UNIL | Université de Lausanne
HEC Lausanne

University of Lausanne
Fall 2023

The more the better ? Forecasting a small open economy using a FAVAR : Evidence from Switzerland

by Viktor Gradoux

Supervisor: Mathieu Grobéty

Expert: Brendan Berthold

Master (MSc) thesis in Economics

Major: Quantitative Economics & Macroeconomic Policy

UNIVERSITY OF LAUSANNE (HEC)

Acknowledgements

I would like to express my sincere thanks to my master thesis supervisor, Mathieu Grobéty (executive director of CREA) who gave me the opportunity to do this master thesis in his institute (CREA, Swiss Institute of Applied Economics) and his expert guidance and support throughout the research process. I am equally thankful to Brendan Berthold, scientific collaborator at CREA and PhD candidate at UNIL, for his generous mentorship in the coding part of my master thesis. His valuable guidance and wealth of knowledge in R have been important in enhancing my coding skills and understanding.

Moreover, I would like also to express my sincere thanks to Suleymane Ceran (head of CEDIF) for his exceptional assistance in automating the dataset update requests.

Additionally, my profound thanks go to Prof. Jean Paul Renne (UNIL), Marco Sotonova (CREA), Adrien Tschopp (UNIL), Adeline Demierre (HFR), Dario Vacchini (UNIL), Lisa Wenger (SNB), Pasha Mammadov (UNIL), Mora Klimberg (IATA), Sébastien Blanco (SNB) and Abel Deillon (SECO) for their valuable input, insightful comments, and suggestions. Their expertise and constructive feedback have enriched the quality of my work.

Abstract

Conducted for CREA (Swiss Institute of Applied Economics of University of Lausanne), we propose a Factor Augmented VAR (FAVAR) tailored for a small economy to forecast real Swiss GDP growth. We use principal components analysis and expectation-Maximization algorithm to extract factors from a diverse set of a large number of variables with missing values and modeled them with the target variable within a VAR framework in a two step estimation approach. Our FAVAR remains structural by dividing the model into two blocks : one for foreign and one for domestic factors. Each block encompasses a real activity, inflation, money supply & interest rates and financial conditions factor, contributing to the model's economic interpretability. To assess its out-of-sample performance, we compare our FAVAR model against six different benchmarks, including univariate and multivariate models, across various forecasting horizons (up to 12 quarters ahead) and over three different evaluation samples. Our key finding suggests that the performance of the FAVAR increases with the forecasting horizon. However, the Covid-19 crisis and the war in Ukraine present challenges to our FAVAR's performance. During this period (2020Q1-2023Q2), we found a decline in forecast performance over the long term as well as a general deterioration across all horizons except when compared to the random walk.

JEL Classification – C32, C38, C53, E32, E37

Keywords – Forecasting Real GDP, Factor Augmented VAR (FAVAR), Small Open Economy, Principal Components Analysis (PCA), Out-of-sample evaluation

Contents

List of Figures **v**

List of Tables **vii**

1 Introduction **1**
1.1 Introduction 1

2 Literature Review **6**
2.1 The birth of FAVAR 6
2.2 Small open economy (SOE) framework 7
2.3 Forecasting with FAVAR and factor models 8

3 Data **10**

4 Methodology **14**
4.1 Econometric framework: FAVAR 14
4.2 A small open economy (SOE) FAVAR 15
4.3 Identification of factors with economic interpretation 16
4.4 Estimation 18
4.5 Diagnostic checking 22

5 Results **26**
5.1 Unconditional forecast 26
5.2 Pseudo relative out-of-sample evaluation 28
5.3 Assessing out-of-sample performance Pre-Covid 34
5.4 Assessing out-of-sample performance in the Covid period and after 37
5.5 Test of forecast unbiasedness and efficiency 42

6 Discussion & Areas of improvement **44**
6.1 Real time forecasting and specification issues 44
6.2 Dealing with the (post) Covid period 45

7 Conclusion **47**

Appendix **52**
A1 Data appendix - Variables table 53
A2 Variables visualisation 58
A3 EM-algorithm 66
A3.1 Implementation in R : PCA EM-algorithm with MissMDA library 67
A4 Algorithm of extraction of domestic factors 69
A4.1 Matrice notation 69
A5 Scree plots PCA (foreign variables) 72
A6 Diagnostic checking - Tests on residuals 76
A7 Out-of-sample forecasts - Hairy graphs 85
A8 Forecast Error Variance Decomposition (FEVD) 87
A9 Structural shocks investigation and IRFs (impulse response functions) . . 88
A9.1 Recursive Approach 88

A9.2 An increase in foreign activity 90
A9.3 An increase in foreign inflation 91
A9.4 An increase in foreign money supply and interest rates 92
A9.5 An increase in foreign financial conditions 93
A9.6 An increase in domestic real activity 94
A9.7 An increase in domestic inflation 95
A9.8 An increase in domestic money supply and interest rates 96
A9.9 An increase in domestic financial conditions 97
A10 Other tables 98

List of Figures

3.1 **Data set shares by type of variable** . . . 11
4.1 Foreign factors . . . 20
4.2 Domestic factors . . . 21
4.3 OLS-CUSUM test (sample : 2000Q1-2023Q2) . . . 23
4.4 Inverse roots of VAR characheristic polynominal . . . 24
5.1 Real Swiss GDP growth forecast (2023Q3-2026Q2), QoQ growth rates . . . 27
5.2 Sequence of out-of-sample forecasts (3-year ahead) for Swiss GDP growth based on a recursive estimation scheme (2010Q1-2023Q2), QoQ growth rates - FAVAR model . . . 29
5.3 Out-of-Sample Assessment - Forecasting Squared Errors . . . 41
A2.1 **Real Activity variables (Switzerland) - Growth Rates (QoQ)** . . . 58
A2.2 **Inflation variables (Switzerland) - Growth Rates (QoQ)** . . . 59
A2.3 **Money & Interest Rates variables (Switzerland) - Growth Rates (QoQ)** . . . 60
A2.4 **Financial Conditions variables (Switzerland) - Growth Rates and First Difference (QoQ)** . . . 61
A2.5 **Real Activity variables (World) - Growth Rates (QoQ)** . . . 62
A2.6 **Inflation variables (World) - Growth Rates (QoQ)** . . . 63
A2.7 **Money and Interest Rates variables (World) - Growth Rates and First Difference (QoQ)** . . . 64
A2.8 **Financial Conditions variables (World) - Growth Rates (QoQ)** 65
A5.1 Scree plot of principal components (in % of the total variance explained) of the foreign real activity variables . . . 72
A5.2 Scree plot of principal components (in % of the total variance explained) of the foreign inflation variables . . . 73
A5.3 Scree plot of principal components (in % of the total variance explained) of the foreign money and interest rates variables . . . 74
A5.4 Scree plot of principal components (in % of the total variance explained) of the foreign financial conditions variables . . . 75
A6.1 **Residuals - Domestic Factor Real Activity regression** . . . 76
A6.2 **Residuals - Domestic Factor Inflation regression** . . . 77
A6.3 **Residuals - Domestic Factor Money and Interest Rates regression** 78
A6.4 **Residuals - Domestic Factor Financial Conditions regression** . . . 79
A6.5 **Residuals - Swiss GDP regression** . . . 80
A6.6 **Residuals - Foreign Factor Real Activity regression** . . . 81
A6.7 **Residuals - Foreign Factor Inflation regression** . . . 82
A6.8 **Residuals - Foreign Factor Money Supply and Interest rates regression** . . . 83
A6.9 **Residuals - Foreign Factor Financial Conditions regression** . . . 84
A7.1 Sequence of out-of-sample forecasts (3-year ahead) for Swiss GDP growth based on a recursive estimation scheme (2010Q1-2023Q2), QoQ growth rates - ARMA(3,4) model . . . 85
A7.2 Sequence of out-of-sample forecasts (3-year ahead) for Swiss GDP growth based on a recursive estimation scheme (2010Q1-2023Q2), QoQ growth rates - FAVAR(3) model without blocks . . . 86

A8.1 **Forecast Error Variance decomposition for Swiss GDP growth - Long run (horizon 9 to 12)** 87
A9.1 Shock to foreign activity on economic variables 90
A9.2 Shock to foreign inflation on economic variables 91
A9.3 Shock to foreign money supply and interest rates on economic variables . 92
A9.4 Shock to foreign financial conditions on economic variables 93
A9.5 Shock to domestic real activity on economic variables 94
A9.6 Shock to domestic inflation on economic variables 95
A9.7 Shock to domestic money supply and interest on economic variables . . . 96
A9.8 Shock to domestic financial conditions on economic variables 97

List of Tables

4.1 Portmanteau Test (asymptotic) on residuals 24
4.2 Normality Test on Residuals (Multivariate) - Jarque Bera Test 25
4.3 Normality Test on Residuals (Multivariate) - Skewness 25
4.4 Normality Test on Residuals (Multivariate) - Kurtosis 25
5.1 Out-of-Sample Assessment: Ratio of Mean Squared Forecast Errors - Univariate benchmarks 33
5.2 Out-of-Sample Assessment: Ratio of Mean Squared Forecast Errors - Multivariate Benchmarks 34
5.3 Out-of-Sample Assessment: Ratio of Mean Squared Forecast Errors - Univariate benchmarks - Covid Excluded 35
5.4 Out-of-Sample Assessment: Ratio of Mean Squared Forecast Errors - Multivariate Benchmarks - Covid Excluded 36
5.5 Out-of-Sample Assessment: Ratio of Mean Squared Forecast Errors - Univariate benchmarks - Covid Only 38
5.6 Out-of-Sample Assessment: Ratio of Mean Squared Forecast Errors - Multivariate Benchmarks - Covid Only 39
5.7 Out-of-Sample Assessment: Ratio of Mean Squared Forecast Errors - Average over all horizons (from 1 to 12) 40
5.8 Test of forecast unbiasedness and efficiency (Mincer-Zarnowitz regression) 43
A10.1 FAVAR lag selection 98
A10.2 Cointegration test - Johansen procedure 98

1 Introduction

1.1 Introduction

The forecasted growth rate of real GDP has drawn attention for many decades, serving as a key indicator of an economy's health. This indicator is constantly tracked by agents, including households and firms, who consider it as a measure of the overall future economic activity. Central banks, in particular, rely on GDP forecasts to assess future inflation levels and estimate the output gap.

In the wake of the advent of online platforms in 2000's where thousands of economic time series data became readily available in real-time, a significant shift occurred with the research of Stock and Watson (2002) who introduced the Dynamic Factor Model (DFM) for forecasting in data-rich environments. This work that follows can be motivated by a question well posed by James Stock and Mark Watson : "Can we move from small models with forecasts adjusted by judgmental use of additional information, to a more scientific system that incorporates as much quantitative information as possible ?" [1]

While the literature exploring large datasets for forecasting macroeconomic variables has shown that these factors models can be successful, using more information has become increasingly evident. In today's data-driven society, professional forecasters recognize the necessity of incorporating more relevant data to enhance the accuracy of their forecasts. In light of these advances, there is a growing need to use this wealth of information for improving the forecasting accuracy of the existing models and enabling more informed economic decision-making.

Thus, our methodology for forecasting Swiss GDP growth over the next three years (up to 12 quarters) is based on a FAVAR model developed by Bernanke et al. (2005). This model is initially developed to assess the impact of monetary policy shocks exploring the advances of dynamic factors models from Stock and Watson (2002) and reconciling with VAR methodology (Sims, 1980). The FAVAR model, along with its extensions such as the Time-Varying Parameter FAVAR and Factor Augmented Error Correction Model

[1] What's New in Econometrics - Time Series Lecture 11: Forecasting and Macro Modeling with Many Predictors, Part I", NBER Summer Institute, 2008

(FECM) has been widely used for forecasting GDP, inflation and other macroeconomic variables (Banerjee et al., 2014; Eickmeier et al., 2011; Faust and Wright, 2013). FAVAR model offers various advantages in data-rich environments. It addresses the curse of dimensionality inherent in standard VAR models and mitigates issues related to omitted variable bias. By incorporating these factors, the FAVAR captures the joint movements and interactions among a wide range of economic variables. In FAVAR models, the information contained in a large number of variables is first expressed using a few latent factors, which are then used in a standard VAR model. Our framework of a FAVAR à la Bernanke et al. (2005) and à la Boivin et al. (2009) is enriched of a small open economy framework in the spirit of Vasishtha and Maier (2013) focused on the international trade with the variables of the main key trading partners of Switzerland taking also into account the specificity of the Swiss economy.

In the context of the standard VAR framework, the target variable for forecasting which is, in this case, the Swiss GDP growth is modelled jointly with factors derived from a wide range of time series data [2]. The specific variables used in the model are detailed in the appendix. Principal Component Analysis (PCA) combined with the Expectation-Maximization (EM) algorithm are employed to extract the factors from the large pool of time series variables. The modelling approach includes the division of variables into two blocks : a domestic block representing Switzerland's economic indicators and a foreign block comprising variables from seven key trading partners of Switzerland. It is important to note that in this framework, the assumption is made that the dynamics of the foreign block are not governed by the domestic block. This assumption is driven by the consideration of Switzerland as a small open economy, where external factors play a significant role in shaping the country's economic dynamics.

Our main contribution to the literature is that our FAVAR remains "structural" as well as for the domestic block which is to our best knowledge not the case in the literature. In each block, there is one factor for real activity, one factor for inflation, one factor for money supply and interest rates and one factor for financial conditions. In line with the approach used by Boivin et al. (2009) and Burri and Kaufmann (2020), we employ an iterative algorithm to estimate the factors in our analysis. This procedure involves two

[2] more than 300 time series carefully selected by economic judgement with the help of the literature, retrieved from Datastream and automated at CREA for future forecasts

key steps aimed at disentangling the effects of foreign and domestic factors. In the first step, we estimate each foreign factor solely using the relevant foreign data associated with it. Specifically, we apply PCA to the foreign real activity variables and extract the foreign factors, assuming that foreign variables load on their respective foreign factors. In the second step, we estimate the domestic factors by removing the variations induced by all the foreign factors and the target observed variable we aim to forecast (Swiss GDP growth) iteratively.

This decomposition in blocks and factors should also allow doing scenario analysis and counterfactuals (what would have happened without the negative COVID shock on foreign output or on domestic output ?) in future research. Indeed, identifying the transmission channels would allow policy-makers to shape better economic policies. Thus, "structural" model in blocks may serve as a valuable resource for CREA, aiding in effective communication of the Swiss economy's key drivers with the intended audience.

One of the feature of the paper is that we strive to make the estimation of the model as simple as possible so that updating remains relatively easy. In fact, we extract the factors by PCA and we model the FAVAR within the VAR framework by OLS in a two-step principal component approach.

