Econometrics: Assignment 2

Forecasting Inflation in Portugal

Group 13 – class 21

Alessandro Caggia, Virginia de Paruta, Luigi Lorenzoni

Data and Variables Description

Our dataset consolidates a comprehensive set of **monthly** macro-financial indicators sourced from major institutional databases including the **European Central Bank** (ECB), **Eurostat**, and the **Federal Reserve Economic Data** (FRED).

The selected variables reflect a broad set of dimensions known to influence price developments, including monetary policy signals, financial market stress, real and external economic activity, and the role of expectations in shaping forward-looking behavior. Specifically, we included:

MONETARY AND FINANCIAL VARIABLES

EONIA, **bank interest rates**, **cost of borrowing, yields on Treasury bonds**, and the **EUR/USD exchange rate**. Together, they capture the stance of monetary policy, financial conditions, liquidity in capital markets. Their inclusion allows us to account for the transmission of interest rates to prices, the cost of credit, and exchange rate pass-through effects that can influence import prices.

ECONOMIC OUTPUT AND ACTIVITY

This group includes **GDP** at different aggregation levels (Euro area, global, Portugal), the **unemployment rate, and government net investment**. These variables are central to measuring cyclical dynamics and output gaps, which directly affect inflation through demand-pull mechanisms and public spending effects.

ENERGY AND COMMODITY PRICES

Incorporating **oil prices**, **natural gas prices**, and a **global energy price index** helps capture external supply shocks and cost-push pressures. These are essential in understanding the direct and indirect effects of energy markets on production costs, inflation expectations, and consumer prices, particularly in energy-importing economies.

SENTIMENT AND EXPECTATION VARIABLES

This category comprises the **Economic Sentiment Indicator,** industrial and services **confidence indicators**, **inflation expectations** and the **consumer sentiment index**. These survey-based indicators summarize the forward-looking views of households and firms, providing timely signals about expected future economic conditions.

PRICE AND INFLATION INDICATORS

This category includes the **Harmonized Index of Consumer Prices (HICP)** — both overall and selected subcomponents — along with the **Producer Price Index (PPI)**. These variables are crucial for capturing both headline and core inflation trends, as well as sector-specific price dynamics. Together, these indicators allow us to monitor the full inflationary pipeline, from production to final consumption, supporting a deeper understanding of inflation persistence and volatility.

Standard OLS

Ordinary Least Squares (OLS) regression is one of the most widely used methods in econometrics, forming the foundation for many forecasting and analytical techniques. The method is fast to compute, easy to interpret, and commonly used as a **benchmark** when evaluating more complex models. Its appeal lies in its simplicity and in the clear economic intuition it provides about the relationship between dependent and explanatory variables.

Despite its simplicity, it often performs well in practice, especially when the underlying assumptions are reasonably satisfied. For this reason, OLS is often the first step in empirical model building, providing a solid baseline before considering more sophisticated techniques. It estimates a linear relationship of the form:

$$y = \beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k + \varepsilon$$

Key Assumptions:

- **Linearity:** The relationship between the dependent and independent variables is assumed to be linear, meaning the effect of a regressor on the outcome is constant across observations.
- **Exogeneity:** The regressors are uncorrelated with the error term. This ensures that the estimated coefficients are unbiased and reflect a true causal relationship.
- **Homoskedasticity**: The variance of the error term is constant across all levels of the independent variables. Violation of this assumption can lead to inefficient estimates and misleading inference.
- **No perfect multicollinearity**: Explanatory variables are not exact linear combinations of each other. High multicollinearity can inflate standard errors and make it difficult to assess the individual effect of each variable.

Main Outputs:

- **Estimated coefficients (β):** Indicate the direction and magnitude of the effect each independent variable has on the dependent variable. Their statistical significance is assessed via p-values and confidence intervals.
- **R^2 (Coefficient of determination)**: Measures the proportion of variation in the dependent variable explained by the model. A higher R^2 indicates a better fit.
- **Prediction intervals:** Provide a range in which future observations are likely to fall, accounting for model uncertainty and residual variance.

Variable	Estimate	Std. Error
Intercept	2.033***	(0.266)
`Price trends over last 12 months`	1.53***	(0.285)
`Major purchases at present`	-0.421*	(0.2)
`Assessment of export order-book levels`	0.506**	(0.186)
`Employment expectations for the months ahead`	0.298+	(0.155)
`exchange rate (USD for 1 euro)`	-0.292+	(0.155)
`Global price of energy index (2016=100)`	0.971+	(0.572)
`Yield on U.S. Treasury Securities at 10-Year Constant Maturity`	0.357+	(0.198)
`Yield on U.S. Treasury Securities at 3-Month Constant Maturity`	-0.532*	(0.212)
`Construction confidence indicator (5%)`	-0.676*	(0.298)

Forward & Backward Variable Selection

The objective of variable selection is to identify the most informative predictors to enhance both **forecasting accuracy** and **model interpretability.** Including too many variables can lead to overfitting and unnecessary complexity, while omitting relevant ones may result in biased estimates. To address this, we adopt forward and backward selection methods, which iteratively assess the statistical significance of each regressor.

