Université Paris Dauphine-PSL
Paris, December 2023

The relationship between short-term interest rates and Gross Fixed Capital Formation in France

Sylvain Moriceau and Baptiste Bachmann
Supervisor: Arthur Thomas

Econometrics project
Major: Economics and Financial Engineering

Abstract

The study examines the relationship between interest rates and investment rates in France from Q1 2005 to Q1 2023 using an econometric approach. Using empirical data, our analysis aims to measure the impact of changes in interest rates on investment levels, while taking into account economic and contextual factors such as the unemployment rate and inflation. Using MA(2) and VAR(5) models, we explore the stationarity, causality and cointegration of these time series. Our results reveal a causal relationship between interest rates and investments, where interest rates are mostly influenced by their own historical shocks while investments are largely influenced by interest rates. The Trace and maximum eigenvalue cointegration tests confirm the existence of two cointegrating relationships, indicating the presence of long-term links between the variables studied.

Contents

List of Figures

1 General Introduction

1.1 Introduction

In the last few decades, 'investments' and 'interest rates' have emerged as fundamental concepts in our societal lexicon. In an era marked by persistent consumption and production, these two global financial constructs significantly influence our day-to-day activities. Central banks, navigating the uncertain economic waters of recent years, have grappled with the challenge of identifying the ideal interest rate that balances the stimulation of investment with the containment of inflation.

Within the realm of investments, accounting recognizes three distinct types:

- **Tangible Investments**: These include the purchase of physical goods and assets, such as buildings and equipment.
- **Intangible Investments**: These are acquisitions that enhance the company's non-physical assets, like goodwill and patents.
- **Financial Investments**: These are the buying of financial instruments, such as shares or bonds, which increase the company's financial portfolio.

Our paper will concentrate on Gross Fixed Capital Formation (GFCF), which the OECD refers to as investment (OECD, 2023b), and INSEE identifies as a measure of material investments made within French territory over a year (INSEE, 2020). From an economic point of view, this is the purchase of durable goods to strengthen a company's capital, typically involving significant assets like buildings or machinery, aimed at enhancing the production of goods and services. In our analysis, GFCF will encompass both tangible and intangible investments, despite the latter being traditionally challenging to quantify and often excluded in the computations.

The topic of interest rates presents a battlefield for central banks, with their decisions to alter the interest rate dynamics being minutely observed by financial entities. Our focus will be on short-term interest rates, acknowledging their direct correlation with the policy interest rate, the rate that dictates borrowing costs for commercial banks, either from the central bank or among themselves. To avoid confusion, we will refer to both the key rate and short-term interest rates, understanding that shifts in the latter are often a direct result of policy rate changes. The OECD defines short-term interest rates as the "rates at which financial institutions lend to one another" (OECD, 2023c), whereas the key interest rate is the rate for borrowing or investing with central banks. We will also discuss the European Central Bank's (ECB) role, as the authority over key interest rates in the Eurozone. Moreover, the central banks' recent quantitative easing measures — the large-scale purchasing of various bonds and assets to stimulate the economy — will not be classified as investments in our study because they refer to financial investments, which we do not consider. Such purchases have led to reduced interest rates, enabling more affordable financing for households, businesses, and governments, thus promoting economic growth and adjusting inflation towards targets that support price stability, as was evident during the Covid-19 crisis.

Why focus on France? As a significant global and European economic force, ranked 7th in the world by GDP according to the International Monetary Fund (IMF, 2023), and the most attractive European nation for foreign investors as per Ernst and Young's annual reports (EY, 2023), France presents a compelling case for study. Despite satisfying results, France faces unique challenges, particularly in job creation from foreign investment projects — foreign investments created fewer jobs (33 per project) compared to the other main host countries of foreign investments in Europe (59 in the UK and 58 in Germany), — which justifies a more in-depth analysis of the evolution of its investments and the impact of interest rates.

Therefore, this paper aims to dissect the relationship between GFCF and short-term interest rates in France by examining their time series data from Q1 2005 to Q1 2023. We question whether there is a causal relationship between these variables: does a decrease in short-term interest rates lead to an increase in investments and vice versa? The study has seven basic components. Firstly, a graphical and macroeconomic analysis of the time series is carried out. Afterwards, a unit root test is applied to the series, followed by a stationarity study. An ARMA model is then used to generate forecasts for the year 2020 for the interest rate series. Furthermore, a VAR model is drawn and the cointegration of the series is studied. Finally, a VECM model is developed.

1.2 Presentation of the selected data

As mentioned in the introduction we focus our analysis on 2 time series with the same time frame: short-term interest rates in the Eurozone and the GFCF investments rate in France. Short-term interest rates for France are the same as for the Eurozone, as it is the European Central Bank (ECB) that is responsible to set up these rates. On the other hand, the rate of the GFCF investments are only focused on France, which makes it an interesting feature to watch while looking at the economic dynamics in a country.

The table of data is to be found in the Excel sheet attached. We have 73 observations for the timeframe Q1 2005 to Q1 2023. As it is also needed to analyze the time series excluding the Covid crisis, for the timeseries we have 60 observations going from Q1 2005 to Q3 2019. This study was carried out using RStudio software.

2 Analysis of graphs, simple and partial autocorrelograms of our time series

2.1 Interest rate

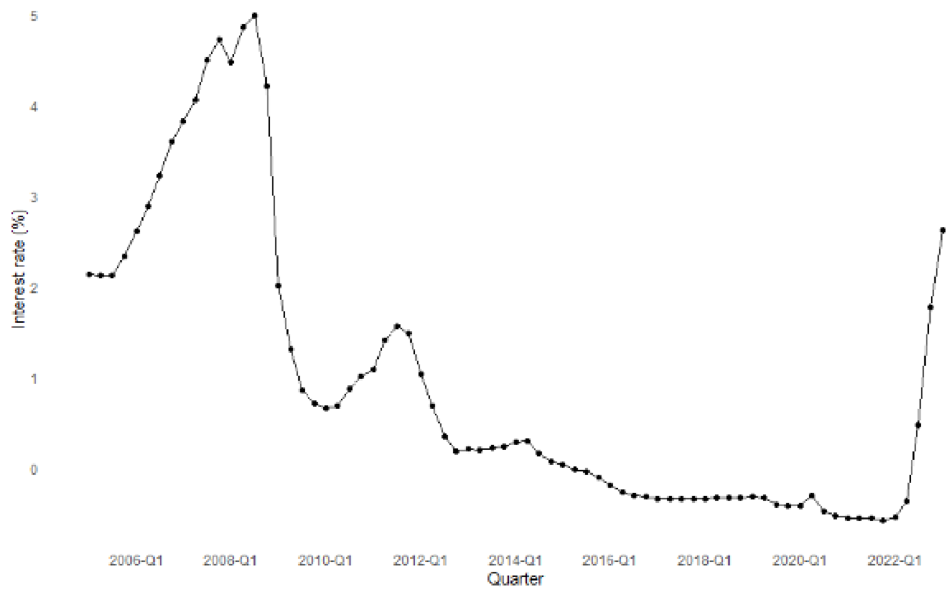


Figure 2.1: Interest Rate evolution from Q1 2005 to Q1 2023

At first glance, it appears that this graphic is composed of 4 different parts:

- from Q1 2005 to Q4 2007
- from Q1 2008 to Q4 2011
- from Q1 2012 to Q4 2021
- from Q1 2022 to Q1 2023

To a have better understanding of these changes and evolution's in interest rates, we will emphasis the following analysis on the timeline decomposition proposed just above.

Beginning in Q1 2005, the interest rate stood at approximately 2%. A few years following the euro's adoption as the national currency, France exhibited a robust economic situation. Notably, the unemployment rate decreased significantly from 9% to about 7% by Q4 2007 (INSEE, 2023b), marking the lowest level until Q1 2023. Globally, the period saw a marked increase in oil prices, which rose from \$37 in Q1 2005 to \$90 in Q4 2007, eventually peaking at \$141 in June 2008 ("Oil Price", 2023) — an approximate 280% rise within a span of three years. Concurrently, the European Central Bank (ECB) executed six increases in short-term interest rates, culminating in a rate close to 5% by the end of 2007 — a 150% hike over two years. These adjustments were largely in response to inflationary pressures within the Eurozone, with rising consumer prices signaling potential inflation. Additional motivations for the rate hikes included the prevention of economic overheating to ensure sustainable growth and the strengthening of the euro, as higher rates tend to attract foreign investment, potentially appreciating the currency. To summarize this period, 2005 was a stable year, while 2006 and 2007 were particularly strong, with GDP growth rates of 2.6% and 2.7% respectively (INSEE, 2023a). However, the last two quarters of 2007 were challenging due to financial instability stemming from the US housing market crisis.

The period following Q1 2008 is etched in the collective memory of financial markets as a time of significant distress. The catalyst was the subprime mortgage crisis in the United States, initiated by unsustainable levels of personal debt. Due to the intertwined global economy, the crisis quickly spread beyond our borders, profoundly affecting France and the rest of the world. In response to this financial maelstrom, central banks globally deployed expansive monetary policies aimed at stabilizing the faltering economic climate and alleviating the repercussions on corporate solvency and household expenditure. The European Central Bank (ECB) has been at the forefront of this battle, orchestrating a spectacular drop in interest rates from almost 5% to just under 1%. This unprecedented reduction in the euro area was a strategic bid to jump-start economic activity amidst a climate of inflation rates that remained stubbornly low, contradicting typical recessionary trends (INSEE, 2023c). Stock markets reflected this economic upheaval, with the CAC40 index, a barometer of French economic health, nosediving by 43% from Q1 2008 to Q1 2009 (TradingView, 2009). Concurrently, the French labor market, which had previously been on an upward trajectory, was hit hard. The unemployment rate experienced a significant uptick, increasing by 2.5 percentage points (INSEE, 2023b).

Emerging from this period of economic strain, the ECB, in an attempt to curtail inflationary pressures, raised interest rates to 1.5% in 2011 — the first increase after a three-year break (Le Monde, 2011). The decision came against a backdrop of sovereign debt crises in Eurozone nations such as Ireland, Portugal, Spain, and notably Greece, which collectively contributed to an environment of heightened inflation concerns (Les Echos, 2019). Despite these macroeconomic headwinds, France experienced an unexpected decline in unemployment rates (source: detailed unemployment data), and the national economy showed signs of resilience with a GDP growth of 1.7% in 2011. This growth mirrored the previous year's rate and marked a cautious recovery from the profound 2009 contraction of 3.1%. By the end of 2011, France's GDP was still slightly shy of its pre-recession peak from the first quarter of 2008 (Statista, 2023). The period thus encapsulated a continued, albeit fragile, economic recovery, with the French economy bearing the aftershocks of these tumultuous years.

The period from Q1 2012 to Q4 2021 for the French economy is characterized by the consequences of the sovereign debt crisis and the persistent efforts of the European Central Bank (ECB) to steer the region towards recovery and growth. During this extended period, the ECB's monetary policy was predominantly accommodating, with interest rates set at historical lows, occasionally dipping into negative territory. This was part of a broader strategy to invigorate the Eurozone's economy, stimulate lending, and preempt deflationary pressures. As observed on the graph, there is a clear trend for this period, at first a consequent drop in interest rates, followed by a long and slight decrease. In France, these measures facilitated a slow yet steady improvement in the economic landscape, especially within the job market. However, the descent in unemployment was gradual, revealing the deep-seated challenges from previous economic downturns and the inherent rigidity within the labor market. For instance, at peak the unemployment rate was at 10% in Q2 2015, dropping to 7.5% in Q2 2020 (INSEE, 2023b). The latter years of this period were overshadowed by the global COVID-19 pandemic, commencing in late 2019, which precipitated a profound economic contraction.The ECB's response, including further monetary easing, and substantial fiscal stimulus measures by national governments, including France, sought to mitigate the pandemic's devastating economic impact. Despite these interventions, the French economy confronted the end of 2020 with a recovery that was still nascent, reflecting the pandemic's deep scars. Economic indicators, including GDP growth, exhibited significant volatility (INSEE, 2023a), symptomatic of the crisis's profound and erratic impact. In summary, the period from Q1 2012 to Q4 2021 was one of recuperation and resilience, as France navigated through a landscape marred by economic crises and extraordinary global upheavals. The policies enacted by the ECB, alongside national french measures, were critical in mitigating the downturns and setting the stage for a post-pandemic economic rebound.

Finally for the last period that starts in Q1 2022, the all dynamic is changing after this trendy decrease in interest rates going 0% to 3%, taking in note that the increase occurred in Q2 2022 after 11 years of non-increase. Price rises in the Eurozone - 6.5% in Q2 2022 (INSEE, 2023c) - continue to gain momentum due to the combined impact of the post-pandemic economic resurgence. The ECB thus closes the era of negative rates begun in 2014, and closes a decade of generous monetary policy that has helped the economy overcome the crises of recent years. Besides, the unemployment rate decreases to its lowest level since Q1 2005 to 6.9% (INSEE, 2023b) which emphasized a changing era.

Visually, from the graph alone, the interest rate series does not appear to be stationary due to the visible trends and changing variance.

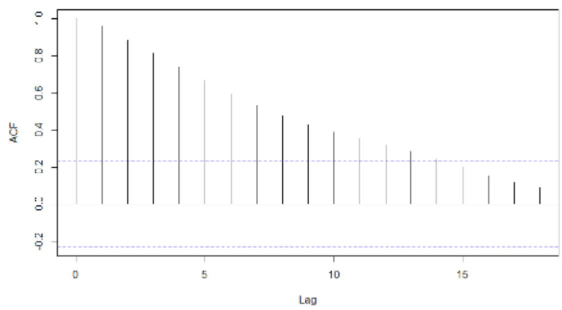


Figure 2.2: Interest Rate Simple Autocorrelogram

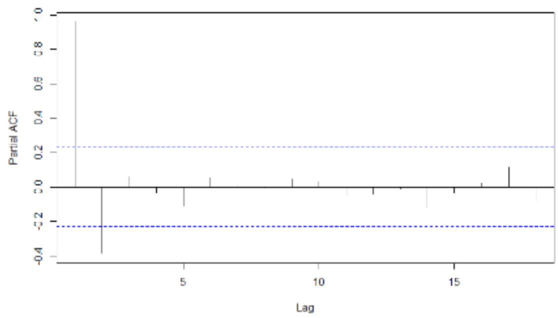


Figure 2.3: Interest Rate Partial Autocorrelogram

Autocorrelations are positive and significant up to lag 14. They are close to 1 for lags 1, 2, 3 and 4 and then decrease towards 0. This suggests that the interest rate is fairly autocorrelated. Nevertheless, it is rather difficult to draw a conclusion on the stationarity or non-stationarity of the interest rate. The autocorrelation profile could correspond to a stationary series. Partial autocorrelations are significant for lags 1 and 2.

2.2 GFCF investment rate

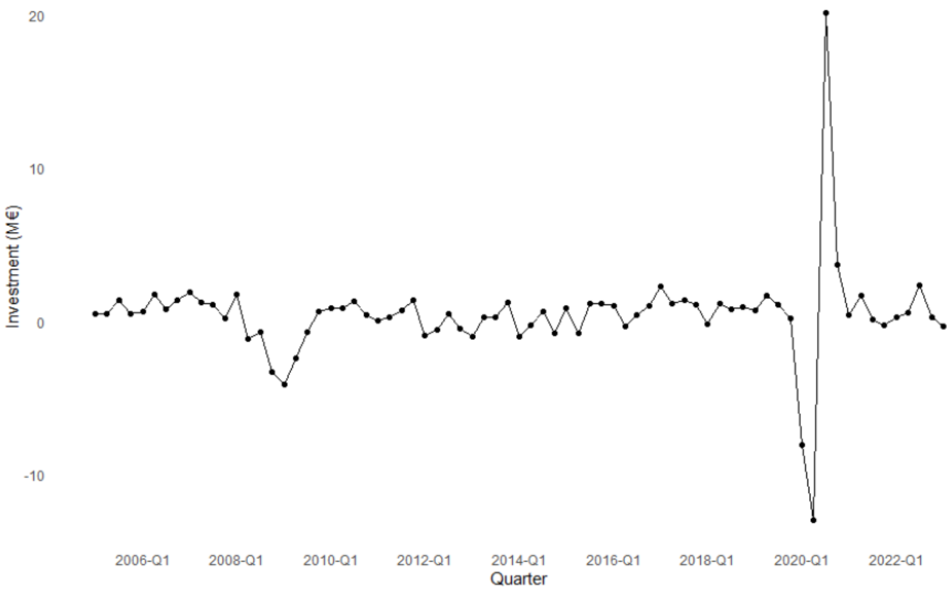


Figure 2.4: GFCG Investment rate from Q1 2005 to Q1 2023

The graph is depicting France's Gross Fixed Capital Formation (GFCF) investment rate from Q1 2005 to Q1 2023 which shows a detailed picture of the country's economic health with facing two major crises: the 2008 Global Financial Crisis and the COVID-19 pandemic of 2020. Starting with the period before 2008, France's investment rate was on a generally positive trajectory, indicative of growth and expansion in the economy. In 2008 triggered by the collapse of the sub-primes crisis in the United States, it had a domino effect on the global financial system, including the Eurozone and therefore France. The rate of investment in France has fallen, as banks have had to contend with the devaluation of assets linked to US real estate. French banks' exposure to these toxic assets resulted in a loss of wealth and a tightening of credit, which had repercussions on the economy. Businesses faced with an uncertain future and a lack of financing put investments on hold, leading to a visible dip in the GFCF rate with a peak in the negative areas at -4% in Q1 2009.

As the French government and the European Central Bank took action, including bailouts and liquidity injections, confidence started to return, and the investment rate saw a gradual recovery. However, this 2010-2011 recovery was not immediate, and did not restore pre-crisis levels. Fast forward to 2020, the graph shows a dramatic plunge in the investment rate due to the COVID-19 pandemic at around -13% rate in Q2 2020. This was a result of unprecedented global worldwide lock downs, which halted economic activity and led to a sharp contraction in GFCF. The uncertainty about the duration and severity of the pandemic led to businesses conserving cash and delaying or canceling investment plans, causing a significant downturn in the GFCF investment rate.

The major peak at almost 20% at the next quarter represents the rebound effect as France began to emerge from lock-downs. The French government's stimulus measures aimed at reviving the economy, including financial support for businesses and investments in public services, led to a surge in GFCF for corporates. This recovery, however, was not smooth and showed signs of volatility with uncertainty regarding the affected economic situation affecting the flow of investments.

By Q1 2023, the investment rate in France appears to show signs of stabilization, although it has not returned to the pre-pandemic levels. The French economy, like many others, continues to navigate the uncertainties post-pandemic, including supply chain issues and changes in the global trade landscape. The path of the GFCF investment rate thus encapsulates the broader story of the French economy's journey through two of the most challenging periods in recent history.

Finally, except these 2 highly challenging periods, the trend is stable with an average positive growth rate between 0% and 2%. It is noticed that for a well developed country like France, it is difficult to set an optimal GFCF investment rate. Moreover, another factor that would be relevant to analyze is the part of GFCF as a percentage of GDP for developed economies - which is around 20% according the OECD (OECD, 2023a).

Looking at the graph, we can see that the investment rate fluctuates around a level that does not appear to have a clear upward or downward trend over the total duration, despite some fluctuations. However, the extreme peak corresponding to the COVID period indicates a potential change in the mean and variance, which could suggest non-stationarity.