Another feature of the paper is that we reconcile mixed frequencies time series (daily, monthly, quarterly) so that we exploit a maximum number of indicators in order to forecast quarterly GDP growth in a simple way. Daily time series are aggregated and quarterly time series are interpolated using a spline cubic interpolation (Miranda-Agrippino and Rey, 2020). Transforming all the time series in high frequency (monthly frequency, 2000M1-2023M4) will allow us to incorporate more data points enhancing the accuracy of our forecasts and helping capture changes in the underlying economic conditions more effectively in a finer-grained analysis (Litterman et al., 1984; Mitchell et al., 2005). Policy-making requires timely information to make informed decisions. In order to meet this need, policymakers are not just supposed to anticipate initial estimates of GDP growth but also seek to understand the underlying dynamics within each quarter. In order to achieve this, spline cubic interpolation is applied to transform quarterly time series into monthly time series. This approach is inspired by Stuart (2018), who employed cubic interpolation to estimate quarterly series from the first log difference (annual growth rate) of real annual GDP starting from 1960. [3]. The derivation of monthly data from cubic interpolation of quarterly series of our dataset, allows us to improve our forecasting accuracy and facilitates informed policy decisions based on up-to-date information (Stuart, 2018).

Regarding the relative performance of our FAVAR model, we compare its out-of-sample accuracy with six benchmark models (three univariate models: a Random Walk, an AR and ARMA model and three multivariate models: a FAVAR without blocks, a VAR in a closed economy and a VAR in an open economy) on 3 different evaluation samples based on Mean Squared Forecast Errors (MSFE). The first evaluation sample is a sample encompassing a period of relative stability with recovery from the 2008 Great Recession as well as Covid-19 and the recent energy crisis (2010Q1-2023Q2). The second evaluation sample where this out-of-sample exercise is carried out is the period before the Covid-19 crisis (2010Q1-2019Q4) and the third evaluation sample includes only the recent period marked by the Covid-19 crisis and the war in Ukraine (2020Q1-2023Q2).

Our main findings highlight the robust performance of our FAVAR model for Switzerland, particularly in the long run, where it consistently demonstrates comparable or even slightly superior performance compared to all benchmark models. This enhanced performance is evident in the forecast accuracy for the third year ahead, across the first two evaluation samples, although the differences are not always statistically significant.

Another finding is that our FAVAR model consistently outperforms the naive benchmark (random walk) over the majority of horizons, with statistically significant ratios validated through the Diebold-Mariano test (Diebold and Mariano, 1995). This indicates that our model effectively exploits all available information in order to provide accurate forecasts. Furthermore, a noteworthy finding is that our FAVAR model shows statistical equivalence to a FAVAR without blocks in all three samples. In other words, the inclusion of structural blocks in our FAVAR framework does not lead to a deterioration in forecast performance relative to the unblocked version.

However, during the evaluation sample characterized by the Covid-19 Crisis and the war in Ukraine, marked by heightened uncertainty, our FAVAR model is on average surpassed

[3] Stuart (2018) required that “the fourth growth rate in the interpolated series is equal to the growth rate of the annual series” and it is important to note that in our evaluation sample, we will rely solely on non-interpolated data to compare with the corresponding realized values

by most of the benchmark models and its performance in the long run is undermined. The ratios are also on average more volatile over the horizons.

Overall, our FAVAR model showcases strengths in capturing long-run dynamics although its long-term performance may deteriorate during major external shocks such as the Covid Crisis and the war in Ukraine.

Regarding the assessment of the model in absolute standards, tests of forecast unbiasedness and efficiency are carried out by a Mincer and Zarnowitz regression (Mincer and Zarnowitz, 1969) over each horizon. We find that our forecasts are unbiased and efficient over each horizon except horizon 3 (3rd quarter).

The remainder of the paper is constructed as follows. Section 2 explores the existing literature related to our topic and our model. Section 3 explains the variables used and the cleaning procedure. Section 4 exposes the methodology used, the estimation approach and the diagnostic checking. Then, section 5 presents the main results of our forecasting model focusing on the relative performance with respect to the different benchmarks and on the efficiency and unbiasedness tests of our forecasts. Finally, section 6 deals with the discussion and the areas of improvement and section 7 concludes.

2 Literature Review

2.1 The birth of FAVAR

In the wake of Sims (1980) work, considerable literature has emerged around vector autoregressions (VAR) methods, which are originally used to understand and measure the effects of monetary policy innovations on macroeconomic variables (L. J. Christiano et al., 1999). However, as L. Christiano and Eichenbaum (1992) put forward, VAR models face some limitations and provide sometimes results in disagreement with standard economic theory. For instance, they find out that a contractionary monetary policy shock through a rise in the federal funds rate resulted in an increase in price level, rather than a decrease. This finding is referred to as the "price puzzle". In order to alleviate the shortcomings of a classic VAR, Bernanke et al. (2005) propose a factor augmented VAR model which is based on Stock and Watson (2002) methodology to extract the factors from an important dataset. Then, the extracted factors are simply integrated into a standard VAR. In their research, Kaufmann and Lein (2012) use FAVAR in order to address the "price puzzle" for the case of Switzerland. They find no evidence of a "price puzzle" in Switzerland at the aggregate level because FAVAR tends to get rid of misspecification.

In fact, the FAVAR is solving two main problems occurring so far with the traditional VAR. The introduction of factors mitigates the curse of dimensionality (Stock and Watson, 2016) as the number of parameters grows with the square of the number of variables. In fact, a VAR with 7 variables and 3 time-lags would have $3 \times 7^2 = 147$ coefficients ($p \times n^2$) (exploding when open economy) to estimate. Indeed, to conserve degrees of freedom, standard VAR rarely incorporates more than eight variables [4]. All the more, FAVAR can prevent omitted variables bias affecting estimates (Stock and Watson, 2016) because we allow the handling of a large set of information. For instance, Bernanke et al. (2005) use a dataset of 120 monthly macroeconomic variables between January 1950 and August 2001 and then help reduce the "price puzzle". Thus, the consideration of more information with FAVAR avoids measurement of policy innovations to be contaminated without losing the statistical advantages of restricting the analysis to a small number of series (Bernanke et al., 2005).

[4]even if Leeper, Sims and Zha (1996) use more variables but apply Bayesian priors

Secondly, choosing specific data series corresponding to the economic concept of "Real activity" such as real GDP or industrial production is often arbitrary and does not reflect it perfectly. The more information on real activity (employment, sales, exports...) we have, the better we can explain real activity dynamics (Bernanke et al., 2005).

Furthermore, the limitations of traditional VAR become more pronounced when it comes to studying the international transmission mechanism as we switch from a closed economy to an open economy because of the increase in the number of countries and therefore relevant variables (Vasishtha and Maier, 2013).

2.2 Small open economy (SOE) framework

Once the foundations of FAVAR were laid, some research applied this framework to a small open economy. The latter refers to the fact that the country takes part in international trade but is too small enough compared to its trading partners to influence their economic variables (prices, incomes, interest rates, GDP...). Although Boivin et al. (2009) study the case of the US which is not a small open economy, they are the first to extend the FAVAR framework by including international factors and therefore disentangling domestic and foreign blocks. The goal of their paper is to determine if the rest of the world matters in the transmission of US monetary policy to the U.S. economy. In contrast, Vasishtha and Maier (2013) examine the case of Canada as a small-open economy. With a foreign block made up of variables from 20 countries allocated in 4 unobserved factors (commodity prices, foreign interest rates, foreign inflation, foreign economy activity), their FAVAR analyzes how Canada is affected by foreign activity and commodity prices. Similarly to the model we propose, their key assumption is that the foreign underlying variables do not react contemporaneously to changes to Canadian factors. On the other hand, Mumtaz and Surico (2009) use FAVAR to understand the interactions and international linkages between the UK economy and the rest of the world treated as a foreign block. Dealing with a large amount of information in their model allows them to measure the impacts on UK variables of a foreign shock to short-term interest rates and to real activity with better accuracy than a classic VAR model. More recently, Madhou et al. (2020) forecast GDP growth in a small open developing economy: Mauritius, using a FAVAR. The main finding in their study is that their FAVAR produces better forecasts than the benchmark:

a Bayesian vector autoregressive (BVAR) model despite data inadequacy.

2.3 Forecasting with FAVAR and factor models

As Eickmeier and Ziegler (2008) point out, "Forecasting with factor models is a relatively new field of research". Moreover, they show that factor models tend to outperform small-scale models (VARs) and highlight that the size of the dataset from which factors are extracted and the factor estimation technique matter for the performance of the factor models. Bai et al. (2016) estimate FAVAR with maximum likelihood instead of the principal component used in Bernanke et al. (2005) for better accuracy despite the fact that it is computationally burdensome. Bernanke et al. (2005) also consider a Bayesian method but notice its burdensome computation procedure of the Markov chain Monte Carlo method in this context.

Bäurle et al. (2021) use a Bayesian vector autoregressive model (BVAR) and a factor model structure for forecasting the production side of GDP (sum of sectoral real value added series) in the euro Area and in Switzerland. Then, the factor and BVAR model are compared to simple benchmarks regarding their out-of-sample performance. The main finding is that the factor model always beats the benchmark in tests. However, our wish to forecast aggregate Swiss GDP instead of its disaggregate components is justified by existing literature. Although most of this literature is applied to forecasting inflation. For instance, Hubrich (2005) carries out out-of-sample forecasts for Euro area inflation and its five sub-components. His main result is that using disaggregate does not lead to significantly better forecasts. The rationale behind this result is that the component errors become highly correlated and are added up rather than cancelled out. Additionally, more disaggregation implies a higher number of parameters to estimate decreasing, therefore, precision. As a consequence, Hendry and Hubrich (2011) propose forecasting aggregate inflation using directly disaggregate information rather than combining disaggregate forecasts.

Stock and Watson (2002) is applied by Ajevskis and Dāvidsons (2008) to forecast Latvia's GDP, and the outcomes are contrasted with those obtained using benchmark AR models. They demonstrate improved forecasting performance for both factor models, although neither increase is statistically significant. On the other hand, other economic studies, such as that by Rahimov et al. (2020) using a factor-augmented vector autoregressive (FAVAR) model to forecast output in Azerbaijan show that almost all of the multivariate models underperform in comparison with the univariate models such as an ARMA process. According to this study, simple benchmarks can be better predictors.

3 Data

Every data series is collected from Datastream [5] and automated by request at CEDIF [6] for future use and updating. The static data series of each country used can be found in the Excel file "nameofthecountry.xslm" in the replication materials. In total, there are 320 variables (152 quarterly, 154 monthly, 14 daily). We invite the reader to look at the data appendix, which lists all the time series used.

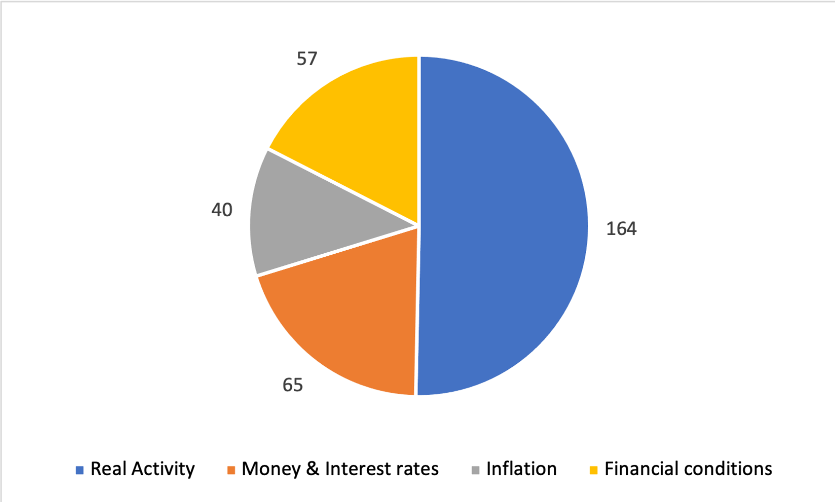
The dataset consists of 155 domestic variables from Switzerland and 172 foreign variables that capture foreign economic activity which is relevant for a small-open economy like Switzerland. The foreign variables encompass variables from the main trading partners in Europe, such as Germany, France, Italy and the United Kingdom, as well as other main trading partners such as the United States, China and Japan.

Then, our dataset contains 4 types of variables as in Mumtaz and Surico (2009) : real activity data including real GDP, gross value added, industrial production, investment, consumption, exports, imports, employment, vacancies, consumer confidence index, retail sales; inflation data including the main price indices (GDP Deflator, CPI, core inflation, PPI, oil prices); money & interest rates data including money supply M1,M2,M3, banknotes in circulation, sight deposits, saving deposits and the different interest rates (policy rates, discount rates, government bond yields, mortgage lending rates. . .) and finally financial conditions variables including house prices, stock prices, uncertainty indexes and effective exchange rates. Swiss variables are more granular, with more precise data by industry and sector, in order to better capture accurately the domestic dynamics. Figure 3.1 presents an overview of how the variables are distributed based on their respective types.

[5]Datastream gathers data series from various sources such as central banks, statistical offices from concerned countries, OCDE, IMF. . .

[6]Centre de documentation et d'initiation financière de l'Université de Lausanne

Figure 3.1: Data set shares by type of variable



The full dataset spans 1980Q1-2023Q2. Despite having enough observations, the choice of the date of start is not trivial as it marks the new era (shift) of the US monetary policy regime, also known as the "Volcker shock" and the beginning of the Great Moderation. However, some variables start after the first quarter of 1980, creating an unbalanced dataset that we will later solve using the PCA EM-algorithm.

Most of the data series, especially the Swiss variables, are not already seasonally adjusted at the source in Datastream. We therefore perform this adjustment where necessary using the X-13ARIMA-SEATS (seasonal) procedure developed by the United States Census Bureau. Where applicable, this is noted in the series description in the data appendix. Naturally, no seasonal pattern in the financial markets is assumed. Consequently, no seasonal adjustment is performed on variables such stock indices. In addition, the series are outlier-adjusted and calendar adjusted.

Then, each series is transformed to be approximately integrated of order 0 by taking the log-difference [7], i.e in growth rates, except for interest rates and yields. Indeed, the latter must have some interpretation, e.g an interest rate growth rate makes no sense and is difficult to interpret. Therefore, we only take the difference for interest rates, i.e percentage point changes (Stock and Watson, 2002). [8]

To check the stationarity of the transformed time series, we apply an augmented Dickey

[7] To ensure consistency regardless of their frequencies, all variables have been transformed into quarter-on-quarter growth rates for quarterly variables and 3-month growth rates for monthly variables .

[8] 3-month differences for monthly variables

Fuller test with a 10% rejection of the null hypothesis H0 [9].Once a first difference of the level/logarithmic-series applied , a few series were not judged stationary with respect to the unit-root test. In these cases, a second difference of the level/logarithmic-series is applied (full details of these transformations are available in the appendix).

The target variable to forecast is approximated using Swiss real GDP from SECO in level already adjusted for seasonal, calendar and sport event effects. In the paper, the target variable Y_t depicts the quarter-on-quarter GDP growth rate corresponding to the change from the previous quarter. It is as defined as :

$$Y_t = 100 \times \log\left(\frac{Y_t^l}{Y_{t-1}^l}\right) \tag{3.1}$$

with Y_t^l the real GDP from SECO in level at quarter t and Y_{t-1}^l at quarter $t-1$. It is a quarterly variable.

Following Stock and Watson (2002) procedure prior to factor analysis, we implement a final preliminary adjustment which consists in standardizing the series to have unit standard deviation (zero mean and unit variance).[10]This transformation will ensure that factor loadings remain unaffected by variance or unit of measurement.