Forward Selection Model

Variable	Estimate	Std. Error
Intercept	2.033***	(0.266)
`Construction confidence indicator (5%)`	-0.676*	(0.298)
`Price trends over last 12 months`	1.53***	(0.285)
`Major purchases at present`	-0.421*	(0.2)
`Assessment of export order-book levels`	0.506**	(0.186)
`Employment expectations for the months ahead`	0.298+	(0.155)
`exchange rate (USD for 1 euro)`	-0.292+	(0.155)
`Global price of energy index (2016=100)`	0.971+	(0.572)
`Yield on U.S. Treasury Securities at 10-Year Constant Maturity`	0.357+	(0.198)
`Yield on U.S. Treasury Securities at 3-Month Constant Maturity`	-0.532*	(0.212)

Forward Selection: Begins with no predictors and adds variables one at a time, choosing those that most improve model fit—typically measured by reductions in the Akaike Information Criterion (AIC)

Backward Elimination Model

Variable	Estimate	Std. Error
Intercept	2.098***	(0.149)
`Services confidence indicator (30 %)`	0.448+	(0.236)
`Construction confidence indicator (5%)`	-0.615***	(0.161)
`The Economic sentiment indicator is a composite measure (average = 100)`	-1.273*	(0.509)
`Financial situation over last 12 months`	1.824***	(0.234)
`Price trends over last 12 months`	1.664***	(0.139)
`Major purchases at present`	-0.416**	(0.137)
`Assessment of export order-book levels`	0.557***	(0.131)
`Assessment of stocks of finished products`	-0.2*	(0.087)
`Employment expectations for the months ahead`	0.318**	(0.108)
`Cost of borrowing for corporations (percent per annum)`	0.439*	(0.215)
PPI	0.745**	(0.279)
`exchange rate (USD for 1 euro)`	-0.293*	(0.125)
`Oil price europe`	-0.706*	(0.309)
`Global price of energy index (2016=100)`	0.903*	(0.369)
`Yield on U.S. Treasury Securities at 10-Year Constant Maturity`	0.33*	(0.136)
`Yield on U.S. Treasury Securities at 3-Month Constant Maturity`	-0.586***	(0.137)

Backward Elimination: Starts with the full set of predictors and iteratively removes the least significant variables, again based on AIC or another selection metric.

Stepwise Selection Model

Variable	Estimate	Std. Error
Intercept	2.098***	(0.149)
`Services confidence indicator (30 %)`	0.448+	(0.236)
`Construction confidence indicator (5%)`	-0.615***	(0.161)
`The Economic sentiment indicator is a composite measure (average = 100)`	-1.273*	(0.509)
`Financial situation over last 12 months`	1.824***	(0.234)
`Price trends over last 12 months`	1.664***	(0.139)
`Major purchases at present`	-0.416**	(0.137)
`Assessment of export order-book levels`	0.557***	(0.131)
`Assessment of stocks of finished products`	-0.2*	(0.087)
`Employment expectations for the months ahead`	0.318**	(0.108)
`Cost of borrowing for corporations (percent per annum)`	0.439*	(0.215)
PPI	0.745**	(0.279)
`exchange rate (USD for 1 euro)`	-0.293*	(0.125)
`Oil price europe`	-0.706*	(0.309)
`Global price of energy index (2016=100)`	0.903*	(0.369)
`Yield on U.S. Treasury Securities at 10-Year Constant Maturity`	0.33*	(0.136)
`Yield on U.S. Treasury Securities at 3-Month Constant Maturity`	-0.586***	(0.137)

Stepwise Selection: A hybrid method that combines forward and backward approaches. At each step, it can add or remove variables depending on whether the change improves the model according to AIC. While it offers flexibility, it also carries a risk of overfitting.

Lasso, Ridge and PCA

LASSO, Elastic Net, and Principal Component Analysis (PCA) are powerful techniques for regularization and dimensionality reduction, particularly useful when the **number of predictors is large** or when **multicollinearity** is present

LASSO (Least Absolute Shrinkage and Selection Operator) imposes a penalty on the absolute size of regression coefficients, effectively shrinking some of them to exactly zero. This results in a simpler model by performing automatic variable selection, helping to reduce overfitting and improve interpretability.

Elastic Net combines the penalties of both LASSO and Ridge regression. It strikes a balance between variable selection (like LASSO) and coefficient shrinkage (like Ridge), making it especially effective when predictors are large in number or highly correlated.

Principal Component Analysis (PCA) reduces dimensionality by transforming the original predictors into a smaller set of uncorrelated principal components that capture the majority of the variance in the data. This not only simplifies the model but also helps mitigate multicollinearity and enhance forecasting stability.

LASSO-Selected OLS Model

Variable	Estimate	Std. Error
Intercept	1.939***	(0.191)
`Retail trade confidence indicator (5%)`	-0.378**	(0.135)
`Construction confidence indicator (5%)`	-0.701***	(0.161)
`Financial situation over last 12 months`	1.501***	(0.239)
`Price trends over last 12 months`	1.53***	(0.212)
`Major purchases at present`	-0.425**	(0.157)
`Assessment of export order-book levels`	0.393**	(0.12)
`Selling price expectations for the months ahead`	0.117+	(0.067)
`Employment expectations for the months ahead`	0.336**	(0.116)
`exchange rate (USD for 1 euro)`	-0.356**	(0.125)
`Yield on U.S. Treasury Securities at 10-Year Constant Maturity`	0.316*	(0.154)
`Yield on U.S. Treasury Securities at 3-Month Constant Maturity`	-0.542**	(0.17)

