons to cause such inconsistency in the predictive power of term spread.

First, to examine the reason for the lost predictive power of term spread, Jardet (2004) performs a multiple structural change test that makes it possible to detect breaks in the correlation between the spread of interest rates and future activities in 1984 for monthly US data. This break is related to the loss of the predictive power of term structure. This work shows that the loss of predictability of the spread is due to a substantial drop in both contributions of monetary policy and supply shocks. Morrel (2018) provides new evidence in the decline in the US term spread's predictive power. The decline could be associated with the changes to the composition of shocks hitting the US economy that has caused term spread to be less reliable of future output growth in recent decades. Dong and Park (2018) examine the stability of the predictive power of term spread for future GDP growth. They find that the predictability has weakened since 1984Q1. They find that the term premium component loses predicting power significantly when they decompose term spread into expectation component and term premium component. The possible reason for this finding is the significant reduction in the volatility of the US macroeconomy. Kuosmanen, Rahko, and Vataja (2019) analyze the predictive power of three financial variables such as term spread, real stock returns, and the real short-term interest rate. Periods with a zero-lower bound of interest rates appear to reduce the predictive ability of stock markets. They also find evidence that persistence inflation increases the predictive content of financial variables. However, Karlsson & Osterholm (2020) examine the stability of predictive relation between term spread and the real economy using the United States data from 1953Q1 to 2018Q2 and applying the B-VAR model allowing drifting parameters and stochastic volatility. They decomposed term spread from the corporate bond yield. The variables under study are term spread, the real GDP growth, and the unemployment rate. Their first finding is that the relationship has been stable. Second, they observe stochastic volatility but do not notice any parameter drift. This means that the usefulness of term spread has not reduced even after the great recession 2008-9.

Second, some studies find that the yield curve augmented with other variables can predict the real economic activity more accurately than the yield curve alone. Chionis and Gogas (2010) examine the European, real GDP deviations from the long-run trend by using the data from the European Union, covering 1994Q1 to 2008Q3. They find that the yield curve augmented with the composite stock index had significant forecasting power in terms of the European Union's real output. Gogas and Pragidis (2011) use data from Germany, France, Italy, Portugal, Spain, Norway, Sweden, and the UK from 1991Q1 to 2009Q3. They find that the yield curve combined with the non-monetary variables has significant forecasting ability in terms of real economic activity, but the results differ qualitatively between the individual economies examined, raising non-trivial policy implications. Kuosmanen and Vataja (2017) re-examine the predictive ability of term spread, short-term interest rates, and the stock returns for real GDP growth in the G – 7 countries. They find that financial variables have regained predictive power since the financial crisis 2008-9, and they suggest that using several financial indicators to forecast GDP growth is preferable. Chen, Valadkhani, and Grant (2016) examine the usefulness of term spread for forecasting growth in the Australian economy from 1969 to 2014. They find evidence that term spread serves as a useful predictor of growth in aggregate output, private assets, the formation of private fixed capital, and inventories, both in-sample and out-of-sample. The predictive content of term spread neither changes with the inclusion of monetary policy variables nor alters when switching to the inflation-targeting regime by the Reserve Bank of Australia in the early 1990s. They provide significant proof to policymakers and economic agents on the usefulness of term spread to forecasting output growth for up to eight quarters ahead.

Third, evidence suggests that the models used in the study can have a role to play in the results. For example, Paya et al. (2004) analyze the non-linear behavior of the information contained in the spread for future real economic activity. They use nine US monthly industrial production series and four UK monthly real industrial production series for the period 1960M1-1999M3. The result shows that the non-linear model predicts more accurately than the linear model does. Based on a consumption-based asset-pricing framework with Generalized Isoelastic preferences, Pena and Rodriguez (2006) present a model that links the behavior of asset prices to the real economy. They use quarterly data from Canada and the United States for the period 1969Q4 to 2003Q3. Besides term spread, their model includes stock market term spread as a new variable, which is the slope of the expected market returns. Empirical results suggest that interest rate term spreads and expected stock market term spreads are significant factors to explain real activity in Canada and, to some extent, in the USA in the pre-technology bubble period. They observe that the predictive power of the two-factor model for Canada and the United States is higher than the one-factor model. Evgenidis and Siripoulos (2014) review the predictive ability of term spread conducting a comparative analysis of forecasting performance of different models by focusing on the last three US recessions: in 1990, in 2001, and in 2007. The results show that although linear models are useful in predicting the 1990 and 2001 decline in economic activity, none of these give the signal of the significant 2007 decline in output. The shape of the yield curve has more predictive power than that of the total term spread. They document that probit models are doing well in signaling the onset of the 2008-9 crisis, although they fail to predict the duration of the crises. Gogas et al. (2015) construct three models for forecasting the positive and negative deviations of real US GDP from its long-run trend over the period from 1976Q3 to 2011Q4. They employ two alternative forecasting methodologies: the probit model and support vector machines approach. Their results show that both methods give 100% out-of-sample forecasting accuracy for recessions. The support vector machine model gives 80% overall forecasting. Gogas et al. (2015b) investigate the forecasting ability of the yield curve in terms of the US real GDP cycle using the Machine Learning Framework. The results show 66.7% accuracy in overall forecasting and 100% accuracy in forecasting recessions. The results are compared to the alternative standard logit and probit model to provide further evidence about the significance of our original model. Gupta et al. (2020) developed a new Keynesian DSGE model to decompose term spread into its unobserved components, such as expected spread and the term premium. They analyze the ability of the whole term spread and then the ability of unobserved components separately to forecast the real economic activity. They estimate the model with Bayesian techniques with 18-time series using the in-sample data from 2000Q1 to 2003Q4 and out-of-sample data from 2004Q1 to 2014Q4. South African Reserve Bank quarterly bulletin and Statistics South Africa are two sources of their data. They find that term spread fails to predict in out-of-sample forecasting; however, in in-sample forecasting, it predicts accurately. To understand the reason for the failure of term spread in forecasting, they decompose term spread into two components, and then they observe the expected spread having the forward-looking component of term spread. They also note that the term premium is responsible for the slope of the curve. However, the scope of their finding applies only to the inflation targeting economy like South Africa. Evgenidis, Papadamou, and Siripoulos (2020) use a meta-analysis framework to deal with the heterogeneity in the results seen in the literature. They suggest considering nonlinearities and monetary policy in modeling the relationship. They argue that term spread is a useful tool in predicting economic activity in many major world economies, the US, Canada, and Europe, especially in financial stress periods. They also note that improvement in the stock market reduces the usefulness of term spread in predicting future economic activity.

4 DATA AND METHODOLOGY

This chapter presents the data used for empirical analysis and the method applied in this study. A short overview of the data is presented, and then the method applied in this study is discussed briefly.

4.1 Data

The sample of the data ranges from 1999Q1 to 2019Q4. Full sample data refers to the data from 1999Q1 to 2019Q4 and set 1, model fitting data set, refers to the sub-sample that ranges from 1999Q1 to 2014Q3. The remaining data, which is the evaluation set, is used for out-of-sample prediction. While doing so, the data converges to somewhere near 2016Q2 since the model fit and our-sample both data are from inside the sample. The data on term spread, GDP growth rate, and economic policy uncertainty was extracted from the OECD statistics on 24 August 2020. The following sections present the overview of economies, variables, and descriptive statistics.

4.1.1 Overview of Economies

The definition of the Euro area is slightly different from the definition of the European Union since the Euro area is a subset of the European Union. The European Union was established in the Maastricht treaty in 1992, while the Euro area was formed in 1999 as a monetary union of some European Union member states that decided to use the euro as their common currency and sole legal tender. The expansion of the number of members in the European Union and in the Euro area is still underway. Based on the recent data, the Euro area consists of 19 member countries: Austria, Belgium, Finland, Cyprus, Estonia, France, Germany, Greece, Ireland, Malta, Italy, Latvia, Lithuania, Luxembourg, the Netherlands, Portugal, Slovakia, Slovenia, and Spain. As compared to the world economy, the Euro area contributes 11.6% of the world GDP in PPP (ECB 2020). It has a 342 million population, covers 2.7 million square km area, and produces $39 thousand GDP per capita that is above the global average. (Eurostat, 2020).

Germany, France, Italy, and Spain are the four major core economies of the Euro area, covering a highly significant portion of the Euro area GDP. For the purposes of the analysis and discussion, this thesis uses these four countries as a set of economies that can be compared to the overall data. Similarly, this work uses another set of small economies, Finland, Ireland, and Belgium, for the purposes of the analysis comparing the overall data. It is interesting to observe and compare these small economies with the overall data as these economies have severe slacks due to the 2008-9 financial crisis.

The economy of Germany is the fifth-largest in the world in PPP terms and Europe's largest exporter of machinery, vehicles, chemicals, and household equipment. Germany is Europe's largest economy, the second-most populous country, and has significant influence in politics and defense. The composition of the German economy consists of industry 25.9 %, public administration, defense, education, human health and social work activities 18.2%, and wholesale and retail trade, transport, accommodation, and food service activities 15.8%. Germany's main export partners are France, the US, and the UK, while its main import partners are the Netherlands, France, and China. The French economy is an advanced industrial economy. Major economic activities contributing to the French economy are automobile manufacture, aerospace, information technology, electronics, chemicals, fashion, and pharmaceuticals. France gets the most visitors in the world, and it maintains the third highest relative income share in the world from tourism. Spain and the US are the main export partners, Belgium and Italy are the main import partners, and Germany belongs to both groups. Italy is a large manufacturer and exporter of a significant variety of products, including machinery, vehicles, pharmaceuticals, furniture, food, clothing, and robots. The economy of Italy is the 8th largest by nominal GDP in the world. It is the eighth largest exporter in the world, with $514 billion exported in 2016. Its closest trade partners are Germany 12.6%, France 11.1%, and the US 6.8%. Italy was among the countries hit worst by the recession of 2008-9 and the following European debt crisis. The economy of Italy was shrunk by 6.76% during the whole period, totaling seven quarters of recession. In 2015, the Italian government's debt was 128% of its GDP, ranking as the second-largest debt ratio of European countries after Greece. The economy of Spain is the world's thirteenth largest (measure in nominal GDP terms), as well as one of the largest in the world in terms of purchasing power parity. Following the financial crisis of 2007 – 2008, the Spanish economy plunged into another recession, entering a cycle of negative macroeconomic performance. (The World Factbook, 2020)

The economy of Finland is highly industrialized. The largest sector of Finland's economy is the service sector that holds 72.7 percent, followed by manufacturing and refining at 31.4 percent. The largest industries are electronics (21.6%), machinery, vehicles, and other engineered metal products (21.1%), forest industry (13.1%), and chemicals (10.9%). Belgium is a modern and capitalist economy. The economy has capitalized on the country's central geographic location, highly developed transport network, and diversified industrial and commercial base. The economy of Belgium has become strong due to its location in Western Europe. This country has a highly skilled and educated workforce. The multilingual nature of the workforce and its industrial emphasis has made the workforce one of the most productive in the world. The Republic of Ireland has a knowledge economy. It focuses on services into high-tech, life sciences, and financial services. Aircraft leasing, the Alcoholic beverage industry, engineering, energy generation, financial services, information and communications technology, medical technologies, and pharmaceuticals are the major sectors in the economy. (The World Factbook, 2020)

4.1.2 Variables in the Empirical Analysis

Term spread, real economic activity, and economic policy uncertainty are the variables used in this thesis. Term spread refers to the difference between short-term and long-term interest rates. Analogously to many previous studies, the term spread is defined in this study as the difference between the ten-year AAA-rated Euro area central government bonds rate minus the 3-month interest rate.