Lastly, before extracting the factors, we apply PCA-EM algorithm to each dataset (quarterly, monthly and daily dataset). The goal of this algorithm is to obtain a balanced dataset - not all variables start from the first quarter of 1980, some of them later (see details in appendix). In fact, Stock and Watson (2002) proposed the EM-algorithm as an imputation method to fill in the missing values. We apply the same algorithm, with a slight difference in the first step by adding a cross-validation procedure as in Josse and Husson (2016) (see details in appendix), also used in Burri and Kaufmann (2020).

Note also that many of the series are only available at quarterly, monthly or daily frequency (mixed frequency series). Having 3 different clean data frames, one at quarterly frequency, one at monthly and one at daily frequency, the aim is to merge all these 3 different dataframes into one with the same frequency in order to have more tractability for our model. We decide to convert all the series to monthly values in order to have high-frequency

[9]Null hypothesis H0 : the time series is not stationary
[10]Note : the target variable Y_t is not scaled because it will be forecast

data. For temporal aggregation for daily data, we take the monthly average as the daily variables are all stock variables. Thus, in our specification, monthly variables $x^m_{i,t}$ are defined as the average of daily observations $x^d_{i,t}$ on the number of days N of each month :

$$x^m_{i,t} = \frac{1}{N}\sum_{t=1}^{N} x^d_{i,t}$$

While quarterly level data are converted into monthly frequency using a shape-preserving piecewise cubic interpolation as in Miranda-Agrippino and Rey (2020)[11].

[11] using spline() command in R

4 Methodology

The econometric framework used in this paper relies on FAVAR Models as proposed by Bernanke et al. (2005) extended to include international factors in the spirit of Boivin et al. (2009), Mumtaz and Surico (2009) and Vasishtha and Maier (2013). Indeed, the model includes two blocks of factors; one for Switzerland (named domestic block) and one for the rest of world (named foreign block), extracted respectively from large sets of Swiss and international time series. The latter refer to economic variables from the main trading partners of Switzerland and having the weight to impact Swiss economy.

4.1 Econometric framework: FAVAR

The FAVAR model was created by Bernanke et al. (2005) as an adaptation of a structural dynamic factor model (Stock and Watson, 2002). The FAVAR depicts the dynamics of a small number of estimated factors, which summarize the common information of a large set of variables, together with the target variable (the Swiss GDP growth rate), within a vector-autoregressive (VAR) model. The FAVAR still has a limited number of variables, but the huge underlying informative set reduces the omitted variables bias (see literature review section).

The observable indicators X_t (with dimension $N \times 1$ and N large) containing information about the economy are related to a set of unobserved (or latent) factors F_t (with dimension $K \times 1$ and K is small) and observed variable Y_t (the target) according to the following measurement equation:

$$X_t = \Lambda^F F_t + \Lambda^Y Y_t + v_t \tag{4.1}$$

where Λ^F and Λ^Y are respectively $N \times K$ (factor loadings) and $N \times 1$ matrices and v_t is a $N \times 1$ vector of zero mean disturbances with a diagonal covariance matrix Σ.

Once F_t extracted, its dynamics as well as Y_t dynamics evolve in a standard VAR model according to the following equation :

$$\begin{pmatrix} Y_t \\ F_t \end{pmatrix} = \Phi(L) \begin{pmatrix} Y_{t-1} \\ F_{t-1} \end{pmatrix} + u_t \tag{4.2}$$

where $\Phi(L)$ is a lag polynomial of finite order p and u_t is an error term with mean zero and a diagonal covariance matrix Ω. Notice that we dropped the constant/trend term for notational convenience.

4.2 A small open economy (SOE) FAVAR

Then, we extend the FAVAR to include two different blocks (one domestic H and one foreign $*$) in the spirit of Boivin et al. (2009). In order to take into consideration the small open economy nature of Switzerland, and finding inspiration in Kamber et al. (2016), we propose the following specification of the measurement equation :

$$\begin{pmatrix} X_t^H \\ X_t^* \end{pmatrix} = \begin{pmatrix} \Lambda_{11}^H & \Lambda_{12}^* \\ 0 & \Lambda_{22}^* \end{pmatrix} \begin{pmatrix} F_t^H \\ F_t^* \end{pmatrix} + \begin{pmatrix} \Lambda_1^Y \\ 0 \end{pmatrix} Y_t + \begin{pmatrix} v_t^H \\ v_t^* \end{pmatrix} \tag{4.3}$$

where domestic and foreign observational time series: $X_t = [X_t^H, X_t^*]$ and domestic and foreign unobserved factors: $F_t = [F_t^H, F_t^*]$. $v_t \sim \mathcal{N}(0, \Sigma)$ is a vector of reduced form errors. As you notice in the measurement equation, we allow the domestic observed variables to load on both domestic and foreign factors as well as target variable Y_t, whereas foreign variables only load on foreign factors. Following Stock and Watson (2002), the factors from equation 4.3 can be retrieved by principal component analysis provided that the informational variables X_t are standardized to have zero mean and unit standard deviation (details in data section).

Furthermore, we also exploit the fact that Switzerland is a small open-economy in the transition equation (restricted VAR model) where foreign factors have an impact on domestic factors and the Swiss target variable but are not governed by them. Following the transition equation in 4.2 :

$$\begin{pmatrix} Y_t \\ F_t^H \\ F_t^* \end{pmatrix} = \begin{pmatrix} \phi_{11}(L) & \phi_{12}(L) & \phi_{13}(L) \\ \phi_{21}(L) & \phi_{22}(L) & \phi_{23}(L) \\ 0 & 0 & \phi_{33}(L) \end{pmatrix} \begin{pmatrix} Y_{t-1} \\ F_{t-1}^H \\ F_{t-1}^* \end{pmatrix} + \begin{pmatrix} u_t^Y \\ u_t^H \\ u_t^* \end{pmatrix} \tag{4.4}$$

with $u_t \sim \mathcal{N}(0, \Omega)$ a vector of reduced form errors.

4.3 Identification of factors with economic interpretation

Following Mumtaz and Surico (2009), we consider four different factors for foreign variables : $F_t^* = [F_{Y,t}^*, F_{\pi,t}^*, F_{MI,t}^*, F_{FC,t}^*]$ where $F_{Y,t}^*$ represents a world [12] real activity factor, $F_{\pi,t}^*$ a world inflation factor, $F_{MI,t}^*$ a world money supply and interest rates factor and $F_{FC,t}^*$ a world financial conditions factor. They are each loaded by their respective observable series $X_t^* = [X_{Y,t}^*, X_{\pi,t}^*, X_{MI,t}^*, X_{FC,t}^*]$ where $X_{Y,t}^*$ represents world observable real activity series, $X_{\pi,t}^*$ world observable inflation series, $X_{MI,t}^*$ world observable money supply and interest rates series and $X_{FC,t}^*$ world observable financial conditions series.

For instance, foreign real activity factor is extracted from all international real activity series (non-Swiss time series). Similarly, the foreign inflation factor is identified as the leading factor that is loaded by all non-Swiss inflation series. The other international factors are identified accordingly. These foreign factors are identified accordingly through Λ_{22}^* (Equation 4.3) which is block diagonal ($N \times 4$ dimension). Thus, foreign factors are extracted with the following equation :

$$\begin{pmatrix} X_{Y,t}^* \\ X_{\pi,t}^* \\ X_{MI,t}^* \\ X_{FC,t}^* \end{pmatrix} = \begin{pmatrix} \Lambda_Y^* & 0 & 0 & 0 \\ 0 & \Lambda_\pi^* & 0 & 0 \\ 0 & 0 & \Lambda_{MI}^* & 0 \\ 0 & 0 & 0 & \Lambda_{FC}^* \end{pmatrix} \begin{pmatrix} F_{Y,t}^* \\ F_{\pi,t}^* \\ F_{MI,t}^* \\ F_{FC,t}^* \end{pmatrix} + \begin{pmatrix} v_{Y,t}^* \\ v_{\pi,t}^* \\ v_{MI,t}^* \\ v_{FC,t}^* \end{pmatrix} \tag{4.5}$$

with $v_t^* \sim \mathcal{N}(0, \Sigma)$ a vector of reduced form errors.

Until now nothing new relatively to the existing literature. The main contribution of

[12] world, foreign and international are interchangeably used

this paper in terms of methodology lies in the domestic block. Contrary to Mumtaz and Surico (2009) who extract the domestic factors [13] from full panel of domestic variables without any structure, we apply in our framework the same structure of foreign block for the domestic block. The rationale behind is that it will allow us to identify properly the structural shocks on each domestic factor and carry out a scenario analysis [14] after forecasting the Swiss GDP growth with our model. Thus, we still consider four different factors for domestic variables : $F_t^H = [F_{Y,t}^H, F_{\pi,t}^H, F_{MI,t}^H, F_{FC,t}^H]$ where $F_{Y,t}^H$ represents a Swiss real activity factor, $F_{\pi,t}^H$ a Swiss inflation factor, $F_{MI,t}^H$ a Swiss money supply and interest rates factor and $F_{FC,t}^H$ a Swiss financial conditions factor. They are each loaded by their respective observable series $X_t^H = [X_{Y,t}^H, X_{\pi,t}^H, X_{MI,t}^H, X_{FC,t}^H]$ where $X_{Y,t}^H$ represents Swiss observable real activity series, $X_{\pi,t}^H$ Swiss observable inflation series, $X_{MI,t}^H$ Swiss observable money supply and interest rates series and $X_{FC,t}^H$ Swiss observable financial conditions series which load Y_t as well. These domestic factors are identified accordingly through Λ_{11}^H (Equation 4.3) which is block diagonal ($N \times 4$ dimension). Note also that the observable domestic series load the foreign factors identified through Λ_{12}^F (Equation 4.3) and is of $N \times 4$ dimension [15]. Thus, domestic factors are extracted with the following equation : with $v_t^H \sim \mathcal{N}(0, \Sigma)$ a vector of reduced form errors.

$$
\begin{aligned}
\begin{pmatrix} X_{Y,t}^H \\ X_{\pi,t}^H \\ X_{MI,t}^H \\ X_{FC,t}^H \end{pmatrix} &= \begin{pmatrix} \Lambda_Y^H & 0 & 0 & 0 \\ 0 & \Lambda_\pi^H & 0 & 0 \\ 0 & 0 & \Lambda_{MI}^H & 0 \\ 0 & 0 & 0 & \Lambda_{FC}^H \end{pmatrix} \begin{pmatrix} F_{Y,t}^H \\ F_{\pi,t}^H \\ F_{MI,t}^H \\ F_{FC,t}^H \end{pmatrix} \\
&+ \begin{pmatrix} \Lambda_{Y,11}^* & \Lambda_{\pi,12}^* & \Lambda_{MI,13}^* & \Lambda_{FC,14}^* \\ \Lambda_{Y,21}^* & \Lambda_{\pi,22}^* & \Lambda_{MI,23}^* & \Lambda_{FC,24}^* \\ \Lambda_{Y,31}^* & \Lambda_{\pi,32}^* & \Lambda_{MI,33}^* & \Lambda_{FC,34}^* \\ \Lambda_{Y,41}^* & \Lambda_{\pi,42}^* & \Lambda_{MI,43}^* & \Lambda_{FC,44}^* \end{pmatrix} \begin{pmatrix} F_{Y,t}^* \\ F_{\pi,t}^* \\ F_{MI,t}^* \\ F_{FC,t}^* \end{pmatrix} \\
&+ \begin{pmatrix} \Lambda_{Y,t}^Y \\ \Lambda_{\pi,t}^Y \\ \Lambda_{MI,t}^Y \\ \Lambda_{FC,t}^Y \end{pmatrix} Y_{it} + \begin{pmatrix} v_{Y,t}^H \\ v_{\pi,t}^H \\ v_{MI,t}^H \\ v_{FC,t}^H \end{pmatrix}
\end{aligned} \tag{4.6}
$$

[13] refering to UK variables as the domestic variables
[14] forthcoming in a next paper
[15] but not block diagonal

4.4 Estimation

As in Stock and Watson (2002), Boivin et al. (2009), Mumtaz and Surico (2009), we estimate our model using a two-step principal component analysis (PCA) approach. The first step consists in extracting principal components from X_t^* and X_t^H and then obtain consistent estimates of the common factors (foreign and domestic factors). In the second step, these estimated factors are included in our standard restricted VAR with our target variable (Swiss GDP growth) Y_t.

Note that, in the first step, the leading principal component is extracted for each of the panel sets of $X_t^* : [X_{Y,t}^*, X_{\pi,t}^*, X_{MI,t}^*, X_{FC,t}^*]$. [16] Therefore, we obtain estimates of the four foreign factors : $[F_{Y,t}^*, F_{\pi,t}^*, F_{MI,t}^*, F_{FC,t}^*]$. Then, we put the restriction that foreign factors are included into the PC for domestic block of the model as well as target variable Y_t. Therefore, if these foreign factors and the target variable are in fact common components, the PC of X_t^H should be able to capture them. Starting from the estimates of factors (F_t^*) of the foreign block of variables X_t^*, iteration proceed through following steps :

- For each panel set of domestic block (4 panel sets in total), we regress X_t^H on F_t^* (estimates of the four foreign factors) and Y_t by OLS and therefore get $\hat{\Lambda}_{12}^0$ and $\hat{\Lambda}^Y$.
- Then, we compute residuals $e_t^0 = X_t^H - \hat{\Lambda}_{12}^0 F_1^* - \hat{\Lambda}_Y Y_t$.
- We extract the first principal component from e_t^0 by PCA in order to get $\hat{F}_t^{H,0}$ and $\hat{\Lambda}_{11}^0$.
- We regress X_t^H on $\hat{F}_t^{H,0}$, F_t^*, and Y_t by OLS and therefore get $\hat{\Lambda}_{11}^1$, $\hat{\Lambda}_{12}^1$, $\hat{\Lambda}_1^Y$.
- We compute residuals $e_t^1 = X_t^H - \hat{\Lambda}_{12}^1 F_t^* - \hat{\Lambda}_Y Y_t$.
- We extract the first principal component from e_t^1 by PCA in order to get $\hat{F}_t^{H,1}$ and $\hat{\Lambda}_{11}^1$.
- If $\hat{\Lambda}_{11}^1 - \hat{\Lambda}_{11}^0 < \epsilon$ we stop the iteration, otherwise we go back to step 3.

with $\hat{\Lambda}_{11}$ representing domestic factor loadings of a panel set, $\hat{\Lambda}_{12}$ representing foreign factor loadings of all the four world panel sets, Y_t denoting the target variable and $\hat{\Lambda}^Y$ indicating loadings associated with the target variable Y_t.

[16] The % variance explained by the first principal component of each panel sets is displayed in the appendix

Thus, four domestic factors are estimated $F_t^H = [F_{Y;t}^H, F_{\pi;t}^H, F_{MI;t}^H, F_{FC;t}^H]$. Therefore, we end up with 8 factors (4 foreign factors and 4 domestic factors). We display a graph of them next pages (Figures 4.1,4.2). Figure 4.1 plots the factors of foreign financial conditions, foreign inflation, foreign money and interest rates and foreign real activity. A few patterns are apparent.