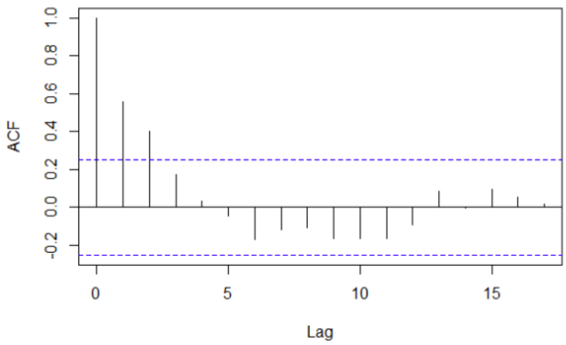


Figure 2.5: GFCF Simple Autocorrelogram

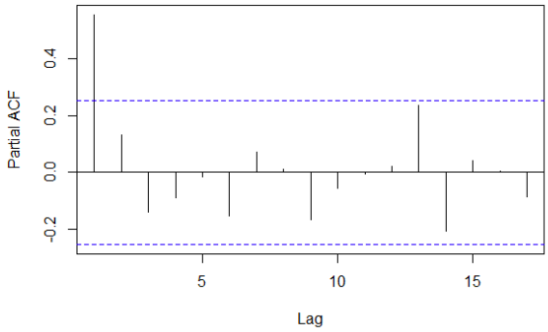


Figure 2.6: GFCF Partial Autocorrelogram

Autocorrelations are positive and significant up to lag 3. The first lag is equal to 1, which indicates that the series is strongly autocorrelated and is typical of a random walk. It then decreases rapidly towards 0, which is a good indicator of stationarity. Partial autocorrelations are significant up to lag 1, which is another possible sign of a random walk.

The observations obtained from the previous graphs are not sufficient to fully understand the stationarity status of the series. Therefore, we will apply statistical tests capable of detecting the presence of a unit root in a time series.

3 Unit root tests and stationarity method unit root test

3.1 Interest rate

3.1.1 Augmented Dickey-Fuller test

Let's test for the presence of unit roots with the augmented Dickey Fuller Test. We implement a strategy Sequential Unit Root Test where both the Unit Root and the deterministic specification of the model are tested.

First, we consider the most general model : model with constant and trend (M3). The Unit Root is tested by applying the critical values of M3, and then an appropriate test is used to verify that the model chosen is the correct one. Otherwise, we start again in a more constrained model : model with constant (M2) and finally in (M1) which is a model with trend.

$$\Delta X_t = c + bt + \rho X_{t-1} + \sum_{j=1}^{g} \delta \Delta X_{t-j} + \varepsilon_t \quad \textbf{(M3)}$$

$$\Delta X_t = c + \rho X_{t-1} + \sum_{j=1}^{g} \delta \Delta X_{t-j} + \varepsilon_t \quad \textbf{(M2)}$$

$$\Delta X_t = \rho X_{t-1} + \sum_{j=1}^{g} \delta \Delta X_{t-j} + \varepsilon_t \quad \textbf{(M1)}$$

The Dickey-Fuller test verifies the following hypothesis:

$$\begin{cases} H_0 : & \rho = 0 \quad \text{(non-stationary)} \\ H_1 : & \rho \neq 0 \quad \text{(stationary)} \end{cases}$$

The rule of thumb is we reject the H_0 hypothesis if $t < t_{ADF}$.
Where t is the test statistic and t_{ADF} is the critical values of the Dickey Fuller table.

Model 3 (constant and trend) : $\Delta X_t = c + bt + \rho X_{t-1} + \sum_{j=1}^{2} \delta \Delta X_{t-j} + \varepsilon_t$ With $\varepsilon_t \sim \left(0, \sigma_\varepsilon^2\right)$

Period $2005 - 2023$:

```
################################################
# Augmented Dickey-Fuller Test Unit Root Test #
################################################

Test regression trend


Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

Residuals:
     Min       1Q   Median       3Q      Max
-1.46976 -0.06753 -0.04249  0.06421  0.91596

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.185061   0.144226   1.283   0.2039
z.lag.1     -0.074486   0.038042  -1.958   0.0544 .
tt          -0.002875   0.002975  -0.966   0.3374
z.diff.lag   0.727823   0.099873   7.287 4.62e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.298 on 67 degrees of freedom
Multiple R-squared:  0.4622,    Adjusted R-squared:  0.4381
F-statistic:  19.2 on 3 and 67 DF,  p-value: 4.311e-09


Value of test-statistic is: -1.958 1.6837 2.4804

Critical values for test statistics:
      1pct  5pct 10pct
tau3 -4.04 -3.45 -3.15
phi2  6.50  4.88  4.16
phi3  8.73  6.49  5.47
```

Figure 3.1: Augmented DF Unit Root Test 2005-2023

Here $t_{obs} = -1.958 > t_{ADF} = -3.45$ so we don't reject the null hypothesis at the 5% threshold.

Study let's test the model. The assumption for this test is as follow:

$$\begin{cases} H_0: & b = 0 \quad \text{(non significance of the time variable)} \\ H_1: & b \neq 0 \quad \text{(significance of the time variable)} \end{cases}$$

The rule of thumb is: we reject H_0 if $|t_{obs}| > |t_{ADF}|$, with the same notations as before. If we reject H_0, we stop and conclude that a process I(1) + C + T (unit root not rejected) or I(0) + C + T (unit root rejected) is concluded. Otherwise, we must consider a more constrained model, model 2.

$|t_{obs}| = 0.966 < |t_{ADF}| = 3.18$ so we don't reject the null hypothesis at the 5% threshold. We must consider a model without determinism.

Period $2005 - 2020$:

```
################################################
# Augmented Dickey-Fuller Test Unit Root Test #
################################################

Test regression trend


Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

Residuals:
     Min       1Q   Median       3Q      Max
-1.45594 -0.07286  0.00156  0.07713  0.79884

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.476672   0.167753   2.842  0.00632 **
z.lag.1     -0.133075   0.040942  -3.250  0.00199 **
tt          -0.011069   0.003975  -2.785  0.00737 **
z.diff.lag   0.633185   0.102995   6.148 9.87e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2829 on 54 degrees of freedom
Multiple R-squared:  0.4442,    Adjusted R-squared:  0.4133
F-statistic: 14.38 on 3 and 54 DF,  p-value: 5.255e-07


Value of test-statistic is: -3.2503 3.6095 5.2952

Critical values for test statistics:
      1pct  5pct 10pct
tau3 -4.04 -3.45 -3.15
phi2  6.50  4.88  4.16
phi3  8.73  6.49  5.47
```

Figure 3.2: Augmented DF Unit Root Test 2005-2020

We have $t_{\text{obs}} = -3.250 > t_{ADF} = -3.45$ so we don't reject the null hypothesis of unit root's presence at the 5% threshold.

The statistical t of the model is equal to $|t_{\text{obs}}| = 2.785 < |t_{ADF}| = 3.18$, therefore we don't reject the null hypothesis at the 5% threshold. We must consider a model without determinism.

Model 2 (constant) $\Delta X_t = c + \rho X_{t-1} + \sum_{j=1}^{2} \delta \Delta X_{t-j} + \varepsilon_t$ With $\varepsilon_t \sim \left(0, \sigma_\varepsilon^2\right)$

Period $2005 - 2023$:

```
###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression drift


Call:
lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)

Residuals:
     Min       1Q   Median       3Q      Max
-1.53411 -0.06689 -0.03559  0.06529  0.86497

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.05140    0.04077   1.261   0.2117
z.lag.1     -0.04468    0.02225  -2.008   0.0486 *
z.diff.lag   0.69340    0.09326   7.435 2.32e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2979 on 68 degrees of freedom
Multiple R-squared:  0.4547,    Adjusted R-squared:  0.4387
F-statistic: 28.35 on 2 and 68 DF,  p-value: 1.109e-09


Value of test-statistic is: -2.0078 2.0608

Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.51 -2.89 -2.58
phi1  6.70  4.71  3.86
```

Figure 3.3: Augmented DF Unit Root Test 2005-2023

Here $t_{obs} = -2.008 > t_{ADF} = -3.45$ so we don't reject the null hypothesis of unit root's presence at the 5% threshold. Concerning the Model 2 test, $|t_{obs}| = 1.261 < |t_{ADF}| = 2.89$ so we don't reject the null hypothesis at the 5% threshold. We must consider a model without constant.

Period $2005 - 2020$:

```
###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression drift


Call:
lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)

Residuals:
     Min       1Q   Median       3Q      Max
-1.60624 -0.04610 -0.01605  0.08833  0.67541

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.02763    0.04904   0.563    0.575
z.lag.1     -0.03853    0.02425  -1.589    0.118
z.diff.lag   0.60218    0.10850   5.550 8.47e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2997 on 55 degrees of freedom
Multiple R-squared:  0.3644,    Adjusted R-squared:  0.3412
F-statistic: 15.76 on 2 and 55 DF,  p-value: 3.877e-06


Value of test-statistic is: -1.5889 1.3684

Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.51 -2.89 -2.58
phi1  6.70  4.71  3.86
```

Figure 3.4: Augmented DF Unit Root Test 2005-2020

In the ADF test, the observed t-statistic (t_{obs}) is -1.589 , which is greater than the critical value at the 5% threshold ($t_{ADF} = -2.89$). Therefore, we don't reject the null hypothesis.

Regarding the Model 2 test, the absolute value of the observed t-statistic $|t_{obs}| = 0.563$, which is less than the absolute value of the critical t-statistic at the 10% threshold ($|t_{ADF}| = 2.89$). As a result, we do not reject the null hypothesis. Therefore, we should consider a model without a constant.

Model 1 (without constant and trend) $\Delta X_t = \rho X_{t-1} + \sum_{j=1}^{2} \delta \Delta X_{t-j} + \varepsilon_t$ With $\varepsilon_t \sim \left(0, \sigma_\varepsilon^2\right)$

Period $2005 - 2023$:

Here $t_{obs} = -1.585 > t_{ADF} = -1.95$ so we don't reject the null hypothesis at the 5% threshold.

```
###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression none


Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)

Residuals:
     Min       1Q   Median       3Q      Max
-1.54762 -0.01317  0.00976  0.08359  0.87098

Coefficients:
           Estimate Std. Error t value Pr(>|t|)
z.lag.1    -0.03071    0.01938  -1.585    0.118
z.diff.lag  0.68556    0.09345   7.336 3.26e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2991 on 69 degrees of freedom
Multiple R-squared:  0.4422,	Adjusted R-squared:  0.426
F-statistic: 27.35 on 2 and 69 DF,  p-value: 1.796e-09


Value of test-statistic is: -1.5845

Critical values for test statistics:
      1pct  5pct 10pct
tau1 -2.6 -1.95 -1.61
```

Figure 3.5: Augmented DF Unit Root Test 2005-2023

Our model is:

$$X_t \sim I(1) \text{ with } \Delta X_t = \varepsilon_t$$

<u>Period $2005 - 2020$:</u>

```
###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression none


Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)

Residuals:
     Min       1Q   Median       3Q      Max
-1.62069 -0.01881  0.00735  0.09777  0.66338

Coefficients:
           Estimate Std. Error t value Pr(>|t|)
z.lag.1    -0.03049    0.01948  -1.565    0.123
z.diff.lag  0.59153    0.10618   5.571  7.5e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2979 on 56 degrees of freedom
Multiple R-squared:  0.3696,	Adjusted R-squared:  0.3471
F-statistic: 16.42 on 2 and 56 DF,  p-value: 2.449e-06


Value of test-statistic is: -1.565

Critical values for test statistics:
      1pct  5pct 10pct
tau1 -2.6 -1.95 -1.61
```

Figure 3.6: Augmented DF Unit Root Test 2005-2020

In the ADF test, the observed t-statistic (t_{obs}) is -1.565 , which is greater than the critical value at the 5% threshold ($t_{ADF} = -1.95$). Therefore, we do not reject the null hypothesis that a unit root is present.

Our model is:

$$X_t \sim I(1) \text{ with } \Delta X_t = \varepsilon_t$$

There is a problem : the tests primarily minimize the error of first kind, i.e. in the previous tests the wrongly rejection of the hypothesis zero from UR (Unit Root). We want to limit the risk to wrongly reject the absence of RU. In order to do so, we reverse the assumptions in the KPSS test.

	H_0 (Reality)	H_1 (Reality)
D_0 (Decision)	$D_0 \mid H_0$ correct decision	$D_0 \mid H_1$ error of second kind
D_1 (Decision)	$D_1 \mid H_0$ error of first kind	$D_1 \mid H_1$ correct decision

3.1.2 KPSS Test

The KPSS test was proposed in 1992 by Kwiatkowski, Phillips, Schmidt and Shin to test the null hypothesis of absence of unit root, while the alternative hypothesis is the presence of a unit root. The results of the KPSS test on interest rates in the M2 are reported here (the KPSS test can only be done in the M2, which is the model closest to the M1 previously used in the sequential strategy).

Period 2005 – 2023 :

```
#######################
# KPSS Unit Root Test #
#######################

Test is of type: mu with 3 lags.

Value of test-statistic is: 1.2323

Critical value for a significance level of:
                10pct  5pct 2.5pct  1pct
critical values 0.347 0.463  0.574 0.739
```

Figure 3.7: KPSS Unit Root Test 2005-2023

$$\begin{cases} H_0 : \quad \rho \neq 0 \quad \text{(stationary)} \\ H_1 : \quad \rho = 0 \quad \text{(non-stationary)} \end{cases}$$

The value of test-statistic is equal to $LM_{obs} = 1.2323$ which is superior to each critical values. We can therefore confirm that the constant is not significant, and that the series is not stationary.

According to the ADF and KPSS tests, the interest rate is a DS process. A DS (Difference Stationary) describes a no stochastic type stationarity: it comes from an accumulation of random shocks in the process. At the level of economic issues, the effect of shocks is permanent on a DS process. At the level of statistical issues, the usual asymptotic properties of estimators and test statistics are no longer valid.

We must differentiate it to stationarize it.

Period 2005 – 2020 :

Same as before, the results of the KPSS test on interest rates in the M2 are reported here (the KPSS test can only be done in the M2, which is the model closest to the M1 previously used in the sequential strategy).

```
#######################
# KPSS Unit Root Test #
#######################

Test is of type: mu with 3 lags.

Value of test-statistic is: 1.1849

Critical value for a significance level of:
                10pct  5pct 2.5pct  1pct
critical values 0.347 0.463  0.574 0.739
```

Figure 3.8: KPSS Unit Root Test 2005-2020

For this test, the t-statistic ($LM_{obs} = 1.1849$) is superior to each critical values. We can therefore confirm that the constant is not not significant, and that the series is not stationary. Therefore, we reject the null hypothesis of stationarity at every threshold. The process is I(1). According to the ADF and KPSS tests, the interest rate is a DS process. We must differentiate it to stationarize it.

3.2 Investment

3.2.1 Augmented Dickey-Fuller test

Model 3 (constant and trend) : $\Delta X_t = c + bt + \rho X_{t-1} + \sum_{j=1}^{2} \delta \Delta X_{t-j} + \varepsilon_t$ With $\varepsilon_t \sim \left(0, \sigma_\varepsilon^2\right)$

Period $2005 - 2023$:

```
###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression trend


Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

Residuals:
     Min       1Q   Median       3Q      Max
-14.3576  -0.6214   0.2409   0.9079  16.2379

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.07072    0.78348   0.090   0.9283
z.lag.1     -1.32586    0.17193  -7.712 7.97e-11 ***
tt           0.01328    0.01861   0.714   0.4779
z.diff.lag   0.25464    0.11824   2.154   0.0349 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.199 on 67 degrees of freedom
Multiple R-squared:  0.5585,	Adjusted R-squared:  0.5388
F-statistic: 28.26 on 3 and 67 DF,  p-value: 6.339e-12


Value of test-statistic is: -7.7117 19.8245 29.7361

Critical values for test statistics:
      1pct  5pct 10pct
tau3 -4.04 -3.45 -3.15
phi2  6.50  4.88  4.16
phi3  8.73  6.49  5.47
```

Figure 3.9: Augmented DF Unit Root Test 2005-2023

Here $t_{\text{obs}} = -7.712 > t_{ADF} = -3.45$ so we don't reject the null hypothesis of unit root's presence at the 5% threshold. $|t_{obs}| = 0.714 < |t_{ADF}| = 3.18$ so we don't reject the null hypothesis that b equals 0 at the 5% threshold. We must consider a model without determinism.

Period $2005 - 2020$:

```
###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression trend


Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

Residuals:
     Min       1Q   Median       3Q      Max
-2.91377 -0.42362  0.09618  0.70273  1.48597

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.012934   0.275905   0.047  0.96278
z.lag.1     -0.397602   0.128616  -3.091  0.00315 **
tt           0.004459   0.008031   0.555  0.58105
z.diff.lag  -0.128071   0.135531  -0.945  0.34889
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.01 on 54 degrees of freedom
Multiple R-squared:  0.2401,	Adjusted R-squared:  0.1979
F-statistic: 5.688 on 3 and 54 DF,  p-value: 0.00185


Value of test-statistic is: -3.0914 3.1867 4.7801

Critical values for test statistics:
      1pct  5pct 10pct
tau3 -4.04 -3.45 -3.15
phi2  6.50  4.88  4.16
phi3  8.73  6.49  5.47
```

Figure 3.10: Augmented DF Unit Root Test 2005-2020

We have $t_{obs} = -3.091 > t_{ADF} = -3.45$ so we don't reject the null hypothesis of unit root's presence at the 5% threshold.

The statistical t of the model is equal to $|t_{obs}| = 0.555 < |t_{ADF}| = 3.18$, therefore we don't reject the null hypothesis at the 5% threshold. We must consider a model without determinism.

Model 2 (constant) $\Delta X_t = c + \rho X_{t-1} + \sum_{j=1}^{2} \delta \Delta X_{t-j} + \varepsilon_t$ With $\varepsilon_t \sim \left(0, \sigma_{\varepsilon}^2\right)$

Period $2005 - 2023$:

```
################################################
# Augmented Dickey-Fuller Test Unit Root Test #
################################################

Test regression drift


Call:
lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)

Residuals:
     Min       1Q   Median       3Q      Max
-13.9863  -0.5044   0.1667   0.9343  16.7022

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.5570     0.3854   1.445   0.1530
z.lag.1      -1.3138     0.1705  -7.707 7.47e-11 ***
z.diff.lag    0.2486     0.1175   2.115   0.0381 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.187 on 68 degrees of freedom
Multiple R-squared:  0.5552,	Adjusted R-squared:  0.5421
F-statistic: 42.44 on 2 and 68 DF,  p-value: 1.092e-12


Value of test-statistic is: -7.7066 29.6963

Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.51 -2.89 -2.58
phi1  6.70  4.71  3.86
```

Figure 3.11: Augmented DF Unit Root Test 2005-2023

Here $t_{\text{obs}} = -7.707 < t_{ADF} = -2.89$, so we reject the null hypothesis of unit root 's presence at the 5% threshold. Concerning the Model 2 test, $|t_{obs}| = 1.445 < |t_{ADF}| = 2.89$ so we don't reject the null hypothesis at the 5% threshold.Therefore, we must consider a model without constant.

Period $2005 - 2020$:

```
################################################
# Augmented Dickey-Fuller Test Unit Root Test #
################################################

Test regression drift


Call:
lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)

Residuals:
    Min      1Q  Median      3Q     Max
-2.9689 -0.4678  0.1257  0.7034  1.5592

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.1446     0.1402   1.031  0.30692
z.lag.1      -0.3861     0.1261  -3.061  0.00341 **
z.diff.lag   -0.1323     0.1345  -0.984  0.32959
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.004 on 55 degrees of freedom
Multiple R-squared:  0.2358,	Adjusted R-squared:  0.208
F-statistic: 8.485 on 2 and 55 DF,  p-value: 0.0006144


Value of test-statistic is: -3.061 4.6849

Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.51 -2.89 -2.58
phi1  6.70  4.71  3.86
```

Figure 3.12: Augmented DF Unit Root Test 2005-2020

In the ADF test, the observed t-statistic (t_{obs}) is -3.061 , which is less than the critical value at the 5% threshold ($t_{ADF} = -2.89$). Therefore, we reject the null hypothesis that a unit root is present. Regarding the Model 2 test, the absolute value of the observed t-statistic $|t_{\text{obs}}| = 1.031$, which is less than the absolute value of the critical t-statistic at the 5% threshold ($|t_{ADF}| = 2.89$). As a result, we don't reject the null hypothesis.