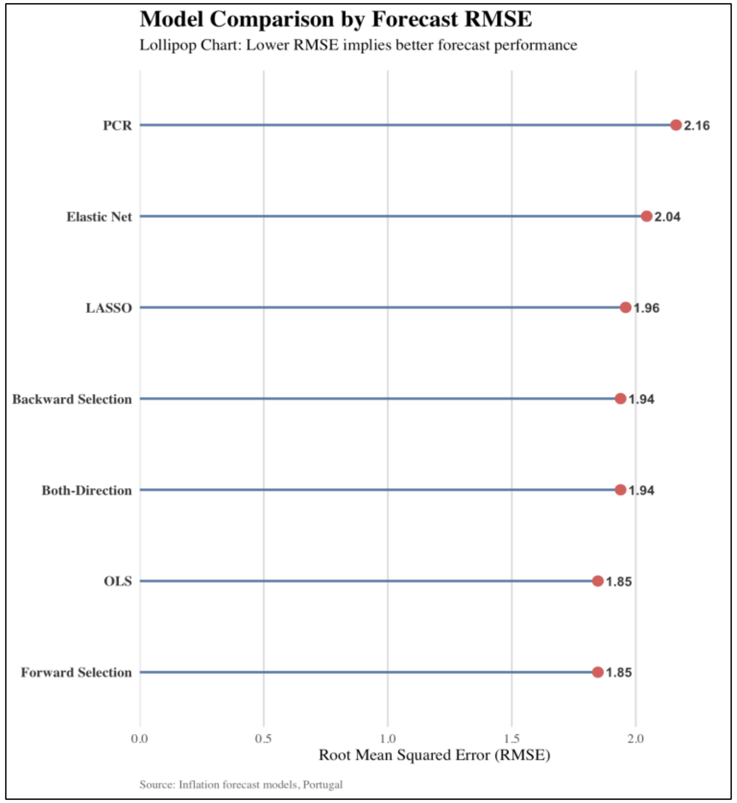
Elastic Net-Selected OLS Model

Variable	Estimate	Std. Error
Intercept	1.985***	(0.169)
`Retail trade confidence indicator (5%)`	-0.371**	(0.138)
`Construction confidence indicator (5%)`	-0.746***	(0.192)
`Financial situation over last 12 months`	1.359***	(0.293)
`Price trends over last 12 months`	1.411***	(0.252)
`Major purchases at present`	-0.391*	(0.156)
`Assessment of export order-book levels`	0.405***	(0.117)
`Selling price expectations for the months ahead`	0.122+	(0.068)
`Employment expectations for the months ahead`	0.337**	(0.115)
PPI	0.499+	(0.288)
`exchange rate (USD for 1 euro)`	-0.335**	(0.105)
`Natural gas price EU`	0.553+	(0.292)
`Yield on U.S. Treasury Securities at 10-Year Constant Maturity`	0.316*	(0.153)
`Yield on U.S. Treasury Securities at 3-Month Constant Maturity`	-0.604***	(0.18)

Principal Component Regression (PCR) Model

Variable	Estimate	Std. Error
Intercept	1.423***	(0.052)
PC1	0.082***	(0.011)
PC2	-0.258***	(0.018)
PC3	0.187***	(0.024)
PC4	-0.126***	(0.032)
PC5	-0.397***	(0.044)
PC7	-0.266***	(0.052)
PC9	0.339***	(0.056)
PC10	0.249***	(0.057)

Overall Performance Assessment

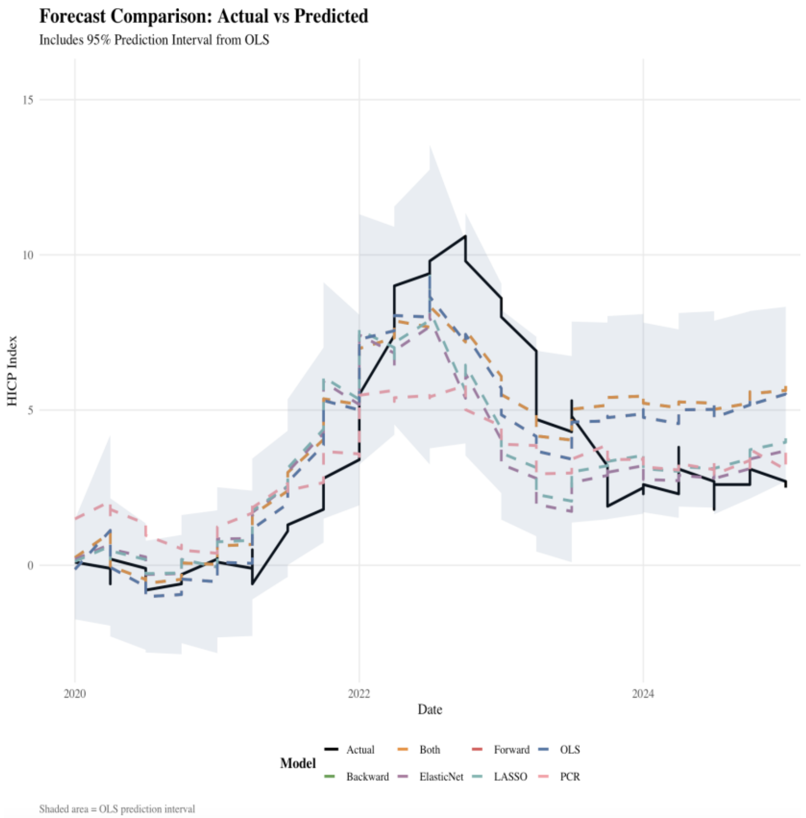


OLS and Forward Selection deliver the lowest RMSE (1.85), indicating the strongest predictive performance among all models.