Since the year 2000, the industrialized world has experienced on average 3 major recessions. These economic downturns include the slowdown associated with the bursting of the dot-com bubble between 2001 and 2002, the recession during the global financial crisis (2007-2009), and the combined impact of the COVID-19 crisis and the subsequent energy crisis (2020-2023). The foreign real activity factor on the bottom right panel shows these episodes of recession. The recent uptick in the measurement of international inflation depicted in the upper right corner is in line with the pattern of elevated inflation experienced by major industrialized nations over this period (2021-). As expected, the foreign financial conditions factors proves to be notably more volatile than other factors, exhibiting pronounced spikes during the height of the global financial crisis in 2008. In the lower left quadrant, the foreign money supply and interest rates factors is less interpretable. Money supply and interest rates are typically negatively correlated. This means that when the money supply increases, interest rates tend to decrease, and when the money supply decreases, interest rates tend to increase. However, it is worth noting that this factor saw a significant surge in 2022, which is consistent with the recent upward trajectory of interest rates observed in most developed countries. Moreover, the aftermath of the 2008 global financial crisis is evident here too, as there was a substantial decrease coinciding with the central banks' decision to bring interest rates down to the zero lower bound.

Figure 4.2 plots the factors of domestic financial conditions, domestic inflation, domestic money and interest rates and domestic real activity. We construct these factors such that we decontaminate them from the influences of the international business cycle. Thus, they are rendered less interpretable graphically.

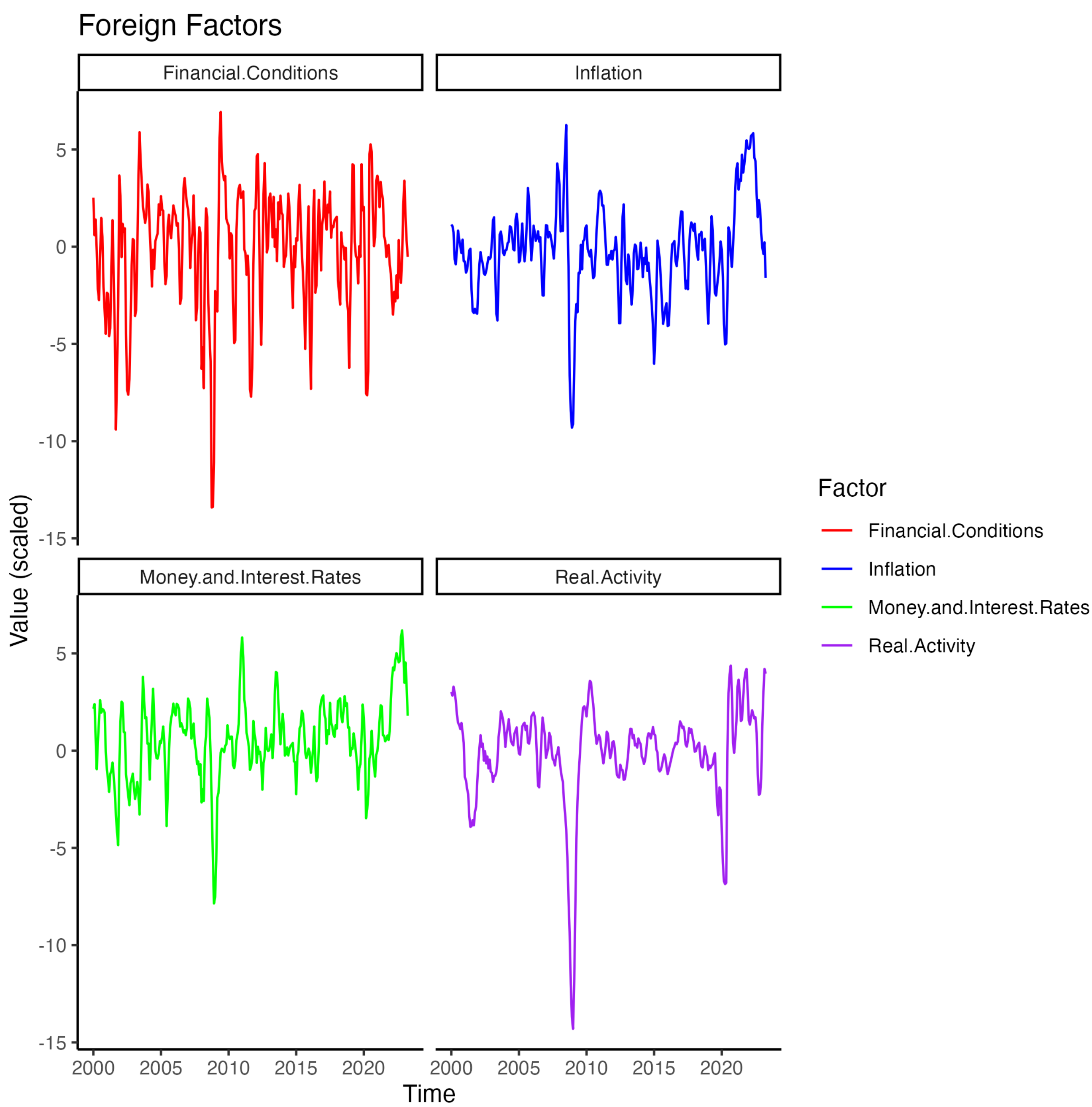


Figure 4.1: Foreign factors

Note: Before extracting the factors all time-series are seasonal, outliers (winzorisation at 95th percentile during Covid period) and calendar adjusted as well as expressed in growth rates. All time-series are also standardized for comparison purposes. The estimation sample starts on the 1st January 2000. Note also that time series are transformed into monthly frequency if in quarterly frequency (cubic interpolation) or in daily frequency (temporal aggregation). In addition, missing values are processed by PCA EM-algorithm by frequency (for more details, see the data section).

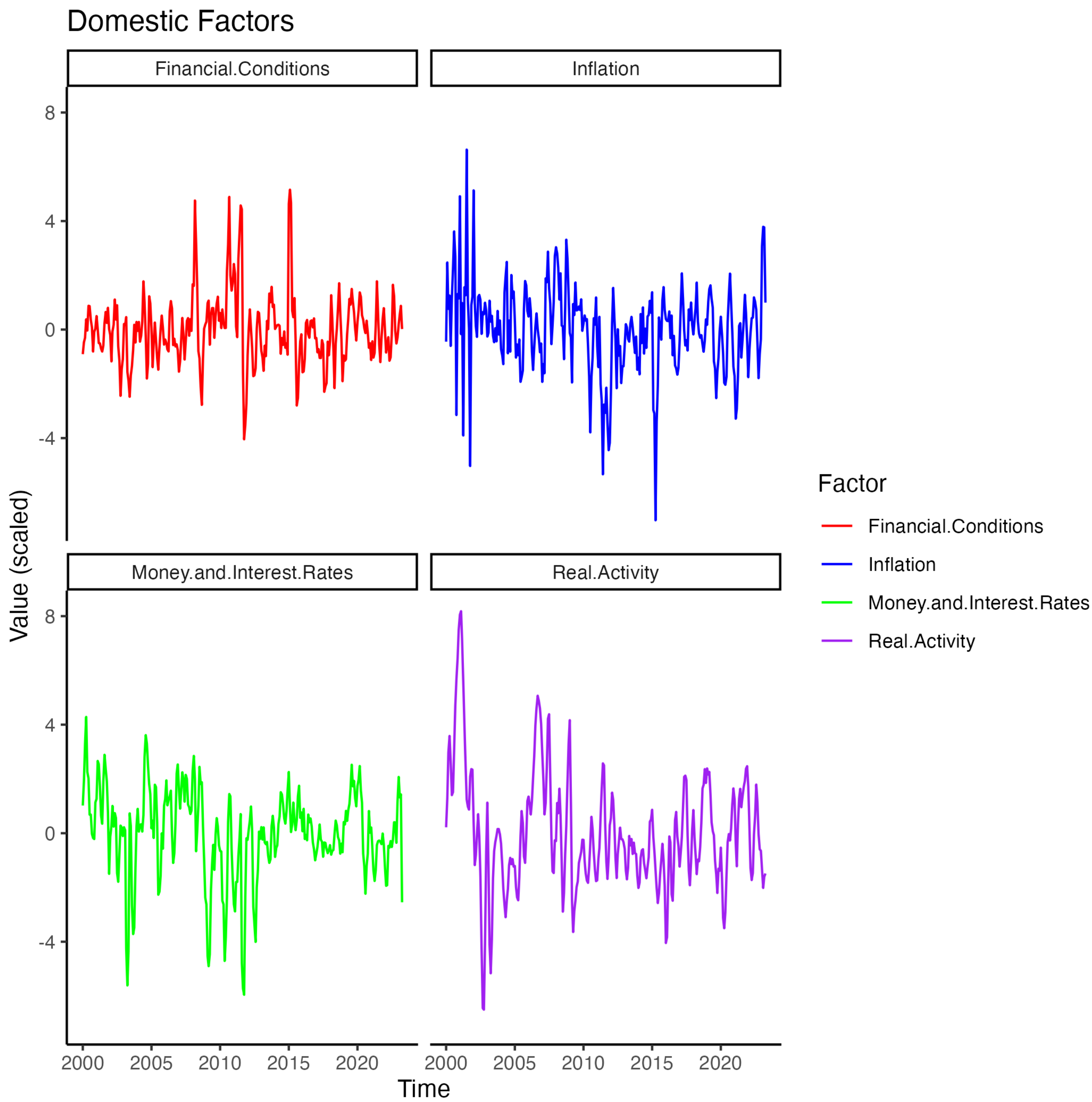


Figure 4.2: Domestic factors

Note: Before extracting the factors all time-series are seasonal, outliers (winzorisation at 95th percentile during Covid period) and calendar adjusted as well as expressed in growth rates. All time-series are also standardized for comparison purposes. The estimation sample starts on the 1st January 2000. Note also that time series are transformed into monthly frequency if in quarterly frequency (cubic interpolation) or in daily frequency (temporal aggregation). In addition, missing values are processed by PCA EM-algorithm by frequency (for more details, see the data section).

Finally, in the second step we estimate our restricted VAR (Equation 4) with 9 variables namely our 4 foreign factors (exogeneous), 4 domestic factors and our target variable Y_t. In order to choose the lag number we rely on four information criteria [17] :

$$\begin{aligned}
\mathrm{AIC}(n) &= \ln\det(\tilde{\Sigma}_u(n)) + \frac{T^2 n K^2}{2} \\
\mathrm{HQ}(n) &= \ln\det(\tilde{\Sigma}_u(n)) + \frac{2\ln(\ln(T))}{T} n K^2 \\
\mathrm{SC}(n) &= \ln\det(\tilde{\Sigma}_u(n)) + \frac{\ln(T)}{T} n K^2 \\
\mathrm{FPE}(n) &= \left(\frac{T+n^*}{T-n^*}\right)^K \det(\tilde{\Sigma}_u(n))
\end{aligned}$$

where $\tilde{\Sigma}_u(n) = T^{-1}\sum_{t=1}^{T} \hat{\boldsymbol{u}}_t\hat{\boldsymbol{u}}_t'$, n^* is the total number of the parameters in each equation, n assigns the lag order, K the number of endogeneous variables and T the number of observations. As the most parsimonious, $\mathrm{SC}(n)$ indicate us that two lags are enough to capture adequately the dynamics of the restricted VAR. Information criteria table is provided in the appendix. Then, we can go for an estimation of parameters via OLS for computational simplicity. Otherwise, we could have used Bayesian methods as in Mumtaz and Surico (2009) to circumvent the dimensionality issue.

4.5 Diagnostic checking

From a forecasting point of view, some robustness checks for our FAVAR are important to ensure consistent estimates and accurate forecasts. First of all, parameter stability is particularly important and procedures should be done for detecting eventual structural breaks. Such past events like Global Financial Crisis or the Covid-19 pandemic may have given rise to a number of structural breaks. That is why it is important to identify them. Therefore, a generalised fluctuation test based on the cumulative sum of the recursive residuals of the model (Brown et al., 1975) may be interesting to implement for identifiying the structural breaks in a graphical way. [18] This CUSUM test suggests that there is parameter instability if the cumulative sum of the recursive residuals breaks one of the two boundaries in red in the graph below (Du Plessis and Kotze, 2012). The test is performed for each equation of our FAVAR model (figure 4.3). Furthermore, after noticing multiple

[17] from vars package in R
[18] stability command in the package vars

breaks in the first sample (1980Q1-2023Q2) we decide to modify the estimation sample and to start in 2000Q1. [19] Our FAVAR is modified accordingly.

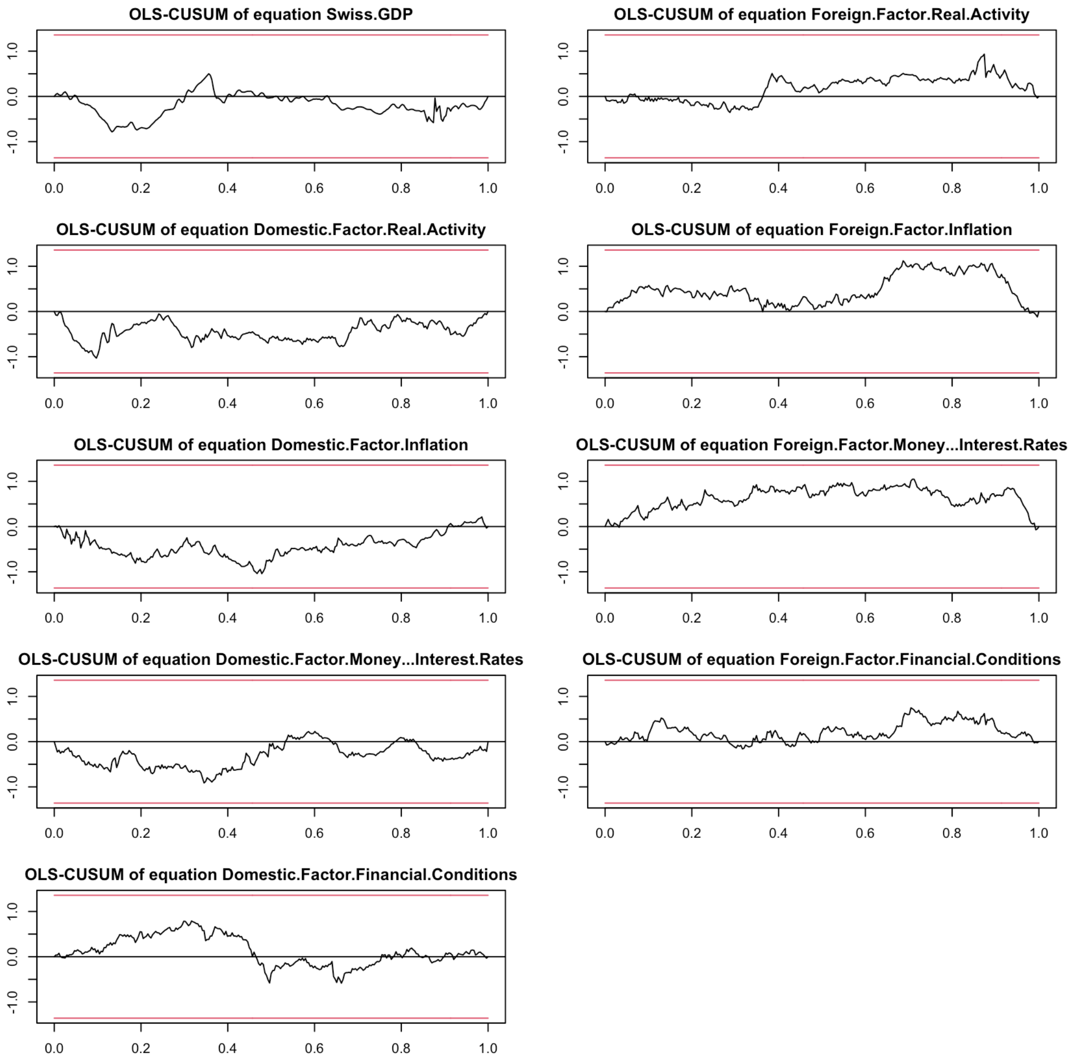


Figure 4.3: OLS-CUSUM test (sample : 2000Q1-2023Q2)

Then, covariance stationarity is an important property for representation, estimation and inference. And for a VAR(p) process, stationarity is verified when all the inverse roots z of $det(I - \phi_1 - ... - \phi_p z^p) = 0$ lie inside the unit circle. In our case with our FAVAR(2), the inverse roots of $det(I - \phi_1 - \phi_2 z^2)$=0 lie inside the unit circle (figure 4.4). Therefore, our FAVAR is covariance stationary.

