Louvain School of Management

Do energy price variations impact the volatility of inflation expectations?

Author(s): Lucas Dupuis & Pierre Houssière
Supervisor(s): Professor L. Iania
Academic year 2024-2025
Dissertation for the Master of Business Engineering
Master subject and focus: Finance
Daytime schedule

Declaration regarding AI tool Usage

During the preparation of this master's thesis, the authors utilized DeepL, GPT-3, for the following purposes:

1. Structure: those tools were used to help articulate ideas in a coherent and logically structured manner, ensuring natural flow of arguments.
2. Reformulation/Translation: those tools provided support in expressing ideas more efficiently while adhering to English academic conventions, as the authors's native language is French.

   After using the above-mentioned AI-tools, the authors diligently reviewed and edited the content produced by the tool and takes full responsibility for the final content presented in this thesis.

By signing this declaration, the authors affirm that the content of this master's thesis reflects his original work, augmented by the responsible use of AI.

Do Energy Price Variations Impact the Volatility of Inflation Expectations?

Master's Thesis

[a] *Professor Leonardo Iania*

[b] *Student Lucas Dupuis*

[c] *Student Pierre Houssiere*

Abstract

As volatility becomes a structural feature of modern economies, uncertainty affects actors at every level, from policymakers to businesses and households. Anticipating macroeconomic risks has thus become a cornerstone of informed decision-making (Adrian et al., 2019; Furno and Giannone, 2024).

This master's thesis investigates whether fluctuations in WTI crude-oil improve rolling-window, one-day-ahead forecasts of the variance of 10-year breakeven inflation rate daily change. Using daily series from the St. Louis Fed, we estimate ARMA-GARCH models with five innovation structures (ARMA(1,1)-GARCH(1,1) model compared with its longer-memory variants ARMA(1,2)-GARCH(1,2); ARMA(1,1)-GJR-GARCH(1,1); AR(1)-EGARCH(1,1), and MA(1)-EGARCH(1,1)) all with a Student's t distribution and include energy-price variables in the conditional mean, variance, or both.

This study shows that including the log-return of crude oil prices in the model does not lead to a statistically significant reduction in the mean-squared error and the mean average error of variance forecasts, as confirmed by the Diebold-Mariano (Diebold and Mariano, 1995) test.

Keywords: Inflation expectations, GARCH models, Energy prices, Volatility forecasting

---

Acknowledgements

We would like to thank Professor Leonardo Iania for guiding us through this thesis with clear advice, timely feedback, and steady support. His expertise helped us stay on track and refine our ideas at every step.

Our thanks also go to the Louvain School of Management for offering such a solid and engaging program. We are grateful to the professors we have met along the way: your passion for teaching and dedication to your fields made our studies even more worthwhile.

To our families and friends: your encouragement and understanding especially during long evenings of writing and research kept us going. We appreciate the practical help, the pep talks, and the occasional reality checks that kept us grounded.

Finally, here's to our partnership. Working together, we brought out each other's strengths and tackled challenges head-on. We are proud of what we have accomplished as a team and thankful for the shared effort that went into this project.

Contents

List of Figures

List of Tables

Introduction

In contemporary economies, price instability has become a persistent reality. Uncertainty weighs on every actor, from central banks setting monetary policy to firms planning investments and households adjusting consumption decisions. Energy-price shocks, driven by geopolitical conflicts (Liu et al., 2019), technological innovations (Baumeister, 2022) and the accelerated shift to renewables (Liu et al., 2022), constantly test our ability to maintain price stability. By directly raising production costs for goods and services, sudden spikes in oil and gasoline prices can generate significant inflationary pressures (Coibion and Gorodnichenko, 2015; Gorodnichenko et al., 2018; Hasenzagl et al., 2022), while sharp declines tend to dampen overall price dynamics.

In this context, inflation expectations are crucial for monetary credibility and economic planning, as they guide wage negotiations, pricing strategies, and investment decisions (Adrian, 2023; Blanchard and Gali, 2007; Bernanke, 2004). Well anchored expectations help to preserve monetary value and support growth, but large energy shocks such as the 1973 oil embargo or the post COVID commodity surge can unanchor them, triggering wage price spirals and costly monetary tightening (Bernanke, 2004; Kilian and Zhou, 2022d).

Because agents react asymmetrically to energy price changes with stronger responses to increases than to decreases and non-linear effects depending on shock size (D'Acunto et al., 2021a) it is therefore crucial to model and forecast the full predictive distribution. Focusing solely on its mean would fail to capture both moderate and extreme outcomes as well as evolving uncertainty (Furno and Giannone, 2024).

It is against this backdrop that the central question of our study arises: Do energy price variations affect the volatility of inflation expectations? We address this question through the specific prism of our own approach.

To develop this approach, we relied on the existing literature, which revealed three key lessons. First, GARCH-type models are effective at capturing persistence and asymmetry in the volatility of macroeconomic series (Bollerslev, 1986; Nelson, 1991; Glosten et al., 1993; Ding et al., 1993). Second, numerous studies show that adding exogenous variables, such as WTI returns or financial uncertainty indices, significantly improves volatility-forecasting accuracy (Bonnier, 2022; Batten et al., 2024). Third, despite these advances, most research focuses on the volatility of energy prices or financial indices, while few or none of them examine how oil shocks differentially impact the conditional variance of the 10-year breakeven inflation rate. These insights led us to refine our research problem into a more targeted question: To what extent does incorporating variations in energy price (WTI) as exogenous variables, enhance the performance of GARCH-type models in predicting the volatility of the 10-year breakeven inflation rate ? Throughout this paper, we will strive to answer this question by following three main steps. First, we perform exploratory data analysis, assess stationarity via KPSS and ADF tests, and examine tail behavior, skewness, volatility clustering and autocorrelation to guide model selection. Second, we verify residual white noise using the ACF and Ljung-Box test, then conduct an In-Sample evaluation, comparing AIC, BIC and Log-likelihood across candidate models. Third, we forecast the final 1245 observations of the sample (Jan 3, 2020-Dec 31 2024) with a rolling-window one-day-ahead procedure, compute MSE and MAE, and apply Diebold-Mariano tests on both loss functions (Diebold and Mariano, 1995).

The empirical findings indicate that incorporating daily WTI shocks into the conditional mean substantially enhances In-Sample performance, evidenced by a higher Log-likelihood, improved AIC and BIC scores, and reduced residual autocorrelation. Whereas their inclusion in the variance equation yields no appreciable gain. In the Out-of-Sample tests, none of the WTI-augmented specifications outperforms the ARMA(1,2)-GARCH(1,2) benchmark: forecast accuracy (MSE, MAE) and Diebold-Mariano (Diebold and Mariano, 1995) statistics remain statistically indistinguishable, underscoring the modest predictive contribution of oil price shocks for ten-year volatility.

The remainder of this dissertation is structured as follows ; The first section provides a clear review of the literature on inflation expectations and details the main GARCH models used to study them. The second section describes the reseach objective and the methodology used. The third section present the data and preliminary analyses underpinning our study. The final sections present our results and respond to key insights regarding hypotheses, limitations and opportunities for future research before concluding.

1 Literature Review

1.1 History and significance of inflation expectation

In developed countries, major central banks generally aim to stabilize prices and support economic growth. Price stability, often regarded as maintaining stable and low inflation, is a key objective of monetary policy. One way to assess whether central bank actions are aligned with this objective is to examine whether inflation expectations are in line with central banks' targets. Inflation expectations refer to the beliefs of market participants and investors about future inflation.

According to Fisher (1911), inflation corresponds to the rate of increase in prices over a given period. It is generally a global measure, reflecting changes in the general price level or cost of living in a country. In addition, the European Central Bank, since its conception, has emphasized the importance of both inflation stabilization and inflation expectations. These expectations significantly influence decisions on consumption, savings, investment and wage-setting.

Historically, inflation expectations were not always integrated into economic models, as illustrated by the original Phillips curve (Adrian, 2023). However, the era of the Great Moderation underscored their significance, particularly during major energy shocks like those in the 1970s (Bernanke, 2004) and in the more recent post-Covid period (Kilian and Zhou, 2022d).

For instance, during the 1973 oil shock, the surge in oil prices led to a steep rise in inflation (Bernanke, 2004). In response to the anticipated long-term reduction in purchasing power, households demanded higher wages (Blanchard and Gali, 2007), while firms increased their prices to offset the growing production costs (Gordon, 1975). This resulting inflationary spiral, fuelled by self-sustaining expectations, caused a challenging price-wage cycle (Alvarez et al., 2022). To counteract these inflation expectations, the Federal Reserve, led by Paul Volcker in the late 1970s, implemented stringent monetary policies, including significant interest rate hikes (Bernanke, 2004; Blanchard, 2016). These measures led to a severe recession in the early 1980s, underscoring the substantial economic consequences of central banks losing credibility in their inflation control efforts (Bernanke, 2004; Kozicki and Tinsley, 2005).

Inflation expectations play a key role in the management of oil shocks as they directly impact the reactions of households and businesses to fluctuations in energy prices. According to Lorenzoni and Werning (2023) and Alvarez et al. (2022), a loss of confidence in central banks' ability to keep inflation under control can lead to a de-anchoring of expectations, thereby reinforcing inflationary pressures. These dynamics underscore the importance of both long-term and short-term inflation expectations as key indicators of central bank credibility (Adrian, 2023; Bernanke, 2004). According to Adrian (2023) and other economists (Kozicki and Tinsley, 2005; Bernanke, 2004), this credibility is fundamental as it affects the views of households and firms on current and future economic stability (Bernanke, 2004; Blanchard and Gali, 2007), guiding their consumption, investments, and wage-setting decisions.

As a result, central banks closely monitor long-term inflation expectations (Adrian, 2023) through surveys and financial market analyses. These tools provide essential insights into the credibility of monetary policy and its transmission through financial channels (Adrian, 2023). However, since the 2008 financial crisis, the ECB and many economists have been concerned that long-term inflation expectations might become too low. This would reduce current inflation and limit central banks' ability to act, potentially resulting in an economic depression, comparable to that of 1929 (Lagarde, 2024).

Businesses and households' expectations have a crucial impact on the economy. Firms use their predictions about inflation to set prices, make investments and manage hiring to maximize profit (Gorodnichenko et al., 2018). Meanwhile, households base their consumption, saving, and financing decisions on their own inflation expectations (Andrade et al., 2020; Duca-Radu et al., 2021). These expectations also affect wage negotiations and labour supply. Market behavior can be further influenced by differences in beliefs among economic actors (Andrade et al., 2020). The post-COVID inflation crisis (2021-2023) clearly demonstrates the impact of inflation expectations on the economy. Households adjusted their spending habits, fearing a loss of purchasing power, while companies raised prices and reduced investments. Rising wage demands have fueled this inflationary spiral, compelling central banks to hike interest rates, which discourages investment and slows economic growth (Aldama et al., 2024).

Additionally, consumers' reactions to price fluctuations play a pivotal role in shaping inflation expectations. When prices of essential goods, like food or energy, go up, consumers tend to react more strongly than when they go down. This asymmetry in price perception further anchors inflation expectations (D'Acunto et al., 2021a,b). Therefore, policymakers need to closely monitor both short-term and long-term inflation expectations, as these are crucial for the credibility of central banks and the effectiveness of monetary policy. The swift disinflation of the 1980s following the Great Inflation period demonstrates how a credible policy can influence expectations and support in achieving price stability.

1.2 Review of existing literature and theoretical framework

Fluctuations in energy prices, particularly those of oil and gasoline, play a key role in the analysis of contemporary economic dynamics. These variations influence not only production costs and consumption, but also inflation expectations (Aidala et al., 2024), which is a key factor in monetary policymaking (Neri et al., 2023; Vatsa and Pino, 2024). Inflation expectations, whether short, medium or long term, significantly influence the decisions of households, businesses and central banks (Alcidi et al., 2022; Adrian, 2023). While the existing literature focuses on how external shocks affect energy prices (Vatsa and Pino, 2024; Kilian and Zhou, 2022d; Hammoudeh and Reboredo, 2018), our study adopts the opposite perspective: we examine how changes in energy prices affect the volatility of inflation expectations (10-year breakeven inflation rate). In other words, instead of measuring the variability of prices themselves, we study the impact of these energy price changes on the degree of uncertainty surrounding economic agents' inflation forecasts. This innovative approach makes it possible to determine whether rises or decreases in energy prices increase or reduce the instability of inflation expectations, a key factor in monetary policy and investment decisions. This literature review aims to explore the different frameworks, methods, and conclusions that have emerged around this topic, highlighting both the empirical results and the theoretical implications.

Given the many geopolitical and economic events of recent years, such as the Covid-19 pandemic (Batten et al., 2024) and Russia's invasion of Ukraine (Kilian and Zhou, 2023), global economies have been strongly affected. The consensus that emerges in most research is that increases in energy prices are a key driver of inflation expectations, especially in the short and medium term (e.g. Boeck and Zörner (2025); Hammoudeh and Reboredo (2018); Abdallah and Kpodar (2023); Kilian and Zhou (2022d)). These fluctuations also have a significant impact on monetary policy, as indicated by several central banks (on Inflation Expectations, 2021). However, understanding how energy variables influence the volatility of these expectations, especially over a long period and across different economic phases, remains a complex topic and is relatively unexplored by peers.

Regarding the impact of energy price shocks on inflation expectations, Neri et al. (2023) show that energy price shocks directly influence inflation and can affect inflation expectations, thus changing economic behaviors and monetary policy decisions (Neri et al., 2023). Since 2021, rising energy prices, amplified by geopolitical factors, have been a major contributor to inflation, accounting for 60% of headline inflation in Q4 2022 depending on the model used (Neri et al., 2023). Furthermore, the high visibility of energy prices increases the responsiveness of households, as they perceive them as one of the most immediate and tangible indicators of inflation (Binder, 2018).

Kilian and Zhou (2022d) take a significantly different approach by demonstrating that oil price increases are also the result of fluctuations in demand in the United States, not just of supply shocks (e.g., (Kilian and Zhou, 2022d; Kilian, 2009; Kilian and Murphy, 2014; Kilian and Zhou, 2020). Understanding this distinction is crucial for central banks to manage inflation expectations and mitigate long-term de-anchoring risks.

More recently, Boeck and Zörner (2025) studied the impact of natural gas price fluctuations in the euro area. Their analysis reveals that gas price rises lead to an instantaneous increase in inflation expectations, peaking about a year after the initial shock. They highlight the indirect effects, where the increase in inflation expectations amplifies the economic repercussions of energy shocks, which consequently reinforce headline inflation and its anchoring in price-wage dynamics.

Interest in a link between oil prices and inflation began in the 1970s due to two major oil shocks and persistent consumer price inflation (Bernanke, 2004)). The two oil shocks of the 1970s refer to the OPEC oil embargo in October 1973 as well as the Iranian revolution of 1979, that led to a sharp fall in oil production in Iran and was aggravated by the Iran-Iraq war in 1980 (Blanchard and Gali, 2007). At the time, it was widely accepted that this persistent inflation was a direct result of sharp increases in oil prices, following events unprecedented in post-war history (Blinder, 1982). This assumption held that oil prices were primarily driven by exogenous supply shocks related to geopolitical events, thereby influencing U.S. inflation. However, subsequent research, including that of Kilian (2009) and, Kilian and Murphy (2014), has challenged this assumption by incorporating the impact of demand and supply changes on global oil markets. These studies show that oil price fluctuations are not exclusively due to supply shocks but are also the result of complex dynamics involving macroeconomic factors.

An alternative explanation, proposed by Barsky and Kilian (2004) and elaborated on by Kilian (2010), links the oil price increases of the 1970s to changes in global monetary policy regimes. According to this approach, expansionary monetary policies would have led to an increase in global demand for commodities, including oil, that would have supported higher headline inflation. Thus, rather than being a cause of inflation, the rise in oil prices would be more a symptom of an economic period marked by excessive monetary expansion. This assumption is supported by the observation that inflation had already started to rise before the oil shocks, partly due to regulatory and contractual constraints that delayed the adjustment of oil prices (Kilian, 2010).

Although not directly included in the basket of goods consumed by households, the price of crude oil plays a central role in the evolution of inflation (Kilian and Zhou, 2023). Indeed, these fluctuations directly influence the costs of goods and services, contributing to changes in consumer price indexes (Chudik and Georgiadis, 2022).

An unexpected and persistent rise in oil prices can cause a broad-based increase in prices, but its impact on inflation is mostly temporary (Vatsa and Pino, 2024). However, some economists, such as Blanchard (1986), warn of the risk of persistent inflation caused by a wage-price spiral as introduced in 1.1 This phenomenon occurs when workers demand wage increases to compensate for the loss of purchasing power due to inflation, prompting companies to raise prices to absorb these additional costs Alvarez et al. (2022); Bernanke (2004). This inflationary dynamic was particularly noticeable in the United Kingdom in the 1970s, where the role of trade unions exacerbated this inflationary loop. However, the wage-price spiral has only been observed to a limited extent in the United States because the protection of workers is much less important (Blanchard and Gali, 2007).

Variations in oil prices undoubtedly influence inflation directly and indirectly, impacting households, businesses and global economic expectations. Managing these fluctuations is therefore a major challenge for economic policymakers and central banks, which must adapt their policies to the dynamics of the oil market in order to ensure price stability and economic growth (Abdallah and Kpodar, 2023).

