Time Series Analysis and Forecasting

Hossein John Ahmadi

Gauss

June 1, 2025

Contents

1 Stationarity in Time Series Analysis

Stationarity is fundamental in time series modeling. When a process is stationary, its statistical properties do not change over time. This ensures that the behavior observed in the past can be used reliably to make inferences about the future.

There are two main types of stationarity considered in practice:

Strict Stationarity

A process $\{X_t\}$ is strictly stationary if the joint distribution of any set of observations is invariant under time shifts. That is, for all $h, k \in Z$:

$$F_{X_{t_1},\ldots,X_{t_k}}(x_1, \ldots, x_k) = F_{X_{t_1+h},\ldots,X_{t_k+h}}(x_1, \ldots, x_k)$$

Implication: All moments (mean, variance, skewness, etc.) remain constant over time, assuming they exist.

Weak Stationarity

Weak stationarity is a less strict form of stationarity that requires only the first two moments to be time-invariant. A process $\{X_t\}$ is weakly stationary if:

$$\begin{cases} E[X_t] = \mu & \text{(constant mean)} \\ \mathrm{Var}(X_t) = \sigma^2 & \text{(constant variance)} \\ \mathrm{Cov}(X_t, X_{t+h}) = \gamma(h) & \text{(covariance depends only on lag } h\text{)} \end{cases}$$

Key Property: The autocorrelation function is well-defined and depends only on the lag h:

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}$$

2 Moving Average (MA)

A Moving Average model of order q, MA(q), is an initial model in time series analysis used to describe a stationary stochastic process. This model expresses the current value of the series y_t as a linear function of past white noise.

Definition of the MA(q) Model

The MA(q) process can be written as:

$$y_t = \mu + u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \cdots + \theta_q u_{t-q}$$

Where:

- μ is the mean of the series,
- $u_t \sim WN(0, \sigma^2)$ is a white noise process,
- $\theta_1, \theta_2, \ldots, \theta_q$ are the parameters of the model.

Rather than relying on realization in the study of MA(q) models, using the lag operator L:

$$y_t = \mu + \theta(L)u_t \quad \text{where } \theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q$$

Properties of an MA(q) model.

- **Stationarity:**

  The MA(q) model is always stationary since it is a finite linear combination of white noise.

- **Mean and Variance:**

$$E(y_t) = \mu$$

$$\text{Var}(y_t) = \sigma^2(1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2)$$

- **Autocovariance and Autocorrelation:**

  The autocovariance function is defined as:

$$\gamma_s = \text{Cov}(y_t, y_{t-s})$$

  For an MA(q) process:

  - $\gamma_s \neq 0$ only for $s \leq q$
  - $\gamma_s = 0$ for $s > q$

  The autocorrelation function (ACF) thus cuts off after lag q, making MA models identifiable by the behavior of their ACF.

3 Autoregressive (AR)

The Autoregressive (AR) model is one of the most common models in time series analysis for stationary processes. It describes a variable using its own past values and a random error term.

An AR(p) model of order p is formally defined as:

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t$$

Where:

- μ is the mean of the process,
- $\phi_1, \ldots, \phi_p$ are autoregressive coefficients,
- $\epsilon_t \sim WN(0, \sigma^2)$ is a white noise process.

An AR(p) can also be expressed using the lag operator L as follows:

$$\Phi(L)y_t = \mu + \epsilon_t$$

Where:

$$\Phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p$$

The stationary solution requires all the roots of the characteristic equation $\Phi(L) = 0$ to lie outside the unit circle (i.e., $|r| > 1$).

Let us discuss some important things about AR(p) models:

4 AR(p) to MA(∞)

In most cases, the AR(p) form of a model can be transformed into an MA(∞) form in time series models with a stationary series, offering several valuable benefits. Most notably, this is exclusively helpful when you want to see the lagged effect of shocks.

The AR(p) Model

In standard notation, an Autoregressive model of order p can be expressed as:

$$y_t = \mu + \sum_{i=1}^{p} \phi_i y_{t-i} + \epsilon_t \quad \text{with } \epsilon_t \sim WN(0, \sigma^2)$$

Where:

- μ is the mean (constant),
- ϕ_i are the autoregressive coefficients,
- ϵ_t is a white noise disturbance term.

Lag Operator

For compactness, we can introduce the lag operator L, so that $L^k y_t = y_{t-k}$. The model can now be expressed as:

$$\Phi(L) y_t = \mu + \epsilon_t \quad \text{where } \Phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p$$

That is, the Lag polynomial $\Phi(L)$ encapsulates the autoregressive behaviour with a short notation.

Invert the Lag Polynomial

Provided that you meet the stationarity assumption, the inverse of the lag polynomial exists as a convergent power series:

$$\Phi(L)^{-1} = \sum_{j=0}^{\infty} \theta_j L^j$$

This is placed on both sides of the equation:

$$y_t = \Phi(L)^{-1}(\mu + \epsilon_t) = \Phi(L)^{-1}\mu + \Phi(L)^{-1}\epsilon_t$$

Because everything that L operates on (i.e., time-varying component) allows us to apply a multiplier on the constant term, we have:

$$\Phi(L)^{-1}\mu = \mu \cdot \sum_{j=0}^{\infty} \theta_j = \mu'$$

MA(∞) in Final Form

We have then impacted the summarized AR(p) process as follows:

$$y_t = \mu' + \sum_{j=0}^{\infty} \theta_j \epsilon_{t-j}$$

We have reached the moving average (of infinite order), which expresses the present value of y_t as a weighted sum of all previous shocks.

Example : AR(2) to MA(∞) Conversion

Now consider an AR(2) model:

$$y_t = 0.6y_{t-1} + 0.2y_{t-2} + \epsilon_t$$

Written with the lag operator:

$$(1 - 0.6L - 0.2L^2)y_t = \epsilon_t$$

We want to represent this as:

$$y_t = \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i}$$

To do this, calculate the MA coefficients ψ_i recursively:

$$\psi_0 = 1$$

$$\psi_1 = 0.6 \cdot \psi_0 = 0.6$$

$$\psi_2 = 0.6 \cdot \psi_1 + 0.2 \cdot \psi_0 = 0.6 \cdot 0.6 + 0.2 = 0.56$$

$$\psi_3 = 0.6 \cdot \psi_2 + 0.2 \cdot \psi_1 = 0.6 \cdot 0.56 + 0.2 \cdot 0.6 = 0.456$$

Thus:

$$y_t = \epsilon_t + 0.6\epsilon_{t-1} + 0.56\epsilon_{t-2} + 0.456\epsilon_{t-3} + \cdots$$

Assume:

$$\epsilon_0 = 1.0, \quad \epsilon_1 = -0.5, \quad \epsilon_2 = 0.3, \quad \epsilon_3 = 0.1$$

Then:

$$y_3 = \epsilon_3 + 0.6\epsilon_2 + 0.56\epsilon_1 + 0.456\epsilon_0$$

$$y_3 = 0.1 + 0.6 \cdot 0.3 + 0.56 \cdot (-0.5) + 0.456 \cdot 1.0 = 0.1 + 0.18 - 0.28 + 0.456 = 0.456$$

5 ARMA(p, q)

The ARMA(p, q) model is a specific time series model that combines Autoregressive (AR) and Moving Average (MA) components to accurately model stationary time series that exhibit both long-run dependencies and short-run shocks.

Model Equation

$$y_t = \mu + \sum_{i=1}^{p} \phi_i y_{t-i} + \epsilon_t + \sum_{j=1}^{q} \theta_j \epsilon_{t-j}$$

where:

- ϕ_i: AR coefficients,
- θ_j: MA coefficients,
- $\epsilon_t \sim WN(0, \sigma^2)$: white noise,
- μ: mean.

Lag Operator Notation

$$\Phi(L) y_t = \mu + \Theta(L) \epsilon_t$$

Assumptions

- AR part is stationary: roots of $\Phi(L) = 0$ lie outside the unit circle.
- MA part is invertible: roots of $\Theta(L) = 0$ lie outside the unit circle.

ACF & PACF Patterns

- Both ACF and PACF taper off gradually.

ACF and PACF Behavior for Model Identification

6 Tests on Sectoral GDP Data

The dataset covers Spain, from 1994 to 2023.

Key variables include:

- Total and sectoral GDP

Model Type	ACF Behavior	PACF Behavior	Model Selection Hint
AR(p)	Tails off gradually	Cuts off after lag p	Use AR model if PACF cuts off
MA(q)	Cuts off after lag q	Tails off gradually	Use MA model if ACF cuts off
ARMA(p,q)	Tails off gradually	Tails off gradually	Use ARMA model if both ACF and PACF tail off

- Investments by sector

Focus variable: Second Sector Investment

Below is the time series graph of Second Sector Investment in Spain:

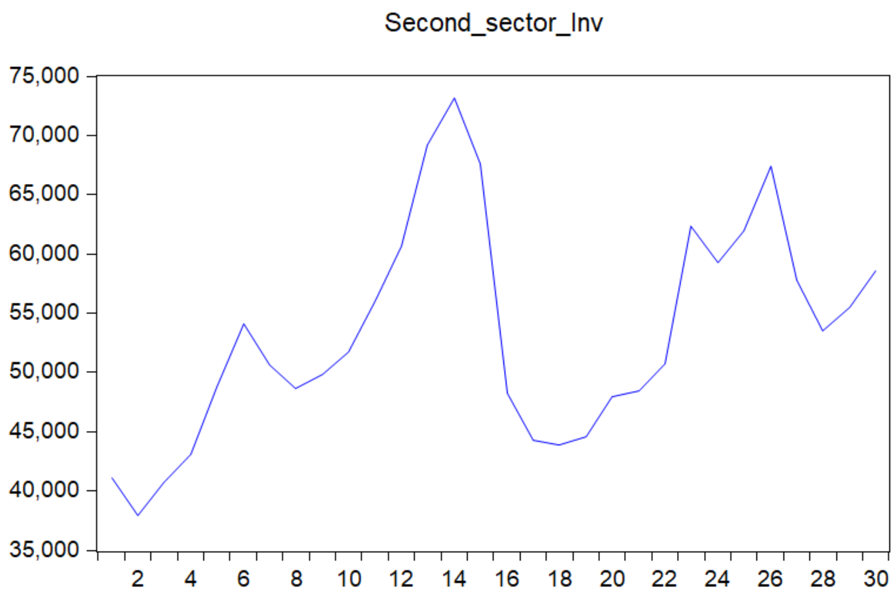


Figure 1: Time series of Second Sector Investment in Spain

The series shows a stationary trend, making it suitable for the following models:

- AR(p) – Autoregressive model
- MA(q) – Moving Average model
- ARMA(p,q) – Autoregressive Moving Average model

We will test these three models to determine which best fits the data.