[19] It can be due to the large number of missing data between 1980Q1 and 2000Q1 treated by the PCA-EM algorithm

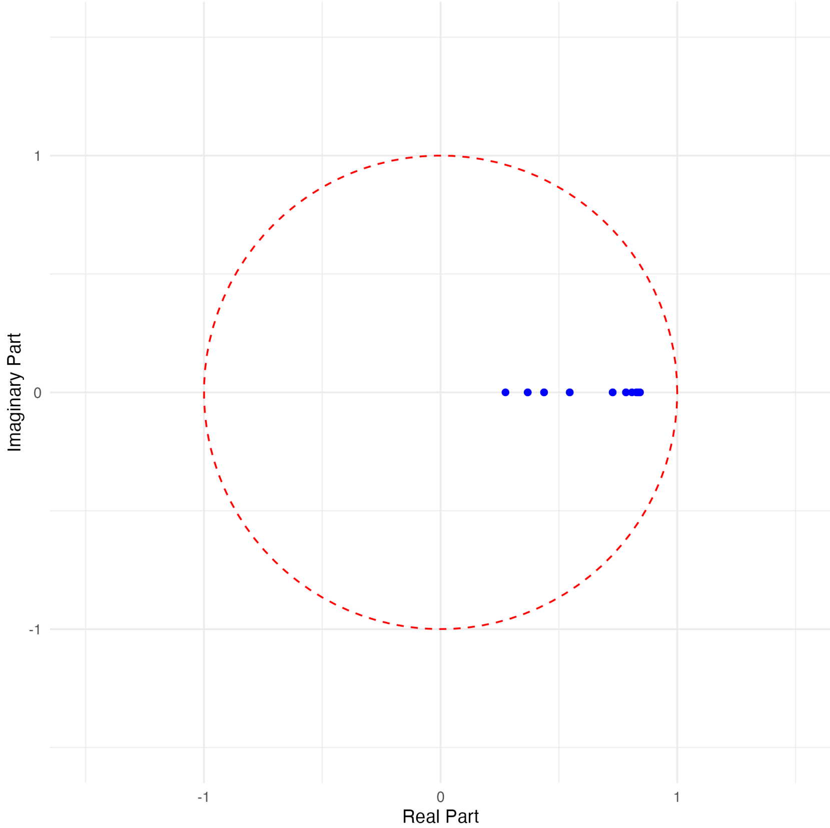


Figure 4.4: Inverse roots of VAR characheristic polynominal

Last but not least, we conduct a Johansen Test (Johansen, 1991) for testing cointegration among the 9 variables within the FAVAR using the trace test version. We reject the null hypothesis H_0 that the rank of cointegration is at least 8 ($r \leq 8$) at 5% level. Therefore, the number of cointegrating vectors r equal to n (the number of variables), i.e 9. Thus, Π has full rank, the model is stable and there is no need to use a Vector Error Correction model (VECM). The table related to the Johansen Test is displayed in the appendix.

Finally, so far, we have assumed that the errors are (multivariate) white noise, namely, uncorrelated and homoskedastic. We test for these properties. To test for serial correlation we can apply a Portmanteau-test on the residuals of the model. We reject the null hypothesis of no serial correlation of errors at 1% level. Indeed, the p-value of the Portmanteau-test is smaller than 1% (table 4.1 below).

Table 4.1: Portmanteau Test (asymptotic) on residuals

Statistic	Chi-Squared	Degrees of freedom	p-value
Portmanteau Test	872.1	729	0.0001968

To assess the distribution of the residuals, we could apply a normality test especially a Jarque Bera-Test as well as a test focusing only on the skewness of the distribution of residuals and another one focusing on the Kurtosis. All the three tests (tables 4.2,4.3, 4.4) reject the null hypothesis of normality of residuals meaning that the residuals are not fairly normally distributed.

Table 4.2: Normality Test on Residuals (Multivariate) - Jarque Bera Test

Statistic	Chi-Squared	Degrees of freedom	p-value
Jarque Bera-Test	11976	18	$< 2.2 \times 10^{-16}$

Table 4.3: Normality Test on Residuals (Multivariate) - Skewness

Statistic	Chi-Squared	Degrees of freedom	p-value
Jarque Bera-Test	254.97	9	$< 2.2 \times 10^{-16}$

Table 4.4: Normality Test on Residuals (Multivariate) - Kurtosis

Statistic	Chi-Squared	Degrees of freedom	p-value
Jarque Bera-Test	11721	9	$< 2.2 \times 10^{-16}$

In order to conduct a more in-depth analysis, we have plotted the distribution of residuals for each regression in the FAVAR model, as well as examined the autocorrelation function (ACF) and partial autocorrelation function (PACF) of these residuals. Upon careful examination, we have observed significant autocorrelations among the residuals, particularly when considering the ACF. Although most of the lags are within the confidence interval (Ljung-Box Test), there are some that go slightly beyond the limit. Notably, at the 3th lag for most of the regressions. In future research, we may have to incorporate the autocorrelation of residuals within a FAVARMA (Factor Augmented VAR with a Moving Average Part) framework like in Dufour and Stevanović (2013). Their work suggest that using FAVARMA may yield enhanced forecast accuracy compared to the conventional FAVAR model. In the appendix, we have included the plots of residuals and their corresponding distributions for each regression. Furthermore, the appendix displays the PACF and ACF plots of the residuals.

5 Results

In this section, we assess the forecasting performance of our proposed FAVAR model for predicting Swiss real activity [20].

5.1 Unconditional forecast

As explained before 8 factors are extracted (4 domestic and 4 foreign factors) and then they are modeled together with the target variable Y_t (Swiss GDP growth rate) in a VAR model with the foreign factors built and set as exogenous. We use iterated h-step ahead forecasts as in Banerjee et al. (2014) in contrast to the use of direct h-step-ahead (Stock and Watson, 2002). In fact, iterated forecasts produce more accurate forecasts except in presence of substantial misspecification (Marcellino et al., 2006). Furthermore, note that our forecasting model allows for a feedback from the target variable Y_t to the factors (only domestic) and for a direct effect of past values of the target variable on its past dynamics (Eickmeier et al., 2011). Thus, transition equation (4.2) can be rewritten as:

$$\begin{pmatrix} Y_{t+h} \\ F_{t+h} \end{pmatrix} = \Phi(L)^h \begin{pmatrix} Y_t \\ F_t \end{pmatrix} + \sum_{j=0}^{h-1} \Phi(L)^j u_{t+h-j} \tag{5.1}$$

Unconditional forecasts as in Eickmeier et al. (2015) are :

$$\mathbb{E} \begin{pmatrix} Y_{t+h} \\ F_{t+h} \end{pmatrix} = \Phi(L)^h \begin{pmatrix} Y_t \\ F_t \end{pmatrix} \tag{5.2}$$

since $\mathbb{E}(u_{t+h}) = 0$ for $h > 0$.

However, note that the FAVAR lag length is set to 2. Thus, unconditional forecasts (i.e the optimal h-step ahead forecasts) can be rewritten as :

$$\mathbb{E} \begin{pmatrix} Y_{t+h} \\ F_{t+h} \end{pmatrix} = \sum_{j=1}^{2} \Phi_j \begin{pmatrix} Y^*_{t-j+h|t} \\ F^*_{t-j+h|t} \end{pmatrix} \tag{5.3}$$

[20] Swiss GDP growth rate

where $Y^*_{t-j+h|t} = Y_{t-j+h}$ and $F^*_{t-j+h|t} = F_{t-j+h}$ for $h \leq j$ [21]

Using equation (5.3), we can forecast the Swiss GDP growth rate in Switzerland for any horizon from 1 (one quarter ahead) to 12 (twelve quarters ahead). Figure 5.1 depicts the prediction made in April 2023 for a horizon of three years. Moreover, a 95% confidence interval is drawn to assess the uncertainty.

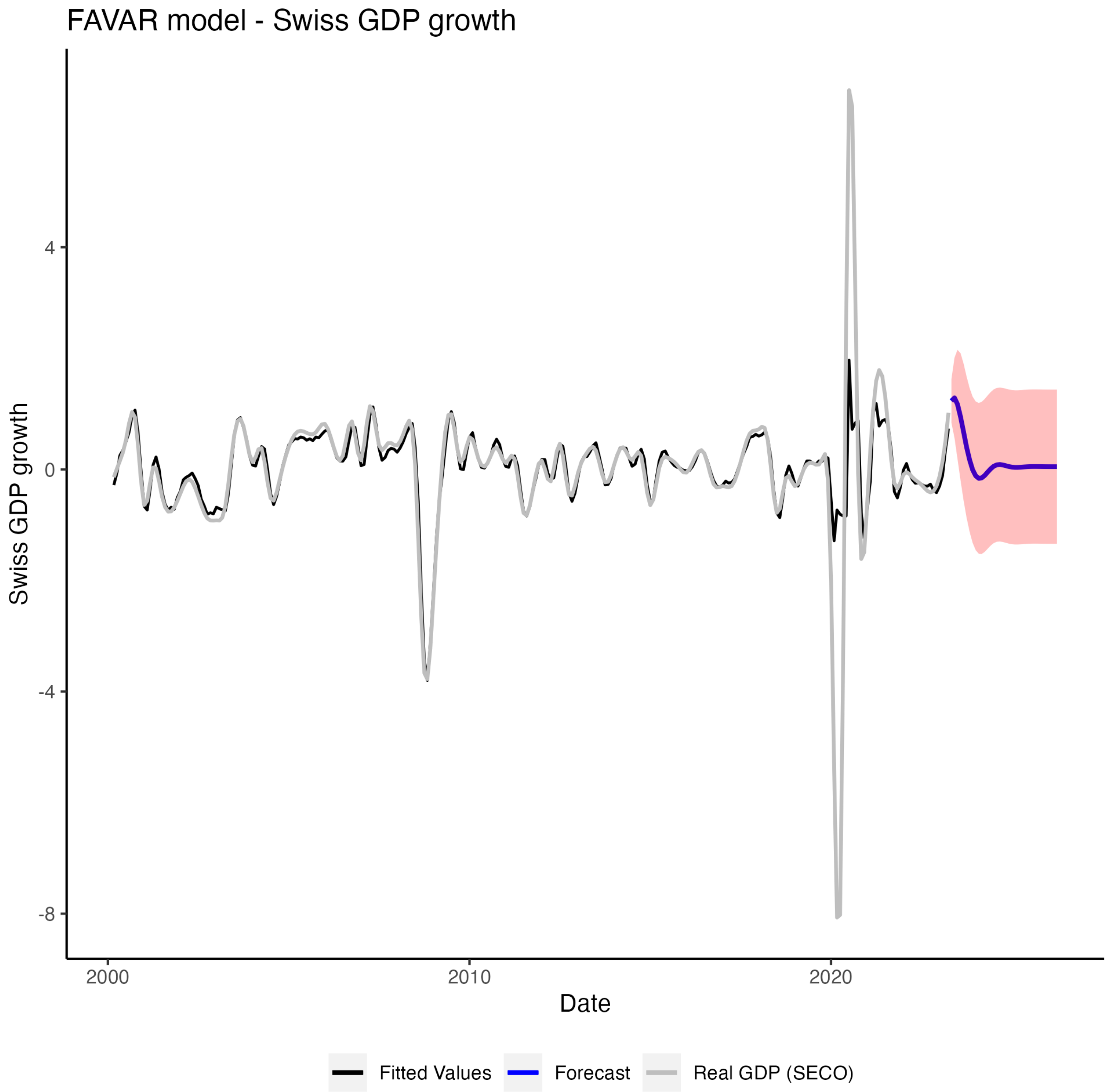


Figure 5.1: Real Swiss GDP growth forecast (2023Q3-2026Q2), QoQ growth rates

Note: This figure displays the Swiss GDP growth rate (QoQ) from 2000Q1 to 2023Q2 (in grey) with the fitted values of our FAVAR model (in black) and the forecast from 2023Q3 to 2026Q2 (in blue). The

[21] * in front of Y_t and F_t refer to an optimal forecast and not to the realized value

shaded area in red represents the 95% confidence interval. In addition, for greater representativeness the values are descaled. Note also that the underlying observed data are winsorized "ex-ante" at the 95th percentile during the Covid period.

5.2 Pseudo relative out-of-sample evaluation

We conduct a pseudo out-of-sample forecast evaluation exercise where the models' accuracy in predicting Swiss GDP growth is assessed recursively. Indeed, in order to assess the predictive accuracy of our FAVAR model regarding Swiss GDP growth, out-of-sample forecasts are produced each quarter for the period between 2010Q1 and 2023Q2 (evaluation sample). In order to mitigate the influence of Covid-related outliers on our results, we implement a winsorization technique on the underlying data ("ex-ante"). Specifically, we apply winsorization at the 95th percentile during all the Covid period. This approach helps ensure that extreme values influenced by the pandemic do not affect our forecasts and is considered as a robust method in this type of context (Jose and Winkler, 2008).

As in Diebold (2017), we denote T as the length of the whole sample with t ranging from 1 to T (from 2000Q1 to 2023Q2), R (2010Q1) denotes the start of our evaluation sample (where $R < T$) and h the horizon in quarter (from $h = 1$ to $h = 12$). The first step consists in estimating our sample (from 1 to R). At time R, forecast is calculated for horizon 1 to 12, using equation (5.3). Then, at time $R + 1$, new information allows the model to be updated and re-estimated and forecasts for horizons one to 12 are calculated using the revised model. Using new data available at $R + t$, the model is updated and reestimated till the end of the sample minus the horizon In practice, as we use monthly frequency data, we produce forecasts for Swiss GDP growth rate at the three, six, nine and twelve months horizon $(t + h)$, then we continue by updating the estimation sample by adding one quarter (i.e three months). With this recursive technique, the sample size grows over time. It is considered as the best technique in absence of structural breaks in the sample (see Stock and Watson, 2007). Additionally, the size of the training data sample (or estimation sample) from 2000Q1 and 2010Q1 is sufficient to produce stable estimation results. Note that factors are estimated over the full sample and not reestimated at each period primarily due to computational considerations.