Backward and Stepwise Selection follow closely with an RMSE of 1.94, suggesting a slight decrease in accuracy—likely due to the inclusion or exclusion of less relevant variables.

LASSO and Elastic Net produce marginally higher RMSEs (1.96 and 2.04, respectively), which may be attributed to penalization bias or weaker variable selection effectiveness in this specific context.

Principal Component Regression (PCR) performs the worst (RMSE of 2.16), likely because the extracted components do not align well with the underlying structure of the target variable.



Overall, **simpler linear models outperform regularized or transformed alternatives** in this forecasting task, emphasizing the strength of traditional approaches when relevant predictors are clearly defined.

The most consistently selected covariates across models include:

- **Exchange rate (USD for 1 euro)**
- **Employment expectations for the months ahead**
- **Price trends over the last 12 months**
- **Major purchases at present**
- **Assessment of export order-book levels**
- **Construction confidence indicator (5%)**

Across models, past price trends and employment expectations show positive effects, while the exchange rate (USD for 1 euro) has a negative effect. Notably, the Yield on U.S. Treasury Securities at 3-Month Constant Maturity consistently shows a negative and statistically significant effect, reinforcing its relevance in explaining the variation in the target variable.

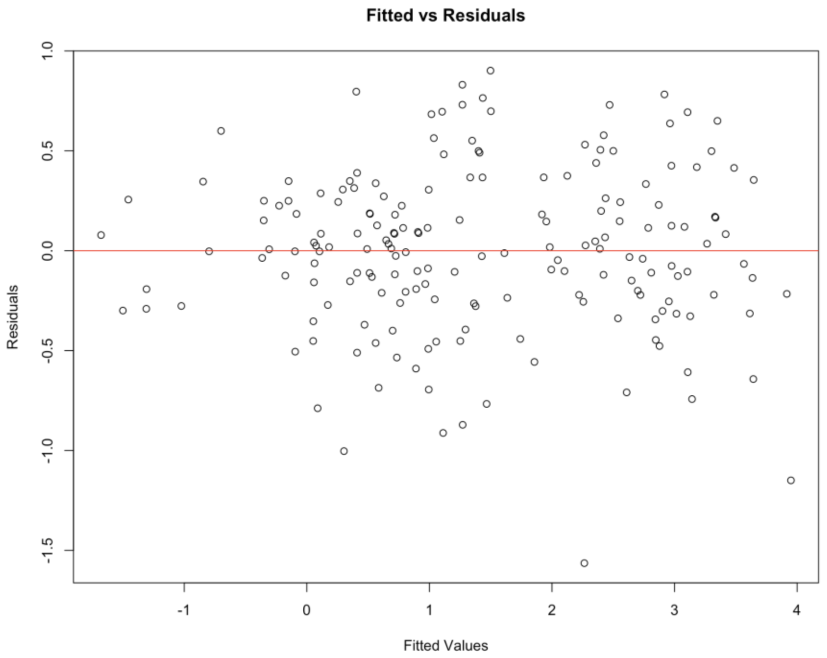
Threats to Model Misspecification

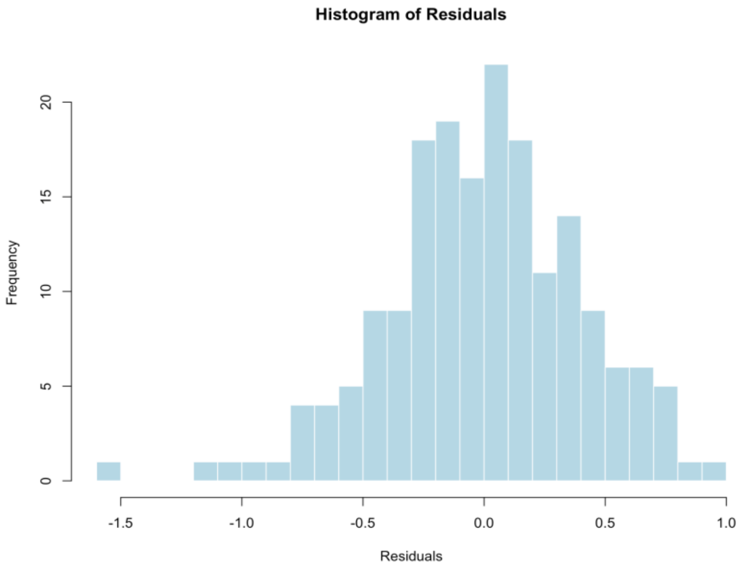
SERIAL CORRELATION

Serial correlation occurs when residuals are correlated across time, violating the OLS assumption of independence. This can distort standard errors and lead to unreliable inference in time series models.

The **Durbin-Watson test** suggests no evidence of first-order autocorrelation (p = 0.80).

However, the **Breusch-Godfrey test** indicates the presence of significant higher-order serial correlation (p = 0.007), implying that residuals are not fully independent across time.





HETEROSKEDASTICITY

Heteroskedasticity refers to non-constant variance in the residuals across observations. It violates the OLS assumption of homoskedasticity and can lead to inefficient estimates and biased inference.

The **Breusch-Pagan and White tests** both yield borderline significance (p ≈ 0.05), indicating mild evidence of heteroskedasticity in the model.