Model Specificationt

To specify the most suitable model for the Second Sector Investment time series, we examined its autocorrelation structure using plots of the ACF and PACF.

The autocorrelation function (ACF):

- has a substantial positive value at lag 1,
- then tails off at subsequent lags.

The partial autocorrelation function (PACF):

- has a pronounced, significant spike at lag 1,
- while the data cuts off after the first lag.

Both of the patterns are consistent with the characteristics of an autoregressive (AR) process, in which there is:

- ACF tails,
- PACF cuts off after lag 1.

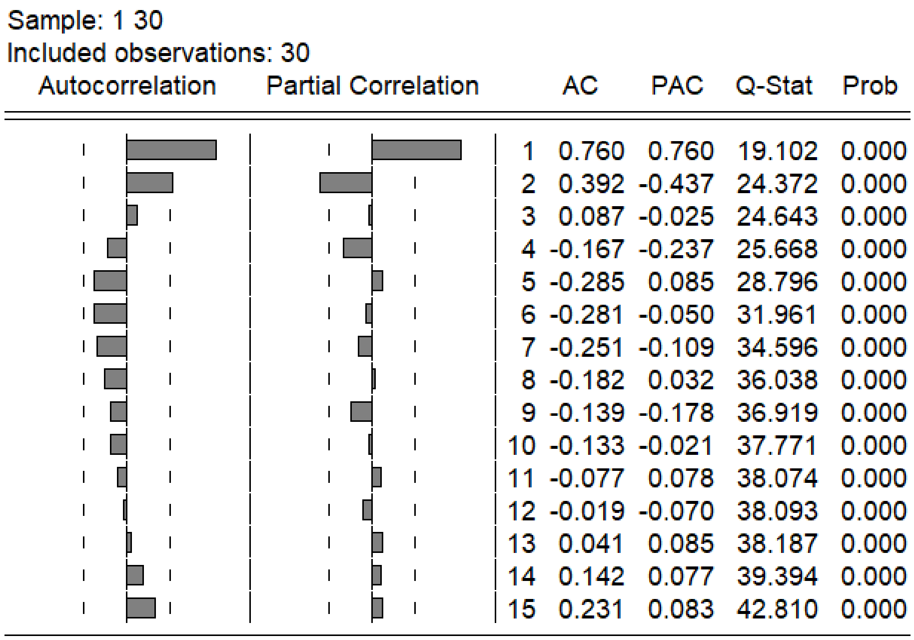
This type of behaviour indicates that an AR process is likely able to capture the process represented in the data effectively.

With these results in mind, we can model the series using an AR model, which is indicative of the sample's statistical characteristics.

AR(1) Model Estimation Results

We have estimated an AR(1) model for Spain's Second Sector Investment. The evidence provides a strong statistical and practical case for the AR model.

The basic results are presented below:



Sample: 1 30
Included observations: 30

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.760	0.760	19.102	0.000
		2	0.392	-0.437	24.372	0.000
		3	0.087	-0.025	24.643	0.000
		4	-0.167	-0.237	25.668	0.000
		5	-0.285	0.085	28.796	0.000
		6	-0.281	-0.050	31.961	0.000
		7	-0.251	-0.109	34.596	0.000
		8	-0.182	0.032	36.038	0.000
		9	-0.139	-0.178	36.919	0.000
		10	-0.133	-0.021	37.771	0.000
		11	-0.077	0.078	38.074	0.000
		12	-0.019	-0.070	38.093	0.000
		13	0.041	0.085	38.187	0.000
		14	0.142	0.077	39.394	0.000
		15	0.231	0.083	42.810	0.000

Figure 2: Autocorrelation and Partial Autocorrelation plots for Second Sector Investment

Coefficients

- AR(1) coefficient = 0.7911; standard error = 0.1178; t-statistic = 6.714
  - The p-value = 0.0000 (the coefficient is statistically significant at a high level).
  - Strong autocorrelation is observed in the time series: current levels of investment depend predominantly upon last year's level of investment.
- Intercept (C) = 52558.40; t-statistic = 11.5088; p-value = 0.0000
  - The base level of second-sector investment when there is a lagged term of 0 is statistically significantly different from 0 and forms a meaningful part of the model.

Goodness-of-Fit

- R-squared = 0.6222
  - Approximately 62.2% of the variation in second-sector investment could be explained by the model.

    - This represents a high degree of explanatory power for the univariate AR(1) model and validates the model as adequate to characterize the underlying dynamics of the series.

Model Stability

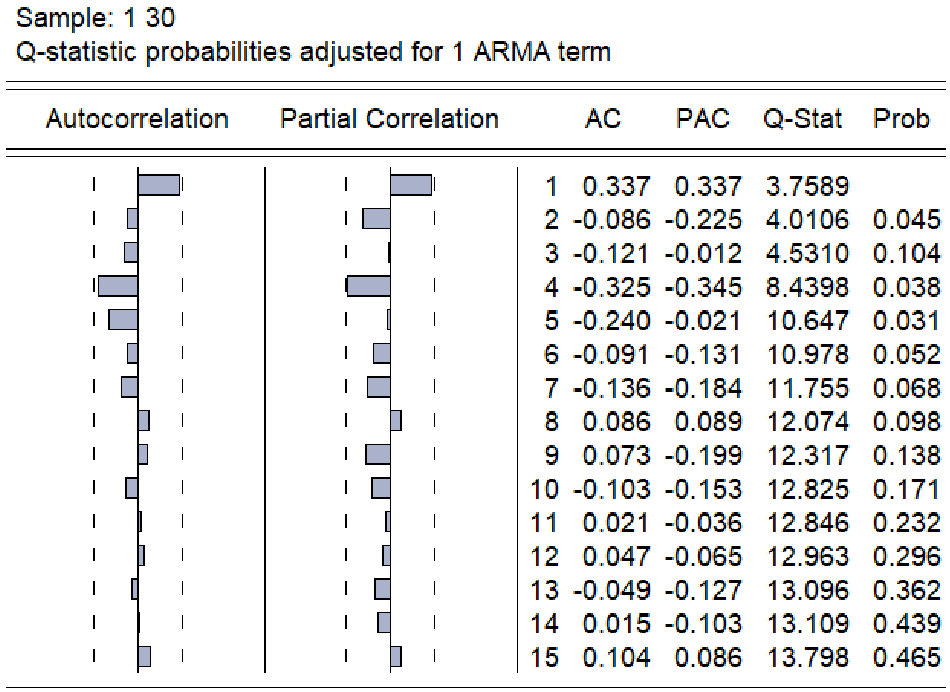
- Inverted AR Root = 0.79
    - Since the absolute value is less than 1, this process is stationary.
    - This meets one of the critical assumptions for AR models and makes longer-horizon forecasts trivial.

Residual Diagnostics of AR(1) Model

In order to examine the AR(1) model's adequacy of fit for Spain's Second Sector Investment data, a Ljung-Box Q-statistic test for the residuals was performed up to lag 15. The purpose of the test is to show whether residuals are autocorrelated enough to violate a model's assumption of white noise if the model is not well specified.

- At lag 2, the Q-statistic = 4.01 and the p-value = 0.045, which shows a statistically significant autocorrelation at the 5% level.
- At lags 4 and 5, p-values = 0.038 and 0.031, respectively — again statistically significant, and also in lag 3.
- Beginning at lag six onward, most of the p-values are greater than 0.05, which signifies decreasing evidence of autocorrelation present at lagged residuals.
- The autocorrelation (AC) values at those lags support the aforementioned findings (e.g., 0.325 at lag 4, 0.240 at lag 5).
- The partial autocorrelations (PAC) suggest there may be an additional structure that is not accommodated in the AR(1) specification.

These results highlight that the residual correlations are not negligible for the early lags only.



Sample: 1 30
Q-statistic probabilities adjusted for 1 ARMA term

	AC	PAC	Q-Stat	Prob
1	0.337	0.337	3.7589	
2	-0.086	-0.225	4.0106	0.045
3	-0.121	-0.012	4.5310	0.104
4	-0.325	-0.345	8.4398	0.038
5	-0.240	-0.021	10.647	0.031
6	-0.091	-0.131	10.978	0.052
7	-0.136	-0.184	11.755	0.068
8	0.086	0.089	12.074	0.098
9	0.073	-0.199	12.317	0.138
10	-0.103	-0.153	12.825	0.171
11	0.021	-0.036	12.846	0.232
12	0.047	-0.065	12.963	0.296
13	-0.049	-0.127	13.096	0.362
14	0.015	-0.103	13.109	0.439
15	0.104	0.086	13.798	0.465

Figure 3: Ljung-Box Q-statistics, autocorrelations, and PACF for residuals of AR(1) model

7 Forecasting Methods: Rolling vs. Recursive

For predicting time series, there are two popular methodologies for model affirmation and training: **rolling forecasting** and **recursive forecasting**.

