Economics School of Louvain - ESL

Economics School of Namur - ESN

Macro-Finance Model of the UK Yield Curve

A Macroeconomic Interpretation using Principal Components Analysis

Author : Marcin Alexandre

Thesis Director : Oikonomou Rigas

Thesis Reader : De Combrugghe De Pickendaele Alain

Academic Year 2023-2024

Master in Economics – 120 credits – Focus : *Professional (ECON2MS/G)*

Université catholique de Louvain - Economics School of Louvain - Place Montesquieu 3, 1348 Louvain-La-Neuve
Université de Namur – Economics School of Namur – Rempart de la Vierge 8, 5000 Namur

Macro-Finance Model of the Yield Curve:
A Macroeconomic Interpretation using Principal Components Analysis

Université Catholique de Louvain & Université de Namur

Economics School of Louvain

Academic Year 2023–2024
Author : Marcin Alexandre
Thesis Director : Oikonomou Rigas

Abstract

We employed Principal Component Analysis (PCA) to model the key factors influencing the British yield curve, integrating these factors with macroeconomic variables (output, inflation, and the monetary rate) within a recursive Vector Autoregressive (VAR) framework. Our analysis explored the dynamics of these factors through impulse response functions and forecast variance decomposition. We constructed various models using both differenced and raw data, examining distinct time periods: 1993-2006 and 2007-2024. This approach allowed us to assess the interaction between yield curve factors and macroeconomic variables, and to evaluate the robustness and predictive power of our models across different economic conditions.

We observed a shift in the monetary policy behaviours in managing the yield curve. Prior to the GFC, monetary policy focused on stabilizing the shape of the yield curve, emphasizing growth and investment rather than inflation control. Post-GFC, the central bank's strategy pivoted towards managing short-term liquidity and accommodating changes in the yield curve's slope and curvature, reflecting a heightened sensitivity to inflation and a constrained ability to influence long-term interest rates.

Acknowledgments

I would like to thank professor Oikonomou Rigas, the supervisor of this master thesis, for his invaluable guidance and insightful advice throughout the research process. His expertise and support were instrumental in shaping the direction and quality of this work. I would also like to express my gratitude to professor de Crombrugghe de Picquendaele Alain, the reader, for dedicating his time to review and evaluate the concluding work of this academic journey.

Merci également à toutes les personnes qui, de près ou de loin, m'ont soutenu au cours de l'écriture de ce mémoire et tout au long de mes études. De mes parents à mes amis les plus proches, chacun de vous a contribué à cet aboutissement, et je vous en remercie sincèrement. Je souhaite adresser un merci tout particulier à Maman, Papa, Guillaume et Talïa pour votre soutien inestimable.

Generative Artificial Intelligence Disclaimer

This master thesis incorporates the use of several Artificial Intelligence (AI) tools for specific tasks to enhance the quality of the work. The following tools were utilized:

- **OpenAI GPT-4**: Used for the reformulation of sentences in English. This tool helped improve clarity and coherence in writing.
  Retrieved from https://www.openai.com
- **DeepL**: Employed for the translation and reformulation of scientific papers in French. This tool ensured accurate translation and interpretation of relevant scientific literature.
  Retrieved from https://www.deepl.com
- **LanguageTool**: Utilized for grammar and style checking to enhance the overall readability and correctness of the text.
  Retrieved from https://www.languagetool.org

These AI tools were used as auxiliary aids and not as primary sources of information or analysis. The substantive content, analysis, and conclusions of this thesis are the sole responsibility of the author, in accordance with academic and scientific integrity practices.

Contents

1 Introduction

The yield curve, a graphical representation of interest rates across different maturities, is a fundamental tool in finance and economics, providing useful insights into the economic landscape. Accurately modelling the yield curve is crucial because it directly influences key financial and economic decisions. A precise model helps in forecasting future interest rates, which is essential for setting monetary policy, guiding investment decisions, and managing financial risk. By understanding the shape and movement of the yield curve, economists, and policymakers can better assess economic conditions, predict economic downturns or expansions, and implement strategies that stabilize or stimulate the economy. For financial professionals, accurate yield curve modelling supports effective risk management and strategic planning, enabling them to optimize investment portfolios and adapt to changing market conditions. In an increasingly complex and interconnected global market, the ability to model the yield curve accurately provides a competitive edge and helps in making informed, strategic decisions.

Over the decades, economists have developed various models to interpret the yield curve, including the Nelson-Siegel (NS) and Dynamic Nelson-Siegel (DNS) models. The NS model provides a parsimonious representation of the yield curve through a few factors (level, slope, and curvature) while the DNS model enhances this framework by incorporating time-varying factors to capture dynamic economic conditions. Despite their utility, these models often struggle with complexity and adaptability in changing economic environments. This thesis explores Principal Component Analysis (PCA) as a simpler yet robust alternative for yield curve modelling. PCA excels in reducing the dimensionality of yield curve data while preserving essential correlations, allowing for a clearer extraction of key underlying factors. By comparing PCA with traditional models, we aim to evaluate its effectiveness in capturing the primary drivers of the yield curve and its interaction with macroeconomic variables.

Our analysis will assess PCA's ability to provide accurate forecasts and insights into yield curve dynamics across different economic periods, highlighting its potential as a practical tool for both theoretical and applied finance.

Among the various models available for yield curve analysis, we opted for a vector autoregressive (VAR) model. This empirical approach enables us to jointly model the financial factors identified earlier alongside key macroeconomic variables. The VAR model's strength lies in its capacity to capture the dynamic relationships between multiple time series, making it well-suited for our study. It allows us to account for the feedback effects and interdependencies between the yield curve components and macroeconomic indicators such as output, inflation, and monetary policy rates. By employing this model, we aim to gain a deeper understanding of how shocks in the financial system or macroeconomic environment propagate and affect the yield curve's shape. Our analysis is conducted using both first-differenced data and non-differenced data, covering two distinct periods: 1993-2006 and 2007-2024. This approach allows us to explore potential structural changes and the impact of different economic conditions on the yield curve dynamics.

The structure of this thesis is organized as follows: Chapter 2 provides an in-depth literature review, covering the foundational theories of the yield curve, various modelling approaches, and the mechanisms identified in previous Macro-Finance models. Chapter 3 details the methodological framework, including the specific models employed, the application of Principal Component Analysis (PCA), and the measurement techniques utilized. Chapter 4 discusses the data set used in the study, presenting an initial statistical analysis of the identified factors and macroeconomic variables. Chapter 5 presents the findings from the different models implemented, highlighting the key differences and similarities observed. Additionally, this chapter evaluates the effectiveness of the PCA method in fitting and forecasting the data. Chapter 6 concludes the thesis by addressing the limitations of the analysis and suggesting areas for future research.

2 Literature Review

Bond market

The bond market plays a pivotal role in the global financial landscape, serving as a crucial platform for governments, corporations, and institutions to raise capital. Bonds are debt securities that represent a loan made by an investor to a borrower, typically a government or a corporation. In essence, when an investor purchases a bond, they are lending money to the issuer in exchange for periodic interest payments and the return of the principal amount at maturity. We assume a rate between the face value (value of the repayment) and the present value (value of the bond today on the secondary market). Therefore, we distinguish two types of rate, the spot and the forward ones.

Spot rates represent the yield on zero-coupon bonds at a specific maturity. These rates are the discount rates used to determine the present value of a future payment. Spot rates are crucial in financial markets as they reflect the current cost of capital, influencing decisions related to borrowing, lending, and investment. Essentially, spot rates provide a snapshot of the current interest rate environment, guiding immediate financial decisions without incorporating future expectations.
Forward rates, on the other hand, are derived from spot rates and represent the expected future interest rates. They are calculated based on the assumption that future loans are influenced by current spot rates. For example, if the current six-month borrowing rate is lower than the one-year rate, a borrower might lock in a favourable forward rate for a future period by strategically borrowing and investing. Forward rates enable financial planning by providing a means to hedge against potential interest rate fluctuations. (Bank of England, 2023)

We focus, in this dissertation, on spot rates for two main reasons. First, as mentioned earlier, forward rates are directly constructed from zero-coupon bonds, within the arbitrage-free framework, making spot rates fundamental to their understanding. Second, almost any interest rate derivative can be constructed using a portfolio of zero-coupon bonds. For instance, interest rate swaps, bond options, and futures contracts often rely on the spot rate curve for pricing and valuation. This makes spot rates a critical component in the analysis and construction of various financial instruments and derivatives. Additionally, spot rates offer a clear view of the yield curve, allowing for a straightforward interpretation of the term structure of interest rates. (Dziwum & Green, 1996)

The yield curve represents the relationship between the interest rates of different maturities. The shape of the latter informs us about the prevailing economic conditions and market expectations. When the yield curve slopes upward, with long-term rates higher than short-term rates, it typically reflects the market's anticipation of economic growth. Conversely, when the yield curve inverts, with short-term rates exceeding long-term rates, it often signals concerns about an economic downturn. A flat yield curve, where short- and long-term rates are similar, may suggest uncertainty. (Reserve Bank of Australia, n.d.) The general shape of the curve is called an upward sloping curve and can be described through several theories.
According to Mishkin, 2007, there are three main stylized facts that characterize the term structure of nominal spot rates. The first one was discussed above, the shape represents the short — long rate differential. But furthermore, all maturities tend to move together and the yield curve "normal" shape is upward sloping. We will now review the theories that explain these stylized facts. After what, we will review some of the most important methods to model this term structure.

Theories about the Yield Curve

The oldest theory issued from economic literature is the expectation theory, originally stated by Irving Fisher in early XX's. Long-term interest rates are formed on the basis of the expectations of future short-term rates. The implications for the yield curve are that its shape reflects market expectations regarding future interest rate movements. A flattening or steepening of the yield curve may be indicative of shifts in these expectations. The Fisher equation states that the nominal interest rate is equal to the real rate plus expected future inflation. The nominal rate then incorporates expectations about future interest rate and future inflation. This theory assumed, for "perfect substitute" securities, that investors don't prefer one maturity over another, because expected returns of holding bonds are the same independently of it. Expectation theory explains why the term structure is able to take different shapes, and why the rates tend to move together (Omondi, 2015). If the longer maturity rates are seen as the average of the shorter ones, the (nominal) rates are expected to rise in the future and inversely if they are lower. However, the Expectations Hypothesis has several significant limitations. Firstly, it does not adequately explain why the yield curve typically has an upward-sloping "normal" shape. According to the theory, if long-term rates were merely an average of expected future short-term rates, the yield curve would be flat unless investors expected rising short-term rates. Additionally, the theory overlooks the presence of a risk premium in long-term rates. It assumes that investors only seek compensation for expected future short-term rates when holding long-term bonds. In reality, investors also demand a premium to compensate for uncertainties about future inflation and interest rate fluctuations. The theory also assumes that investors have no preferences for bonds of certain maturities over others. This neglects the fact that investors might favour specific maturities, which can influence the yield curve independently of expectations about future short-term rates. This leads us to consider alternative theories that address these shortcomings.

The shape of the yield curve can be described through the view of the liquidity preference theory, originally stated by John M. Keynes and later adapted by Modigliani, 1944. Investors require a premium to hold longer maturities securities against shorter ones since they are less liquid, which may be more challenging to sell or convert to cash quickly, especially in times of market stress. The theory leaves more place to the investors' risk preferences, explaining the presence of a term premium. It is the additional yield required by investors to hold long-term maturity bonds against shorter ones. The greater the maturity is, the greater the uncertainty about the future economic conditions should be. The overall risk of holding these long-term maturities is higher, giving interpretation to the upward-sloping shape, but not why the shape can become inverted.

These theories allow us to compute the term premium. We are able to decompose it in two components, in the inflation premium and in the liquidity premium. The inflation one can be calculated by testing the Fisher equation econometricaly, and the other by regressing inflation compensation on a liquidity measure (Gürkaynak et al., 2010). TIPS are inflation-indexed US bonds, the inflation compensation thus being the implied inflation rate of holding TIPS against nominal bonds. The authors showed that the inflation compensation is equal to inflation expectation plus inflation premium minus liquidity premium.

The segmented markets theory provides an alternative explanation for the shape of the yield curve, focusing on the independent determination of interest rates across different maturities. According to Culbertson, 1957, four key factors influence these variations: the bond's liquidity (liquidity premium), expected future changes in the bond's price (inflation premium), changes in the maturity structure, and the cost of acquiring and managing the security. In this framework, investors are primarily concerned with expected yields and are willing to invest in bonds of any maturity based on these expectations, leading to interest rates being driven by the supply and demand dynamics specific to each maturity. The theory is built on three core assumptions: the market is segmented by maturity, with investors operating within preferred sectors; investors are generally indifferent to shifts in maturity preferences, fostering stability within these segments; and stable demand and supply within each sector result in independent dynamics, with sector-specific factors predominantly influencing yields. While this theory effectively explains the yield curve's shape by highlighting the role of supply-demand conditions and maturity premiums, it does not account for the tendency of interest rates across different maturities to move together in response to economic shocks.

The preferred habitat theory, as proposed by Modigliani and Shiller, 1973, extends the expectation hypothesis by introducing the concept of investor preferences for certain maturity segments. According to this theory, while investors may have preferred maturities, they are not strictly confined to these preferences. Instead, they are willing to shift to other maturities when the yield differential and perceived risks justify such a move. This theory allows for a more flexible understanding of the term structure, where the long-term interest rate is viewed as a linear combination of expected future short-term rates, factoring in forecast errors from previous periods. Thus, the yield curve is influenced not only by expectations of future rates but also by the yield differential across maturities.

However, the preferred habitat theory alone does not fully explain the observed behaviour of yield curves. If the theory were solely accurate, yield curves would be dominated by segmented demand and supply, potentially leading to significant differences between neighbouring maturities. In reality, we observe that yields across different maturities tend to move together in response to economic shocks, suggesting the presence of a term premium. Vayanos and Vila, 2021 highlight the role of arbitrageurs in this process, showing how they help transmit shocks across different maturity segments, thus influencing the term premium. Arbitrage is crucial because it helps to correct mismatches in yield differentials across maturities, ensuring that the yield curve remains consistent with the principles of financial theory. An arbitrage-free yield curve is essential for maintaining market efficiency, as it aligns with the efficient market hypothesis, and supports the consistent modelling of the general equilibrium in financial markets.

Model the Yield Curve

In financial literature, three main types of models are commonly used to describe the term structure of interest rates: equilibrium models, affine no-arbitrage models, and econometric/parametric models. An early example of an equilibrium model is provided by Vasicek, 1977, which offers a formula to determine interest rates for various maturities based on specific factors at a given time. This model belongs to the family of Ornstein-Uhlenbeck processes, a type of stochastic process characterized by mean-reverting behaviour. The Vasicek model utilizes this mean reversion and assumes a term premium, resulting in interest rates that converge to a long-term average. Another foundational equilibrium model is the Cox-Ingersoll-Ross (CIR) model, developed by Cox et al., 2005, which builds on the Vasicek model by incorporating additional factors such as investor expectations and risk aversion. The CIR model integrates the term structure within a broader intertemporal asset pricing equilibrium, aligning with the preferred habitat theory by accounting for

preferences among different maturities and assets.
The second category, affine no-arbitrage models, builds upon equilibrium models by incorporating an arbitrage-free bond pricing framework. Hull and White, 1990 extend the Vasicek and CIR models to better align with market prices and interest rate volatilities, allowing for the accurate pricing of interest rate derivatives. Another significant contribution in the no-arbitrage framework is the Heath-Jarrow-Morton (HJM) model (Heath et al., 1992), which provides a methodology for valuing contingent claims—derivatives that depend on uncertain future events. The no-arbitrage condition, a key feature of these models, ensures that there are no opportunities for risk-free profits, in line with the efficient market hypothesis. This principle dictates that the asset price must equal the expected value of its discounted future cash flows.

This review covers only a few of the most significant and pioneering models, though there are numerous others tailored to specific markets, situations, derivatives, or financial instruments. While equilibrium models effectively capture the dynamics of the yield curve, they require strong assumptions about premiums to be solved. Affine no-arbitrage models, on the other hand, can accurately fit interest rates at a particular point in time but struggle to model the temporal dynamics of the yield curve. Consequently, the focus of this thesis will be on the third category of models: econometric/parametric models.
One of the earliest and most widely used models in this category is the Nelson-Siegel (NS) model, introduced by Nelson and Siegel, 1987. This model is favoured in both academic research and by central banks due to its flexibility and the ease with which it can generate economically interpretable factors. The NS model assumes that the yield curve can be decomposed into three distinct factors, typically representing the long-term, medium-term, and short-term components of the yield curve, which are then estimated based on a predetermined functional form over various maturities. These factors offer a parsimonious and intuitive representation of the yield curve, enhancing its utility for economic analysis. Additionally, the NS model provides better in-sample predictions of the actual yield curve compared to earlier affine models. The model was later extended by Svensson, 1995 to include a fourth factor, further increasing the model's flexibility in estimating the curvature of the

yield curve.

The Nelson-Siegel (NS) model, in its original form, features a static expression with a constant decay rate, which represents the speed at which a variable approaches zero over time. However, Diebold and Li, 2006 introduced a dynamic version of the Nelson-Siegel model (DNS) that incorporates a time-varying decay rate and factors. This dynamic adaptation is achieved using a Kalman filter, which estimates evolving parameters over time by iteratively updating the decay factor based on new observations. This method is particularly effective for capturing changes in the underlying dynamics of the yield curve. To model the yield curve dynamics, the DNS approach employs an autoregressive function AR(1) applied individually to each factor, rather than using a Vector Autoregression (VAR) model for all factors. This choice is motivated by two considerations: first, an unrestricted VAR model tends to suffer from poor out-of-sample performance due to overfitting; second, the factors exhibit minimal cross-correlation, making the separate AR process more appropriate for accurate forecasting.

The limitation of such model is that we impose factors that are taken from the economic theory, not from the data. There is often high correlation between the different yields, thus between the factors, after some market events, which leads to difficulties in the estimation of the parameters (Oprea, 2022). As done by Litterman and Scheinkman, 1991, we can also use machine learning techniques to derive these factors from the data and interpret them accordingly. The method use was a principal component analysis (PCA), and showed that only three factors are sufficient to describe most of the volatility (95%) of the bond returns. These results show that, from an empirical perspective, the yield curve can reasonably be decomposed in factors, what was only supposed by Nelson and Siegel, 1987. Moreover, the factors are now supposed uncorrelated together by construction, which avoid multicollinearity problems. Even if it was not the purpose of the paper, we see that the results leave a little place to macroeconomic variables in the variability of the yield curve. We will come back to it in the following section. Derived the factors directly from market data allow us to describe the curve in a more parsimonious manner (Novosyolov & Satchkov, 2008), which mean a model that capture most of the data variability with the fewest factors as possible. The factors derived can also be more meaningful. Taking the example from the paper, the level factor is most of the time described as a parallel shift, although in reality it is not perfectly parallel. Such little dynamic would be captured by the principal component and not by the DNS factor as constructed in the original paper.

The Principal Component Analysis (PCA) method offers several advantages, including simplicity in implementation and the generation of orthogonal (uncorrelated) factors. However, it is based on certain assumptions, such as linear relationships between yields of different maturities and the stationarity of the time series—assumptions that are rarely perfectly met in practice. Additionally, PCA is a data-driven approach that facilitates factor interpretation but is highly sensitive to outliers and the specific period used for analysis. Consequently, the factor loadings derived from a yield curve may vary significantly between different time periods.
While the Dynamic Nelson-Siegel (D)NS model is extensively employed by central banks across Europe (e.g., Belgium, Germany, Switzerland) (BIS, 2005), PCA remains prevalent in risk management, hedge funds, and trading institutions (e.g., Goldman Sachs, Moody's) (Oprea, 2022). The Nelson-Siegel model leverages economic theory to derive time series for each factor, making it particularly useful for policymakers. Conversely, PCA is often more effective for predicting long-term maturities, which is valuable for managing interest rate risk. Dodampe Gamage, 2021 demonstrated that the (D)NS model tends to perform better for short maturities, whereas the PCA method excels in predicting mid- to long-term maturities, as evidenced by comparisons of out-of-sample RMSE for both models.

We can also mention other methods, such as Independent Components Analysis (ICA) (Chen et al., 2019), or the Gaussian Process Regression (Gonzalvez et al., 2019). These other machine learning methods are more recent and less popular in the financial literature, but may present interesting features compared to traditional methods. "The forecasting performance of the SG-ICA is remarkably better than the traditional parametric DNS model in an era of quantitative easing with zero policy rates" (Chen et al., 2019).

Macroeconomic interactions

One of the first paper to propose a model combining macro-variables and yields is Estrella and Mishkin, 1997. They used a VAR model with yields of different maturities, output, and inflation. We see here one of the first attempt in the literature to model the dynamics between the yield curve and the macroeconomic world, but with evident limitations. The model only takes into account a few maturities, and the dynamics of the whole curve cannot be inferred due to the use of an unrestricted VAR (Ang & Piazzesi, 2003). These last authors proposed instead a no-arbitrage affine model that involve 2 factors representing proxies for output and inflation (each factor being the first principal component of a large set of variable representing the macro-variable), and 3 unobserved (latent) factors of the yield curve. They restricted the model to follow the Taylor rule and price the bonds according to a risk-neutral affine model, ensuring no-arbitrage in a general bond pricing kernel, following Duffie and Kan, 1996 (also called “Duffie-Kan” conditions in the literature). The originality of the paper is to incorporate directly the macro-factors in the pricing kernel. They found that including macro-factors improve the forecasting abilities of the model, compared to a “yield-only” model. Computing the Forecast Error Variance Decomposition (FEVD), the macro-factors account to 85% of the variations in bond yields.