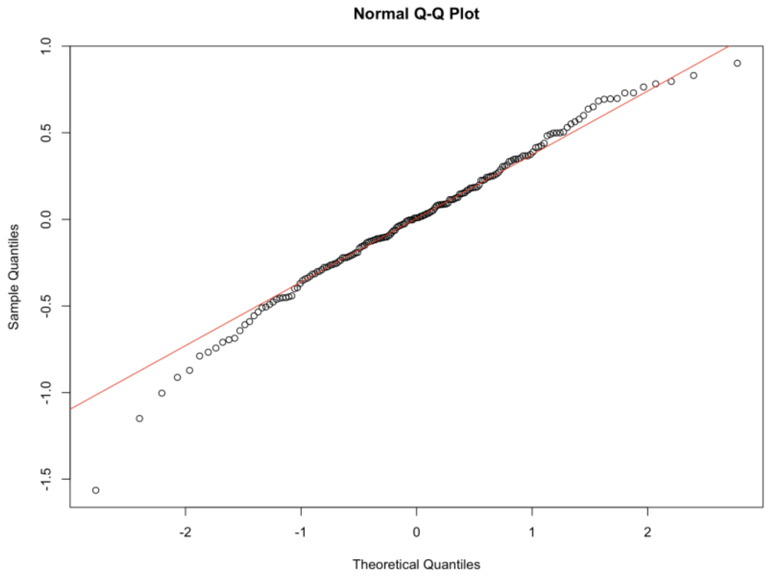
To correct for this and ensure consistent standard errors, we apply **Newey-West standard errors**, which adjust for both heteroskedasticity and autocorrelation, improving the reliability of inference.

NORMALITY

Normality of residuals is not strictly required for estimation but is important for valid hypothesis testing (e.g., t-tests and F-tests). Deviations from normality can weaken inference, especially in small samples.

The **Shapiro-Wilk test** (p = 0.08), along with visual inspection of the Q-Q plot and histogram, suggests that residuals are approximately normally distributed.

Although the test is slightly above conventional significance levels, the distribution appears symmetric and bell-shaped, indicating that the normality assumption is reasonably satisfied..



Arima: Stationarity and Structural Breaks

STATIONARITY

We assess the statistical properties of **Portugal's HICP (Overall Index)** to verify whether the series is stationary — a key requirement for reliable ARIMA modeling. Stationarity ensures that the model produces unbiased estimates and accurate forecasts.

To determine this, we apply two complementary tests:

- The **Augmented Dickey-Fuller (ADF) test**, where the null hypothesis is the presence of a unit root (i.e., non-stationarity)
- The **KPSS test**, where the null hypothesis is that the series is **stationary**

These tests help inform the required number of differencing operations:

- d: number of regular differences
- D: number of seasonal differences

The ADF test offers mild support for stationarity, while the KPSS test rejects the null, pointing to non-stationarity. Adopting a conservative approach, we proceed under the **assumption that the series is non-stationary.**

Based on these results:

- We set d = 1 (one regular difference)
- And D = 0 (no seasonal differencing)

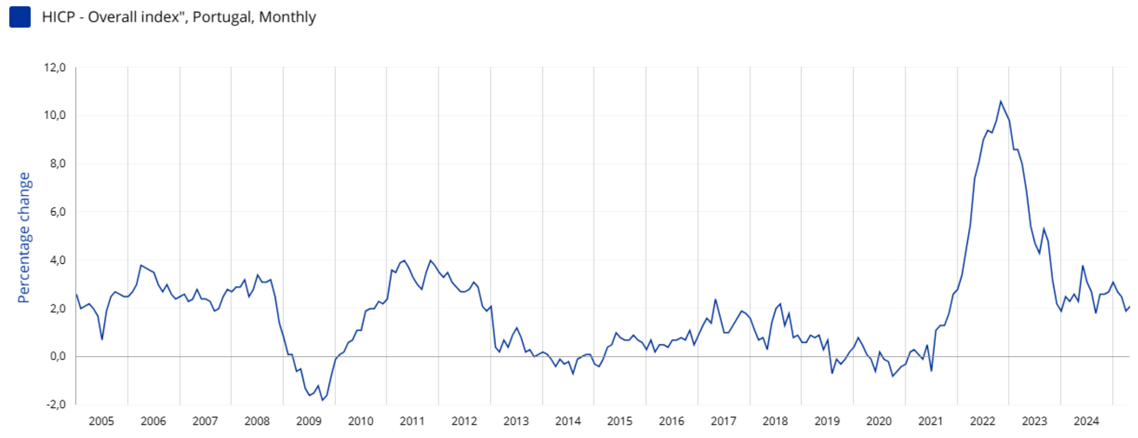
A first difference is applied to remove the stochastic trend. Seasonal ARIMA terms will later be introduced to capture seasonal variation, ensuring a parsimonious yet robust model specification.

STRUCTURAL BREAKS

To assess the presence of structural breaks, we applied the **Bai-Perron test**, which allows for the identification of statistically significant changes in the mean or trend without requiring pre-specified break dates. The test detected a structural break around January 2013, indicating a significant shift in the average level of inflation.'

This shift is likely related to macroeconomic adjustments following the **European sovereign debt crisis,** such as fiscal tightening, structural reforms, and changes in monetary policy frameworks. These events may have altered the underlying inflation dynamics, resulting in a permanent change in the data-generating process.

The presence of such a break can have important implications for model specification, as it may affect the stability of parameter estimates and lead to biased residuals if not properly accounted for. Moreover, without explicit adjustment, seasonal components might absorb part of the structural change, blurring the distinction between genuine seasonality and long-term regime shifts.