In our study, we implemented the GARCH model (Bollerslev, 1986) as it is widely recognized in finance and economics for capturing market fluctuations and predicting financial instabilities (Batten et al., 2024; Byun and Cho, 2013). GARCH-type specifications, notably the EGARCH (exponential GARCH) model (Nelson, 1991) and the GJR-GARCH (Glosten et al., 1993) model, extend the basic framework by allowing asymmetry in the transmission of positive and negative volatility shocks. These extensions are particularly useful for modelling energy markets, where price increases and decreases can have very different effects. In addition, we incorporate an ARMA(p,q) (Yule, 1927; Slutsky, 1937; Box and Jenkins, 1970) component into the averaging equation to capture the possible linear autocorrelation of returns and prevent volatility estimates from being contaminated by persistent non-volatile dynamics. This ARMA-GARCH framework, enriched by exogenous variables such as WTI, follows the methodology of Bonnier (2022) and Batten et al. (2024), who show that including these external indicators significantly improves the accuracy of energy price volatility forecasts. This exogenous approach makes it easier to capture the effect of external shocks such as geopolitical tensions or financial crises on market dynamics. However, as no previous study has explored this perspective, we will be able to rely only on their methodology but not on their findings that greatly differ from ours (Bonnier, 2022; Kilian and Zhou, 2022d; Hammoudeh and Reboredo, 2018; Vatsa and Pino, 2024).

1.3 Hypotheses

H1: Energy price fluctuations significantly influence the volatility of 10-year breakeven inflation rate daily change. This hypothesis reflects the idea that oil price shocks have immediate effects on inflation expectations, due to their direct impact on production costs and household consumption (Neri et al., 2023; Kilian and Zhou, 2022d; Neri et al., 2023).

H2: Including energy prices (WTI) as exogenous variables simultaneously in both the conditional mean and the conditional variance equations in our model will yield superior In-Sample fit and Out-of-Sample forecasting accuracy compared to specifications that include these exogenous regressors only in the mean equation or only in the variance equation (Diop and Kengne, 2021; Batten et al., 2024).

H3: The impact of energy price volatility on 10-year breakeven inflation rate daily change, varies across economic regimes (e.g. crisis vs. growth periods). The correlation is stronger during periods of high macroeconomic uncertainty (e.g. post-COVID, financial crises) and weaker during stable growth periods (Batten et al., 2024; Kilian and Zhou, 2023).

2 Methodology

2.1 Research Objective

The objective of this master's thesis is to highlight the contribution of energy price shocks to the dynamics of volatility on the 10-year breakeven inflation. More specifically, it aims to determine the extent to which the inclusion in the GARCH models of daily variations in the prices of crude oil (WTI) as exogenous variables improves the quality of volatility forecasts at ten-year horizons.

Our empirical approach starts with data diagnostics and feature transformation, followed by stationarity testing and distributional diagnostics. It then deals with residual autocorrelation checks and white-noise validation, proceeds to In-Sample model estimation and comparison, and culminates in Out-of-Sample forecasting and accuracy evaluation.

2.2 GARCH Models

In this study, we adopt the GARCH (Generalized Autoregressive Conditional Heteroskedasticity) model to capture the conditional volatility of the daily change in the 10-year breakeven inflation rate. The specification of the GARCH model is based on Engle (1982) ARCH model by allowing the conditional variance to depend not only on past squared shocks but also on its own lagged values, hereby limiting the amount of coefficients to estimate and preventing overfitting.

The standard GARCH(p,q) model specifies

$$\begin{cases} r_t^Y = \mu + \varepsilon_t, \\ \varepsilon_t = \sigma_t z_t, \quad z_t \sim D(0,1), \\ \sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2, \end{cases} \tag{1}$$

subject to the stationarity condition

$$\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1. \tag{2}$$

Here,

$$r_t^Y = (Y_t - Y_{t-1}), \quad Y_t \in \{\text{T10YIE}\}. \tag{3}$$

with $\omega > 0, \quad \alpha_i \geq 0, \quad \beta_j \geq 0$. Where μ represents the constant mean of the first differences series in our case, the daily change of 10-year breakeven inflation rate, while the innovation $\varepsilon_t = r_t^Y - \mu$ captures the unpredictable deviation of the actual first differences from this mean. The term ω serves as the baseline level of conditional variance, and the coefficients α_i and β_j quantify how past squared innovations and past variances, respectively, feed into today's volatility. Finally, z_t is an i.i.d. standardized random variable drawn from a distribution $D(0,1)$, representing the innovation distributions considered in our study (Student's t).

However, the Ljung-Box test highlights that our sample shows a non-negligible autocorrelation, which indicates that the constant average was not able to capture all of the observed time dynamics. We have therefore included the AR (autoregressive), MA (moving average) and ARMA (Autoregressive-Moving Average) component in our mean, in order to reduce the autocorrelation of the residuals ε_t (Tsay, 2010; Hamilton, 1994; Box et al., 2008).

To capture the linear dependence of past first differences in our mean forecast, we introduce the autoregressive component. Define the one-step-ahead forecast,

$$r_t' = \mu + \sum_{i=1}^{p} \phi_i \, r_{t-i}^Y, \tag{4}$$

so that the innovation is,

$$\varepsilon_t = r_t^Y - r_t'. \tag{5}$$

The full AR-GARCH(p,q) is defined as:

$$\begin{cases} r_t^Y = r_t' + \varepsilon_t, \\ \varepsilon_t = \sigma_t \, z_t, \quad z_t \sim D(0,1), \\ \sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \, \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \, \sigma_{t-j}^2. \end{cases} \tag{6}$$

In this AR-GARCH specification, the AR extension introduces two new elements: the one-step-ahead forecast r_t', which replaces the constant mean in the prediction of inflation expectation first differences, and the autoregressive coefficients ϕ_i, which link past first differences r_{t-i}^Y to today's forecast.

The full MA-GARCH(p,q) is defined as:

$$\begin{cases} r_t^Y = \mu + \sum_{\ell=0}^{Q} \theta_\ell \, \varepsilon_{t-\ell}, \quad \text{where } \theta_0 = 1, \\ \varepsilon_t = \sigma_t \, z_t, \quad z_t \sim D(0,1), \\ \sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \, \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \, \sigma_{t-j}^2. \end{cases} \tag{7}$$

The full ARMA-GARCH(p,q) is defined as:

$$\begin{cases} r_t^Y = r_t' + \sum_{\ell=1}^{q} \theta_\ell \, \varepsilon_{t-\ell} + \varepsilon_t, \\ \varepsilon_t = \sigma_t \, z_t, \quad z_t \sim D(0,1), \\ \sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \, \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \, \sigma_{t-j}^2. \end{cases} \tag{8}$$

The daily logarithmic return[1] of the exogenous variable is defined as

$$x_t^P = \ln\left(\frac{P_t}{P_{t-1}}\right). \tag{9}$$

where P_t is the series price at time t Bollerslev (1986); Engle and Patton (2001).

In addition, to analyse the impact of 10-year breakeven inflation rate daily change on energy prices (WTI), we extend the usual ARMA-GARCH framework by introducing the exogenous variable into both the mean and variance equations. Concretely, following the approach of several authors (Engle, 2001; Kur et al., 2021; Diop and Kengne, 2021; Tsay, 2010), we consider three specifications:

- ARMA-X-GARCH, where x_t^P enters only the conditional-mean equation, aiming to quantify the direct impact of energy shocks on inflation expectations;
- ARMA-GARCH-X, where x_t^P enters only the conditional-variance equation;
- ARMA-X-GARCH-X, where x_t^P appears in both the conditional-mean and conditional-variance equations.

[1] see appendix section 5.3 for further explanations

Compared to the baseline ARMA-GARCH specification, the ARMA-X-GARCH, ARMA-GARCH-X and ARMA-X-GARCH-X models introduce the exogenous series x_t^P in different parts of the system:

ARMA-X-GARCH (mean-augmented):

$$r_t^Y = \mu + \sum_{i=1}^{p} \phi_i \, r_{t-i}^Y + \sum_{\ell=1}^{q} \theta_\ell \, \varepsilon_{t-\ell} + \delta \, x_t^P + \varepsilon_t. \tag{10}$$

where x_t^P enters only the conditional mean via δ, and the variance equation is unchanged from the standard GARCH.

ARMA-GARCH-X (variance-augmented):

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \, \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \, \sigma_{t-j}^2 + \psi x_t^P, \tag{11}$$

where x_t^P enters only the conditional variance via ψ, and the mean equation remains the AR-GARCH baseline.

ARMA-X-GARCH-X (mean and variance augmented):

$$\begin{cases} r_t^Y = \mu + \sum_{i=1}^{p} \phi_i \, r_{t-i}^Y + \sum\limits_{\ell=1}^{q} \theta_\ell \, \varepsilon_{t-\ell} + \delta \, x_t^P + \varepsilon_t, \\ \sigma_t^2 = \omega + \sum\limits_{i=1}^{p} \alpha_i \, \varepsilon_{t-i}^2 + \sum\limits_{j=1}^{q} \beta_j \, \sigma_{t-j}^2 + \psi x_t^P, \end{cases} \tag{12}$$

where x_t^P influences both the conditional mean (through δ) and the conditional variance (through ψ).

2.3 Other GARCH Models

In order to test for asymmetry (the leverage effect), we will employ two extensions of the standard GARCH model: the GJR-GARCH (Glosten et al., 1993) in which we incorporate an ARMA component in the mean equation, and the EGARCH model (Nelson, 1991), for which we specify separate AR and MA components.

Firstly, the GJR-GARCH model extends the standard GARCH model by an asymmetry term γ_i

via an indicator variable (Glosten et al., 1993), which multiplies the square of the error term of the previous period in the variance equation. This indicator equals to 1 when the previous shock is negative, thereby activating an additional variance component, and 0 otherwise, that effectively capturing the leverage effect.

The GJR-GARCH(p,q) model is defined as:

$$h_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}^2 + \sum_{i=1}^{p} \gamma_i \varepsilon_{t-i}^2 \mathbb{I}_{\{\varepsilon_{t-i}<0\}} \tag{13}$$

Where the indicator variable:

$$\mathbb{I}_{\{\varepsilon_{t-i}<0\}} = \begin{cases} 1, & \text{if } \varepsilon_{t-i} < 0, \\ 0, & \text{otherwise.} \end{cases} \tag{14}$$

Here, if $\mathbb{I}_{\{\varepsilon_{t-i}<0\}} = 1$ and $\gamma_i > 0$, the model places extra weight on the impact of a negative shock amplifying its effect on future volatility and thus capturing the leverage effect.

Secondly the EGARCH model, makes it possible to take into account the leverage effect where negative shocks have a greater impact on volatility than positive shocks of the same intensity (Nelson, 1991). A key characteristic of the EGARCH model is that, rather than directly modelling the conditional variance, it models the logarithm of this conditional variance.

The EGARCH(p,q) model is mathematically defined by:

$$\log h_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \left(\frac{\varepsilon_{t-i}}{h_{t-i}}\right) + \sum_{j=1}^{q} \beta_j \log h_{t-j}^2 + \sum_{i=1}^{p} \gamma_i \left[\left|\frac{\varepsilon_{t-i}}{h_{t-i}}\right| - \mathbb{E}\left(\left|\frac{\varepsilon_{t-i}}{h_{t-i}}\right|\right)\right] \tag{15}$$

Here, if $\gamma_i < 0$, negative shocks increase log-variance more than positive shocks of equal size.

2.4 Model Specifications

In the light of the different autocorrelation dynamics between the daily change in 10-year breakeven inflation rate and additionnal changes caused by adding the log returns of exogenous energy price variables, we estimate four tailored GARCH models. For each case, we change the way and the place where the exogenous variable x_t^P enters the system (mean equation, variance equation, or both):

- **ARMA(1,1)-GARCH(1,1)**: we estimate three specifications in which the exogenous log-return x_t^P is included: only in the conditional mean, only in the conditional variance, or in both equations.

- **ARMA(1,2)-GARCH(1,2)**: based on the previous model, we extend the GARCH component to order (1,2) to allow for longer memory in conditional variance, The exogenous regressor x_t^P is tested in the mean and joint mean-variance setups.
- **ARMA(1,1)-GJR-GARCH(1,1)**: this variant replaces the GARCH term with a GJR-GARCH specification to accommodate potential asymmetry (leverage effects) in volatility. As ARMA(1,1)-GARCH(1,1), x_t^P may enter the mean, the variance, or both.
- **AR(1)-EGARCH(1,1)**: here, the conditional mean is simplified to an AR(1) component because the ARMA component did not offer significant results for this model. In this case, like for ARMA(1,2)-GARCH(1,2), the exogenous regressor x_t^P is tested in the mean and joint mean-variance setups.
- **MA(1)-EGARCH(1,1)**: similar to the AR(1)-EGARCH(1,1) model, this specification uses a MA(1) process for the conditional mean, combined with EGARCH(1,1) for the variance. Again, the inclusion of x_t^P follows the same scenarios as the AR(1)-EGARCH(1,1) model.

By comparing these four specifications, we assess not only the direct effects of energy shocks on inflation expectations but also how they propagate through conditional volatility under different dynamic assumptions.

2.5 Estimation

2.5.1 Ljung-Box Tests and ACF Diagnostics

During the In-Sample period, we complement residual diagnostics with Ljung-Box tests (Ljung and Box, 1978) on standardized residuals and their squares (up to lag 20) to detect any remaining autocorrelation and confirm that the residuals behave like white noise. In parallel, we inspect the sample autocorrelation function ACF (Box et al., 2008) of both series to visually verify the absence of significant serial dependence and white-noise characteristics after model fitting.

2.5.2 Log-likelihood, AIC and BIC

We report the maximized Log-likelihood for each fitted model to assess absolute fit. To balance goodness-of-fit against complexity, we compute the Akaike Information Criterion (AIC; Akaike, 1974) and Bayesian Information Criterion (BIC; Schwarz, 1978) where a lower AIC and BIC values indicate a more parsimonious model that still adequately explains the data.

$$\ell(\hat{\theta}) = \max_{\theta} \ell(\theta), \quad \text{AIC} = -2\,\ell(\hat{\theta}) + 2k, \quad \text{BIC} = -2\,\ell(\hat{\theta}) + k\ln(n) \tag{16}$$

2.5.3 Forecast Accuracy: MSE and MAE

Our Out-of-Sample performance is evaluated via Mean Squared Error (MSE; Gauss, 1823) and Mean Absolute Error (MAE; Armstrong and Collopy, 1992) between the variance forecasts σ^2_{t+1} and the predicted first difference r'^{Y}_{t+1}. Then the MAE and MSE for the variance forecasts are computed as:

$$\text{MAE}_{\text{var}} = \frac{1}{|\mathscr{T}|} \sum_{t \in \mathscr{T}} \left| (r^{Y}_{t+1})^2 - \sigma^2_{t+1} \right|, \quad \text{MSE}_{\text{var}} = \frac{1}{|\mathscr{T}|} \sum_{t \in \mathscr{T}} \left((r^{Y}_{t+1})^2 - \sigma^2_{t+1} \right)^2. \tag{17}$$

Where $\mathscr{T} = \{4237, 4238, \dots, 5281\}$ designates the set of time indices for which forecasts are made.

2.5.4 Diebold-Mariano Test

To formally compare predictive accuracy between competing models on an Out-of-Sample dataset of $T = 1245$ observations, we apply the Diebold-Mariano (DM; Diebold and Mariano, 1995) test to the loss-differential series. Denoted by

$$d_t^{(i,j)} = L\big(e_t^{(i)}\big) \; - \; L\big(e_t^{(j)}\big), \tag{18}$$

where $e_t^{(i)}$ is the forecast error of model i at time t and $L(\cdot)$ is a loss function (e.g. $L(e) = e^2$ for MSE or $L(e) = |e|$ for MAE). The DM statistic for comparing model i (row i in our result matrix : see Figure 18) against model j (column j) is then

$$\text{DM}_{i,j} = \frac{\overline{d^{(i,j)}}}{\sqrt{\widehat{\text{Var}}\big(\overline{d^{(i,j)}}\big)}} \;\approx\; \frac{\frac{1}{T}\sum_{t=1}^{T} d_t^{(i,j)}}{\sqrt{\frac{1}{T}\sum_{t=1}^{T}\big(d_t^{(i,j)} - \overline{d^{(i,j)}}\big)^2}}, \tag{19}$$

which under the null hypothesis of equal predictive accuracy follows approximately a standard normal distribution. In our implementation, we organize the DM statistics and corresponding p-values into a 4×4 matrix, where the entry in row i, column j reflects the comparison of model i versus model j.

3 Data Collection

Following the methodological approach, we present here the data and preliminary analyses underpinning our study. We start by describing the dependent variables, 10-year breakeven inflation rate (section 3.1). Next, we introduce the exogenous regressor, WTI crude oil (section 3.2). Finally, we conduct descriptive and statistical diagnostics, including measures of tail behavior, skewness, volatility clustering, and autocorrelation, and perform stationarity tests to guide model selection (section 3.3.1).

To ensure robust modeling and proper alignment between variables, our analysis is based on two slightly offset time ranges to enable the calculation for the exogenous variable while securing perfect alignment with the first differences of our dependent variables.

- Dependent variables (10-year breakeven inflation rate) span from January 2, 2003 trough December 31, 2024, i.e. 5482 observations and therefore 5481 first differences,
- Exogenous variable (spot price: WTI) covers the same period, so that its daily log-return series aligns exactly with the 5 481 changes in the breakeven rate.

These aligned and high-frequency datasets consequently provide a comprehensive foundation for our GARCH analysis of the volatility of inflation expectations.

3.1 Dependent Variables

As the dependent variable, we adopt the 10-year break-even inflation rate (T10YIE), which is a market-derived indicator of average annual inflation expectations over the coming decade. Specifically, it measures the difference, in percentage points, between the yield on a nominal 10-year Treasury security with a constant maturity (BC_10YEAR) and the yield on a 10-year Treasury security with a constant maturity indexed to inflation (TC_10YEAR). By definition, this measure represents the average inflation rate that would equalise the yields on nominal bonds with those on real bonds. From this series, we calculate the daily change in breakeven (in percentage points) as the main time series for volatility modelling, hereby isolating short-term fluctuations and avoiding the trend and unit root problems inherent in level data.