However, we observe some contradictions in the findings. As previously mentioned by Litterman and Scheinkman, 1991, the first three principal components of the yield curve account for approximately 95% of its variations. In contrast, Ang and Piazzesi, 2003 found that macroeconomic factors jointly explain 85% of the variations in yields. According to Bikbov and Chernov, 2010, this discrepancy arises from differences in measurement approaches. The variations explained in macro-finance models are often computed from the Forecast Error Variance Decomposition (FEVD) of a VAR model, which is typically Cholesky identified. However, FEVD is sensitive to the ordering of variables and the identification scheme used. To address this, Bikbov and Chernov, 2010 propose using the coefficient of determination, R^2 , as a more robust measure. This approach calculates the proportion of variance in the yield factors that can be directly explained by macroeconomic factors.

Ang and Piazzesi, 2003 developed a model incorporating macroeconomic factors into bond pricing, assuming these factors were determined independently. To provide a more nuanced analysis, Diebold et al., 2006 used the Dynamic Nelson-Siegel (DNS) model to extract yield factors, isolating them from the yield curve itself, and incorporated these factors into a VAR model with macroeconomic variables. This method allows for direct interpretation of the financial factors relative to the yield curve separately from macroeconomic variables, and the examination of their relationships through Impulse Response Functions (IRFs).
Using U.S. data from 1972 to 2000, including yield curves, the Federal Funds rate, inflation, and capacity utilization, Diebold et al., 2006 noted that the results are influenced by the specific time period analysed. Shifts in monetary policy, financial crises, and other significant events can impact the behaviour of these variables, providing insights into how macroeconomic factors interact with financial factors and how these relationships evolve over time.

First, the authors found a close connection between the level factor and inflation, which is interpreted as the market's perception of long-term inflation expectations. A shock to inflation tends to permanently increase the level factor, while a shock to the level factor can lead to a permanent rise in both the policy rate and inflation, aligning with expectations of increased long-term inflation. However, Rudebusch and Wu, 2008 suggests that the level factor may also reflect the central bank's implicit midterm inflation target as perceived by the market.
Another mechanism discussed by Diebold et al., 2006 is the relationship between monetary policy and the level factor. A contractionary monetary policy, anticipated by the market due to an announced lower inflation target, is expected to decrease the level factor. Conversely, a contractionary policy in response to high inflation tends to increase the level factor. Kagraoka and Moussa, 2013 demonstrated that during the Zero Lower Bound (ZLB) period—when the policy rate was constrained near zero in response to the 2008 financial crisis—quantitative easing led to a reduction in the level factor in Japan. Additionally, the impact of monetary policy on the level factor can be influenced by the central bank's credibility and the effectiveness of its monetary transmission mechanisms. For instance, Yan and Guo, 2015 found that the minimal impact on the level factor in China was attributed to the weak

transmission channel of Chinese monetary policy.

Secondly, there appears to be a strong connection between the slope factor and macroeconomic variables, particularly with the policy rate. This relationship can be interpreted through both the accommodation of monetary policy to changes in yields and the bond market's anticipation of monetary policy responses to shocks in inflation or output. Dewachter and Lyrio, 2006 link the slope factor to the business cycle and the monetary policy's response to fluctuations in real economic activity. For instance, an increase in the policy rate in response to an inflation shock tends to raise the slope factor nearly one-to-one. This reflects a rise in short-term interest rates, with a more muted increase in long-term rates if inflation is expected to decrease in the future, leading to a flattening of the yield curve. A similar effect is observed with shocks to output.
The slope factor also provides insights into expected future growth. If market participants anticipate a recession, they are likely to invest in long-term bonds to safeguard their savings, which in turn increases the slope factor (Sowmya & Prasanna, 2018). While the level factor is a robust predictor of expected inflation, the slope factor serves as a strong indicator of expected economic growth. However, both factors are influenced by the behaviour of monetary authorities.

Finally, Diebold et al., 2006 did not find any significant reaction of the curvature factor to macroeconomic variables, nor any inverse relationship. The curvature factor, which reflects midterm maturities, manifests as a hump in the yield curve. According to Dewachter and Lyrio, 2006, the curvature remains largely unaffected in developed countries, serving as a reflection of the "current stance of monetary policy" and representing real interest rate movements that are uncorrelated with macroeconomic variables.
However, Mönch, 2012 challenged this view, finding that shocks to the curvature factor do significantly impact macroeconomic variables, notably decreasing output and inflation. This suggests that the curvature does interact with macroeconomic variables, albeit in a one-way direction. Furthermore, Moller, 2014 demonstrated that the curvature factor has greater predictive value for forecasting economic growth than the slope factor. However, this study employed a "non-structural predictive

regression" approach which, as noted by the author, does not provide an economic explanation for the observed relationships.

A key limitation of the parametric method, as noted by Diebold et al., 2006, is the absence of no-arbitrage constraints. To address this, Christensen et al., 2011 developed a no-arbitrage dynamic Nelson-Siegel model. However, subsequent research has questioned the added value of these restrictions for parametric models. Including no-arbitrage constraints does not necessarily enhance the forecasting performance of a VAR model, as demonstrated by Duffee, 2011, who found out that such constraints did not improve the predictive accuracy for either PCA (3-factors) or DNS methods. According to the author, the linear relationships imposed by VAR models are sufficient to minimize forecasting errors. Despite this, no-arbitrage restrictions remain crucial for purposes such as pricing and hedging. They are deemed irrelevant only in the context of forecasting using yield curve information. Coroneo et al., 2011 found that factors derived from models incorporating no-arbitrage constraints are statistically consistent with those obtained from an "arbitrage-free model using the Nelson–Siegel factors as exogenous factors", suggesting that no-arbitrage restrictions may not be necessary to achieve financially consistent factors with such methods.

3 Methodologies in Use

3.1 Nelson-Siegel Model

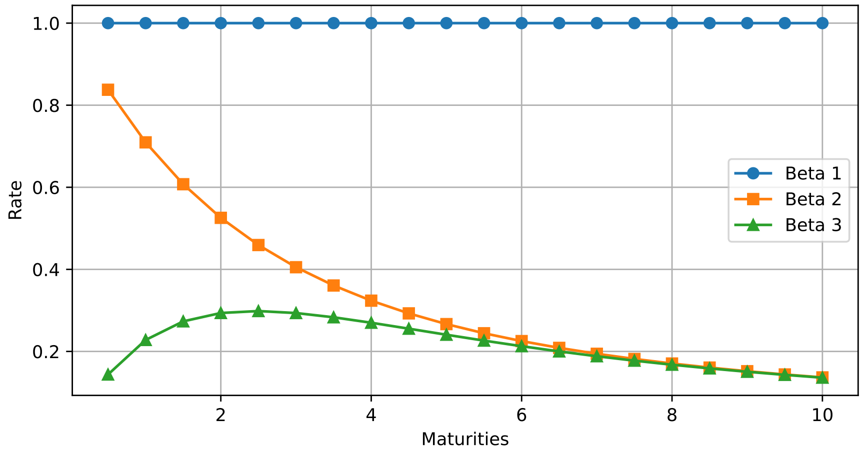
Nelson and Siegel, 1987 developed a method that decomposes the term structure of interest rates into 3 distinct factors using the following expression :

$$y_t(\tau) = \beta_{0,t} + \beta_{1,t}(\frac{1 - e^{-\lambda\tau}}{\lambda\tau}) + \beta_{2,t}(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}) \tag{3.1}$$

Where $y_t(\tau)$ represents the range of interest rates by maturity τ (in year), at time t, and with the decay rate λ. The beta coefficients are obtained through OLS regression of the yields on these factors. These coefficients are time series representations of the factors, enabling economic and statistical interpretations. The Loadings can be visually depicted in Figure 3.1, and the Beta's series with the data presented in the following section in Annex 6.1.

These factors represent, respectively, the level, slope, and curvature of the yield curve. In the static version of the model, the decay rate is invariant and fixed at $\lambda = 0.0609$, as in the original paper. This value was determined based on the maturity where the curvature factor is maximized. However, assuming an invariant decay rate is unrealistic and does not fit the data well. Diebold and Li, 2006 developed a dynamic version that selects the value of the decay rate at every point in time to maximize the likelihood of the function. This method is widely used in the literature for two main reasons. First, it has satisfactory out-of-sample performance and is widely accepted as an effective method to fit the yield curve. Second, the factor loadings are easily interpretable through their definitions in 3.1. The level factor reflects the overall path of the term structure, assigning equal weight to shocks across all maturities and indicating shifts in the general level of interest rates. In contrast, the slope factor assigns decreasing weights to longer maturities, reflecting

Figure 3.1: Nelson-Siegel model Loadings for $\lambda = 0.0609$



changes in the steepness of the yield curve and is often associated with market expectations of future interest rate movements. Notably, positive shocks to the slope factor typically lead to a flattening of the curve, while negative shocks result in a steepening. The curvature factor, on the other hand, focuses on midterm maturities and captures deviations from a purely linear yield curve, often manifesting as a hump or trough. The decay rate determines how quickly the impact of a shock diminishes across different maturities of the yield curve. A higher decay rate implies that shocks to interest rates have a more immediate effect on short-term maturities, with the impact diminishing rapidly for longer-term maturities. Conversely, a lower decay rate indicates a slower decline in the impact of shocks, leading to a more prolonged influence on longer-term interest rates.

However, the dynamic optimization method is far from being trivial. It requires the use of a Kalman filter to maximize the likelihood function and determine the optimal decay rate over time. Another drawback of the method is the high correlation between these factors, which lead to difficulties when estimating the parameters. The purpose of this master thesis will be to use an alternative method to model these factors, and to evaluate it usefulness in understanding the dynamics with the macroeconomic environment. We will evaluate it with respect to the Nelson-Siegel factors to interpret the new ones and to keep some comparison basis with the literature.

3.2 Principal Components Analysis

Principal Component Analysis (PCA) is an unsupervised machine learning technique used to reduce the dimensionality of a dataset by minimizing noise while retaining as much of the original variability as possible. This is achieved by computing principal components, which are linear combinations of the original variables designed to capture the maximum variance in the data.
Consider the dataset represented by the matrix $X \in \mathbb{R}^{(n,p)}$, where n is the number of observations (dates in the context of time series) and p is the number of variables (maturities). PCA typically involves centering and standardizing the variables to make them comparable. However, since our data are already in a consistent format (interest rates in percentages) and directly interpretable, we have chosen not to apply further transformations.

The unit vectors are defined as orthonormal:

$$\begin{aligned} U_\alpha \perp U_\beta \iff U'_\alpha U_\beta = 0, \quad \forall \alpha \neq \beta \in [1, 2, \ldots, p] \\ ||U_\alpha|| = ||U_\beta|| = 1, \quad \forall \alpha, \beta \in [1, 2, \ldots, p] \end{aligned} \tag{3.2}$$

By defining these vectors, we seek to find new axes to represent X, ensuring that the dimensionality does not exceed p. The line generated by U_1, denoted as $\psi_1 U_1$, is referred to as the first factorial axis. With this, we can reconstruct the original space as follows:

$$\{\psi_1 U_1 + \psi_2 U_2 + \cdots + \psi_p U_p \mid \psi_\alpha \in \mathbb{R},\, U_\alpha \in \mathbb{R}^n\}. \tag{3.3}$$

We project all rows $X_n (\in \mathbb{R}^P)$ on each factorial axis such that $\psi_n = X'_n \cdot U_\alpha$. We seek the subspace U that minimize the sum of squared errors of these projections. The goal here is to find the subspace that minimize the total projection errors of the data.
After few calculations, we find that the optimal unit vectors are the eigenvectors of the covariance matrix of X (denoted Σ), that we order in function of their eigenvalues. Formally, with any positive semi-definite covariance matrix :

$$\Sigma = X'X \iff X'XU_\alpha = \lambda_\alpha U_\alpha \tag{3.4}$$

Where U_α is the αth eigenvector, and λ_α its corresponding eigenvalue. We order the eigenvector in function of their eigenvalues such that $\lambda_{\alpha=1} \geq \lambda_{\alpha=2} \geq ... \geq \lambda_p$. The principal components are the original dataset projected on each of these eigenvectors : $\psi = XU$, where U is the space generated by a specific amount of eigenvectors. We see that if we choose a subsample of lower dimension than the dataset ($k < p$), we have a fewer amount of principal components than variables. The dimension reduction is completed.

These principal components are interesting because they are uncorrelated, since we assumed the eigenvectors to be orthogonal to each others (definition 3.2). Their lack of inter-dependence allow us to represent information in a more efficient way. Each ψ_α has a variance, which is equal to the associated λ_α. The diagonal of the variance-covariance matrix represents the variance of the component α such that $\psi'_\alpha \psi_\beta = \lambda_\alpha$ when $\alpha = \beta$. The first component has therefore the greatest variance, and the last as the least. The covariances are null, because $\psi'_\alpha \psi_\beta = 0$ for $\alpha \neq \beta$. The ordering of eigenvectors by their eigenvalues ensures that we retain the most significant dimensions, capturing the maximum variance in the data with the lowest dimensionality as possible. To select a parsimonious amount of components (k), we calculate the proportion of inertia explained by the reconstruction U $\in \mathbb{R}^k$ on the total inertia for k = p. This shows the percentage of variance explained by the addition of the following component in the model, and allow us to choose the most parsimonious amount of components.

Formally :

$$T_k = \frac{\sum_{\alpha=1}^{k}(\lambda_\alpha)}{\sum_{\alpha=1}^{p}(\lambda_\alpha)} * 100 \tag{3.5}$$

The Scores ($\psi_\alpha \in \mathbb{R}^{(n,1)}$) represent the projection of each X columns in the subspace generated by the component principal. They are the evolution through time of the corresponding factor. With these series, we have the overall pattern of the term structure for the corresponding factor. But we don't see how the different maturities

will be affected. To understand these dynamics, we have to look at the Loadings ($U_\alpha \in \mathbb{R}^{(1,p)}$), which is the eigenvector of Σ for the factor α.
This matrix serves as a fundamental tool for interpreting our factors. The first three components encapsulate the factors that account for the highest variance in the original data. These loadings may be interpreted with the same logic that the Nelson-Siegel factors of figure 3.1. This method provides time series of the factors (ψ_α), and the corresponding loadings (U_α) to interpret them. The difference with NS is that the loadings are constructed to give series that are uncorrelated together, instead to be restricted to a meaningful economic expression only. The goal is to capture the driving patterns of the yield curve directly from the data without any predetermined structure.

We are also able to reconstruct original data, but with only a few principal components. We define the reconstruction as :

$$\hat{X}_k = \psi_k * V_k' \tag{3.6}$$

The reconstruction will only depend on the amount of components that are reused. It allows us to compare with the original data and determine the usefulness and the accuracy of the factorial decomposition in preserving the underlying patterns of X.

3.3 Modeling process

3.3.1 ARMA process

The univariate Auto-Regressive Moving-Average (ARMA) process only consider past values of a time series in the determination of its current one. Formally, it can be described as :

$$ARMA(p,q) := y_t = \mu + \sum_{i=1}^{p}(\phi_i y_{t-i}) + \sum_{i=1}^{q}(\theta_i \varepsilon_{t-i}) + \varepsilon_t \tag{3.7}$$

Where N is the amount of observations in the process, p the amount of lag and q the amount of lagged past errors incorporated in the model. For it to be consistent, several assumptions have to be made on the process:

- $y \in \mathbb{R}^{(N,1)}$ has to be stationary, meaning that its mean (μ) and variance (σ_y^2) are finite and time-invariant.
- The stochastic components $\varepsilon \in \mathbb{R}^{(N,1)}$ are assumed to follow a white noise process, meaning they have an expected value of zero $E(\varepsilon_t) = 0$, a finite and constant variance $\text{Var}(\varepsilon_t) = \sigma_\varepsilon^2 < \infty$, and no autocorrelation over time $\text{Cov}(\varepsilon_t, \varepsilon_s) = 0 \, \forall \, t \neq s$.
- More generally, the model assumes the process to be linear in terms of the coefficients of past observations and errors.

Stationarity may statically be tested with a Dickey-Fuller test, which test the presence of an unit root in the process. An unit root is defined as the full incorporation of past values in the determination of the current one, showing no-mean reversion behaviour, which automatically leads to an undefined unconditional mean and variance. If the p-value is less than 5%, the serie is considered to be stationary.

$$\begin{aligned} y_t = \rho_1 y_{t-1} + \varepsilon_t \\ H_0 : \rho = 1 \\ H_1 : \rho < 1 \text{ and } \varepsilon_t \sim \mathcal{N}(0, \sigma^2) \end{aligned} \tag{3.8}$$

Normality can be tested with a Jarque-Bera test, which jointly test the Kurtosis and the Skewness of the empirical distribution against the normal one. If the p-value of the JB-stat against the Chi-squared distribution < 0.05, the empirical distribution is considered as non-normally distributed.

$$JB = \frac{n-k}{6} * (S^2 + \frac{(K-3)^2}{4}) \sim \chi^2 \tag{3.9}$$

Finally, autocorrelation will be tested using a Lljung-Box test.

$$Q_m = T(T+2) \sum_{s=1}^{m} \left(\frac{r_s^2}{T-j} \right) \sim \chi_m^2, \text{ where } r_s = \frac{\sum_{t=1}^{T} X_t X_{t-s}}{\sum_{t=1}^{T} X_t} \tag{3.10}$$

r is the sample autocorrelation, and the null hypothesis (H_0) is that data are independently distributed : $r_1 = r_2 = \ldots = r_s = 0$. The test is designed to incorporate a specific amount of m lags, and shows evidence for auto-correlation if the p-value < 0.05.

The amount of lags p and q will be chosen through a parsimonious comparison of the AIC, BIC and HQ tests. We keep the combination of p and q with the lowest scores for each criterion. The likelihood $\hat{\Sigma}$ is an empirical measure of how well the model fit the data. The higher the measure is, the better will be the goodness-of-fit. We add a penalty corresponding to the amount of parameters (p and q) and the number of observations (T), representing its complexity.

$$\begin{aligned} BIC(p,q) &= \ln(N) * (p+q) - 2 * \ln(\hat{\Sigma}) \\ AIC(p,q) &= 2 * (p+q) - \ln(\hat{\Sigma}) \\ HQ(p,q) &= 2 * (p+q)\ln(\ln(N)) - 2 * \ln(\hat{\Sigma}) \end{aligned} \tag{3.11}$$

3.3.2 VAR process

A Vector Auto-Regressive (VAR) model analyse the joint behaviour of multiple variables over time. It is a multivariate extension of the univariate autoregressive (AR) model, allowing for the simultaneous modelling of the dynamic relationships between several variables.

$$VAR(T) \coloneqq Y_t = C + \sum_{i=1}^{T} (A_i Y_{t-i}) + v_t \tag{3.12}$$

With $Y_i \in \mathbb{R}^{(1,P)}$, N the amount of observations, P the amount of endogenous variables to fit the model and T the number of lag taken into account. $A_i \in \mathbb{R}^{(P,P)}$ is the coefficient matrix capturing the lagged effects of the variables in t-i, and $v_t \in \mathbb{R}^{(1,P)}$ represents the error vector at time t. The underlying assumptions of such model are similar than previously :

- The VAR model assumes a linear relationship among the variables and their lagged values. If the true underlying relationship is nonlinear, the model may not capture the dynamics accurately.

- The error vector follows a white noise behaviour : $E(v_t) = 0$, $E(v_t v_t') = \Sigma$ (any symmetric and positive semi-definite matrix), and $E(v_{t-i} v_{t-j}') = 0, \forall i \neq j$
- The system in itself is stationary, requiring the variables to be separately and jointly stationary.

We can rewrite the reduced form VAR(T) into a simple VAR(1) by using the companion form :

$$\begin{pmatrix} Y_t \\ Y_{t-1} \\ Y_{t-2} \\ \vdots \\ Y_{t-P+1} \end{pmatrix} = \begin{pmatrix} c \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} A_1 & A_2 & \cdots & A_p & 0 \\ I_N & 0 & \cdots & 0 & 0 \\ 0 & I_N & & \vdots & \vdots \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I_N & 0 \end{pmatrix} \begin{pmatrix} Y_{t-1} \\ Y_{t-2} \\ \vdots \\ Y_{t-P} \end{pmatrix} + \begin{pmatrix} v_t \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{3.13}$$

Which can be rewritten : $\Gamma_t = C + A\Gamma_{t-1} + u_t$, and (if the VAR is stationary, which requires all the eigenvalues of Γ to be between in the unit circle [-1,1]) in the VMA(∞) form (with L the lag operator) :

$$\begin{aligned} \Gamma_t = (I - A_i L - A_2 L^2 - ... - A_T L^T)^{-1} C + \sum_{i=0}^{\infty} (A^i u_{t-i}) \\ (I - A_i L - A_2 L^2 - ... - A_T L^T)\Gamma_t = C + v_t \\ \Gamma_t = \mu + \Psi(L) v_t \end{aligned} \tag{3.14}$$

Once the VAR model is estimated, we are able to reconstruct the original series separately using our results.

$$\hat{Y}_t = A_1 Y_{t-1} + A_2 Y_{t-2} + \ldots + A_T Y_{t-T} \tag{3.15}$$

To assess the performance of the VAR model, we can use the Root Mean Squared Error (RMSE). It measures the average magnitude of the errors between predicted values and observed values. We can use it both in-sample and out-of-sample, which means reducing the amount of observation in the initial model and trying to predict them.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Y_i - \hat{Y}_i)^2} \tag{3.16}$$

3.3.3 Cholesky identification

The main drawback of using an unrestricted VAR model is that it does not allow us to distinguish the causal effect of one variable on another. It assumes well that the errors v_t are uncorrelated over time, but it does not impose any restrictions on their correlation at a specific point in time. To address this issue, we need to impose restrictions on the model to identify the causal impacts. Specifically, by assuming that the variance-covariance matrix of the residuals is positive semi-definite, we can always decompose it using the Cholesky theorem, such that $\Sigma_v = BB'$, where B is a lower triangular matrix.

$$B = \begin{pmatrix} b_{1,1} & 0 & 0 & ... & 0 \\ b_{2,1} & b_{2,2} & 0 & ... & 0 \\ \vdots & & & & \vdots \\ b_{n-1,1} & b_{n-1,2} & \dots & b_{n-1,n-1} & 0 \\ b_{n,1} & b_{n,2} & \dots & b_{n,n-1} & b_{n,n} \end{pmatrix} \tag{3.17}$$

With this decomposition, we can express v_t as a linear combination of orthogonal shocks, $v_t = B\varepsilon_t$, where ε_t is distributed as $\varepsilon_t \sim WN(0, I)$. These structural shocks are orthogonal because their variance-covariance matrix is the identity matrix, meaning the covariances are zero. However, this method imposes certain restrictions on the structural relationships because B is a lower triangular matrix. The first row of the matrix represents the contemporaneous effects of the structural shocks on the first endogenous variable, the second row represents the effects on the second variable, and so forth. Importantly, the first variable does not react contemporaneously to the other variables, responding only to the first shock, while the last variable responds to all shocks. Consequently, the ordering of the endogenous variables in Y_t is crucial, as they need to be placed in order of exogeneity.