Real GDP growth is used as the measure of real economic activity. The real economic growth measures economic growth as it relates to the gross domestic product from one period to another, adjusted for inflation, and expressed in real terms as opposed to nominal terms. Using real GDP growth is more appropriate than using the nominal GDP for the purposes of this thesis. The real GDP is a more accurate gauge of the change in production levels from one period to another, whereas nominal GDP is a better gauge of consumer purchasing power. The industrial production index covers a part of the real GDP. The real GDP measures the price paid by the end-user, so it includes value-added in the retail sector, which the industrial production index ignores.

The European level data is used to represent the Euro area economic policy uncertainty. Policy uncertainty is an economic risk in which the future path of government policy is uncertain, increasing risk premia and making businesses and individuals to delay consumption and investment until this uncertainty has been resolved. Increases in the economic policy uncertainty index imply a rise in the systematic risk and also an increase in the cost of capital in the economy. As a result, higher economic policy uncertainty lowers investment, mainly because of the irreversibility of investment. The higher trade policy uncertainty can have adverse effects on GDP and investment. The website named www.policyuncertainty.com releases a monthly index of Global Economic Policy Uncertainty (GEPU) (Baker et al., 2016,) that runs from January 1997 to the present. The GEPU Index is a GDP-weighted average of national EPU indices for 21 countries. The data on the news based economic policy uncertainty index for the Euro area were extracted from www.policyuncertainty.com on 24 August 2020.

4.1.3 Descriptive Statistics

Term spread, the GDP growth rate, the EPU index for the Euro Area from 1999Q1 to 2019Q4 is graphically illustrated in Figure 1. Term spread increased from 0.90 in 1999Q1 to 2.35 in 1999Q3, and it decreased to 0.24 in 2001Q1. Again, the increase in the spread reached 2.27 in 2004Q2 before it plunged to the lowest point in the sample period -0.38 in 2007Q4. It remained on the floor for some quarters. The spread started to widen again from 2008Q3, which increased rapidly to a peak of 3.40 in 2010Q1. The high level was roughly maintained for 18 quarters. As the negative interest rate policy came into effect in the Euro Area, the spread decreased close to 1. Since then, the spread has been low and positive until 2019Q4.

The GDP growth rate was 2.15% in 1999Q1, which increased in 6 quarters to 4.5% and decreased strongly to 0.322% in 2003Q2. The growth rate reached 3.81% again in 2006Q4, and then it fell sharply to -5.65% in 2009Q1. In 2011Q1, the growth rate reached 2.92%, which again jumped to a negative region for six quarters until 2013Q2. After 2013Q2, the growth rate has stayed positive, hoovering between 1% to 2%.

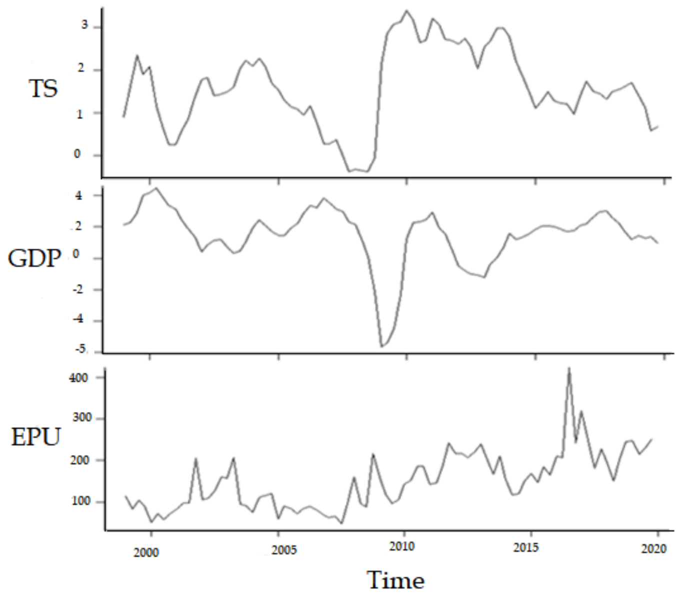


Figure 1: Three Variables Used in Empirical Analysis

In 1999Q1, the uncertainty level was at 113.21, which decreased to 51.95 in 2000Q1. The uncertainty index level was recorded at its lowest level (47.69) in 2007Q3, just before the global financial crisis started to erupt. In general, the index has been extremely volatile throughout the sample period. The highest level of uncertainty was recorded at 424.38 in 2016Q3.

The following table presents the descriptive statistics of three variables for the Euro area.

Table 1: Descriptive Statistics and Augmented Dickey-Fuller Test

	Descriptive Statistics					
	Min	Mean	Max	SD	Kurt	Skew
TS	-0.38	1.59	3.40	0.94	2.44	-0.14
GDP	-5.65	1.45	4.5	1.85	7.14	-1.73
EPU	47.69	150.25	424.38	68.89	4.44	0.90

	Augmented Dickey-Fuller Test			
	P-value	Test Stat.	Critical value	CI
dTS	2.574 e^{-08}	-5.768	-3.51	0.99
GDP	7.033 e^{-11}	-4.248	-3.51	0.99
dEPU	2.2 e^{-16}	-8.626	-3.51	0.99

Term spread is denoted as 'TS', the quarterly real GDP growth rate is denoted as 'GDP', and Economic policy uncertainty is denoted as 'EPU'. For the first differenced notation a 'd' is added as a prefix.

The minimum spread was -0.38, while the maximum spread was 3.40. The average spread during the sample period has been 1.59, with a standard deviation of 0.94. The distribution of the data slightly peaked as the kurtosis value is 2.44, and the symmetry of the data is slightly negatively skewed since the skewness is -0.14. The average quarterly growth rate of the real GDP is 1.45%, with 1.85 standard deviation. The distribution is highly peaked as the kurtosis value is 7.14, and the distribution of the data is negatively skewed. The average EPU index level in the Euro Area is 150.25, with a standard deviation of 68.89. The distribution is peaked as the kurtosis is 4.44, and the data experiences positive excess skewness.

The VAR model estimation's first step is to ensure the stationarity of the series analyzed since it is necessary for the time series methods assuming stationarity of the scrutinized variables for valid statistical inferences. One of the most widely used ways to test a time series stationarity is to perform the Augmented Dickey-Fuller (ADF) test. The ADF test is performed to check the stationarity of the time series under study. The test function ur.df () is applied in R studio. The stationarity test result shows that the p-value of the quarterly real GDP growth rate series for the Euro Area is $7.033e^{-11,}$ and the value of the test statistics is -4.2477, which is greater than the critical value, in absolute terms, at a 99% confidence interval. Similarly, the first differenced term spread series for the Euro area is stationary as the p-value of the series is $2.574e^{-08,}$ and the value of the test statistics is -5.7684, which is greater than the critical value, in absolute terms, at a 99% confidence interval. Furthermore, the first differenced EPU index series for the Euro area is also stationary since the p-value of the series is $2.2e^{-16,}$ and the value of the test statistics is -8.6262, which is greater, in absolute terms than the critical value at 99% confidence interval.

4.2 Methodology

Vector autoregression (VAR) models are natural tools for time series forecasting. In this framework, the current values of a set of variables are partly explained by past values of other variables involved. VAR models can be used in economic analysis. Structural VAR representations investigate the structural economic hypothesis. The relationships among variables in VAR models can be observed deeply using some methods such as impulse response, historical decompositions, the analysis of forecast scenarios, and forecast error variance decompositions. (Luetkepohl, 2011).

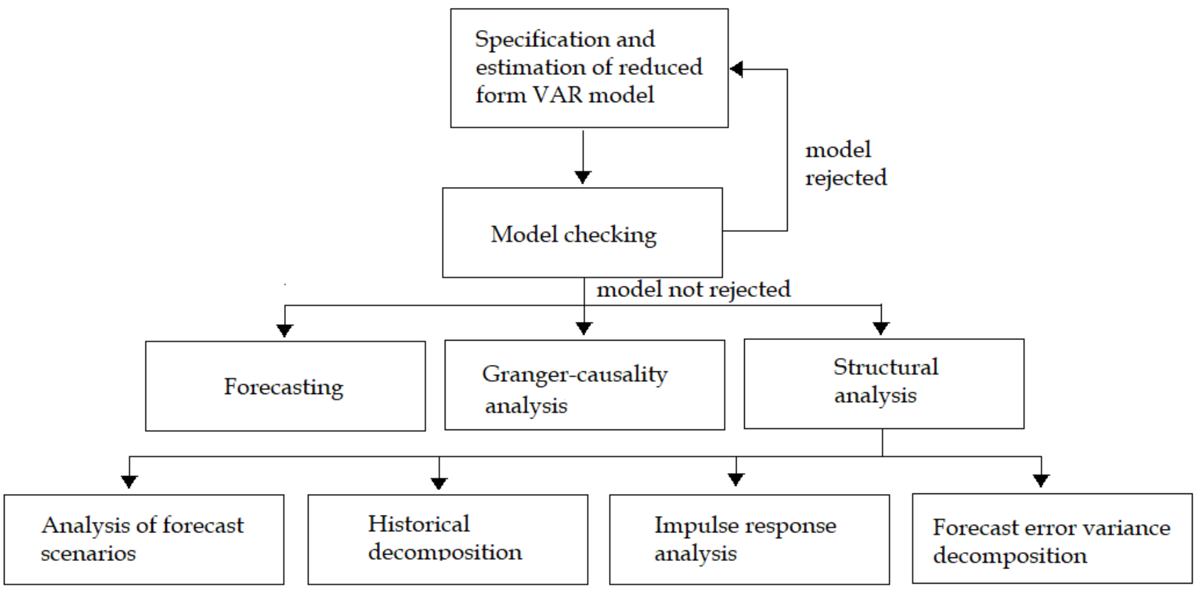


Figure 2: VAR Model Estimation Process (Luetkepohl, 2011)

The impulse response analysis is one of the ways to interpret the estimated VAR model. Generally, an impulse response refers to the reaction of any dynamic system in response to some external change. The impulse response function of VAR is utilized for analyzing the dynamic effects of the system when the model receives the impulse. In practice, the impulse responses are computed from the estimated VAR coefficients, and the bootstrap method is used to construct confidence intervals (CIs), which reflect the estimation uncertainty. In some cases, alternative bootstrap approaches may also be implemented. (Luetkepohl, 2011) The variance decomposition method is useful to evaluate how shocks reverberate through a system. In other words, variance decomposition is a way to quantify how important each shock is in explaining the variation in each of the variables in the system. It is equal to the fraction of the forecast error variance of each variable due to each shock at each horizon. (Luetkepohl, 2010) Sims (2011) states the following equation 10, explaining the forecast error variance of a variable j due to shock j at horizon h. Here ω_{i}, j(h) is the forecast error variance of variable i due to shock j at horizon h.

$$\omega_{i,j}(h) = \sum_{k=0}^{h} C_{i,j}\,(k)^2 \qquad (10)$$

$$\phi_{i,j}\,(h) = \frac{\omega_{i,j}\,(h)}{\Omega_i(h)} = \frac{\sum_{k=0}^{h} C_{i,j}(k)^2}{\sum_{k=0}^{h}\sum_{j=1}^{n} C_{i,j}(K)^2} \qquad (11)$$

To get the part of the forecast error variance of variable i due to shock j at horizon h, denoted $\varphi_{i,j}$ (h), the value obtained based on equation (10) is divided by the total forecast error variance.