We obtain daily breakeven observations from the Federal Reserve Economic Data (Fred; Federal Reserve Bank of St. Louis, 2025a), with the understanding that these yields have been directly supplied by the U.S. Department of the Treasury. The reason we adopt a daily sample frequency is the requirement for a rich data set in volatility modelling. As demonstrated by Bollerslev (1986), the GARCH model requires at least about 1,000 observations to ensure stable identification of the dynamics of conditional heteroskedasticity. In addition, we focus on daily variations (rather than on the level) as our interest lies in the variability of expectations and because the level of the series may contain trends or unit roots.

The literature strongly supports the use of market-based ten-year breakeven inflation rates for the analysis of long-term inflation expectations (e.g. Lejsgaard Autrup and Grothe (2014), Pincheira (2013)). Unlike consumer price indices or survey measures, the ten-year breakeven inflation rate provides real-time forward-looking information derived directly from transactions in both nominal and inflation-indexed Treasury bills, and is less likely to be subjected to mechanical revisions or anchoring biases (Kose et al., 2019). Autrup and Grothe (2014) furthermoreshow that a GARCH(1,1) model applied to daily changes in the break-even point over ten years effectively captures both volatility clustering and volatility spikes associated with macroeconomic shocks.

3.2 Indenpendent Variables

In addition to the 10-year break-even inflation rate, energy prices represent a key element in our analysis, as they are potential exogenous factors in the volatility of inflation expectations. To capture their influence and represent energy market mouvements, we apply the daily log return of the West Texas Intermediate (WTI) crude-oil spot price from FRED (series DCOILWTICO; Federal Reserve Bank of St. Louis (2025b)).

WTI is a key benchmark in the North American oil market. Its price influences both production and transmission costs, as well as broader energy prices (Batten et al., 2024). It is widely referenced in energy inflation literature due to its strong correlation with global inflationary shocks and its central role in commodity markets (Bernanke et al., 1997; Harris et al., 2009; Peersman and Van Robays, 2009; Coibion and Gorodnichenko, 2015; Aastveit et al., 2023; Wong, 2015; Bjørnland et al., 2018).

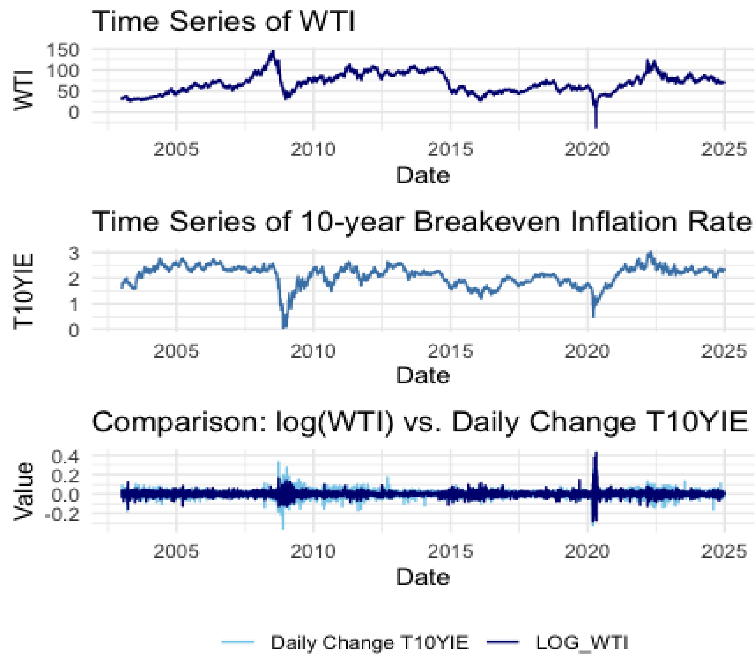

Figure 1: Daily time series of study variables, 2 Jan 2003 - 1 Jul 2024).

3.3 Additionnal Tests on Data

Following the presentation of the data and the description of the transformations applied, it is crucial to understand the distribution and statistical properties of our variables before examining the analysis results in section 4. The descriptive statistics offer a preliminary overview of the trends, volatility and asymmetries found in the data studied.

The descriptive statistics for our independent variables (energy price: WTI) and dependent variables (10-year breakeven inflation rate) are displayed in Table 10, corresponding to the gross and transformed values descriptive statistics. The statistics are calculated across four sub-periods (2003-2006, 2007-2009, 2010-2019, 2020-2024) as well as for the full sample (2003-2024).

Our methodological approach is inspired by Batten et al. (2024), which explores the interplay between inflation and energy price volatility. Our tables' structure and content are based on their framework, particularly when it comes to the relevance of specific distributional measures and temporal segmentation of the data.

The division of the sample into four distinct periods is based on graphical evidence from the raw and transformed series (see Figure 1). These visualizations reveal heightened variability, particularly during the 2007-2009 and 2020-2024 periods.

This segmentation enables the detection of potential volatility clusters through comparative analysis of standard deviations across sub-periods. Such clustering is indicative of heteroscedasticity, an important feature in financial time series, that justifies the application of conditional volatility models. This analytical strategy, in line with Batten et al. (2024) position, offers insights into the dynamic behavior of inflation expectations and energy prices.

We focus on six statistical descriptors (Min, Mean, Stdev, Max, Skew, Kurt) due to their usefulness in identifying macro-financial data. The standard deviation that measures volatility, is particularly interesting, as is Skewness and Kurtosis' information on the asymmetry and tail behavior of the distributions. Additionally, we conduct the Jarque-Bera test to formally assess deviations from normality. These properties are critical when modelling non-linearities and asymmetries, especially using GARCH-type models that account for fat tails and leverage effects. Furno and Giannone (2024), recent literature aligns with these methodology choices., emphasizing the need to capture such features for robust forecasting.

Finally, this comparison facilitates a preliminary assessment of whether fluctuations in energy price volatility are mirrored in inflation expectation measures. In section 3.3.4.1 and 3.3.4.2, we will formally test for stationarity using the Augmented Dickey-Fuller (ADF) and KPSS tests on all transformed series to rule out unit roots, and in section 4 we will conduct ARCH effect tests to detect volatility clustering and justify the use of GARCH models.

3.3.1 Summary Statistics

3.3.2 Volatility Clusters and Sub-Period Variability

Our analysis of Table 10 descriptive statistics reveals several critical insights into the volatility and influence of energy prices.

In the first place, the analysis of standard deviations (Stdev) reveals significant variations across the different sub-periods and the entire sample, confirming the presence of volatility clusters. This insight validates the application of GARCH models, particularly for both variables, which exhibits fluctuating volatility depending on the period.

For instance, the standard deviation of the T10YIE variable is notably higher in the 2020-2024 (0.60) and 2007-2009 periods (0.41) than between 2011-2019 (0.30), indicating a more volatile environment. More specifically, the strong variability of inflation data during economic crises or periods of instability (e.g. the 2008 financial crisis and post-2020 inflationary shocks) has been identified. (e.g., the 2008 financial crisis and post-2020 inflation shocks). This pronounced variability during economic crises or instability (e.g. such as the 2008 financial crisis and the post-2020 inflation surge; see figure 1) illustrates temporal heterogeneity in volatility for both WTI and T10YIE, underscoring the suitability of GARCH models to capture these dynamics.

3.3.3 Distributional Characteristics

We will add to this investigation an analysis of the higher-order moments (kurtosis and skewness) for the transformed data (Log return WTI and Daily Change of 10-year breakeven inflation rate).

As shown in Table 10, across all four sub-samples, the analysis of kurtosis (Kurt)[2] reveals clearly fat tails for Log Return WTI (31.89), Daily Change T10YIE (12.61). We also applied the Jarque-Bera test to confirm this result in the descriptive statistics, which show a p-value ≈ 0, which indicates extreme tailing behaviour that is incompatible with normality.

We find that the WTI Log return series has a slight positive skewness (0.67) and extreme kurtosis (31.89), while the daily T10YIE changes have a slight negative skewness (-0.37) and high kurtosis (12.61). The positive skewness of Log WTI suggests that large upward shocks to oil yields are more frequent than downward shocks, while the negative skewness of inflation changes indicates heavier downward tails for equilibrium rate shocks. The high kurtosis of both series, however, confirms the presence of fat tails.

In conclusion, the extreme flattening and asymmetry that we observe in the upward spikes in oil yields and the heavier downward tails in inflation expectations demonstrate that occasional large-scale shocks occur much more frequently than expected from a Gaussian distribution. To account for this, our GARCH models use a Student t-distribution. In addition, we also perform a rigorous stationarity test on each transformed series using ADF and KPSS tests prior to estimation, ensuring that our volatility models produce reliable and unbiased results.

Variables	**N**	**Skewness**	**Kurtosis**	**JB**	**p-value**
LOG WTI	5481	0.67	31.89	232701	<0.0001
Daily change T10YIE	5481	−0.37	12.61	36421.76	<0.0001

Table 1: Jarque–Bera test on full sample

[2]Note that the values reported in Table 10 correspond to *excess* kurtosis (i.e., raw kurtosis minus 3). Any positive value indicates heavier tails than a Normal distribution.

3.3.4 Stationarity of the Variables

The analysis of the descriptive statistics in our data series revealed the need of assessing their stationarity, and verifying whether their transformation into first differences and Log returns maintains this property. Stationarity is crucial for the application of GARCH models, which rely on time series exhibit consistent statistical properties (mean, variance, and, autocorrelation) in the long term.

To study the stationarity of our selected variables, we applied two complementary statistical tests: the KPSS test (Kwiatkowski et al., 1992) and the ADF test (Dickey and Fuller, 1979), with its ADF extension Augmented by Said and Dickey (1984)). Following the Statsmodels library manual's approach, we combined these tests to provide a compehensive analysis of the series' stationarity.

3.3.4.1 KPSS

The KPSS test (Kwiatkowski et al., 1992) is used to determine whether a time series is stationary around a mean or a deterministic trend. Unlike the ADF test, that checks the presence of a unit root and tries to confirm non-stationarity, the KPSS test uses a contrasting approach. Its null hypothesis asserts that the series is stationary, and tries to prove wrong. Therefore, if the test statistic exceeds the critical values, the stationarity hypothesis is rejected, indicating that the series is non-stationary. Conversely, if the statistic is below the critical values, the stationarity assumption holds, and the series can be considered stationary.

In this study, we have applied the type μ variant of the KPSS unit root test to our series. The results of this test give us KPSS statistics for Log WTI (0.052) and for daily variation T10YIE (0.032), which are well below the critical values of 10%, 5%, 2.5% and 1% (0.347, 0.463, 0.574 and 0.739). We therefore do not reject the null hypothesis of stationarity for the two series, which confirms that the transformed data are stationary.

Variables	**Test statistic**	**10%**	**5%**	**2.5%**	**1%**
LOG WTI	0.052	0.347	0.463	0.574	0.739
Daily change T10YIE	0.032	0.347	0.463	0.574	0.739

Table 2: KPSS unit root test on full sample (type: μ, 10 lags)

3.3.4.2 ADF

The Dickey-Fuller test (Dickey and Fuller, 1979) and its Augmented ADF version (Said and Dickey, 1984) examine the presence of a unit root in a time series, indicating non-stationarity Table 3. Unlike the KPSS test, which aims to verify stationarity, the ADF test starts null hypothesis that the series is non-stationary and attempts to reject it. Therefore, if the test statistic is lower than the critical values, the null hypothesis is rejected, confirming stationarity. Conversely, if the statistic exceeds the critical values, the hypothesis of non-stationarity stands. We therefore use the drift specification with 10 lags to integrate any autocorrelation in the residuals.

The results from ADF tests for both series (-51.90 for Log WTI and -51.62 for daily change T10YIE) are significantly below the 1% critical values (-3.43). Consequently, we definitively reject the null hypothesis of a unit root at all conventional levels of significance. These results are consistent with those of the KPSS test, which also demonstrated the stationarity of both series. This therefore confirms that both the Log-returns of WTI prices and the daily changes in the 10-year breakeven inflation rate are stationary, justifying their direct use in GARCH-type volatility models and ensuring robust, reliable estimation and forecasting.

Variable	**Test statistic**	**1%**	**5%**	**10%**
LOG WTI	−51.90	−3.43	−2.86	−2.57
Daily change T10YIE	−51.62	−3.43	−2.86	−2.57

Table 3: Augmented Dickey–Fuller test on full sample (drift, 10 lags)

3.3.4.3 ARCH Effect Test

Finally, to check for conditional heteroskedasticity, we apply the Engle ARCH test to the squared residuals of the daily change 10-year breakeven inflation rate series. This involves regressing ε_t^2 on a constant and its own lags, then testing whether all lag coefficients are jointly zero.

The results presented in Table 4 indicate that both the LM statistic and the F-statistic exceed the usual 5% or 1% critical thresholds, meaning that we can reject the null hypothesis of no ARCH effect (constant conditional variance). This implies that the squared residuals display significant autocorrelation, in other words that large shocks tend to be followed by large shocks and small shocks by small shocks, resulting in a clustering characteristic of volatility.

Test	Value
LM statistic	881.40
LM p-value	0.00
F-statistic	87.36
F-test p-value	0.00

Table 4: Results of ARCH Effect Test

4 Results

4.1 In-Sample

Before delving into the In-Sample results, please note that the data and figures corresponding to each section of this analysis are provided in the Appendix (from Figures 3 to Figure 17), organized in the same order as they appear in the main text.

4.1.1 ARMA(1,1)-GARCH(1,1)-Student

The estimation of the four model variants demonstrates that the AR(1) coefficient ϕ_1 and the MA(1) coefficient θ_1 are consistently significant ($p < 0.05$) across all specifications. In the base model, ϕ_1= -0.28 (p=0.047) and θ_1= 0.34 (p=0.014). When the exogenous variable is introduced into the mean equation (models "XM" and "XM_XV"), these estimates become ϕ_1= -0.3871 (p=0.023) and θ_1 = 0.42 (p=0.011), indicating an improved fit; the exogenous term itself has a coefficient of δ= 0.37 (p=0), underlining its high significance. Conversely, including the exogenous variable in the variance equation (models "XV" and "XM_XV") yields a coefficient of zero with p=1.000, indicating no detectable effect on conditional volatility. This consistent significance of the ARMA coefficients confirms our decision to extend the model in order to minimize any residual autocorrelation.

When we apply the Ljung-Box test to the standardized residuals and their squares at 20 lags, the base model shows no significant autocorrelation in the residuals themselves (Q(20)=25.75, p=0.17) but strong autocorrelation in the squared residuals (Q(20)=51.17, p=0.0002), revealing persistent volatility clustering. As we introduce exogenous regressors, the test statistic on squared residuals steadily falls, reaching (Q(20)=33.18, p=0.0322) in the "XM" model and (Q(20)=33.13,p=0.0327) in "XM_XV", approaching the threshold of non-significance. This pattern highlights that each additional exogenous variable brings us closer to the theoretical limit of a perfectly "whitened" series.

All models confirm a robust GARCH(1,1) structure. The constant term ω is extremely close to zero yet remains significant (p≈0.0), indicating a low baseline volatility. The ARCH parameter $\alpha_1 \approx 0.06$ and the GARCH parameter $\beta_1 \approx 0.93$ add up to approximately 0.99, reflecting high volatility persistence. The shape parameter $\nu \approx 8$ under the Student-t distribution highlights the presence of heavy tails, that captures extreme movements in the series and confirms our choice of the Student-t distribution for this model.

The Log-likelihood and information criteria for each model are as follows: the base model yields a LogLik = 11733.9, AIC = -4.28, and BIC = -4.27. Introducing the exogenous variable in the mean ("XM") improves the Log-likelihood to 11986.6 and lowers AIC and BIC to -4.32 and -4.36, respectively. The "XV" model remains identical to the base model in AIC and BIC, confirming the ineffectiveness of the variance exogenous. The combined "XM_XV" specification does not outperform "XM" alone, as the exogenous in the variance remains insignificant.

Given the consistent parameter significance, the improvements in information criteria, and the satisfactory residual diagnostics, the ARMA(1,1)-GARCH(1,1) model with the exogenous variable included only in the mean equation ("XM") emerges as the best-performing In-Sample specification. It combines a significant additional mean adjustment (δ= 0.37, p=0) with high volatility persistence ($\alpha_1 + \beta_1 \approx 0.99$), and the "XM" specification further brings the standardized residuals closer to white noise. This makes it a strong candidate for our Out-of-Sample analysis.

Model	**Log-likelihood**	**AIC**	**BIC**
ARMA(1,1)-GARCH(1,1) : Base model	11 733.86	−4.28	−4.27
ARMA(1,1)-X-GARCH(1,1) : XM	11 986.58	−4.37	−4.36
ARMA(1,1)-GARCH(1,1) -X : XV	11 733.86	−4.28	−4.27
ARMA(1,1)-X-GARCH(1,1) -X : XM_XV	11 986.58	−4.37	−4.36

Table 5: In-Sample Performance of the ARMA(1,1)–GARCH(1,1)–Student

4.1.2 ARMA(1,2)-GARCH(1,2)-Student

The estimation of our ARMA(1,2)-GARCH(1,2) variants confirms that the AR(1) coefficient ϕ_1 and both MA terms θ_1 and θ_2 are highly significant in every specification. In the base model, we find ϕ_1 = 0.60, θ_1 = -0.54 and θ_2 = -0.067 for which each p-value≈0, underlining the importance of a richer moving-average component. When the exogenous regressor is added to the mean equation ("XM" and "XM_XV"), the ARMA part remains significant and the exogenous coefficient itself is δ = 0.37 (p≈0.001), demonstrating that it continues to bring substantial explanatory power; the exogenous variable in the variance equation of the "XM_XV" model, however, remains always insignificant (p = 1).

Turning to the Ljung-Box tests at 20 lags, the base ARMA(1,2)-GARCH(1,2) model shows no remaining autocorrelation in its residuals (Q(20) = 26.64, p=0.15) but still strong autocorrelation in the squared residuals ($Q^2(20)$ = 48.25, p = 0.0004), indicating persistent volatility clustering. Once we include the exogenous regressor in the mean equation ("XM"), the statistic on squared residuals falls to ($Q^2(20)$=29.94 p = 0.071), thus satisfying the white-noise criterion. This demonstrates that the combination of the additional MA(2) term and the external regressor is sufficient to "whiten" both the levels and the conditional variance In-Sample. Compared with the ARMA(1,1)-GARCH(1,1) specification, this richer MA structure and extra GARCH coefficient capture heteroscedasticity more effectively.