Rolling Window Forecasting

In the rolling or sliding procedure, the model is fit on a fixed-size window of historical data considered relevant to the forecast. On each occasion of getting a new observation, the window advances one step by dropping the oldest observation and adding the newest one. The method is designed to evaluate how well a model forecasts by relying solely on the most recent data.

Example:

- Train on periods 1–85 → Forecast period 86
- Train on periods 2–86 → Forecast period 87

- Etc.

The rolling window-based method works best when recent data is considered more useful than older data. However, if the chosen window is too short, many longer-term patterns may be lost.

Recursive Forecasting

Recursive (sometimes expanding-window) forecasting utilises an initial training set. The new observations are then added to this training set one by one without dropping any observations. This training set grows in size as each new prediction is made and observed. One-step-ahead prediction is carried out at each iteration.

Example:

- Train on periods 1–85 → Forecast period 86
- Train on periods 1–86 → Forecast period 87
- Etc.

A recursive methodology is preferable when stable conditions prevail and the former information remains useful, at least for a while. If the structure of the data changes, its effectiveness may be limited.

8 Conditional Forecasting

Given that the objective in conditional forecasting is the prediction of the next value of a series, one must use all information available at present. This includes the series itself and potentially other external or contextual variables.

The usual form of conditional forecasting is given as follows:

$$\hat{x}_{t+1} = E(x_{t+1} \mid \mathcal{I}_t)$$

Where:

- $\hat{x}_{t+1}$ is the forecast for the next period,
- $\mathcal{I}_t$ represents all information known at time t, including past observations, and includes additional information from external entities that may influence the variable of interest.

For example, in sales forecasting, it involves sales data up to the current time, with restrictions on any known events that might affect sales in the following time period, such as holidays or promotions.

9 Model Evaluation

Standard Error

In assessing the forecasting efficiency of the AR(1) model of Second Sector Investment, one first examines the standard error of the regression. This informs the user, on average, how far each observed value is from the corresponding predicted value.

Definition

The standard error is the square root of the residual variance:

$$SE = \sqrt{\frac{\sum(y_t - \hat{y}_t)^2}{n - k}}$$

Where:

- y_t is the actual observed value,
- $\hat{y}_t$ is the predicted value of the model,
- n is the number of observations,
- k is the number of estimated parameters, including the intercept.

Interpretation

- The lower the value for the standard error, the closer are the predictions of the model to the actual values on average.
- It is also a measure of how much the residuals spread or deviate from the regression line.
- It does not measure bias, but it is an essential measure of forecast accuracy in the sample.

The standard error is a preliminary measure ahead of RMSE, MAE, or out-of-sample forecast errors.

10 Diebold-Mariano (DM) Test

To compare the predictive accuracy of two competing time series models, we apply the Diebold-Mariano test, a formal statistical method for determining whether one model outperforms another in forecasting.

Step 1: Select Competing Models

Model 1: AR(1) — a simple autoregressive model:

$$\hat{y}_t^{(1)} = \phi_1 y_{t-1} + \varepsilon_t$$

Model 2: ARMA(2,2) — a more complex model incorporating two autoregressive and two moving average terms:

$$\hat{y}_t^{(2)} = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \varepsilon_t$$

Step 2: Calculate Forecast Errors

For each model, compute the forecast error:

$$e_t^{(1)} = y_t - \hat{y}_t^{(1)} \quad ; \quad e_t^{(2)} = y_t - \hat{y}_t^{(2)}$$

Step 3: Choose a Loss Function

A common choice is the squared error:

$$g(e_t) = e_t^2$$

This penalizes larger errors more heavily.

Step 4: Compute the Loss Differential

For each period:

$$d_t = g(e_t^{(1)}) - g(e_t^{(2)})$$

This represents how much better or worse Model 1 performs relative to Model 2 at time t.

Step 5: Average the Loss Differential

Compute the sample mean of the loss differential over H forecast points:

$$\bar{d} = \frac{1}{H} \sum_{t=1}^{H} d_t$$

A positive $\bar{d}$ suggests Model 2 outperforms Model 1 on average.

Step 6: Estimate the Variance of $\bar{d}$

If loss differentials are uncorrelated:

$$\mathrm{Var}(\bar{d}) \approx \frac{\mathrm{Var}(d_t)}{H}$$

Otherwise, autocorrelation should be accounted for using:

$$\text{Var}(\bar{d}) = \frac{1}{H}\left(\gamma_0 + 2\sum_{j=1}^{q}\gamma_j\right)$$

where γ_j is the autocovariance at lag j.

Step 7: Compute the DM Statistic

$$DM = \frac{\bar{d}}{\sqrt{\text{Var}(\bar{d})}}$$

This statistic follows an approximate standard normal distribution under the null hypothesis.

Step 8: Hypothesis Testing

- **Null Hypothesis** (H_0): Both models have equal predictive accuracy.
- **Decision Rule**:
  - If $|\text{DM}| > 1.96$, reject H_0 at the 5% significance level.
  - If $|\text{DM}| \leq 1.96$, accept H_0 — no significant difference.

11 Forecast Combination

By combining forecasts from various models, an attempt is made to reduce forecast error in general, allowing different models to utilize their respective strengths. Below is a mathematical framework that illustrates how forecast combinations and errors are constructed and optimized.

1. Combined Forecast Equation

The weighted average of two forecasts, f_1 and f_2, is:

$$f = \omega f_1 + (1-\omega) f_2$$

Where:

- $\omega \in [0, 1]$: Weight for forecast f_1
- f_1, f_2: Forecasts from two distinct models

2. Forecast Errors

Let y be the actual observed value:

$$e_1 = y - f_1, \quad e_2 = y - f_2, \quad e(\omega) = \omega e_1 + (1 - \omega)e_2$$

3. Variance of Forecast Error

The variance of the combined error $e(\omega)$ is:

$$\mathrm{Var}(e(\omega)) = \omega^2\sigma_1^2 + (1 - \omega)^2\sigma_2^2 + 2\omega(1 - \omega)\sigma_{12}$$

Where:

- σ_1^2, σ_2^2: Variance of forecast errors from models 1 and 2

4. Optimal Weight to Minimize Error Variance

The optimal weight ω^* that minimizes $\mathrm{Var}(e(\omega))$ is given by:

$$\omega^* = \frac{\sigma_2^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}}$$

This guarantees a minimum combined forecast variance, considering the variance of each constituent model and their mutual correlation in the process.

12 Forecast Visualization

Actual vs Forecast Line Plot

A comparison of actual values versus forecasted values corresponding to observations 25 to 30. This particular plot demonstrates how well the selected model performs in tracking and assessing forecast accuracy visually.

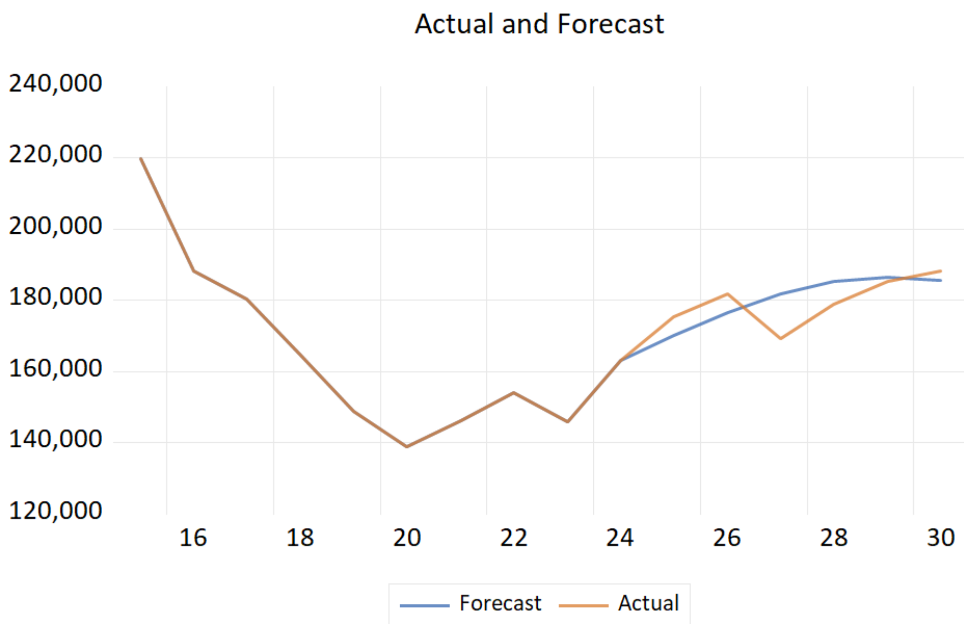


Figure 4: Actual vs Forecast Plot

Forecast Comparison Across ARMA Models

Forecast trajectories of various ARMA(p, q) models are drawn. The chosen specification, ARMA(2,1), is displayed with a thick red line for easy side-by-side comparison into robustness and smoothness.