With this identification scheme, the Structural VAR model is now expressed as :

$$SVAR(T) \coloneqq Y_t = C + \sum_{i=1}^{T}(A_i Y_{t-i}) + B\varepsilon_t \tag{3.18}$$

3.3.4 Dynamic analysis

We can analyse the dynamic in work inside a (S)VAR model with two main methods, Impulse Response Functions (IRF) and Forecast Error Variance Decomposition (FEVD).

The IRF captures the dynamic response of the system Y_t to a one-time shock in another variable $\varepsilon_{j,t}$. Taking the VMA(∞) of the structural form : $\Gamma_t = \mu + \Psi(L)B\varepsilon_t$, which van be developed to :

$$\Gamma_t = \mu + \Psi_1 B\varepsilon_{t-1} + \Psi_2 B\varepsilon_{t-2} + ... \tag{3.19}$$

The IRF is defined as $\frac{\delta Y_{t+i}}{\delta \varepsilon_{j,t}} = \Psi_i B\varepsilon_t$, where we define a synthetic shock vector $\varepsilon_t = (0, d, 0, ..., 0)'$ with a shock d at the impulse variable (in this case j = 2), 0 otherwise. With such method, we are able to observe the impact of a specific shock on the entire system and observe how the effects propagate over time.

The FEVD shows what portion of the variance of the VAR's forecast errors (at a given horizon h) is due to each structural shock. We compute it as the cumulative effect of structural shocks up to the horizon t+h.

$$FE_{t+h}(\varepsilon_j) = Y_{t+h} - E_{t-1}(Y_{t+h}) = \sum_{i=0}^{h} A^{h-i}B\varepsilon_{j,t+i} \tag{3.20}$$

To find the proportion attributed to a specific shock, we calculate the ratio of the variance of the forecast error attributed to ε_j to the total variance of the forecast error at horizon t+h:

$$FEVD_{j,t+h} = \frac{\text{Var}(FE_{t+h}(\varepsilon_j))}{\sum_{j=1}^{P}\text{Var}(FE_{t+h}(\varepsilon_j))} \tag{3.21}$$

4 Dataset

4.1 Yield Curve

4.1.1 Descriptive Statistics of Yield-Curve Level

The yield curve data utilized in this study were sourced from the Bank of England's official website[1]. The dataset comprises the values of nominal spot rates for all semi-annual maturities up to ten years. Table 4.1 presents descriptive statistics for key maturities, while Figure 4.1 illustrates the historical rates. The dataset spans the period from January 4, 1993, to December 29, 2023. This timeframe was deliberately chosen to concentrate on the era of inflation-targeting monetary policy adopted by the Bank of England, which began in September 1992 (King, 2005) and remains in effect. This selection also serves to exclude the transitional period, which is often characterized by substantial data gaps.

Table 4.1: Summary Statistics of Yields of some maturities

(%)	mean	std	min	max	skewness	kurtosis
0.5	3.355	2.406	-0.157	7.411	-0.198	-1.566
1	3.397	2.407	-0.150	7.297	-0.191	-1.551
2.5	3.611	2.395	-0.158	8.303	-0.129	-1.360
5	3.900	2.295	-0.134	8.952	-0.035	-0.997
7.5	4.099	2.187	-0.035	9.061	0.040	-0.689
10	4.240	2.094	0.111	8.991	0.106	-0.443

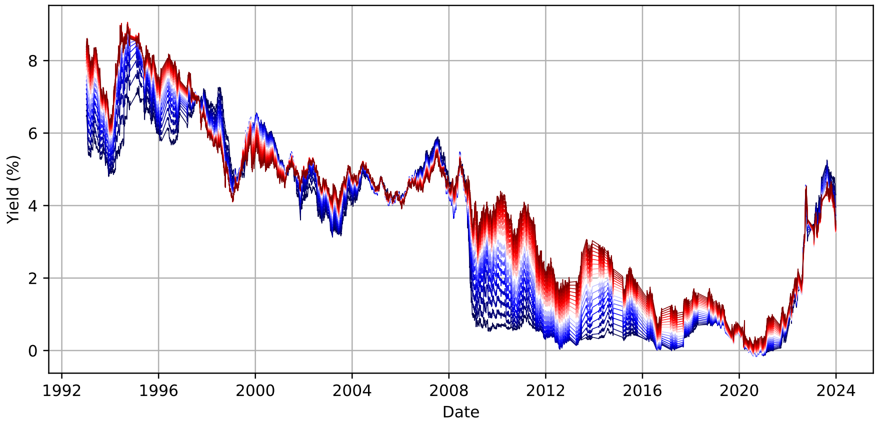
The data indicate that the mean yield increases with maturity, while volatility decreases. This pattern suggests that investors demand higher yields for long-term maturities compared to short-term ones, a phenomenon typically attributed to a combination of liquidity and inflation premiums. This supports the stylized

[1] https://www.bankofengland.co.uk/statistics/yield-curves

fact that the "normal" shape of the yield curve is upward sloping. The observed decline in volatility with longer maturities implies that long-term interest rates are less susceptible to external shocks than short-term rates. These observations are illustrated in Figure 4.2, which displays the average yield curve along with its upper and lower bounds (representing one standard deviation). The curve begins at 3.35% and rises monotonically to 4.24%, with the boundaries clearly showing the tendency for rate volatility to decrease as maturity lengthens.

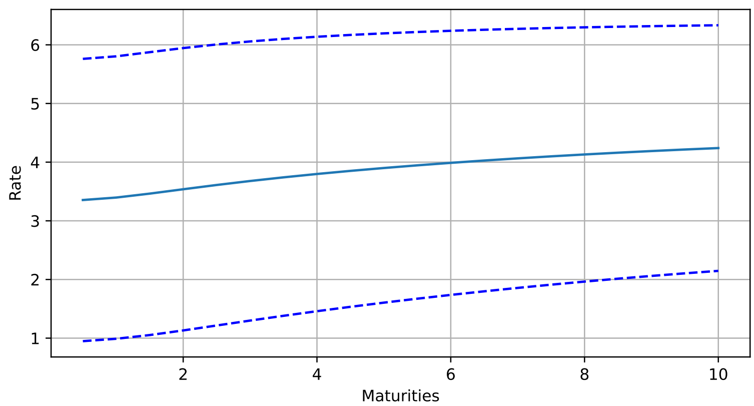
The kurtosis values for the yield distributions are consistently negative and increase with maturity, suggesting that longer-term yields have distributions with lower peaks and narrower tails. Typically, a kurtosis value closer to 3 would suggest a more Gaussian distribution, characterized by a sharper peak and thicker tails. However, rather than observing these characteristics, the distributions (Figure 4.3) exhibit a bimodal structure, reflecting greater variability in concentration across different yield rates. Specifically, shorter maturities tend to cluster around the first peak, longer maturities around the second peak, while midterm maturities display a more balanced distribution between the two peaks. This bimodal pattern challenges the expectation of flatter distributions for longer maturities. Although volatility decreases as maturity increases, the negative kurtosis values reveal that longer-term yields still have wide, flat distributions. Typically, a higher kurtosis value indicates a more concentrated distribution with fewer extreme outcomes. However, negative kurtosis means that these distributions are broader, with thinner peaks and wider tails. This implies that even though longer-term yields experience less volatility overall, the chances of extreme outcomes (large fluctuations in rates) do not decrease as much as we might expect. In other words, while yields become less volatile, they still maintain a considerable level of unpredictability.
Similarly, the skewness values show a negative trend for maturities up to 6.5 years, turning positive for longer maturities. Given the twin-peak structure of the distributions, the skewness can be interpreted as reflecting the degree to which yields are more likely to cluster around one peak over the other, indicating the relative weight of these distributions at different maturities. It is however very challenging to give a rigorous interpretation of these two components due to the twin peak distributions.

Figure 4.1: Historical series of UK nominal spot rates by maturity (1993-2024)



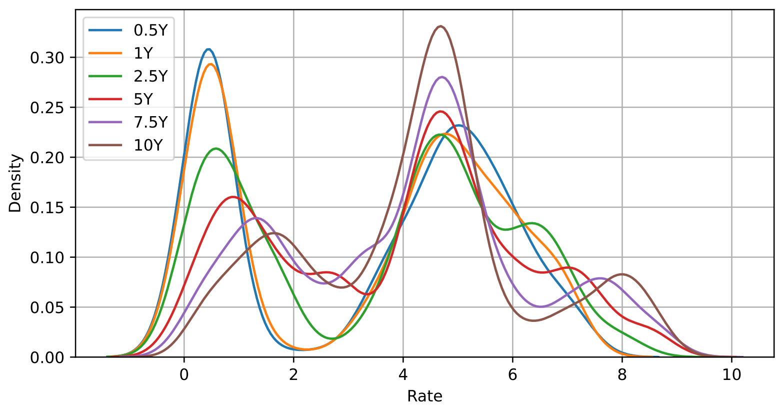
Note : Maturities go from 0.5Y (dark blue) to 10Y (dark red)

Figure 4.2: Average Yield curve



Note: plain line represents the mean rate, stripped ones represent ± 1 standard error

Figure 4.3: Kernel distributions of Yields



This presents an opportune moment to show the sensitivity of short-term interest rates to monetary policy compared to longer-term rates. By examining the period from 1993 to the present (Figure 4.1), we can discern three distinct phases: pre-Zero Lower Bound (ZLB), ZLB, and post-ZLB. Prior to the ZLB era, the UK policy rate was subject to frequent adjustments in response to prevailing macroeconomic conditions and business cycles. Monetary policymakers regularly altered interest rates to manage inflation, stimulate economic growth, or address other economic challenges. During the ZLB period, following the global financial crisis, the UK policy rate remained at historically low levels, typically around or near 0.5%, for an extended duration. This was part of the monetary policy response aimed at mitigating the economic downturn and supporting recovery efforts in the aftermath of the crisis. After the ZLB phase, as economic conditions improved, the UK policy rate gradually began to normalize, albeit at a gradual pace. This period witnessed a shift in the monetary policy stance, with interest rates adjusted upwards to reflect improvements in economic fundamentals and to guard against potential inflationary pressures.

This sequence of monetary policy actions underscores the pronounced reactivity of short-term interest rates to changes in the policy rate. Short-term rates closely track movements in the policy rate, essentially mirroring its adjustments. In contrast, longer-term interest rates exhibit a more subdued response, with the "low-rate" peak observed in short-term rates appearing to shift to higher levels as maturity lengthens. Additionally, longer-term rates appear to experience a decrease in density, reflecting a more gradual adjustment process influenced by a broader array of economic factors beyond just monetary policy.

4.1.2 Descriptive Statistics of Yield-Curve Returns

We look now to the daily changes of the interest rates by maturities. The same statistics as previously can be seen in Table 4.2.

Differentiating the yield data transforms each value into the change from the previous day's yield. This operation produces several important changes in the statistical properties of the dataset. First, the mean of the differentiated yields across all maturities becomes zero, indicating that, on average, there is no directional trend in daily yield changes. This zero mean reflects a balanced mix of positive and negative changes, with no prevailing upward or downward movement. Additionally, the standard errors across maturities become more uniform, signalling a reduction in the variability of daily changes. This suggests that yield fluctuations are more stable and consistent across the maturity spectrum. However, despite this uniformity in standard errors, the pattern of increasing volatility with longer maturities, observed earlier, remains true.

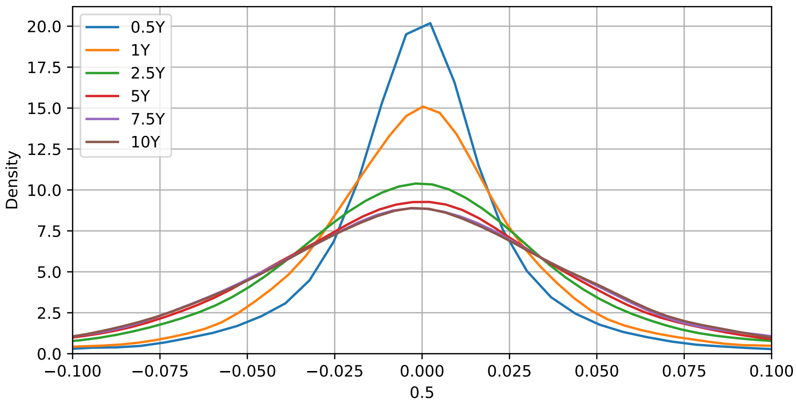
Table 4.2: Summary Statistics of Differentiated Yields

	mean	std	min	max	skewness	kurtosis
0.5	-0.000	0.039	-0.586	0.749	1.241	47.248
1	-0.000	0.043	-0.453	0.458	0.455	14.223
2.5	-0.001	0.053	-0.346	0.548	0.516	7.967
5	-0.001	0.056	-0.385	0.497	0.329	6.676
7.5	-0.001	0.057	-0.427	0.473	0.128	6.095
10	-0.001	0.057	-0.443	0.471	0.001	5.958

The kernel density estimates in Figures 4.4 demonstrate that differentiation leads to a more uniform distribution across maturities, resulting in a single peak for each distribution. This uniformity is particularly beneficial for Principal Components Analysis (PCA), as it requires the variables to be interpretable on a similar scale and is sensitive to outliers. Currently, both skewness and kurtosis exhibit distinct trends. The skewness decreases monotonically as maturity increases, indicating that longer-term yields tend to have more symmetric distributions. For shorter maturities, skewness is positive, suggesting a higher likelihood of positive yield changes, meaning investors might expect short-term gains more frequently. In contrast, longer-term maturities have distributions centred more evenly around the mean, reflecting a more balanced expectation of gains and losses.

Regarding the kurtosis, we observe positive values across all maturities, with a sharp decline as maturity increases. The shortest maturities, such as 0.5Y, show extremely high kurtosis (47.24), indicating distributions with pronounced heavy tails, and meaning a greater likelihood of extreme yield changes. However, as the

Figure 4.4: Kernel distributions of differentiated Yields



maturity lengthens, the kurtosis decreases rapidly, signalling that the distributions become less extreme and closer to normality over time. This reduction in kurtosis suggests that longer-term yields are less prone to extreme values, contributing to a more stable yield distribution.

4.1.3 Factorial Analysis

We now proceed with the principal component analysis (PCA) of the monthly yield returns, as outlined in Section 3.2. The analysis was performed on the returns rather than the raw yields to address the issue of non-stationarity, a crucial property required for subsequent sections, as well as for the reasons discussed in the previous section. Corresponding figures for the raw yields, including the explained variance (Figure 6.2), scores (Figure 6.3), and Loadings (Figure 6.4), are available in the Annexes. A stationarity test using the Dickey-Fuller test (Equation 3.8) is reported in Annexe 6.1, revealing that at the 0.05 significance level, only the third component is stationary. To ensure that the results of the subsequent tests and methods are interpretable in the context of the models computed in this dissertation, we use monthly returns. Further methodological details are provided in Section 4.2.
When calculating the explained variance for each model (Equation 3.5), we found that the first three principal components account for 99.52% of the total variability (Figure 4.5). Specifically, the first component captures 89.97%, the second 8.21%, and the third 1.56%. The explained variances for the fourth component and beyond

Figure 4.5: Components Selection - yield returns

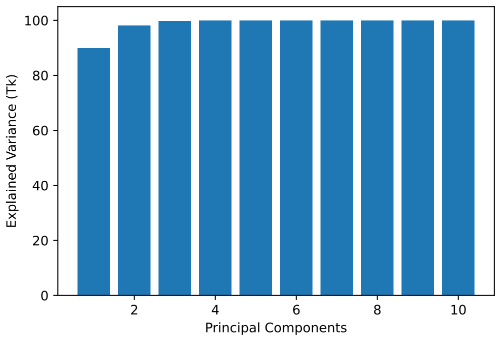
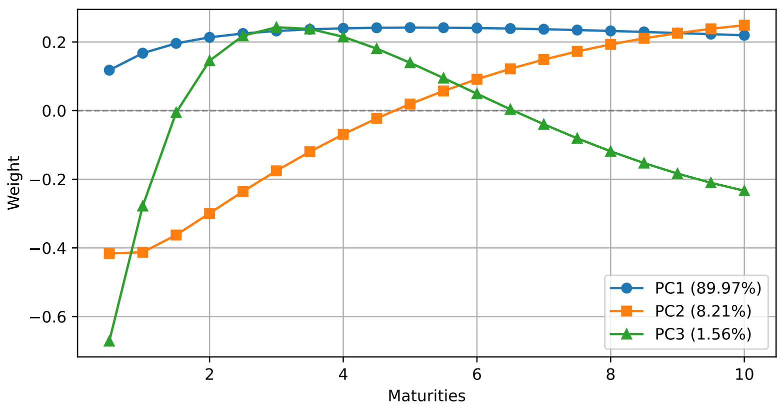
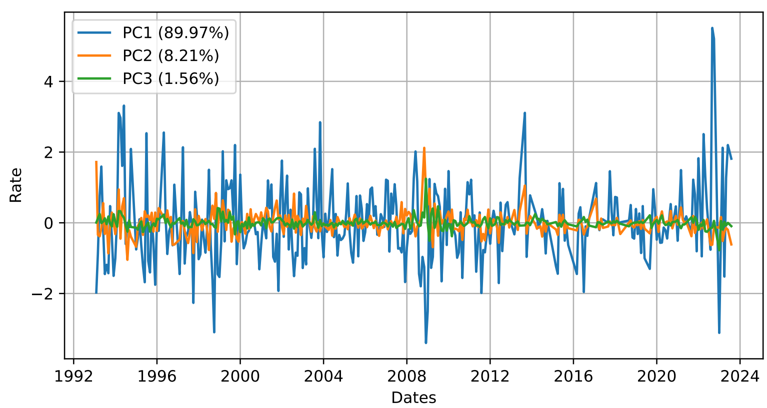


Figure 4.6: Loadings - Yields returns



Note : Values in parenthesis represent the explained variance of the factor. The PC1 represents the loadings of the level, PC2 the slope and PC3 the curvature

Figure 4.7: Scores - Yields returns



Note : Values in parenthesis represent the explained variance of the factor. The PC1 represents the scores of the level, PC2 the slope and PC3 the curvature are negligible (0.23% for the fourth, 0.03% for the fifth, and so on). We suppose then that accounting of only the three first factors is sufficient for our model.

The loadings for the first three principal components are illustrated in Figure 4.6. We observe that these components exhibit similarities to the factors identified in the Nelson-Siegel model, as shown in Figure 3.1. A shock in the level factor (PC1) causes a positive shift in yields across all maturities, though the shift is not uniform; the impact varies across different maturities, resulting in a non-parallel movement with a pronounced effect around the 4-year maturity mark. Conversely, a shock in the slope factor (PC2) leads to a decrease in rates for maturities up to 3 years, while rates for longer maturities rise, effectively sharpening the yield curve. However, it's important to note that this behaviour is the inverse of the traditional slope factor defined in the Nelson-Siegel model. The curvature factor (PC3) behaves similarly to its traditional counterpart, with a shock causing an increase in midterm rates (1.5 to 6.5 years), peaking around the 3-year maturity. These findings align with previous empirical studies (Bolder et al., 2004; Sekyere Asare, 2019; Wellmann & Trück, 2018), reinforcing the robustness of the observed factor structures.

We are able to interpret these factors the same way we interpret traditional factors. However, an important point to highlight is that the slope has to be interpreted as an opposite shock. The other part of the factorial analysis are the scores, that are reported in Figure 4.7.
The principal components analysis is supposed to return an uncorrelated factor, which is the case, as we can see in Annexe 6.2. It shows an absence of linear relationship between them and allow us to really interpret the factor separately, and avoid the multicollinearity problem highlighted by Diebold and Li, 2006. If we compute these correlations for the previously presented static NS factors, we obtain much higher results (Annexe 6.3).