Model 1 (without constant and trend) $\Delta X_t = \rho X_{t-1} + \sum_{j=1}^{2} \delta \Delta X_{t-j} + \varepsilon_t$ With $\varepsilon_t \sim \left(0, \sigma_{\varepsilon}^2\right)$

Period $2005 - 2023$:

Here $t_{obs} = -7.511 < t_{ADF} = -1.95$ so we reject the null hypothesis of unit root's presence at the 5%

```
###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression none


Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)

Residuals:
     Min       1Q   Median       3Q      Max
-13.2440   0.0762   0.6709   1.4393  17.7565

Coefficients:
           Estimate Std. Error t value Pr(>|t|)
z.lag.1     -1.2665     0.1686  -7.511 1.57e-10 ***
z.diff.lag   0.2249     0.1173   1.917   0.0593 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.212 on 69 degrees of freedom
Multiple R-squared:  0.5415,    Adjusted R-squared:  0.5282
F-statistic: 40.75 on 2 and 69 DF,  p-value: 2.067e-12


Value of test-statistic is: -7.5109

Critical values for test statistics:
      1pct  5pct 10pct
tau1 -2.6 -1.95 -1.61
```

Figure 3.13: Augmented DF Unit Root Test 2005-2023

threshold. Our model is :

$$Y_t \sim I(0) \text{ with } \Delta Y_t = \rho Y_{t-1} + \varepsilon_t$$

<u>Period $2005 - 2020$:</u>

```
###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression none


Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)

Residuals:
     Min       1Q   Median       3Q      Max
-1.62069 -0.01881  0.00735  0.09777  0.66338

Coefficients:
           Estimate Std. Error t value Pr(>|t|)
z.lag.1    -0.03049    0.01948  -1.565    0.123
z.diff.lag  0.59153    0.10618   5.571  7.5e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2979 on 56 degrees of freedom
Multiple R-squared:  0.3696,    Adjusted R-squared:  0.3471
F-statistic: 16.42 on 2 and 56 DF,  p-value: 2.449e-06


Value of test-statistic is: -1.565

Critical values for test statistics:
      1pct  5pct 10pct
tau1 -2.6 -1.95 -1.61
```

Figure 3.14: Augmented DF Unit Root Test 2005-2020

$t_{obs} = -7.511 < t_{ADF} = -1.95$ so we reject the null hypothesis of unit root's presence at the 5% threshold. Our model is :

$$Y_t \sim I(0) \text{ with } \Delta Y_t = \rho Y_{t-1} + \varepsilon_t$$

3.2.2 KPSS Test

The results of the KPSS test on investment in the M2 are reported here (the KPSS test can only be done in the M2, which is the model closest to the M1 previously used in the sequential strategy).

<u>Period 2005 – 2023 :</u>

```
#######################
# KPSS Unit Root Test #
#######################

Test is of type: mu with 3 lags.

Value of test-statistic is: 0.0969

Critical value for a significance level of:
                10pct  5pct 2.5pct  1pct
critical values 0.347 0.463  0.574 0.739
```

Figure 3.15: KPSS Unit Root Test 2005-2023

The value of test statistic is equal to 0.0969 which is inferior to each critical values. Therefore we do not reject the null hypothesis of stationarity for every threshold.

<u>Period 2005 – 2020 :</u>

Same as before, the results of the KPSS test on investment in the M2 are reported here.

```
#######################
# KPSS Unit Root Test #
#######################

Test is of type: mu with 3 lags.

Value of test-statistic is: 0.1733

Critical value for a significance level of:
                10pct  5pct 2.5pct  1pct
critical values 0.347 0.463  0.574 0.739
```

Figure 3.16: KPSS Unit Root Test 2005-2020

For this test, the t-statistic ($LM_{obs} = 0.1733$) is inferior to each critical values. We can therefore confirm that the constant is significant and we can reject that the series is stationary.

Unlike the first two tests, the KPSS test accepts the null hypothesis that the series investment is stationary. We therefore have a doubt around these series in terms of stationarity. This can be explained by the presence of ruptures in this series. Indeed, the ADF test becomes fragile in the face of time series with ruptures. The test of Zivot and Andrews (1992) take into account the presence of ruptures. The Zivot and Andrews (1992) test takes into account the presence of breaks. Given the low number of observations in our series, we chose to discard this test. We then accept the hypothesis according to which all our series are integrated of order 1.

4 Stationarity

In order to determine the most appropriate estimation method, we checked in a the first step the stationarity of the series. Indeed, a series is said to be stationary if it does not contain a unit root. We also checked to see if the series are integrated into the same order should be used to consider performing a cointegration test. For example, we say that an X_t series is integrated of order 1, if it is stationary in first difference. The two time series studied here are DS processes without drifts to stationarize them,we will use the first difference.

4.1 Interest rate

Period 2005 – 2023 :

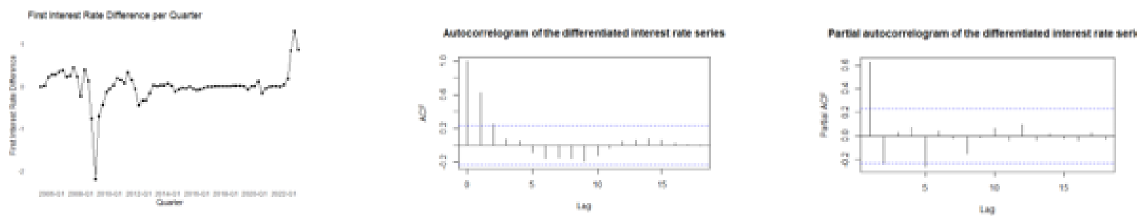


Figure 4.1: Plots, Simple and Partial Autocorrelograms of the differentiated series 2005 - 2023

Period 2005 – 2020 :

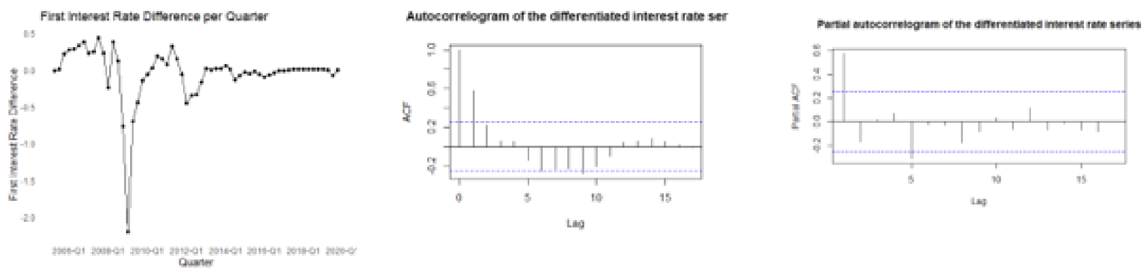


Figure 4.2: Plots, Simple and Partial Autocorrelograms of the differentiated series 2005 - 20020

With and without COVID, they seem to correspond to those of a stationary series, we repeat the tests carried out previously in order to test this hypothesis.

We verify that the time series is stationary with the ADF test with the Model 1 (M1)

4.1.1 Augmented Dickey-Fuller

Model 1 (without constant and trend)

Period 2005 – 2023 :

```
###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression none


Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)

Residuals:
     Min       1Q   Median       3Q      Max 
-1.56936 -0.03657  0.00156  0.05467  0.88151 

Coefficients:
           Estimate Std. Error t value Pr(>|t|)    
z.lag.1     -0.4200     0.1094  -3.841 0.000272 ***
z.diff.lag   0.2124     0.1299   1.635 0.106657    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3009 on 68 degrees of freedom
Multiple R-squared:  0.1799,	Adjusted R-squared:  0.1558 
F-statistic:  7.46 on 2 and 68 DF,  p-value: 0.001178


Value of test-statistic is: -3.8409 

Critical values for test statistics: 
      1pct  5pct 10pct
tau1 -2.6 -1.95 -1.61
```

Figure 4.3: Augmented DF Unit Root Test 2005-2023

$$t_{obs} = -3.841 < t_{ADF} = -1.95;$$

The critical test values of Dickey-Fuller at the 1% threshold, 5% and 10% are respectively $-2.6, -1.95$ and -1.61. We can therefore reject the hypothesis H_0 and conclude that the series is stationary.

Period 2005 – 2020 :

```
###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression none


Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)

Residuals:
     Min       1Q   Median       3Q      Max 
-1.65740 -0.04363  0.00076  0.03999  0.67597 

Coefficients:
           Estimate Std. Error t value Pr(>|t|)    
z.lag.1     -0.4857     0.1211  -4.011 0.000184 ***
z.diff.lag   0.1700     0.1329   1.279 0.206367    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3026 on 55 degrees of freedom
Multiple R-squared:  0.2304,	Adjusted R-squared:  0.2025 
F-statistic: 8.235 on 2 and 55 DF,  p-value: 0.000744


Value of test-statistic is: -4.011 

Critical values for test statistics: 
      1pct  5pct 10pct
tau1 -2.6 -1.95 -1.61
```

Figure 4.4: Augmented DF Unit Root Test 2005-2020

Same as before, $t_{obs} < t_{ADF}$, therefore we reject the null hypothesis of stationarity. The series is stationary.

4.2 Investment

Period 2005 – 2023 :

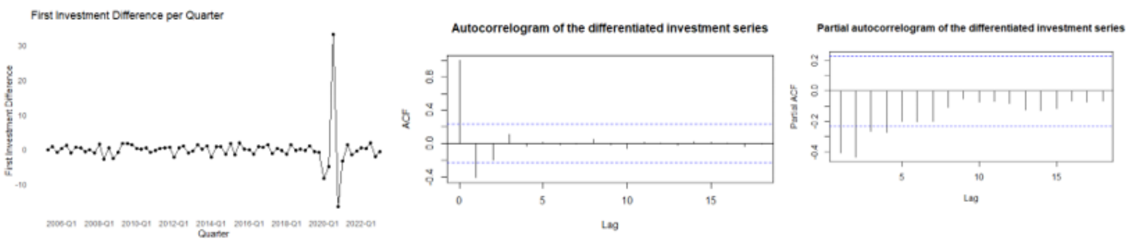


Figure 4.5: Plots, Simple and Partial Autocorrelograms of the differentiated series 2005 - 2023

Period 2005 – 2020 :

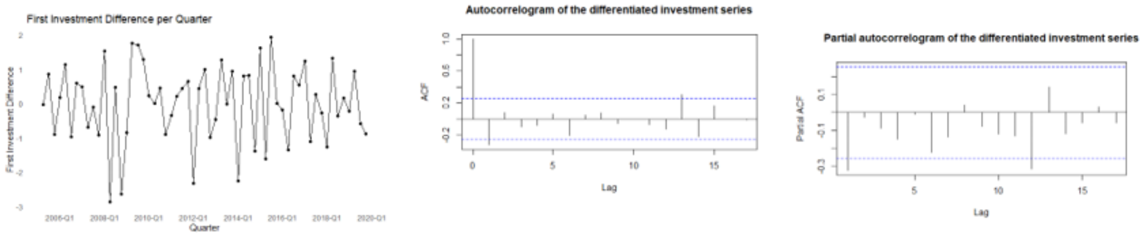


Figure 4.6: Plots, Simple and Partial Autocorrelograms of the differentiated series 2005 - 20020

With and without COVID, they seem to correspond to those of a stationary series, we repeat the tests carried out previously in order to test this hypothesis. We verify that the time series is stationary with the ADF test with the Model 1 (M1)

4.2.1 Augmented Dickey-Fuller

Model 1 (without constant and trend)

Period 2005 – 2023 :

```
###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression none


Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)

Residuals:
     Min       1Q   Median       3Q      Max 
-10.1471  -0.6674  -0.1183   0.5667  26.5577 

Coefficients:
           Estimate Std. Error t value Pr(>|t|)    
z.lag.1     -2.0273     0.1832 -11.063  < 2e-16 ***
z.diff.lag   0.4392     0.1092   4.022 0.000147 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.92 on 68 degrees of freedom
Multiple R-squared:  0.7611,	Adjusted R-squared:  0.754 
F-statistic: 108.3 on 2 and 68 DF,  p-value: < 2.2e-16


Value of test-statistic is: -11.0633 

Critical values for test statistics: 
      1pct  5pct 10pct
tau1 -2.6 -1.95 -1.61
```

Figure 4.7: Augmented DF Unit Root Test 2005-2023

$$t_{obs} < t_{ADF}$$

The critical test values of Dickey-Fuller at the 1% threshold, 5% and 10% are respectively $-2.6, -1.95$ and -1.61. We can therefore reject the hypothesis H_0 and conclude that the series is stationary.

Period $2005 - 2020$:

```
###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression none


Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)

Residuals:
     Min       1Q   Median       3Q      Max
-2.56068 -0.63750  0.04285  0.80673  2.26993

Coefficients:
           Estimate Std. Error t value Pr(>|t|)
z.lag.1    -1.36826    0.22118  -6.186 8.04e-08 ***
z.diff.lag  0.03135    0.13554   0.231    0.818
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.079 on 55 degrees of freedom
Multiple R-squared:  0.6634,    Adjusted R-squared:  0.6512
F-statistic: 54.21 on 2 and 55 DF,  p-value: 9.865e-14


Value of test-statistic is: -6.1861

Critical values for test statistics:
      1pct  5pct 10pct
tau1 -2.6 -1.95 -1.61
```

Figure 4.8: Augmented DF Unit Root Test 2005-2020

Same as previously, $t_{obs} < t_{ADF}$ The critical test values of Dickey-Fuller at the 1% threshold, 5% and 10% are respectively $-2.6, -1.95$ and -1.61. We can therefore reject the hypothesis H_0 and conclude that the series is stationary.

5 Modelling interest rates using an ARMA model

We decided to take the interest rate time series. As the interest rates influence directly the investments, we thought it would be more interesting to modelling interest rates using an ARMA model. First we will identify the autocorrelation, and continue by estimating the model. Then we will perform diagnostic tests. Finally we will able to produce forecasts for the time series.

5.1 Autocorrelation identification

We choose to base the model on a stable period before a significant disruption like the Covid-19 pandemic. Thus, this allows the model to capture more representative patterns and relationships in the data.

Period 2005 – 2020 :

From the simple and partial autocorrelograms, we identify an AR(1) process. We plot the information criteria of the ARMA(p,q) processes estimated with a constant for p = 0, . . . 4 and q = 0, . . . , 4.

```
> print(mat_aic)
         q=0      q=1      q=2      q=3      q=4
p=0 51.86426 33.28168 28.35065 29.12263 30.48363
p=1 30.17871 30.63767 28.98952 30.95163 32.19295
p=2 30.42442 32.41772 30.89388 32.98184 33.00848
p=3 32.41141 33.13142 32.09275 33.97756 34.99952
p=4 34.14774 34.20121 33.78995 32.00700 31.45912
> print(mat_bic)
         q=0      q=1      q=2      q=3      q=4
p=0 56.01934 39.51429 36.66080 39.51032 42.94885
p=1 36.41133 38.94782 39.37721 43.41686 46.73571
p=2 38.73457 42.80540 43.35911 47.52460 49.62878
p=3 42.79910 45.59665 46.63551 50.59786 53.69736
p=4 46.61296 48.74397 50.41025 50.70484 52.23450
```

Figure 5.1: AIC and BIC criteria 2005-2020

According to the AIC criteria, we identify a MA(2) contrary to the BIC criteria which identifies à AR(1) process.

```
Series: interest_rate_without_diff_df$Interest
ARIMA(0,0,2) with non-zero mean

Coefficients:
         ma1     ma2     mean
      0.7667  0.3824  -0.0415
s.e.  0.1449  0.1474   0.0790

sigma^2 = 0.08609:  log likelihood = -10.18
AIC=28.35   AICc=29.09   BIC=36.66

Training set error measures:
                        ME      RMSE       MAE      MPE     MAPE      MASE        ACF1
Training set -0.0003143314 0.2858553 0.1486511 249.2598 389.0136 0.9183547 -0.04562931
```

Figure 5.2: MA(2) with non-zero mean 2005-2020

Our model is written as :

$$X_t = m + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}$$

```
Series: interest_rate_without_diff_df$Interest
ARIMA(1,0,0) with non-zero mean

Coefficients:
         ar1     mean
      0.5695  -0.0417
s.e.  0.1044   0.0876

sigma^2 = 0.0907:  log likelihood = -12.09
AIC=30.18   AICc=30.62   BIC=36.41

Training set error measures:
                        ME      RMSE       MAE      MPE     MAPE      MASE     ACF1
Training set -0.0003293562 0.2960115 0.1366014 237.8072 355.5726 0.8439125 0.108631
```

Figure 5.3: AR(1) with non-zero mean 2005-2020

Our model is written as :

$$X_t = c + \phi_1 X_{t-1} + \varepsilon_t \qquad \text{AR(1)}$$

We need to calculate the constant $\hat{c} = \widehat{m}\left(1 - \widehat{\phi_1}\right)$ with the estimators of the coefficients and $\widehat{m}$ the estimator of the mean. We replace by the obtained values : $\hat{c} = -0.01796274$.

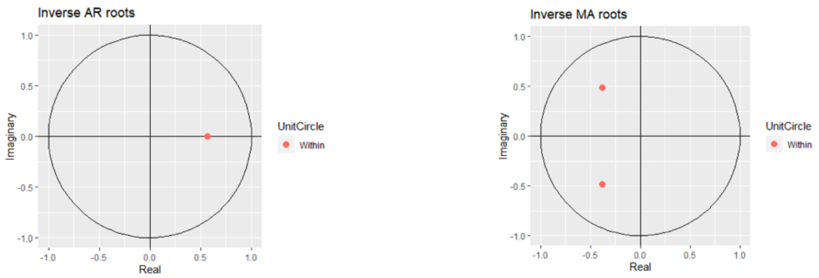


Figure 5.4: Inverse AR and MA roots in the complex unit disk 2005-2020

The inverse of the unit roots lies inside the complex unit disk, therefore the estimated processes are stationary.

5.2 Model estimation

5.2.1 AR(1)

```
z test of coefficients:

           Estimate Std. Error z value Pr(>|z|)
ar1        0.569524   0.104434  5.4535 4.94e-08 ***
intercept -0.041728   0.087582 -0.4764   0.6338
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 5.5: AR(1) Test of coefficients

The coefficient ϕ_1 is highly significant with a p-value of $4.94\text{e} - 08$, well below the usual statistical significance thresholds. The intercept is not significant with a p-value of 63.38%. It can be removed. Therefore, we repeat the coefficients test with a model without constant: $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t$ We add an AR(2) term and then a MA(1) term to check that they are not significant. After comparing models that include and exclude this constant, we conclude that models without a constant are preferable. This decision is supported by lower AIC and BIC values for models without a constant, indicating a better quality of fit compared to models with a constant. Consequently, we will exclude the constant in our further tests.

Let's test the significance of the AR(2) model coefficients : $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t$

```
z test of coefficients:

    Estimate Std. Error z value  Pr(>|z|)
ar1  0.67290    0.12704  5.2970 1.178e-07 ***
ar2 -0.16436    0.12618 -1.3026    0.1927
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 5.6: AR(2) Test of coefficients

The term AR(2) is not significant because its p-value is well above the usual statistical significance thresholds.