ARMA(2,1): Model Output

Coefficients and Significance:

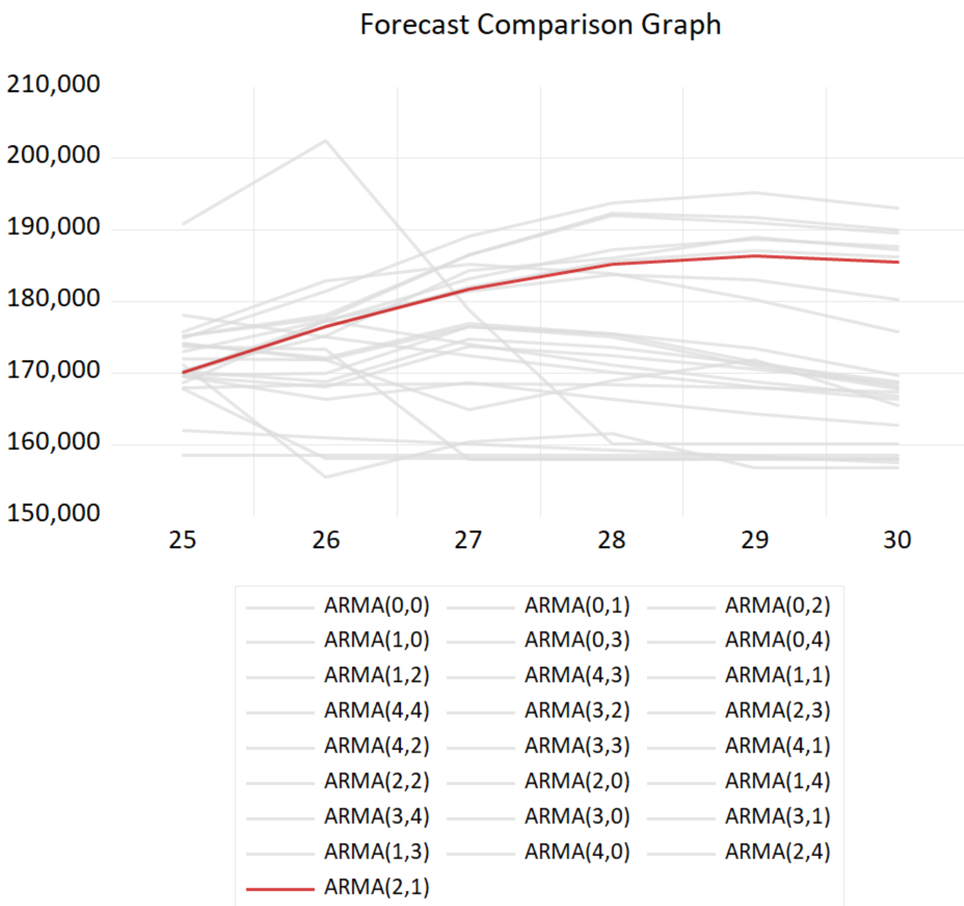


Figure 5: Forecast Comparison Across ARMA Models

- AR(1) = 1.848 $(p < 0.01)$
- AR(2) = -0.946 $(p < 0.01)$
- MA(1) = -0.645 (p = 0.0086)

All terms are statistically significant at a 1% level and hence contribute meaningfully to the model.

Basis of Model Fit:

- $R^2 = 0.945 \rightarrow$ About 94.5% of the variation in the log investment is explained by the model.
- Adjusted $R^2 = 0.9338 \rightarrow$ Adjusts for the number of predictors.

– Durbin-Watson statistic = 1.7965 → Acceptable degree of autocorrelation.

Information Criteria:

– AIC = -2.4946

– BIC = -2.2492

– HQ = -2.4295

Lower values indicate a better model fit.

Sample: 1 24
Included observations: 24
Convergence achieved after 12 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	12.01938	0.043632	275.4726	0.0000
AR(1)	1.848063	0.062255	29.68559	0.0000
AR(2)	-0.946405	0.063088	-15.00137	0.0000
MA(1)	-0.645171	0.220416	-2.927056	0.0086
SIGMASQ	0.002677	0.000967	2.768353	0.0122

R-squared	0.945378	Mean dependent var	11.97355
Adjusted R-squared	0.933878	S.D. dependent var	0.226137
S.E. of regression	0.058149	Akaike info criterion	-2.494650
Sum squared resid	0.064245	Schwarz criterion	-2.249222
Log likelihood	34.93580	Hannan-Quinn criter.	-2.429538
F-statistic	82.21104	Durbin-Watson stat	1.796510
Prob(F-statistic)	0.000000		

Inverted AR Roots	.92+.30i	.92-.30i
Inverted MA Roots	.65	

Figure 6: Estimation Output for ARMA(2,1)

Model Selection Comparison

A complete set of ARMA(p, q) models have been estimated and ranked by the AIC, BIC, and HQ. Model (2,1) has the lowest AIC value, thus representing the best trade-off in forecast accuracy between fit and parsimony.

Sample: 1 24
Included observations: 24

Model	LogL	AIC*	BIC	HQ
(2,1)	34.935799	-2.494650	-2.249222	-2.429538
(2,4)	37.863122	-2.488594	-2.095909	-2.384414
(4,0)	35.453541	-2.454462	-2.159948	-2.376327
(1,3)	34.953541	-2.412795	-2.118282	-2.334661
(3,1)	34.947312	-2.412276	-2.117763	-2.334141
(3,0)	33.937960	-2.411497	-2.166069	-2.346385
(3,4)	37.879367	-2.406614	-1.964844	-2.289412
(1,4)	35.827908	-2.402326	-2.058727	-2.311169
(2,0)	32.812813	-2.401068	-2.204725	-2.348978
(2,2)	34.688429	-2.390702	-2.096189	-2.312568
(4,1)	35.466235	-2.372186	-2.028587	-2.281029
(3,3)	36.434125	-2.369510	-1.976826	-2.265331
(4,2)	36.145206	-2.345434	-1.952749	-2.241254
(2,3)	35.094778	-2.341231	-1.997632	-2.250075
(3,2)	34.965848	-2.330487	-1.986888	-2.239330
(4,4)	37.876203	-2.323017	-1.832161	-2.192793
(1,1)	30.747911	-2.228993	-2.032650	-2.176903
(4,3)	35.206620	-2.183885	-1.742115	-2.066683
(1,2)	31.170250	-2.180854	-1.935426	-2.115742
(0,4)	30.943046	-2.078587	-1.784074	-2.000453
(0,3)	29.479664	-2.039972	-1.794544	-1.974860
(1,0)	27.078469	-2.006539	-1.859282	-1.967472
(0,2)	23.826714	-1.652226	-1.455884	-1.600136
(0,1)	15.834615	-1.069551	-0.922295	-1.030484
(0,0)	2.134964	-0.011247	0.086924	0.014798

Figure 7: Model Selection Table

A bar chart of the top 20 models, ranked by AIC, serves as an intuitive model comparison.

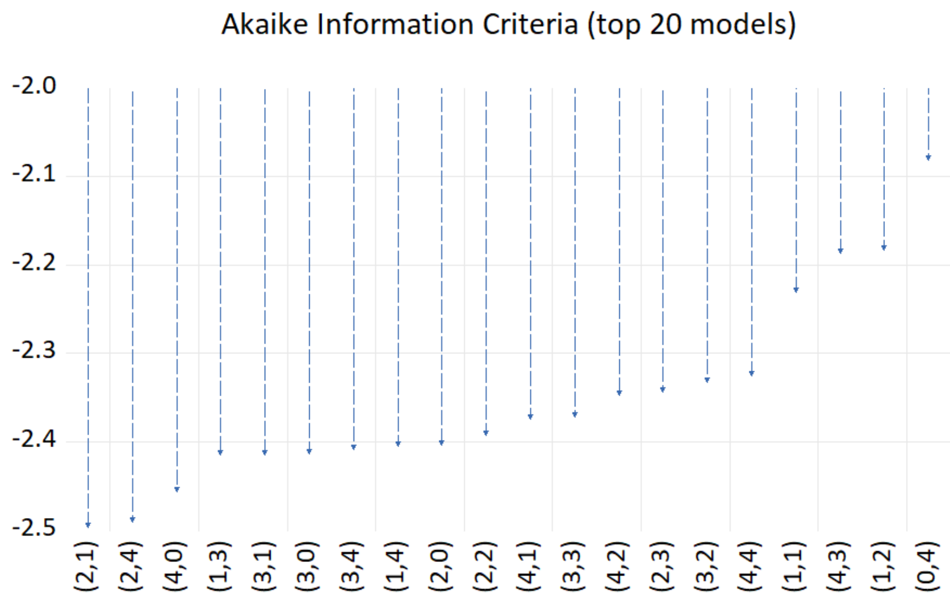


Figure 8: Top 20 ARMA Models Ranked by AIC

13 Forecast Improvement Through Model Combination

For these given criteria, ARMA(2,1), ARMA(2,4), and ARMA(4,0) were the best-performing models for assessing Spain's third-sector investment. Each of these models captures distinct features in the data. From the perspective of forecast combination, it is possible to obtain a more stable and accurate forecast.

Using model averaging techniques, such as:

– Equal weighting, or

– Inverse MSE (Mean Squared Error) weighting,

we can allow the models to complement each other and reduce the weaknesses of individual models.

However, to verify whether this combined forecast performs better, we can apply the Diebold-Mariano (DM) test. If large DM values are found, we can provide statistical evidence supporting the claim that the combined model performs better than each model individually, thereby supporting forecast combination in applied economic modelling.

14 Why Stationarity Matters

Stationarity is assumed by most classical time series models including AR, MA, ARMA, and ARIMA. If one applies such models to non-stationary data, parameter estimates will be biased, conclusions will be misleading, and forecasts will likely be inappropriate.

The autoregressive process AR(1) tries to explain the dynamic behavior of the observed series:

$$X_t = \mu + \phi X_{t-1} + u_t$$

- If $|\phi| < 1$, the process is stationary: the impact of shocks diminishes over time.
- If $\phi = 1$, the process follows a random walk: shocks have permanent effects, and the process is non-stationary.

Treatment of Non-Stationary Series

A non-stationary time series typically exhibits trend-based behavior. Two types of trends can be identified:

Deterministic Trend

For example:

$$X_t = \alpha + \beta t + u_t$$

To make the series stationary:

$$\text{Detrended series} = X_t - (\alpha + \beta t)$$

Stochastic Trend (Unit Root)

A process like:

$$X_t = X_{t-1} + \mu + u_t$$

can be made stationary via differencing:

$$\Delta X_t = X_t - X_{t-1} = \mu + u_t$$

For example, the stock price series X_t is usually non-stationary, whereas the return series ΔX_t is often stationary.