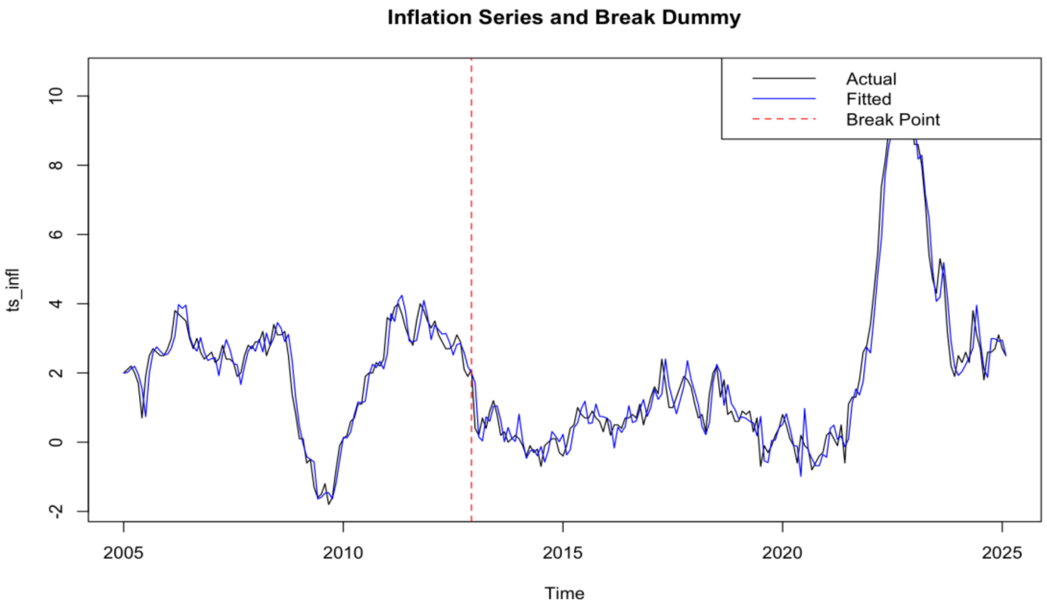
Arima: Stationarity and Structural Breaks

We estimate a Seasonal ARIMA model for Portugal's HICP inflation using automatic model selection with an exogenous break dummy, resulting in an ARIMA(1,1,1)(0,0,2)(12) specification. This includes a **dummy variable** to account for a structural shift around 2013.

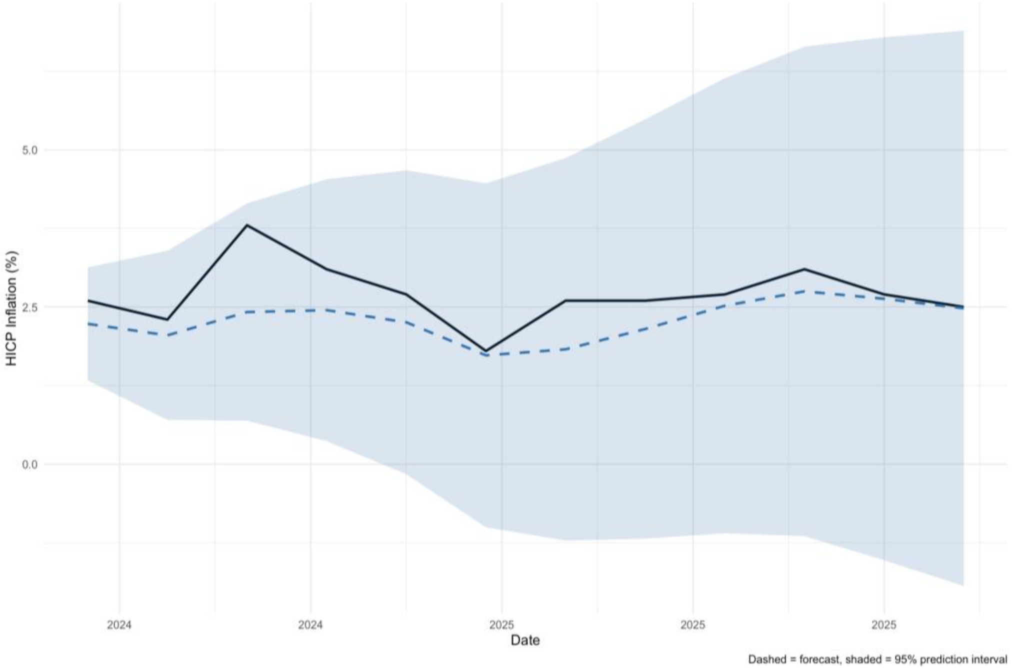
- **The non-seasonal component** shows strong dynamics, with an autoregressive term **(AR = 0.868)** indicating high persistence, and a negative moving average term **(MA = –0.769)** suggesting effective short-term shock correction.
- **Seasonality** is captured by two moving average terms at lags 12 and 24 (coefficients –0.52 and –0.138), reflecting annual and biannual cyclical adjustments.
- The **structural break dummy** has a positive but statistically insignificant effect (0.160, s.e. = 0.364), implying limited impact on the inflation path post-2013.

The model satisfies stationarity and invertibility conditions, with all estimated coefficients well below unity in absolute value.

.