Let's test the significance of the ARMA(1,1) model coefficients : $X_t = \phi_1 X_{t-1} - \theta_1 \varepsilon_{t-1} + \varepsilon_t$

```
z test of coefficients:

    Estimate Std. Error z value Pr(>|z|)
ar1  0.42890    0.17034  2.5179  0.01181 *
ma1  0.23346    0.17102  1.3651  0.17221
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 5.7: ARMA(1,1) Test of coefficients

The term MA(1) is not significant. The addition of the MA(1) term also reduces the significance of the AR(1) term. The combination of AR and MA effects at a single lag does not capture the structure of the time series very well.

5.2.2 MA(2)

Remember that the AIC criteria suggest an MA(2) process. Let's test the significance of the MA(2) model coefficients: $X_t = m + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}$

```
Series: interest_rate_without_diff_df$Interest
ARIMA(0,0,2) with non-zero mean

Coefficients:
         ma1     ma2     mean
      0.7667  0.3824  -0.0415
s.e.  0.1449  0.1474   0.0790

sigma^2 = 0.08609:  log likelihood = -10.18
AIC=28.35   AICc=29.09   BIC=36.66

Training set error measures:
                        ME      RMSE       MAE      MPE     MAPE      MASE        ACF1
Training set -0.0003143314 0.2858553 0.1486511 249.2598 389.0136 0.9183547 -0.04562931
```

```
z test of coefficients:

           Estimate Std. Error z value  Pr(>|z|)
ma1        0.766704   0.144858  5.2928 1.204e-07 ***
ma2        0.382351   0.147380  2.5943  0.009478 **
intercept -0.041485   0.079032 -0.5249  0.599640
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 5.8: MA(2) with non-zero mean and test of coefficients

The coefficient θ_1 is highly significant with a p-value of $1.204\mathrm{e}-07$, well below the 1% thresholds which suggests a significant influence of the first lag of errors on the present value of the series. The same applies to the MA(2) coefficient, which has a value of 0.382351 and a p-value of less than 5%. The intercept was not statistically significant, with a p-value of 0.599640, exceeding the usual threshold of 0.05. It can be removed.

Let's test the significance of the MA(3) model coefficients : $X_t = \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \theta_3 \varepsilon_{t-3}$

```
z test of coefficients:

    Estimate Std. Error z value Pr(>|z|)
ma1  0.74348    0.12380  6.0053 1.91e-09 ***
ma2  0.32358    0.13903  2.3274  0.01994 *
ma3 -0.13110    0.12200 -1.0746  0.28255
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 5.9: MA(3) Test of coefficients

The term MA(3) is not significant because its p-value is well above the usual statistical significance thresholds. The addition of the MA(3) term also reduces the significance of the MA(2) term.

Let's test the significance of the ARMA(1,2) model coefficients : $X_t = \phi_1 X_{t-1} - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}$

```
z test of coefficients:

    Estimate Std. Error z value  Pr(>|z|)
ar1 -0.30373    0.22976 -1.3220 0.1861762
ma1  1.03649    0.19235  5.3885 7.105e-08 ***
ma2  0.56124    0.14793  3.7938 0.0001483 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 5.10: ARMA(1,2) Test of coefficients

The AR(1) coefficient is not statistically significant, as the associated p-value is 0.186762, which is above the common threshold of 0.05. This means that we cannot reject the null hypothesis that the AR(1) coefficient is equal to zero.

The coefficients for the MA(1) and MA(2) terms are 1.03649 and 0.56124 respectively, both with extremely low p-values (significantly below 0.001). This means that these coefficients are statistically significant, and that past errors have a significant influence on the current value of the series. The inclusion of the AR(1) term in this ARMA(1,2) model did not improve the significance of the model.

Consequently, our models are an AR(1) and MA(2), both without constant.

AR(1) without constant :

```
z test of coefficients:

    Estimate Std. Error z value  Pr(>|z|)
ar1  0.57527    0.10386   5.539 3.042e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
Series: interest_rate_without_diff_df$Interest
ARIMA(1,0,0) with zero mean

Coefficients:
         ar1
      0.5753
s.e.  0.1039

sigma^2 = 0.08946:  log likelihood = -12.2
AIC=28.4   AICc=28.62   BIC=32.56

Training set error measures:
                      ME      RMSE       MAE      MPE    MAPE      MASE      ACF1
Training set -0.01832104 0.2965481 0.1315144 42.15586 117.857 0.8124858 0.1037384
```

Figure 5.11: AR(1) without constant test of coefficients and with zero mean

MA(2) without constant :

```
z test of coefficients:

    Estimate Std. Error z value  Pr(>|z|)
ma1  0.77196    0.14387  5.3658 8.059e-08 ***
ma2  0.38758    0.14656  2.6446   0.00818 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
Series: interest_rate_without_diff_df$Interest
ARIMA(0,0,2) with zero mean

Coefficients:
         ma1     ma2
      0.7720  0.3876
s.e.  0.1439  0.1466

sigma^2 = 0.08496:  log likelihood = -10.31
AIC=26.62   AICc=27.06   BIC=32.86

Training set error measures:
                      ME      RMSE       MAE      MPE     MAPE      MASE        ACF1
Training set -0.01967737 0.2864888 0.1455913 40.00809 149.3572 0.8994519 -0.05003038
```

Figure 5.12: MA(2) without constant test of coefficients and with zero mean

The AIC criterion favors the MA(2) model without constant, with a value of 26.62, compared with the AR(1) model without constant, which has an AIC of 28.4. This difference indicates that the MA(2) model has a better fit. Therefore we will use a MA(2) model for the following diagnostic tests.

5.3 Diagnostic Tests

Let's verify that residuals:

- are not autocorrelated with Ljung-Box test on residuals
- are homoscedastic via the Engle test
- are normally distributed with the Jarque-Bera test.

MA(2):

$$X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}$$

$$\widehat{X}_t = \varepsilon_t - 0.7720 \times \varepsilon_{t-1} - 0.3876 \times \varepsilon_{t-2}$$

5.3.1 Ljung Box Test

Residuals are tested for autocorrelation. The test hypothesis is as follow:

$$\begin{cases} H_0: & \rho_i = 0, \forall i \in \{1, \dots, 10\} \\ H_1: & \exists i \in \{1, \dots, 10\} \text{ such that } \rho_i \neq 0 \end{cases}$$

The decision rule is that we reject H_0 if the Ljung-Box statistic is greater than the values of the distribution of the $\chi^2(8)$ law. For probabilities of $10\%, 5\%$ and 1%, the χ^2 law takes the values: $13.36, 15.51, 20.09$. It can be seen that $Q = 7.7607$ is strictly below these values. The null hypothesis of no autocorrelation of the regression residuals cannot therefore be rejected. Besides, the p-value is equal to $45.72\% > 5\%$.

```
        Ljung-Box test

data:  Residuals from ARIMA(0,0,2) with zero mean
Q* = 7.7607, df = 8, p-value = 0.4572

Model df: 2.   Total lags used: 10
```

Figure 5.13: Ljung-Box test

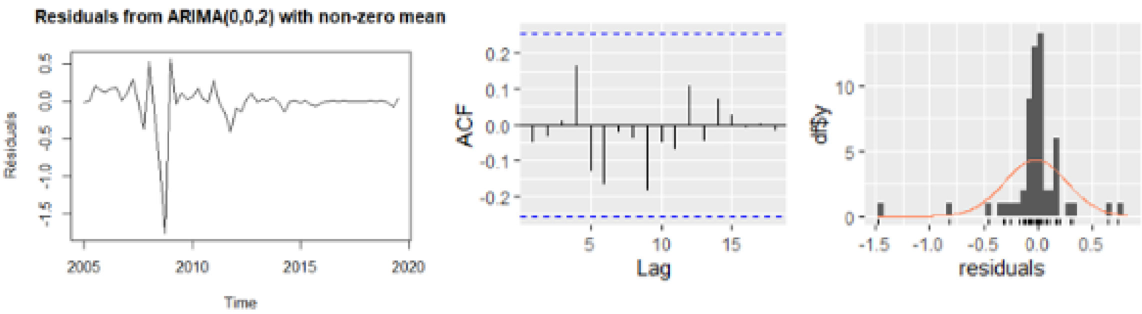


Figure 5.14: Plot of interest rate residuals

Examination of the autocorrelations reveals that individual autocorrelations are not statistically significant at the 5% level, even up to order 20. Notably, there is a clear presence of high residuals during the tumultuous period of the 2008-09 recession. ARIMA models have limitations when it comes to forecasting rare events, particularly those characterized by substantial changes in monetary policy, such as a pronounced interest rate cut orchestrated by the European Central Bank (ECB). These models assume stationarity, and may fail in the presence of unforeseen exogenous shocks or structural breaks. Consequently, more complex models are needed to effectively capture these anomalies, as illustrated by GARCH models, which are better equipped to encapsulate the inherent volatility of financial time series during periods of crisis.

5.3.2 Engle Test : ARCH effects test

We will now study this form of heteroscedasticity. The test hypothesis is as follows:

$$\begin{cases} H_0 : & \alpha_i = 0, \forall i \in \{1, \dots, 4\} \\ H_1 : & \overline{H_0} \end{cases}$$

The decision rule is that we reject H_0 if the Engle statistic is greater than the values of the distribution of the $\chi^2(4)$ law.

```
        ARCH LM-test; Null hypothesis: no ARCH effects

data:  MA2_no_intercept_without$residuals
Chi-squared = 13.823, df = 4, p-value = 0.007883
```

Figure 5.15: ARCH LM-test

$$LM_{obs} = 13.823 > \chi^2_{4,0.95} = 9.94$$

Besides, the p-value is equal to $0.8\% < 5\%$. For probabilities of $10\%, 5\%$ and 1%, the χ^2 law takes the values: $7.78, 9.94, 13.28$ which are strictly below the t-statistic. Therefore, the null hypothesis can be rejected, there are 4th-order ARCH effects in the residuals. This means that there is conditional heteroscedasticity in the residuals: the volatility of the errors varies over time, which is characteristic of periods of financial crisis, such as 2008.

5.3.3 Jarque-Bera Test

We do a Jarque-Bera test to verify the normality of the residuals. The test hypothesis is as follows :

$$\begin{cases} H_0 : & S = 0 \text{ and } K = 3 \quad \text{(normality of residuals)} \\ H_1 : & S \neq 0 \text{ and } K \neq 3 \quad \text{(lack of normality)} \end{cases}$$

The decision rule is that we reject H_0 if the Jarque Bera statistic is greater than the values of the distribution of the $\chi^2(2)$ law.

```
        Jarque Bera Test

data:  MA2_no_intercept_without$residuals
X-squared = 336.18, df = 2, p-value < 2.2e-16
```

Figure 5.16: Jarque-Bera test

$$JB_{obs} = 336.18 > \chi^2_{2,0.95} = 5.99$$

$JB_{\text{obs}} = 336.18$ is strictly superior than the values of the $\chi^2(2)$ law at the $10\%, 5\%$ and 1% thresholds (which are respectively $4.61, 5.99$ and 9.21). Moreover, the p-value is $2.2\text{e} - 16$. Therefore residuals are not normally distributed.

5.3.4 Adding a Dummy variable

The non-normality of the residuals is corrected by adding an indicator variable in Q4 2008 to capture the trough of the recession of 2008-09. The AIC and BIC criteria are minimized at 33.98 and 45.37 respectively.

```
Series: interest_rate_without_diff_df$Interest
Regression with ARIMA(0,0,2) errors

Coefficients:
         ma1     ma2  intercept    xreg
      1.1630  0.4309    -0.0548  0.9401
s.e.  0.2119  0.1477     0.0903  0.2651

sigma^2 = 0.07864:  log likelihood = -7.38
AIC=24.75   AICc=25.89   BIC=35.14

Training set error measures:
                        ME      RMSE      MAE      MPE     MAPE      MASE        ACF1
Training set -0.0009064273 0.2707552 0.155913 183.7591 412.6407 0.9632185 -0.05584201
> coeftest(MA2_DUM)

z test of coefficients:

           Estimate Std. Error z value  Pr(>|z|)
ma1        1.162963   0.211911  5.4880 4.066e-08 ***
ma2        0.430924   0.147657  2.9184 0.0035183 **
intercept -0.054803   0.090318 -0.6068 0.5440002
xreg       0.940058   0.265061  3.5466 0.0003903 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> checkresiduals(MA2_DUM,lag=10)

        Ljung-Box test

data:  Residuals from Regression with ARIMA(0,0,2) errors
Q* = 8.6995, df = 8, p-value = 0.3683

Model df: 2.   Total lags used: 10

> ArchTest(MA2_DUM$residuals,lags=4)

        ARCH LM-test; Null hypothesis: no ARCH effects

data:  MA2_DUM$residuals
Chi-squared = 7.0562, df = 4, p-value = 0.1329

> jarque.bera.test(MA2_DUM$residuals)

        Jarque Bera Test

data:  MA2_DUM$residuals
X-squared = 210.95, df = 2, p-value < 2.2e-16
```

Figure 5.17: Diagnostic tests with the dummy variable

Following the addition of a dummy variable to capture the rupture in 2008 , it is observed that the p-value of the Ljung-Box statistical tests is smaller, which means that the introduction of this variable has helped to better explain the variance in the data.

The t-statistic of the Engle test is smaller. In fact, it is equal to $LM_{obs} = 7.0562 < \chi^2_{4,0.95} = 9.94$. For probabilities of $10\%, 5\%$ and 1%, the χ^2 law takes the values: $7.78, 9.94, 13.28$ which are strictly below the t-statistic. Therefore, the null hypothesis cannot be rejected, there are not 4th-order ARCH effects in the residuals. This means that there is no conditional heteroscedasticity in the residuals. The addition of the dummy variable has corrected the non-normality of the residuals and eliminated the effects of conditional heteroscedasticity, indicating a significant improvement in model specification. This means that some of the residual volatility that seemed to change over time was in fact due to the impact of the recession, which is now captured by the dummy.

For the Jarque-Bera test, there are no significant changes. The reduction in AIC and BIC after the addition of the dummy also indicates an overall improvement in model quality.

In conclusion, the inclusion of the dummy variable to account for the rupture in 2008 has led to significant improvements in the model.

6 Forecast

6.1 Forecast without the dummy model

Finally, we use the MA(2) process estimated from Q1 2005 to Q4 2019 to forecast the period from Q1 2020 to Q4 2020.

	Point Forecast	Lo 95	Hi 95
60	0.01393785	-0.5573361	0.5852118
61	0.02096791	-0.7007207	0.7426565
62	0.00000000	-0.7548893	0.7548893
63	0.00000000	-0.7548893	0.7548893

Figure 6.1: MA(2) process estimation without the dummy

- Q1 2020 Forecast: 0.013 vs. Actual: -0.406
- Q2 2020 Forecast : 0.021 vs. Actual : -0.300
- Q3 2020 Forecast : 0.0 vs. Actual : -0.4717
- Q4 2020 Forecast : 0.0 vs. Actual : -0.5227

The previsions are way above than the observed values.

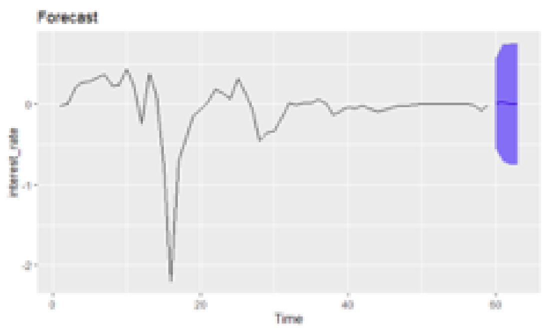


Figure 6.2: MA(2) forecast without the dummy

The confidence intervals are fairly wide, indicating considerable uncertainty around the forecasts. The MA(2) model appears not to have accurately captured the trend in interest rates over the forecast period. We will redo the forecast with the dummy to see if volatility justifies this discrepancy.

6.2 Forecast with the dummy model

The forecast is revised by incorporating a dummy variable for Q4 2008, aiming to enhance the forecast's accuracy and reliability.

```
   Point Forecast       Lo 95     Hi 95
60     0.03862299 -0.5110059 0.5882519
61    -0.01330920 -0.8563195 0.8297011
62    -0.05480256 -0.9304528 0.8208477
63    -0.05480256 -0.9304528 0.8208477
```

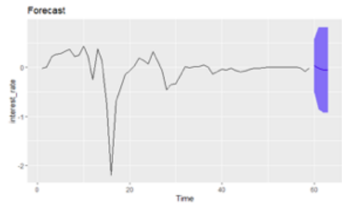


Figure 6.3: MA(2) process estimation and forecast the dummy

The addition of the dummy variable seems to have improved the quality of the forecast compared with previous estimates, which did not include the impact of the 2008 recession. The forecast becomes negative, which corresponds to the direction of the real rates you provided, but not to the amplitude. The confidence intervals are wide, indicating significant uncertainty in the forecast.

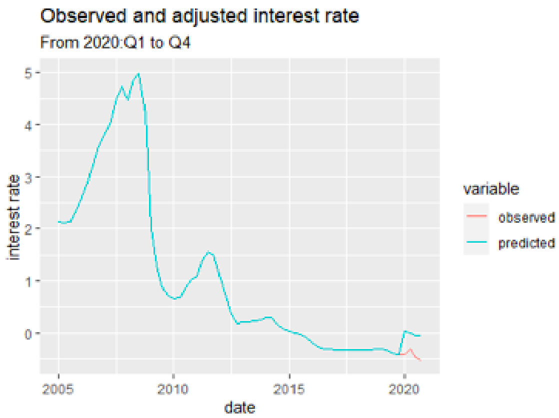


Figure 6.4: Observed and adjusted interest rate

```
> sprintf("MAE out of sampe %f", mae(df_forc$observed,df_forc$forecast))
[1] "MAE out of sampe 0.025255"
> sprintf("RMSE out of sampe %f", mae(df_forc$observed,df_forc$forecast))
[1] "RMSE out of sampe 0.025255"
```

Figure 6.5: Out-of-sample MAE and RMSE

The model tends to overestimate the future evolution of the interest rate in relation to its observed values. Out-of-sample MAE and RMSE are higher (=0.253 both) than their values in the training set. The out-of-sample values allow us to assess more precisely the predictive quality of the estimated model.

6.3 Forecast with dummies on the Great Recession

As already mentioned and presented during this paper, the period from Q4 2007 to Q2 2009 was experienced as a real economic and financial distress in France. The subprime crisis of 2008, triggered by risky mortgage lending in the US (La finance pour tous, 2008), had less severe repercussions in France due to the strength of the public sector and social safety nets. The French economy faced challenges such as slowing growth, rising public debt and high unemployment, but these were cushioned by a solid banking system and government stimulus measures (UK Essays, 2008). The economic slowdown was cushioned by domestic consumption and infrastructure investment, which prevented the crisis from triggering as deep a recession as in other countries.