The GARCH(1,2) volatility specification remains robust: in the base model α_1=0.072, β_1 = 0.66, β_2 = 0.26 (all p = 0), and $\nu \approx 8.22$ under the Student-t law. These values testify to a highly persistent and heavy-tailed volatility process, closely mirroring what we observed for the simpler GARCH(1,1) and GJR-GARCH(1,1) models.

Information criteria are little changed compared to the ARMA(1,1)-GARCH(1,1) case: the base ARMA(1,2) model yields LogLik = -11737.87, AIC = -4.28, BIC = -4.27, whereas the "XM" version improves to LogLik = -11992.45, AIC = -4.37, BIC = -4.36. The "XV" and "XM_XV" variants deliver virtually identical criteria to the base and "XM" models, respectively, since the variance exogenous remains non-significant (p = 1).

Overall, the ARMA(1,2)-GARCH(1,2) with exogenous in the mean ("XM") stands out as our favorite In-Sample specification. It retains all the strengths of the ARMA(1,1)-GARCH(1,1) specification parameter significance, volatility persistence, heavy tails. While the additional MA(2) term and the mean exogenous variable jointly succeed in fully removing autocorrelation from both the residuals and their squares. The "XM_XV" version does not confer any further advantage, as the exogenous regressor in the variance equation remains irrelevant.

Model	**MAE**	**MSE**
ARMA(1,1)-GARCH(1,1) :Base model	0.2710	0.0015
ARMA(1,1)-X-GARCH(1,1) : XM	0.2710	0.0015
ARMA(1,2)-GARCH(1,2) : Base model	0.2640	0.0014
ARMA(1,2)-X-GARCH(1,2) : XM	0.2640	0.0014

Table 6: Out-Sample Performance of the Models

4.1.3 ARMA(1,1)-GJR-GARCH(1,1)-Student

In our four ARMA(1,1)-GJR-GARCH(1,1) specifications (base, XM, XV, and XM_XV), we observe that the asymmetry coefficient γ_1 is consistently very close to zero and not significant (p-value between 0.47 and 0.70). In other words, we detect no leverage effect: positive and negative shocks have the same impact on the volatility of our series.

In the absence of any contribution from the asymmetry term, the variance equation of our GJR-GARCH collapses to the symmetric form: our variance equations and parameter estimates exactly match up with those of a standard GARCH(1,1) model.

Since γ_1 provides no additional adjustment, the GJR extension does not alter the volatility dynamics: our GJR-GARCH models yield the same parameter estimates, information criteria and diagnostics as their GARCH counterparts.

Model	**Log-likelihood**	**AIC**	**BIC**
ARMA(1,1)-GJRGARCH(1,1): Base model	11734.11	–4.28	–4.27
ARMA(1,1)-X-GJRGARCH(1,1): XM	11986.66	–4.37	–4.36
ARMA(1,1)-GJRGARCH(1,1)-X: XV	11734.11	–4.28	–4.27
ARMA(1,1)-X-GJRGARCH(1,1)-X: XM_XV	11986.66	–4.37	–4.36

Table 7: In-Sample Performance of the ARMA(1,1)-GJRGARCH(1,1)-Student

4.1.4 AR(1)-EGARCH(1,1)-Student and MA(1)-EGARCH(1,1)-Student

For the MA(1)-EGARCH(1,1), AR(1)-EGARCH(1,1) and their extensions with exogenous regressors, all mean coefficients (AR and MA) as well as the volatility parameters (ω, β_1 and the asymmetry parameter γ) are significant, and the Student-t distribution remains fully justified ($\nu \approx 8$, p = 0). Adding the exogenous variable in the mean confirms its contribution (δ = 0.367, p = 0), whereas the exogenous variable in the variance is systematically non-significant (p $\approx$ 1.000) in the "XM_XV" versions.

However, when we examine the Ljung-Box tests on the *squared* standardized residuals, *none* of the EGARCH specifications, even those augmented with exogenous regressors, attain the required threshold of insignificance (all $Q^2(20)$ p-values remain below 0.05). In other words, conditional variance remains autocorrelated In-Sample despite the model extensions.

Model	**Log-likelihood**	**AIC**	**BIC**
AR(1)-EGARCH(1,1) :Base Model	11732.72	−4.28	−4.27
AR(1)-X-EGARCH(1,1) :XM	11986.06	−4.37	−4.36
AR(1)-X-EGARCH(1,1) -X :XM_XV	11998.85	−4.38	−4.36
MA(1)-EGARCH(1,1) :Base Model	11733.36	−4.28	−4.27
MA(1)-X-EGARCH(1,1) :XM	11986.24	−4.37	−4.36
MA(1)-X-EGARCH(1,1) -X :XM_XV	11999.02	−4.38	−4.36

Table 8: In-Sample Performance of the AR(1)& MA(1)-EGARCH(1,1)-Student

4.1.5 Conclusion

Across all specifications, we consistently find that incorporating the exogenous regressor in the mean equation (“XM”) substantially improves In-Sample fit and brings the standardized residuals closer to white noise. In particular, the ARMA(1,1)-GARCH(1,1)-Student “XM” model combines significant AR and MA coefficients, a highly persistent GARCH structure ($\alpha_1 + \beta_1 \approx 0.99$), and a meaningful exogenous effect ($\gamma \approx 0.37$, p≈0.01). Extending to ARMA(1,2)-GARCH(1,2) further “whitens” the squared residuals (Ljung-Box $Q^2(20) = 29.94$, p=0.071), thanks to the extra MA(2) term alongside the regressor. Neither GJR-GARCH nor EGARCH variants, however, deliver any additional In-Sample benefit: the asymmetry (only GJR) and variance exogenous terms remain insignificant, and volatility clustering in the squares cannot be eliminated.

These In-Sample findings guide our choice of models for the Out-of-Sample exercise. We will compare following models in order to assess their forecasting performance and robustness:

- ARMA(1,1)-GARCH(1,1)
- ARMA(1,1)-X-GARCH(1,1) (exogenous in mean)
- ARMA(1,2)-GARCH(1,2)
- ARMA(1,2)-X-GARCH(1,2) (exogenous in mean)

This selection reflects the clear In-Sample advantage of mean exogenous extensions and, in the case of ARMA(1,2), the added value of a richer moving-average component in mitigating autocorrelation.

4.2 Out-of-Sample

Out-of-Sample evaluation begins on January 1, 2024, covering the final 1 245 observations of our 5 482-point series. At each step, we re-estimate the models on a fixed rolling window of 4 000 observations and produce a one-step-ahead forecast. This design ensures a constant amount of recent history and spans both high- and low volatility regimes, allowing us to assess the contribution of exogenous variables. A fixed-length window also simulates a true "real-time" environment and maintains comparability across models, while limiting the sensitivity of GARCH parameter estimates to sample size (cf. Ng and Lam, 2006).

With this framework being set, we compare the predictive performance of our four candidate specifications. The models ARMA(1,1)-GARCH(1,1), ARMA(1,1)-X-GARCH(1,1), ARMA(1,2)-GARCH(1,2) and ARMA(1,2)-X-GARCH(1,2) present virtually identical MAE (0.2710 vs. 0.2640) and MSE (0.0015 vs. 0.0014) (see Table 9). Diebold-Mariano tests on the MSE show ARMA(1,1)-GARCH(1,1) vs. ARMA(1,1)-X-GARCH(1,1): t_{DM}=-1.074 (p = 0.2826) and ARMA(1,2)-GARCH(1,2) vs. ARMA(1,2)-X-GARCH(1,2): t_{DM}=-1.227 (p = 0.2197), while Diebold-Mariano tests on the MAE do not reveal any significant differences (all p > 0.18) (see Figure 18). Notably, the only statistically significant difference across all pairwise comparisons is observed between ARMA(1,1)-GARCH(1,1) and ARMA(1,2)-GARCH(1,1), indicating that increasing the ARMA order can improve predictive accuracy. In other words, adding the exogenous input has no significant impact on Out-of-Sample MSE or MAE, and model order of the ARMA component is the primary driver of any detectable performance change.

To sum up, while exogenous variables can have a positive effect on In-Sample benefits by whitening squared residuals, they do not improve, and even slightly worsen, quadratic forecast accuracy Out-of-Sample. Their practical impact on day-to-day point forecasts is therefore limited, and one should be cautious about relying on them to reduce quadratic loss beyond the estimation sample.

Model	**MAE**	**MSE**
ARMA(1,1)-GARCH(1,1) :Base model	0.2710	0.0015
ARMA(1,1)-X-GARCH(1,1) : XM	0.2710	0.0015
ARMA(1,2)-GARCH(1,2) : Base model	0.2640	0.0014
ARMA(1,2)-X-GARCH(1,2) : XM	0.2640	0.0014

Table 9: Out-Sample Performance of the Models

5 Discussion

5.1 Key findings

The aim of this thesis was to highlight the contribution of energy price shocks to the dynamics of ten-year breakeven inflation volatility, by examining more precisely whether the inclusion of daily WTI crude oil price changes as an exogenous variable in GARCH models improved long-term volatility forecasts.

In the In-Sample analysis, adding WTI to the mean equation resulted in an increase in Log-likelihood and a simultaneous decrease in the AIC and BIC, demonstrating a better fit-complexity trade-off, and produced a clear reduction in residual autocorrelation according to the Ljung-Box test. Including WTI in the variance equation, in constrast, had no significant effect on Log-likelihood or on the AIC/BIC, nor did it reduce the autocorrelation of the conditional variance. A comparison between simpler and richer ARMA structures underscored the joint benefit of an additional MA(2) term and the exogenous variable in approaching white noise. Asymmetric extensions (GJR-GARCH and EGARCH) brought no further improvement: neither shock asymmetry (only GJR) nor a variance exogenous term enhanced the model.

Out-of-Sample, it was precisely the ARMA(1,2)-GARCH(1,2) specification, with its extra MA(2) term and second GARCH coefficient β_2, that delivered a slight improvement in forecast quality, while the inclusion of the oil price did not significantly reduce forecasting errors.

In sum, daily WTI shocks clearly improve In-Sample fit by shifting the conditional mean, but they neither modify volatility persistence nor enhance Out-of-Sample forecast accuracy. Our approach, however, remains fairly limited, having tested only two ways of incorporating the exogenous oil price (in the mean and in the variance). Alternative specifications or additional channels of transmission would most probably have yielded different insights. These methodological limitations are discussed in the Limitations section.

5.2 Hypotheses discussion

We now evaluate our empirical findings in light of the three hypotheses established at the beginning.

5.2.1 H1:

Energy price fluctuations significantly influence the volatility of 10-year breakeven inflation rate daily change. Our In-Sample results provide clear support for H1: introducing WTI price changes as an exogenous regressor in the mean equation consistently improved model fit (higher Log-likelihood, lower AIC/BIC) and reduced residual autocorrelation. This effect, however, does not fully translate Out-of-Sample, where the inclusion of WTI in the mean yielded no significant reduction in forecasting errors. This suggests that, while oil shocks clearly shape historical volatility dynamics, their real-time predictive power for daily breakeven inflation changes at a ten-year horizon is more limited (Boeck and Zörner, 2025; Hammoudeh and Reboredo, 2018; Abdallah and Kpodar, 2023; Kilian and Zhou, 2022d).

5.2.2 H2:

Including energy prices (WTI) as exogenous variables simultaneously in both the conditional mean and the conditional variance equations will generate superior In-Sample fit and Out-of-Sample forecasting accuracy. Our evidence does not validate H2. While the "XM" specification (exogenous in the mean) delivered significant In-Sample gains, adding WTI to the variance equation ("XM_XV") neither improved information criteria nor reduced variance autocorrelation. Out-of-Sample, models with dual inclusion did not outperform those with mean-only exogenous terms. Thus, contrary to Diop and Kengne (2021) or Batten et al. (2024), the variance channel of oil shocks appears inactive in our framework (Hamilton, 1989).

5.2.3 H3:

The impact of energy price volatility on daily changes in 10-year breakeven inflation varies across economic regimes, being stronger in high-uncertainty periods. In principle, mean squared error (MSE) is particularly sensitive to large forecast deviations, often arising during crisis episodes, making it a useful proxy for assessing model performance under high-uncertainty regimes (Patton and Timmermann, 2010). In our Out-of-Sample tests, however, we observed no significant differences in MSE between models with or without WTI exogenous terms, even though crises should amplify large errors. Consequently, we cannot claim support for H3: the absence of MSE improvements suggests that our models do not capture a stronger oil-price effect in turbulent periods. A dedicated regime-switching or time-varying framework would be required to further explore this hypothesis. This limitation will be discussed below.

5.3 Limitations

Several limitations should be acknowledged to contextualize our conclusions and guide future research.

First, our analysis was constrained by the availability of freely accessible databases, limiting us to ten-year breakeven inflation expectations. However, existing literature indicates that energy price shocks have a more pronounced effect on shorter horizons (one- and five-year expectations) because they more immediately affect production costs and household consumption (Kilian and Zhou, 2022d). We therefore attempted to extend our GARCH specifications to these shorter maturities using monthly data (approximately 500 observations). This allowed us to include a wider range of energy-related exogenous regressors, such as natural gas prices, composite energy indices, and oil production measures, available only at a monthly frequency, alongside daily WTI oil prices. Nonetheless, these estimations remained statistically weak, demonstrating that the dataset's frequency and size, rather than the GARCH framework itself, limited our ability to validate horizon-dependent effects.

Second, our non-exogenous GARCH benchmarks failed to fully "whiten" the squared residuals, leaving significant autocorrelation that indicates persistent volatility clustering (Engle, 1982). As a result, Out-of-Sample comparisons rely on reference models that do not satisfy the white-noise condition, complicating the attribution of performance differences solely to the inclusion of exogenous energy variables. We also attempted to estimate FIGARCH models to capture long-memory effects, but standard optimization routines in Python (arch) and R (rugarch) returned NaN parameter estimates, preventing convergence.

Finally, the linear incorporation of WTI Log-returns into the conditional variance equation may oversimplify the transmission of energy shocks. Alternative specifications, such as Markov regime-switching models, may more effectively capture nonlinear and state-dependent dynamics.

Future studies that addresses these data and model limitations may offer deeper insights into the influence of energy price shocks on inflation expectations across different time horizons.

Conclusion

This dissertation aimed at answering the central question: "To what extent do daily WTI oil price shocks influence the daily volatility of ten-year breakeven inflation expectations ?" To address this, we employed a range of GARCH-type models (both symmetric and asymmetric), comparing specifications with and without the inclusion of daily WTI returns in the conditional mean and variance equations, and evaluating performance both In-Sample and Out-of-Sample using information criteria, autocorrelation tests, and forecast metrics (MSE, MAE, Diebold-Mariano test).

The results show that adding the oil shock to the mean equation markedly improves In-Sample fit (higher Log-likelihood, lower AIC/BIC, reduced residual autocorrelation), whereas its insertion into the variance equation yields no significant effect. Out-of-Sample, no WTI-augmented model consistently outperforms the ARMA(1,2)-GARCH(1,2) benchmark: forecast errors and Diebold-Mariano statistics remain comparable, underscoring the limited predictive power of oil shocks for long-term volatility.

In conclusion, this study offers a nuanced perspective on the transmission channels from energy markets to long-term inflation: it confirms the historical impact of oil shocks on the level of expectations while demonstrating their weak real-time predictive role for future volatility. These findings pave the way for promising avenues of further research: testing nonlinear or regime-switching specifications (e.g., Markov-switching, FIGARCH), examining shorter maturities (one- and five-year breakevens), and incorporating composite energy indices (natural gas, broader energy mix) to better capture the complex interaction between energy markets and financial stability.

Bibliography

Aastveit, K.A., Bjørnland, H.C., Cross, J.L., 2023. Inflation expectations and the pass-through of oil prices. Review of Economics and Statistics 105, 733–743. doi:10.1162/rest_a_01073.

Abdallah, C., Kpodar, K., 2023. How large and persistent is the response of inflation to changes in retail energy prices? Journal of International Money and Finance 132, 102806. doi:10.1016/j.jimonfin.2023.102806.

Adrian, T., 2023. The role of inflation expectations in monetary policy. https://www.imf.org/en/News/Articles/2023/05/15/sp-role-inflation-expectations-monetary-policy-tobias-adrian.

Adrian, T., Boyarchenko, N., Giannone, D., 2019. Vulnerable growth. American Economic Review 109, 1263–1289. doi:10.1257/aer.20161923.

Aidala, F., Armantier, O., Koşar, G., Somerville, A., Topa, G., van der Klaauw, W., 2024. Gasoline price changes and consumer inflation expectations: Experimental evidence. Journal of Economic Behavior & Organization 220, 66–80. doi:10.1016/j.jebo.2024.01.027.

Akaike, H., 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control 19, 716–723. doi:10.1109/TAC.1974.1100705.

Alcidi, C., Gros, D., Shamsfakhr, F., 2022. Inflation Expectations: Models and Measures. Technical Report. Policy Department for Economic, Scientific and Quality of Life Policies, European Parliament. https://www.europarl.europa.eu/RegData/etudes/IDAN/2022/703338/IPOL_IDA(2022)703338_EN.pdf.

Aldama, P., Le Bihan, H., Le Gall, C., 2024. What Caused the Post-Pandemic Inflation? Replicating Bernanke and Blanchard (2023) on French Data. Technical Report Working Paper No. 967. Banque de France. https://www.banque-france.fr/en/publications-and-statistics/publications.

Alvarez, J., Bluedorn, J., Hansen, N.J., Huang, Y., Pugacheva, E., Sollaci, A., 2022. Wage-Price Spirals: What is the Historical Evidence? Technical Report WP/22/221. International Monetary Fund. URL: https://www.imf.org/en/Publications/WP/Issues/2022/11/11/Wage-Price-Spirals-What-is-the-Historical-Evidence-525073.