Granger Causality is a method to investigate causality between two variables as time series (Granger, 1969). Granger Causality is a way to find patterns of correlation, and the results from the test are useful to know if a variable comes before another in the time series. The assumption in Granger causality is that the data generating processes in any time series are independent variables, and the data sets are analyzed to see if they are correlated. The null hypothesis for the Granger causality test is that lagged x-values do not explain the variation in y. In other words, it assumes that x(t) does not Granger cause y(t).

5 EMPIRICAL RESULTS

This chapter presents the empirical results obtained from the VAR model estimation, the linear model estimation with dummy variables, and the Granger causality test results.

5.1 The VAR Model

The following table summarizes the results from the estimated regression models in this section. First, the bivariate VAR model is estimated, and then the trivariate VAR model is estimated. As the trivariate VAR model outperforms the bivariate VAR model, another trivariate VAR model is estimated only for the period before the negative interest rate in the Euro area.

The bivariate VAR model consists of the first differenced term spread series with the quarterly real GDP growth series for the full sample data from 1999Q1 to 2019Q4. Akaike Information Criteria (AIC) is used to find the optimal lag length selection. Assuming that there is no 'true model' in the candidate set, Yang (2005) argues that the AIC is asymptotically optimal for selecting the model with the least squared error. Another reason to prefer the AIC over other information criteria is that it has been extensively used in VAR model estimation in several previous research works. As suggested by the AIC, lag order two is applied for this VAR model estimation. The GDP growth rate is a dependent variable that depends on its own two lags and the two lags of the first differenced term spread. Similarly, the first differenced term spread depends on its own two lags and the two lags of the GDP growth rate. The intercept is reported as a constant in both equations.

In the results presented in Table 2, the log-likelihood value of the model is -96.028, which is found to be relatively poor as compared with the log-likelihood value of the tri-variate VAR model to be estimated below in this subchapter. In general terms, lower log-likelihood values are considered as better; it means that the bigger negative numbers are better when comparative models have negative numbers. In the equation GDP growth rate, the estimated coefficients of GDP_{t-1} and GDP_{t-2} and the estimated coefficients of dTS_{t-1} are highly statistically significant, whereas the estimated coefficient of dTS_{t-2} is statistically significant at 95% confidence. The equation's residual standard error is 0.5855, the adjusted R^2 value is 0.903, and the p-value is smaller than $2.2e^{-16}$. Generally, these figures indicate that the model is good, but confirmation can be made only after the diagnostic tests.

Table 2: The VAR Estimation Results for Equation GDP_t

	Bivariate – VAR	Trivariate - VAR	
	Sample 1999Q1:2019Q4 Equation: GDP_t	Sample 1999Q1:2019Q4 Equation: GDP_t	Sample 1999Q1:2014Q3 Equation: GDP_t
Observations	81	81	60
dTS_{t-1}	-0.57**	-0.53**	-0.43
	(0.19)	(0.19)	(0.24)
dTS_{t-2}	0.41*	0.33	0.33
	(0.41)	(0.20)	(0.25)
GDP_{t-1}	1.45***	1.44***	1.46***
	(0.09)	(0.09)	(0.11)
GDP_{t-2}	-0.61***	-0.60***	-0.61***
	(0.09)	(0.08)	(0.10)
$dEPU_{t-1}$		-0.00	-0.00
		(0.00)	(0.00)
$dEPU_{t-2}$		-0.00	-0.00
		(0.00)	(0.00)
Constant	0.22	0.22*	0.18
	(0.09)	(0.09)	(0.11)
Residual SE	0.58	0.58	0.66
Adj. R^2	0.903	0.904	0.905
Log Likelihood	-96	-512.53	-373.79
ARCH test	P = 0.00036	P = 0.1964	P = 0.34
JB test (normality)	$P < 2.2\ e^{-16}$	$P < 2.2\ e^{-14}$	$P < 1.936\ e^{-09}$
BG-LM test	P = 0.409	P = 0.67	P = 0.17
OLS CUSUM	##	##	##

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

Note: OLS-based CUSUM test for stability see Appendix 1

In the table, dTS_{t-1} and dTS_{t-2} are two lags of term spread, GDP_{t-1} and GDP_{t-2} are two lags of GDP growth rate, and $dEPU_{t-1}$, $dEPU_{t-2}$ are two lags of EPU. OLS based diagnostic result is depicted in Appendix 1.

The degree of reliability of the information based on the estimated models can be supported or challenged after performing diagnostic tests. Therefore, diagnostics tests are performed in this study. ARCH test, normality test, serial test, and stability test are performed as diagnostic tests for the residuals of the estimated VAR models. The ARCH test result shows that the Chi-squared value 71.821 at 36 degrees of freedom and p-value of 0.0003575. The ARCH test's null hypothesis is that there is no existing autoregressive conditional heteroskedasticity in the residuals. As the p-value is below 0.05, the null hypothesis is rejected at a 95% confidence level, which means an ARCH effect is present in the residuals. The presence of the ARCH effect indicates that the inferences based on this model may not be reliable. This is the main reason why the trivariate VAR model, in which the ARCH effect is disappeared, has outperformed this bivariate model. The result from the JB test for normality shows that the Chi-squared value is 388.22, and the p-value is less than $2.2e^{-16}$. The null hypothesis of the normality test is that the sample distribution is normal. The result shows that the p-value is lower than 0.05, which means that the null hypothesis is rejected. So, it is confirmed that the full sample distribution is other than the normal distribution. The Breusch-Godfrey LM test for serial correlation of residuals shows that the Chi-squared value is 16.643, and the p-value is 0.409. The null hypothesis of the serial test is that there is no serial correlation in the residuals. The null is not rejected as the p-value is higher than 0.05, which means no serial correlation of residuals. The OLS-based CUSUM test for stability of the empirical fluctuation process in residuals shows that both equations in the VAR system do not touch the boundary line by exploding right after a shock in residuals. In conclusion, the results of the diagnostic tests show that the estimated bivariate VAR model has an ARCH effect and non-normality. Further analysis based on this model can be misleading in such a context, mainly due to the ARCH effect. Therefore, a variable 'EPU' is added to the model to estimate the trivariate VAR model.

In the trivariate VAR model, the equation for term spread and GDP growth rate is the same as in the bivariate model since the optimum lag length is two in both cases. The added variable, the first differenced EPU index, depends on its own two lags, and two lags of GDP growth rate, and the first differenced term spread. The intercept is reported as a constant in all three equations. Like in the bivariate VAR model, Akaike Information Criteria is used to find the optimal lag length. The AIC suggested that the optimal lag length be two.

In the results from the GDP_t equation, the estimated coefficients of GDP_{t-1} , GDP_{t-2} , and dTS_{t-1} are highly statistically significant. The equation's residual standard error is 0.58, the adjusted R-Squared value is 0.904, and the p-value is smaller than $2.2e^{-16}$.

ARCH test, normality test, serial test, and stability test are performed as diagnostic tests for the trivariate VAR model residuals. The ARCH test result for Euro Area shows that the Chi-squared value 158.29 at 144 degrees of freedom and p-value of 0.1964. The ARCH test's null hypothesis is that there is no existing autoregressive conditional heteroskedasticity in the residuals. As the p-value is above 0.05, the null hypothesis is not rejected at 95% confidence, which means no ARCH effect in the residuals is found. The opposite was the case in the bivariate model explained in the earlier part of this section. The result from the JB test for normality shows that the Chi-squared value is 305.51, and the p-value is less than $2.2e^{-16}$. The null hypothesis of the normality test is that the sample distribution is normal. The p-value is lower than 0.05, which means that the null hypothesis is rejected. It is confirmed that the sample is not normally distributed. The Breusch-Godfrey LM test for serial correlation of residuals shows that the Chi-squared value is 31.77, and the p-value is 0.67. The null hypothesis of the serial test is that there is no serial correlation in the residuals. The null is not rejected as the p-value is higher than 0.05, which means no serial correlation of residuals. The OLS-based CUSUM test for stability of the empirical fluctuation process in residuals shows that all three equations in the VAR system do not explode right after a shock in residuals. Thus, the trivariate model is considered as a relatively better model than the bivariate model.

Based on the trivariate VAR, in-sample model fitting is performed to reflect the explanatory power of the variables under study. The following figure shows the in-sample fitness of the model.

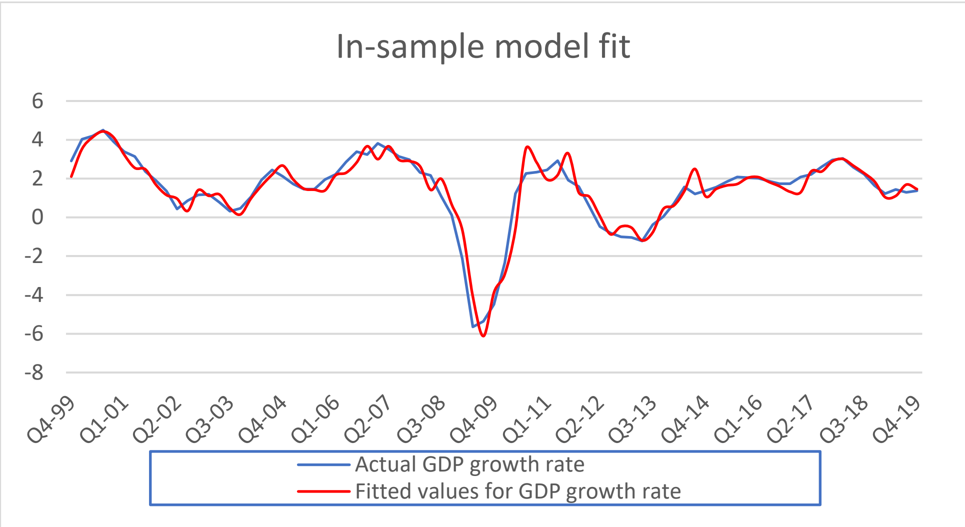


Figure 3: In-sample Model Fit for Euro Area

In Figure 3, the fitted values from the estimated VAR model suggest that the model is relatively better to capture the true development of the dependent variable. Most of the time, GDP growth rate's fitted values are very close to the actual values. Based on the in-sample model fit, the model's average prediction error is 44.12%, and the median prediction error of the model is 14.79%. Some outliers, for example, 2008Q2, 2009Q2, 2010Q1, 2011Q2, and 2014Q2, present in the residuals, seem to worsen the model fit. When five observations with the highest residual error are omitted, the remaining observations' average residual error would be 22.77%. The change in the residual error due to outliers is significant.