We define individual indicator variables for the quarters of the period and add them to the MA(2) model.

```
Series: interest_rate_without$Interest[1:60]
Regression with ARIMA(0,1,2) errors

Coefficients:
         ma1     ma2  DU_07Q4  DU_08Q1  DU_08Q2  DU_08Q3  DU_08Q4  DU_09Q1  DU_09Q2
      1.1999  0.8296   0.0365   0.0980   1.1341   1.9112   1.7999   0.2528   0.1262
s.e.  0.0832  0.0865   0.1114   0.2455   0.3506   0.3790   0.3498   0.2432   0.1093

sigma^2 = 0.01651:  log likelihood = 40.77
AIC=-61.55   AICc=-56.96   BIC=-40.77

Training set error measures:
                      ME      RMSE        MAE      MPE     MAPE      MASE      ACF1
Training set -0.01319897 0.1173118 0.08929286 23.32615 48.83591 0.4875267 0.2217926
```

```
z test of coefficients:

        Estimate Std. Error z value  Pr(>|z|)    
ma1     1.199906   0.083155 14.4297 < 2.2e-16 ***
ma2     0.829552   0.086544  9.5853 < 2.2e-16 ***
DU_07Q4 0.036464   0.111402  0.3273  0.743425    
DU_08Q1 0.097958   0.245520  0.3990  0.689907    
DU_08Q2 1.134128   0.350615  3.2347  0.001218 ** 
DU_08Q3 1.911229   0.378981  5.0431 4.581e-07 ***
DU_08Q4 1.799924   0.349806  5.1455 2.668e-07 ***
DU_09Q1 0.252827   0.243201  1.0396  0.298536    
DU_09Q2 0.126171   0.109296  1.1544  0.248339    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 6.6: Regression of the MA(2) model and test of coefficients from Q4 2007 to Q2 2009

The MA(2) coefficients are significant and their order of magnitude is close to that of the model without indicators. Only the Q2 2008, Q3 2008 and Q4 2008 dummies are significant at 1%. The AIC and BIC criteria are much smaller than for the regression without dummies. They go from 26.62 to -61.55 and 32.86 to -40.77 respectively.

We only select dummies from Q2 2008 to Q4 2008 for the next model.

```
Series: interest_rate_without$Interest[1:60]
Regression with ARIMA(0,1,2) errors

Coefficients:
         ma1     ma2  DU_08Q2  DU_08Q3  DU_08Q4
      1.2109  0.8474   0.9634   1.6754   1.4948
s.e.  0.0801  0.0949   0.0888   0.1208   0.0870

sigma^2 = 0.01558:  log likelihood = 40.11
AIC=-68.22   AICc=-66.61   BIC=-55.76

Training set error measures:
                      ME      RMSE        MAE     MPE     MAPE      MASE      ACF1
Training set -0.01301801 0.1184318 0.09018664 22.6679 48.08848 0.4924066 0.2073901
```

```
z test of coefficients:

        Estimate Std. Error z value  Pr(>|z|)    
ma1     1.210895   0.080063 15.1243 < 2.2e-16 ***
ma2     0.847433   0.094863  8.9332 < 2.2e-16 ***
DU_08Q2 0.963373   0.088778 10.8515 < 2.2e-16 ***
DU_08Q3 1.675351   0.120754 13.8741 < 2.2e-16 ***
DU_08Q4 1.494774   0.087031 17.1751 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 6.7: Regression of the MA(2) model and test of coefficients from Q2 2008 to Q4 2008

All variables are highly significant. Their p-values are all below 0.1%. The AIC and BIC criteria have been further reduced compared with the previous model. The RMSE and MAE reported are the values for the training set. We will compare them with their out-of-sample values by calculating forecasts for observations not used to estimate the model.

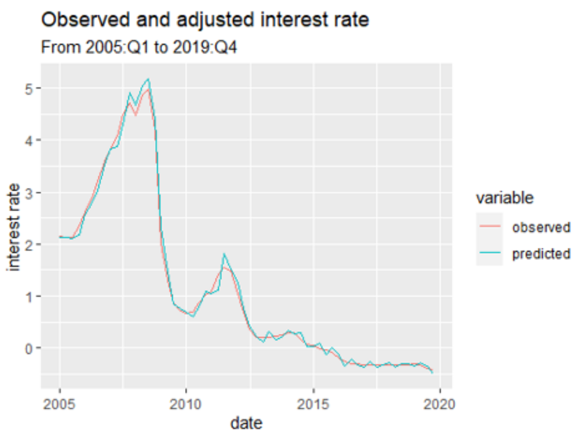
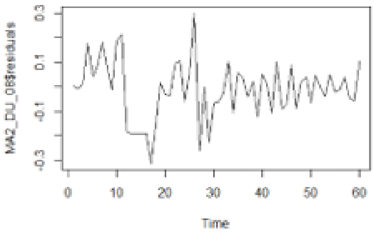


Figure 6.8: Observed and adjusted interest rate from Q1 2005 to Q4 2019

The observed values and the values predicted by the model are shown below. We can see that the two curves are fairly close, which suggests that the quality of the fit is not too bad.




```
        Jarque Bera Test

data:  MA2_DU_08$residuals
X-squared = 0.30936, df = 2, p-value = 0.8567

        ARCH LM-test; Null hypothesis: no ARCH effects

data:  MA2_DU_08$residuals
Chi-squared = 12.468, df = 4, p-value = 0.01419
```

Figure 6.9: Jarque Bera and ARCH LM Tests on the model with dummies from Q2 2008 to Q4 2008

Residuals appear to fluctuate around zero, which is a good sign for a fitted model. There are no obvious signs of cluster volatility (heteroscedastic volatility). This could indicate that the residuals have a constant variance. The stability of the residuals after the introduction of the dummies could indicate that the impact of the crisis has been effectively incorporated into the model. However, to be on the safe side, we performed the Jarque Bera and Engle diagnostic tests.

$$LM_{obs} = 12.468 > \chi^2_{4,0.95} = 9.94$$

Besides, the p-value is equal to 1.42% < 5. Therefore, the null hypothesis can be rejected, there are 4th-order ARCH effects in the residuals. ARCH effects appears when adding significant dummies variables representing the 2008 financial crisis. This indicates that the conditional variance of the residuals changes with the introduction of this variable.

We do not reject the null hypothesis of a normal distribution (Jarque and Bera test) for an error of first kind of 1%.

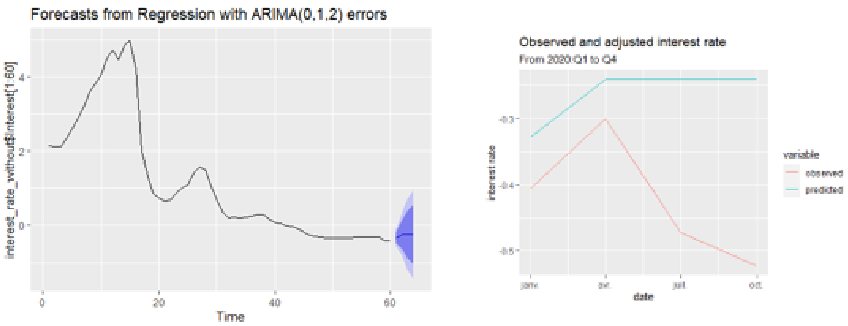


Figure 6.10: Forecasts from Regression with MA(2) and observed and adjusted interest rate from Q1 to Q4 2020

Surprisingly, the RMSE and MAE are greater in forecasts obtained from the model without dummies. However, as with the previous model, predicted values are higher than observed values and the confidence interval is reduced.

7 VAR model

We use a VAR model to study the relationship between short term interest rates and GFCF (investment) in France from Q1 2005 to Q1 2023.

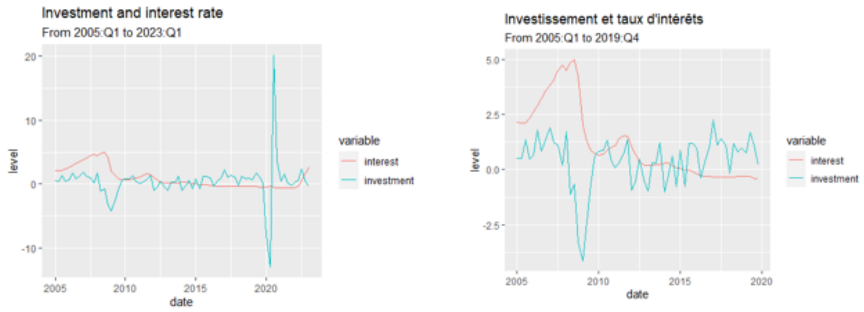
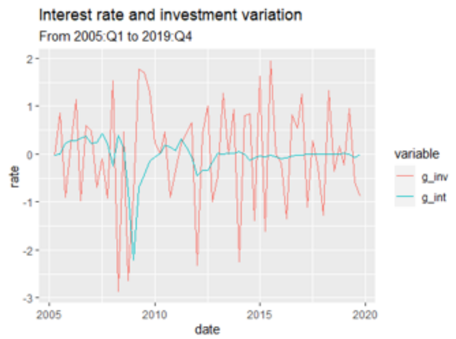


Figure 7.1: Investment and Interest rate in France from Q1 2005 to Q1 2023 and from Q1 2005 to Q4 2019

Due to the specificity of the covid period, we limit ourselves to the period Q1 2005 to Q4 2019 for the following test.

Stopping the analysis before the COVID-19 pandemic avoids distortions resulting from structural breaks and anomalies, facilitating a clearer understanding of the underlying economic relationships.

7.1 First difference




```
> summary(df_g_short[c("g_int","g_inv")])
     g_int                g_inv
 Min.   :-2.862948   Min.   :-2.2029670
 1st Qu.:-0.864628   1st Qu.:-0.0692666
 Median : 0.163474   Median : 0.0003334
 Mean   :-0.005004   Mean   :-0.0431062
 3rd Qu.: 0.819927   3rd Qu.: 0.0970000
 Max.   : 1.934798   Max.   : 0.4356670
> sapply(df_g_short[c("g_int","g_inv")],sd)
    g_int     g_inv
1.1183445 0.3661318

        Pearson's product-moment correlation

data:  df_g_short$g_int and df_g_short$g_inv
t = 0.24234, df = 57, p-value = 0.8094
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.2258559  0.2858160
sample estimates:
       cor
0.03208191
```

Figure 7.2: Plots, descriptive statistics and correlation

The average of the 2 series is close to 0, showing no clear upward or downward trend. This shows a lack of trend in the relative changes in interest rates and investment over time. However, for a more in-depth analysis, it would be useful to visualise the data using graphs and to carry out stationarity tests.

The p-value associated with the correlation test is 80.94%. This means that the correlation observed between these two series is not statistically significant at the 95% confidence level, with a correlation close to 0.

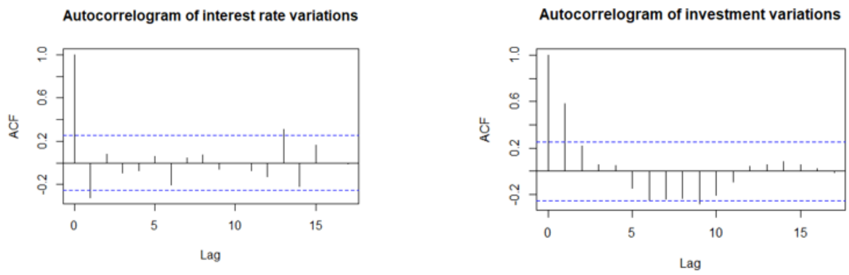


Figure 7.3: Autocorrelograms of the differences series

7.2 VAR modelling

We choose a maximum number of lags equal to 8 (corresponding to two years of lag for quarterly data) and we estimate a VAR model with a constant (the average of the s is not null).

```
               1          2          3           4           5          6          7          8
AIC(n) -2.092336 -2.1708442 -2.2888366 -2.32233703 -2.33482318 -2.3090718 -2.2562095 -2.2344516
HQ(n)  -2.005488 -2.0260976 -2.0861914 -2.06179313 -2.01638064 -1.9327306 -1.8219697 -1.7423132
SC(n)  -1.865063 -1.7920549 -1.7585316 -1.64051622 -1.50148663 -1.3242195 -1.1198415 -0.9465679
FPE(n)  0.123432  0.1142255  0.1017385  0.09877907  0.09816963  0.1016636  0.1085648  0.1128732
> pselect$selection # Nombre de retards optimal selon les critères de sélection
AIC(n)  HQ(n)  SC(n) FPE(n) 
     5      3      1      5 
```

Figure 7.4: Selection of the optimum number of lags

The optimum number of lags varies according to the criterion: AIC recommends 4 lags, the HQ criterion recommends 2 lags, the SC criterion recommends 1 lags, the optimal number of lags is nevertheless less than lag.max. The AIC criterion is known to overestimate the number of delays. We decide to estimate the VAR model with 1 lag.

7.2.1 Estimation of a VAR(1) model

```
VAR Estimation Results:
=========================
Endogenous variables: g_int, g_inv 
Deterministic variables: const 
Sample size: 58 
Log Likelihood: -93.004 
Roots of the characteristic polynomial:
0.5348 0.2839
Call:
VAR(y = df_g_short[, c("g_int", "g_inv")], type = "const", lag.max = 1)


Estimation results for equation g_int: 
======================================
g_int = g_int.l1 + g_inv.l1 + const 

         Estimate Std. Error t value Pr(>|t|)  
g_int.l1 -0.32131    0.12645  -2.541   0.0139 *
g_inv.l1 -0.48754    0.38419  -1.269   0.2098  
const    -0.02276    0.14158  -0.161   0.8729  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 1.071 on 55 degrees of freedom
Multiple R-Squared: 0.131,	Adjusted R-squared: 0.09944 
F-statistic: 4.147 on 2 and 55 DF,  p-value: 0.02101 

Estimation results for equation g_inv: 
======================================
g_inv = g_int.l1 + g_inv.l1 + const 

         Estimate Std. Error t value Pr(>|t|)    
g_int.l1  0.06578    0.03511   1.874   0.0663 .  
g_inv.l1  0.57224    0.10666   5.365 1.66e-06 ***
const    -0.01921    0.03931  -0.489   0.6269    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 0.2972 on 55 degrees of freedom
Multiple R-Squared: 0.3751,	Adjusted R-squared: 0.3523 
F-statistic:  16.5 on 2 and 55 DF,  p-value: 2.429e-06 



Covariance matrix of residuals:
        g_int   g_inv
g_int 1.14608 0.08399
g_inv 0.08399 0.08833

Correlation matrix of residuals:
      g_int g_inv
g_int 1.000 0.264
g_inv 0.264 1.000
```

Figure 7.5: VAR(1) estimation

```
> roots(var.1lag)
[1] 0.5347754 0.2838522
```

Figure 7.6: VAR(1) inverse of unit root values

g_int.li and g_inv.li represent g_int and g_inv respectively with a lag of i periods.

Interest rate equation g_int :

- The coefficient of $g\_int.l1$ is significant and negative for an error of first kind of 1%.
- We tend to observe an alternating acceleration and deceleration in the quarterly rate of interest rates.
- The coefficient of $g\_inv.l1$ is negative and not significant.
- The intercept is not significant.
- The adjusted R^2 is equal to 10%.

Investment rate equation g_inv :

- The coefficient of $g\_inv.l1$ is positive and significant for a first order kind error of 10%; an acceleration in investment rates have the effect of accelerating interest rates.
- The coefficient of $g\_inv.l1$ is positive and significant for a first order kind error of 1%; investment rates are positively correlated with its past values.
- The adjusted R^2 is equal to 35%: the explanatory power of the regression is greater than for g_int.

The inverse of the unit roots have a modulus strictly less than 1: the stationarity condition is therefore satisfied. We are going to test a model with two lags and if the latter improves the fit statistics compared with the one-lag model (such as a higher adjusted R^2 and better significance of the coefficients), this could justify the choice of two lags.

7.2.2 Estimation of a VAR(2) model

```
VAR Estimation Results:
=========================
Endogenous variables: g_int, g_inv
Deterministic variables: const
Sample size: 57
Log Likelihood: -85.909
Roots of the characteristic polynomial:
0.5754 0.5754 0.4158 0.4158
Call:
VAR(y = df_g_short[, c("g_int", "g_inv")], type = "const", lag.max = 2)


Estimation results for equation g_int:
======================================
g_int = g_int.l1 + g_inv.l1 + g_int.l2 + g_inv.l2 + const

         Estimate Std. Error t value Pr(>|t|)
g_int.l1 -0.44840    0.13111  -3.420  0.00123 **
g_inv.l1  0.47646    0.46818   1.018  0.31354
g_int.l2 -0.08646    0.13182  -0.656  0.51480
g_inv.l2 -1.54761    0.46060  -3.360  0.00147 **
const    -0.05899    0.13264  -0.445  0.65834
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 0.9918 on 52 degrees of freedom
Multiple R-Squared: 0.2873,     Adjusted R-squared: 0.2325
F-statistic: 5.242 on 4 and 52 DF,  p-value: 0.001267
```

```
Estimation results for equation g_inv:
======================================
g_inv = g_int.l1 + g_inv.l1 + g_int.l2 + g_inv.l2 + const

         Estimate Std. Error t value Pr(>|t|)
g_int.l1  0.07031    0.03986   1.764   0.0836 .
g_inv.l1  0.60779    0.14233   4.270 8.31e-05 ***
g_int.l2  0.03147    0.04007   0.785   0.4358
g_inv.l2 -0.09874    0.14002  -0.705   0.4838
const    -0.02321    0.04032  -0.576   0.5673
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 0.3015 on 52 degrees of freedom
Multiple R-Squared: 0.3918,     Adjusted R-squared: 0.345
F-statistic: 8.374 on 4 and 52 DF,  p-value: 2.719e-05



Covariance matrix of residuals:
        g_int   g_inv
g_int 0.98363 0.07409
g_inv 0.07409 0.09090

Correlation matrix of residuals:
       g_int  g_inv
g_int 1.0000 0.2478
g_inv 0.2478 1.0000
```

Figure 7.7: VAR(2) estimation

Interest rate equation g_int :

- The coefficient of $g\_int.l1$ is significant and negative for a first order kind of 1%: interest rates are negatively correlated with its past value.
- The other coefficients are not statistically significant.
- The coefficient for $g\_inv.l2$ is negative and significant: interest rates depend negatively on investments delayed by two quarters.

- The coefficient for $g\_inv.l1$ is negative and not significant.
- The constant is non significant.
- The adjusted R^2 for g_int is 23%, which is an improvement on the adjusted R^2 of 10% in the previous model, indicating a better fit of the model with two lags than with a single lag.

Investment rate equation g_inv :

- The coefficient of $g\_inv.l1$ is positive and significant: an acceleration in investment rates have the effect of accelerating interest rates.
- The coefficient of $g\_inv.l1$ is positive and significant for a first order kind of 1%: investment rate positively correlated with its past value.
- The other coefficients, including those for two-period lags, are not significant.
- The adjusted R^2 does not change.

```
> roots(var.2lag)
[1] 0.5754251 0.5754251 0.4157974 0.4157974
```

Figure 7.8: VAR(2) inverse of unit root values

The inverses of the unit roots have a modulus strictly less than 1 : the stationarity condition is therefore satisfied. The adjusted R2 improves from 10% to 23% for the interest rates and remains equal for investment rates.

We test a model with 3 lags and if it improves the fit statistics compared with the 2 lag model we select 3 lags, otherwise we choose a VAR(2) model.