Impulse Response and Stationarity

Impulse response analysis indicates how a one-time shock evolves over time in a time series. In a stationary AR(1) process, a shock at time t diminishes with future observations.

For example, with $\phi = 0.7$:

Lag	**Effect**
0	1.0000
1	0.7000
2	0.4900
3	0.3430

This exponential decay highlights the temporary nature of shocks in stationary processes. In contrast, for non-stationary processes, shocks are permanent at the level of the series.

15 Stationarity Testing Methods

Before proceeding with time series modelling, it is essential to determine whether the data is stationary. If the series is not stationary, any forecasts or regressions can become unreliable—non-stationary data often leads to misleading results.

AR(1) Model and Unit Root

To start, consider the AR(1) (autoregressive of order 1) model:

$$Y_t = \phi Y_{t-1} + \varepsilon_t$$

Here, today's value depends on yesterday's value plus a random shock. If $\phi = 1$, the process becomes a random walk, meaning the effects of shocks never fade away—this is a classic case of non-stationarity.

To test for this, rewrite the model by subtracting Y_{t-1} from both sides:

$$\Delta Y_t = \delta Y_{t-1} + \varepsilon_t \quad \text{where } \delta = \phi - 1$$

Testing whether $\delta = 0$ is equivalent to checking if the original series has a unit root.

15.1 Augmented Dickey-Fuller (ADF) Test

This leads to the Dickey-Fuller test, which tests the null hypothesis:

$$H_0 : \delta = 0 \quad \text{(unit root, non-stationarity)} H_1 : \delta < 0 \quad \text{(stationarity)}$$

The test statistic is compared with critical values specific to this test.

However, real-world data often shows autocorrelation in residuals. To address this, the Augmented Dickey-Fuller (ADF) test expands the Dickey-Fuller test by including lagged differences:

$$\Delta Y_t = \mu + \beta t + \delta Y_{t-1} + \sum_{i=1}^{p} \alpha_i \Delta Y_{t-i} + \varepsilon_t$$

Where:

- μ: constant (drift),
- βt: deterministic time trend,
- δ: key parameter to test stationarity,
- α_i: coefficients on lagged differences,
- p: number of lags chosen to remove residual autocorrelation.

ADF Test Variants

There are three standard forms of the ADF test:

- Without constant or trend — if the data fluctuates around zero.
- With a constant — if the data fluctuates around a non-zero mean.
- With both constant and trend — if the series exhibits a linear trend.

Choosing the appropriate form depends on both visual inspection and statistical characteristics of the series.

Lag Selection Rule

Choosing the correct number of lags p helps ensure white noise residuals. A practical guideline comes from Schwert's rule:

$$p_{\max} = \left\lfloor 12 \left(\frac{T}{100} \right)^{1/4} \right\rfloor$$

For instance, if your sample size is $T = 40$, then $p_{\max} = 8$ lags.

Assessing stationarity using these tests is a crucial step before moving forward with modelling, as it directly impacts the reliability of your results.

15.2 ADF Test for Stationarity

One must conduct the ADF test to determine if the series of total GDP for Spain is stationary. It is a test for the presence of a unit root, indicating non-stationary behaviour in the data.

1. Quantity Against Null

- H_0: The series has a unit root.
- H_1: The series is stationary.

2. Resultant Important Findings

The value of the ADF test statistic was –1.9122.

Since this value is not more negative than the 10%, 5%, and 1% critical values, we fail to reject the null hypothesis H_0.

The p-value is 0.6225, which is very high compared to any conventional rejection level such as 0.05 or 0.10. Hence, the series **TOTAL_GDP** is non-stationary in the level form.

3. Supporting Output

The regression used for testing the hypothesis contained a constant and a trend. The coefficients estimated and the test results reject the null hypothesis of stationary behaviour and confirm that the series has a unit root.

Null Hypothesis: TOTAL_GDP has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=7)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-1.912155	0.6225
Test critical values:	1% level	-4.309824	
	5% level	-3.574244	
	10% level	-3.221728	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(TOTAL_GDP)
Method: Least Squares
Date: 05/29/25 Time: 17:51
Sample (adjusted): 2 30
Included observations: 29 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
TOTAL_GDP(-1)	-0.214991	0.112434	-1.912155	0.0669
C	183780.1	84244.69	2.181503	0.0384
@TREND("1")	2808.612	1876.226	1.496947	0.1464

R-squared	0.133621	Mean dependent var	17807.43
Adjusted R-squared	0.066976	S.D. dependent var	36079.44
S.E. of regression	34850.27	Akaike info criterion	23.85321
Sum squared resid	3.16E+10	Schwarz criterion	23.99465
Log likelihood	-342.8715	Hannan-Quinn criter.	23.89751
F-statistic	2.004973	Durbin-Watson stat	1.724658
Prob(F-statistic)	0.154955		

Figure 9: ADF test result output for `TOTAL_GDP` (Spain)

15.3 DF-GLS Test for Stationarity

To examine the stationarity of the `TOTAL_GDP` series for Spain more closely, we conducted the Elliott-Rothenberg-Stock DF-GLS test. This test incorporates a Generalised Least Squares detrending method prior to the Dickey-Fuller test and is recognised for its enhanced power in smaller samples.

1. Hypothesis Setup

- Null Hypothesis (H_0): The series contains a unit root (non-stationary).
- Alternative Hypothesis (H_1): The series is stationary.

2. Key Findings

The DF-GLS test statistic came in at -1.8199.

This value is not more negative than any of the critical values for the 10%, 5%, or 1% significance levels:

- 1% = –3.7700
- 5% = –3.1900
- 10% = –2.8900

The p-value is 0.0795, which means we do not have enough evidence to reject the null hypothesis.

As a result, the DF-GLS test indicates that the `TOTAL_GDP` series is non-stationary at level.

3. Supporting Output

The regression for the test was performed with a constant and trend, using GLS-detrended residuals. The coefficient on the lagged residual (–0.1979) was not statistically significant at traditional levels.

15.4 Phillips-Perron (PP) Test for Stationarity

To further assess whether Spain's `TOTAL_GDP` series is stationary, we performed the Phillips-Perron (PP) test. This test is practical because it accounts for possible autocorrelation and heteroskedasticity in the error term, all without introducing lagged difference terms, unlike the ADF test.

Test Details

- **Null Hypothesis:** `TOTAL_GDP` has a unit root (non-stationary)

Null Hypothesis: TOTAL_GDP has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=7)

		t-Statistic
Elliott-Rothenberg-Stock DF-GLS test statistic		-1.819919
Test critical values:	1% level	-3.770000
	5% level	-3.190000
	10% level	-2.890000

*Elliott-Rothenberg-Stock (1996, Table 1)
Warning: Test critical values calculated for 50 observations
and may not be accurate for a sample size of 29

DF-GLS Test Equation on GLS Detrended Residuals
Dependent Variable: D(GLSRESID)
Method: Least Squares
Date: 05/29/25 Time: 18:22
Sample (adjusted): 2 30
Included observations: 29 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GLSRESID(-1)	-0.197898	0.108740	-1.819919	0.0795

R-squared	0.103858	Mean dependent var	1642.200
Adjusted R-squared	0.103858	S.D. dependent var	36079.44
S.E. of regression	34154.51	Akaike info criterion	23.74905
Sum squared resid	3.27E+10	Schwarz criterion	23.79620
Log likelihood	-343.3612	Hannan-Quinn criter.	23.76382
Durbin-Watson stat	1.693058		

Figure 10: DF-GLS test result output for `TOTAL_GDP` (Spain)

- **Exogenous Terms:** Constant and linear trend
- **Bandwidth Selection:** Newey-West automatic with Bartlett kernel

Results Summary

- **PP Test Statistic:** -1.9193
- **MacKinnon One-Sided P-Value:** 0.6188

Critical Values (for comparison)

- **1% Level:** -4.3098
- **5% Level:** -3.5742
- **10% Level:** -3.2217

The PP test statistic is not more negative than any of the critical values at commonly used significance levels. Furthermore, the p-value is significantly higher than the 0.05

threshold. These findings lead us to **fail to reject the null hypothesis**, indicating that the GDP series is non-stationary.

Note

Residual variance has also been adjusted using HAC (Bartlett kernel), and details on the regression statistics can be found in the figure below.

Null Hypothesis: TOTAL_GDP has a unit root
Exogenous: Constant, Linear Trend
Bandwidth: 3 (Newey-West automatic) using Bartlett kernel

		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		-1.919337	0.6188
Test critical values:	1% level	-4.309824	
	5% level	-3.574244	
	10% level	-3.221728	

*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	1.09E+09
HAC corrected variance (Bartlett kernel)	1.10E+09

Phillips-Perron Test Equation
Dependent Variable: D(TOTAL_GDP)
Method: Least Squares
Date: 05/29/25 Time: 18:35
Sample (adjusted): 2 30
Included observations: 29 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
TOTAL_GDP(-1)	-0.214991	0.112434	-1.912155	0.0669
C	183780.1	84244.69	2.181503	0.0384
@TREND("1")	2808.612	1876.226	1.496947	0.1464

R-squared	0.133621	Mean dependent var	17807.43
Adjusted R-squared	0.066976	S.D. dependent var	36079.44
S.E. of regression	34850.27	Akaike info criterion	23.85321
Sum squared resid	3.16E+10	Schwarz criterion	23.99465
Log likelihood	-342.8715	Hannan-Quinn criter.	23.89751
F-statistic	2.004973	Durbin-Watson stat	1.724658
Prob(F-statistic)	0.154955		

Figure 11: Phillips-Perron test output for `TOTAL_GDP` (Spain)

16 Time Series Cointegration

Definition

Cointegration is defined by the presence of two or more non-stationary time series (often referred to as integrated of order one, i.e., $I(1)$) when there is also a stationary combination that exists. In particular, if:

$$X_t \sim I(1)$$

$$Y_t \sim I(1)$$

but

$$Z_t = Y_t - \beta X_t \sim I(0)$$

then X_t and Y_t are said to be cointegrated. This establishes a long-run equilibrium relationship between the two series without contradicting the stochastic trends (unit roots) that exist individually.