In addition to the in-sample model fit, this study also performs out-of-sample prediction using the sample's first 60 observations. Those first 60 observations are called sample set 1, which covers the period before the Euro area's negative interest rate period. So, the data ranges from 1999Q1 to 2014Q3.

Using the set 1 sample, the out-of-sample prediction is performed. The out-of-sample prediction is compared to the actual data from 2014Q4 to 2019Q4. A new VAR system is created by binding the first differenced term spread series, the quarterly real GDP growth series, and the first differenced economic policy uncertainty index series for the set 1 sample. As suggested by the HQ and FPE, lag order two is applied for VAR model estimation. Lag length two is applied for this model, considering the parsimony of the model. However, AIC suggests a much higher lag length, which does not seem to be practical.

In the GDP_t equation, the estimated coefficients of GDP_{t-1} and GDP_{t-2} of lag 1 and 2 are highly statistically significant. The equation's residual standard error is 0.66, the adjusted R-Squared value is 0.9057, and the p-value is smaller than $2.2e^{-16}$.

ARCH test, normality test, serial test, and stability test are performed as diagnostic tests for the Euro area VAR model's residuals. The ARCH test result shows that the Chi-squared value 150.54 at 144 degrees of freedom and p-value of 0.34. The ARCH test's null hypothesis is that there is no existing autoregressive conditional heteroskedasticity in the residuals. As the p-value is much higher than 0.05, the null hypothesis is not rejected at 95% confidence, which means no ARCH effect in the residuals. The result from the JB test for normality shows that the Chi-squared value is 51.919, and the p-value is less than $1.936e^{-09}$. The null hypothesis of the normality test is that the sample distribution is normal. The p-value is lower than 0.05, which means that the null hypothesis is rejected. It is confirmed that the sample is significantly different than normal. The Breusch-Godfrey LM test for serial correlation of residuals shows that the Chi-squared value is 44.149, and the p-value is 0.17. The null hypothesis of the serial test is that there is no serial correlation in the residuals. The null is not rejected as the p-value is higher than 0.05, which means a serial correlation of residuals is not present. The OLS-based CUSUM test for stability of the empirical fluctuation process in residuals shows that both equations in the VAR system do not explode right after a shock in residuals.

The following figure shows the out-of-sample prediction of the model for the Euro area.

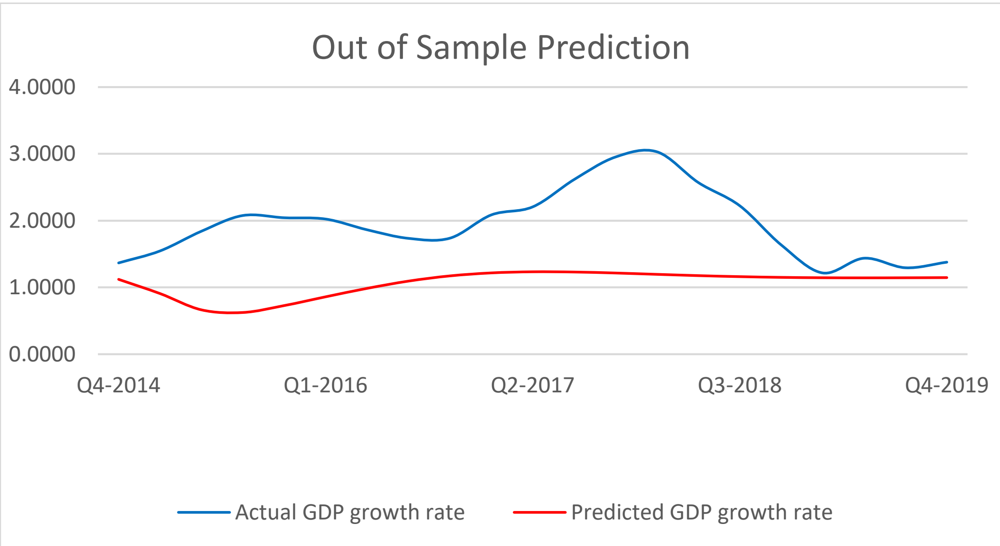


Figure 4: Out-of-sample Prediction of the Model

In Figure 4, this model's prediction error suggests that this model has low predictive power compared to the explanatory power observed in-sample model fitness. Based on the out-of-sample prediction, the model's average prediction error is 41.71%, and the median prediction error of the model is 44.06%. When five observations with the highest prediction error are omitted, the average prediction error would be 34.90%. The changes in prediction error are not significantly big, meaning that the average prediction error is sensible.

5.2 Model Tuning

In this section, a linear model is estimated for further empirical treatments to observe the model's changes in predictive power while adding three dummy variables. The financial crisis, negative interest rate period, and high uncertainty period are assigned as dummy variables.

Financial crisis refers to the financial crisis of 2008-9; its dummy variable takes 1 for the period from 2008Q4 to 2009Q4; otherwise, it takes 0. The negative interest rate period is a dummy variable, which takes 0 for the period from 1999Q1 to 2014Q3, and it takes 1 for the period from 2014Q4 to 2019Q4. The high uncertainty period refers to the quarters that have an above-average economic policy uncertainty index. When the economic policy uncertainty index for the quarter in question is above average, it takes 1; otherwise, 0.

In this way, a new linear regression model is estimated whose equation is as follows.

$$GDP_t = GDP_{t-1} + GDP_{t-2} + dTS_t + dTS_{t-1} + dTS_{t-2} + dEPU_t + dEPU_{t-1} + dEPU_{t-2} + NIRP + FC + HUP + constant \ldots\ldots \quad (12)$$

GDP_t = quarterly real GDP growth rate, GDP_{t-1} = quarterly real GDP growth rate for lag one, GDP_{t-2} = quarterly real GDP growth rate lag two, dTS_t = the first differenced term spread, dTS_{t-1} = the first differenced term spread for lag one, dTS_{t-2} = the first differenced term spread for lag two, $dEPU_t$ = the first differenced economic policy uncertainty index, $dEPU_{t-1}$ = the first differenced economic policy uncertainty index for lag one, $dEPU_{t-2}$ = the first differenced economic policy uncertainty index for lag two, NIRP = Negative interest rate period, FC = Financial crisis 2008-9, and HUP = High uncertainty period.

In equation 12, the Euro area's GDP growth rate is a dependent variable that depends on its own two lags, the two lags of first differenced term spread, and the first differenced EPU index of the Euro area, the negative interest rate period, the financial crisis 2008-9, and the high uncertainty period.

The following table shows the results for the equation GDP_t obtained from the linear model with the full sample, the VAR model before the negative interest rate period, and the linear model before the negative interest rate period. Furthermore, the table also presents the estimate coefficients as well as the model-specific statistical information.

Table 3: Results From the Linear Model With Dummy Variables

	Linear Model Sample 1999Q1:2019Q4 Equation: GDP_t	Linear Model Sample 1999Q1:2014Q3 Equation: GDP_t
Observations	81	60
dTS_t	-0.33	-0.24
	(0.19)	(0.24)
dTS_{t-1}	0.33	0.37
	(0.18)	(0.23)
dTS_{t-2}	0.09	0.05
	(0.18)	(0.22)
GDP_{t-1}	1.20***	1.17***
	(0.09)	(0.11)
GDP_{t-2}	-0.47***	-0.44***
	(0.08)	(0.10)
$dEPU_t$	0.00	0.00
	(0.00)	(0.00)
$dEPU_{t-1}$	-0.00	-0.00
	(0.00)	(0.00)
$dEPU_{t-2}$	-0.00	-0.00
	(0.00)	(0.00)
FC	-1.68***	-1.82***
	(0.39)	(0.46)
HUP	-0.45**	-0.56**
	(0.15)	(0.21)
NIP	0.33*	NA
	(0.16)	NA
Constant	0.59***	0.65***
	(0.12)	(0.15)
Residual SE	0.51	0.56
Adj. R^2	0.927	0.930

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

The notations of the variables are the same as the notations used in Table 2. The dummy variables: FC refers to the financial crisis 2008-9, HUP refers to the high uncertainty period, and NIP refers to the negative interest period.

In GDP_t equation, the estimated coefficients of GDP_{t-1} and GDP_{t-2} lag 1 and 2, and the FC are highly statistically significant. The estimate coefficients for dTS_t is -0.33, which is statistically significant at 90% confidence interval, dTS_{t-1} is 0.33, which is also statistically significant at 90% confidence level, and dTS_{t-2} is 0.09. None of them are statistically significant at 95% confidence level in explaining the GDP_t equation in this model, while both lags of GDP growth rate have statistically significant estimate coefficients. The estimate coefficient for the financial crisis is -1.68, high uncertainty period and negative interest rate period are all statistically significant at at-least-95% confidence interval.

The financial crisis, the high uncertainty period, and the negative interest rate period are also statistically significant. The equation's residual standard error is 0.507, the adjusted R-Squared value is 0.93, and the p-value is smaller than $2.2e^{-16}$. The following figure shows the in-sample model fitness for the Euro area.

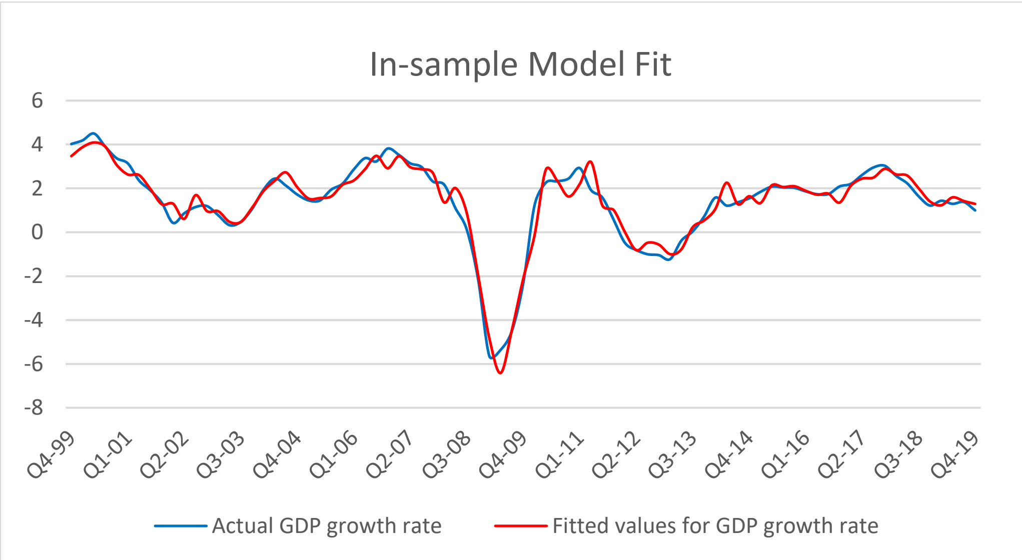


Figure 5: In Sample Model Fit

In Figure 5, the model's data generation process resembles that this model is slightly better than the trivariate VAR model discussed in the previous section. Most of the time, the GDP growth rate's fitted values are close to the actual values. Based on the in-sample model fitness, the model's average residuals error is 37.11%, and the model's median residuals error is 15.42%. When five observations with the highest prediction error are omitted, the model's average residuals error would be 19.18%.
The linear model for the period before the negative interest rate period is estimated. In the model, the model's adjusted R^2 is 0.93, and the residual standard error is 0.56. The statistically significant estimate coefficients are GDP_{t-1}, GDP_{t-2}, FC, and HUP. The estimate coefficients for dTS_t is -0.24, dTS_{t-1} is 0.37, and dTS_{t-2} is 0.05. None of them are

statistically significant at a 95% confidence level in explaining GDP_t equation in this model, while both lags of GDP growth rate have statistically significant estimate coefficients. Simultaneously, the financial crisis and the high uncertainty period are both statistically significant at at-least 95% confidence interval.