The figure 5.2 represents the sequence of out-of-sample forecasts 12 quarters ahead based on this recursive scheme :

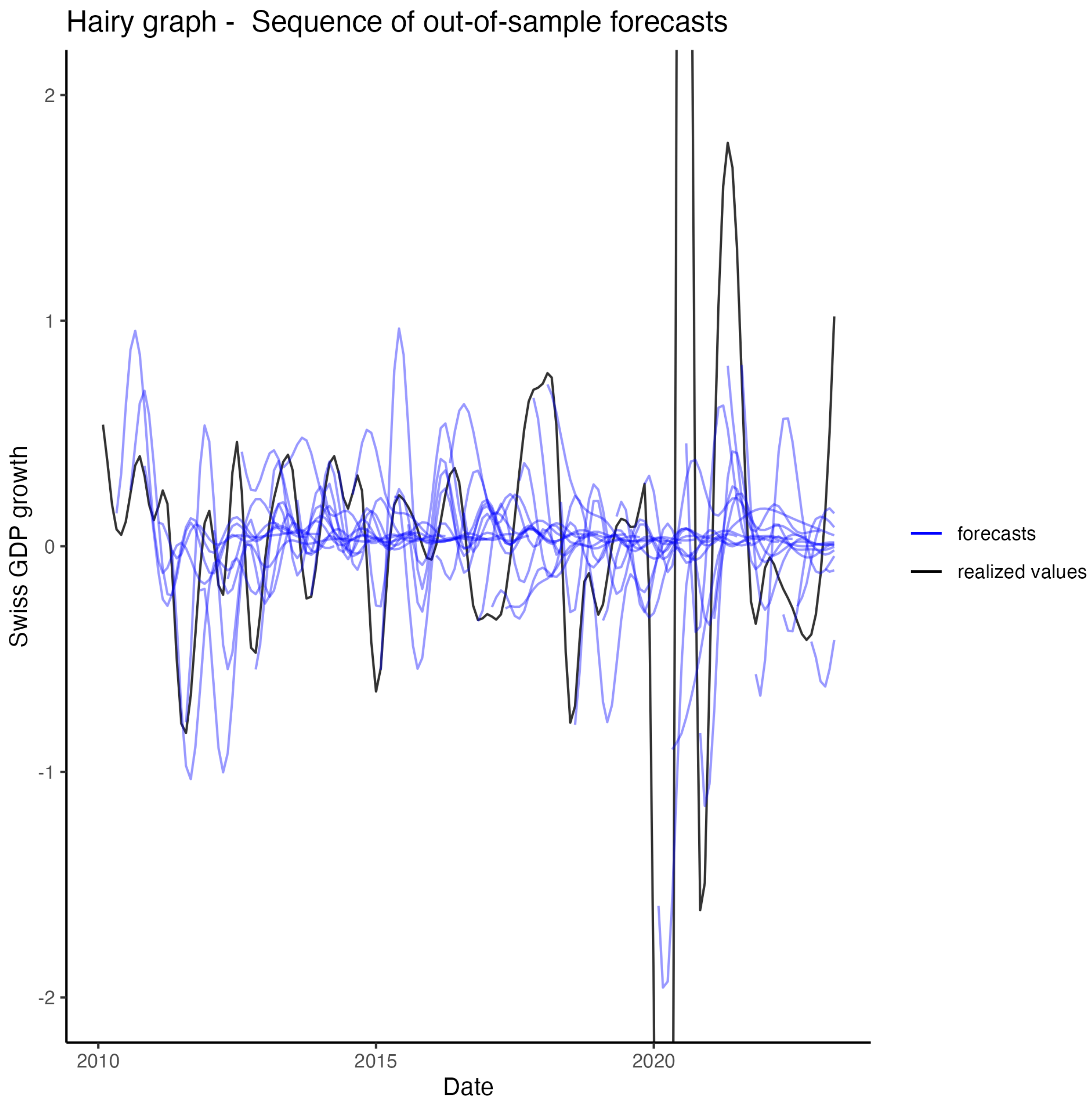


Figure 5.2: Sequence of out-of-sample forecasts (3-year ahead) for Swiss GDP growth based on a recursive estimation scheme (2010Q1-2023Q2), QoQ growth rates - FAVAR model

Thus, in order to assess quantitatively the predictive accuracy of the model regarding Swiss GDP growth, one can compute its mean squared forecast errors (MSFE) and compare them with those of six different benchmark models. For the univariate benchmarks, there are an auto-regressive process of order three : AR(3), a naive random walk : RW and an autoregressive moving average of order (3,4) : ARMA(3,4) [22]. We also employ several multivariate benchmarks for comparative analysis, including a non-block FAVAR(2) and

[22]optimal p, q are computed by information criteria (AIC and BIC)

two VAR models, one for a closed economy and another for an open economy.
The non-block FAVAR(2) comprises the target variable Y_t alongside 5 factors extracted through principal component analysis from our entire dataset, without distinguishing between blocks or factor types.

For the VAR in a closed economy (Switzerland), the vector of variables $\mathbf{Y}_t = [\ Y_t, SWIPI_t, SWCPI_t, SARON_t, SWIMI_t]$ consists of the target variable Y_t, which is the Swiss GDP growth rate, and four leading indicators: $SWIPI_t$ (Swiss industrial production index), $SWCPI_t$ (Swiss consumer price index), $SARON_t$ (Swiss average rate overnight) and $SWIMI_t$ (Swiss market index). These indicators represent respectively factors related to domestic real activity, domestic inflation, domestic money supply and interest rates, and domestic financial conditions.

As for the VAR in a closed economy, VAR in open economy includes the same variables but is augmented with an exogenous vector $\mathbf{Y}_t^+$ of US variables. Since the US is Switzerland's main trading partner, the $\mathbf{Y}_t^+$ vector includes key US variables: $USIPI_t$ (US industrial price index), $USCPI_t$ (US consumer price index), $USFEDRATE_t$ (US federal funds rate), and $S\&P500_t$ (Standard & Poor's 500, a US stock market index). These US variables represent respectively factors associated with foreign real activity, foreign inflation, foreign money supply and interest rates, and foreign financial conditions. The selection of these variables was informed by their significance as leading indicators, as identified in prior research by Miranda-Agrippino and Rey (2020).

Using all the past notation, squared forecast errors can be computed from $R + 1$ for each horizon in order to evaluate the accuracy of the model, namely:

$$e_{t,h}^2 = (y_{t+h} - \widehat{y}_{t+h|t})^2 \tag{5.4}$$

where y_{t+h} is the realized value of target variable and $\widehat{y}_{t+h|t}$ the forecast of the target variable made in time t for horizon h.

Then, Mean Squared Forecast Errors (MSFE) are computed as follow :

$$MSFE_h = \frac{1}{P} \sum_{t=R}^{T-h} e_{t,h}^2 \tag{5.5}$$

where $P = T - h - R + 1$.

Therefore, in order to compute the performance of our model relatively to the six benchmarks, we compute the ratio of the relative MSFE, which is defined as :

$$\frac{MSFE_h}{MSFE_{benchmark,h}} \tag{5.6}$$

Table 5.1 and 5.2 display the ratio of MSFE for each horizon (from $h = 1$ to $h = 12$, h is a quarter) and relatively to each of the six benchmarks (against univariate and multivariate benchmarks). We assess the performance of our model in comparison to the benchmark using the Mean Squared Forecast Error (MSFE) ratio. When the MSFE ratio is smaller than one, it indicates that our model outperforms the benchmark. Conversely, when the ratio is larger than one, the benchmark performs better than our model. A ratio equal to one signifies that both models perform equally well. To visually represent this comparison, we shade the MSFE ratios below 1 and those equal to one in green, meaning that our model is at least as good as the benchmark in these instances.

A Diebold-Mariano test (Diebold and Mariano, 1995) is also performed for each horizon h in order to determine whether the differences of MSFE between the model and the benchmark are statistically significant. The null hypothesis of equal statistical predictive accuracy is therefore tested. The DM-test statistic is defined as follows :

$$DM = \sqrt{P}\left(\frac{\overline{d_h}}{\widehat{\sigma_d}}\right) \tag{5.7}$$

where $\overline{d_h} = \frac{1}{P}\sum_{t=R}^{T-h} L(e_{1;t,h}) - L(e_{2;t,h})$ with $L(e_{i;t,h}) = (e_{i;t,h})^2$, i.e the squared forecasting error of model $i = 1$ (our FAVAR) and model $i = 2$ (the benchmark). P denotes the length of the evaluation sample, $\widehat{\sigma_d^2}$ is the Heteroskedasticity and Autocorrelation Consistent (HAC) estimator of the long-run variance estimated with the Bartlett kernel and using Newey and West (1986) estimator in order to take into account the serial correlation of the forecast errors. Note that we implement the modified DM-test proposed by Harvey et al. (1997). Indeed, they suggest a DM -test statistic corrected for small-sample (HLN-DM = $\sqrt{\frac{T+1-2h+h(h-1)}{T}DM}$) and compared to a Student-t distribution with $T-1$ degrees of freedom.[23]

In the out-of-sample period from 2010Q1 to 2023Q2, the FAVAR model demonstrates a modest advantage of approximately 1% over two prominent univariate benchmarks, namely the AR(3) and ARMA(3,4) models for the horizon periods of 9 to 11 (long run). Note that our FAVAR model consistently outperforms the random Walk model across all horizons, with statistical significance observed at the 1% level in the majority of cases.

However, it is worth noting that in the short-run, particularly during horizon 1 and 3 (one quarter and three quarter ahead), the AR(3) and ARMA(3,4) models exhibit a clear, though not statistically significant, competitive performance against the FAVAR model. Despite this short-term phenomenon, the FAVAR model's forecast accuracy remains commendable and difficult to outperform when compared to these simple univariate benchmarks.

Upon a global assessment over multiple horizons, the ratios of forecast mean squared prediction errors (MSFE) between the FAVAR model and the three univariate benchmarks remain relatively close to 1. This observation implies a strong forecasting ability of the FAVAR model when put in competition against these univariate models, which are difficult to beat across the literature.

[23]performed with forecast library and dm.test() command

Table 5.1: Out-of-Sample Assessment: Ratio of Mean Squared Forecast Errors - Univariate benchmarks

Horizon	AR(3)	RW	ARMA(3,4)
1	1.06	0.86	0.93
2	0.98	0.84*	1.03
3	1.04	0.98*	1.06
4	0.96**	0.91***	0.97
5	1.00	0.93***	1.00
6	1.01	0.92**	1.00
7	1.03	1.03***	1.00
8	1.00	0.82***	1.00
9	0.99*	0.89***	0.99
10	0.99*	0.93***	0.99
11	0.99*	0.98***	0.99
12	1.00	0.97	1.00

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: This table displays the ratios of the MSFE of the model FAVAR relative to three univariate benchmarks: An auto-regressive process of order 3, a random walk, and an auto-regressive moving average process of order 3,4 (the optimal p,q are computed by information criteria over the sample). 12 quarterly forecast horizons are considered. The models are estimated on the estimation sample that increases at every period. The models are re-estimated every period, as the information set increases (recursive out-of-sample methodology). The estimation sample is from 2000Q1 to 2010Q1 and the out-of-sample from 2010Q1 to 2023Q2 (full out-of-sample).

In our comparison of the FAVAR model's performance with that of various multivariate benchmarks, namely $FAVAR_{without\ blocks}$, $VAR_{closed\ economy}$, and $VAR_{open\ economy}$, we have observed strikingly similar forecast accuracy. The ratios of mean squared forecast errors (MSFE) for the FAVAR model relative to these multivariate benchmarks consistently are consistently nearly equal to 1 across the vast majority of forecast horizons.

Notably, our FAVAR model exhibits superior performance compared to both the VAR in closed economy and the VAR in open economy, with statistical significance observed at the 5% level during horizons 4 and 6.

Moreover, over the long run, specifically during horizons 9 to 12, our FAVAR model demonstrates an impressive level of accuracy, with the ratio of MSFE being very close to 1 and mostly at 0.99. This finding further showcases the FAVAR model's accuracy and reliability.

Table 5.2: Out-of-Sample Assessment: Ratio of Mean Squared Forecast Errors - Multivariate Benchmarks

Horizon	FAVAR_without blocks	VAR_closed economy	VAR_open economy
1	1.00	1.01	0.98
2	1.03	1.01	1.08
3	1.00	1.02	0.97
4	1.00	0.98**	0.97**
5	1.02	0.99	0.99
6	1.01	0.97**	0.97*
7	0.99	1.05	1.01
8	0.98	1.03	1.01
9	0.99	1.00	1.00
10	1.00	0.99	0.99
11	1.00	1.00	0.99*
12	0.99	0.99	0.99

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: This table displays the ratios of the MSFE of the model FAVAR relative to three univariate benchmarks: A FAVAR without blocks of lag 2, a VAR in closed economy of lag 2 and a VAR in open economy of lag 2 (lag 2 are settled for comparative purposes). 12 quarterly forecast horizons are considered. The models are estimated on the estimation sample that increases at every period. The models are re-estimated every period, as the information set increases (recursive out-of-sample methodology). The estimation sample is from 2000Q1 to 2010Q1 and the out-of-sample from 2010Q1 to 2023Q2 (full out-of-sample).

On the whole, our out-of-sample assessment from 2010Q1 to 2023Q2 reveals that the FAVAR model exhibits good forecasting abilities, particularly in the long run (horizon 9 to 12), performing relatively well against various benchmarks. Notably, it outperforms the naive random walk benchmark across all horizons and maintains accuracy similar to a FAVAR model without blocks, indicating that incorporating blocks does not lead to forecast accuracy loss. However, in the short run, results are more nuanced and depend on the specific forecast horizon.

5.3 Assessing out-of-sample performance Pre-Covid

Fisher et al. (2002) and Stock and Watson (2007) show how the forecast predictive power of a forecasting model depends highly on the evaluated period. That is why we assess the out-of-sample performance over a sub-sample marked by a period of a relative stability without Great Recession, the Covid crisis and the ongoing war in Ukraine (2010Q1-2019Q4). In this period, the FAVAR model demonstrates a slight advantage over the three univariate benchmarks (AR, RW, ARMA) starting from horizon 5. Notably, the

FAVAR model consistently outperforms the RW model across all horizons. As we focus on the last two horizons ($h = 11$, $h = 12$), the FAVAR model exhibits superior performance compared to ARMA(3,4) model with the mean squared forecast errors (MSFE) ratios being significantly lower than 1 at the 1% significance level. The comparison of the MSFE ratios for the FAVAR model versus the three univariate benchmarks is presented below (Table 5.3).