Andrade, P., Gautier, E., Mengus, E., 2020. What Matters in Households' Inflation Expectations? Technical Report. Federal Reserve Bank of Cleveland. URL: https://www.clevelandfed.org/-/media/project/clevelandfedtenant/clevelandfedsite/events/2020/inflations-drivers-and-dynamics/philippe-andrade-paper.pdf. preliminary working paper; using survey data on French households.

Armstrong, J.S., Collopy, F., 1992. Error measures for generalizing about forecasting methods: Empirical comparisons. International Journal of Forecasting 8, 69–80.

Barsky, R.B., Kilian, L., 2004. Oil and the macroeconomy since the 1970s. Journal of Economic Perspectives 18, 115–134. doi:10.1257/0895330042632708.

Batten, J.A., Mo, D., Pourkhanali, A., 2024. Can inflation predict energy price volatility? Energy Economics 129, 107158. doi:10.1016/j.eneco.2023.107158.

Baumeister, C., 2022. Pandemic, War, Inflation: Oil Markets at a Crossroads? Working Paper No. 31496. National Bureau of Economic Research. doi:10.3386/w31496.

Bernanke, B.S., 2004. Oil and the economy. Remarks at the Distinguished Lecture Series, Darton College, Albany, Georgia. URL: https://www.federalreserve.gov/boarddocs/speeches/2004/20041021/default.htm.

Bernanke, B.S., Gertler, M., Watson, M.W., 1997. Systematic monetary policy and the effects of oil price shocks. Brookings Papers on Economic Activity , 91–157.

Binder, C.C., 2018. Inflation expectations and the price at the pump. Journal of Macroeconomics 58, 1–18. doi:10.1016/j.jmacro.2018.08.006.

Bjørnland, H.C., Larsen, V.H., Maih, J., 2018. Oil and macroeconomic (in)stability. American Economic Journal: Macroeconomics 10, 128–151. doi:10.1257/mac.20150171.

Blanchard, O., 2016. The phillips curve: Back to the '60s? American Economic Review: Papers & Proceedings 106, 31–34. doi:10.1257/aer.p20161003.

Blanchard, O.J., 1986. The wage price spiral. Quarterly Journal of Economics 101, 543–565. URL: https://www.jstor.org/stable/1885696, doi:10.2307/1885696.

Blanchard, O.J., Gali, J., 2007. The macroeconomic effects of oil shocks: Why are the 2000s so different from the 1970s? Technical Report Working Paper No. 13368. National Bureau of Economic Research. URL: http://www.nber.org/papers/w13368.

Boeck, M., Zörner, T.O., 2025. Natural gas prices, inflation expectations, and the pass-through to euro area inflation. Energy Economics 141, 108061. doi:10.1016/j.eneco.2024.108061.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307–327. doi:10.1016/0304-4076(86)90063-1.

Bollerslev, T., 1987. A conditionally heteroskedastic time series model for speculative prices and rates of return. The Review of Economics and Statistics 69, 542–547. URL: https://www.jstor.org/stable/1925546.

Bonnier, J., 2022. Forecasting crude oil volatility with exogenous predictors: As good as it gets? Energy Economics 111, 106059. doi:10.1016/j.eneco.2022.106059.

Box, G.E.P., Jenkins, G.M., 1970. Time Series Analysis: Forecasting and Control. Holden–Day, San Francisco.

Box, George E. P., Jenkins, G.M., Reinsel, G.C., 2008. Time Series Analysis: Forecasting and Control. 4 ed., Wiley, Hoboken, NJ.

Byun, S.J., Cho, H., 2013. Forecasting carbon futures volatility using garch models with energy volatilities. Energy Economics 40, 207–221. doi:10.1016/j.eneco.2013.06.017.

Campbell, J.Y., Lo, A.W., MacKinlay, A.C., 1997. The Econometrics of Financial Markets. Princeton University Press, Princeton, NJ.

Chudik, A., Georgiadis, G., 2022. Estimation of impulse response functions when shocks are observed at a higher frequency than outcome variables. Journal of Business & Economic Statistics 40, 965–979. doi:10.1080/07350015.2021.1889567.

Coibion, O., Gorodnichenko, Y., 2015. Is the phillips curve alive and well after all? inflation expectations and the missing disinflation. American Economic Journal: Macroeconomics 7, 197–232. doi:10.1257/mac.20130306.

Dickey, D.A., Fuller, W.A., 1979. Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association 74, 427–431. doi:10.2307/2286348.

Diebold, F.X., Mariano, R.S., 1995. Comparing predictive accuracy. Journal of Business & Economic Statistics 13, 253–263.

Ding, Z., Granger, C.W.J., Engle, R.F., 1993. A long memory property of stock market returns and a new model. Journal of Empirical Finance 1, 83–106. doi:10.1016/0927-5398(93)90006-D.

Diop, M.L., Kengne, W., 2021. Inference and model selection in general causal time series with exogenous covariates. ArXiv preprint.

Duca-Radu, I., Kenny, G., Reuter, A., 2021. Inflation expectations, consumption and the lower bound: Micro evidence from a large multi-country survey. Journal of Monetary Economics 118, 120–134. doi:10.1016/j.jmoneco.2020.03.005.

D'Acunto, F., Malmendier, U., Ospina, J., Weber, M., 2021a. Exposure to grocery prices and inflation expectations. Journal of Political Economy 129, 1615–1639. doi:10.1086/713192.

D'Acunto, F., Malmendier, U., Weber, M., 2021b. Gender roles produce divergent economic expectations. Proceedings of the National Academy of Sciences 118, e2008534118. doi:10.1073/pnas.2008534118.

Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. Econometrica 50, 987–1007. URL: https://www.jstor.org/stable/1912773.

Engle, R.F., 2001. An Introduction to the Use of ARCH/GARCH Models in Applied Econometrics. Technical Report. NYU Stern.

Engle, R.F., Patton, A.J., 2001. What good is a volatility model? Quantitative Finance 1, 237–245. doi:10.1016/B978-075066942-9.50004-2.

Federal Reserve Bank of St. Louis, 2025a. 10–year breakeven inflation rate (t10yie). URL: https://fred.stlouisfed.org/series/T10YIE.

Federal Reserve Bank of St. Louis, 2025b. Daily crude oil prices: West texas intermediate (dcoilwtico). URL: https://fred.stlouisfed.org/series/DCOILWTICO.

Fisher, I., 1911. The Purchasing Power of Money: Its Determination and Relation to Credit, Interest and Crises. The Macmillan Company, New York, NY.

Furno, F., Giannone, D., 2024. Nowcasting recession risk. SSRN Electronic Journal doi:10.2139/ssrn.4706700.

Gauss, C.F., 1823. Theoria Combinationis Observationum Erroribus Minimis Obnoxiae. F. Perthes und I. H. Besser, Göttingen.

Glosten, L.R., Jagannathan, R., Runkle, D.E., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. The Journal of Finance 48, 1779–1801. doi:10.1111/j.1540-6261.1993.tb05128.x.

Gordon, R.J., 1975. Alternative responses of policy to external supply shocks. Brookings Papers on Economic Activity 6, 183–206.

Gorodnichenko, Y., Coibion, O., Kumar, S., 2018. How do firms form their expectations? new survey evidence. American Economic Review 108, 2671–2713. doi:10.1257/aer.20151299.

Hamilton, J.D., 1989. A new approach to the economic analysis of nonstationary time series and the business cycle 57, 357–384. doi:10.2307/1912559.

Hamilton, J.D., 1994. Time Series Analysis. Princeton University Press, Princeton, NJ.

Hamilton, J.D., 2009. Causes and consequences of the oil shock of 2007–08. URL: https://www.nber.org/papers/w15002.

Hammoudeh, S., Reboredo, J.C., 2018. Oil price dynamics and market-based inflation expectations. Energy Economics 75, 484–491. doi:10.1016/j.eneco.2018.09.011.

Harris, E.S., Kasman, B.C., Shapiro, M.D., West, K.D., 2009. Oil and the Macroeconomy: Lessons for Monetary Policy. Conference Report. U.S. Monetary Policy Forum, Initiative on Global Markets, University of Chicago Booth School of Business; Rosenberg Institute of Global Finance, Brandeis International Business School. URL: https://public.websites.umich.edu/~shapiro/papers/USMPF-2009.pdf. revised November 2009.

Hasenzagl, T., Koester, G., Malmendier, U., 2022. Inflation expectations and the financial cycle: Evidence from the euro area. Journal of Monetary Economics 127, 1–17. doi:10.1016/j.jmoneco.2021.11.005.

Hyung, N., Granger, C.W.J., 2008. Do asymmetric garch models help in forecasting volatility? Journal of Forecasting 27, 179–192. doi:10.1002/for.1034.

on Inflation Expectations, E.W.S., 2021. Inflation expectations and their role in Eurosystem forecasting. Technical Report 264. European Central Bank. URL: https://ssrn.com/abstract=3928285.

Kilian, L., 2009. Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. American Economic Review 99, 1053–1069. doi:10.1257/aer.99.3.1053.

Kilian, L., 2010. Oil price shocks, monetary policy and stagflation, in: Fry, R., Jones, C., Kent, C.T. (Eds.), Inflation in an Era of Relative Price Shocks. Reserve Bank of Australia, Sydney, pp. 60–84.

Kilian, L., Murphy, D.P., 2014. The role of inventories and speculative trading in the global market for crude oil. Journal of Applied Econometrics 29, 454–478. doi:10.1002/jae.2322.

Kilian, L., Zhou, X., 2020. Oil Prices, Gasoline Prices and Inflation Expectations: A New Model and New Facts. Technical Report Research Department Working Paper No. 2025. Federal Reserve Bank of Dallas. doi:10.24149/wp2025.

Kilian, L., Zhou, X., 2022a. A Broader Perspective on the Inflationary Effects of Energy Price Shocks. Technical Report Research Department Working Paper No. 2224. Federal Reserve Bank of Dallas. doi:10.24149/wp2224.

Kilian, L., Zhou, X., 2022b. Disentangling demand and supply drivers of oil prices: A frequency-domain analysis. Journal of Applied Econometrics 37, 605–623. doi:10.1002/jae.2912.

Kilian, L., Zhou, X., 2022c. The impact of rising oil prices on u.s. inflation and inflation expectations in 2020–23. Energy Economics 113, 106228. doi:10.1016/j.eneco.2022.106228.

Kilian, L., Zhou, X., 2022d. Oil prices, gasoline prices, and inflation expectations. Journal of Applied Econometrics 37, 867–881. doi:10.1002/jae.2911.

Kilian, L., Zhou, X., 2023. Oil Price Shocks and Inflation. Technical Report Working Paper No. 2312. Federal Reserve Bank of Dallas. doi:10.24149/wp2312.

Kose, M.A., Matsuoka, H., Panizza, U., Vorisek, D., 2019. Inflation Expectations: Review and Evidence. Policy Research Working Paper 8785. World Bank, Macroeconomics, Trade and Investment Global Practice. URL: https://documents1.worldbank.org/curated/en/723151553110873395/pdf/WPS8785.pdf.

Kozicki, S., Tinsley, P.A., 2005. What do you expect? imperfect policy credibility and tests of the expectations hypothesis. Journal of Monetary Economics 52, 421–447. doi:10.1016/j.jmoneco.2004.05.004.

Kur, A.P., Ngesa, O., Sarguta, R., 2021. Exponential garch model with exogenous covariate for ssp-usd exchange rate volatility: On the effects of conflict on volatility. Journal of Mathematical Finance 11.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., Shin, Y., 1992. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? Journal of Econometrics 54, 159–178. doi:10.1016/0304-4076(92)90104-Y.

Lagarde, C., 2024. Remarks at the ecb forum on inflation expectations. URL: https://www.ecb.europa.eu/press/key/date/2024/html/ecb.sp240920~7edac96692.fr.html. european Central Bank.

Lejsgaard Autrup, S., Grothe, M., 2014. Economic Surprises and Inflation Expectations: Has Anchoring of Expectations Survived the Crisis? Working Paper 1671. European Central Bank. URL: https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1671.pdf.

Liu, J., Ma, F., Tang, Y., Zhang, Y., 2019. Geopolitical risk and oil volatility: A new insight. Energy Economics 84, 104548. doi:10.1016/j.eneco.2019.104548.

Liu, Y., Sharma, P., Jain, V., Shukla, A., Shabbir, M.S., Tabash, M.I., Chawla, C., 2022. The relationship among oil prices volatility, inflation rate, and sustainable economic growth: Evidence from top oil importer and exporter countries. Resources Policy 77, 102674. doi:10.1016/j.resourpol.2022.102674.

Ljung, G.M., Box, G.E.P., 1978. On a measure of a lack of fit in time series models. Biometrika 65, 297–303. doi:10.1093/biomet/65.2.297.

Lorenzoni, G., Werning, I., 2023. Inflation is Conflict. Technical Report Working Paper No. 31099. National Bureau of Economic Research. URL: https://www.nber.org/papers/w31099.

Nelson, D.B., 1991. Conditional heteroskedasticity in asset returns: A new approach. Econometrica 59, 347–370. URL: https://www.jstor.org/stable/2938260.

Neri, S., Busetti, F., Conflicts, C., Corsello, F., Delle Monache, D., Tagliabracci, A., 2023. Energy Price Shocks and Inflation in the Euro Area. Technical Report 792. Bank of Italy. URL: https://www.bancaditalia.it/pubblicazioni/qef/2023-0792/QEF_792_23.pdf.

Ng, H. S.R., Lam, K. P., 2006. How does sample size affect garch models?, in: Proceedings of the 9th Joint International Conference on Information Sciences (JCIS-06), Atlantis Press. URL: https://doi.org/10.2991/jcis.2006.139, doi:10.2991/jcis.2006.139.

Patton, A.J., Timmermann, A., 2010. Why do forecasters disagree? lessons from the term structure of cross-sectional dispersion. Journal of Monetary Economics 57, 803–820. doi:10.1016/j.jmoneco.2010.07.001.

Peersman, G., Van Robays, I., 2009. Oil and the euro area economy. Economic Policy 24, 603–651. doi:10.1111/j.1468-0327.2009.00233.x.

of Philadelphia, F.R.B., 2024. Survey of professional forecasters: Fourth quarter 2024. URL: https://www.philadelphiafed.org/-/media/FRBP/Assets/Surveys-And-Data/survey-of-professional-forecasters/spf-documentation.pdf?sc_lang=en&hash=8408A4F1BF351A3C268B40F6BC7B95AA.

Pincheira, P.M., 2013. Interventions and Inflation Expectations in an Inflation Targeting Economy. Working Paper 427. Bank for International Settlements. URL: https://www.bis.org/publ/work427.pdf.

Said, S.E., Dickey, D.A., 1984. Testing for unit roots in autoregressive-moving average models of unknown order. Biometrika 71, 599–607. doi:10.1093/biomet/71.3.599.

Schwarz, G., 1978. Estimating the dimension of a model. The Annals of Statistics 6, 461–464. doi:10.1214/aos/1176344136.

Slutsky, E., 1937. The summation of random causes as the source of cyclic processes. Econometrica 5, 105–146.

Tsay, R.S., 2010. Analysis of Financial Time Series. 3rd ed., Wiley.

Vatsa, P., Pino, G., 2024. Do petrol prices affect inflation and inflation expectations? evidence from new zealand. Energy Economics doi:10.1016/j.eneco.2024.107939.

Wilks, S.S., 1938. The large-sample distribution of the likelihood ratio for testing composite hypotheses. Annals of Mathematical Statistics 9, 60–62. doi:10.1214/aoms/1177732360.

Wong, B., 2015. Do inflation expectations propagate the inflationary impact of real oil price shocks? evidence from the michigan survey. Journal of Money, Credit and Banking 47, 1673–1689. doi:10.1111/jmcb.12288.

Yule, G.U., 1927. On a method of investigating periodicities in disturbed series, with special reference to wolfer's sunspot numbers. Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character 226, 267–298.

Appendix

Glossary

Term	Definition
T10YIE	10-year Breakeven inflation rate.
WTI	West Texas Intermediate crude oil price.
AR	Autoregressive.
MA	Mean Average.
ARMA	Autoregressive - Mean Average.
ARCH	Autoregressive Conditional Heteroskedasticity.
ARCH	Generalised Autoregressive Conditional Heteroskedasticity.
MAE	Mean Absolute Error.
MSE	Mean Squared Error.
LogLik	Log-likelihood
AIC	Akaike Information Criterion.
BIC	Bayesian Information Criterion.
KPSS	Kwiatkowski–Phillips–Schmidt–Shin stationarity test.
ADF	Augmented Dickey–Fuller unit-root test.
DM	Diebold Mariano.
Base	GARCH models without exogenous variable.
XM	GARCH models with exogenous variable in conditional mean.
XV	GARCH models with exogenous variable in conditional variance.
XM_XM	GARCH models with exogenous variable in conditional mean and variance.

Data

Explanation of Logarithmic returns of WTI

By taking logarithmic returns rather than price levels, we eliminate long-term trends and generate a series that fluctuates around zero, which is necessary for reliable GARCH estimation. Logarithmic returns also have a clear economic interpretation. This transformation makes the series stationary, eliminating long-term trends in price levels. What's more, thanks to the additivity property of logarithmic returns, volatility over a multi-period horizon is simply calculated as the sum of elementary returns. Finally, the logarithm dampens extreme price movements, making our volatility models more robust and more in line with current financial practice (Campbell et al., 1997). This approach is now a standard in financial econometrics in the GARCH literature (Bollerslev, 1986; Engle and Patton, 2001).