7.2.3 Estimation of a VAR(3) model

```
VAR Estimation Results:
=========================
Endogenous variables: g_int, g_inv
Deterministic variables: const
Sample size: 56
Log Likelihood: -77.659
Roots of the characteristic polynomial:
0.7565 0.7565 0.7392 0.7392 0.4996 0.4996
Call:
VAR(y = df_g_short[, c("g_int", "g_inv")], type = "const", lag.max = 3)


Estimation results for equation g_int:
======================================
g_int = g_int.l1 + g_inv.l1 + g_int.l2 + g_inv.l2 + g_int.l3 + g_inv.l3 + const

         Estimate Std. Error t value Pr(>|t|)
g_int.l1 -0.54094    0.14482  -3.735  0.00049 ***
g_inv.l1  0.48591    0.47205   1.029  0.30837
g_int.l2 -0.18734    0.15389  -1.217  0.22928
g_inv.l2 -1.06777    0.54503  -1.959  0.05581 .
g_int.l3 -0.04755    0.13399  -0.355  0.72420
g_inv.l3 -0.84628    0.50844  -1.664  0.10240
const    -0.06424    0.13423  -0.479  0.63437
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 0.9917 on 49 degrees of freedom
Multiple R-Squared: 0.3212,     Adjusted R-squared: 0.2381
F-statistic: 3.865 on 6 and 49 DF,  p-value: 0.003097
```

```
Estimation results for equation g_inv:
======================================
g_inv = g_int.l1 + g_inv.l1 + g_int.l2 + g_inv.l2 + g_int.l3 + g_inv.l3 + const

         Estimate Std. Error t value Pr(>|t|)
g_int.l1  0.09549    0.04183   2.283  0.02681 *
g_inv.l1  0.56327    0.13635   4.131  0.00014 ***
g_int.l2  0.09249    0.04445   2.081  0.04269 *
g_inv.l2 -0.26900    0.15743  -1.709  0.09383 .
g_int.l3  0.10504    0.03870   2.714  0.00915 **
g_inv.l3  0.22767    0.14686   1.550  0.12750
const    -0.02735    0.03877  -0.705  0.48396
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 0.2864 on 49 degrees of freedom
Multiple R-Squared: 0.4782,     Adjusted R-squared: 0.4143
F-statistic: 7.484 on 6 and 49 DF,  p-value: 1.01e-05



Covariance matrix of residuals:
        g_int   g_inv
g_int 0.98346 0.09476
g_inv 0.09476 0.08205

Correlation matrix of residuals:
       g_int  g_inv
g_int 1.0000 0.3336
g_inv 0.3336 1.0000
```

Figure 7.9: VAR(3) estimation

Interest rate equation g_int :

- The coefficient for $g\_int.l1$ is negative and significant, indicating negative autocorrelation, meaning that higher interest rates in the previous quarter tend to be followed by lower rates in the current period.

- The other coefficients for g_int and g_inv at different lags are not significant, with the exception of $g\_int.l2$ which has a p-value just at the significance level.
- The adjusted R^2 for this equation is 24%, indicating that the model explains around 24% of the variation in interest rates.

Investment rate equation g_inv :

- The coefficient for $g\_inv.l1$ is significant and positive, suggesting that investment tends to persist, i.e., a high rate of investment in the previous quarter is followed by a high rate in the next quarter.
- The coefficients for $g\_int.l1$ and $g\_int.l2$ are positive and significant, suggesting that higher interest rates in the two previous periods have a positive effect on the investment rates. In an economic climate, "companies can intensify their investment activities as a preventive strategy against anticipated increases in borrowing costs." according to Agnès Bénassy-Quéré, Laurence Boone and Virginie Coudert in their book "Les taux d'intérêt" (p.11-32 - Agnès Bénassy-Quéré, 2014). This foresight can be seen in the positive correlation between current investment levels and future interest rates, highlighting a proactive response to expected economic conditions, because the firms try to anticipate the potential rise in interest rates.
- The adjusted R^2 is 41%, which is an improvement on the VAR(2) model and means that the model explains around 41% of the variation in investment rates.

```
> roots(var.3lag)
[1] 0.7565018 0.7565018 0.7392088 0.7392088 0.4996492 0.4996492
```

Figure 7.10: VAR(3) inverse of unit root values

The inverse of the unit roots are all less than 1, indicating that the model is stationary. We test a model with 4 lags and if it improves the fit statistics compared to the 3 lag model we select 4 lags, otherwise we choose a VAR(3) model.

7.2.4 Estimation of a VAR(4) model

```
VAR Estimation Results:
=========================
Endogenous variables: g_int, g_inv
Deterministic variables: const
Sample size: 55
Log Likelihood: -71.988
Roots of the characteristic polynomial:
0.8056 0.8056 0.7264 0.6707 0.6707 0.5938 0.5938 0.5825
Call:
VAR(y = df_g_short[, c("g_int", "g_inv")], type = "const", lag.max = 4)


Estimation results for equation g_int:
======================================
g_int = g_int.l1 + g_inv.l1 + g_int.l2 + g_inv.l2 + g_int.l3 + g_inv.l3 + g_int.l4 + g_inv.l4 + const

         Estimate Std. Error t value Pr(>|t|)    
g_int.l1 -0.58781    0.15665  -3.752  0.00049 ***
g_inv.l1  0.71789    0.53747   1.336  0.18823    
g_int.l2 -0.27251    0.18620  -1.463  0.15014    
g_inv.l2 -1.13785    0.56152  -2.026  0.04855 *  
g_int.l3 -0.14544    0.17039  -0.854  0.39777    
g_inv.l3 -0.60968    0.58495  -1.042  0.30272    
g_int.l4 -0.12120    0.14911  -0.813  0.42052    
g_inv.l4 -0.42865    0.55902  -0.767  0.44713    
const    -0.05864    0.13902  -0.422  0.67512    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Estimation results for equation g_inv:
======================================
g_inv = g_int.l1 + g_inv.l1 + g_int.l2 + g_inv.l2 + g_int.l3 + g_inv.l3 + g_int.l4 + g_inv.l4 + const

         Estimate Std. Error t value Pr(>|t|)   
g_int.l1  0.12411    0.04292   2.892  0.00583 **
g_inv.l1  0.49971    0.14725   3.394  0.00143 **
g_int.l2  0.14625    0.05101   2.867  0.00623 **
g_inv.l2 -0.23584    0.15384  -1.533  0.13212   
g_int.l3  0.14031    0.04668   3.006  0.00428 **
g_inv.l3  0.08418    0.16026   0.525  0.60189   
g_int.l4 -0.00591    0.04085  -0.145  0.88561   
g_inv.l4  0.33839    0.15315   2.209  0.03216 * 
const    -0.02446    0.03809  -0.642  0.52387   
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 0.2775 on 46 degrees of freedom
Multiple R-Squared: 0.5341,	Adjusted R-squared: 0.453 
F-statistic: 6.591 on 8 and 46 DF,  p-value: 1.04e-05



Covariance matrix of residuals:
       g_int  g_inv
g_int 1.0258 0.1087
g_inv 0.1087 0.0770

Correlation matrix of residuals:
       g_int  g_inv
g_int 1.0000 0.3869
g_inv 0.3869 1.0000
```

Figure 7.11: VAR(4) estimation

Interest rate equation g_int :

- The coefficient for $g\_int.l1$ is negative and significant, indicating negative autocorrelation.
- The other coefficients for g_int and g_inv at different lags are not significant, with the exception of $g\_int.l2$ which has a p-value just at the significance level.
- The adjusted R^2 for this equation is 24%, indicating that the model explains around 23.81% of the variation in interest rates.

Investment rate equation g_inv :

- The coefficient for $g\_inv.l1$ is significant and positive, indicating positive autocorrelation.
- The coefficients for $g\_int.l1$ and $g\_int.l2$ are positive and significant, suggesting that higher interest rates in the previous two periods have a positive effect on investment rates.
- The adjusted R^2 is 41%, which is an improvement on the VAR(3) model and means that the model explains around 41% of the variation in investment.

```
> roots(var.4lag)
[1] 0.8056144 0.8056144 0.7263943 0.6706608 0.6706608 0.5937659 0.5937659 0.5824695
```

Figure 7.12: VAR(4) inverse of unit root values

The inverse of the unit roots are all less than 1, indicating that the model is stationary. We test a model with 5 lags and if it improves the fit statistics compared to the 4 lag model we select 5 lags, otherwise we choose a VAR(4) model.

7.2.5 Estimation of a VAR(5) model

```
VAR Estimation Results:
=========================
Endogenous variables: g_int, g_inv
Deterministic variables: const
Sample size: 54
Log Likelihood: -66.499
Roots of the characteristic polynomial:
0.8618 0.8618 0.7902 0.7902 0.7452 0.7452 0.6686 0.6686 0.5243 0.5243
Call:
VAR(y = df_g_short[, c("g_int", "g_inv")], type = "const", lag.max = 5)


Estimation results for equation g_int:
======================================
g_int = g_int.l1 + g_inv.l1 + g_int.l2 + g_inv.l2 + g_int.l3 + g_inv.l3 + g_int.l4 + g_inv.l4 + g_int.l5 +
g_inv.l5 + const

         Estimate Std. Error t value Pr(>|t|)
g_int.l1  -0.6118     0.1578  -3.877 0.000357 ***
g_inv.l1   0.9879     0.5712   1.730 0.090892 .
g_int.l2  -0.3418     0.2052  -1.666 0.102924
g_inv.l2  -1.3605     0.5883  -2.313 0.025589 *
g_int.l3  -0.2879     0.2115  -1.361 0.180512
g_inv.l3  -0.5597     0.5761  -0.971 0.336735
g_int.l4  -0.1830     0.1898  -0.964 0.340358
g_inv.l4  -0.2408     0.5845  -0.412 0.682414
g_int.l5   0.1351     0.1468   0.920 0.362479
g_inv.l5  -0.7547     0.5916  -1.276 0.208897
const     -0.1009     0.1376  -0.733 0.467551
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 0.9891 on 43 degrees of freedom
Multiple R-Squared: 0.396,      Adjusted R-squared: 0.2555
F-statistic: 2.819 on 10 and 43 DF,  p-value: 0.008837


Estimation results for equation g_inv:
======================================
g_inv = g_int.l1 + g_inv.l1 + g_int.l2 + g_inv.l2 + g_int.l3 + g_inv.l3 + g_int.l4 + g_inv.l4 + g_int.l5 +
g_inv.l5 + const

          Estimate Std. Error t value Pr(>|t|)
g_int.l1  0.120934   0.042829   2.824 0.007167 **
g_inv.l1  0.572835   0.155037   3.695 0.000618 ***
g_int.l2  0.134629   0.055681   2.418 0.019922 *
g_inv.l2 -0.340402   0.159661  -2.132 0.038758 *
g_int.l3  0.116170   0.057401   2.024 0.049228 *
g_inv.l3  0.109624   0.156371   0.701 0.487047
g_int.l4 -0.002778   0.051501  -0.054 0.957230
g_inv.l4  0.355119   0.158637   2.239 0.030414 *
g_int.l5  0.068886   0.039834   1.729 0.090924 .
g_inv.l5 -0.150672   0.160574  -0.938 0.353313
const    -0.034765   0.037349  -0.931 0.357148
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 0.2685 on 43 degrees of freedom
Multiple R-Squared: 0.5862,     Adjusted R-squared:  0.49
F-statistic: 6.092 on 10 and 43 DF,  p-value: 1.066e-05


Covariance matrix of residuals:
       g_int   g_inv
g_int 0.9783 0.08390
g_inv 0.0839 0.07207

Correlation matrix of residuals:
      g_int g_inv
g_int 1.000 0.316
g_inv 0.316 1.000
```

Figure 7.13: VAR(5) estimation

Interest rate equation g_int :

- The coefficient for $g\_int.l1$ remains negative and significant, indicating negative autocorrelation.
- The other coefficients for g_int and g_inv at different lags remain insignificant.
- $g\_inv.l1$ becomes significant with a p-value just above 10% and $g\_inv.l2$ is significant for a first order kind of 1%.
- The adjusted R^2 for this equation is 26%, indicating an improvement in the variation in interest rates explained.

Investment rate equation g_inv :

- The coefficient for $g\_inv.l1$ remains positive and significant.
- The coefficients for $g\_int.l1$ and $g\_int.l2$ remain positive and significant.
- $g\_int.l3$ becomes positive and significant for a first order kind of 1%.
- The R^2 increases to 47%, which is an improvement on the VAR(4) model.

```
> roots(var.4lag)
[1] 0.8056144 0.8056144 0.7263943 0.6706608 0.6706608 0.5937659 0.5937659 0.5824695
```

Figure 7.14: VAR(5) inverse of unit root values

The inverse of the unit roots are all less than 1 indicating that the model is stationary. We therefore select the VAR(5) as recommended by the final prediction criterion (FPE) and the AIC.

We can then write our VAR(5) as follow:

$$X_t = c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \phi_3 X_{t-3} + \phi_4 X_{t-4} + \phi_5 X_{t-5} + v_t \leftrightarrow \phi(L) X_t = c + v_t$$

With $c = \begin{bmatrix} a_1^0 \\ a_2^0 \end{bmatrix}, \phi_i = \begin{bmatrix} a_{11}^i & a_{12}^i \\ a_{21}^i & a_{22}^i \end{bmatrix}$ for i $= 1, \ldots, 5$ $\quad \phi(L) = \mathrm{I}_N - \sum_{i=1}^{5} \phi_i L^i, \phi_5 \neq 0_N$ and $v_t \sim BB(0, \Sigma)$

The VAR(5) process can be written as a VAR(1) process, but of dimension 2p (instead of 2). Let the VAR(5) process be of dimension $2\phi(L)X_t = \tilde{\phi}_0 + \phi_n + v_t$

$$Y_t = \begin{bmatrix} X_t \\ X_{t-1} \\ X_{t-2} \\ X_{t-3} \\ X_{t-4} \end{bmatrix}$$

We can rewrite $Y_t = \phi Y_{t-1} + \widetilde{\phi}_0 + \tilde{v}_t$ with $\widetilde{\phi}_0 = \begin{bmatrix} c \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \tilde{v}_t = \begin{bmatrix} v_t \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \phi = \begin{bmatrix} \phi_1 & \phi_2 & \phi_3 & \phi_4 & \phi_5 \\ \mathrm{I}_2 & 0 & 0 & 0 & 0 \\ 0 & \mathrm{I}_2 & 0 & 0 & 0 \\ 0 & 0 & \mathrm{I}_2 & 0 & 0 \\ 0 & 0 & 0 & \mathrm{I}_2 & 0 \end{bmatrix}$

Here are the coefficient matrices $\phi_i = \begin{bmatrix} a^i_{11} & a^i_{12} \\ a^i_{21} & a^i_{22} \end{bmatrix}$ for $i = 1, \ldots, 5$:

$$\phi_1 = \begin{bmatrix} a^1_{11} & a^1_{12} \\ a^1_{21} & a^1_{22} \end{bmatrix} = \begin{bmatrix} -0.612 & 0.988 \\ 0.121 & 0.573 \end{bmatrix},$$
$$\phi_2 = \begin{bmatrix} a^2_{11} & a^2_{12} \\ a^2_{21} & a^2_{22} \end{bmatrix} = \begin{bmatrix} -0.342 & -1.360 \\ 0.135 & -0.340 \end{bmatrix},$$
$$\phi_3 = \begin{bmatrix} a^3_{11} & a^3_{12} \\ a^3_{21} & a^3_{22} \end{bmatrix} = \begin{bmatrix} -0.288 & -0.560 \\ 0.116 & 0.110 \end{bmatrix},$$
$$\phi_4 = \begin{bmatrix} a^4_{11} & a^4_{12} \\ a^4_{21} & a^4_{22} \end{bmatrix} = \begin{bmatrix} -0.183 & -0.241 \\ -0.003 & 0.355 \end{bmatrix},$$
$$\phi_5 = \begin{bmatrix} a^5_{11} & a^5_{12} \\ a^5_{21} & a^5_{22} \end{bmatrix} = \begin{bmatrix} 0.135 & -0.755 \\ 0.069 & -0.151 \end{bmatrix},$$

```
> A<-Acoef(var.5lag)
> A
[[1]]
          g_int.l1  g_inv.l1
g_int -0.6118447 0.9879233
g_inv  0.1209340 0.5728351

[[2]]
          g_int.l2   g_inv.l2
g_int -0.3418423 -1.3604724
g_inv  0.1346292 -0.3404016

[[3]]
          g_int.l3   g_inv.l3
g_int -0.2878955 -0.5597100
g_inv  0.1161702  0.1096241

[[4]]
            g_int.l4   g_inv.l4
g_int -0.182950156 -0.2407847
g_inv -0.002778138  0.3551188

[[5]]
          g_int.l5   g_inv.l5
g_int 0.13508746 -0.7547479
g_inv 0.06888607 -0.1506720
```

Figure 7.15: VAR(5) coefficients

7.3 Granger causality test

We test whether at least one of the lags in the g_investment series has a causal effect on the g_interest series. The test hypotheses are:

$$\begin{cases} H_0 : & a_{12}^i = 0, i = 1, \cdots, p \leftrightarrow g\_ \text{ investment does not cause } g\_ \text{ interest} \\ H_1 : & a_{12}^i \neq 0, i = 1, \cdots, p \leftrightarrow g\_ \text{ investment causes } g\_ \text{ interest} \end{cases}$$

where a_{12}^i is the coefficient of row 1 and column 2 of ϕ_i. The test results are shown below:

```
$Granger

        Granger causality H0: g_inv do not Granger-cause g_int

data:  VAR object var.5lag
F-Test = 3.6927, df1 = 5, df2 = 86, p-value = 0.004479


$Instant

        H0: No instantaneous causality between: g_inv and g_int

data:  VAR object var.5lag
Chi-squared = 4.9024, df = 1, p-value = 0.02682
```

Figure 7.16: Granger causality test, investment rates

The critical thresholds are taken from an $F(5, 50)$ distribution. For the 5% probability, the critical value is 2.4, which is lower than the test statistic of 3.7. We therefore reject the null hypothesis of no causality between investment and interest rates. The critical probability of the test is equal to 0.00%, which leads to the same conclusions.

We test whether interest rates "cause" investment rates in Granger's sense. The test hypotheses are :

$$\begin{cases} H_0 : a_{21}^i = 0, \; i = 1, \cdots, p \leftrightarrow g\_ \text{ interest does not cause } g\_ \text{ investment} \\ H_1 : a_{21}^i \neq 0, \; i = 1, \cdots, p \leftrightarrow g\_ \text{ interest causes } g\_ \text{ investment} \end{cases}$$

where a_{12}^i is the coefficient of row 2 and column 1 of ϕ_i. The test results are shown below:

```
$Granger

        Granger causality H0: g_int do not Granger-cause g_inv

data:  VAR object var.5lag
F-Test = 3.5806, df1 = 5, df2 = 86, p-value = 0.005458


$Instant

        H0: No instantaneous causality between: g_int and g_inv

data:  VAR object var.5lag
Chi-squared = 4.9024, df = 1, p-value = 0.02682
```

Figure 7.17: Granger causality test, interest rates

The critical threshold for the usual 5% $F(5, 50)$ distribution is 2.4 , which is lower than the test statistic of 3.6. We therefore reject the null hxpothesis of no causality between the of interest rates and the grouxth rate of investment. The critical probability of the test is equal to 0.00%, which leads to the same conclusions.

7.3.1 Conclusion

We can see that investment rates cause interest rates in Granger's sense, and reciprocally. Causality operates in both directions, and we speak of a feedback effect.

7.4 Impulse-Response Function

The response function of a VAR traces the effect of a shock at t to one of the variables on the present and future values of that variable and the other variables in the VAR.

As each of the two VAR variables causes the other, causality tests in the Granger sense do not allow us to rank the variables. Nevertheless, it is reasonable to assume that shocks to interest rates have an instantaneous impact on investment rates, rather the inverse. We decide to "place" the interest rate $g_$interest in first position and the investment $g_$investment in second.