Understanding the dynamics of differentiated series through visual inspection alone can indeed be challenging and incomplete. To gain deeper insights, we adopt a more rigorous approach by analysing the correlations with empirical reconstructions. This methodology enables us to directly assess whether the interpretation of the scores aligns with that of the loadings. The formula of the following empirical

Figure 4.8: PC1 against empirical counterparts

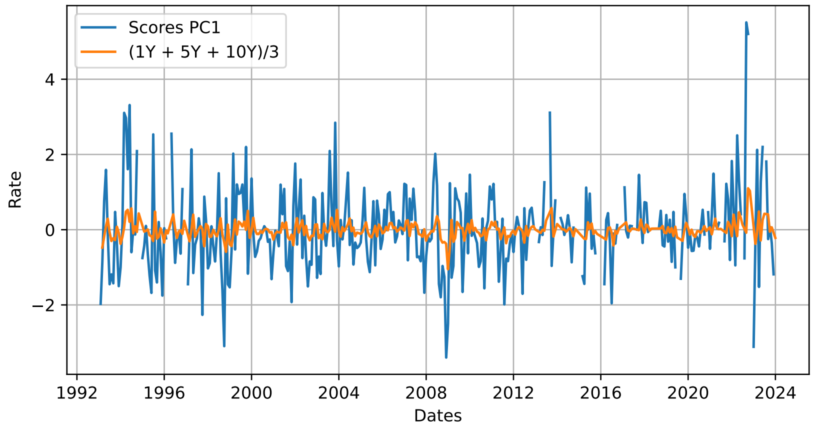


Figure 4.9: PC2 against empirical counterparts

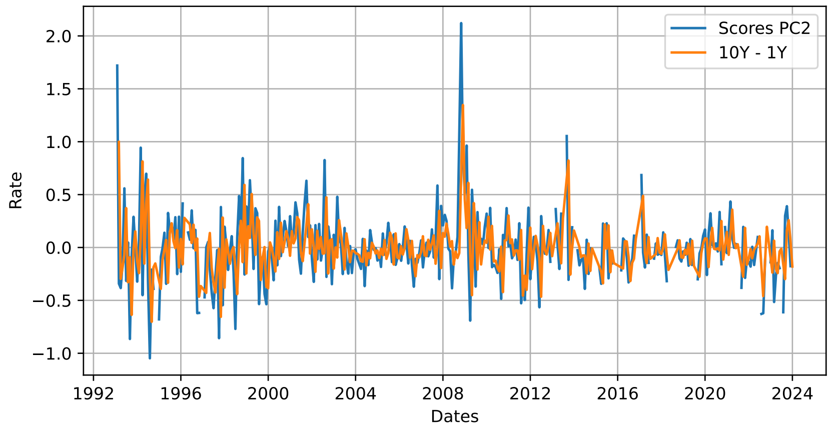
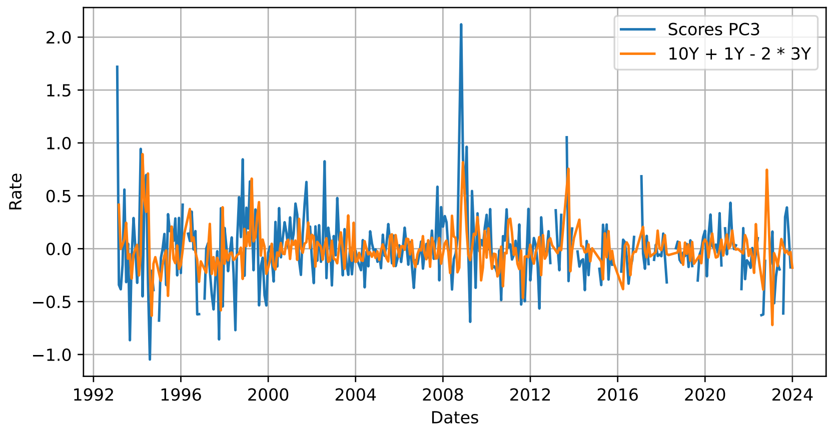


Figure 4.10: PC3 against empirical counterparts



counterparts of the three factors are directly taken from Diebold et al., 2006.
For instance, the first factor is supposed to follow long-term maturities. We evaluate its effectiveness by computing the correlation with the average between short (1Y), midterm (5Y) and the longest maturities in our dataset, 10Y. The resulting correlation of 98.14%, as depicted in Figure 4.8, indicates a remarkably strong relationship. This suggests that the level factor serves as an excellent representation of the yield curve's level dynamics (in the sense of the counterpart formula). Moving on to the second factor, which represents the slope, we compare it with the spread between long-term yields (10Y) and short-term yields (1Y). Figure 4.9 showcases this comparison, revealing a high correlation of 96.43%. This indicates that the curvature factor effectively captures the curvature dynamics of the yield curve. Finally, we examine the curvature component, which is reproduced using the spread of the half of the sum between long-term yields (10Y) and short-term yields (1Y) minus midterm yields (3Y). Figure 4.10 displays these comparisons, revealing a correlation of 79.27%. We see that the relationship of this last factor is not as straightforward as the two precedents, even if the value of the correlation is still high enough to be significant, and to provide a correct proxy for our third factor.

4.2 Macroeconomic variables

The variables that we include in the Macro-Finance model are supposed to be the bare minimum to capture basic macroeconomic dynamics : Real activity, inflation, and the policy rate.

We use the UK Index of Industrial Production (excluding construction) as a measure of real economic activity, as shown in Figure 4.11. This dataset covers the period from January 1993 to November 2023 and is available at a monthly frequency[2]. This index is essential for understanding production trends, assessing economic health, and identifying business cycles by signalling periods of expansion or contraction in industrial activity. It is particularly relevant for analysing the slope factor in our model, which is closely tied to these business cycles. Additionally, this metric offers valuable insights into inflationary pressures and serves as a leading indicator for

[2]https://fred.stlouisfed.org/series/GBRPROINDMISMEI

monetary policy decisions.Inflation is measured using the Harmonized Consumer Price Index (CPI)[3], which tracks the price changes of a basket of consumer goods and services. The data represents the year-over-year growth rate for each month and spans from January 1993 to November 2023, as depicted in Figure 4.12. Lastly, the policy rate data is sourced directly from the Bank of England's website[4]. The policy rate, commonly referred to as the "Bank Rate," is the central bank's main policy instrument. The data, reported monthly, covers the period from January 1993 to August 2023, as illustrated in Figure 4.13.

From a Dickey-Fuller test (Equation 3.8), we see that only the policy rate cannot be considered as stationary at a 0.05 threshold (Annexe 6.4). We decide therefore to take the first difference of all of these variables. The motivation comes first to ensure stationary for all variables, but also to make all the variables comparable with the yield factors that are already differentiated. Taking all variables as variations only is important to ensure more accurate comparison and interpretation of their relationships within the model. This step ensures consistency across the dataset and enhances the robustness of our analysis, ultimately leading to more reliable insights into the dynamics of the macroeconomic environment and its impact on yield movements.

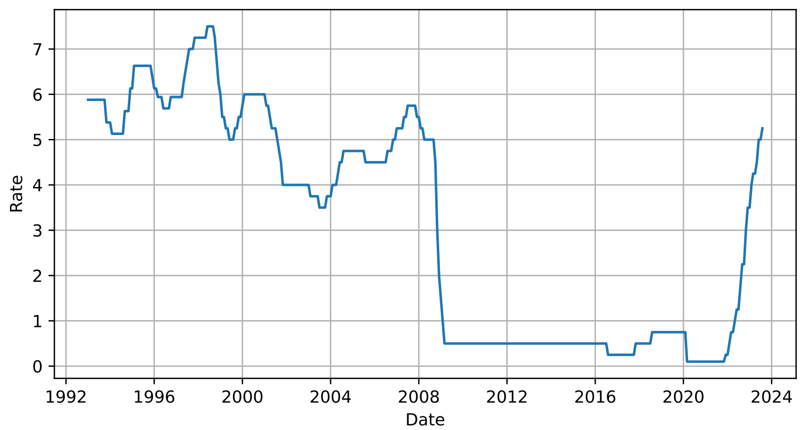
To ensure consistency with the monthly frequency of our data, we opt to retain only the first day of each month from the yield data before proceeding with the PCA analysis. This methodology, advocated in Diebold et al., 2006, allows us to establish precedence for the factors over the macroeconomic variables. This sequencing is important for constructing the VAR model identification scheme, ensuring that the factors precede the variables within the system. Importantly, this prepossessing step has no impact on the interpretability of the factors; both scores and loadings dynamics remain unaltered. When comparing with previous correlation of the term structure factors, we see that their values are much higher (Annexe 6.2). However, they remain largely sufficiently low to not contradict our previous remarks.

[3] https://fred.stlouisfed.org/series/CPHPTT01GBM659N
[4] https://www.bankofengland.co.uk/monetary-policy/the-interest-rate-bank-rate

Figure 4.11: UK Economic Activity: Industry (Except Construction) (1993-2023)

Figure 4.12: UK Harmonized CPI (YoY) (1993-2023)

Figure 4.13: Bank rate (1993-2023)



5 Results

5.1 Yield-only ARMA

The first model that we use is an ARMA, separately on each principal component. The idea comes from Diebold and Li, 2006, which shows that separate AR(1) models of each factor are more effective (from a forecasting performance point of view) than a VAR(1). However, because they didn't discuss the choice of the AR(1) further than choosing the simplest model and because we don't use the same factorial decomposition, we decide to remake the analysis with ARMA models. The following Table 5.1 displays the results of an ARMA (Equation 3.7) on each component. The lags p and q are chosen through the AIC, BIC and HQ test (Equation 3.11). The results of the best lag order are reported in Annexe 6.5. The models' final selection choices are made with the objective to be as more parsimonious as possible, with the most significant variables as possible.

In our analysis, the constant μ of the model is expected to be null, as the process is assumed to be stationary. This expectation is met, as the estimated values are very close to zero, with a p-value near 1 indicating no significant deviation from zero. The unconditional variance σ^2 decreases across the factors, a pattern anticipated by the principal components' construction. Importantly, the factors do not adhere to a uniform process. Specifically, the first and third factors are modelled as following an ARMA(2,3) process, while the second factor adheres to an ARMA(1,1) model.

For the level factor (PC1) and the curvature factor (PC3), we observe significant dependencies on their past values extending up to two lags and on past errors extending up to three lags. Notably, for the level factor, the first past value lacks significance in both the autoregressive and moving-average components. These findings suggest that these factors exhibit notable persistence, with a delayed effect for the level factor. Specifically, changes in the yield curve's level require approximately

Table 5.1: ARMA process

Parameter	PC1	PC2	PC3
μ	-0.0036 (0.079)	0.0019 (0.029)	0.0004 (0.012)
ϕ_1	0.0057 (0.114)	0.6841^{***} (0.127)	1.5879^{***} (0.066)
ϕ_2	-0.7350^{***} (0.120)	-	-0.8046^{***} (0.068)
θ_1	0.1738 (0.125)	-0.5383^{***} (0.146)	-1.4992^{***} (0.079)
θ_2	0.6738^{***} (0.127)	-	0.6154^{***} (0.108))
θ_3	0.2802^{***} (0.048)	-	0.1883^{***} (0.049
σ^2	1.1425 (0.065)	0.1090 (0.005)	0.0204 (0.001)

For significance levels: *** indicates $p < 0.01$, ** indicates $p < 0.05$, * indicates $p < 0.1$. Standard errors of the parameters are shown in parentheses.

one month to fully manifest, aligning with the notion that market participants adjust their expectations gradually in response to economic indicators and policy changes. This indicates a stable nature of the level factor, reflecting gradual shifts in market expectations. The persistence in the curvature factor (PC3) reveals that changes in the yield curve's shape also reflect longer-term adjustments in market perceptions, with a delayed impact extending up to three months. This implies that shifts in curvature are not immediate reactions to new information, but rather reflect ongoing adjustments in market sentiment and economic conditions. In contrast, the slope factor (PC2) displays short-term dynamics, showing dependence only up to one lag. This factor is the least persistent, being influenced primarily by its immediate past values and past errors. This suggests that the slope factor reacts more quickly to short-term fluctuations, rather than showing the longer-term persistence observed in the other factors.

While these results provide valuable insights into the persistence and dynamics of the factors, they do not offer a comprehensive analysis of the yield curve's persistence. Instead, they illustrate the distinct behaviours of the factors used in subsequent analyses. Each factor incorporates both autoregressive and moving-average components, indicating that market participants often base their expectations on historical trends (AR component) and are also influenced by past disturbances or shocks (MA component). However, the interpretation of these results relies on several assumptions about the time series and residuals. The ARMA model assumes normally distributed residuals without autocorrelation. The Jarque-Bera test (Equation 3.9) results indicate that none of the models' residuals are normally distributed, with p-values very close to zero for all factors. Additionally, the presence of autocorrelation (Lljung-Box test, Equation 3.10) up to lag 10 is not conclusively established for all factors (Annexe 6.6). Consequently, the violation of the normality assumption introduces potential biases in the estimators, complicates inference, and may degrade forecasting performance. The kernel distributions of the residuals, shown in Annexe 6.5, visually illustrate these deviations from normality.

Understanding the persistence and responsiveness of the factors offers insights into how long-term fundamentals and short-term fluctuations, such as unexpected shocks, impact the yield curve. Nevertheless, these separate ARMA models do not capture the inter-dynamics between factors. To explore the relationships between factors and their underlying fundamentals, a Vector Autoregressive (VAR) model is required. The variance in ARMA processes observed across factors highlights the potential utility of VAR models in providing deeper insights into the dynamics of each factor within the term structure.

5.2 Yield-only VAR

We now proceed to jointly model the factors using a VAR(1) model, as specified in Equation 3.13. The lag length is determined through a BIC test, and the resulting coefficients are detailed in Table 5.2. In this table, rows represent the lagged values of the variables, while columns correspond to their current values. For instance, the coefficient -0.485 signifies a negative impact of the lagged slope factor on the current level factor. It is noteworthy that all relationships are statistically significant, except for the effect of the lagged curvature factor on the current level factor.
To explore causality within this framework, we apply a Cholesky decomposition (Equation 3.17) to the variance-covariance matrix, which is transformed into a lower triangular matrix. This decomposition allows us to identify causal relationships, as shown in Table 5.3. For example, the element 0.026 in the second row of the B matrix indicates that a shock to PC2 results in a small positive effect on PC1, with PC1 increasing by 0.026 units for each unit shock in PC2, given that the standard deviation of PC2 is 0.312. This suggests a modest contemporaneous influence, where movements in PC2 may lead to subsequent changes in PC1. Conversely, the

Table 5.2: Coefficient Matrix A for the Yield-Only

	const	**L1.PC1**	**L1.PC2**	**L1.PC3**
PC1	0.011 (0.061)	0.150^{***} (0.054)	-0.485^{***} (0.180)	-0.082 (0.413)
PC2	-0.008 (0.017)	-0.050^{***} (0.016)	0.156^{***} (0.052)	0.256^{**} (0.118)
PC3	0.001 (0.008)	-0.007 (0.007)	0.148^{***} (0.023)	0.093^{*} (0.052)

For significance levels: *** indicates $p < 0.01$, ** indicates $p < 0.05$, * indicates $p < 0.1$. Standard errors of the parameters are shown in parentheses. L1 represents the lagged variable.

Table 5.3: Lower triangular B matrix for the Yield-Only

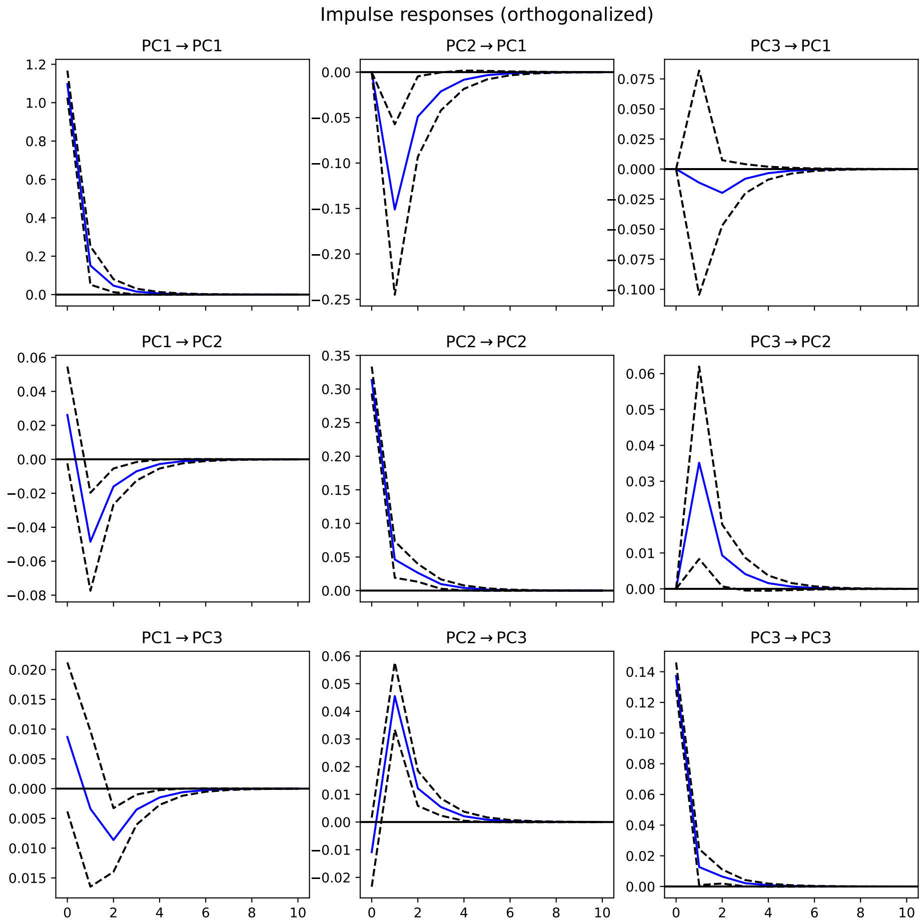
	PC1	**PC2**	**PC3**
PC1	1.092	0.000	0.000
PC2	0.026	0.312	0.000
PC3	0.009	-0.011	0.137

value 0.000 in the third column indicates that, according to this model, shocks to PC2 do not immediately affect PC3.

However, it is difficult to see the dynamics only with the coefficient matrix. We plot therefore the IRF (Equation 3.19). We see from these figures (Figure 5.1) that a shock from any factor is shown to have a significant impact on all other factors, except in one case. Indeed, a shock from the curvature seems to not significantly impact the level at all. All these results show that even if the different factors are uncorrelated together by construction, they seem to still be connected to each others. We will interpret them column by column, trying to interpret them through the term structure transmission mechanisms.

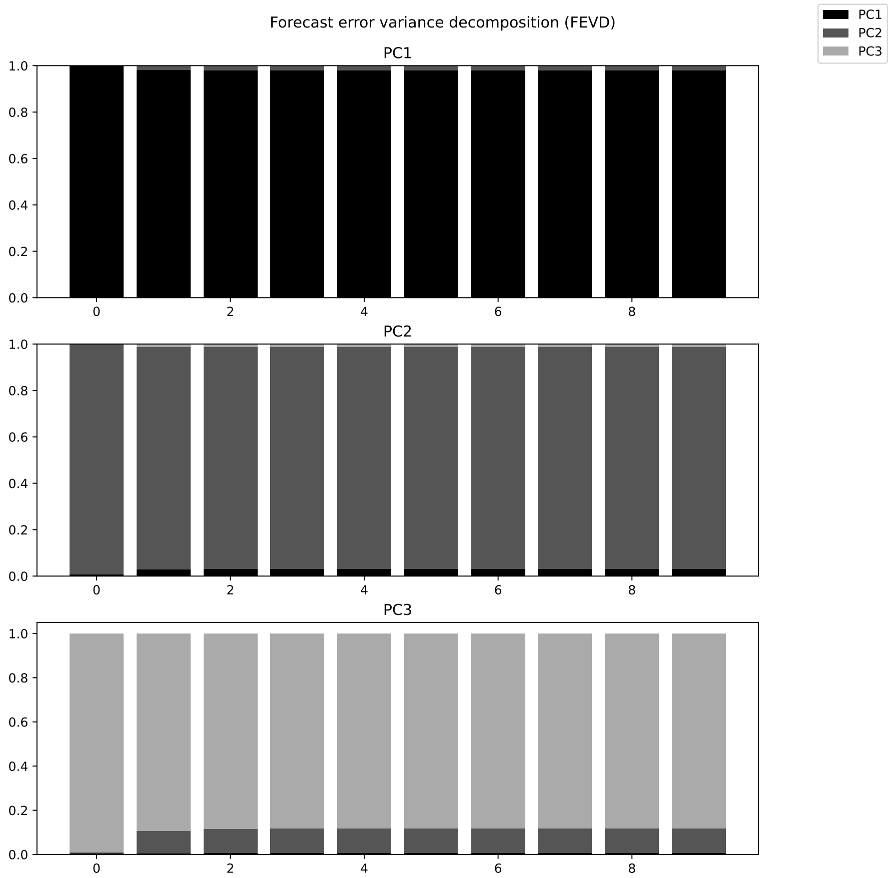
The level factor, which is theoretically linked to expected future inflation, exhibits significant effects on the slope and curvature factors. A permanent shock to the level factor, driven by market expectations of future inflation, results in a decrease in both the slope and curvature factors after several lags. Specifically, a reduction in the slope factor translates to an increase in short-term rates and a decrease in long-term rates, leading to a flattening of the yield curve. This observation is consistent with the financial markets' anticipation of the central bank raising the policy rate in response to inflationary pressures. Consequently, the rise in short-term rates tends to flatten the yield curve, while the effects on midterm and long-term maturities are less clear.

Figure 5.1: IRF - Yield-only VAR



Note : Blue lines represent the response function. The dashed lines represent the 90% confidence intervals. PC1 is the factor representing the level of the yield curve, PC2 the (reversed) slope and PC3 the curvature.