The out-of-sample prediction is performed to assess the predictive power of Term spread. The following figure shows the out-of-sample prediction of the model for the Euro area.

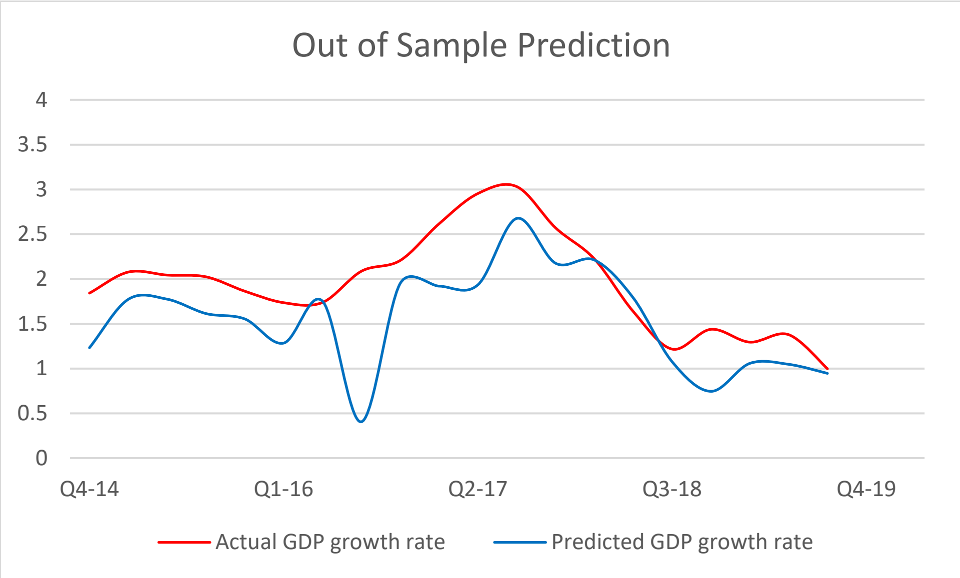


Figure 6: Out of Sample Prediction

In Figure 6, the out of sample prediction result suggests that the model has gained its predictive power compared to the model's predictive power without dummy variables. Based on the out-of-sample prediction, the model's average prediction error is 20.70%, and the median prediction error of the model is 15.37%. When five observations having the highest prediction error are omitted, the average prediction error would be 13.24%. In the figure, the prediction looks slightly pessimistic as the blue line is consistently under the red lines most of the times, resulting smaller predicted values than the actual values.

5.3 Granger Causality

This section presents the detailed Granger Causality results, based on the full sample VAR model, from the Euro area. First, the Granger causality between Term spread and GDP growth rate is tested. The null hypothesis is that the first differenced term spread does not Granger cause GDP growth rate. The following table shows the Granger causality test results for the Euro area.

Table 4: **G**ranger Causality Results for Euro Area

Null Hypothesis	
Term Spread does not Granger Cause GDP growth rate	F Test = 6.158 P value = 0.003
GDP growth rate does not Granger Cause term spread	F Test = 5.842 P value = 0.004
EPU does not Granger Cause GDP growth rate	F Test = 2.677 P value = 0.072
Optimum lag length = 2	

Table 4 shows that the result from the optimum lag two is statistically significant to reject the null hypothesis, which means that term spread Granger causes GDP growth rate. Besides, in the optimum lag two, term spread Granger causes in lag up to six. The causation strength decays as the lag length rises, and it disappears in the higher lags than six.

Before the negative interest rate period, term spread Granger causes up to three lags only, and the causation strength is much weaker than that from the full sample. GDP growth rate also Granger causes term spread in the optimum lag two. So, the bi-directional causality is observed between term spread and the GDP growth rate. GDP growth rate's causation to term spread is much weaker than that from term spread to GDP growth rate. Furthermore, the economic policy uncertainty does not Granger cause GDP growth rate at any lags at all. However, before the negative interest rate period, EPU Granger causes GDP from the second lag to the fifth lag.

6 DISCUSSION OF THE MAIN RESULTS

The general aim of this study is to re-examine the power of term spread in predicting the real economic activity in the Euro area during the negative interest rates era. The study begins with the three down-to-earth research objectives such as to test the predictive power of term spread in the negative interest rate period in the Euro area, to examine the joint forecasting power of term spread and the economic policy uncertainty, and to reveal the Granger causality between the variables under study.

Term spread used in this study refers to the difference between the three months interest rate and the ten-year government bond rate. The GDP growth rate is calculated based on the log difference between the GDP level of the present quarter and the previous quarter. The EPU index is retrieved from its official website (www.policyuncerainty.com), derived from the comprehensive text mining of the major historical news articles. Term spread and GDP growth rate are major variables; however, EPU is also added to improve the model.

The full sample of this thesis ranges from 1999Q1 to 2019Q4. The beginning of the sample refers to the establishment of the Euro area and the end of the sample period refers to the quarter right before the COVID 19 outbreak. A fraction of the sample that ranges from 1999Q1 to 2014Q3 is termed as the sample set 1, in which the out of sample prediction is based. The set 1 sample contains the observations until the commencement of the negative interest rate policy in the Euro area. The predicted values for 21 quarters, starting from 2014Q4, are compared to the actual values from 2014Q4 to 2019Q4, while the in-sample model fitness is generated from the full sample data.

The bivariate VAR model is estimated using term spread and GDP growth rate before estimating the tri-variate VAR model, including an additional variable: EPU. Although the R^2 value of the bivariate model is not significantly different from that of the tri-variate model, the tri-variate model is considered better than the bivariate model since the latter has an arch effect in the residual. Furthermore, a linear model is estimated with some additional information such as the financial crisis 2008-9, high uncertainty period, and a negative interest rate period.

At the end of the empirical analysis, the Granger causality is also tested among the variables under study. The results obtained from the model reveal the explanatory and the predictive power of term spread before and during the negative interest rate period.

6.1 GDP Growth Rate and Term Spread in the Euro Area

The predictive power of term spread is discussed using the results obtained from the in-sample and out-of-sample analysis.

6.1.1 GDP Growth Rate and Term Spread Model Fit

The determination coefficient of the adjusted R^2 estimates the amount of variability in GDP growth rate that can be explained by all independent variables mentioned in equation 12. The increased value of the determination coefficient of the adjusted R^2 can be interpreted as the increased explanatory power of the overall model at hand. Even though it is interpreted that the higher the adjusted $R^{2,}$ the better the explanatory power of the model, sometimes the higher adjusted R^2 value can also be obtained just because of overfitting of the model. The models discussed here are not overfitted models since they neither have too many variables nor too many parameters associated with variables. The adjusted R^2 for the linear model is increased to 0.93 from the VAR model's 0.90. This means that the linear model is better than the VAR model to explain the Euro area's GDP growth rate as well as the real economic activities.

Although the adjusted R^2 for the linear model indicates that the model's explanatory power is outstanding, term spread's estimate coefficients are not found to be impressive as none of them are statistically significant at a 95% confidence interval. In the full sample trivariate VAR model, dTS_{t-1} is -0.53 and dTS_{t-2} is 0.33. In the full sample linear model, dTS_t, dTS_{t-1}, and dTS_{t-2} are -0.33, 0.33, and 0.09 respectively. Only the dTS_{t-1} from the VAR model is statistically significant at a 95% confidence interval, while dTS_t and dTS_{t-1} from the linear model are significant at a 90% confidence interval. Based on the adjusted R^2 and the estimated coefficient for term spread, it is confirmed that the explanatory power of term spread in explaining the GDP growth rate is poor. The following table presents the residual errors from the VAR model and the linear model.

Table 5: The Residual Error Analysis

	VAR Model Sample 1999Q1:2019Q4	Linear Model Sample 1999Q1:2019Q4
Average Residual Error	44.12%	37.11%
Median Residual Error	14.79%	15.42%
Average Residual Error Without Five Outliers	22.77%	19.8%

The following findings further justify the low explanatory power of term spread. The average residual error from the in-sample model fit in the VAR model is 44.2% and in the linear model is 37.11%. In either case, the explanatory power does not look good. This is in line with the statistically insignificant values of the estimate coefficients for term spread in explaining GDP growth rate. A higher residual error refers to the poor performance of the model.

Unlike the average residual errors, the median values of the residual errors give a slightly different result. The median residual error obtained using the VAR model is 14.79% and it using the linear model is 15.42%. A possible reason for such a big difference in the average and the median residual error in both models can be due to outliers that are responsible for making the sample distribution skewed. Digging deeper, when five observations with the highest residual error are omitted from the sample, the remaining observations' average residual error found to be 22.77% in the VAR model and 19.8% in the linear model. The average residual error change is significant, indicating that the outliers have played a considerable role in increasing the residual errors. Mostly the outliers are not from the consecutive quarters of a particular period but from the quarters before and after the financial crisis 2008-9.

The financial crisis 2008-9 gave a huge set back to the Euro area economy. Rising economic policy uncertainty badly affected the economy too. Also, the negative interest rate policy period has complicated the issues, worsening the economy even worse. Three dummy variables representing the abovementioned three issues are added to the linear model, searching for more useful information. Nevertheless, the dummy variables' selection is the author's subjective decision, justified by the estimate coefficients of the dummy variables and the improved model's performance. For example, the linear model is estimated for the equation GDP_t, in which adjusted R^2 is 0.93, which was 0.90 in the tri-variate VAR model, showing significant improvement in the model due to the inclusion of the dummy variables.

Moreover, the estimated coefficient for the financial crisis has a coefficient value of -1.68 is highly significant; having the highest estimate coefficient among all the variables in the model. With the coefficient value of -0.45, economic policy uncertainty is also significant at the 99% confidence interval. Similarly, the estimate coefficient, whose value is 0.33, for the negative interest rate policy is also significant at the 95% confidence interval. The estimate coefficient for the financial crisis is higher than high uncertainty and negative interest rate; this means that the dummy variable helped the model to explain more accurately. The negative interest rate period seems to have less impact on the model's performance among the three of them. In contrast, high uncertainty also has a significant impact on model performance. In subchapter 6.3, further discussion about the predictive power of term spread with respect to the dummy variable is presented.