Table 5.3: Out-of-Sample Assessment: Ratio of Mean Squared Forecast Errors - Univariate benchmarks - Covid Excluded

Horizon	AR(3)	RW	ARMA(3,4)
1	1.07	0.67***	1.10
2	1.10	0.69**	1.17
3	1.00	0.56*	1.05
4	0.98	0.61***	1.00
5	0.98	0.69	0.97
6	0.93	0.73	0.93
7	0.98	0.59**	0.99
8	0.96	0.55**	0.97
9	0.98*	0.64**	0.98
10	0.99	0.62**	0.99
11	0.98	0.66**	0.99***
12	0.98	0.71***	0.98***

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: This table displays the ratios of the MSFE of the model FAVAR relative to three univariate benchmarks: An auto-regressive process of order 3, a random walk, and an auto-regressive moving average process of order 3,4 (the optimal p,q are computed by information criteria over the sample). 12 quarterly forecast horizons are considered. The models are estimated on the estimation sample that increases at every period. The models are re-estimated every period, as the information set increases (recursive out-of-sample methodology). The estimation sample is from 2000Q1 to 2010Q1 and the out-of-sample from 2010Q1 to 2019Q4 (Covid period excluded).

Upon comparing the performance of our FAVAR model with that of the multivariate benchmarks (namely, $FAVAR_{without\ blocks}$, $VAR_{closed\ economy}$, and $VAR_{open\ economy}$), the FAVAR model demonstrates pretty good forecast accuracy. The ratios of mean squared forecast errors (MSFE) for the FAVAR model relative to the multivariate benchmarks are consistently smaller than 1 in the vast majority of forecast horizons. This indicates that the FAVAR model outperforms the $VAR_{closed\ economy}$ and $VAR_{open\ economy}$ models, except for horizon 2,3,7 (short-run forecasts) against the $VAR_{closed\ economy}$, and horizon 4 against the $VAR_{open\ economy}$. Furthermore, the performance of our FAVAR model remains quite similar to the FAVAR model without blocks across most horizons. However, note its strong improvement over the the first three horizons. This is a favorable result, as it shows that the inclusion of blocks in our FAVAR model does not lead to a significant loss in forecast accuracy. Remarkably, at the final horizon $h = 12$, the FAVAR model achieves a slight edge over all three multivariate benchmarks, indicating its robustness and predictive power even in the long run (outperforming even $FAVAR_{without\ blocks}$ and $VAR_{open\ economy}$ at 1% level). The Table 5.4 below presents the detailed ratios of MSFE for our FAVAR model compared to the three multivariate benchmarks.

Table 5.4: Out-of-Sample Assessment: Ratio of Mean Squared Forecast Errors - Multivariate Benchmarks - Covid Excluded

Horizon	FAVAR_without blocks	VAR_closed economy	VAR_open economy
1	0.92	0.77	0.72**
2	0.90***	1.11	0.90
3	0.83**	1.08	0.81**
4	1.07	0.99	1.03
5	1.04	0.98	0.94
6	0.96	0.97	0.92
7	1.01	1.02	0.95
8	0.99	0.98	0.98
9	1.00	0.99	0.97
10	0.99	0.99	1.00
11	1.00	0.99	0.97***
12	0.99***	0.98*	0.96***

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: This table displays the ratios of the MSFE of the model FAVAR relative to three univariate benchmarks: A FAVAR without blocks of lag 2, a VAR in closed economy of lag 2 and a VAR in open economy of lag 2 (lag 2 are settled for comparative purposes). 12 quarterly forecast horizons are considered. The models are estimated on the estimation sample that increases at every period. The models are re-estimated every period, as the information set increases (recursive out-of-sample methodology). The estimation sample is from 2000Q1 to 2010Q1 and the out-of-sample from 2010Q1 to 2019Q4 (Covid period excluded).

All things considered that our FAVAR model exhibits commendable performance compared to the six benchmarks (univariate and multivariate) during this out-of-sample period. Most notably, our FAVAR model's robustness becomes evident in the long-run forecasts (horizons 8 to 12) where it achieves relatively stable ratios lower than one in majority of cases.

5.4 Assessing out-of-sample performance in the Covid period and after

Now, we examine the out-of-sample performance of our FAVAR model during the Covid-19 and its following challenging period (2020Q1-2023Q2) in order to see if we find consistent results with the full evaluation sample (2010Q1-2023Q2) and the subsample (2010Q1-2019Q4). The results reveal a more nuanced picture. Although our FAVAR model exhibits substantial superiority over the naive benchmark (RW) across various horizons with a statistical significance validated at the 1% level in the majority of cases - It is important to highlight that the FAVAR model's performance falls short compared to the Autoregressive (AR) process and the Autoregressive Moving Average (ARMA) model in the long-term and, more notably, during the last two horizons. This contrasts with its performance before the Covid period (2010Q1-2019Q4), where it achieved superiority over the vast majority of horizons and in long term. Furthermore, it is important to recognize the increased volatility in the performance ratios across horizons in this evaluation sample. Table 5.5 displays the ratios of MSFE relatively to the univariate benchmarks in this evaluation sample.

Table 5.5: Out-of-Sample Assessment: Ratio of Mean Squared Forecast Errors - Univariate benchmarks - Covid Only

Horizon	AR(3)	RW	ARMA(3,4)
1	1.05	0.88***	0.91
2	1.23	0.90	1.38
3	0.79	0.55**	0.85
4	0.93	0.53***	0.84
5	0.89***	0.33***	1.00
6	1.26	0.30***	1.14
7	1.08	0.37***	0.87
8	1.03	0.42***	1.11
9	1.00	0.24***	0.78***
10	0.99	0.39*	0.94***
11	1.03	1.22***	1.65
12	1.05	0.24	1.06

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: This table displays the ratios of the MSFE of the model FAVAR relative to three univariate benchmarks: An auto-regressive process of order 3, a random walk, and an auto-regressive moving average process of order 3,4 (the optimal p,q are computed by information criteria over the sample). 12 quarterly forecast horizons are considered. The models are estimated on the estimation sample that increases at every period. The models are re-estimated every period, as the information set increases (recursive out-of-sample methodology). The estimation sample is from 2000Q1 to 2019Q4 and the out-of-sample from 2020Q1 to 2023Q2 (Covid period and after).

When comparing our FAVAR model to multivariate benchmarks, we find that FAVAR without blocks consistently outperforms our "structural" FAVAR over the majority of horizons (9 out of 12). Although the differences in performance are not statistically significant, they are noticeable in the short term with relatively high ratios (e.g., 1.12 for horizon 2, 1.20 for horizon 4 and 1.22 for horizon 6). Now, when we compare our FAVAR against other multivariate benchmarks such as VAR in closed economy and VAR in open economy, our FAVAR model demonstrates better performance. It manages to outperform these benchmarks, especially in the earlier horizons. However, in the long run, particularly in the last horizons, our model is outperformed by them. Table 5.6 displays the ratios of MSFE relatively to the multivariate benchmarks in this evaluation sample.

Table 5.6: Out-of-Sample Assessment: Ratio of Mean Squared Forecast Errors - Multivariate Benchmarks - Covid Only

Horizon	FAVAR_without blocks	VAR_closed economy	VAR_open economy
1	1.00	1.05	1.00
2	1.12	1.21	1.29
3	1.02	0.59	0.72
4	1.20	0.84***	0.79***
5	1.09	0.90**	0.85**
6	1.22	0.75***	0.81***
7	0.88	1.02	0.88
8	1.01*	0.93***	0.96***
9	1.05**	0.87	0.97
10	0.99	0.94	1.02
11	1.03	1.05	1.00
12	0.96	1.03	0.95

* p < 0.1, ** p < 0.05, *** p < 0.01.

Note: This table displays the ratios of the MSFE of the model FAVAR relative to three univariate benchmarks: A FAVAR without blocks of lag 2, a VAR in closed economy of lag 2 and a VAR in open economy of lag 2 (lag 2 are settled for comparative purposes). 12 quarterly forecast horizons are considered. The models are estimated on the estimation sample that increases at every period. The models are re-estimated every period, as the information set increases (recursive out-of-sample methodology). The estimation sample is from 2000Q1 to 2019Q4 and the out-of-sample from 2020Q1 to 2023Q2 (Covid period and after).

All things considered that it becomes apparent that our FAVAR model exhibits certain weaknesses during this Covid period, especially when compared to other benchmarks. Surprisingly, these weaknesses are more pronounced in the long-run, contrary to the findings we previously obtained for the other two evaluation samples.

To gauge the extent of its underperformance in this particular sample, we do an average of the Mean Squared Forecast Errors (MSFE) ratios over all horizons between our FAVAR and each other benchmarks for the two distinct periods: before the Covid outbreak (2010Q1-2019Q4) and after (2020Q1-2023Q2). The findings are presented in Table 5.7.

What emerges from these results is that the average of the MSFE ratios were higher in the Post Covid period compared to the Pre-Covid period for 4 out of 6 benchmarks. Surprinsingly, our FAVAR's average performance even surpassed the value of 1, indicating it performed worse in average than an AR process, an ARMA process, and a FAVAR without blocks, which was not the case during the pre-Covid period.

Table 5.7: Out-of-Sample Assessment: Ratio of Mean Squared Forecast Errors - Average over all horizons (from 1 to 12)

Horizon	AR	RW	ARMA	$\text{FAVAR}_{without\ blocks}$	$\text{VAR}_{closed\ economy}$	$\text{VAR}_{open\ economy}$
Pre-Covid	0.99	0.64	1.01	0.97	0.99	0.93
Post-Covid	1.03	0.53	1.04	1.05	0.93	0.94

Indeed, the results depicted in figure 5.3 clearly illustrate that our FAVAR model exhibits significantly higher squared forecasting errors during the Covid period (and even after) compared to the Pre Covid period. This observation holds true for the various horizons considered, namely horizon 1 (short run), horizon 6 (medium run) and horizon 12 (long run).

Additionally, it is worth noting that the disparity in forecasting squared errors between the short run (horizon 1) and the long run (horizon 12) across the out-of-sample is relatively minor. This suggests a consistent predictive performance of our FAVAR model across different forecasting horizons. This robustness underscores the reliability of our model's in predicting Swiss GDP in long run.

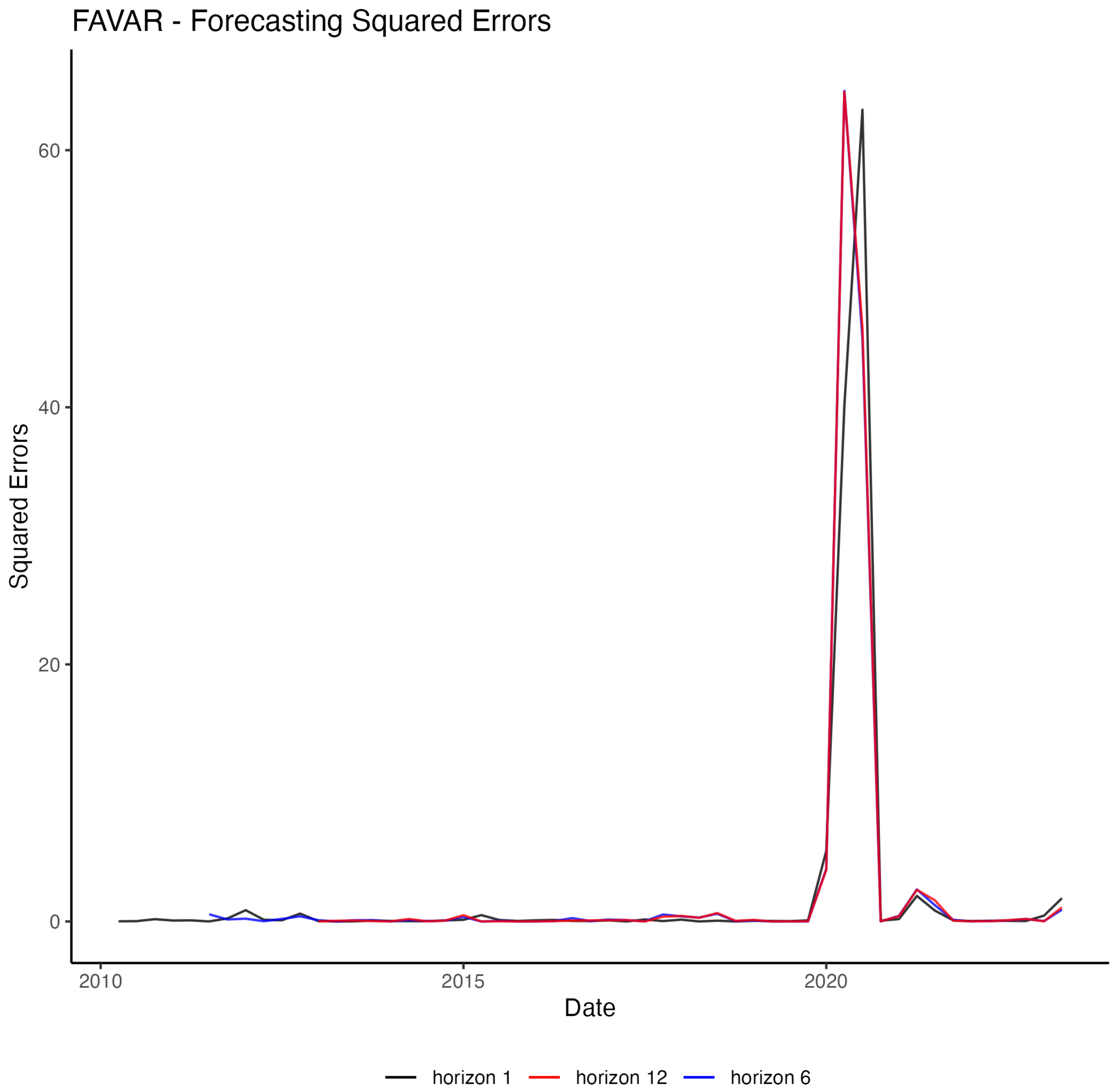


Figure 5.3: Out-of-Sample Assessment - Forecasting Squared Errors

5.5 Test of forecast unbiasedness and efficiency

Based also on a recursive technique on all the evaluation sample (2010Q1-2023Q2) we investigate in this section whether the out-of-sample forecasts of our FAVAR are efficient for all the horizons. In order to check whether forecasts are unbiased and (strongly) efficient, it is common to test by the mean of a robust version of the $F-test$ the null hypothesis H_0 : $\alpha = 0$, $\beta = 1$ in the following Mincer-Zarnowitz (MZ) regression [24] (Mincer and Zarnowitz, 1969) :

$$y_{t+h} = \alpha + \beta y_{t+h|t} + \varepsilon_{t+h} \tag{5.8}$$

where $t = R, ..., T-H$ ($R = 2010Q1$) and autocorrelation of ε_{t+h} is taken into account in the HAC variance estimation.

We rewrite MZ regression as follows :

$$e_{t+h|t} = \tau + \varepsilon_{t+h} \tag{5.9}$$

where $\tau = \alpha + (\beta - 1)y_{t+h|t}$

Thus, series of $\left\{e_{t+h|t}\right\}_{t=R}^{T-h}$ are computed for each horizon (from $h = 1$ to $h = 12$), i.e for each quarter. Then, the null hypothesis H_0: $\tau = 0$ is tested for each horizon with a robust version of the t-test.