Figure 5.2: Forecast Error Variance Decomposition - Yield-only VAR



A positive shock to the slope factor, which leads to a steepening of the yield curve, also influences the level and curvature factors. When the slope factor increases, it indicates that the difference between short-term and long-term rates is widening. This steepening can prompt the central bank to adjust downward its policy rate to maintain economic stability. Consequently, this adjustment might lead to a decrease in the level factor. Additionally, a steepening curve positively affects the curvature factor, especially impacting midterm maturities. This happens because the slope factor differentiates between short/midterm and long-term rates, while the curvature factor focuses on the midterm segment. A shock to the slope factor, therefore, is reflected in changes in the curvature factor, as the midterm segment adjusts in response to the changes in the slope.

The curvature factor shows no significant immediate effect on the level factor, but has a positive effect on the slope factor. This lack of impact on the level suggests that changes in midterm maturities, while noticeable, are not substantial enough to influence overall inflation expectations or long-term economic fundamentals. We recall that the curvature factor only explains 1.56% of the total variance, when the level explains 89.97% of it. The positive effect on the slope indicates that shifts in midterm maturities can influence the yield curve's steepness. However, the significance of this relationship is weaker compared to the direct impact of the slope factor, also possibly due to the curvature factor's lower explained variation.

The Forecast Error Variance Decomposition (FEVD) analysis (see Annex 5.11, Equation 3.21) reveals that a substantial portion of the variance unexplained by each factor is attributed to the slope factor. However, the percentage contribution is relatively low, likely due to the uncorrelated nature of the Principal Components. Residual analysis (see Annex 6.6) shows that although the residuals are not normally distributed, their densities are smaller compared to those from the ARMA models, indicating a better fit of the VAR model to the data.

It is important to approach these interpretations with caution. The observed dynamics illustrate the overall relationships between the factors but do not specify the exact impact across different maturity ranges (short, mid, or long term). The assumed effects, based on the principal component loadings, may not fully capture the nuanced impacts across different maturities. Additionally, it is unclear whether the relationships are driven by economic transmission or structural linkages between the factors. While the level factor's influence on other factors appears to be due to economic transmission, the interaction between the slope and curvature factors is less straightforward. Comparing these results with those from the "Macro-Yield" model will provide further insights into the differences and enrich our understanding of the yield curve dynamics.

5.3 Macro-Yield VAR

We now model the dynamics of the Yield Curve jointly with the macroeconomic variables. The goal is now to highlight the new mechanisms provided by these additional variables for the dynamics between the factors and these macroeconomic variables. We order the variables based on their level of exogeneity, assuming that GDP has no contemporaneous effect on either inflation or the policy rate, and that inflation likewise has no contemporaneous effect on the policy rate. We also assume that the yield curve factors are exogenous to the real economy. Table 5.4 represents the VAR(1) coefficients of the A matrix and the Cholesky decomposition of the Variance-Covariance matrix (B) (Table 5.5) that we will use for the IRF's.

"In returns" Analysis
In analysing the IRFs (Figure 5.3), we first compare the results of the macroeconomic model with those from the "Yield-Only" model. We observe a notable change in the interactions among the yield curve factors. The level factor now has an immediate positive effect on both the slope and curvature factors, contrasting with the previous model where only delayed negative effects were observed. Although the negative impact on the slope factor remains, it is now insignificant. The slope factor continues to exert an instantaneous negative effect on the level, but now also has a new negative impact on the curvature factor, which was previously positive. Additionally, the curvature factor shows a significant negative effect on the level and a delayed positive effect on the slope. Overall, while the macroeconomic model does not alter the fundamental directions of these relationships, it does impact the confidence intervals and the magnitude of these effects. Specifically, the amplitude

Table 5.4: Coefficient Matrix A for the Macro-Yield model (in returns)

A	Variables					
	PC1	**PC2**	**PC3**	**growth**	**inflation**	**policy**
const	0.021 (0.059)	-0.014 (0.017)	-0.001 (0.007)	0.068 (0.093)	0.010 (0.018)	-0.004 (0.008)
L1.PC1	0.079 (0.055)	-0.019 (0.016)	0.008 (0.007)	0.018 (0.088)	0.018 (0.017)	0.015 (0.008)
L1.PC2	-0.029 (0.199)	-0.007 (0.056)	0.061** (0.024)	0.056 (0.316)	0.036 (0.062)	-0.153 (0.028)
L1.PC3	0.646 (0.423)	0.026 (0.120)	-0.033 (0.051)	-0.012 (0.671)	-0.019 (0.132)	-0.222 (0.059)
L1.growth	-0.019 (0.035)	0.008 (0.010)	-0.003 (0.004)	0.024 (0.056)	-0.014 (0.011)	-0.005 (0.005)
L1.inflation	0.467*** (0.178)	0.027 (0.050)	-0.020 (0.022)	0.241 (0.282)	0.189*** (0.055)	0.020 (0.025)
L1.policy	1.775*** (0.408)	−0.707*** (0.115)	−0.353*** (0.049)	1.189* (0.648)	0.021 (0.127)	0.314*** (0.056)

For significance levels: *** indicates $p < 0.01$, ** indicates $p < 0.05$, * indicates $p < 0.1$. Standard errors of the parameters are shown in parentheses. L1 represents the lagged variable.

Table 5.5: Lower triangular B for the Macro-Yield model (in returns)

B	**PC1**	**PC2**	**PC3**	**Growth**	**Inflation**	**Policy**
L1.PC1	0.051	0.000	0.000	0.000	0.000	0.000
L1.PC2	0.023	-0.032	0.000	0.000	0.000	0.000
L1.PC3	-0.075	-0.087	0.120	0.000	0.000	0.000
L1.Growth	-0.024	-0.026	-0.048	1.656	0.000	0.000
L1.Inflation	0.022	-0.057	-0.022	0.012	0.321	0.000
L1.Policy	0.022	-0.057	-0.022	0.012	0.013	0.128

of reactions, especially between the level and curvature factors, has increased. This suggests that macroeconomic variables have amplified the yield curve reactions, but have not fundamentally changed the underlying dynamics between the yield curve factors.

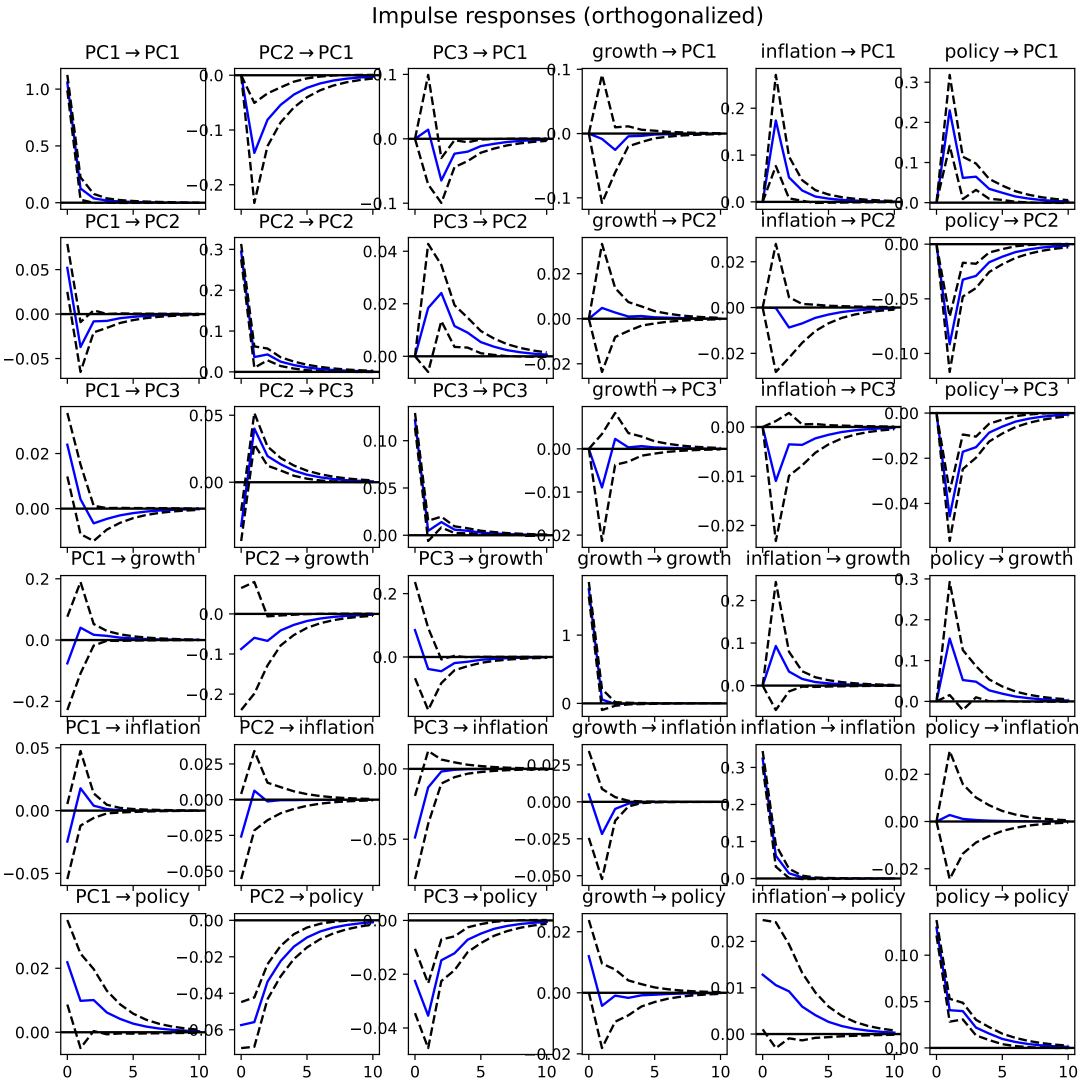
When examining the macroeconomic variables alone, we find that monetary policy does not significantly respond to surprises in output or inflation. Despite some minor significance in reaction to inflation surprises, the policy rate does not show a substantial response (as detailed in Table 5.4). This indicates that the central bank does not seem to react to permanent macroeconomic changes. Furthermore, growth appears to increase inflation, but inflation does not significantly impact growth. These results suggest that, in the long term, the central bank does not directly adjust its policy rate based on real economic variables.

The policy rate significantly affects all the yield curve factors, reflecting its strong influence on short-term rates and thus on the entire term structure. For example, an increase in the policy rate leads to higher short-term rates and a flatter yield curve, which decreases the slope factor. This interpretation also applies to the curvature factor, indicating that short-term rates are heavily influenced by changes in the monetary policy rate.

Interestingly, all yield curve factors trigger a reaction from the central bank. For instance, a permanent shock to the slope factor (resulting in a steeper curve with lower short/midterm rates and higher long-term rates) leads to a decrease in the policy rate. This suggests that the central bank adjusts its monetary policy to align with changes in short-term rates induced by yield curve factors. Similar adjustments are observed for the level and curvature factors. The weaker significance of the policy rate's response to shocks in the level factor, compared to the other factors, might reflect the central bank's limited response to changes in expected inflation. However, the relationships between the level and the other factors (PC1 $\rightarrow$ PC2 and PC1 $\rightarrow$ PC3) indicate that changes in the policy rate affect the short-term rates, which are then transmitted to the entire term structure.

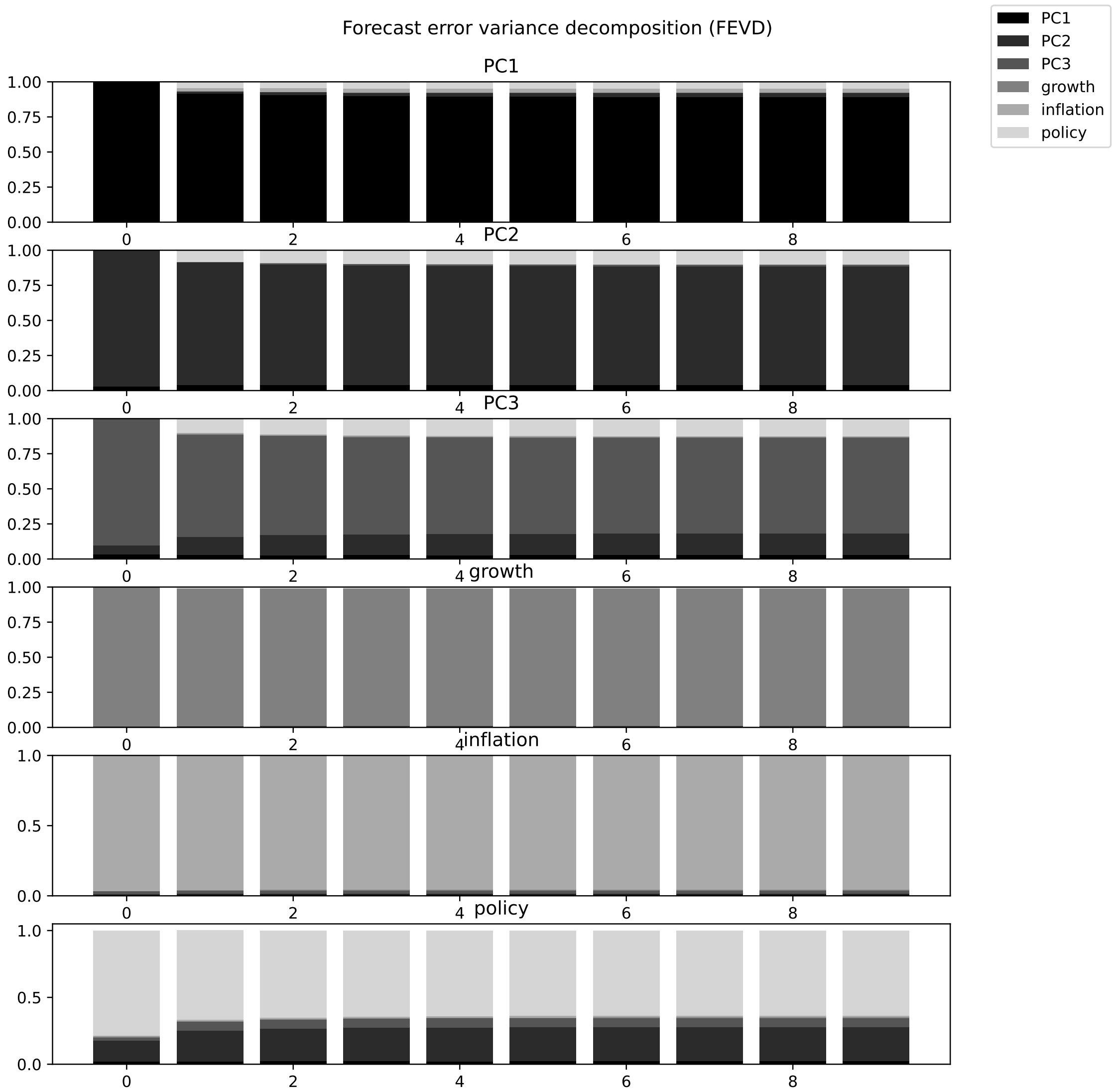
Inflation has a significant impact on the level factor (PC1), consistent with the

Figure 5.3: IRF - Macro-Yield VAR (returns)



Note : Blue lines represent the response function. The dashed lines represent the 90% confidence intervals. PC1 is the factor representing the level of the yield curve, PC2 the (reversed) slope and PC3 the curvature.

Figure 5.4: Forecast Error Variance Decomposition - Macro-Yield VAR (returns)



assumption that this factor reflects expectations of future inflation. However, inflation does not significantly affect the slope or curvature factors, which contradicts the expectation that the slope factor should respond to changes in the price level. Similarly, growth does not significantly impact any yield curve factors, suggesting no direct long-term relationship between the yield curve and economic growth. This finding challenges the notion that the yield curve should directly reflect economic growth fluctuations.
On the other hand, only the curvature has a significant impact on inflation. It is the only factor to have such impact neither on growth nor inflation.

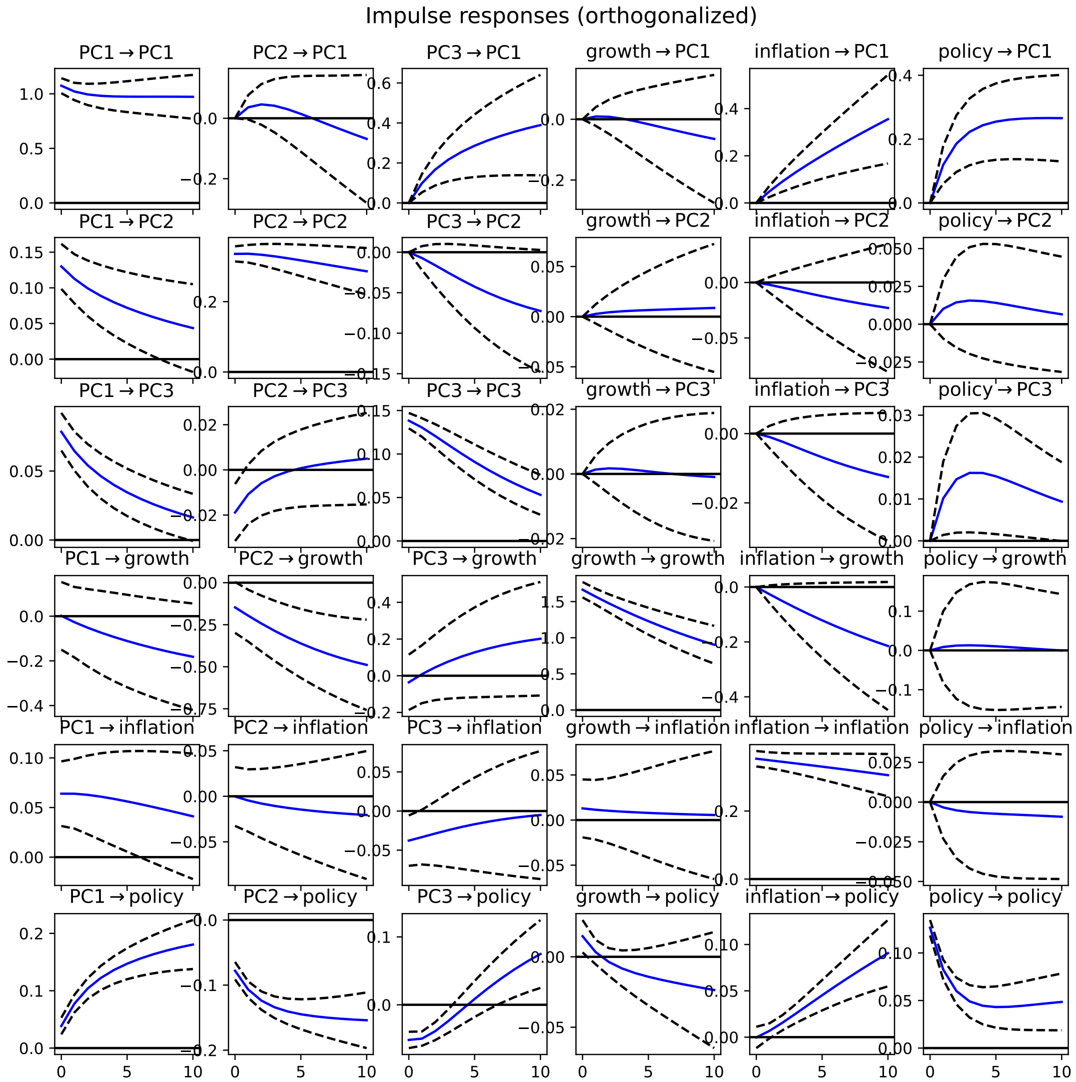
The FEVD analysis (Figure 5.6) shows that inflation and growth are primarily explained by their own dynamics. The yield curve factors, however, incorporate the policy rate into their forecast error variance, indicating the strong role of monetary policy in shaping the curve. The slope factor has a substantial impact on the curvature factor, highlighting how changes in the difference between short-term and long-term rates affect the overall shape of the yield curve. The policy rate is strongly influenced by the yield curve factors, with its importance increasing over time.

"In level" Analysis

We only saw here the long-term relationships between the variables. It could be interesting to also look at these without differentiated variables, to see the reaction to temporary shocks only and thus short/midterm dynamics. However, we recall the different remarks that we made about the use of these data in a PCA and a VAR model. The results might not reflect exactly the true dynamics of the economic relationships. However, we will voluntarily ignore these limitations to still try an interpretation and see if we can highlight valuable information from the model. We will refer to this model as the "level model", in opposition to the previous one referred to the "returns model".