Economic Interpretation

In economics or finance, many indicators grow over time due to factors like inflation, technological progress, or other trends. However, some of these important variables move together in a consistent long-run ratio, reflecting underlying economic fundamentals.

Example: House Income and Selling Prices

Let:

- X_t: Average household income (non-stationary)

- Y_t: Average house prices (non-stationary)

Both variables are expected to increase over time. However, if housing remains proportionally affordable in the long term, we would expect:

$$Y_t - \beta X_t \sim I(0)$$

This indicates that the gap between house prices and income remains stable, suggesting cointegration. Economically, this implies that affordability is preserved on average over time.

17 Why Cointegration Can Be a Disaster for OLS Regression

Spurious Regressions When Cointegration is Absent

Ordinary least squares regression estimates a model of relationships between variables and yields results. However, when OLS is applied to non-stationary time series that are not cointegrated, the results can produce spurious regressions. This means that while estimated relationships appear statistically significant, they are actually misleading—variables may trend together over time, giving a false sense of connection rather than genuine economic significance.

Consider the regression:

$$Y_t = \alpha + \beta X_t + \varepsilon_t$$

When estimated on non-stationary data that are not cointegrated, we may observe:

- High R^2 values
- Statistically significant coefficients

- Very low p-values

Even though there is no true statistical relationship, this yields a false sense of interconnection between the two indicators. Time series with trend movements driven by separate factors—yet not truly correlated—will produce spurious results.

Technical Issues with OLS When Cointegration Exists

When OLS is applied to non-stationary series that are actually cointegrated—but without acknowledging this fact—the following technical problems arise:

1. Endogeneity The regressor X_t may be correlated with the error term ε_t:

$$E[\varepsilon_t \mid X_t] \neq 0$$

This violates the OLS assumption of exogeneity, resulting in biased and inconsistent estimators.

2. Serial Correlation in Errors For trending series, residuals tend to be autocorrelated:

$$\mathrm{Corr}(\varepsilon_t, \varepsilon_{t-1}) \neq 0$$

OLS assumes no serial correlation in errors. When this assumption is violated, standard errors and test statistics become unreliable, leading to incorrect inferences.

These issues highlight the need for proper cointegration testing and robust estimation methods, such as Dynamic OLS (DOLS), to avoid misleading conclusions and biased parameter estimates.

The Solution — Dynamic OLS (DOLS)

Dynamic OLS (DOLS) is one of the mainstream methods used to address endogeneity and serial correlation, which are often encountered when testing cointegrated time series. DOLS accomplishes this by incorporating leads and lags of the first difference of independent variables, while preserving the long-run equilibrium relationship.

DOLS Model Specification

The DOLS model can be written as:

$$Y_t = \alpha + \beta X_t + \sum_{j=-q}^{q} \gamma_j \Delta X_{t-j} + u_t$$

where:

- ΔX_{t-j} represents the first differences of X_t with multiple leads and lags.
- q indicates the appropriate number of leads and lags chosen.

Benefits and Corrections

This method corrects for the following:

- **Residual autocorrelation**: Ensures that there is no serial correlation in the error terms, which would otherwise bias the estimation.
- **Endogeneity**: Handles cases where X_t is correlated with the error term u_t, preventing biased and inconsistent parameter estimates.

Key Advantages

DOLS delivers consistent and efficient estimates of the long-run relationship coefficient β, and allows researchers to use standard t-tests and confidence intervals for valid inference.

This method has been frequently applied in economics and finance to model relationships such as consumption functions, term structures of interest rates, and purchasing power parity, making it an important tool for analyzing cointegrated time series.

18 Cointegration Tests

18.1 Engle-Granger Cointegration Test

The Engle-Granger method tests for a long-run equilibrium relationship even in cases where the series are individually non-stationary but may be cointegrated.

1. Theoretical Underpinning

Both series must be $I(1)$, meaning they are non-stationary, but their first differences are stationary. If a linear combination of these $I(1)$ series is stationary (i.e., $I(0)$), then they are cointegrated.

For example:

$$Y_t = \alpha + \beta X_t + u_t$$

If u_t is stationary, then Y_t and X_t are said to be cointegrated.

2. Testing Procedure

1. **Unit Root Test:** Apply a univariate test (e.g., ADF) on each series.
2. **OLS Regression on Levels:**

   $$Y_t = \alpha + \beta X_t + u_t$$

   Save the residuals $\hat{u}_t$.
3. **ADF Test on Residuals:** Fit the ADF test on $\hat{u}_t$.

3. Interpretation

- **Stationary Residuals:** Indicates that there exists a long-run relationship between the dependent and independent variables.

- **Non-Stationary Residuals (p-value $\geq$ 0.05):** No cointegration; the relationship is spurious.

This method allows us to distinguish between genuine and spurious regressions in time series analysis.

18.2 Phillips-Ouliaris Cointegration Test

The Phillips-Ouliaris approach tests for cointegration between non-stationary time series by examining the residuals of a cointegrating regression, similar in spirit to the Engle-Granger method but using different test statistics.

1. Theoretical Foundation

This test is designed for variables that are integrated of order one, $I(1)$. If a linear combination of these series is stationary, cointegration is present.

For two series, Y_t and X_t:

$$Y_t = \alpha + \beta X_t + u_t$$

If u_t is stationary (i.e., $I(0)$), then Y_t and X_t are cointegrated.

2. Testing Procedure

1. Confirm that both Y_t and X_t are $I(1)$ using unit root tests.
2. Estimate the cointegrating regression and obtain the residuals $\hat{u}_t$.
3. Apply the Phillips-Ouliaris test statistic (either the Phillips-Perron test statistic or the Z_t statistic) to $\hat{u}_t$.

3. Interpretation

- If the test statistic is significant, reject the null hypothesis of no cointegration, indicating a long-run relationship exists.

- If the test statistic is not significant, conclude that no cointegration is present.

This test complements the Engle-Granger approach and is often used in empirical work to validate long-run relationships in economic data.

19 Results of Engle-Granger Cointegration Test

The Engle-Granger Two-step Cointegration Test was applied to First Sector GDP and First Sector Investment in Spain to check if they share a long-run equilibrium relationship.

Graphical Representation

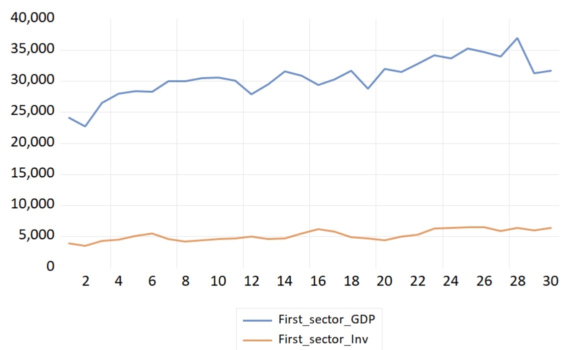


Figure 12: Time series of First Sector GDP and First Sector Investment in Spain

Both series appear to be increasing over time, which could indicate non-stationarity and potentially cointegration.

Test Setup

- **Sample:** 30 observations
- **Dependent Variables:** First Sector GDP and First Sector Investment
- **Cointegrating Relation:** With a constant and trend

– **Lag Selection:** Largest eigenvalue selected, lag length chosen using the Schwarz criterion (max lag = 6)

Test Results

Series: FIRST_SECTOR_GDP FIRST_SECTOR_INV
Sample: 1 30
Included observations: 30
Null hypothesis: Series are not cointegrated
Cointegrating equation deterministics: C @TREND
Automatic lags specification based on Schwarz criterion (maxlag=6)

Dependent	tau-statistic	Prob.*	z-statistic	Prob.*
FIRST_SECTOR_	-3.820894	0.0879	-20.92564	0.0551
FIRST_SECTOR_I	-3.818653	0.0898	-31.29383	0.0010

*MacKinnon (1996) p-values.

Intermediate Results:

	FIRST_SE	FIRST_SECTOR_INV
Rho - 1	-0.721574	-0.674381
Rho S.E.	0.188849	0.176602
Residual variance	2482813.	169452.4
Long-run residual variance	2482813.	465412.6
Number of lags	0	1
Number of observations	29	28
Number of stochastic trends**	2	2

**Number of stochastic trends in asymptotic distribution

Figure 13: Engle-Granger Test Output for First Sector GDP and First Sector Investment

Interpretation

Dependent variable (First Sector GDP):

– Tau-statistic p-value 0.0879 → Weak evidence of cointegration at the 10% level

– Z-statistic p-value 0.0551 → Marginal significance at the 5% level

Second Sector Investment as dependent variable:

– Tau-statistic p-value 0.0898 → Again indicates weak evidence of cointegration at the 10% level

– Z-statistic p-value 0.0010 → Strong evidence for cointegration at the 1% level

Even if the direction is inconsistent, one side of all pendulums indicates a potential long-term equilibrium relationship.

Implications for Additional Examinations

Based on this finding, the Engle-Granger test suggests a long-run relationship between First Sector GDP and First Sector Investment. This could motivate further analysis using models such as Dynamic OLS (DOLS) or Error Correction Models (ECM). Low-order stationarity and directional consistency suggest that these indicators may share a stable, long-run relationship.

20 Dynamic OLS (DOLS) Estimation

Following the potential cointegration between First Sector GDP and First Sector Investment, we applied the Dynamic OLS (DOLS) estimation to establish a single long-run relationship between the two. The DOLS method corrects for endogeneity and serial correlation by including leads and lags of the differenced independent variable.