Table 5.8 below shows that the forecasts are strongly efficient for each horizon except horizon 3 (the third quarter). Indeed, the null hypothesis H_0: $\tau = 0$ is rejected at 5% for horizon 3 but not for other horizons at 10% level. For longer horizons, the MZ p-value is even very high confirming that forecast errors are uncorrelated with forecasts and hence unforecastable using their information sets. Note that for computing the standard errors we consider HAC variance-covariance estimator proposed by Newey and West (1986).

[24] y_{t+h} represents the realized value and $y_{t+h|t}$ the predicted value

Table 5.8: Test of forecast unbiasedness and efficiency (Mincer-Zarnowitz regression)

Horizon	$\widehat{\tau}$	MZ p-value
1	-0.13	0.17
	(0.10)	
2	-0.20	0.10
	(0.12)	
3	-0.23	0.04**
	(0.11)	
4	-0.20	0.08
	(0.11)	
5	-0.14	0.20
	(0.11)	
6	-0.15	0.20
	(0.11)	
7	-0.09	0.39
	(0.11)	
8	-0.09	0.37
	(0.10)	
9	-0.08	0.49
	(0.12)	
10	-0.09	0.46
	(0.12)	
11	-0.10	0.46
	(0.13)	
12	-0.11	0.41
	(0.13)	

Note: This table reports for each horizon h the p-values of the Mincer and Zarnowitz (1969) forecast efficiency test ("labeled as "MZ p-value"), the estimate of $\widehat{\tau}$ in this following regression $e_{t+h|t} = \tau + \varepsilon_{t+h}$ together with in parentheses below of each estimate the Heteroskedasticity and Autocorrelation consistent (HAC) standard errors. We perform the implementation of the HAC variance-covariance estimator proposed by Newey and West (1986). On R, the library "sandwich" as well as the function NeweyWest are used to compute these standard errors.

6 Discussion & Areas of improvement

6.1 Real time forecasting and specification issues

We acknowledge the potential discrepancy between our analyzed information set and the data analyzed by professional forecasters. Our use of revised data from Datastream does not ensure us that we are as close as possible to the information available to forecasters at the time of their analysis. In our future research, we aim to work with non-revised data to reflect the information available at the time of forecasting. This approach would allow us to carry out the forecast based on the same information set as the forecasters. However, due to the unavailability of real-time vintages for all series in our panel, we acknowledge that our simulated out-of-sample analysis is a "pseudo" real-time evaluation (Giannone et al., 2008). Then, we recognize the need to reestimate the factors over each period when extracting them so as to use all the data available at the time of the forecast. Additionally, in order to address missing values in the data we emphasize the importance of performing the EM algorithm across different blocks of activity (real activity, inflation, money and interest rates, financial conditions) rather than per frequency (daily, monthly, quarterly), even if the missing values represent less than 5% of the total observations in the time span 2020Q1-2023Q2 and would not affect the accuracy of our forecasts in a significant way.

Regarding the specification, we would propose going to a FAVARMA framework that takes into account the autocorrelation of the residuals, which is a issue in our current model. Dufour and Stevanović (2013) show that FAVARMA provides better forecast accuracy compared to standard factors models for several key macroeconomic aggregates. However, it is important to highlight that FAVARMA models involve the estimation of factors using methods such as maximum likelihood and nonlinear least squares. Thus, these techniques require nonlinear optimization, which might not be feasible with the large number of parameters to estimate (Dufour and Stevanović, 2013). As we continue to enhance our dataset, we believe it would be valuable to incorporate the improvements discussed into a time-varying FAVAR (TV-FAVAR) framework (Eickmeier et al., 2011). This approach allows for the variation of factor loadings, factor dynamics, and the variance-covariance matrix of innovations over time. By considering the time-varying nature of these factors, we can capture a more comprehensive understanding of economic developments, particularly in the aftermath of the 2007 Great Recession and the COVID-19 crisis. However, as in FAVARMA case it requires high computational resources.

6.2 Dealing with the (post) Covid period

The Covid-19 pandemic in 2020 and the ongoing Ukraine conflict have introduced significant uncertainties and disruptions to the global supply chain between the first quarter of 2020 and now. The repercussions of Covid-19 and now the consequences of the war in Ukraine hit every part of the world, leaving no region untouched. During this period, our FAVAR model's predictive performance suffers more in average compared to other benchmarks (especially univariate benchmarks) and relatively to the period before Covid. Indeed, in period of relative stability, such as from the first quarter of 2010 to the fourth quarter of 2019, as well as in the overall out-of-sample period from the first quarter of 2010 to the third quarter of 2023, our FAVAR model consistently perform relatively well in long-run. To address this issue, we would propose for further research extending our FAVAR model to account for the unique impact of foreign factors on Swiss GDP growth (Y_t) and on the domestic factors (F_t^H) in the Covid-19 and post-Covid-19 period, following the approach of Boivin et al. (2009) and Mumtaz and Surico (2009) who apply this methodology to deal with the post-2000 era and its turning point. This enhancement involves augmenting our VAR system within the FAVAR framework with a dummy variable that interacts with the lagged foreign factors, including real activity, inflation, money and interest rates, and financial conditions factors. To be specific, we suggest estimating the following system:

$$\begin{pmatrix} Y_t \\ F_t^H \\ F_t^* \end{pmatrix} = \begin{pmatrix} \phi_{11}(L) & \phi_{12}(L) & \phi_{13}(L) \\ \phi_{21}(L) & \phi_{22}(L) & \phi_{23}(L) \\ 0 & 0 & \phi_{33}(L) \end{pmatrix} \begin{pmatrix} Y_{t-1} \\ F_{t-1}^H \\ F_{t-1}^* \end{pmatrix} + \begin{pmatrix} \phi_{13}^d(L) \\ \phi_{23}^d(L) \\ \phi_{33}^d(L) \end{pmatrix} d_t F_{t-1}^* + \begin{pmatrix} u^Y \\ u^H \\ u^* \end{pmatrix} \tag{6.1}$$

Here, the dummy variable d_t takes the value 0 during the period from the first quarter of 2000 to the fourth quarter of 2019 and 1 thereafter (from the first quarter of 2020 to the second quarter of 2023). For example, for the F_t^H (domestic factor) regression the coefficients $\phi_{23}(L)$ capture the effects of lagged foreign factors from 2000 Q1 to 2019 Q4, and the coefficients $\phi_{23}(L) + \phi_{23}^d(L)$ the period from 2020 Q1 to 2023 Q2. To minimize the burden of parameter estimation and preserve degrees of freedom, we can consider reducing the lag length of our FAVAR model to 1. By incorporating this refined FAVAR model, we aim to better capture the dynamics of Swiss GDP growth in the context of the Covid-19 period and the ongoing Ukraine conflict.

7 Conclusion

On the whole, our Factor Augmented VAR (FAVAR) model tailored for forecasting real Swiss GDP growth has shown promising results. Thanks to PCA and EM algorithm, we successfully extract factors from a diverse set of variables with missing values, allowing us to model them alongside the target variable within a VAR framework.

One of the key strengths of our FAVAR is its structural approach, where we divide the model into two blocks: one for foreign factors and another for domestic factors. Each block encompasses essential economic indicators such as real activity, inflation, money supply & interest rates and financial conditions summarized by factors. This partitioning contributes to the model's economic interpretability, enabling us to gain valuable insights into the underlying driving forces of the Swiss economy.

In assessing the out-of-sample performance, our FAVAR has demonstrated considerable forecasting abilities, outperforming several benchmarks, particularly in longer forecasting horizons even if not always significantly.

Furthermore, in every evaluation sample and across the vast majority of cases our FAVAR performs better than a random walk. This suggests that our model is well-equipped to make intelligent forecasts by effectively using all available information. Moreover, the addition of blocks and factors in our FAVAR does not compromise its statistical accuracy.

However, it is essential to acknowledge the impact of extraordinary events such as the Covid-19 crisis and the war in Ukraine. During this period, we observe a decline in forecast performance over the long term and a general deterioration across all horizons, albeit still outperforming the random walk benchmark. Such challenges highlight the importance of continuously adapting and improving the model to address unexpected shocks and uncertainties in the global economy.

In light of these findings, our FAVAR model represents a valuable tool for the Swiss Institute of Applied Economics of the University of Lausanne (CREA) for risk assessment and scenario analysis in future research. As the economic landscape evolves, the FAVAR's ability to provide accurate and interpretable forecasts will be crucial in supporting informed decision-making.

References

Ajevskis, V., & Dāvidsons, G. (2008). Dynamic factor models in forecasting latvia's gross domestic product. *Department of the Bank of Latvia*, *2*, 24.

Bai, J., Li, K., & Lu, L. (2016). Estimation and inference of favar models. *Journal of Business & Economic Statistics*, *34*(4), 620–641.

Banerjee, A., Marcellino, M., & Masten, I. (2014). Forecasting with factor-augmented error correction models. *International Journal of Forecasting*, *30*(3), 589–612.

Bäurle, G., Steiner, E., & Züllig, G. (2021). Forecasting the production side of gdp. *Journal of Forecasting*, *40*(3), 458–480.

Bernanke, B. S., Boivin, J., & Eliasz, P. (2005). Measuring the effects of monetary policy: A factor-augmented vector autoregressive (favar) approach. *The Quarterly journal of economics*, *120*(1), 387–422.

Berthold, B. (2023). The macroeconomic effects of uncertainty and risk aversion shocks. *European Economic Review*, *154*, 104442.

Boivin, J., Giannoni, M. P., & Mihov, I. (2009). Sticky prices and monetary policy: Evidence from disaggregated us data. *American economic review*, *99*(1), 350–384.

Brown, R. L., Durbin, J., & Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society: Series B (Methodological)*, *37*(2), 149–163.

Burri, M., & Kaufmann, D. (2020). A daily fever curve for the swiss economy. *Swiss Journal of Economics and Statistics*, *156*, 1–11.

Christiano, L., & Eichenbaum, M. S. (1992). Liquidity effects and the monetary transmission mechanism.

Christiano, L. J., Eichenbaum, M., & Evans, C. L. (1999). Monetary policy shocks: What have we learned and to what end? *Handbook of macroeconomics*, *1*, 65–148.

Diebold, F. X., & Mariano, R. S. (1995). Com paring predictive accu racy. *Journal of Business and Economic Statistics*, *13*(3), 253–263.

Du Plessis, S., & Kotze, K. (2012). Trends and structural changes in south african macroeconomic volatility. *Economic Research Southern Africa Working Paper*, *297*.

Dufour, J.-M., & Stevanović, D. (2013). Factor-augmented varma models with macroeconomic applications. *Journal of Business & Economic Statistics*, *31*(4), 491–506.

Eickmeier, S., Lemke, W., & Marcellino, M. (2015). Classical time varying factor-augmented vector auto-regressive models—estimation, forecasting and structural analysis. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 493–533.

Eickmeier, S., Lemke, W., & Marcellino, M. G. (2011). Classical time-varying favar models-estimation, forecasting and structural analysis.

Eickmeier, S., & Ziegler, C. (2008). How successful are dynamic factor models at forecasting output and inflation? a meta-analytic approach. *Journal of Forecasting*, *27*(3), 237–265.

Faust, J., & Wright, J. H. (2013). Forecasting inflation. In *Handbook of economic forecasting* (pp. 2–56). Elsevier.

Fisher, J. D., Liu, C. T., & Zhou, R. (2002). When can we forecast inflation? *Economic Perspectives-Federal Reserve Bank of Chicago*, *26*(1), 32–44.

Giannone, D., Reichlin, L., & Small, D. (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of monetary economics*, *55*(4), 665–676.

Harvey, D., Leybourne, S., & Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of forecasting*, *13*(2), 281–291.

Hendry, D. F., & Hubrich, K. (2011). Combining disaggregate forecasts or combining disaggregate information to forecast an aggregate. *Journal of business & economic statistics*, *29*(2), 216–227.

Hubrich, K. (2005). Forecasting euro area inflation: Does aggregating forecasts by hicp component improve forecast accuracy? *International Journal of Forecasting*, *21*(1), 119–136.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica: journal of the Econometric Society*, 1551–1580.

Jose, V. R. R., & Winkler, R. L. (2008). Simple robust averages of forecasts: Some empirical results. *International journal of forecasting*, *24*(1), 163–169.

Josse, J., Chavent, M., Liquet, B., & Husson, F. (2012). Handling missing values with regularized iterative multiple correspondence analysis. *Journal of classification*, *29*, 91–116.

Josse, J., & Husson, F. (2016). Missmda: A package for handling missing values in multivariate data analysis. *Journal of statistical software*, *70*, 1–31.

Kamber, G., Karagedikli, Ö., Ryan, M., et al. (2016). International spill-overs of uncertainty shocks: Evidence from a favar.

Kaufmann, D., & Lein, S. (2012). Is there a swiss price puzzle? *Swiss Journal of Economics and Statistics*, *148*, 57–75.

Litterman, R. B., et al. (1984). Above-average national growth in 1985 and 1986. *Quarterly Review*, *8*(Fall).

Luciani, M. (2014). Large-dimensional dynamic factor models in real-time: A survey. *Available at SSRN 2511872*.

Madhou, A., Sewak, T., Moosa, I., & Ramiah, V. (2020). Forecasting inflation in a small open developing economy. *Applied Economics*, *52*(20), 2123–2134.

Marcellino, M., Stock, J. H., & Watson, M. W. (2006). A comparison of direct and iterated multistep ar methods for forecasting macroeconomic time series. *Journal of econometrics*, *135*(1-2), 499–526.

Mincer, J. A., & Zarnowitz, V. (1969). The evaluation of economic forecasts. In *Economic forecasts and expectations: Analysis of forecasting behavior and performance* (pp. 3–46). NBER.

Miranda-Agrippino, S., & Rey, H. (2020). Us monetary policy and the global financial cycle. *The Review of Economic Studies*, *87*(6), 2754–2776.

Mitchell, J., Smith, R. J., Weale, M. R., Wright, S., & Salazar, E. L. (2005). An indicator of monthly gdp and an early estimate of quarterly gdp growth. *The Economic Journal*, *115*(501), F108–F129.

Mumtaz, H., & Surico, P. (2009). The transmission of international shocks: A factor-augmented var approach. *Journal of Money, Credit and Banking*, *41*, 71–100.

Newey, W. K., & West, K. D. (1986). A simple, positive semi-definite, heteroskedasticity and autocorrelationconsistent covariance matrix.

Rahimov, V., Guliyev, N., & Ahmadov, V. (2020). *Modeling azerbaijan's inflation and output using a factor-augmented vector autoregressive (favar) model* (tech. rep.). Graduate Institute of International and Development Studies Working Paper.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica: journal of the Econometric Society*, 1–48.

Stock, J. H., & Watson, M. W. (2002). Forecasting using principal components from a large number of predictors. *Journal of the American statistical association*, *97*(460), 1167–1179.

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

Stock, J. H., & Watson, M. W. (2016). Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics. In *Handbook of macroeconomics* (pp. 415–525). Elsevier.

Stuart, R. (2018). A quarterly phillips curve for switzerland using interpolated data, 1963–2016. *Economic Modelling*, *70*, 78–86.

Vasishtha, G., & Maier, P. (2013). The impact of the global business cycle on small open economies: A favar approach for canada. *The North American Journal of Economics and Finance*, *24*, 191–207.