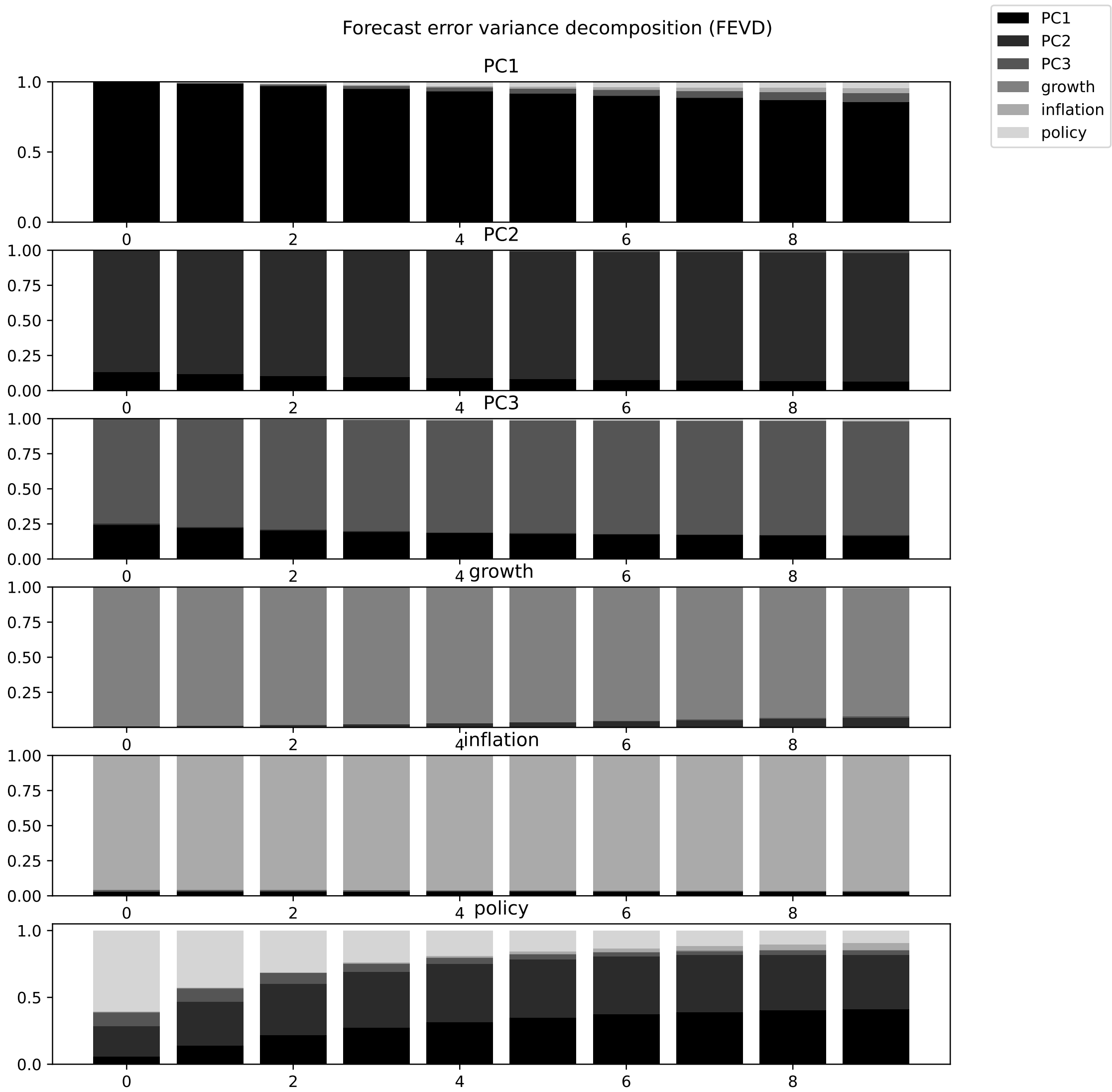
The IRFs depicted in Figure 5.5 reveal significant changes in the relationships between variables. Notably, the policy rate no longer significantly impacts the slope of the yield curve, and its effect on curvature is now more subdued. Conversely,

Figure 5.5: IRF - Macro-Yield VAR (levels)



Note : Blue lines represent the response function. The dashed lines represent the 90% confidence intervals. PC1 is the factor representing the level of the yield curve, PC2 the (reversed) slope and PC3 the curvature.

Figure 5.6: Forecast Error Variance Decomposition - Macro-Yield VAR (levels)



the level factor significantly influences the policy rate's evolution, suggesting that monetary authorities consider the overall level of interest rates in their short-term decision-making. Additionally, there is a notable response to inflation shocks, indicating that monetary policy is more attuned to inflation and inflation expectations in the short to midterm.

To further understand these dynamics, we consider three scenarios: inflation shocks, output shocks, and monetary shocks.
In the case of an inflation shock, there is a direct impact on both the level factor and the bank rate, with no significant effect on other yield curve factors or economic growth. The overall rise in interest rates suggests that investors demand higher yields to offset anticipated inflation. The relationship between the level and other financial factors implies that the shock transmits through the term structure, indirectly influencing the slope and curvature. A sharp increase in the slope factor, likely due to higher long-term interest rates, could lead to a steeper term spread, affecting midterm maturities and reducing output. This outcome contrasts with the typical expectation that a sharper spread, offering lower short-term rates and higher long-term rates, would stimulate economic activity.

An output shock, in contrast, shows no significant linkage to yield curve factors or inflation. The primary observable effect is an increase in the policy rate, though the significance and magnitude of this relationship are debatable. The only consistent relationship is a unidirectional impact from the slope, which appears contradictory to economic expectations.

For monetary shocks, the analysis shows a direct influence on the level and curvature factors, resulting in increased inflation. The yield curve transmission mechanism also indicates a positive impact on the slope, leading to decreased output and higher inflation. This finding contradicts traditional economic theory, suggesting that contractionary monetary policy should reduce inflation.

The Forecast Error Variance Decomposition (FEVD) analysis (Figure 5.6) offers a deeper understanding of the role of yield curve factors in influencing real economic variables. While these factors do not exhibit a strong immediate effect on real variables, the slope factor's increasing significance for output becomes apparent after a few lags. Similarly, the level factor substantially contributes to the variability of other factors, although its influence wanes over time. Conversely, the curvature factor initially has a considerable impact, but its significance decreases over time. Notably, the policy rate increasingly responds to yield curve factors over time, with their influence eventually surpassing the policy rate's own intrinsic determinants. This dynamic reveals an evolving influence of all yield curve factors, particularly highlighting the emergence and growing importance of the level factor, which starts from an initially negligible position at time zero. Meanwhile, the curvature factor's influence diminishes, and the slope factor maintains a consistently strong and increasing impact from the outset.

This model show us that the link between the monetary policy and the real economy, through the interpretation of the yield curve transmission mechanism. Despite the complexities and the limitations noted, it offers some insights into how shocks propagate through the economy via interest rate channels. The monetary policy behaviour exhibits an increasing responsiveness to the overall level of interest rates, particularly the slope and level factors of the yield curve, while displaying an evolving, yet inconsistent, relationship with inflation and output shocks.

5.4 Macro-Yield VAR before and after GFC

There are certain relationships observed in the data that remain unexplained. For instance, it is unclear why a steepening of the yield curve slope correlates with a decrease in output, contradicting theoretical expectations. Similarly, the behaviour of monetary policy appears to be reactive primarily to short-term rate movements, exhibiting inconsistent impacts on inflation. To explore these dynamics further, we consider two distinct periods: before (1993–2006, Annex 6.9) and after (2007–2024, Annex 6.11) the Great Financial Crisis (GFC). The sensitivity of parametric models to the selected time span, coupled with differences in monetary policy between these periods, could influence the observed relationships, potentially reflecting the evolving objectives of central banks. Despite these differences, the factor loadings in both periods maintain the same interpretative framework as previously described.

Our analysis begins with the "in returns" model, focusing on permanent shocks.

"In returns" Models Analysis

The IRFs are reported in the Annexes 6.9 & 6.11. In both periods, the relationship between growth and the yield curve, as well as between growth and inflation, remains insignificant. However, a notable divergence is observed in the response to monetary surprises: prior to the GFC, such surprises do not significantly impact growth, whereas post-GFC, it does. We also note the impact of an inflation surprise on the level factor, insignificant prior-GFC but well after. This change suggests that during the crisis, investors were more concerned about the impact of current inflation on long-term investments, likely due to heightened uncertainty and the limitations faced by conventional monetary policy during that period, which made financial markets more sensitive to inflation as an economic indicator.

Post-GFC, inflation becomes more reactive to the yield curve, with both the slope and curvature factors significantly influencing inflation. The transmission mechanisms through the yield curve are markedly more pronounced after 2007, with the combined impact of the factors being greater and more significant in the second period. This could be attributed to the introduction of unconventional monetary tools such as quantitative easing and forward guidance, which may have led to more interconnected factors over the long run.

The policy rate's response to level movements becomes less significant after 2007, and the level factor does not show significant reactiveness to monetary surprises. The overall pattern of reactions to the other factors remains largely consistent. These observations suggest that the central bank's focus post-GFC was less on the overall level of the term structure and more on accommodating short-term maturities, influenced primarily by the slope and curvature factors. This shift could be linked to the constraints imposed by the Zero Lower Bound (ZLB) on the policy rate, which was near zero for a prolonged period post-crisis. Such constraints may have limited the central bank's ability to influence the overall level of interest rates, prompting a greater reliance on unconventional policies aimed at shaping the slope and curvature rather than the level. This highlights the challenges faced by the Bank of England in lowering long-term rates using conventional monetary tools over the long term.

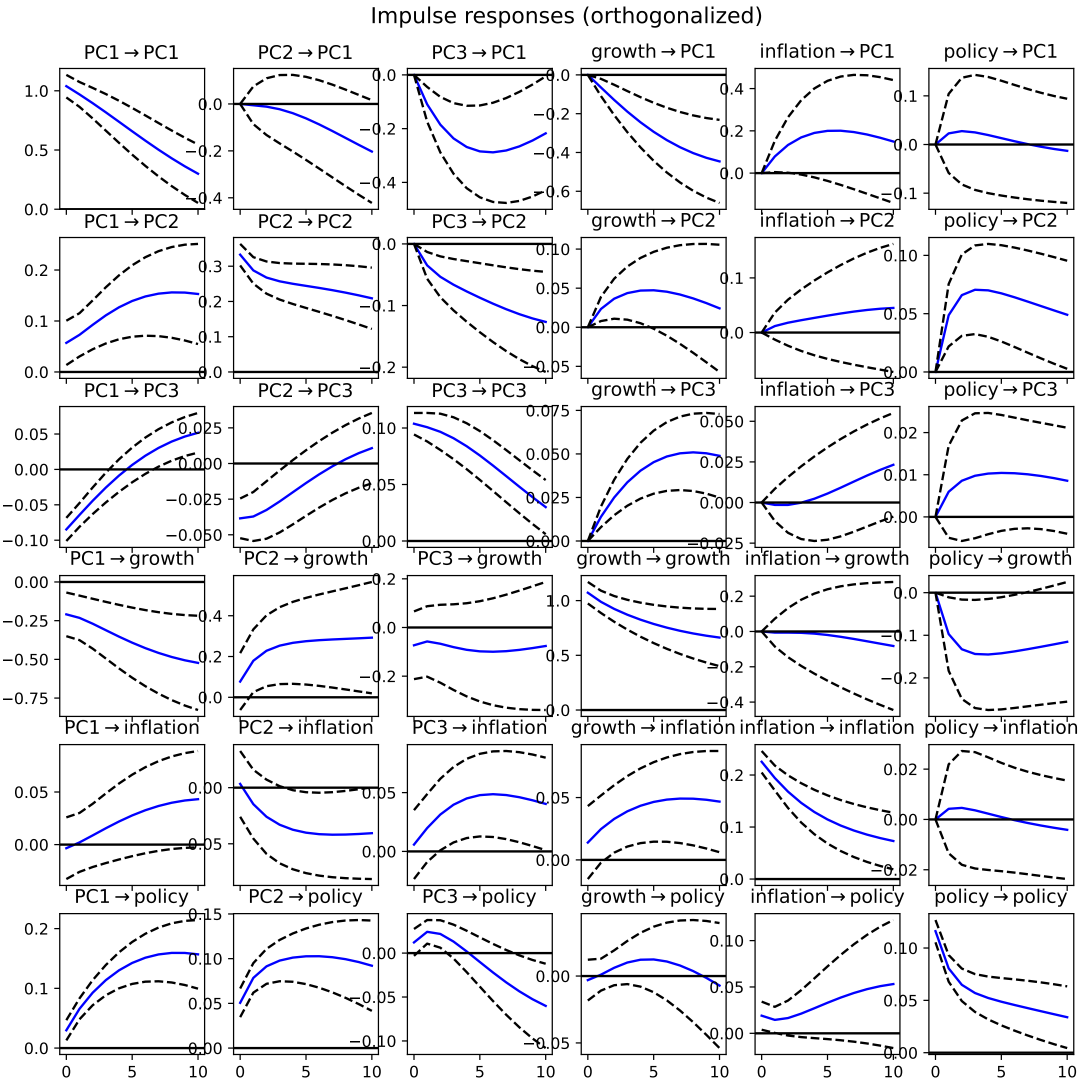
The Forecast Error Variance Decomposition (FEVD) analyses (Annexes 6.10 & 6.12) reveal differences between the pre- and post-GFC periods. Before the GFC, the monetary rate contributed more significantly to the forecast error variance (FEV) of the slope and curvature compared to after the GFC. The most pronounced difference lies in the amplitude of the slope's contribution to the FEV of the monetary rate, which is substantially larger after 2007. This underscores the increased importance of the slope factor in informing monetary policy decisions in the long run. Other contributions remain relatively consistent with the overall model (Figure 5.11), indicating some stability in the relationships despite the changing economic environment.

Pre-GFC "In level" Model Analysis (1993–2006)

We will now pay more attention to the models in level, where we see some interesting differences. In the period before the Great Financial Crisis, the analysis of level models (Figure 5.7) reveals that economic growth had a substantial influence on the yield curve factors, significantly affecting the level, slope, and curvature. Specifically, an increase in growth led to a decrease in the level factor and an increase in both the slope and curvature. This indicates that during this period, growth reduced the level while increasing the difference between short-term and long-term rates (slope) and the curvature of the yield curve. Such results being not seen in the 'in returns' models, we assume the impact of growth on the yield curve being only temporary.

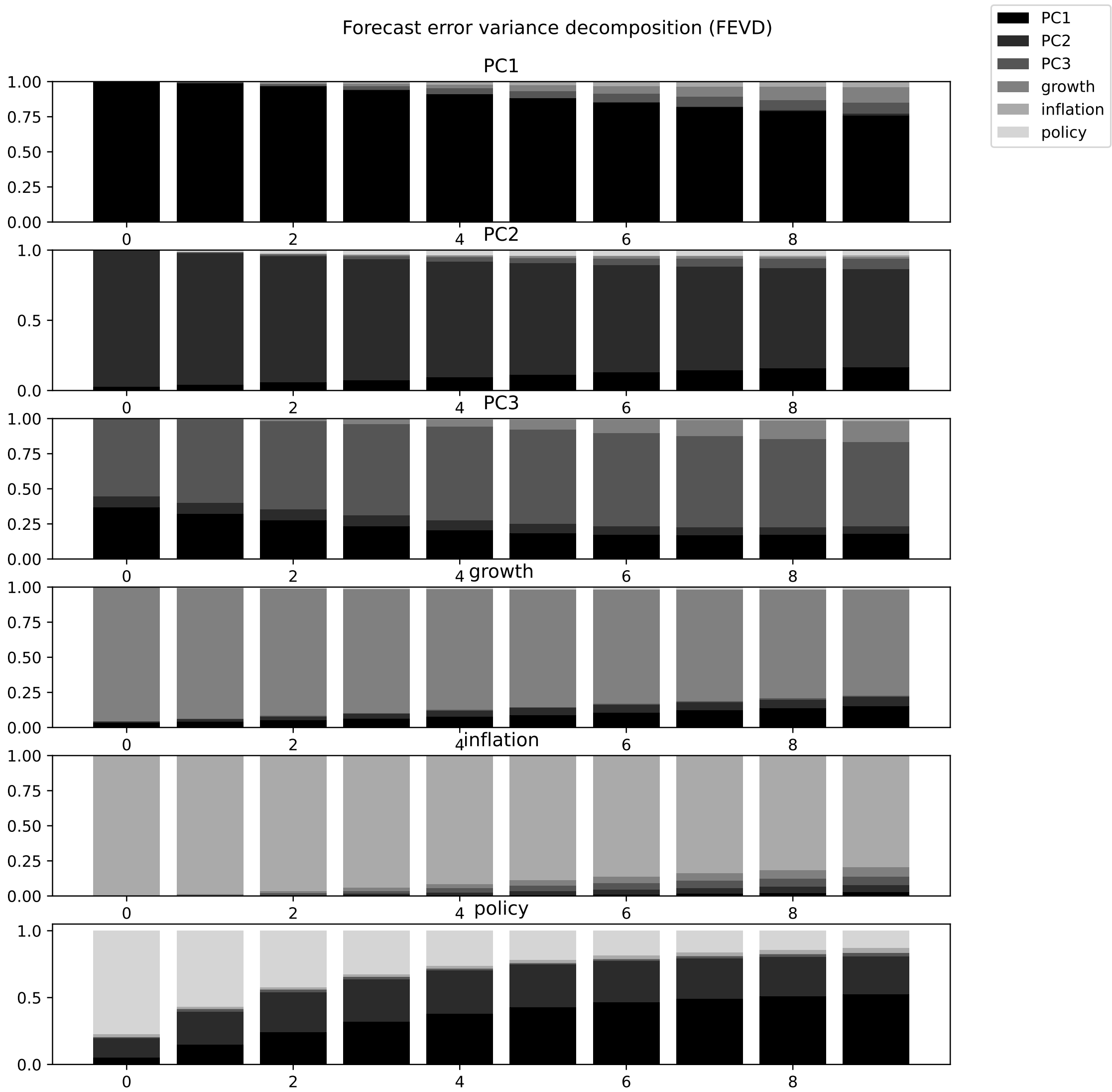
Conversely, both the level and slope factors had a significant impact on growth, with the slope's impact on growth becoming more consistent with economic theory. A steepening slope, traditionally interpreted as a signal of future economic growth, seemed to stimulate investment and, consequently, midterm economic growth. However, inflation's influence on the level factor was negligible, suggesting that the level was not strongly tied to inflation expectations during this period. Instead, the yield curve's impact on inflation dynamics shifted, with inflation becoming more responsive to changes in the slope and curvature rather than the level. This shift underscores the growing importance of midterm maturities in determining inflation

Figure 5.7: IRF - Macro-Yield VAR (1993–2006, level)



Note : Blue lines represent the response function. The dashed lines represent the 90% confidence intervals. PC1 is the factor representing the level of the yield curve, PC2 the (reversed) slope and PC3 the curvature.

Figure 5.8: Forecast Error Variance Decomposition - Macro-Yield VAR (1993-2006, level))



expectations. The main driver of the yield curve seems to be real activity rather than inflation prior to the GFC.

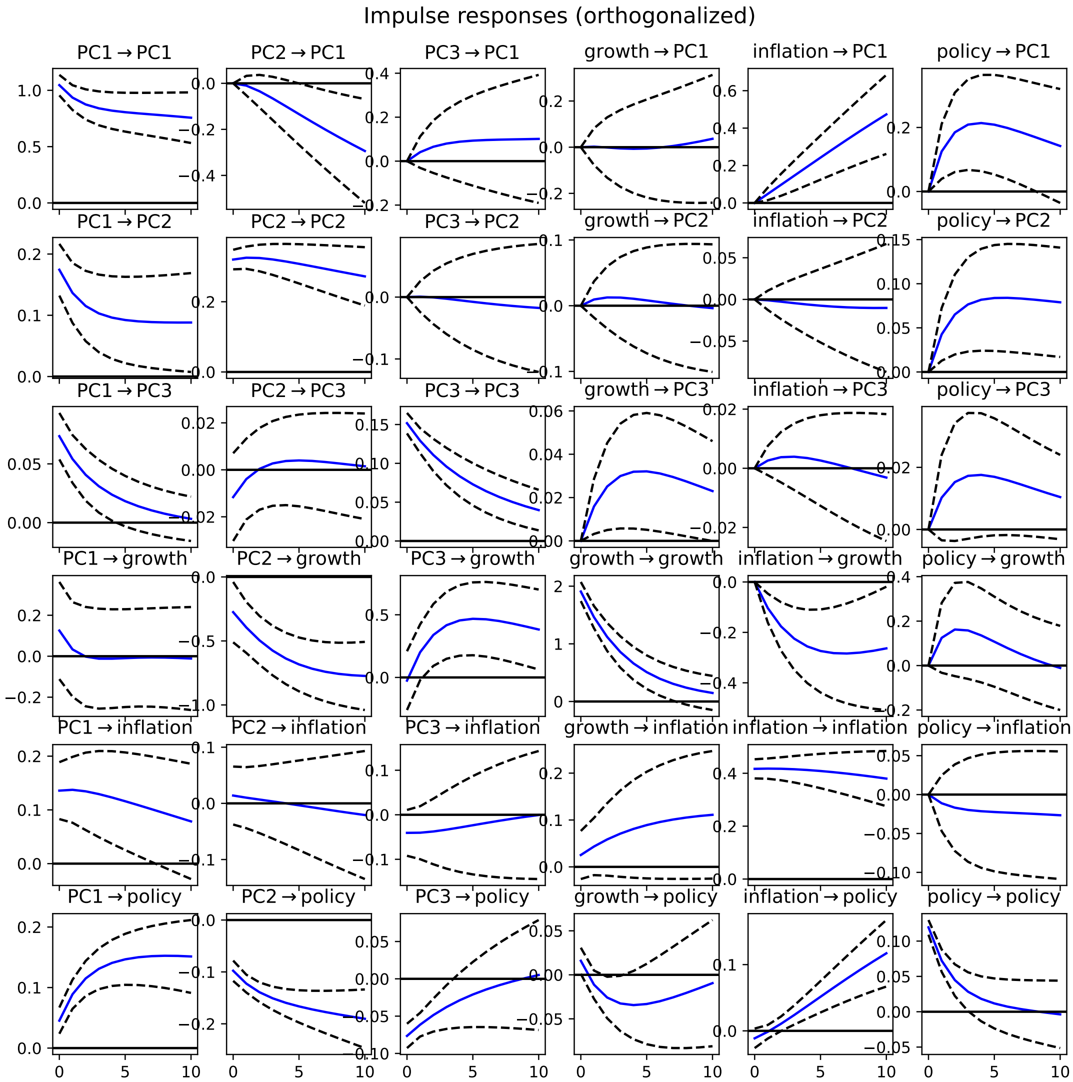
The monetary policy during this period also displayed distinct characteristics. The policy rate's response to changes in the slope shifted direction, indicating a move away from accommodating short-term rate movements. The reaction to curvature shocks similarly reversed, suggesting that the central bank aimed to stabilize the yield curve's shape rather than target a specific level of interest rates. This stabilization effort reflects a broader focus on supporting economic growth over inflation control, as evidenced by the monetary authority's direct response primarily to growth and slope changes. We can make a parallel with the importance of the growth factor, suggesting a deeper concern with the macroeconomic factors that drive economic growth, such as investment and consumption, rather than immediate inflation pressures.