6.1.2 Predictive Power of Term Spread

After confirming that the model is significantly useful to make an out-of-sample prediction, the models' prediction error is estimated to assess the model's predictive power.

Table 6: The Prediction Error Analysis

	VAR Model Sample 1999Q1:2014Q3	Linear Model Sample 1999Q1:2014Q3
Average Prediction Error	41.71%	20.70%
Median Prediction Error	44.06%	15.37%
Average Prediction Error without Five Outliers	34.90%	13.24%

The models estimated for the out-of-sample prediction are based on set 1 data. The VAR model's adjusted R^2 is 0.93, which is 0.90 in the tri-variate VAR model, and the residual standard error is 0.56, which is 0.66 in the tri-variate VAR model. The changes clearly show that the linear model is much better than the tri-variate VAR model in explaining GDP growth rate. Based on the out of sample prediction, the model's average prediction error is 20.70%, which is 41.71% in the tri-variate VAR model, and the median prediction error of the model is 15.37%, which is 44.06% in the tri-variate VAR model. When five observations having the highest prediction error are omitted, the average prediction error of the remaining observations is 13.24%, which is 34.90% in the tri-variate VAR model. The changes between the prediction error values of the VAR model and the linear model show that the linear model is better at explaining the predictive power for GDP growth rate.

Even the models are excellent, the estimate coefficients for term spread are still poor, as were in in-sample-analysis. In the VAR model, the estimate coefficients for dTS_{t-1} is -0.43, which is statistically significant at a 90% confidence level, and dTS_{t-2} is 0.33. In the linear model, the estimate coefficients for dTS_t is -0.24, for dTS_{t-1} is 0.36, and dTS_{t-2} is 0.05. None of term spread's estimate coefficients are statistically significant, showing that term spread's low predictive power in predicting the GDP growth rate.

6.2 Term Spread and Predictive Power in Other EU Economic Area

Germany, France, Italy, and Spain are the core economies for the Euro area. The predictive power of Term spread in the core economies is presented in APPENDIX 1. The average residual errors for Italy and Spain are extremely high, while the adjusted R^2 for Italy is the highest among all the core countries, and the adjusted R^2 for Spain is the lowest. So, the models for Italy and Spain seem to have poor predictive performances. Germany has a moderate level of adjusted R^2 and high average residual error, which means this model is also poor. Thus, the predictive power of term spread is not strong in Germany, Italy, and Spain.

The models for France are much better in predicting the real economic activities than any other models from the core economies. The adjusted R^2 from the VAR model for France is 0.88, and it is the same from the linear model too. Not only the adjusted R^2 of the model is good, but also the estimate coefficients for term spread is statistically significant. Based on the linear model, term spread has good predictive power in France since the estimate coefficients for lag 1 is 0.73, which is statistically significant at a 98% confidence interval. Based on the VAR model, the estimate coefficient for lag 1 of term spread is 0.38, which is also statistically significant at a 95% confidence interval.

Based on the in-sample analysis, the average residual error for France is nearly 50%, while the median residual error for France is nearly 20%, showing close similarities with the Euro area's result from the in-sample analysis. When the five observations having the highest residual error are omitted, the average residual error drops to 27%. Even the predictive power of term spread is low in the Euro area, the predictive power of term spread is good in France.

Based on the out-of-sample analysis, the model for France outperforms all other models from the core economies. The average prediction error for France is 30.81%, and the median residual error is 23.73%. These figures are relatively better than figures from Italy, Spain, and Germany. When the five observations having the highest residual error are omitted, the average residual error drops to 19.91%. In the case of the model for France, the result from the out of sample analysis is consistent with the result from the in-sample analysis. Thus, the predictive power of term spread is good in France. However, the predictive power of lags of GDP growth rate is even stronger than that of term spread.

Among the small economies in the Euro area, Ireland has the lowest adjusted R^2, Belgium has the highest adjusted R^2, and Finland has a moderate adjusted R^2 value. Since the average residual error for Ireland is too high, Ireland's model is of no use in making a prediction. The model for Finland is also not so good since it has moderate adjusted R^2 with a reasonably high average residual error. Only the model for Belgium looks good. The average residual error is nearly 43%, and the median residual error is below 20%. When the five observations having the highest residual error are omitted, the average residual error drops to 12%. For Belgium's models, results from two different models are not consistent, making suspicious to confirm that the Belgium's model is also like France's model. The models for Belgium and the Euro area are little similar in the in-sample analysis in terms of term spread's predicting power.

6.3 Implications

The linear regression model with dummy variables outperforms the VAR model, having higher adjusted R^2 and lower residual and prediction errors. Moreover, the linear model provides flexibility to adjust the independent variables, such as testing the model without the lags of GDP growth rate and testing the model using only term spread. While adjusting the independent variables to observe the model's strength, this study finds three pieces of evidence to justify the relatively low predictive power of term spread in the model.

Robustness tests are performed using the linear model. The results from the robustness test led to a few implications, which are presented in the following section. The following table shows the changing adjusted R^2 while altering the exogeneous variables.

Table 7: Robustness Test Results

Linear Model
Endogenous Variable = GDP_t
Lag length Selected = 2

Exogenous Variables	Adjusted R^2
$dTS_t + dTS_{t-1} + dTS_{t-2}$	0.40
$dTS_t + dTS_{t-1} + dTS_{t-2} + dEPU_t + dEPU_{t-1} + dEPU_{t-2} + FC + NIP + HUP$	0.61
$dTS_t + dTS_{t-1} + dTS_{t-2} + dEPU_t + dEPU_{t-1} + dEPU_{t-2} + FC + NIP + HUP + GDP_{t-1} + GDP_{t-2}$	0.93

In table 7, FC refers to financial crisis 2008-9, NIP refers to negative interest rate period, and HUP refers to high uncertainty period. Changing the exogeneous variables results significantly different adjusted R^2 values. The residual errors drastically increase as the adjusted R^2 decreases in the models presented in table 7, indicating that the predictive power of term spread is low based on either model used from robustness tests.

First, the adjusted R^2 of the model deteriorates to 0.40 as only term spread is assigned to the role of the independent variable removing GDP growth rate and all dummy variables. The model's strength indicates that term spread alone in the model is not too good and not completely useless. At this level of adjusted R^2 the predictive error is significantly high, indicating that the predictive power of term spread alone is low. As a result, even the estimated model looks good, the predictive performance of term spread in explaining GDP growth rate is low.

Second, the adjusted R^2 of the linear model is 0.93 when term spread, GDP growth rate, the economic policy uncertainty, and the dummy variables are set as independent variables, but the adjusted R^2 of the model drops to 0.61 when the lags of GDP growth rate are removed from the independent variables. Thus, the drastic reduction in the adjusted R^2 of the model while removing the lags of GDP growth rate shows that the model significantly depends on the lags of GDP growth rate, but not on term spread. Like the adjusted R^2, the estimate coefficient values also indicate that the predictive power of term spread is relatively poor than that of GDP growth rate. For example, the previous two lags of GDP growth rate, namely GDP_{t-1} and GDP_{t-2}, can predict GDP growth rate more accurately than term spread. In the VAR model, the estimate coefficient for GDP_{t-1} is 1.46 and for GDP_{t-2} is -0.62, being statistically significant in explaining the real economic activities. In the linear model, the estimate coefficient for GDP_{t-1} is 1.17 and for GDP_{t-2} is -0.47, again being statistically significant in explaining the real economic activities. When the predictive power of term spread and GDP growth rate are compared to each other, the predictive power of term spread in predicting the real economic activities is relatively low. Such a relatively low predictive power of term spread is common in the major economic areas such as Germany, Italy, France, and Spain. The models' comparative predictive power from small economies such as Ireland, Belgium, and Finland are also not different from those of the major economies, being low. Thus, it is found that the previous lags of the GDP growth rate have higher estimate coefficients than term spread's in the Euro area along with the other seven countries analyzed in this work.

Third, the adjusted R^2 of the linear model changes from 0.93 to 0.92 when only term spread is removed away from that model, suggesting that there is no significant role of term spread to make the model better or worse. On the other hand, the lags of GDP growth rate make a big difference in the model's adjusted R^2. Since the reduction of the adjusted R^2 is not significant when term spread is removed from the model, the reduction of the adjusted R^2 is highly significant when GDP growth rate is removed. Meanwhile, term spread's estimate coefficients are also not statistically significant at a 95% confidence interval, while those of GDP growth rate are significant. Hence, term spread does not have a strong influence on the model compared to GDP growth rate.

It is also observed that the predictive power of term spread during the financial crisis 2008-9 is low but it is not different from average predictive power during whole sample period. This study does not find evidence for lost predictive power of term spread during the crisis. As GDP growth rate turned negative from the fourth quarter of 2008, and it lasted until the last quarter of 2009 in the Euro area. During those five quarters, the average residual error was 10.65%, and the median error was 9%, which is not much different from the full sample median of 15.42%. It suggests that the model is consistent even during the financial crisis period. Term spread has low estimate coefficients, but it is consistent during the financial crisis 2008-9. The case is different in case of right before and right after the crisis since the model fails to maintain its perfection, resulting in extremely higher residual errors. In particular, the extreme residual errors are found in the third quarter of 2008 and the first quarter of 2010, which are one quarter before the GDP growth rate turned negative and the first positive quarter after the recession.

Unlike the financial crisis 2008-9, the recession caused by the European sovereign debt crisis that turned the GDP growth rate negative from the first quarter of 2012 to the second quarter of 2013 yielded completely different results in terms of residual errors. The average residual error for the six quarters, when the GDP growth rates were negative, is 54.18%, and the median error is 49%, showing that the decreasing explanatory power significantly in compared to the previous recession of 2008-9. On the other hand, the term spread turned negative right before the financial crisis, but it stayed positive before and during the recession caused by the debt crisis. A possible reason for a different level of predictive errors during two different recessions could be associated with the long-term interest rate movement. The term spread turned negative for six quarters starting from the third quarter of 2007, but the term spread did not turn negative before or during the recession caused by the debt crisis. After the recovery of the financial crisis 2008-9, investors were already anticipating that the central banks would come up with recovery plans, and, therefore, they were not as hopeless as they were during the financial crisis 2008-9, the confidence that kept term spread non-negative before and during the recession caused by the debt crisis.

During the high uncertainty period, which refers to the quarters with above-average economic policy uncertainty index, the average residual error is 35.72%, and the median residual error is 19.06%, which is slightly different from the full sample median of 15.42%. But when compared to the errors from the period of the financial crisis 2008-9, the errors in the period of high uncertainty increase moderately, suggesting that the high uncertainty period affected the model more than the financial crisis 2008-9 did. The predictive power of term spread weakens as uncertainty rises since the median residual error is slightly higher than the full sample median residual error. Exploring the possible reasons for those changes in predictive power during high uncertainty period is not within the scope of this thesis. Identifying the possible reasons for such weakening of the predictive power of term spread during high uncertainty periods could be an exciting topic for further research.