Data - Additional Tests

Variable	Min	Mean	Stdev	Max	Skew	Kurt
2003–2006						
WTI	25.25	48.81	14.49	77.05	0.10	1.67
T10YIE	1.57	2.34	0.28	2.76	−1.16	3.68
log(WTI)	−0.15	0.00	0.02	0.12	−0.48	6.93
Daily Change T10YIE	−0.18	0.00	0.03	0.10	−0.24	5.26
2007–2009						
WTI	30.28	78.02	25.41	145.31	0.64	2.76
T10YIE	0.04	1.95	0.60	2.57	−1.43	4.16
LOG WTI	−0.13	0.00	0.03	0.16	0.05	6.53
Daily Change T10YIE	−0.36	0.00	0.05	0.33	−0.30	11.19
2010–2019						
WTI	26.19	72.49	21.95	113.39	0.03	1.63
T10YIE	1.18	1.99	0.30	2.64	−0.12	2.29
LOG WTI	−0.11	0.00	0.02	0.14	0.13	6.57
Daily Change T10YIE	−0.17	0.00	0.03	0.17	−0.17	5.71
2020–2024						
WTI	−36.98	71.24	20.49	123.64	−0.52	3.80
T10YIE	0.50	2.18	0.41	3.02	−1.33	4.82
LOG WTI	−0.28	0.00	0.04	0.43	1.17	36.70
Daily Change T10YIE	−0.32	0.00	0.04	0.25	−0.57	10.41
Full sample (2003–2024)						
WTI	−36.98	68.68	23.07	145.31	0.23	2.56
T10YIE	0.04	2.09	0.40	3.02	−1.23	5.59
LOG WTI	−0.28	0.00	0.03	0.43	0.67	31.89
Daily Change T10YIE	−0.36	0.00	0.03	0.33	−0.37	12.61

Table 10: Descriptive statistics for WTI, T10YIE, log WTI, and daily changes in T10YIE, by sub-period and for the full sample

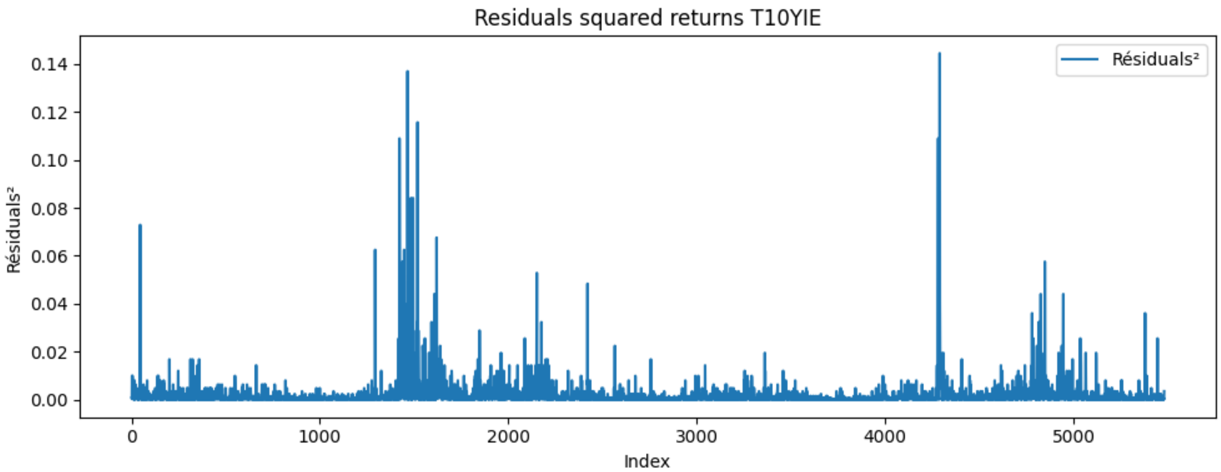

Figure 2: Daily Squared Changes of breakeven inflation rate over the full sample

In Sample Results

ARMA(1,1)-GARCH(1,1)

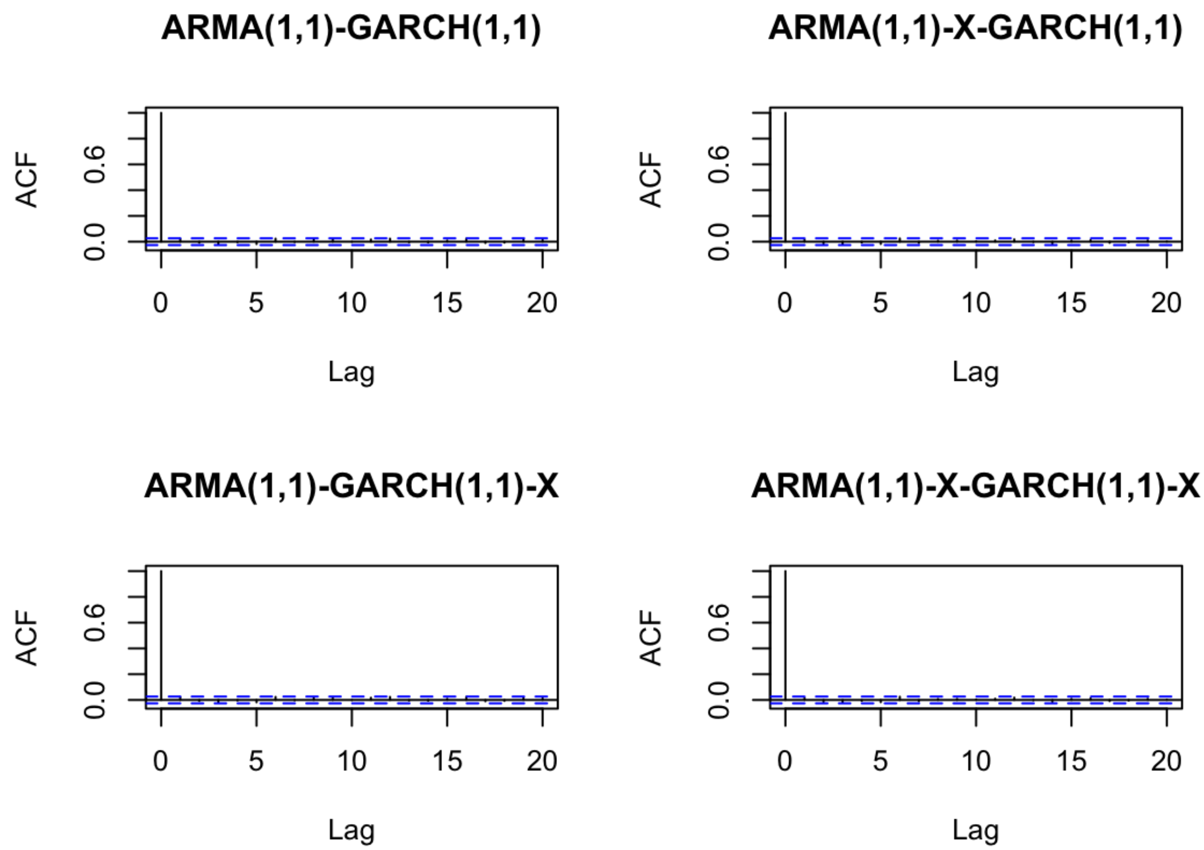

Figure 3: ARMA(1,1)-GARCH(1,1) ACF

```
--- base ---
       Estimate StdError  z.value p.value
mu       0.0003   0.0003   0.9421  0.3461
ar1     -0.2840   0.1429  -1.9884  0.0468
ma1      0.3442   0.1396   2.4651  0.0137
omega    0.0000   0.0000   2.7784  0.0055
alpha1   0.0583   0.0077   7.5379  0.0000
beta1    0.9349   0.0083 112.8266  0.0000
shape    8.1459   0.7701  10.5779  0.0000
LogLik = 11733.86 | AIC = -4.28 | BIC = -4.27
base: Ljung-Box Q(20) sur résidus      = 25.75, p = 0.1742
base: Ljung-Box Q²(20) sur résidus^2 = 51.17, p = 0.0002

--- xm ---
       Estimate StdError  z.value p.value
mu       0.0000   0.0003   0.1434  0.8859
ar1     -0.3871   0.1705  -2.2699  0.0232
ma1      0.4246   0.1671   2.5411  0.0111
mxreg1   0.3661   0.0161  22.7617  0.0000
omega    0.0000   0.0000   3.2378  0.0012
alpha1   0.0665   0.0080   8.3394  0.0000
beta1    0.9241   0.0072 128.5312  0.0000
shape    8.0762   0.7758  10.4100  0.0000
LogLik = 11986.58 | AIC = -4.37 | BIC = -4.36
xm: Ljung-Box Q(20) sur résidus      = 17.54, p = 0.6179
xm: Ljung-Box Q²(20) sur résidus^2 = 33.18, p = 0.0322
```

Figure 4: ARMA(1,1)-GARCH(1,1) Specifications: (base) No Exogenous Variable; (xm) Exogenous in Conditional Mean

```
--- xv ---
       Estimate StdError  z.value p.value
mu       0.0003   0.0003   0.9388  0.3478
ar1     -0.2841   0.1429  -1.9886  0.0467
ma1      0.3442   0.1396   2.4654  0.0137
omega    0.0000   0.0000   2.7752  0.0055
alpha1   0.0583   0.0081   7.1510  0.0000
beta1    0.9349   0.0086 108.6343  0.0000
vxreg1   0.0000   0.0002   0.0000  1.0000
shape    8.1433   0.7762  10.4908  0.0000
LogLik = 11733.86 | AIC = -4.28 | BIC = -4.27
xv: Ljung-Box Q(20) sur résidus      = 25.75, p = 0.1742
xv: Ljung-Box Q²(20) sur résidus^2 = 51.17, p = 0.0002

--- xm_xv ---
       Estimate StdError  z.value p.value
mu       0.0000   0.0003   0.1430  0.8863
ar1     -0.3885   0.1702  -2.2827  0.0224
ma1      0.4260   0.1667   2.5552  0.0106
mxreg1   0.3661   0.0161  22.7611  0.0000
omega    0.0000   0.0000   3.2264  0.0013
alpha1   0.0665   0.0082   8.1217  0.0000
beta1    0.9240   0.0072 127.5391  0.0000
vxreg1   0.0000   0.0002   0.0000  1.0000
shape    8.0619   0.7823  10.3054  0.0000
LogLik = 11986.58 | AIC = -4.37 | BIC = -4.36
xm_xv: Ljung-Box Q(20) sur résidus      = 17.54, p = 0.6177
xm_xv: Ljung-Box Q²(20) sur résidus^2 = 33.13, p = 0.0327
```

Figure 5: ARMA(1,1)-GARCH(1,1) Specifications: (xv) Exogenous in Conditional Variance; (xm_xv) Exogenous in Mean and Variance

ARMA(1,2)-GARCH(1,2)

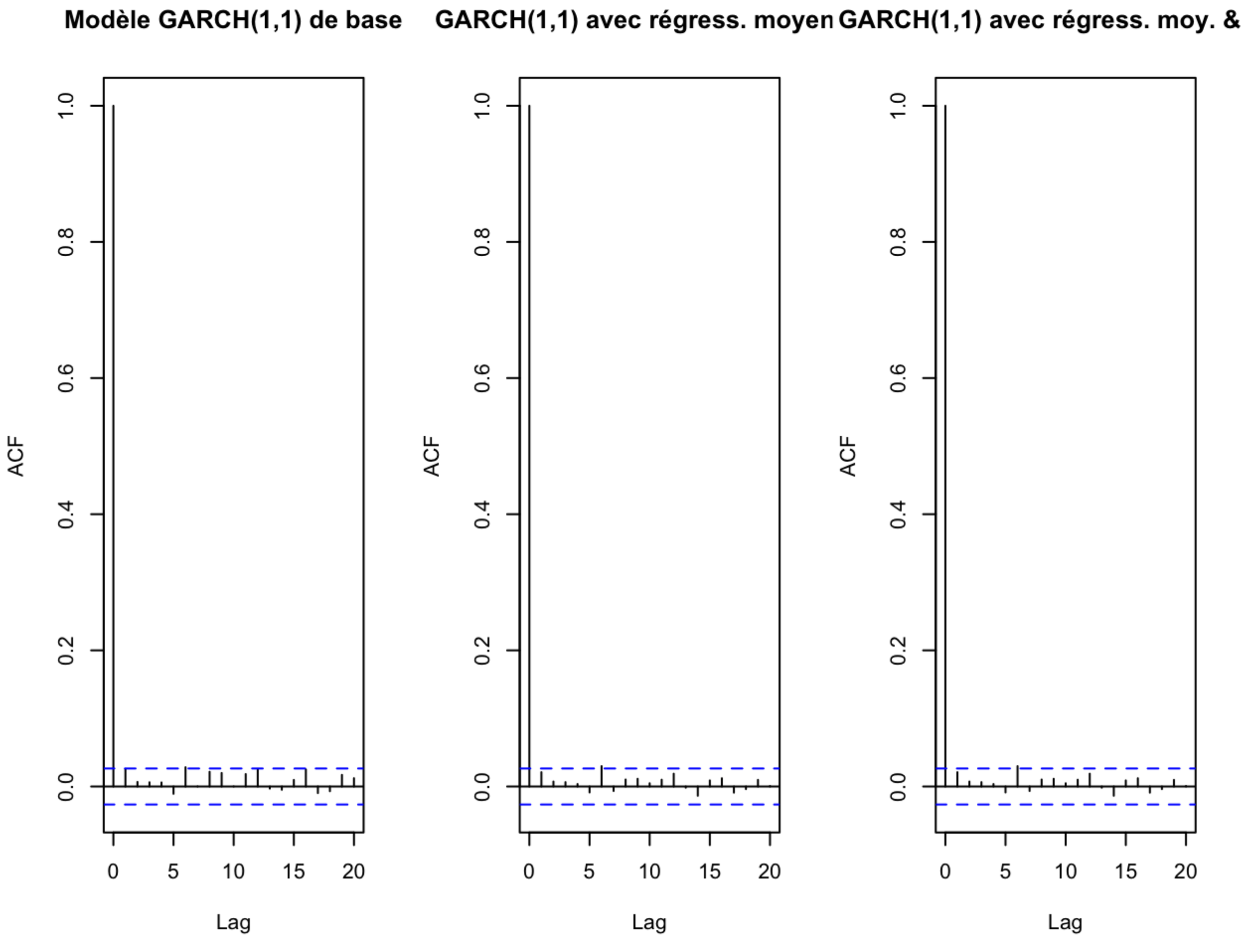

Figure 6: ARMA(1,2)-GARCH(1,2) ACF

```
--- base ---
       Estimate StdError z.value p.value
mu       0.0003   0.0003  1.0293  0.3033
ar1      0.5981   0.1632  3.6656  0.0002
ma1     -0.5407   0.1631 -3.3143  0.0009
ma2     -0.0673   0.0143 -4.7014  0.0000
omega    0.0000   0.0000  3.1348  0.0017
alpha1   0.0722   0.0086  8.4319  0.0000
beta1    0.6608   0.0101 65.5722  0.0000
beta2    0.2586   0.0098 26.4731  0.0000
shape    8.2169   0.7840 10.4802  0.0000
LogLik = 11737.87 | AIC = -4.28 | BIC = -4.27
base: Ljung-Box Q(20) sur résidus      = 26.64, p = 0.1456
base: Ljung-Box Q²(20) sur résidus^2 = 48.25, p = 0.0004

--- xm ---
       Estimate StdError z.value p.value
mu       0.0001   0.0003  0.1927  0.8472
ar1      0.5932   0.1732  3.4240  0.0006
ma1     -0.5620   0.1733 -3.2433  0.0012
ma2     -0.0572   0.0139 -4.1245  0.0000
mxreg1   0.3671   0.0161 22.8644  0.0000
omega    0.0000   0.0000  3.1169  0.0018
alpha1   0.0836   0.0114  7.3179  0.0000
beta1    0.6137   0.0575 10.6714  0.0000
beta2    0.2910   0.0555  5.2387  0.0000
shape    8.1432   0.7995 10.1859  0.0000
LogLik = 11992.45 | AIC = -4.37 | BIC = -4.36
xm: Ljung-Box Q(20) sur résidus      = 15.81, p = 0.7285
xm: Ljung-Box Q²(20) sur résidus^2 = 29.94, p = 0.0708
```

Figure 7: ARMA(1,2)-GARCH(1,2) Specifications: (base) No Exogenous Variable; (xm) Exogenous in Conditional Mean

```
--- xm_xv ---
        Estimate StdError z.value p.value
mu        0.0001   0.0003  0.1926  0.8473
ar1       0.5933   0.1732  3.4249  0.0006
ma1      -0.5621   0.1733 -3.2443  0.0012
ma2      -0.0572   0.0139 -4.1242  0.0000
mxreg1    0.3671   0.0161 22.8637  0.0000
omega     0.0000   0.0000  3.0541  0.0023
alpha1    0.0836   0.0119  7.0472  0.0000
beta1     0.6138   0.0579 10.6042  0.0000
beta2     0.2909   0.0558  5.2135  0.0000
vxreg1    0.0000   0.0003  0.0000  1.0000
shape     8.1443   0.8159  9.9816  0.0000
LogLik = 11992.45 | AIC = -4.37 | BIC = -4.36
xm_xv: Ljung-Box Q(20) sur résidus      = 15.81, p = 0.7285
xm_xv: Ljung-Box Q²(20) sur résidus^2 = 29.94, p = 0.0708
```

Figure 8: ARMA(1,2)-GARCH(1,2) Specifications: (xv) Exogenous in Conditional Variance; (xm_xv) Exogenous in Mean and Variance

ARMA(1,1)-GJR-GARCH(1-1)