The Forecast Error Variance Decomposition (FEVD) analysis (Figure 5.8) further supports these findings. It shows that macroeconomic variables like growth are predominantly self-explanatory but increasingly influenced by the level factor over time. The level factor's significant role in explaining the variance of the curvature factor also highlights its growing importance, although this influence diminishes in favour of growth over time. The policy rate's variability, initially driven by both the level and slope factors, shows an increasing emphasis on the level factor, indicating a monetary policy focus on long-term interest rate levels and overall yield curve shape stabilization.

Post-GFC "In level" Model Analysis (2007–2024)

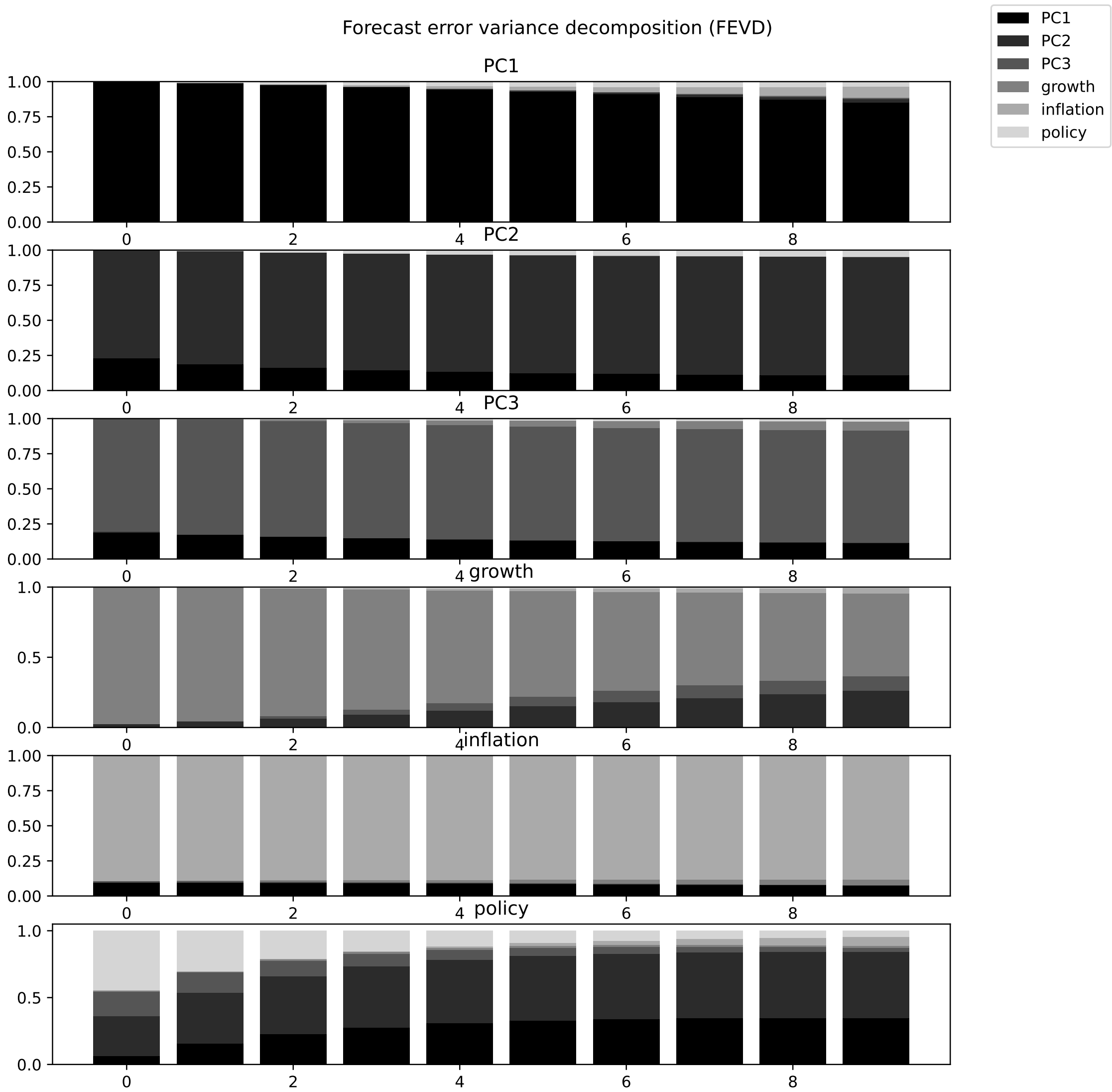
The period following the GFC tells a markedly different story (Figure 5.9). Here, inflation regained its influence on the level factor and the policy rate, while the level factor lost its previous significant impact on growth. This change reflects a shift in the central bank's focus, likely due to the constraints imposed by the Zero Lower Bound (ZLB) on interest rates, which characterizes a liquidity trap scenario. In such a trap, traditional monetary policy tools lose effectiveness as nominal interest rates approach zero, limiting the central bank's ability to stimulate the economy through rate cuts.

Figure 5.9: IRF - Macro-Yield VAR (2007–2024, level)



Note : Blue lines represent the response function. The dashed lines represent the 90% confidence intervals. PC1 is the factor representing the level of the yield curve, PC2 the (reversed) slope and PC3 the curvature.

Figure 5.10: Forecast Error Variance Decomposition - Macro-Yield VAR (2007-2024, level)



Post-GFC, the analysis aligns with the "in returns" models regarding growth, which no longer significantly impacts the yield curve. Instead, inflation emerges as the dominant influence, notably affecting the level factor, growth, and the policy rate. An inflation surprise raises the level factor, reflecting increased expectations of future inflation, while simultaneously reducing growth and prompting a rise in the policy rate. These relationships are consistent with economic theory, highlighting the central bank's focus on controlling inflation due to its broad economic implications.

The monetary policy response post-GFC reflects a strategic pivot. During the crisis, the central bank prioritized short-term liquidity and the stabilization of short-term rates as the money market tightened, shifting focus away from the entire yield curve. This is evident in the response to slope and curvature shocks; the central bank became less accommodating to short-term rate movements (slope) and did not significantly impact curvature. Indeed, the answers to these two factors of the policy rate is inverted compared to prior to the GFC. The slope's influence on growth, which traditionally indicated economic optimism, began to reflect the higher risk premiums investors demanded due to increased uncertainty. This shift suggests that a steepening yield curve no longer signalled economic optimism but rather investor caution, contributing to the paradox where a steepening curve was associated with reduced output growth post-crisis.

In this environment, characterized by a liquidity trap, the central bank's ability to influence long-term rates was constrained. As a result, unconventional monetary tools, such as quantitative easing and forward guidance, became more prominent. These tools aimed to influence the yield curve's slope and curvature, attempting to stimulate economic activity by lowering long-term interest rates even when short-term rates were near zero.

The FEVD analysis (Figure 5.10) post-GFC highlights the increased influence of the slope factor on the policy rate, reflecting the central bank's shift in focus towards managing the shape of the yield curve rather than its overall level. The level factor's impact on inflation became more pronounced, while growth dynamics became significantly influenced by the slope and curvature factors. This shift indicates that during the post-GFC period, the economy became more sensitive to changes in the yield curve's shape and the spread between different maturities. This sensitivity underscores the heightened economic uncertainty and the critical role of financial conditions in driving economic performance during and after the GFC.

Summary

Before the Global Financial Crisis (GFC), monetary policy was focused on stabilizing the yield curve's shape rather than targeting specific interest rate levels. The central bank responded to changes in the slope and curvature of the yield curve, reflecting a concern with supporting economic growth over controlling inflation. This approach indicated a priority on managing growth and investment through the yield curve's structure, rather than addressing short-term inflationary pressures.

With the financial crisis, the monetary policy underwent a notable transformation. Prior to the GFC, the central bank's strategy involved adjusting the policy rate to stabilize economic growth and inflation, focusing on both the yield curve's level and its slope. During this period, economic growth significantly influenced the yield curve's factors, with a steepening slope indicating optimism and stimulating midterm growth. However, post-GFC, traditional tools became less effective due to the Zero Lower Bound (ZLB) on interest rates. The central bank's focus shifted to unconventional monetary policies, such as quantitative easing and forward guidance, which aimed to influence the yield curve's slope and curvature. The results show that post-GFC, inflation regained its impact on the level factor and the policy rate, while growth no longer significantly affected the yield curve. The policy response became less accommodating to changes in the slope and curvature, reflecting a strategic pivot towards managing inflation and stabilizing short-term rates in a constrained environment. This shift underscores the central bank's adaptation to new economic realities and the limitations of traditional monetary policy tools.

5.5 Models Comparison

Here we will compare the different main models that we used above. We will only keep the model using differentiated data from the whole time span (1993-2024), assuming them to have better forecasting performance than models in levels due to VAR and ARMA assumptions.

Table 5.6: Skewness and Kurtosis of the residuals

Skewness	PC1	PC2	PC3	Growth	Inflation	Policy
Yield-Only ARMA	0.698	1.071	1.754	-	-	-
Yield-Only VAR	0.756	0.682	0.463	-	-	-
Macro-Yield VAR	0.708	0.671	-0.887	-1.660	0.366	-1.721

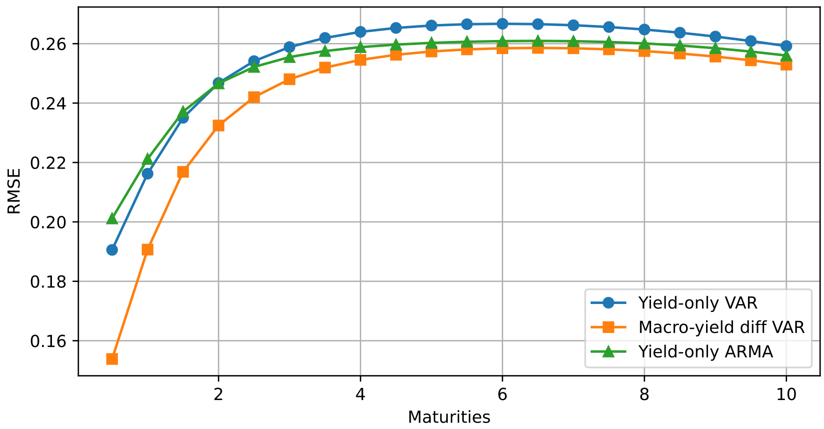
Kurtosis	PC1	PC2	PC3	Growth	Inflation	Policy
Yield-Only ARMA	2.518	5.853	18.157	-	-	-
Yield-Only VAR	2.568	4.297	9.016	-	-	-
Macro-Yield VAR	2.028	5.161	8.859	26.409	3.983	11.480

Note : PC1 is the factor representing the level of the yield curve, PC2 the (reversed) slope and PC3 the curvature.

Table 5.6 displays the measure of skewness and of kurtosis from all residuals of the different models. You can find the kernel distributions of these residuals in the annexe. The objective of comparing them is to see if we can reach values nearer the normal ones (0 for the Skewness, and 3 for the Kurtosis). Indeed, residuals that are closer to normality generally indicate that the model's assumptions are more valid, which can enhance the reliability of forecasts and policy implications. We see that the Yield-only VAR model has “more Gaussian” residuals compared to separate ARMA models, except with PC1. However, we see that PC1 residuals are better affected with the inclusion of macroeconomic variables, when the change is not straightforward for the two other factors.

We can also look to the in-sample RMSE (Equation 3.18) analysis. We compute it for each maturities to compare which model provide the most accurate forecasts across different maturity levels.

Figure 5.11: In-Sample RMSE of the different models (PCA)



We observe that the Macro-Yield model consistently outperforms the other two models across all maturities. Notably, its performance is remarkably strong for short-term maturities, although the differences between models diminish for longer maturities. The Yield-only VAR model also shows slightly superior results for short-term maturities compared to the ARMA models. These findings suggest that cross-relationships and macroeconomic conditions are crucial for modelling short-term maturities, where their impact is significant. Conversely, for mid- to long-term maturities, the intrinsic dynamics of the factors become more prominent. Nevertheless, incorporating macroeconomic variables appears to be highly useful for modelling the term structure of interest rates.

6 Conclusion and Limitations

In this master thesis, we aimed to develop a model of the term structure of interest rates in conjunction with selected macroeconomic variables. We created several models using principal components of yield curves (giving the level, the slope, and the curvature factors), ranging from simple to more complete, with the primary objective of interpreting the relationships revealed by our findings.

In the ARMA models, we analyse the dynamics of the factors independently of each others. We have interpreted the results, showing that all factors are influenced both by historical values and external shocks. We have also seen that the slope has the less persistence compared to the two other factors, and that the level is shown to have a delayed adjustment reaction.
In the "Yield-only" VAR model, we have shown that the yield curve factors, even if uncorrelated together by construction, have influence on each others. We tried to interpret these relationships in the light of the loadings of each factor.
Finally, the "Macro-Yield" model shows that the integration of macroeconomic variables into the yield curve model revealed enhanced interactions and response magnitudes among the yield curve factors, without altering their fundamental dynamics. The central bank's policy response shifted notably post-GFC, focusing more on managing inflation and the yield curve's shape rather than stimulating economic growth, reflecting the constraints of the Zero Lower Bound and the reliance on unconventional monetary tools. This highlights the evolving influence of macroeconomic conditions on monetary policy and yield curve dynamics over time.
The RMSE comparison of these models showed that the "Macro-Yield" model has better in-sample forecasting performance for all maturities than the two other models. We also highlighted that both models exhibit better results for shorter term maturities than longer ones.

However, our method has important limitations. While Principal Component Analysis is powerful, it may be too simplistic to yield the most economically meaningful factors. Comparing its forecasting performance with other models, such as the Dynamic Nelson-Siegel model, could provide a better basis for evaluation. The linearity assumption of the PCA may also be damaging in a context where economic relationships are often non-linear and more complex. Additionally, this method does not account for potential structural breaks, outliers, or regime changes in the economic data, which could significantly impact the results. Exploring non-linear models or incorporating regime-switching mechanisms might provide a more accurate representation of the underlying dynamics. The use of VAR models also presents certain limitations. One major issue is that VAR models require numerous parameters to be estimated, which can lead to overfitting, especially with limited data. This can reduce the model's forecasting accuracy. Additionally, VAR models assume that the relationships between variables are linear and stable over time, which may not hold true in real-world economic contexts where non-linear dynamics and structural breaks are common. We have seen that all the models including different time periods return different results. Furthermore, the identification of shocks in VAR models can be challenging, and the interpretation of impulse response functions may be sensitive to the chosen identification scheme. Future research could explore alternative modelling approaches or non-linear models that can better capture the complexities and potential regime changes in the data.

The inclusion of macroeconomic variables is also subjected to an arbitrary choice. We could include more variables such as the exchange rate, the shadow rate or other monetary policy indicator depending on the subject of analysis. We would also include other economic areas such as the European Union or the United States of America to see if the relationships can be translated or not, and why.

Furthermore, enhancing the robustness of the models by incorporating out-of-sample testing and validation techniques would strengthen the findings. This could involve using rolling window analysis or expanding the dataset to include more recent data, ensuring the models remain relevant and reliable in changing economic conditions.

List of Tables

List of Figures

Bibliography

Ang, A., & Piazzesi, M. (2003). A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary economics*, *50*(4), 745–787.

Bank of England. (2023). *Terminology and concepts - yield curves*. https://www.bankofengland.co.uk/statistics/yield-curves/terminology-and-concepts#notes

Bikbov, R., & Chernov, M. (2010). No-arbitrage macroeconomic determinants of the yield curve. *Journal of Econometrics*, *159*(1), 166–182.

BIS, M. (2005). Economic department: Zero-coupon yield curves: Technical documentation. *BIS papers*, (25).

Bolder, D. J., Johnson, G., & Metzler, A. (2004). An empirical analysis of the canadian term structure of zero-coupon interest rates.

Chen, Y., Niu, L., Chen, R.-B., & He, Q. (2019). Sparse-group independent component analysis with application to yield curves prediction. *Computational statistics & data analysis*, *133*, 76–89.

Christensen, J. H., Diebold, F. X., & Rudebusch, G. D. (2011). The affine arbitrage-free class of nelson–siegel term structure models. *Journal of Econometrics*, *164*(1), 4–20.

Coroneo, L., Nyholm, K., & Vidova-Koleva, R. (2011). How arbitrage-free is the nelson–siegel model? *Journal of Empirical Finance*, *18*(3), 393–407.

Cox, J. C., Ingersoll Jr, J. E., & Ross, S. A. (2005). A theory of the term structure of interest rates. In *Theory of valuation* (pp. 129–164). World Scientific.

Culbertson, J. M. (1957). The term structure of interest rates. *The quarterly journal of economics*, *71*(4), 485–517.

Dewachter, H., & Lyrio, M. (2006). Macro factors and the term structure of interest rates. *Journal of Money, Credit and Banking*, 119–140.

Diebold, F. X., & Li, C. (2006). Forecasting the term structure of government bond yields. *Journal of econometrics*, *130*(2), 337–364.

Diebold, F. X., Rudebusch, G. D., & Aruoba, S. B. (2006). The macroeconomy and the yield curve: A dynamic latent factor approach. *Journal of econometrics*, *131*(1-2), 309–338.

Dodampe Gamage, R. (2021). Empirical analysis and forecasting of yield curves.

Duffee, G. R. (2011). Forecasting with the term structure: The role of no-arbitrage. *manuscript, Johns Hopkins University.*

Duffie, D., & Kan, R. (1996). A yield-factor model of interest rates. *Mathematical finance*, *6*(4), 379–406.

Dziwum, I. R., & Green, E. M. (1996). Interest rate expectations and the shape of the yield curve.

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*(7), 1375–1401.

Gonzalvez, J., Lezmi, E., Roncalli, T., & Xu, J. (2019). Financial applications of gaussian processes and bayesian optimization. *arXiv preprint arXiv:1903.04841*.

Gürkaynak, R. S., Sack, B., & Wright, J. H. (2010). The tips yield curve and inflation compensation. *American Economic Journal: Macroeconomics*, *2*(1), 70–92.

Heath, D., Jarrow, R., & Morton, A. (1992). Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica: Journal of the Econometric Society*, 77–105.

Hull, J., & White, A. (1990). Pricing interest-rate-derivative securities. *The review of financial studies*, *3*(4), 573–592.

Kagraoka, Y., & Moussa, Z. (2013). Quantitative easing, credibility and the time-varying dynamics of the term structure of interest rate in japan. *Journal of International Financial Markets, Institutions and Money*, *25*, 181–201.

King, M. (2005). The inflation target ten years on. *Available at SSRN 709004*.

Litterman, R. B., & Scheinkman, J. (1991). Common factors affecting bond returns. *The journal of fixed income*, *1*(1), 54–61.

Mishkin, F. (2007). Money, banking and financial markets. *New Horizons, Paris, France.*

Modigliani, F. (1944). Liquidity preference and the theory of interest and money. *Econometrica, journal of the econometric society*, 45–88.

Modigliani, F., & Shiller, R. J. (1973). Inflation, rational expectations and the term structure of interest rates. *Economica*, *40*(157), 12–43.

Moller, S. V. (2014). Gdp growth and the yield curvature. *Finance Research Letters*, *11*(1), 1–7.

Mönch, E. (2012). Term structure surprises: The predictive content of curvature, level, and slope. *Journal of Applied Econometrics*, *27*(4), 574–602.

Nelson, C. R., & Siegel, A. F. (1987). Parsimonious modeling of yield curves. *Journal of business*, 473–489.

Novosyolov, A., & Satchkov, D. (2008). Global term structure modelling using principal component analysis. *Journal of Asset Management*, *9*, 49–60.

Omondi, O. (2015). Term structure of interest rates. review of a theory of the term structure of interest rates (cir). *Review of a Theory of the Term Structure of Interest Rates (CIR)(August 26, 2015).*

Oprea, A. (2022). The use of principal component analysis (pca) in building yield curve scenarios and identifying relative-value trading opportunities on the romanian government bond market. *Journal of Risk and Financial Management*, *15*(6), 247.

Reserve Bank of Australia. (n.d.). *Bonds and the yield curve.* https://www.rba.gov.au/education/resources/explainers/bonds-and-the-yield-curve.html

Rudebusch, G. D., & Wu, T. (2008). A macro-finance model of the term structure, monetary policy and the economy. *The Economic Journal*, *118*(530), 906–926.

Sekyere Asare, I. (2019). *Yield curve modelling: A comparison of principal components analysis and the discrete-time vasicek model.* [Doctoral dissertation, Concordia University].

Sowmya, S., & Prasanna, K. (2018). Yield curve interactions with the macroeconomic factors during global financial crisis among asian markets. *International Review of Economics & Finance*, *54*, 178–192.

Svensson, L. E. (1995). Estimating forward interest rates with the extended nelson & siegel method. *Sveriges Riksbank Quarterly Review*, *3*(1), 13–26.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of financial economics*, *5*(2), 177–188.

Vayanos, D., & Vila, J.-L. (2021). A preferred-habitat model of the term structure of interest rates. *Econometrica*, *89*(1), 77–112.

Wellmann, D., & Trück, S. (2018). Factors of the term structure of sovereign yield spreads. *Journal of International Money and Finance*, *81*, 56–75.

Yan, Y., & Guo, J. (2015). The sovereign yield curve and the macroeconomy in c hina. *Pacific Economic Review*, *20*(3), 415–441.