The negative interest rate period has been in effect since the fourth quarter of 2014. The average residual error during the negative interest rate period is 11.31%, while the median residual error is 6%. The full sample median residual error is 15.42%, which is considerably higher than the median residual error during the negative interest rate period. The negative interest rate period has observed a lower residual error than the residual errors of the major economic events such as the financial crisis 2008-9, the recession caused by the debt crisis, and the high uncertainty periods. Term spread's estimate coefficient is not statistically significant to justify the strong predictive power of term spread during the sample period. This study confirms that the predictive power of term spread has been low during the whole sample period, but the negative interest rate is not a cause of the low predictive power since the model performs better in this period than before the negative interest rate period. Thus, term spread's predictive power has slightly improved during the negative interest rate period; however, term spread's predictability has not reached the level of statistical significance. This finding is somehow similar to the findings of Hyozdenska (2015b), which studies the cases for European countries such as the UK, Iceland, Norway, Switzerland, and Russia. Hyozdenska (2015b) examined the relationship between term spread and the economic activities of selected countries between 2000 to 2013, and found the different predictive power of term spread before and after the financial crisis. Term spread is the only variable used as the independent variable in simple linear regression to explain the real economic activities, while this thesis covers much wide and deep scope in terms of variables and models applied. The common finding from Hyozdenska (2015b) and this study is that both studies find that the predictive power of term spread has been low, and it has been changing, particularly increasing. However, the reference point of change and the set of countries studied are different in this thesis.

While re-examining the predictive ability of term spread, short-term interest rates, and the stock returns for the real GDP growth in the G – 7 countries, Kuosmanen and Vataja (2017) concluded that the financial variables have regained predictive power since the financial crisis. Since Germany, Italy, and France are in the G7 and the Euro area both, their finding is somehow look still relevant, at least for term spread in the Euro area. Indeed, term spread is gaining its predictive power in Euro area. In the other hand, this thesis also discovers slightly different result from Kuosmanen and Vataja (2017) because the predictive power of term spread had decreased while the recession caused by the debt crisis, and that period falls after the financial crisis. This difference might be due to influence of different set of countries, mainly the USA. Despite the small difference, the common point is that the predictive power of term spread is improving. This thesis also comes to the similar conclusion that even the predictive power is still low, the predictive power of term spread has slightly increased during the negative interest rate period.

In line with Hyozdenska (2015b) and Kuosmanen and Vataja (2017), Hyozdenska (2015a) noted improved predictive power of term spread in her second half of the sample. She examined the relationship between the term spread and the economic activity of the United Kingdom, Iceland, Switzerland, Norway, and Russia between the years 2000 and 2013. She divided the sample into two parts: before 2008 and after 2008. She observed the poor predictive power of the yield curve in the first part of the sample, and it increases after 2008 in Iceland, Russia, and Great Britain. Since this thesis also finds the improving predictive power lately, including during the negative interest rate period.

This study does not find a piece of strong evidence for the increased predictive power of term spread when augmented with EPU, instead, the ARCH effect observed on the model without EPU disappeared with the inclusion of EPU. Except that, EPU has no significant contribution in the model.

Based on the full sample data, the first differenced term spread Granger causes the GDP growth rate at the optimum lag length of two. The significant causation can be seen up to three lags, while weak causation is observed for fourth, fifth, and sixth lags too. The GDP growth rate poorly Granger causes term spread for the optimum lag length of two. So, the bidirectional Granger causality is seen; however, the causation's strength differs significantly depending on the direction. This result is consistent with the time before and during the negative interest rate period. Term spread strongly Granger causes the GDP growth rate at shorter lags, optimally for lag two, and the GDP growth rate poorly Granger causes term spread. EPU does not Granger cause the GDP growth rate at any lag.

7 CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

Motivated by the unique context developed by series of major economic events such as financial crisis 2008-9, the recession caused by the debt crisis, high uncertainty period, and negative interest rate period in the Euro area, this study re-examines the predictive power of term spread in the Euro area. Term spread and GDP growth rate are two key variables; however EPU is also added to the model for model's improvement. The full sample of this thesis ranges from 1999Q1 to 2019Q4, in which the in-sample model fit is based on. A fraction of the sample that ranges from 1999Q1 to 2014Q3 is termed as the sample set 1, in which the out of sample prediction is based on. First, the VAR model is estimated using the endogenous variables; term spread, GDP growth rate, and EPU, and then the linear model is estimated with three dummy variables such as financial crisis 2008-9, high uncertainty period, and negative interest rate period. Then, the Granger causality is tested among the variables under study.

Term spread has low predictive power in the Euro area since the estimate coefficients for term spread are mostly not found to be statistically significant. The model's adjusted R^2 does not change by much when term spread is removed from the independent variables, but the adjusted R^2 drops from 0.93 to 0.61 as the lags of GDP are removed from the independent variables, indicating that the real economic activities in the Euro area can be better predicted by GDP growth rate's lags than by term spread. Estimate coefficients for EPU are almost null, indicating that the inclusion of EPU in the model does not improve the predictive power of the model.

The predictive power of term spread worsens further during the recession caused by the debt crisis and during the high uncertainty period. The result is slightly different during the financial crisis 2008-9 since the predictive power is not found to be as bad as during the recession caused by the debt crisis and the high uncertainty period. The negative interest rate period has no role in deteriorating the predictive power of term spread. Surprisingly, term spread has relatively better predictive power during the negative interest rate period; however, the estimate coefficients are not statistically significant yet.

Term spread Granger causes the GDP growth rate in up to three lags before the negative interest rate period; however, term spread Granger causes the GDP growth rate up to six lags during the negative interest rate period. This means that the relationship between term spread and the GDP growth rate is slightly improving during the negative interest rate period. A very weak bidirectional Granger causality between term spread and the GDP growth rate is observed, while EPU does not Granger cause the GDP growth rate at all.

Indeed, this study is not free from limitations, as many other scientific studies are. So, limitations that are identified by the author are discussed in this section. First, the sample of this study excludes the recent observations after 2019Q4. The exclusion of the observation is due to the COVID 19 pandemic, assuming the extreme changes in GDP growth rate during the pandemic could lead to a misleading conclusion of the study. Second, term spread used in this study is the difference of ten years and three months interest rates; however, there can be many variations of term spread, leading to different results than the result obtained from this study.

Moreover, this study completely ignores whether other variations of term spread could have different predictive power than the predictive power of term spread used in this study. Similarly, this study also does not consider whether the explanatory and predictive power of only short rate instead of term spread could be different, deriving a completely different conclusion than this study has made. Third, this study partially justifies the theoretical reasoning for the chosen dummy variables. However, it is also true that the selection decision is partly based on the author's decision, which certainly leaves the possibility of many or fewer dummy variables that could have been added or removed from the study. Fourth, this study could also have been conducted using other empirical methods than the VAR model. The results could have been different to some extent, as the previous literature suggests that the model chosen for the study can also impact the findings of the study to some extent.

This study can be extended to several paths; some of them are discussed in this section. One possible path could be that term spread's predictive power can be tested for the Euro area considering the COVID 19 pandemic period. The new study, including the COVID19 period, certainly can offer more useful results for the stakeholders than this study does, as it includes the bigger sample size and additional information in the model. Doing so can improve the quality of this work to a higher level and contribute significantly to academia, providing precise information about term spread's predictive power during multiple extreme contexts.

Another possible path is to test whether term spread of the four European core economies could predict the real economic activity for the whole Euro area since the four core economies cover significant population, size, and GDP of the Euro area. The finding could help to understand the changing context of integration of country-level-economies to the Euro area economy. Nevertheless, another interesting issue to observe could be to assess whether the core economies are still as significant in driving the Euro area's economic activities as they had been.

The third possible path to extend this study is to test term spread's predictive power in the US economy and the Euro area economy. During Donald Trump's presidency, the United States economy has been affected by extreme political decisions such as stricter foreign policies and been affected by sore trade relations with major economies, and the economic consequences of COVID 19 pandemic has increased pressure in the economy. Such a newly developed context in the United States could have altered the predictive power of the US term spread in predicting their real economic activities. Since the US economy impacts other significant economies globally, the comparative study of term spread's predictive power between the Euro area and the US could lead to some useful findings.

REFERENCES

Anderson, N. & Sleath, J, 2001. New Estimates of the UK real and nominal yield curves. Bank of England working papers.

Argyropoulos, E. & Tzavalis, E. 2016. Forecasting economic activity from yield curve factors. North American Journal of Economics and Finance, Volume 36, 293-311.DOI: 10.1016/J.NAJEF.2016.02.003.

Baker, S. R. & Bloom, N.&Davis, S. J. 2016. Measuring Economic Policy Uncertainty. The Quarterly Journal of Economics, Volume 131, Issue 4, November 2016, Pages 1593-1636, https://doi.org/10.1093/qje/qjw024

Bismans, F. & Majetti, R. 2011. Forecasting recessions using financial variables: the French case. Empirical Economics, 44, 419-433. DOI: https://doi.org/10.1007/s00181-012-0550-z.

Caldara, Dario, Matteo Iacoviello, Patrick Molligo, Andrea Prestipino, and Andrea Raffo (2019). "Does Trade Policy Uncertainty Affect Global Economic Activity?," FEDS Notes. Washington: Board of Governors of the Federal Reserve System, September 4, 2019, https://doi.org/10.17016/2380-7172.2445.

Chen, G., Valadkhani, A. & Grant, B. 2016. How useful is the term spread as a predictor of growth in Australia? Journal of Economic Studies, Volume 43(2), 222-241. DOI: https://doi.org/10.1108/JES-09-2014-0159.

Chionis, D., Gogas, P. & Pragidis, I. 2010. Predicting European Union Recessions in the Euro Era: The Yield Curve as a Forecasting Tool of Economic Activity. International Advances in Economic Research, 16, 1-10. DOI: https://doi.org/10.1007/s11294-009-9247-2

Dar, A. B., Samantaraya, A. & Shah, F. A. 2014. The predictive power of term spread: evidence from wavelet analysis. Empirical Economics, 46, 887-901. DOI: 10.1007/s00181-013-0705-6.

De Boor, C. , De Boor, C., Mathematicien, E.U., De Boor, C., & De Boor, C., 1978. A Practical Guide to Splines (Vol. 27, p. 325). New York: springer-verlag. Revised Edition 2003.

Dong, K. & Park, E. 2018. Has the Predictability of the Term spread Changed? Seoul Journal of Economics, Volume 31(4), 449-463.

ECB 2020. Structure of the euro area economy https://www.ecb.europa.eu/mopo/eaec/html/index.en.html#:~:text=Key%20characteristics,the%20United%20States%20and%20China.

Estrella, A. & Hardouvelis, G. A. 1991. The term structure as a predictor of real economic activity. Journal of Finance, 46, 555-576.

Estrella, A. & Mishkin, F. S. 1997. The predictive power of the term structure of interest rates in Europe and the United States: Implications for the European Central Bank. European Economic Review, 41, 1375-1401.