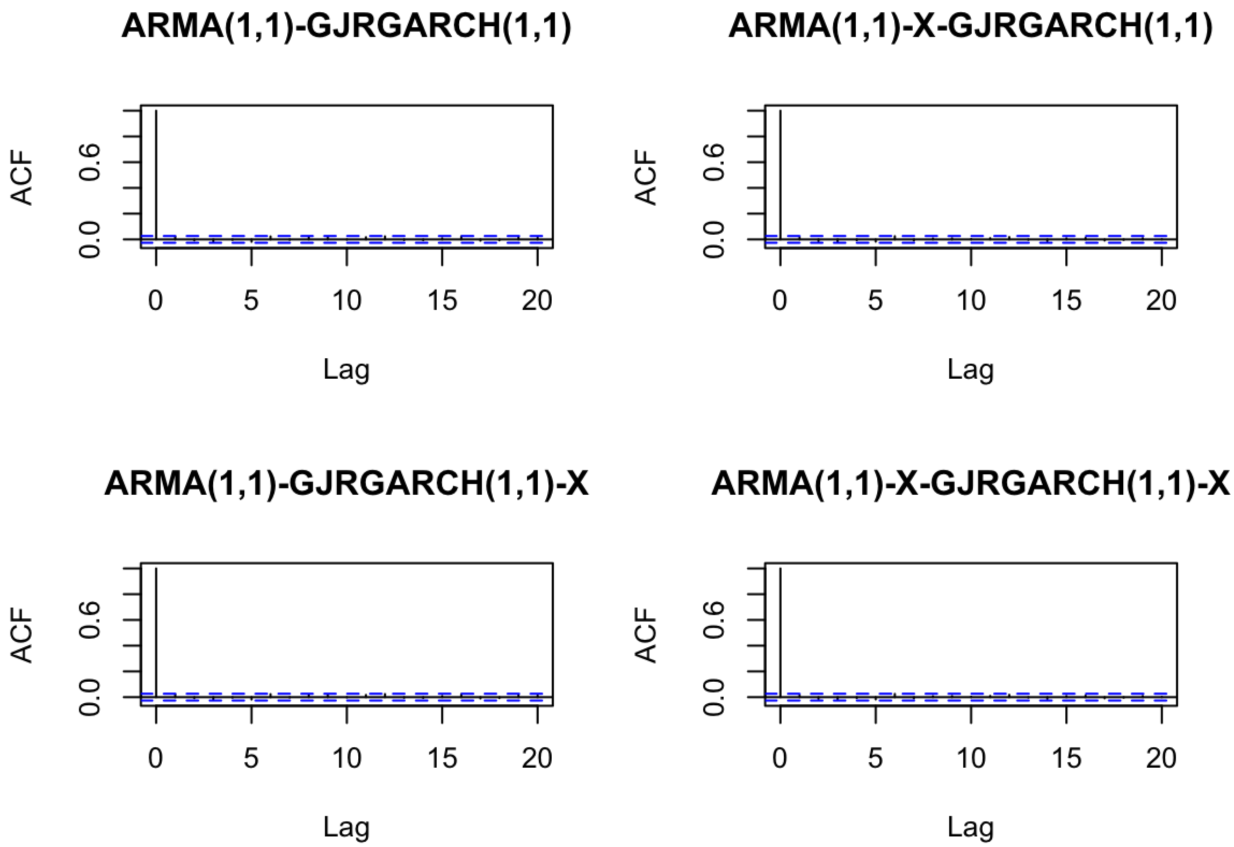

Figure 9: ARMA(1,1)-GJR-GARCH(1,1) ACF

```
--- base ---
       Estimate StdError  z.value p.value
mu       0.0003   0.0004   0.8339  0.4043
ar1     -0.2851   0.1426  -1.9992  0.0456
ma1      0.3453   0.1394   2.4777  0.0132
omega    0.0000   0.0000   2.7074  0.0068
alpha1   0.0542   0.0094   5.7629  0.0000
beta1    0.9357   0.0085 109.7912  0.0000
gamma1   0.0066   0.0091   0.7223  0.4701
shape    8.1765   0.7764  10.5319  0.0000
LogLik = 11734.11 | AIC = -4.28 | BIC = -4.27
base: Ljung-Box Q(20) sur résidus      = 25.31, p = 0.1897
base: Ljung-Box Q²(20) sur résidus^2 = 48.68, p = 0.0003

--- xm ---
       Estimate StdError  z.value p.value
mu       0.0000   0.0003   0.0730  0.9418
ar1     -0.3874   0.1702  -2.2754  0.0229
ma1      0.4249   0.1668   2.5473  0.0109
mxreg1   0.3660   0.0161  22.7601  0.0000
omega    0.0000   0.0000   3.2292  0.0012
alpha1   0.0641   0.0097   6.6220  0.0000
beta1    0.9245   0.0073 127.1333  0.0000
gamma1   0.0042   0.0106   0.3921  0.6950
shape    8.1031   0.7824  10.3565  0.0000
LogLik = 11986.66 | AIC = -4.37 | BIC = -4.36
xm: Ljung-Box Q(20) sur résidus      = 17.41, p = 0.6260
xm: Ljung-Box Q²(20) sur résidus^2 = 32.52, p = 0.0381
```

Figure 10: ARMA(1,1)-GJR-GARCH(1,1) Specifications: (base) No Exogenous Variable; (xm) Exogenous in Conditional Mean

```
--- xv ---
        Estimate StdError  z.value p.value
mu        0.0003   0.0004   0.8107  0.4175
ar1      -0.2851   0.1426  -1.9997  0.0455
ma1       0.3453   0.1393   2.4785  0.0132
omega     0.0000   0.0000   2.7006  0.0069
alpha1    0.0541   0.0094   5.7538  0.0000
beta1     0.9358   0.0088 106.8499  0.0000
gamma1    0.0066   0.0101   0.6541  0.5130
vxreg1    0.0000   0.0002   0.0000  1.0000
shape     8.1782   0.7819  10.4594  0.0000
LogLik = 11734.11 | AIC = -4.28 | BIC = -4.27
xv: Ljung-Box Q(20) sur résidus      = 25.31, p = 0.1897
xv: Ljung-Box Q²(20) sur résidus^2 = 48.74, p = 0.0003

--- xm_xv ---
        Estimate StdError  z.value p.value
mu        0.0000   0.0003   0.0746  0.9405
ar1      -0.3875   0.1703  -2.2754  0.0229
ma1       0.4250   0.1669   2.5471  0.0109
mxreg1    0.3660   0.0161  22.7597  0.0000
omega     0.0000   0.0000   3.2117  0.0013
alpha1    0.0641   0.0097   6.6207  0.0000
beta1     0.9245   0.0073 126.1528  0.0000
gamma1    0.0041   0.0110   0.3753  0.7074
vxreg1    0.0000   0.0002   0.0000  1.0000
shape     8.1039   0.7899  10.2596  0.0000
LogLik = 11986.66 | AIC = -4.37 | BIC = -4.36
xm_xv: Ljung-Box Q(20) sur résidus      = 17.42, p = 0.6257
xm_xv: Ljung-Box Q²(20) sur résidus^2 = 32.53, p = 0.0380
```

Figure 11: ARMA(1,1)-GJR-GARCH(1,1) Specifications: (xv) Exogenous in Conditional Variance; (xm_xv) Exogenous in Mean and Variance

AR(1)-EGARCH(1,1)

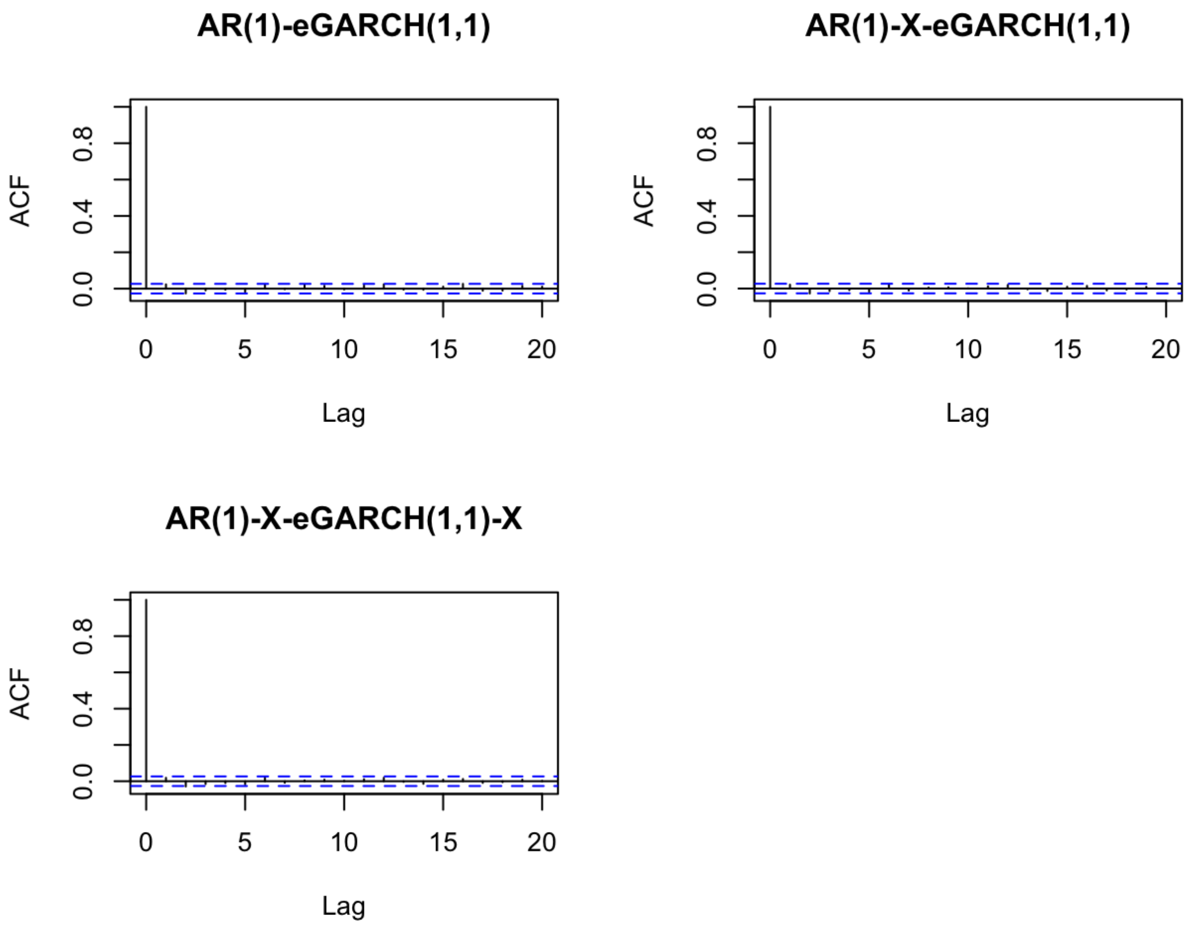

Figure 12: AR(1)-EGARCH(1,1) ACF

```
--- base ---
        Estimate StdError   z.value p.value
mu        0.0003   0.0004    0.7607  0.4468
ar1       0.0578   0.0135    4.2691  0.0000
omega    -0.0544   0.0050  -10.8911  0.0000
alpha1   -0.0161   0.0067   -2.3887  0.0169
beta1     0.9923   0.0007 1393.0862  0.0000
gamma1    0.1193   0.0167    7.1518  0.0000
shape     8.0629   0.8874    9.0857  0.0000
LogLik = 11732.72 | AIC = -4.28 | BIC = -4.27
base: Ljung-Box Q(20) sur résidus      = 28.97, p = 0.0884
base: Ljung-Box Q²(20) sur résidus^2 = 80.06, p = 0.0000

--- xm ---
        Estimate StdError   z.value p.value
mu        0.0000   0.0003    0.0780  0.9378
ar1       0.0281   0.0144    1.9473  0.0515
mxreg1    0.3675   0.0164   22.3719  0.0000
omega    -0.0644   0.0051  -12.5093  0.0000
alpha1   -0.0134   0.0075   -1.7869  0.0740
beta1     0.9911   0.0007 1388.0620  0.0000
gamma1    0.1376   0.0129   10.6528  0.0000
shape     8.0691   0.8165    9.8819  0.0000
LogLik = 11986.06 | AIC = -4.37 | BIC = -4.36
xm: Ljung-Box Q(20) sur résidus      = 20.82, p = 0.4075
xm: Ljung-Box Q²(20) sur résidus^2 = 59.82, p = 0.0000
```

Figure 13: AR(1)-EGARCH(1,1) Specifications: (base) No Exogenous Variable; (xm) Exogenous in Conditional Mean

```
--- xm_xv ---
       Estimate StdError   z.value p.value
mu       0.0001   0.0004    0.4043  0.6860
ar1      0.0275   0.0152    1.8134  0.0698
mxreg1   0.3691   0.0162   22.7811  0.0000
omega   -0.0527   0.0048  -11.0371  0.0000
alpha1  -0.0030   0.0073   -0.4091  0.6825
beta1    0.9927   0.0007 1467.1741  0.0000
gamma1   0.1200   0.0159    7.5651  0.0000
vxreg1  -1.4611   0.2779   -5.2584  0.0000
shape    8.5565   1.0260    8.3397  0.0000
LogLik = 11998.85 | AIC = -4.38 | BIC = -4.36
xm_xv: Ljung-Box Q(20) sur résidus      = 19.96, p = 0.4604
xm_xv: Ljung-Box Q²(20) sur résidus^2 = 39.76, p = 0.0053
```

Figure 14: AR(1)-EGARCH(1,1) Specifications: (xm_xv) Exogenous in Mean and Variance

MA(1)-EGARCH(1-1)

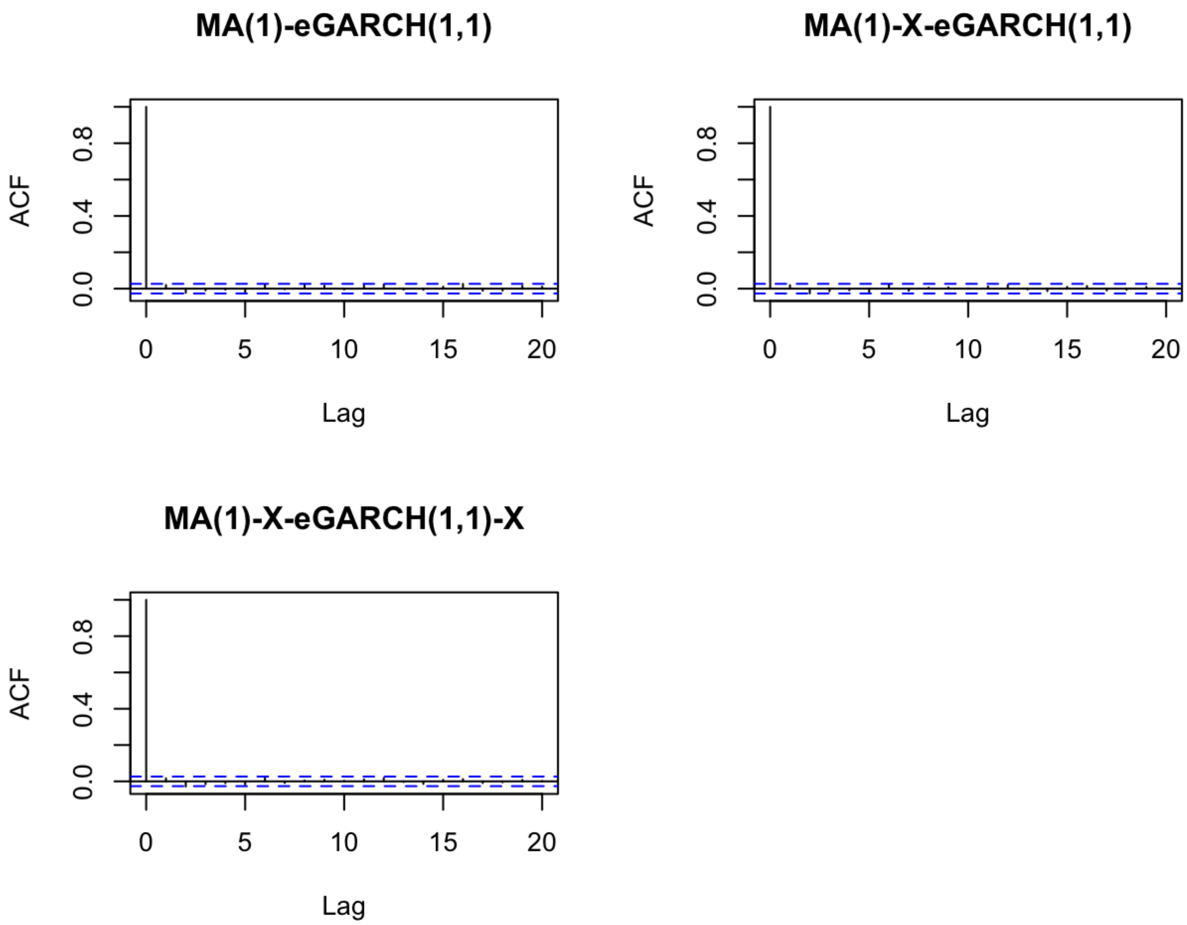

Figure 15: MA(1)-EGARCH(1,1) ACF

```
--- base ---
       Estimate StdError   z.value p.value
mu       0.0003   0.0004    0.7556  0.4499
ma1      0.0618   0.0140    4.4234  0.0000
omega   -0.0543   0.0050  -10.9148  0.0000
alpha1  -0.0161   0.0067   -2.3934  0.0167
beta1    0.9924   0.0007 1397.2123  0.0000
gamma1   0.1194   0.0166    7.1780  0.0000
shape    8.0697   0.8880    9.0876  0.0000
LogLik = 11733.36 | AIC = -4.28 | BIC = -4.27
base: Ljung-Box Q(20) sur résidus      = 27.15, p = 0.1311
base: Ljung-Box Q²(20) sur résidus^2 = 79.88, p = 0.0000

--- xm ---
       Estimate StdError   z.value p.value
mu       0.0000   0.0003    0.0753  0.9400
ma1      0.0306   0.0142    2.1514  0.0314
mxreg1   0.3671   0.0159   23.0595  0.0000
omega   -0.0643   0.0051  -12.5237  0.0000
alpha1  -0.0135   0.0075   -1.7893  0.0736
beta1    0.9911   0.0007 1390.2524  0.0000
gamma1   0.1375   0.0129   10.6489  0.0000
shape    8.0559   0.8133    9.9051  0.0000
LogLik = 11986.24 | AIC = -4.37 | BIC = -4.36
xm: Ljung-Box Q(20) sur résidus      = 20.10, p = 0.4514
xm: Ljung-Box Q²(20) sur résidus^2 = 60.18, p = 0.0000
```