Model Details

- **Dependent variable:** First Sector GDP
- **Independent variable:** FIRST_SECTOR_INV
- **Observations after adjustments:** 27
- **Leads and lags included:** One lead and one lag

Estimation Output

Interpretation

FIRST_SECTOR_INV Coefficient: 3.242

This means that First Sector Investment tends to increase First Sector GDP by an average of 3.242 units for every one-unit rise.

t-Statistic: 7.75

Dependent Variable: FIRST_SECTOR_GDP
Method: Dynamic Least Squares (DOLS)
Date: 05/30/25 Time: 19:17
Sample (adjusted): 3 29
Included observations: 27 after adjustments
Cointegrating equation deterministics: C
Fixed leads and lags specification (lead=1, lag=1)
Long-run variance estimate (Bartlett kernel, Newey-West fixed bandwidth = 3.0000)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
FIRST_SECTOR_INV	3.241984	0.418249	7.751325	0.0000
C	14084.90	2211.029	6.370292	0.0000

R-squared	0.730912	Mean dependent var	31066.63
Adjusted R-squared	0.681986	S.D. dependent var	2512.664
S.E. of regression	1416.959	Sum squared resid	44171013
Long-run variance	1974342.		

Figure 14: DOLS Estimation Output for First Sector GDP and First Sector Investment

This significant value indicates a strong and statistically significant relationship.

p-Value: 0.0000

This indicates that the relationship is highly unlikely to have occurred by chance.

Intercept (C): 14,085

This is also significant, representing the level of First Sector GDP when investment is zero.

Model Fit

- **R-squared:** 0.731

  Approximately 73.1% of the variation in First Sector GDP is accounted for by this model, indicating a strong relationship.

- **Adjusted R-squared:** 0.682

  Although adjusted for degrees of freedom, it still shows a strong relationship.

The DOLS results provide robust evidence that First Sector Investment drives First Sector GDP in the long run. The significant t-statistics and low p-values confirm that this relationship is statistically significant. This aligns with the observation that stable investment in this sector is crucial for long-term economic growth.

21 Seasonal Time Series Analysis

Definition

Seasonal time series refers to data that exhibit repeating patterns at regular intervals — for example, higher sales in December or peak tourist seasons during the summer. Explicitly accounting for seasonality helps in identifying underlying trends and improving forecasting accuracy.

Additive Seasonal Model

In most cases where seasonality is present, we model it using an additive seasonal model:

$$Y_t = \mu + \sum_{s=1}^{S-1} \delta_s D_{st} + Z_t$$

where:

- Y_t: Observation at time t
- μ: Global average (non-seasonal component)
- δ_s: Effect that season s has on the series
- D_{st}: Time-invariant seasonal dummy that equals 1 if time t is in season s and 0 otherwise
- Z_t: Residual component

Applying the Model

1. Create seasonal dummies for all periods except one (to avoid overlap). For example, with quarterly data, create dummies for Q1, Q2, and Q3; Q4 acts as the base period.

2. Use regression to estimate how each season affects the variable of interest.
3. Inspect the residuals: if they are stationary, it indicates that the model is effectively capturing seasonality.

22 Seasonal Trigonometric Models

Many seasonal cycles are not sharp or discrete, but gradual, such as temperature changes or energy demand ramping. Fourier decomposition or trigonometric seasonal models employ sine and cosine terms to capture smooth seasonal effects.

Model Structure

$$Y_t = \sum_{i=1}^{S/2} \left(\alpha_i \cos\left(\frac{2\pi it}{S}\right) + \beta_i \sin\left(\frac{2\pi it}{S}\right) \right) + Z_t$$

where:

- Y_t: Value being observed
- μ: Mean level
- α_i, β_i: Seasonal coefficients
- S: Number of seasons (e.g., 4 for quarterly data)
- Z_t: Random noise

Example in Quarterly Data (S=4)

Usually, one or two harmonic terms are sufficient:

- Seasonal wave pattern: captured by cosine terms
- Timing adjustment: captured by sine terms

Why Use It?

- Works well for smooth seasonal effects
- More computationally efficient than using N dummy variables
- Helps isolate the trend for more accurate forecasting

These terms enable effective modeling of time series seasonality, making it easier to interpret, analyze, and forecast.

23 Seasonal Unit Root Tests

Seasonal unit root tests provide an avenue for testing whether a time series has a unit root at seasonal frequencies (e.g., quarterly or monthly cycles). The presence of a seasonal unit root implies that the observed data contains a time-varying seasonal cycle, making effective prediction challenging.

Overview of Key Tests

Several important tests are commonly applied:

- **Dickey-Fuller (1984) Test:** Tests for the seasonal unit root at a specific frequency, such as quarterly data. The null hypothesis is that the seasonal unit root equals zero, implying that the seasonal effect is not constant over time.
- **Hegy Test (1990):** Decomposes the unit root process into regular and seasonal components. This test helps determine whether the series exhibits a seasonal unit root, a regular unit root, or both. It is more flexible than the Dickey–Hasza–Fuller method.
- **Canova and Hansen Test (1995):** Tests for seasonal trend stability, i.e., whether the seasonal pattern remains constant over time. The null hypothesis is that the seasonal pattern is stable, implying no seasonal unit root.

– **Hodrick–Prescott (HP) Test:** Similar to the Canova and Hansen test, but emphasizes testing whether the seasonal pattern is stable over time.

Implications of a Seasonal Unit Root

– **If present:** Seasonal differencing is needed to stationarize the data before further modeling.

– **If absent:** Seasonality can be modeled using dummies or trigonometric terms without differencing.

24 SARMA (Seasonal ARMA) Models

SARMA models are an extension of the classic ARMA setup for seasonal time series. They are particularly useful for series with repeating cycles, such as quarterly GDP, monthly temperatures, or annual sales peaks.

Model Notation

SARMA$(p, q)(P, Q)_s$ typically notates the model structure:

– p and q: Non-seasonal AR and MA components, capturing short-term dependencies.

– P and Q: Seasonal AR and MA orders.

– s: Seasonal period (e.g., 4 for quarterly, 12 for monthly).

Stationarity and Model Building

Unlike SARIMA, SARMA does not assume differencing is required, as it typically models stationary data with repeating seasonal cycles and tame trends. However, it is essential to check for stationarity before fitting a SARMA model. Use the ACF and KPSS tests to assess stationarity.

Once the data is confirmed to be stationary, interpret the ACF and PACF plots to determine the appropriate orders for each model component. After fitting the model, the residuals should be checked for randomness—this ensures that most of the time series' information has been captured.

25 SARIMA Models

The SARIMA (Seasonal AutoRegressive Integrated Moving Average) model extends the basic ARIMA framework to account for repeating seasonal patterns in time series data. SARIMA is ideal for modeling and predicting data with trends and seasonality, such as quarterly sales or monthly temperatures.

Notation

SARIMA$(p, d, q)(P, D, Q)_s$ is used to specify the model:

- p, d, q: Non-seasonal autoregressive order (p), differencing of degree (d), and moving average order (q).
- P, D, Q: Seasonal autoregressive order (P), seasonal differencing (D), and seasonal moving average order (Q).
- s: Seasonal period (e.g., 4 for quarterly data, 12 for monthly data).

This notation allows combining short-term dynamics with seasonal oscillations. Differencing (both regular and seasonal) helps convert non-stationary data into a stationary series appropriate for modeling.

Application

SARIMA models are widely used in economic and financial analysis to address:

- Data with a gradual trend over time.

– Recurring seasonality with a fixed period (e.g., monthly or quarterly).

– Complex interactions involving both trend and seasonal components.

In practice, the specification of (p, d, q, P, D, Q) is typically guided by:

– Examining the autocorrelation and partial autocorrelation functions (ACF and PACF).

– Model selection criteria such as the Akaike Information Criterion (AIC) or the Bayesian Information Criterion (BIC).

26 STL Decomposition

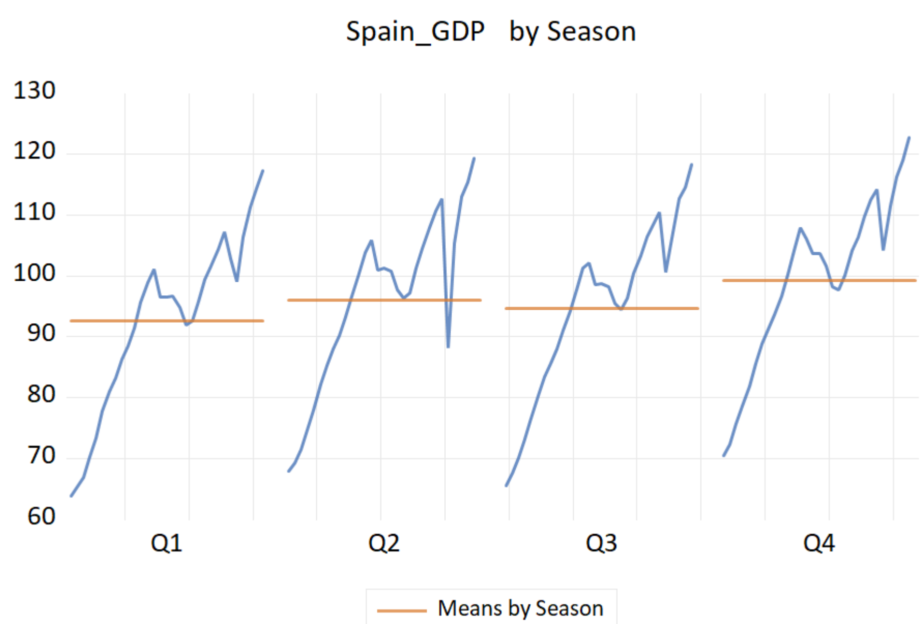


Figure 15: Spain's GDP by Quarter with Seasonal Means

This section presents a single comprehensive figure for analyzing the STL (Seasonal-Trend decomposition using Loess) of Spain's quarterly GDP from 1995-Q1 to 2025-Q1, all in one graph. STL decomposition elucidates the fundamental principles of time series by breaking it down into three interpretable components: Trend, Seasonality, and Remainder, enabling the identification of underlying driving forces.