Estrella, A. & Mishkin, F.S. 1998. Predicting US recessions:financial variables as leading indicators. The Review of Economics and Statistics, 80, 45-61.

Eurostats 2020 Data Retrievd on 24 August 2020 from the Europa official webpage. https://ec.europa.eu/eurostat/tgm/refreshTableAction.do;jsessionid=9ea7d07 e30dd3bf0a52b9a8a474c872db039e243c026.e34OaN8Pc3mMc40Lc3aMaNyTa3eQe0?tab=table&plugin=1&pcode=tec00001&language=en

Evgenidis, A. & Siriopoulos, C. 2014. Does the term spread retain its forecasting ability during the 2007 recession? A comparative analysis. Applied Economic Letters, Volume 21(12), 817-822. https://doi.org/10.1080/13504851.2014.884694.

Evgenidis, A., Papadamou, S. & Siriopoulos, C. 2020. The yield spread's ability to forecast economic activity: What have we learned after 30 years of studies? Journal of Business Research, 106, 221-232. DOI: 10.1016/j.jbusres.2018.08.041.

Fulllerton, T. M., Saenz-Rojo, E. D. & Walke, A. G. 2017. Yield spreads, currency movements, and recession predictability for southern border economies in the United States. Applied Economics, 49:30, 2910-2921. DOI: 10.1080/00036846.2016.1251556.

Gogas, P. & Pragidis, I. 2011. GDP trend deviations and the term spread: the case of eight EU countries. Journal of Economics and Finance, 36, 226 – 237. DOI: 10.1007/s12197-011-9176-9

Gogas, P., Papadimitriou, T. & Chrysanthidou, E. 2015a. Yield curve point triplets in recession forecasting. International Finance, 18(2), 207-226. DOI: 10.1111/infi.12067.

Gogas, P., Papadimitriou, T., Matthaiou, M. & Chrysanthidou, E. 2015b. Yield Curve and Recession Forecasting in a Machine Learning Framework. Computational Economics, 45, 635-645.

Granger, C.W.J. 1969. Investigating Causal Relations by Econometric Models and Crossspectral Methods. The Econometrica, Vol. 37, No. 3, 424-438. https://doi.org/10.2307/191291.

Gupta, R., Hollander, H., and Steinbach, R. 2020. Forecasting output growth using a DSGE-based decomposition of the South African yield curve. Empirical Economics 58:351-378. DOI: 10.1007/s00181-018-1607-4.

Harvey, C. 1988. The real Term Structure and Consumption Growth. Journal of Financial Economics, 22(2), 305-333. DOI: https://doi.org/10.1016/0304-405X(88)90073-6

Hasegawa, M. & Fukuta, Y. 2009. An empirical analysis of information in the term spread on future recessions in Japan. Applied Economics, Volume 43, Issue 15. DOI: 10.1080/00036840902780136

Hännikäinen, J. 2017. When does the yield curve contain predictive power? Evidence from a data-rich environment. International journal of Forecasting, 33, 1044-1064. DOI: https://doi.org/10.1016/j.ijforecast.2017.05.006

Hvozdenska, J. 2015a. The Yield Curve as a Predictor of Gross Domestic Product Growth in Nordic Countries. Procedia Economics and Finance, Volume 26, 438-445. https://doi.org/10.1016/S2212-5671(15)00871-0

Hvozdenska, J. 2015b. The Application of Sovereign Bond Spreads: The Case of United Kingdom, Iceland, Norway, Switzerland and Russia. Procedia Economics and Finance, Volume 23, 1227-1232.DOI: 10.1016/S2212-5671(15)00476-1

Jardet, C. 2004. Why did the term structure of interest rates lose its predictive power? Economic Modelling, Volume 21, Issue 3, Pages 509-524. DOI: https://doi.org/10.1016/S0264-9993(03)00042-7

Karlsson, S. and Österholm, P. 2020. The relation between the corporate bond-term spread and the real economy: Stable or time-varying? Economic Letters 186 (108883). DOI: https://doi.org/10.1016/j.econlet.2019.108883.

Kozicki, S. 1997. Predicting real growth and inflation with the term spread. Economic Review, Federal Reserve Bank of Kansas City, Fourth Quarter, 82(4).

Kuosmanen, P. & Vataja, J. 2017. The return of financial variables in forecasting GDP growth in the G-7. Economic Change and Restructuring, Volume 50(3), 259-277. DOI: https://doi.org/10.1007/s10644-017-9212-7.

Kuosmanen, P., Rahko, J. & Vataja, J. 2019. Predictive ability of financial variables in changing economic circumstances. North American Journal of Economics and Finance, Volume 47, 37-47. DOI: https://doi.org/10.1016/j.najef.2018.11.012.

Lutkepohl, H. 2010. Macroeconometrics and Time Series Analysis. The New Palgrave Economics Collection, 369-371.

Lutkepohl. H. 2011. "Vector Autoregressive Models," Economics Working Papers ECO2011/30, European University Institute.

Mishkin, F. S. 2011. The economics of money, banking, and financial markets. 4th Canadian edition. Page 113-27.

Mishkin, F. S. 2011. The economics of money, banking, and financial markets. 4th Canadian edition. Page 661-55.

Moersch, M. & Pohl, A. 2011. Predicting recessions with the term spread – recent evidence from seven countries.Applied Economics Letters, Volume 18, Issue 13. DOI: 10.1080/13504851.2010.534055.

Morell, J. 2018. The decline in the predictive power of the US term spread: A structural interpretation. Journal of Macroeconomics, Volume 55, 314-331. DOI: 10.1016/j.jmacro.2017.12.003.

Nelson, C. R. & Siegel, A. F. 1987. Parsimonious Modeling of Yield Curves. The Journal of Business, Vol. 60, No. 4 (Oct., 1987), pp. 473-489. DOI: 10.1086/296409 .

Nymand-Andersen, P. 2018. Yield curve modelling and a conceptual framework for estimating yield curves: Evidence from the European Central Bank's yield curves, ECB Statistics Paper, No. 27

OECD 2020. OECD data, finance, interest rates. Retrieved on 24 August 2020. https://data.oecd.org/finance.htm#profile-Interest%20rates

Papadamou, S. 2009. Term spreads and real economic activity in East European transition economies. Applied Economics Letters, Volume 16, Issue 5. DOI: 10.1080/13504850601032123.

Paya, I., Venetis, I. A. & Peel, D. A. 2004. Asymmetry in the Link between the Term spread and Industrial Production: Threshold effects and Forecasting. Journal of Forecasting, 23, 373-384. DOI: 10.1002/for.921.

Pena, J. I. & Rodriguez, R. 2006. On the Economic Link Between Asset Prices and Real Activity. Journal of Business Finance & Accounting, 34(5) & (6), 889-996. DOI: 10.2139/ssrn.726301

Schunk, D. 2011. Probability predictions of rising real GDP growth and inflation: the usefulness of monetary indicators. Applied Economics, Volume 40, Issue 9. DOI: 10.1080/00036840600771247 .

Sims, E. 2011. Graduate Macro Theory II: Notes on time series. University of Notre Dame.

Stuart, R. 2020. The term structure, leading indicators, and recessions: evidence from Switzerland, 1974–2017. Swiss Journal of Economics and Statistics, 156:2. DOI: 10.1186/s41937-019-0044-4 .

Svensson, L. 1994. Estimating and interpreting forward interest rates: sweden 1992-1994. Working paper 4871. NBER. DOI: 10.3386/w4871

Taylor, J. B. 1993. Discretion versus policy rules in practice. Carnegie-Rochester Conference Series on Public Policy 39, 195 – 214. DOI: https://doi.org/10.1016/0167-2231(93)90009-L

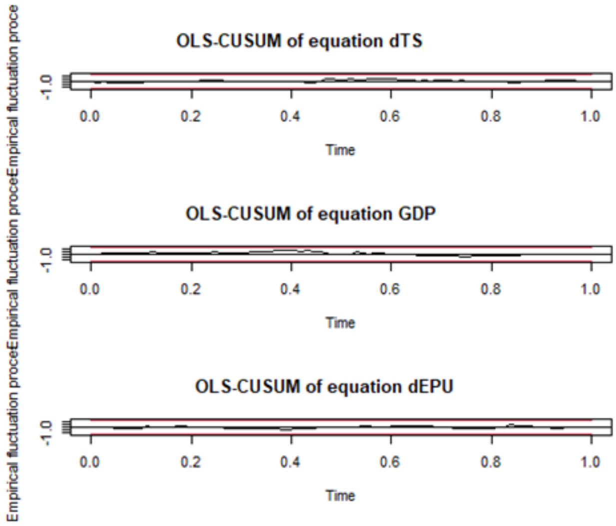
The world fact book 2020. https://www.cia.gov/library/publications/the-world-factbook/geos/gm.html

Waggoner, Daniel F. (1997) : Spline methods for extracting interest rate curves from coupon bond prices, Working Paper, No. 97-10, Federal Reserve Bank of Atlanta, Atlanta, GA.

Yang (2005), Can the strengths of AIC and BIC be shared?, Biometrika, 92: 937-950. DOI: 10.1093/biomet/92.4.937 .

APPENDIX 1

1. OLS based CUSUM test for Stability (See page 32, Table 2)



2. The Residual Error Analysis of Selected Countries

	Germany		France		Italy		Spain	
	VAR	LM	VAR	LM	VAR	LM	VAR	LM
ARE	62.24%	136.6%	50.42%	50.63%	424.7%	455.2%	157.2%	192.3%
MRE	32.49%	33.10%	19.76%	20.93%	33.15%	27.22%	65.73%	78.23%
ARE*	45.40%	54.8%	27.04%	27.04%	50.28%	54.81%	103.7%	142.3%

	Ireland		Belgium		Finland	
	VAR	LM	VAR	LM	VAR	LM
ARE	113.7%	125.7%	43.26%	43.25%	81%	81%
MRE	94.5.5%	96.36%	12.41%	19.84%	28.62%	28.62%
ARE*	94.19%	97.18%	19.84%	12.40%	55.34%	55.33%

ARE = Average Residual Error, MRE = Median Residual Error, ARE* = Average Residual Error Without 5 Outliers

3. The Prediction Error Analysis of Selected Countries

	Germany		France		Italy		Spain	
	VAR	LM	VAR	LM	VAR	LM	VAR	LM
APE	114%	83.43%	48.75%	30.81%	51397%	108%	86.3%	175.5%
MPE	44.13%	55.58%	42.93%	23.73%	492%	44.5%	88.7%	78.05%
APE*	39.43%	40.15%	32.57%	19.91%	385%	37.2%	74.2%	85.68%

	Ireland		Belgium		Finland	
	VAR	LM	VAR	LM	VAR	LM
APE	99.3%	94.43%	82.46%	17.2%	156%	394%
MPE	100.1%	93.93%	91.70%	17.25%	103.4%	139%
APE*	98.03%	68.33%	72.31%	13.37%	87.18%	98.65%

APE = Average Prediction Error, MPE = Median Prediction Error, ARE* = Average Prediction Error Without 5 Outliers