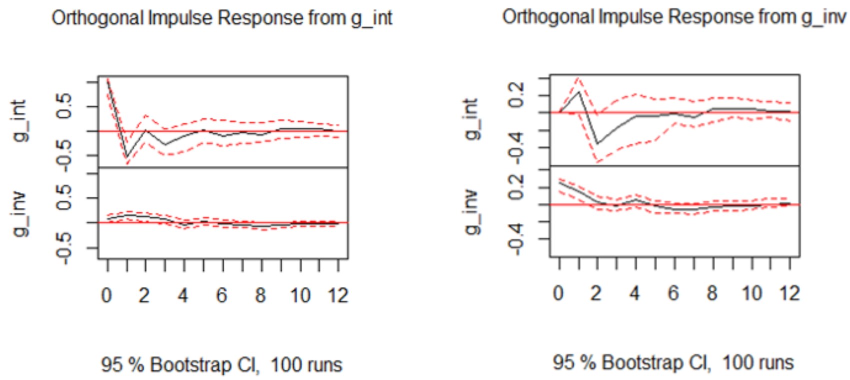


Figure 7.18: Impulse-Response Function

The first graph shows the response of $g_$interest and $g_$investment to an orthogonal shock in $g_$interest.

- For $g\_interest$: there is an initial fall, followed by recovery and fluctuation around zero. This indicates that the effect of the shock on the interest rate fades fairly quickly and the variable returns to its initial level.
- For $g\_investment$: the shock in $g\_interest$ appears to have a negative short-term effect on $g\_investment$, but the 95% confidence interval includes zero, indicating that the effect is not statistically significant at all time horizons.

The second graph shows the response of $g_$interest and $g_$investment to an orthogonal shock in $g_$investment.

- For $g\_investment$: the shock in $g\_investment$ leads to an immediate increase, indicating positive autocorrelation. However, the effect diminishes rapidly, and the confidence intervals become wide, suggesting uncertainty about the long-term impact.
- For $g\_interest$: the shock does not appear to have a significant effect, as the response line and confidence intervals contain zero at all time horizons.

The problem with the previous approach is that the way in which X_t evolves may not be adequately represented by the VAR(5). As a result, this introduces a bias into the estimation of responses, and this bias becomes increasingly pronounced as the horizon lengthens, as errors propagate from one period to the next. An alternative is to use local regressions.

7.5 Estimating IRF using local projections

Here's how the responses are calculated, using Jorda's method of local projections. Structural shocks are determined using a Cholesky decomposition.

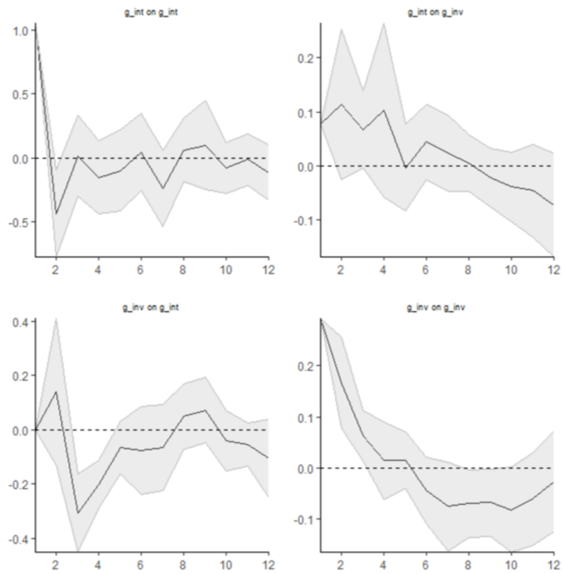


Figure 7.19: IRF plots

Impact of a shock on g_interest on $g_$investment : it appears that the impact of the shock in the case of local projections shows greater variability over the time horizon, with wide confidence intervals suggesting uncertainty as to the sustainability of the shock's impact. The confidence intervals encompass zero throughout the horizon, indicating that there is no statistical evidence of a lasting effect.

Here are the values of impulse-response functions for shocks on $g_$investment:

```
          [,1]       [,2]       [,3]       [,4]         [,5]       [,6]        [,7]       [,8]
[1,] 1.05160247 -0.4348236 0.01701101 -0.1506653 -0.097061008 0.04640449 -0.23540516 0.06236350
[2,] 0.07706784  0.1132764 0.06777860  0.1027272 -0.002967354 0.04471968  0.02426116 0.00584664
            [,9]       [,10]       [,11]       [,12]       [,13]
[1,]  0.10083663 -0.07676010 -0.01265764 -0.11313316 -0.07250357
[2,] -0.02179559 -0.03865337 -0.04603776 -0.07286099 -0.02849778
```

Figure 7.20: Impulse-response values for shocks on g_interest

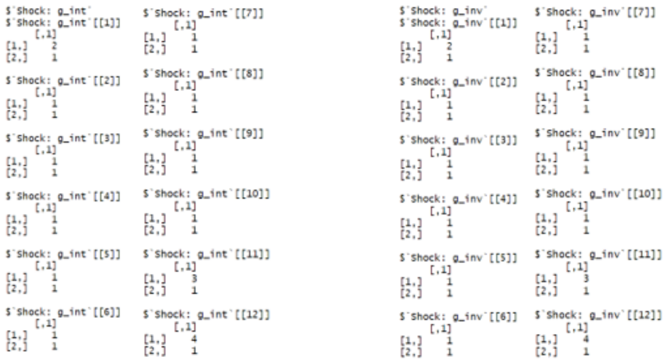
Impact of a shock on $g_$investment on $g_$interest : in the local projections, there appears to be a more dynamic initial effect, which quickly dissipates and becomes insignificant. This could indicate that local projections capture a more immediate response to shocks, but provide no evidence of long-term effects.

Here are the values of impulse-response functions for shocks on $g_$interest :

```
          [,1]      [,2]        [,3]        [,4]        [,5]        [,6]        [,7]        [,8]
1,] 0.0000000 0.1391034 -0.30780160 -0.20234748 -0.06603277 -0.07739771 -0.06534866  0.04731528
2,] 0.2919506 0.1670650  0.06268545  0.01422203  0.01511898 -0.04395882 -0.07509532 -0.06992763
           [,9]       [,10]       [,11]      [,12]      [,13]
1,]  0.07198342 -0.04063095 -0.05536502 -0.1055269 0.11824042
2,] -0.06757027 -0.08175280 -0.06019826 -0.0264063 0.01326671
```

Figure 7.21: Impulse-response values for shocks on g_investment

The display below shows that the number of optimal lags included in the local projection varies according to the horizon variable, the variable being shocked and the variable whose dynamics are being studied.



```
$`Shock: g_int`
$`Shock: g_int`[[1]]
     [,1]
[1,]    2
[2,]    1

$`Shock: g_int`[[2]]
     [,1]
[1,]    1
[2,]    1

$`Shock: g_int`[[3]]
     [,1]
[1,]    1
[2,]    1

$`Shock: g_int`[[4]]
     [,1]
[1,]    1
[2,]    1

$`Shock: g_int`[[5]]
     [,1]
[1,]    1
[2,]    1

$`Shock: g_int`[[6]]
     [,1]
[1,]    1
[2,]    1

$`Shock: g_int`[[7]]
     [,1]
[1,]    1
[2,]    1

$`Shock: g_int`[[8]]
     [,1]
[1,]    1
[2,]    1

$`Shock: g_int`[[9]]
     [,1]
[1,]    1
[2,]    1

$`Shock: g_int`[[10]]
     [,1]
[1,]    1
[2,]    1

$`Shock: g_int`[[11]]
     [,1]
[1,]    3
[2,]    1

$`Shock: g_int`[[12]]
     [,1]
[1,]    4
[2,]    1
```

```
$`Shock: g_inv`
$`Shock: g_inv`[[1]]
     [,1]
[1,]    2
[2,]    1

$`Shock: g_inv`[[2]]
     [,1]
[1,]    1
[2,]    1

$`Shock: g_inv`[[3]]
     [,1]
[1,]    1
[2,]    1

$`Shock: g_inv`[[4]]
     [,1]
[1,]    1
[2,]    1

$`Shock: g_inv`[[5]]
     [,1]
[1,]    1
[2,]    1

$`Shock: g_inv`[[6]]
     [,1]
[1,]    1
[2,]    1

$`Shock: g_inv`[[7]]
     [,1]
[1,]    1
[2,]    1

$`Shock: g_inv`[[8]]
     [,1]
[1,]    1
[2,]    1

$`Shock: g_inv`[[9]]
     [,1]
[1,]    1
[2,]    1

$`Shock: g_inv`[[10]]
     [,1]
[1,]    1
[2,]    1

$`Shock: g_inv`[[11]]
     [,1]
[1,]    3
[2,]    1

$`Shock: g_inv`[[12]]
     [,1]
[1,]    4
[2,]    1
```

Figure 7.22: Shock on g_interest (left) and on g_investment (right)

7.6 Variance decomposition

We examine the forecast error variance decomposition (FEVD) for interest rates and investment rates. This method decomposes the forecast error variance of each of these time series into shares attributable to its own shocks and those of the other variable in a VAR model.

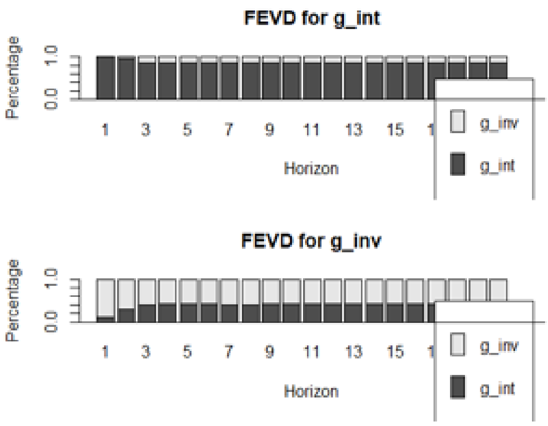


Figure 7.23: FEVD

The FEVD results for interest rates show that most of the variance in interest rate forecasts is explained by its own past shocks, with little or no variation attributable to shocks in investment levels. This suggests that interest rates are largely determined by their own factors and are less sensitive to investment dynamics.

There appears to be a significant contribution from past interest rate shocks to the variance in investment level forecasts. This explained variance increases with the horizon. This highlights the importance of considering the long-term impact of monetary policies on real economic activity.

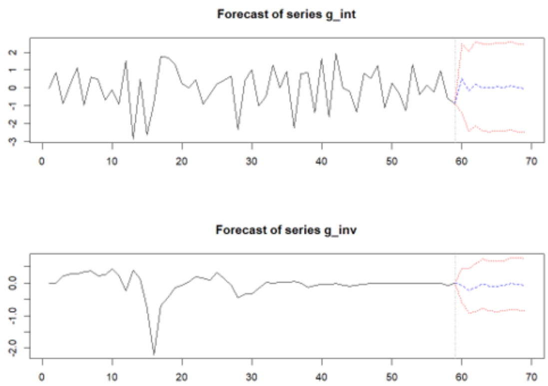


Figure 7.24: Forecast of series g_int (on top) and g_inv (underneath)

7.7 Forecast calculation

The graphs show atypical values for the 2nd , 3rd and 4th quarters of 2008, due to the crisis. We introduce 3 indicator variables to take this particular observation into account.

```
> pselect$selection # nombre de retards optimal selon les critères de sélec
AIC(n)  HQ(n)  SC(n) FPE(n)
     7      4      1      4
```

Figure 7.25: Criteria selection for the optimum number of lags

The AIC and FPE criteria recommend a VAR(4).

7.8 Estimation of VAR(4) with indicators 2008:Q2, 2008:Q3, 2008:Q4

```
VAR Estimation Results:
=========================
Endogenous variables: g_int, g_inv
Deterministic variables: const
Sample size: 52
Log Likelihood: -41.927
Roots of the characteristic polynomial:
0.9146 0.9146 0.8603 0.8603 0.8498 0.8498 0.8369 0.8369 0.8092 0.8092 0.7411 0.7411 0.6216 0.5982
Call:
VAR(y = df_g_short[, c(2, 3)], type = "const", exogen = cbind(df_g_short$DUM_08Q2,
    df_g_short$DUM_08Q3, df_g_short$DUM_08Q4), lag.max = 8)


Estimation results for equation g_int:
======================================
g_int = g_int.l1 + g_inv.l1 + g_int.l2 + g_inv.l2 + g_int.l3 + g_inv.l3 + g_int.l4 + g_inv.l4 + g_int.l5 + g_inv.l5
+ g_int.l6 + g_inv.l6 + g_int.l7 + g_inv.l7 + const + exo1 + exo2 + exo3

         Estimate Std. Error t value Pr(>|t|)
g_int.l1 -0.79060    0.17390  -4.546  6.6e-05 ***
g_inv.l1  1.66051    0.62140   2.672  0.01149 *
g_int.l2 -0.70854    0.21409  -3.310  0.00222 **
g_inv.l2 -0.65360    0.66771  -0.979  0.33456
g_int.l3 -0.57949    0.22851  -2.536  0.01598 *
g_inv.l3 -0.88650    0.66686  -1.329  0.19258
g_int.l4 -0.57562    0.24927  -2.309  0.02714 *
g_inv.l4  0.10357    0.58051   0.178  0.85945
g_int.l5 -0.22512    0.22156  -1.016  0.31677
g_inv.l5 -0.74224    0.58837  -1.262  0.21571
g_int.l6 -0.33036    0.18339  -1.801  0.08052 .
g_inv.l6 -0.55757    0.57885  -0.963  0.34223
g_int.l7 -0.20312    0.14854  -1.367  0.18048
g_inv.l7  0.02075    0.55079   0.038  0.97016
const     0.09061    0.13654   0.664  0.51139
exo1     -1.45368    0.98276  -1.479  0.14830
exo2     -1.60404    1.09922  -1.459  0.15367
exo3     -3.90639    1.10449  -3.537  0.00119 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 0.8791 on 34 degrees of freedom
Multiple R-Squared: 0.6157,     Adjusted R-squared: 0.4236
F-statistic: 3.204 on 17 and 34 DF,  p-value: 0.001877


Estimation results for equation g_inv:
======================================
g_inv = g_int.l1 + g_inv.l1 + g_int.l2 + g_inv.l2 + g_int.l3 + g_inv.l3 + g_int.l4 + g_inv.l4 + g_int.l5 + g_inv.l5
+ g_int.l6 + g_inv.l6 + g_int.l7 + g_inv.l7 + const + exo1 + exo2 + exo3

         Estimate Std. Error t value Pr(>|t|)
g_int.l1  0.08665    0.04993   1.735  0.09172 .
g_inv.l1  0.51755    0.17842   2.901  0.00648 **
g_int.l2  0.10560    0.06147   1.718  0.09491 .
g_inv.l2 -0.22880    0.19172  -1.193  0.24097
g_int.l3  0.16353    0.06561   2.492  0.01772 *
g_inv.l3 -0.21975    0.19147  -1.148  0.25910
g_int.l4  0.06078    0.07157   0.849  0.40172
g_inv.l4  0.32557    0.16668   1.953  0.05905 .
g_int.l5  0.12892    0.06361   2.027  0.05060 .
g_inv.l5 -0.10601    0.16893  -0.628  0.53451
g_int.l6  0.09574    0.05266   1.818  0.07785 .
g_inv.l6 -0.03669    0.16620  -0.221  0.82662
g_int.l7  0.05042    0.04265   1.182  0.24537
g_inv.l7  0.20103    0.15814   1.271  0.21229
const    -0.04602    0.03920  -1.174  0.24863
exo1      0.54310    0.28217   1.925  0.06267 .
exo2      0.03494    0.31361   0.111  0.91250
exo3     -0.69210    0.31713  -2.182  0.03608 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 0.2524 on 34 degrees of freedom
Multiple R-Squared: 0.6963,     Adjusted R-squared: 0.5445
F-statistic: 4.586 on 17 and 34 DF,  p-value: 7.889e-05


Covariance matrix of residuals:
        g_int   g_inv
g_int 0.77282 0.09497
g_inv 0.09497 0.06371

Correlation matrix of residuals:
      g_int g_inv
g_int 1.000 0.428
g_inv 0.428 1.000
```

Figure 7.26: VAR(4) estimation with dummies from Q2 2008 to Q4 2008

Only the indicator variable Q4 2008 is significant in the equation $g_$interest. The indicator variable Q2 2008 is significant at the 10% threshold and Q4 2008 at the 5% threshold in the equation $g_$investment.

We keep only the significant Q4 2008 dummy and repeat the selection tests.

```
> pselect$criteria # Affichage des valeurs des critère de sélection
                1          2          3           4          5           6           7           8
AIC(n) -2.2315211 -2.2901571 -2.4819752 -2.54224026 -2.5167229 -2.51026720 -2.46949337 -2.40514976
HQ(n)  -2.1157238 -2.1164611 -2.2503807 -2.25274704 -2.1693311 -2.10497669 -2.00630421 -1.88406196
SC(n)  -1.9284896 -1.8356099 -1.8759123 -1.78466158 -1.6076285 -1.44965705 -1.25736748 -1.04150813
FPE(n)  0.1074343  0.1014725  0.0840157  0.07950381  0.0821874  0.08362877  0.08841566  0.09616243
> pselect$selection
AIC(n)  HQ(n)  SC(n) FPE(n)
     4      4      1      4
```

Figure 7.27: Criteria selection for the optimum number of lags

AIC, HQ and FPE recommend a VAR(4). We therefore estimate a VAR(4) including the Q4 2008 dummy.

7.9 Estimation of a VAR(4) with dummy on Q4 2008

```
VAR Estimation Results:
=========================
Endogenous variables: g_int, g_inv
Deterministic variables: const
Sample size: 55
Log Likelihood: -63.436
Roots of the characteristic polynomial:
0.8104 0.8104 0.6949 0.6764 0.6741 0.6741 0.6722 0.6722
Call:
VAR(y = df_g_short[, c(2, 3)], type = "const", exogen = cbind(df_g_short$DUM_08Q4),
    lag.max = 8)


Estimation results for equation g_int:
======================================
g_int = g_int.l1 + g_inv.l1 + g_int.l2 + g_inv.l2 + g_int.l3 + g_inv.l3 + g_int.l4 + g_inv.l4 + const + exo1

         Estimate Std. Error t value Pr(>|t|)
g_int.l1 -0.68817    0.14353  -4.795 1.82e-05 ***
g_inv.l1  0.87926    0.48463   1.814  0.07630 .
g_int.l2 -0.51446    0.18102  -2.842  0.00672 **
g_inv.l2 -0.64846    0.52326  -1.239  0.22167
g_int.l3 -0.18447    0.15334  -1.203  0.23527
g_inv.l3 -1.25100    0.55644  -2.248  0.02951 *
g_int.l4 -0.16396    0.13440  -1.220  0.22884
g_inv.l4 -0.38170    0.50193  -0.760  0.45094
const     0.01611    0.12661   0.127  0.89931
exo1     -3.74419    1.07632  -3.479  0.00113 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 0.9091 on 45 degrees of freedom
Multiple R-Squared: 0.4759,     Adjusted R-squared: 0.3711
F-statistic: 4.541 on 9 and 45 DF,  p-value: 0.0002764

Estimation results for equation g_inv:
======================================
g_inv = g_int.l1 + g_inv.l1 + g_int.l2 + g_inv.l2 + g_int.l3 + g_inv.l3 + g_int.l4 + g_inv.l4 + const + exo1

          Estimate Std. Error t value Pr(>|t|)
g_int.l1  0.100515   0.040692   2.470 0.017362 *
g_inv.l1  0.537647   0.137401   3.913 0.000305 ***
g_int.l2  0.089370   0.051323   1.741 0.088460 .
g_inv.l2 -0.120786   0.148352  -0.814 0.419827
g_int.l3  0.131132   0.043475   3.016 0.004199 **
g_inv.l3 -0.066586   0.157760  -0.422 0.674980
g_int.l4 -0.015963   0.038103  -0.419 0.677256
g_inv.l4  0.349431   0.142304   2.456 0.017995 *
const    -0.006889   0.035896  -0.192 0.848675
exo1     -0.880253   0.305154  -2.885 0.005995 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 0.2577 on 45 degrees of freedom
Multiple R-Squared: 0.6068,     Adjusted R-squared: 0.5281
F-statistic: 7.716 on 9 and 45 DF,  p-value: 9.402e-07



Covariance matrix of residuals:
        g_int   g_inv
g_int 0.82640 0.05892
g_inv 0.05892 0.06643

Correlation matrix of residuals:
       g_int  g_inv
g_int 1.0000 0.2515
g_inv 0.2515 1.0000
```

Figure 7.28: VAR(4) estimation with dummy on Q4 2008

Comparing the results of the VAR(5) and VAR(4) models with a dummy variable for Q4 2008, we observe that the inclusion of a dummy variable for Q4 2008 in the VAR(4) model appears to significantly improve the adjusted R-squared. For g_int, it rises to 37%, and for g_inv to 53%, suggesting that the model with the dummy and fewer lags provides a better fit.