Figure 16: MA(1)-EGARCH(1,1) Specifications: (base) No Exogenous Variable; (xm) Exogenous in Conditional Mean; (xm_xv) Exogenous in Mean and Variance

```
--- xm_xv ---
       Estimate StdError   z.value p.value
mu       0.0001   0.0003    0.5769  0.5640
ma1      0.0296   0.0141    2.1040  0.0354
mxreg1   0.3688   0.0162   22.7700  0.0000
omega   -0.0527   0.0047  -11.0940  0.0000
alpha1  -0.0030   0.0073   -0.4139  0.6789
beta1    0.9927   0.0007 1475.1221  0.0000
gamma1   0.1199   0.0158    7.5952  0.0000
vxreg1  -1.4604   0.2773   -5.2659  0.0000
shape    8.5538   1.0241    8.3527  0.0000
LogLik = 11999.02 | AIC = -4.38 | BIC = -4.36
xm_xv: Ljung-Box Q(20) sur résidus      = 19.34, p = 0.4999
xm_xv: Ljung-Box Q²(20) sur résidus^2 = 39.97, p = 0.0050
```

Figure 17: MA(1)-EGARCH(1,1) Specifications: (xm_xv) Exogenous in Mean and Variance

Out-of-Sample

```
=== Diebold-Mariano t-statistics (MSE) ===
                        ARMA(1,2)-GARCH(1,2)  ARMA(1,2)-X-GARCH(1,2)  ARMA(1,1)-X-GARCH(1,1)  ARMA(1,1)-GARCH(1,1)
ARMA(1,2)-GARCH(1,2)                     NaN               -1.227401               -1.376952             -2.161917
ARMA(1,2)-X-GARCH(1,2)              1.227401                     NaN               -0.791584              0.831886
ARMA(1,1)-X-GARCH(1,1)              1.376952                0.791584                     NaN              1.074437
ARMA(1,1)-GARCH(1,1)                2.161917               -0.831886               -1.074437                   NaN

=== Diebold-Mariano p-values (MSE) ===
                        ARMA(1,2)-GARCH(1,2)  ARMA(1,2)-X-GARCH(1,2)  ARMA(1,1)-X-GARCH(1,1)  ARMA(1,1)-GARCH(1,1)
ARMA(1,2)-GARCH(1,2)                     NaN                0.219672                0.168527              0.030625
ARMA(1,2)-X-GARCH(1,2)              0.219672                     NaN                0.428603              0.405473
ARMA(1,1)-X-GARCH(1,1)              0.168527                0.428603                     NaN              0.282627
ARMA(1,1)-GARCH(1,1)                0.030625                0.405473                0.282627                   NaN

=== Diebold-Mariano t-statistics (MAE) ===
                        ARMA(1,2)-GARCH(1,2)  ARMA(1,2)-X-GARCH(1,2)  ARMA(1,1)-X-GARCH(1,1)  ARMA(1,1)-GARCH(1,1)
ARMA(1,2)-GARCH(1,2)                     NaN               -0.731638                0.238612            -11.788411
ARMA(1,2)-X-GARCH(1,2)              0.731638                     NaN                1.187939             -0.587466
ARMA(1,1)-X-GARCH(1,1)             -0.238612               -1.187939                     NaN             -1.344447
ARMA(1,1)-GARCH(1,1)               11.788411                0.587466                1.344447                   NaN

=== Diebold-Mariano p-values (MAE) ===
                        ARMA(1,2)-GARCH(1,2)  ARMA(1,2)-X-GARCH(1,2)  ARMA(1,1)-X-GARCH(1,1)  ARMA(1,1)-GARCH(1,1)
ARMA(1,2)-GARCH(1,2)                     NaN                0.464390                0.811406              0.000000
ARMA(1,2)-X-GARCH(1,2)              0.464390                     NaN                0.234857              0.556891
ARMA(1,1)-X-GARCH(1,1)              0.811406                0.234857                     NaN              0.178804
ARMA(1,1)-GARCH(1,1)                0.000000                0.556891                0.178804                   NaN
```

Figure 18: Performance of Diebold Mariano Test

Code Used

```
# stationarity_tests_extended.R

install.packages("readxl")
install.packages("urca")

library(readxl)    # pour lire Excel
library(urca)      # pour ur.df() et ur.kpss()

fichier_excel <- "Downloads/DATA␣FINAL.xlsx"
df <- read_excel(path = fichier_excel, sheet = 1)

y_full <- df$LOG_WTI
y      <- na.omit(y_full)

adf_urca <- ur.df(y, type = "drift", selectlags = "AIC")
cat("===␣ADF␣(ur.df)␣test␣sur␣LOG_WTI␣===\n")
cat("Statistique␣de␣test␣:", adf_urca@teststat, "\n")
cat("Valeurs␣critiques␣:\n")
print(adf_urca@cval)

kpss_urca <- ur.kpss(y, type = "mu", lags = "short")
cat("\n===␣KPSS␣(ur.kpss)␣test␣sur␣LOG_WTI␣===\n")
summary(kpss_urca)

plot.ts(y,
        main = "Log-Return␣WTI␣(sans␣NAs)",
        ylab = "LOG_WTI",
        xlab = "Index")

z_full <- df$`Daily Change T10YIE`
z      <- na.omit(z_full)

adf_urca_z <- ur.df(z, type = "drift", selectlags = "AIC")
cat("\n===␣ADF␣(ur.df)␣test␣sur␣Daily␣Change␣T10YIE␣===\n")
cat("Statistique␣de␣test␣:", adf_urca_z@teststat, "\n")
```

```r
cat("Valeurs␣critiques␣:\n")
print(adf_urca_z@cval)

kpss_urca_z <- ur.kpss(z, type = "mu", lags = "short")
cat("\n===␣KPSS␣(ur.kpss)␣test␣sur␣Daily␣Change␣T10YIE␣===\n")
summary(kpss_urca_z)

plot.ts(z,
        main = "Daily␣Change␣T10YIE␣(sans␣NAs)",
        ylab = "Daily␣Change␣T10YIE",
        xlab = "Index")
```

Listing 1: stationarity_tests_extended.R

```r
# descriptives_subperiods_extended.R

required_pkgs <- c(
  "readxl", "dplyr", "lubridate",
  "moments", "knitr", "tibble", "tidyr"
)
for(pkg in required_pkgs) {
  if (!requireNamespace(pkg, quietly = TRUE)) {
    install.packages(pkg)
  }
}

library(readxl)
library(dplyr)
library(lubridate)
library(moments)
library(knitr)
library(tibble)
library(tidyr)

fichier <- "Downloads/DATA␣FINAL.xlsx"

df_transformed <- read_excel(fichier, sheet = 1) %>%
```

```
  mutate(observation_date = as.Date(observation_date))
df_wti <- read_excel(fichier, sheet = 3) %>%
  mutate(observation_date = as.Date(observation_date))
df_t10 <- read_excel(fichier, sheet = 4) %>%
  mutate(observation_date = as.Date(observation_date))

periods <- tribble(
  ~period,     ~start,        ~end,
  "2003-2006", "2003-01-02", "2006-12-31",
  "2007-2009", "2007-01-01", "2009-12-31",
  "2010-2019", "2010-01-01", "2019-12-31",
  "2020-2024", "2020-01-01", "2024-12-31"
) %>%
  mutate(
    start = as.Date(start),
    end   = as.Date(end)
  )

calc_stats <- function(df, value_var) {
  df %>%
    crossing(periods) %>%
    filter(observation_date >= start,
           observation_date <= end) %>%
    group_by(period) %>%
    summarise(
      N        = sum(!is.na(.data[[value_var]])),
      mean     = mean(.data[[value_var]], na.rm = TRUE),
      sd       = sd(.data[[value_var]], na.rm = TRUE),
      skewness = skewness(.data[[value_var]], na.rm = TRUE),
      kurtosis = kurtosis(.data[[value_var]], na.rm = TRUE),
      min      = min(.data[[value_var]], na.rm = TRUE),
      q1       = quantile(.data[[value_var]], 0.25, na.rm = TRUE),
      median   = median(.data[[value_var]], na.rm = TRUE),
      q3       = quantile(.data[[value_var]], 0.75, na.rm = TRUE),
      max      = max(.data[[value_var]], na.rm = TRUE),
      .groups  = "drop"
```

```r
    )
}

full_stats <- function(df, value_var, label = "Full␣sample") {
  x <- na.omit(as.numeric(df[[value_var]]))
  data.frame(
    period   = label,
    N        = length(x),
    mean     = mean(x),
    sd       = sd(x),
    skewness = skewness(x),
    kurtosis = kurtosis(x),
    min      = min(x),
    q1       = quantile(x, 0.25),
    median   = median(x),
    q3       = quantile(x, 0.75),
    max      = max(x)
  )
}

stats_wti_sub       <- calc_stats(df_wti,           "DCOILWTICO")
stats_t10_sub       <- calc_stats(df_t10,           "T10YIE")
stats_log_wti_sub   <- calc_stats(df_transformed,   "LOG_WTI")
stats_daily_t10_sub <- calc_stats(df_transformed,   "Daily␣Change␣
   T10YIE")

stats_wti_full       <- full_stats(df_wti,           "DCOILWTICO")
stats_t10_full       <- full_stats(df_t10,           "T10YIE")
stats_log_wti_full   <- full_stats(df_transformed,   "LOG_WTI")
stats_daily_t10_full <- full_stats(df_transformed,   "Daily␣Change␣
   T10YIE")

cat("\n===␣Descriptives␣WTI␣(DCOILWTICO)␣par␣sous-p riode␣+␣full␣
   sample␣===\n")
kable(bind_rows(stats_wti_sub, stats_wti_full), digits = 4)

```

```r
cat("\n=== Descriptives T10YIE (T10YIE) par sous-p riode + full sample ===\n")
kable(bind_rows(stats_t10_sub, stats_t10_full), digits = 4)

cat("\n=== Descriptives LOG WTI par sous-p riode + full sample ===\n")
kable(bind_rows(stats_log_wti_sub, stats_log_wti_full), digits = 4)

cat("\n=== Descriptives Daily Change T10YIE par sous-p riode + full sample ===\n")
kable(bind_rows(stats_daily_t10_sub, stats_daily_t10_full), digits = 4)

jb_full <- function(x) {
  x <- na.omit(as.numeric(x))
  n <- length(x)
  s <- skewness(x)
  k <- kurtosis(x)
  JB <- n/6 * (s^2 + k^2/4)
  pval <- 1 - pchisq(JB, df = 2)
  data.frame(
    n        = n,
    skew     = s,
    kurtosis = k,
    JB       = JB,
    p.value  = pval
  )
}

jb_log_wti      <- jb_full(df_transformed$LOG_WTI)
jb_daily_t10ie  <- jb_full(df_transformed$`Daily Change T10YIE`)

cat("\n=== JarqueBera sur l' chantillon complet ===\n")
cat("\nLOG WTI :\n");      print(jb_log_wti)
cat("\nDaily Change T10YIE :\n"); print(jb_daily_t10ie)
```

Listing 2: descriptives_subperiods_extended.R

```
# 1) Installer et charger les packages
install.packages(c("rugarch", "readxl", "forecast"))
library(rugarch)
library(readxl)
library(forecast)

# 2) Charger les donn es
fichier <- "Desktop/RStudio/DATA␣FINAL.xlsx"
df <- read_excel(fichier, sheet = 1)
df <- df[complete.cases(df$Daily_Change_T10YIE, df$LOG_WTI), ]
y <- df$Daily_Change_T10YIE
x <- df$LOG_WTI
x_mat <- matrix(x, ncol = 1)  # rugarch attend une matrice

stopifnot(all(complete.cases(x_mat)))  # V rifie qu'il n'y a pas de NA

# 3) Fonction pour cr er une sp cification selon les options
create_spec <- function(xm = FALSE, xv = FALSE) {
  mean_reg <- if (xm) x_mat else NULL
  var_reg  <- if (xv) x_mat else NULL
  ugarchspec(
    variance.model = list(model = "sGARCH", garchOrder = c(1, 1),
                          external.regressors = var_reg),
    mean.model = list(armaOrder = c(1, 1), include.mean = TRUE,
                      external.regressors = mean_reg),
    distribution.model = "std"
  )
}

# 4) Estimer les 4 mod les
models <- list(
  base   = ugarchfit(spec = create_spec(FALSE, FALSE), data = y),
  xm     = ugarchfit(spec = create_spec(TRUE,  FALSE), data = y),
  xv     = ugarchfit(spec = create_spec(FALSE, TRUE),  data = y),
```

```
  xm_xv  = ugarchfit(spec = create_spec(TRUE,  TRUE),  data = y)
)

# 5) Affichage des r sultats
print_model <- function(name, fit) {
  cat(sprintf("\n---␣%s␣---\n", name))
  coefs <- coef(fit)
  se <- sqrt(diag(vcov(fit)))
  z <- coefs / se
  pval <- 2 * pnorm(-abs(z))
  res <- data.frame(Estimate = round(coefs, 4),
                    StdError = round(se, 4),
                    z.value = round(z, 4),
                    p.value = round(pval, 4))
  print(res)
  cat(sprintf("LogLik␣=␣%.2f␣|␣AIC␣=␣%.2f␣|␣BIC␣=␣%.2f\n",
              likelihood(fit), infocriteria(fit)[1], infocriteria(fit
                )[2]))
}

# 6) Diagnostic Ljung-Box + ACF des r sidus standardis s
print_diag <- function(name, fit) {
  stdres <- residuals(fit, standardize = TRUE)
  lb <- Box.test(stdres, lag = 20, type = "Ljung-Box")
  cat(sprintf("%s:␣Ljung-Box␣Q(20)␣=␣%.2f,␣p␣=␣%.4f\n", name, lb$
     statistic, lb$p.value))
}

# 7) ACF avec corrections
model_titles <- list(
  base = "Mod le␣GARCH(1,1)␣de␣base",
  xm = "GARCH(1,1)␣avec␣r gress.␣moyenne",
  xv = "GARCH(1,1)␣avec␣r gress.␣variance",
  xm_xv = "GARCH(1,1)␣avec␣r gress.␣moy.␣&␣var."
)

```

```r
par(mfrow = c(2, 2))  # Disposition des graphiques
for (name in names(models)) {
  fit <- models[[name]]
  print_model(name, fit)
  print_diag(name, fit)

  # ACF corrig e : as.numeric + titre + taille du titre r duite
  acf(as.numeric(residuals(fit, standardize = TRUE)),
      lag.max = 20,
      main = model_titles[[name]],
      cex.main = 0.9)
}
par(mfrow = c(1, 1))  # R initialise la disposition

# 8)  tude  Out-of-Sample
n_forecast <- 100
y_in <- y[1:(length(y) - n_forecast)]
x_in <- x_mat[1:(length(y) - n_forecast), , drop = FALSE]
y_out <- y[(length(y) - n_forecast + 1):length(y)]
x_out <- x_mat[(length(y) - n_forecast + 1):length(y), , drop = FALSE]

forecast_one_step <- function(spec, y_all, x_all, start_idx) {
  if ((start_idx + 1) > nrow(x_all)) return(NA)

  fit <- tryCatch({
    ugarchfit(spec, data = y_all[1:start_idx])
  }, error = function(e) return(NA))

  if (inherits(fit, "uGARCHfit")) {
    pred <- tryCatch({
      ugarchforecast(fit, n.ahead = 1,
                     external.forecasts = list(
                       mregfor = x_all[start_idx + 1, , drop = FALSE],
                       vregfor = x_all[start_idx + 1, , drop = FALSE]
```

```
                     ))
    }, error = function(e) return(NA))

    if (!is.na(pred)) {
      return(fitted(pred))
    } else {
      return(NA)
    }
  } else {
    return(NA)
  }
}

# Pr voir avec chaque mod le
oos_results <- list()

for (model_name in names(models)) {
  cat(sprintf("Pr visions OOS pour mod le %s\n", model_name))
  xm <- model_name %in% c("xm", "xm_xv")
  xv <- model_name %in% c("xv", "xm_xv")
  spec <- create_spec(xm = xm, xv = xv)

  forecasts <- numeric(n_forecast)
  for (i in 1:n_forecast) {
    idx <- length(y_in) - n_forecast + i
    forecasts[i] <- forecast_one_step(spec, y, x_mat, idx)
  }

  errors <- y_out - forecasts
  mse <- mean(errors^2, na.rm = TRUE)
  mae <- mean(abs(errors), na.rm = TRUE)
  oos_results[[model_name]] <- list(forecast = forecasts,
                                    error = errors,
                                    MSE = mse,
                                    MAE = mae)
  cat(sprintf("MAE = %.4f | MSE = %.4f\n", mae, mse))
```

```r
}

# 9) Test de Diebold-Mariano entre mod les
model_pairs <- combn(names(oos_results), 2)
cat("\nTest␣de␣Diebold-Mariano␣entre␣mod les␣(MSE):\n")
for (i in 1:ncol(model_pairs)) {
  m1 <- model_pairs[1, i]
  m2 <- model_pairs[2, i]
  e1 <- oos_results[[m1]]$error
  e2 <- oos_results[[m2]]$error
  valid_idx <- complete.cases(e1, e2)
  dm <- dm.test(e1[valid_idx], e2[valid_idx],
                alternative = "two.sided", h = 1, power = 2)
  cat(sprintf("%s␣vs␣%s:␣DM␣Stat␣=␣%.4f,␣p-value␣=␣%.4f\n",
              m1, m2, dm$statistic, dm$p.value))
}
```

Listing 3: ARMA-GARCH.R