ARIMA: Forecast vs Actual (past 12 months)

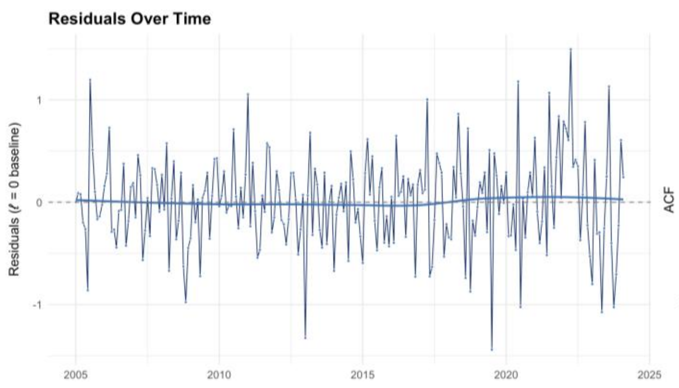


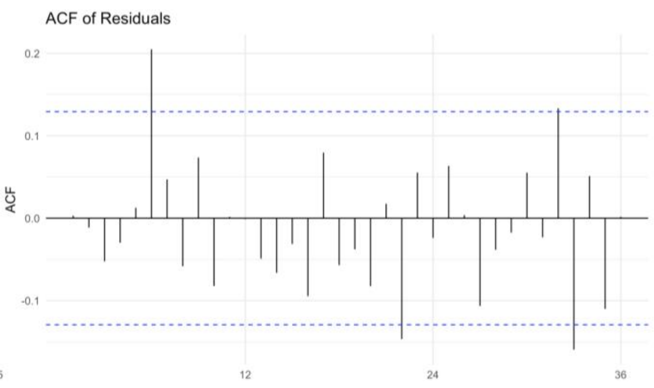
Arima: Residual Diagnostics

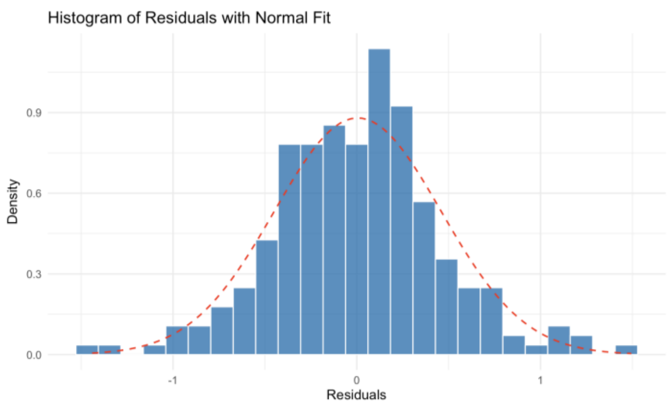
Overall fit and forecast accuracy: Residual diagnostics confirm that the ARIMA(1,1,1)(0,0,2)[12] model provides a statistically sound fit to the inflation data.

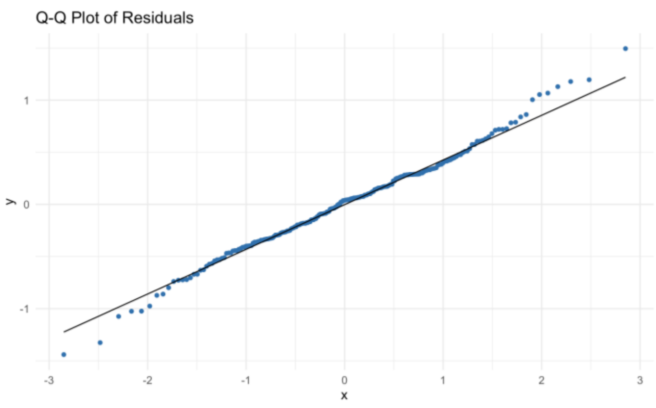
- No autocorrelation in residuals: $ACF_1 = –0.012$
- Low residual variance: $\sigma^2 = 0.212$
- No systematic bias: Mean Error ≈ 0
- Strong in-sample forecast performance: RMSE = 0.455; MASE = 0.195

These metrics indicate a clear improvement over naive models, and suggest the model effectively captures the trend and seasonal components of Portuguese inflation.









Statistical Residual Tests: Ljung-Box test ($p = 0.06$) and Breusch-Godfrey test ($p = 0.20$) both fail to reject the null of no serial correlation. This confirms that autocorrelation has been successfully removed from the residuals.

ARCH LM test ($p = 0.80$) shows no evidence of conditional heteroskedasticity, supporting the assumption of stable variance over time.

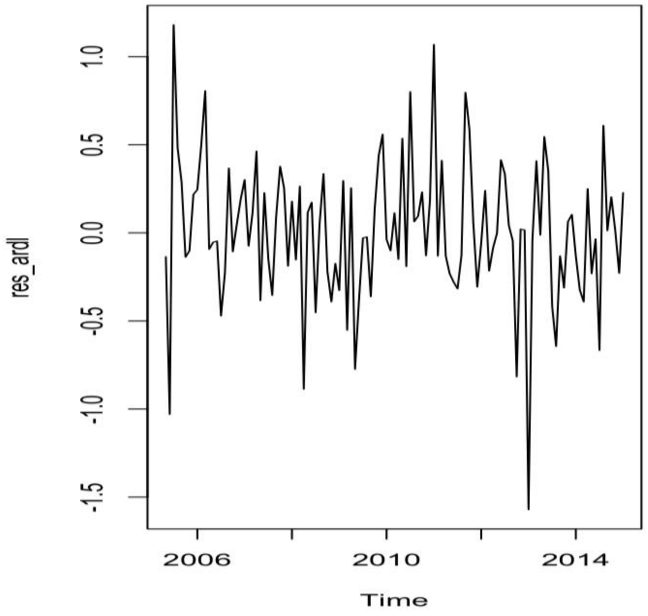
Jarque-Bera test ($p = 0.07$) indicates mild non-normality, which is common in macroeconomic data and generally not problematic for forecasting accuracy.