The coefficient of the dummy variable is significantly negative for g_int, indicating a notable impact of Q4 2008 on this variable. F-statistics are also better in the VAR(4) model with dummy, reinforcing the idea that the model is more robust. New variables also become significant.

8 Cointegration

8.1 Characterization of the two series

First, we need to check that the two series are indeed non-stationary processes by applying the augmented Dickey-Fuller test. We had obtained the results that the interest rate series is $X_t \sim I(1)$ with $\Delta X_t = \varepsilon_t$ and the investment series is $Y_t \sim I(0)$ with $\Delta Y_t = \rho Y_{t-1} + \varepsilon_t$ according to the Augmented Dickey-Fuller test. However, for the investment rate series, the KPSS tests did not reject the null hypothesis of stationarity.

We will therefore consider 2 cases:

$$1^{\text{st}} \text{ case } : X_t \sim I(1) \text{ and } Y_t \sim I(0) \rightarrow aX_t + bY_t \sim I(1) \text{ where } a, b \in \mathbb{R} \text{ and } b \neq 0$$
$$2^{\text{nd}} \text{ case } : X_t \sim I(1) \text{ and } Y_t \sim I(1) \rightarrow aX_t + bY_t \sim I(1) \text{ where } a, b \in \mathbb{R} \text{ and } b \neq 0$$

Our two series are therefore integrated of order 1. We can therefore envisage the possible existence of a cointegration relationship between the time series. The concept of cointegration was introduced by Granger (1981). The central notion of cointegration is that, although variables may show divergent movements in the short term, there may be a stable long-term relationship between them. This approach suggests that variables may have temporary fluctuations that cause them to move in opposite directions in the short term, but that they tend to converge or share a coherent relationship in the longer term. We adopt Johansen's approach, which is a multivariate method for detecting cointegration. This method assumes that the time series in question are integrated of the same order, as previously established when applying unit root tests.

To assess the number of cointegration relationships, denoted h (also known as the cointegration rank), Johansen introduced two statistics: the trace test and the largest eigenvalue test. When the results of the two tests diverge, it is generally preferable to favor the findings of the trace test, which is often considered more robust than the largest eigenvalue test.

8.2 Trace test

The t-statistic is given below:

$$Q(r) = -T \sum_{i=r+1}^{n} \ln\left(1 - \hat{\lambda}_i\right)$$

Where T is the number of observations and λi is the ith largest eigenvalue of the matrix. Test hypothesis:

$$\begin{cases} H_0(r): & \text{rang}(\Pi) \leq r \text{ (at most } r \text{ cointegration relationships)} \\ H_1(r): & \text{rang}(\Pi) > r \end{cases}$$

Its null hypothesis is rejected when $Q(r)$ is above the critical threshold.

```
                 1           2           3           4           5
AIC(n) -2.63427190 -2.59614720 -2.59167515 -2.48550837 -2.41258018
HQ(n)  -2.54795719 -2.45228936 -2.39027417 -2.22656425 -2.09609292
SC(n)  -2.40912839 -2.22090803 -2.06634031 -1.81007786 -1.58705399
FPE(n)  0.07178962  0.07464922  0.07514116  0.08387157  0.09075346
                 6          7          8
AIC(n) -2.38368974 -2.3057851 -2.2265504
HQ(n)  -2.00965935 -1.8742116 -1.7374337
SC(n)  -1.40806788 -1.1800676 -0.9507372
FPE(n)  0.09422794  0.1030973  0.1134123
> pselect$selection
AIC(n)  HQ(n)  SC(n) FPE(n)
     1      1      1      1
```

Figure 8.1: Criteria selection

Thus, we choose one lag.

```
######################
# Johansen-Procedure #
######################

Test type: trace statistic , without linear trend and constant in cointegration

Eigenvalues (lambda):
[1]  4.166038e-01  5.836284e-02 -1.567916e-17

Values of teststatistic and critical values of test:

          test 10pct  5pct  1pct
r <= 1 |  3.49  7.52  9.24 12.97
r = 0  | 34.74 17.85 19.96 24.60

Eigenvectors, normalised to first column:
(These are the cointegration relations)

               interest.l1 investment.l1   constant
interest.l1       1.000000     1.0000000  1.0000000
investment.l1     6.598650    -0.2306189 -0.6853908
constant         -4.490856    -0.6582051 -7.7142266

Weights W:
(This is the loading matrix)

             interest.l1 investment.l1      constant
interest.d    0.01934238   -0.03653976 -1.047544e-19
investment.d -0.10993109   -0.10096351  2.901529e-17
```

Figure 8.2: Criteria selection

We start with the hypothesis:

$$\begin{cases} H_0(0) : r = \text{rang}(\Pi) \leq 0 \\ H_1(0) : r = \text{rang}(\Pi) > 0 \end{cases}$$

The t-statistic of Trace test is :

$$Q(0) = -T \sum_{i=1}^{2} \ln\left(1 - \hat{\lambda}_i\right) = 34.74$$

The rejection threshold for a first-kind order of 1% is equal to 24.60 . We therefore have $Q(0) = 34.74 > 24.60$. We therefore reject the null hypothesis $H_0(0) : r = \text{rang}(\Pi) \leq 0$ that the number of cointegration relationships is equal to 0 for a first- kind order $\alpha = 1\%$. We proceed to the following assumptions:

$$\begin{cases} H_0(1) : r = \text{rang}(\Pi) \leq 1 \\ H_1(1) : r = \text{rang}(\Pi) > 1 \end{cases}$$

The t-statistic of Trace test is :

$$Q(1) = -T \ln\left(1 - \hat{\lambda}_2\right) = 3.49$$

The rejection threshold for a first-kind order of 1% is equal to 12.97 . We therefore have $Q(1) = 3.49 < 12.97$. We therefore do not reject the null hypothesis $H_0(0) : r = \text{rang}(\Pi) \leq 1$ that the number of cointegration relationships is equal to 1 for a first-kind order $\alpha = 1\%$.

We conclude that there is one cointegration relationship between the interest rates and investment rates.

8.3 Test for maximum eigenvalue

The test statistic is :

$$Q_{\max}(r \mid r+1) = -T \ln\left(1 - \hat{\lambda}_{r+1}\right) = Q(r+1) - Q(r) \approx T\hat{\lambda}_{r+1}$$

Test hypotheses:

$$\begin{cases} H_0(1): & r = \text{rang}(\Pi) \leq 1 \text{(at most r cointegration relationships)} \\ H_1(1): & r = \text{rang}(\Pi) > 1 \text{(at most r+1 cointegration relationships)} \end{cases}$$

Its null hypothesis is rejected when $Q_{\max}(r \mid r+1)$ is above the critical threshold.

```
######################
# Johansen-Procedure #
######################

Test type: maximal eigenvalue statistic (lambda max) , without linear trend and constant in cointegration

Eigenvalues (lambda):
[1]  4.166038e-01  5.836284e-02 -1.567916e-17

Values of teststatistic and critical values of test:

          test 10pct  5pct  1pct
r <= 1 |  3.49  7.52  9.24 12.97
r = 0  | 31.26 13.75 15.67 20.20

Eigenvectors, normalised to first column:
(These are the cointegration relations)

              interest.l1 investment.l1   constant
interest.l1      1.000000     1.0000000  1.0000000
investment.l1    6.598650    -0.2306189 -0.6853908
constant        -4.490856    -0.6582051 -7.7142266

Weights W:
(This is the loading matrix)

             interest.l1 investment.l1      constant
interest.d    0.01934238   -0.03653976 -1.047544e-19
investment.d -0.10993109   -0.10096351  2.901529e-17
```

Figure 8.3: Maximum eigenvalue test

We start with the hypothesis:

$$\begin{cases} H_0 : r = \text{rang}(\Pi) = 0 \\ H_1 : r = \text{rang}(\Pi) = 1 \end{cases}$$

The maximum eigenvalue test statistic is :

$$Q(0 \mid 1) = -T \sum_{i=1}^{1} \ln\left(1 - \hat{\lambda}_i\right) = 31.26$$

The test rejection threshold for a first-kind order of 1% is $20.20 < Q(0 \mid 1) = 31.26$: we therefore reject the null hypothesis $H_0 : r = \text{rang}(\Pi) \leq 0$ that the number of cointegration relationships is equal to 0 for a first-species risk $\alpha = 1\%$. We proceed to the following assumptions :

$$\begin{cases} H_0 : r = \text{rang}(\Pi) = 1 \\ H_1 : r = \text{rang}(\Pi) = 2 \end{cases}$$

The maximum eigenvalue test statistic is :

$$Q(1 \mid 2) = -T \ln \left(1 - \hat{\lambda}_2\right) = 3.49$$

The test rejection threshold for a first-kind order of 1% is $12.97 > Q(1 \mid 2) = 3.49$: the null hypothesis is therefore rejected. $H_0 : r = \text{rang}(\Pi) = 1$ that the number of cointegration relationships is equal to 1 for a first kind order $\alpha = 1\%$.

There are therefore two cointegration relationships between the interest rates and investment rates. This confirms the conclusions we obtained with the trace test. We have strong evidence that there is a cointegration relationship between the two variables.

8.4 VECM

We conclude that there is a cointegrating relationship between interest and investment rates. The VECM model is estimated by imposing a cointegrating relationship.

```
Response interest.d :

Call:
lm(formula = interest.d ~ ect1 + interest.dl1 + investment.dl1 - 
    1, data = data.mat)

Residuals:
     Min       1Q   Median       3Q      Max 
-1.52787 -0.05590 -0.00927  0.13026  0.52966 

Coefficients:
               Estimate Std. Error t value Pr(>|t|)  
ect1           0.019342   0.008327   2.323   0.0239 *
interest.dl1   0.293251   0.159160   1.842   0.0708 .
investment.dl1 0.009172   0.041427   0.221   0.8256  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.2842 on 55 degrees of freedom
Multiple R-squared:  0.4364,	Adjusted R-squared:  0.4057 
F-statistic:  14.2 on 3 and 55 DF,  p-value: 5.741e-07


Response investment.d :

Call:
lm(formula = investment.d ~ ect1 + interest.dl1 + investment.dl1 - 
    1, data = data.mat)

Residuals:
    Min      1Q  Median      3Q     Max 
-3.1966 -0.5458  0.1182  0.6321  1.4585 

Coefficients:
                Estimate Std. Error t value Pr(>|t|)    
ect1           -0.109931   0.027647  -3.976 0.000206 ***
interest.dl1    1.140668   0.528434   2.159 0.035263 *  
investment.dl1 -0.001116   0.137545  -0.008 0.993558    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9437 on 55 degrees of freedom
Multiple R-squared:  0.3248,	Adjusted R-squared:  0.2879 
F-statistic: 8.817 on 3 and 55 DF,  p-value: 7.229e-05
```

Figure 8.4: VECM estimation

We have the two VECM equations:

- Interest.d: Represents the first difference of the interest variable, calculated as *interest* at time t minus *interest* at time $t-1$.
- Investment.d: Similarly denotes the first difference of the investment variable.
- Interest.dli: Refers to the *interest.d* variable shifted by i lags. The same concept applies to *investment.d*.
- ect 1: Indicates the lagged deviation from the cointegrating relationship.

Equation for interest.d:

- The coefficient (speed of adjustment) of ect 1 is positive and significant at 5%. This indicates that the cointegration relationship affects the dynamics of interest.d. A positive sign suggests that when ect > 0, interest rates at time $t-1$ are below their long-term value, and the error correction mechanism tends to correct this.

- The coefficient *interest.dl1* is significant at 10% with a positive sign.
- The coefficient *investment.dl1* is not significant.

<u>Equation for investment.d:</u>

- The coefficient (speed of adjustment) of ect1 is negative and significant at 1%. This implies that the cointegration relationship influences the dynamics of investment.d. A positive sign for ect1 indicates that when it is greater than 0, investment rates at time $t-1$ are higher than their long-term value, and the error correction mechanism works to adjust this towards the equilibrium.
- The coefficient *interest.dl1* is significant at 5% with a positive sign.
- The coefficient *investment.dl1* is not significant.

The variables and the cointegration relationship can be plotted using the plot function.

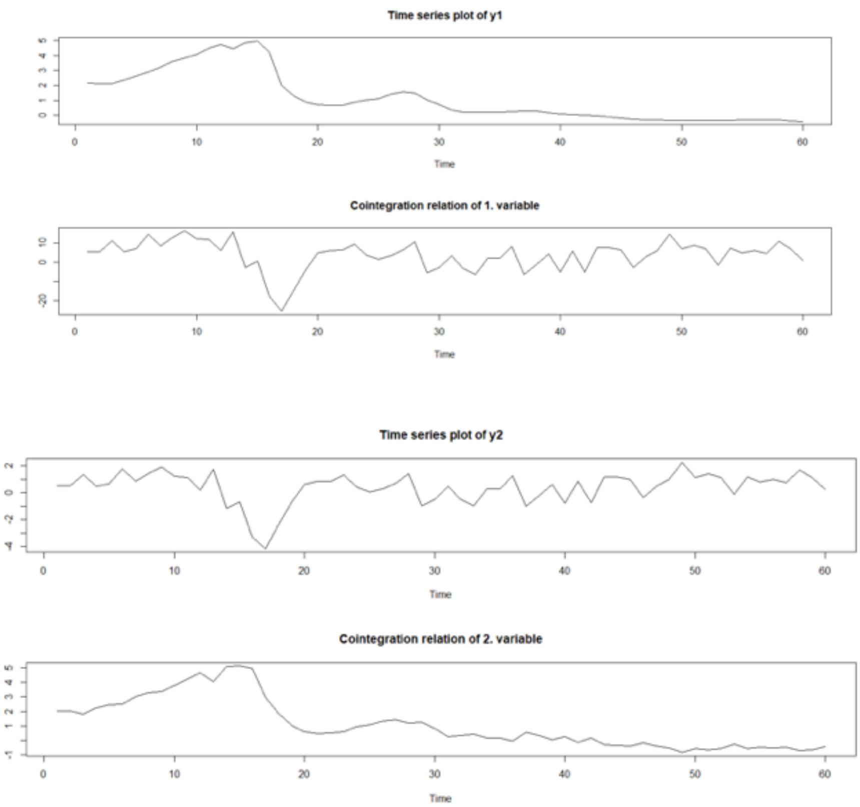


Figure 8.5: Graphs of the cointegration relationship

.

9 Conclusion

In summary, our analysis shows that tangible and intangible investments, which refer to GFCF, and short-term interest rates are an integral part of France's economic dynamics. Indeed, by exploring the relationship between GFCF rates and short-term interest rates on a quaterly basis, our study recognises the close relationship between these rates. Moreover, it also confirms the major impact of the key rate set by the European Central Bank. This relationship is essential, especially given the ECB's role in managing key interest rates in the Eurozone and its quantitative easing measures during periods of economic downturn, such as the recent Covid-19 crisis. However, it is important to note that our study does not classify these quantitative easing measures as investments, since they concern financial investments. Using a sequential unit root strategy, we demonstrated that these two variables are stationary processes with no drift. The analyses, including the significance of the coefficients and examination of the residuals, led us to select an MA(2) model for the interest rate series with dummy over the Q4 2008 period as the best possible representation, with a minimised AIC and BIC criterion. We then performed forecasting tests with the MA(2) by adding dummies over the period Q2 2008 to Q4 2008 to predict the year 2020. The predicted values are higher than the observed values, but fairly close.

The use of a VAR(5) model and the Granger test revealed a bidirectional causal relationship between interest rates and investment rates. However, the results of the FEVD analysis showed that interest rate forecasts are mainly influenced by their own past shocks, while shocks in investment levels have a limited impact on interest rates. By comparing the VAR(5) and VAR(4) models with a dummy variable for Q4 2008, we found that the inclusion of this dummy variable in the VAR(4) model significantly improves the adjusted coefficient of determination.

Besides, cointegration tests of the Trace and the maximum eigenvalue revealed the existence of one cointegration relationship. This suggests that, despite their possible individual non-stationarity, interest rates and investment rates tend to evolve jointly over the long term. This finding is crucial for understanding the long-term dynamics of these economic indicators. Although this may strengthen our model by indicating a lasting relationship, it is important to note that cointegration does not necessarily prove a causal relationship, but it does offer an additional perspective for better understanding the interactions between variables over a prolonged period.

References

Agnès Bénassy-Quéré, V. C., Laurence Boone. (2014). *Les taux d'intérêt.*

EY. (2023). *Baromètre de l'attractivité de la france 2023.* https://assets.ey.com/content/dam/ey-sites/ey-com/fr_fr/topics/attractiveness/barometre-de-l-attractivite-de-la-france-2023/ey-barometre-attractivite-france-2023-episode-1-20230525.pdf

IMF. (2023). *World economic outlook.* https://www.imf.org/external/datamapper/NGDPD@WEO/OEMDC/ADVEC/WEOWORLD

INSEE. (2020). *Formation brute de capital fixe / fbcf / investissement (comptabilité nationale).* https://www.insee.fr/fr/metadonnees/definition/c1371

INSEE. (2023a). *Gdp in france.* https://www.insee.fr/en/statistiques/7615290

INSEE. (2023b). *L'essentiel sur le chômage en france.* https://www.insee.fr/fr/statistiques/4805248

INSEE. (2023c). *Taux d'inflation.* https://www.insee.fr/fr/statistiques/2122401

La finance pour tous. (2008). *La crise des subprimes (2007-2008).* https://www.lafinancepourtous.com/decryptages/economie/crise-economique/la-crise-des-subprimes

Le Monde. (2011). *La bce augmente son taux directeur à 1,25%.* https://www.lemonde.fr/economie/article/2011/04/07/la-bce-augmente-son-taux-directeur-a-1-25_1504437_3234.html

Les Echos. (2019). *Il y a dix ans, cette crise grecque qui faillit tout emporter.* https://www.lesechos.fr/idees-debats/editos-analyses/il-y-a-dix-ans-cette-crise-grecque-qui-faillit-tout-emporter-1155383

OECD. (2023a). *Investment by sector.* https://data.oecd.org/gdp/investment-by-sector.htm

OECD. (2023b). *Investment gfcf.* https://data.oecd.org/gdp/investment-gfcf.htm

OECD. (2023c). *Taux d'intérêt - taux d'intérêt à court terme.* https://data.oecd.org/fr/interest/taux-d-interet-a-court-terme.htm

Oil price. (2023). https://prixdubaril.com

Statista. (2023). *Évolution du pib réel de la france.* https://fr.statista.com/statistiques/479446/evolution-annuelle-du-pib-en-volume-france/

TradingView. (2009). *Cac 40 index chart and quotes.* https://www.tradingview.com/symbols/TVC-CAC40/

UK Essays. (2008). *Impact of the global financial crisis on france economics essay.* https://www.ukessays.com/essays/economics/impact-global-financial-crisis-france-9386.php