Annexes

Figure 6.1: Nelson-Siegel model Scores for $\lambda = 0.0609$

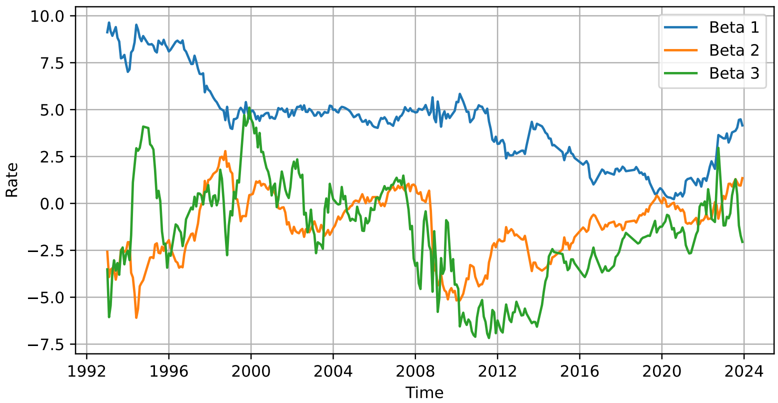


Table 6.1: Dickey-Fuller test for PC's of raw yield

Principal Component	P-Value
PC1	0.466
PC2	0.110
PC3	0.009

Figure 6.2: Components Selection - Raw Data

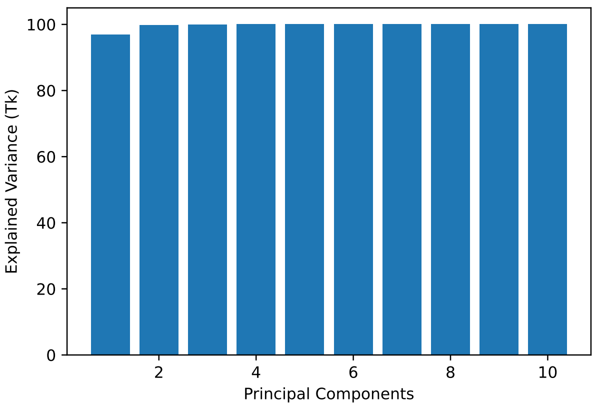


Figure 6.3: Scores - Raw Data

Figure 6.4: Loadings - Raw Data

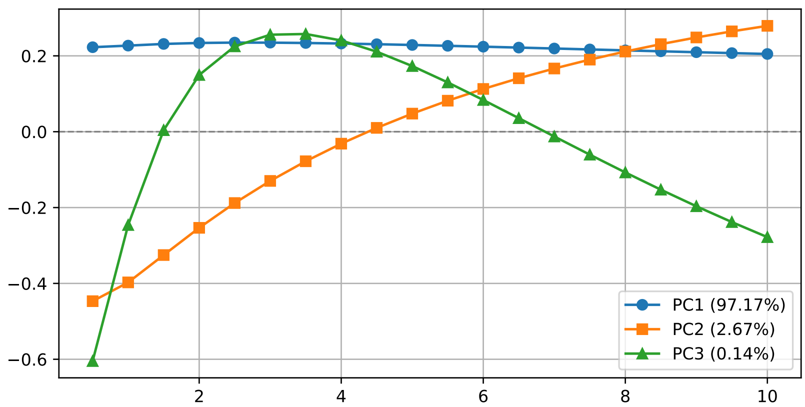


Table 6.2: Correlation for PC1, PC2, and PC3 (PCA returns)

	PC1	PC2	PC3
PC1	100%	-	-
PC2	2.75e-14%	100%	-
PC3	1.054733e-14%	8.98e-17%	100%

Table 6.3: Correlation Matrix of NS factor ($\lambda = 0.0609$)

	Beta 1	Beta 2	Beta 3
Beta 1	100%	-	-
Beta 2	-68.91%	100%	-
Beta 3	11.09%	-46.00%	100%

Table 6.4: Dickey-Fuller test for Macro-variables

Variable	P-Value
GDP	0.4029225
Inflation	0.0111306
Bank Rate	0.0856185

Table 6.5: Best ARIMA Lag Orders (p, 0, q) for Principal Components

Component	BIC	AIC	HQ
PC1	(1, 0, 1)	(2, 0, 3)	(3, 0, 0)
PC2	(1, 0, 0)	(3, 0, 4)	(1, 0, 1)
PC3	(0, 0, 0)	(2, 0, 3)	(2, 0, 3)

Table 6.6: Ljung-Box test of ARMA residuals (p-value)

Lag	PC1	PC2	PC3
1	0.955	0.769	0.595
2	0.998	0.872	0.835
3	0.997	0.528	0.715
4	0.999	0.637	0.637
5	0.980	0.312	0.329
6	0.975	0.374	0.336
7	0.990	0.411	0.102
8	0.996	0.362	0.141
9	0.998	0.440	0.184
10	0.999	0.414	0.190

Figure 6.5: Kernel density of the residuals - Yield-only ARMA

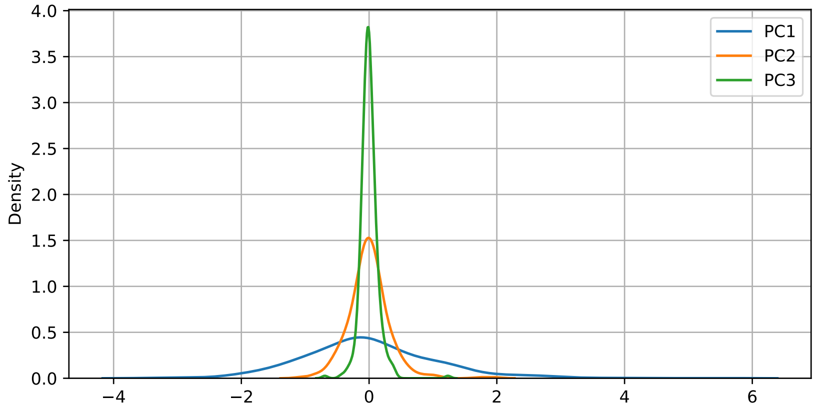


Figure 6.6: Kernel density of the residuals - Yield-only VAR

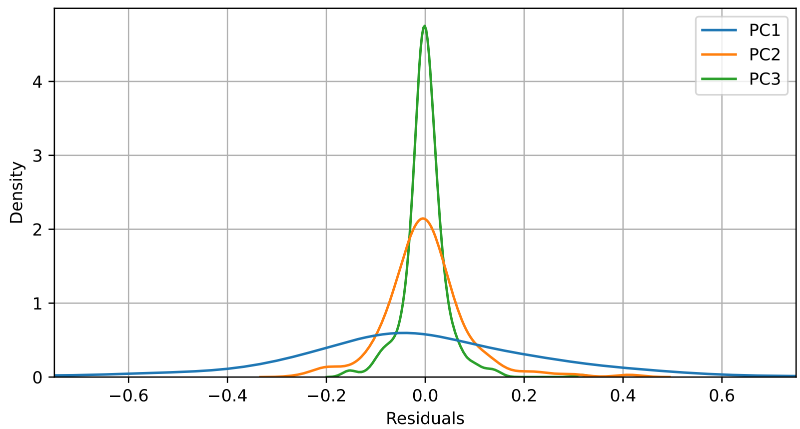


Figure 6.7: Residuals of Yield Curve factors - Macro-Yield VAR (returns)

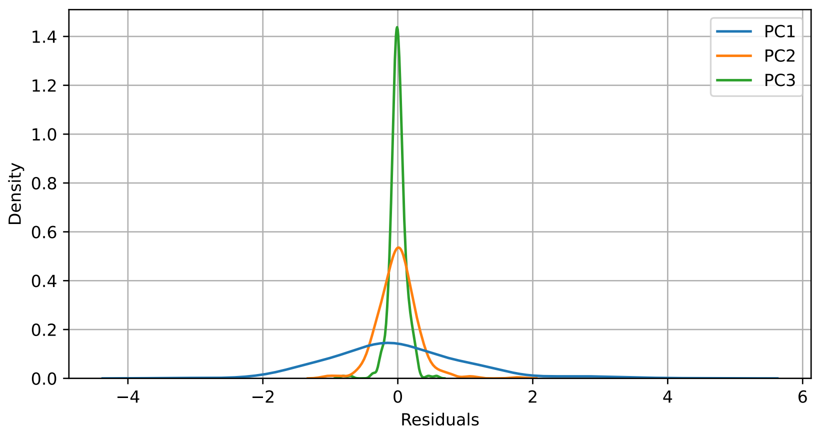


Figure 6.8: Residuals of macroeconomic variables - Macro-Yield VAR (returns)

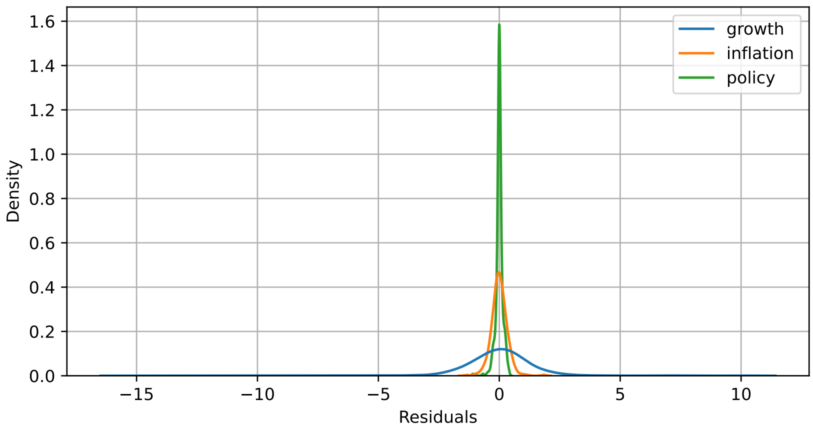
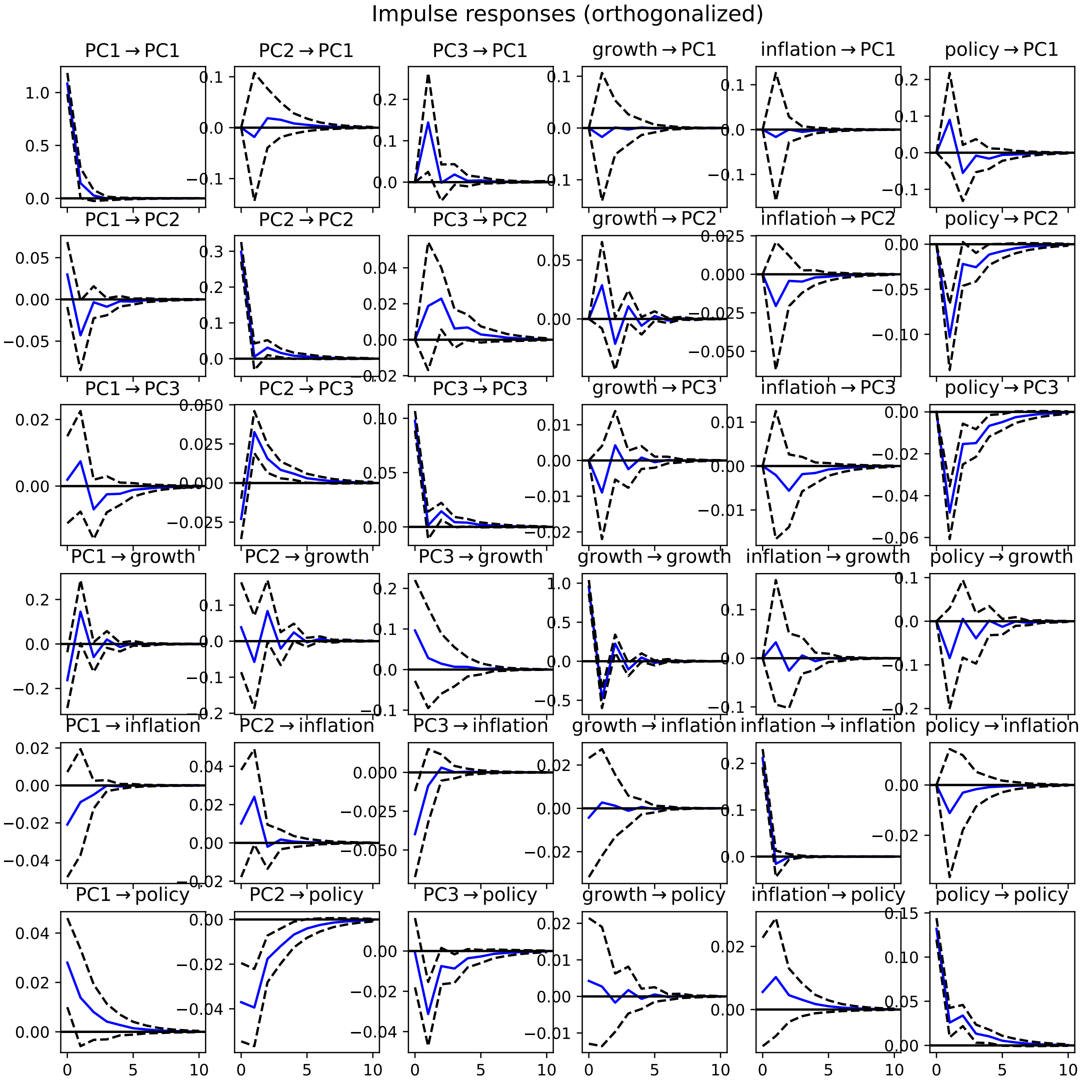


Figure 6.9: IRF - Macro-Yield VAR in returns (1993-2006)



Note : Blue lines represent the response function. The dashed lines represent the 90% confidence intervals. PC1 is the factor representing the level of the yield curve, PC2 the (reversed) slope and PC3 the curvature.

Figure 6.10: Forecast Error Variance Decomposition - Yield-only VAR in returns (1993-2006)

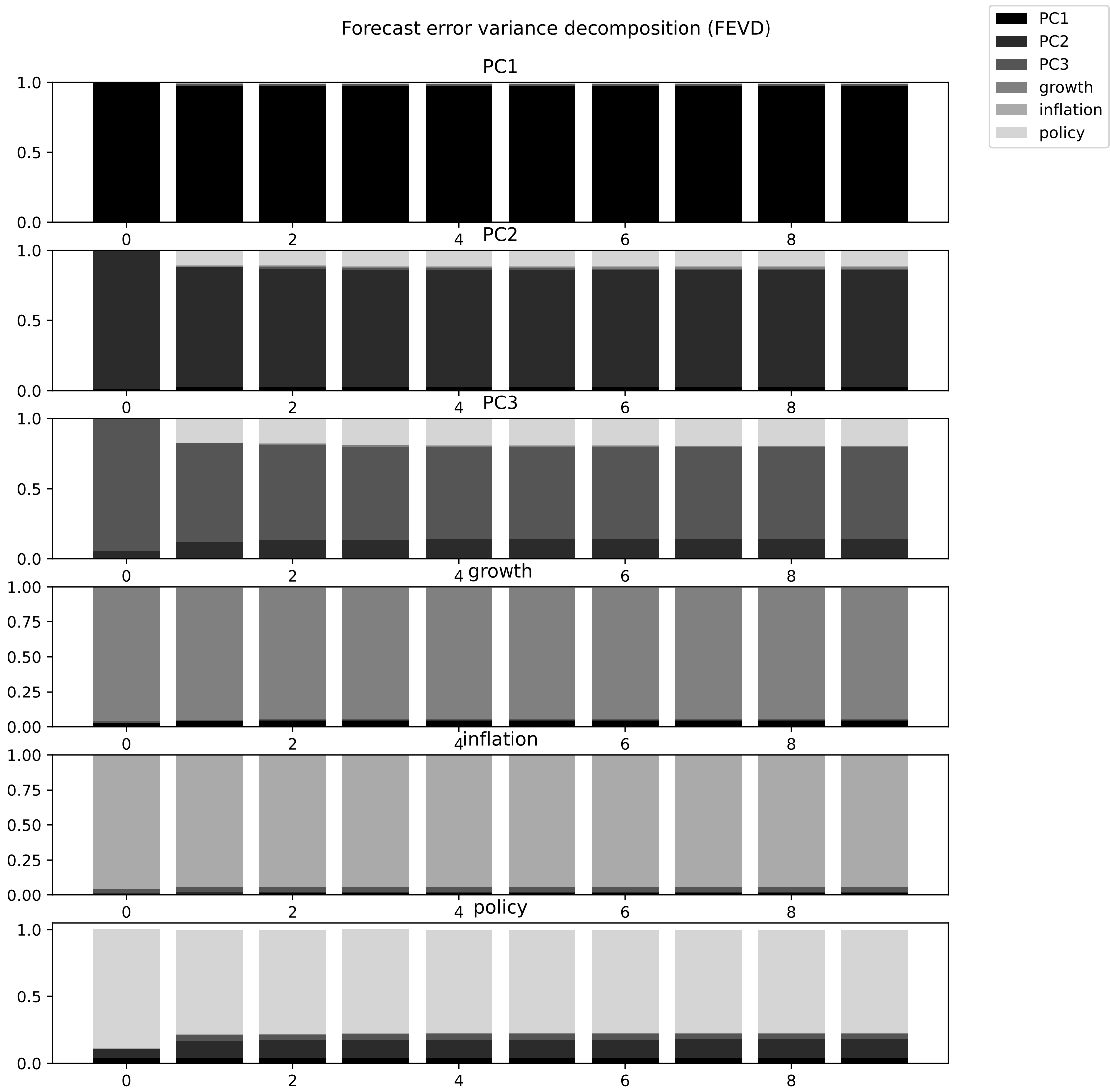
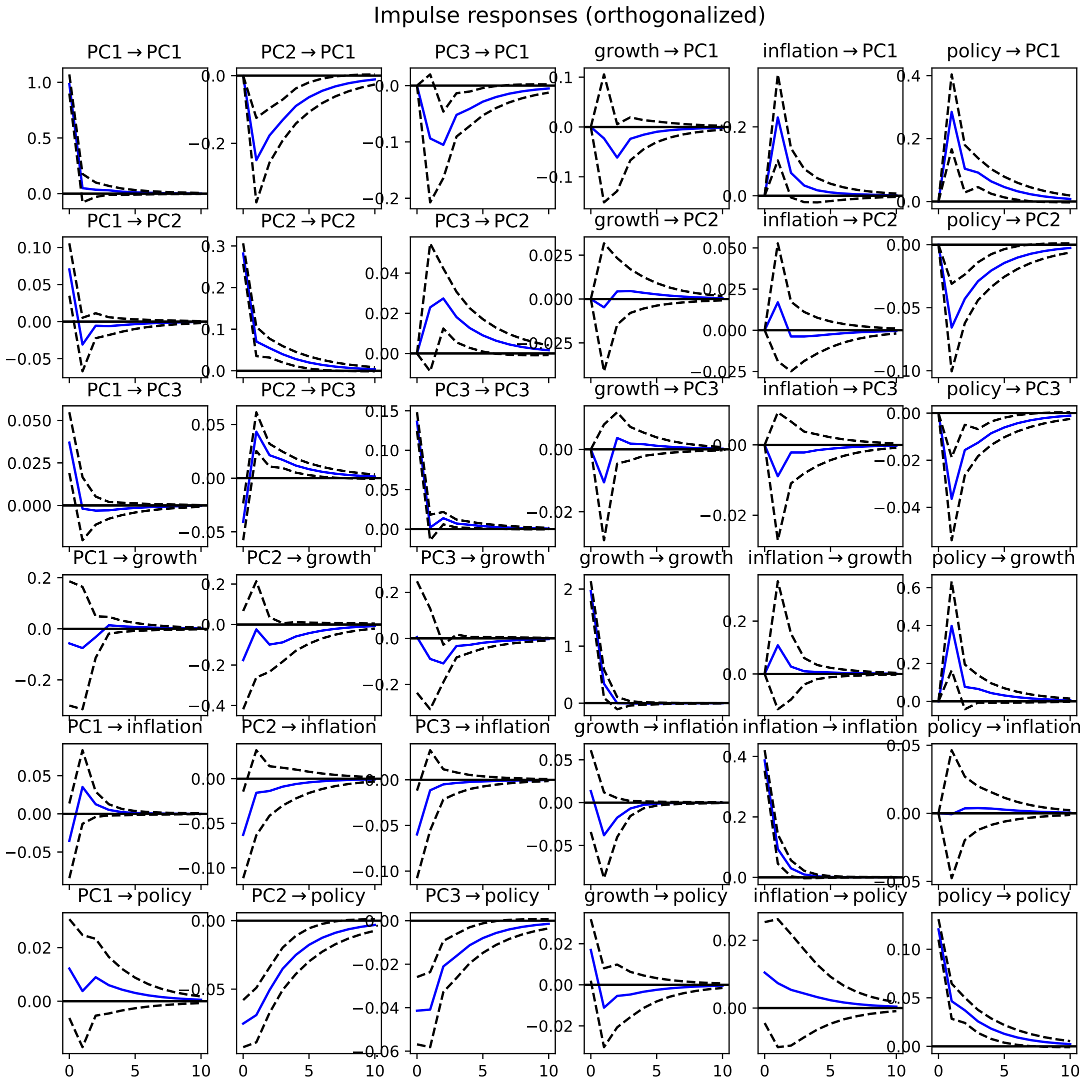


Figure 6.11: IRF - Macro-Yield VAR in returns (2007–2024)



Note : Blue lines represent the response function. The dashed lines represent the 90% confidence intervals. PC1 is the factor representing the level of the yield curve, PC2 the (reversed) slope and PC3 the curvature.

Figure 6.12: Forecast Error Variance Decomposition - Yield-only VAR in returns (2007-2024)