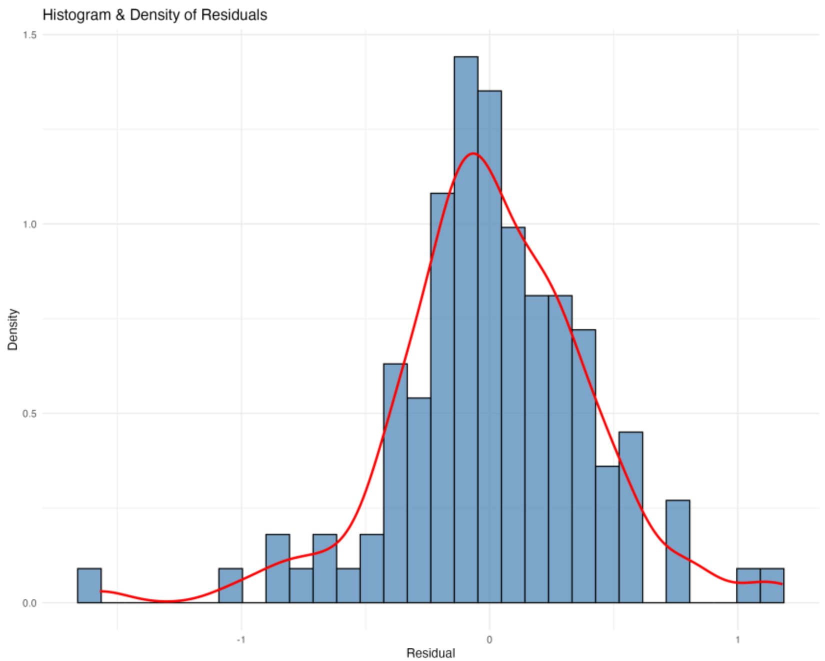
Conclusion: Taken together, the residuals behave approximately as **white noise** — they are uncorrelated, homoskedastic, and nearly normal. This validates the model's **adequacy for both inference and short-term inflation forecasting**.

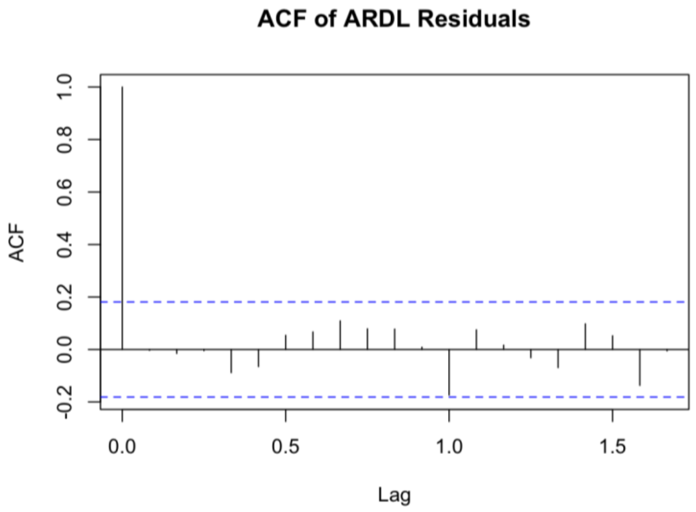
ARDL Model

For completeness, we also run an **ARDL(4,4)** model, so as to try to strike a balance between forecasting ability and economic interpretability.

We include as covariates: unemployment, inflation expectations and energy prices. This should mimick some of the classical covariates which the literature regards as drivers of inflation (e.g. Bernanke and Blanchard).







Residuals look as being almost **white noise** and- despite a mean slightly below zero- almost normally distributed. The ACF plot shows not much autocorrelation except for a spike at 0.

To further investigate this, we run Ljung-Box test and Breusch-Pagan test and find that we fail to **reject the null hypothesis of both serial correlation and homoskedasticity**, respectively with p-values equal to 0.85 and 0.22.

Jarque-Bera test however **rejects the normality** of residuals. This, though quite common in macro data, might be an issue for inference, which could be solved by using bootstrapping to derive forecast confidence intervals and make valid inference.

The model's fundamental assumption do not seem to be systematically violated.

One Step Ahead Forecast

We run dynamic one step ahead simulations for the ARDL model and for the AR(2). The forecast are made at time t for time t+1 estimating the model on the whole set of available information at time t. We compare the forecast to AR(2) as a benchmark.

A general observation is that both models track reasonably well inflation, with a worsening in forecasting ability in the last periods. This is in line with inflation settling down from a period of sustained inflation and likely varying structural relations. We compute RMSE and MAE for both models.

- RMSE ARDL: 0.63
- RMSE AR(2): 0.6124862
- MAE ARDL: 0.4711654
- MAE AR(2): 0.4738239

All errors computed on the training sample.

AR(2) seems to perform better than ARDL. A more parsimonious approach would suggest going for a simpler AR(2) process, rather than overhitting with a more complex model. Notably, our coefficient on inflation in the month before is around 1, a critical threshold in AR estimation.

Even though the forecast of these two models look quite good, both these two models underperform the previously estimated ARIMA in terms of MAE and RMSE. **ARIMA thus ultimately emerges as our preferred model.**