Method Overview

- **Trend (T):** Captures the slow, long-term change in real GDP over many years.
- **Seasonality (S):** Represents the recurring quarterly fluctuations within each year.
- **Remainder (R):** Residual variations, including shocks and noise.

STL is highly flexible and relatively robust, making it an ideal candidate for the quarterly economic dataset, where seasonality may evolve smoothly.

Graphical Decomposition Results

Top row: The original series, showing Spain's GDP with an upward trend and clear seasonal pattern.

Trend component: Displays persistent growth and key economic recessions, including the 2008 financial crisis and the 2020 pandemic.

Seasonal component: Highlights the stable seasonal fluctuations observed in each quarter over the years.

Remainder component: Captures all unexpected shocks, such as the large dip associated with the COVID-19 shock.

Bottom row: The seasonally adjusted series (GDP minus seasonal) reveals the underlying economic trajectory.

To sum up this decomposition provides a comprehensive view of the structural evolution of Spain's economic growth, including seasonal GDP cycles and unusual events. This insight is essential for improving forecasts and informing policy planning.

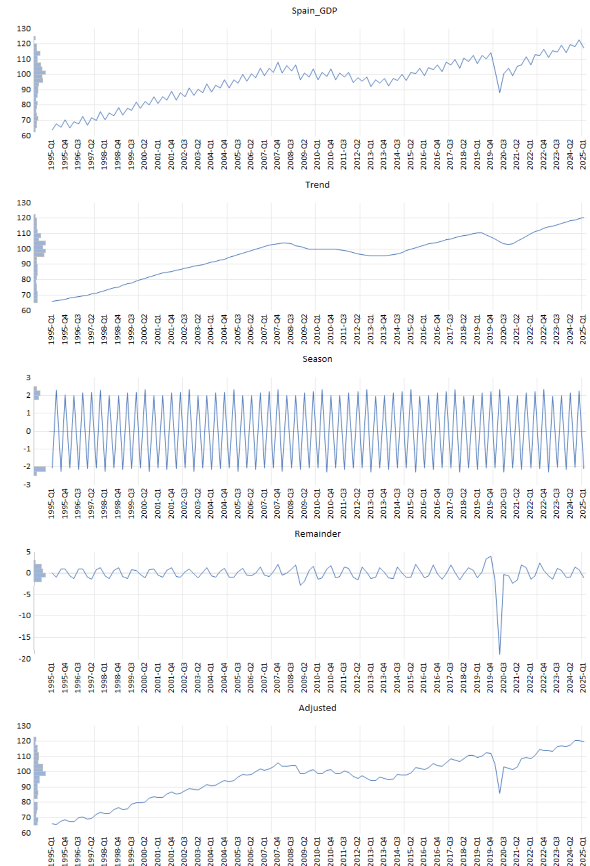


Figure 16: STL Decomposition of Spain's Quarterly GDP: Original Series, Trend, Seasonal Component, Remainder, and Seasonally Adjusted Series

27 HEGY Test for Seasonal Unit Roots

This section presents the application of the HEGY (Hylleberg, Engle, Granger, and Yoo) test to Spain's quarterly GDP data. The HEGY test is a specialized method to detect seasonal unit roots — persistent, non-stationary patterns that repeat at seasonal frequencies (e.g., quarterly or monthly).

Purpose of the HEGY Test

The test determines whether Spain's GDP requires seasonal differencing to achieve stationarity. If seasonal unit roots are present, seasonal differencing is necessary; otherwise, standard modeling without differencing is appropriate.

Methodology Overview

The HEGY test involves regressing the quarterly differenced GDP on constructed lag terms that correspond to different seasonal frequencies. These terms include:

- A non-seasonal unit root component (annual trend)
- Seasonal unit root components (at semi-annual and quarterly frequencies)

Test Results

The output table from the HEGY regression is provided above. It contains the estimated coefficients, standard errors, t-statistics, and p-values for each component tested.

Interpretation of Key Findings

- **Non-Seasonal Unit Root (C(1)):** The coefficient has a t-statistic of approximately -2.13 and a p-value of 0.0351, indicating statistical significance at the 5% level. Therefore, we reject the null hypothesis of a non-seasonal unit root — Spain's GDP is stationary at the annual level.
- **Seasonal Unit Roots (C(2), C(3), C(4)):** All coefficients have highly significant p-values (all ¡ 0.05). This means there is no evidence of seasonal unit roots at quarterly frequencies — seasonal shocks do not persist over time.
- **Model Fit:** The regression shows an R^2 of approximately 0.66, indicating a good fit.

Overall, the HEGY test results suggest that Spain's GDP is both trend-stationary and seasonally stationary. This implies that seasonal differencing is not required before modeling. Standard ARIMA or SARIMA models can be applied without seasonal differencing (i.e., $D = 0$). These results support reliable forecasts without the risk of over-differencing.

Dependent Variable: D(SPAIN_GDP,0,4)
Method: Least Squares (Gauss-Newton / Marquardt steps)
Date: 05/31/25 Time: 23:23
Sample (adjusted): 1996Q1 2025Q1
Included observations: 117 after adjustments
D(SPAIN_GDP,0,4)=C(1)*S1-C(2)*S2+C(3)*S3-C(4)*S4+C(5)+C(6)
(@SEAS(1)-0.25)+C(7)(@SEAS(2)-0.25)+C(8)*(@SEAS(3)-0.25)
+C(9)*@TREND

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	-0.022372	0.010483	-2.134186	0.0351
C(2)	-0.413846	0.070325	-5.884797	0.0000
C(3)	0.330496	0.080896	4.085447	0.0001
C(4)	-0.601894	0.080663	-7.461817	0.0000
C(5)	7.387931	3.136936	2.355143	0.0203
C(6)	-9.336796	1.247926	-7.481849	0.0000
C(7)	-1.629903	0.852865	-1.911092	0.0586
C(8)	-5.372959	1.390457	-3.864167	0.0002
C(9)	0.028264	0.016104	1.755081	0.0821

R-squared	0.661469	Mean dependent var	1.796077
Adjusted R-squared	0.636392	S.D. dependent var	4.022222
S.E. of regression	2.425395	Akaike info criterion	4.683670
Sum squared resid	635.3147	Schwarz criterion	4.896145
Log likelihood	-264.9947	Hannan-Quinn criter.	4.769932
F-statistic	26.37814	Durbin-Watson stat	2.007744
Prob(F-statistic)	0.000000		

Figure 17: HEGY Test Output for Spain's Quarterly GDP

28 Automatic ARIMA Selection

This section presents the automatic ARIMA model selection results using EViews for Spain's quarterly GDP data.

Purpose of Automatic ARIMA Selection

The ARIMA model is a powerful approach for capturing trend and seasonal dynamics within time series data. The EViews selection algorithm tests multiple combinations of ARIMA models and automatically selects the best-fitting model.

Reason to Use Auto ARIMA Selection

Automatic ARIMA selection tests:

- Unseasonal autoregressive (AR) and moving average (MA) lags.
- Seasonal AR terms (quarterly lags) and seasonal MA terms.
- Differencing needed to achieve stationarity.

It chooses the model based on information criteria, such as the Akaike Information Criterion (AIC), to balance model fit and complexity.

Summary of the Selected Model

Recommended Model by EViews:

- ARIMA$(2, 1, 1)(4, 0, 4)[4]$
- Non-seasonal components: AR(2) and MA(1)
- Quarterly seasonal components: SAR(4) and SMA(4)
- Differencing order: $d = 1$ to address non-stationary trend.

Model Estimation Output

Highlights in the Output

- **AR(1):** $p < 0.01$ — highly significant, capturing short-term autocorrelation.

Dependent variable: D(SPAIN_GDP,1)
Method: ARMA Maximum Likelihood (BFGS)
Date: 05/31/25 Time: 23:42
Sample: 1997Q2 2025Q1
Included observations: 112
Convergence achieved after 149 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.404685	0.295182	1.370969	0.1733
AR(1)	0.729281	0.151136	4.825319	0.0000
AR(2)	0.201843	0.127114	1.587895	0.1153
SAR(4)	0.998596	0.005179	192.8275	0.0000
MA(1)	-1.000000	15.68805	-0.063743	0.9493
SMA(4)	-0.917326	0.129284	-7.095411	0.0000
SIGMASQ	6.232209	3.225507	1.932164	0.0560

R-squared	0.743825	Mean dependent var	0.449393
Adjusted R-squared	0.729186	S.D. dependent var	4.954504
S.E. of regression	2.578311	Akaike info criterion	4.869401
Sum squared resid	698.0074	Schwarz criterion	5.039307
Log likelihood	-265.6865	Hannan-Quinn criter.	4.938338
F-statistic	50.81268	Durbin-Watson stat	1.980506
Prob(F-statistic)	0.000000		

Inverted AR Roots	1.00	.94	-.00+1.00i	-.00-1.00i
	-.21	-1.00		
Inverted MA Roots	1.00	.98	-.00+.98i	-.00-.98i
	-.98			

Figure 18: Automatic ARIMA Model Selection Output for Spain's GDP

- **AR(2):** $p \approx 0.1153$ — marginally insignificant but included to capture potential longer-term lag effects.
- **MA(1):** $p \approx 0.9493$ — not statistically significant, suggesting weaker short-term shock impact.
- **Seasonal SMA(4):** Almost significant at $p < 0.01$ — indicates presence of quarterly seasonality.

Model Fit Statistics

- $R^2 = 0.7438$ — the model explains approximately 74% of the variance in the differenced series.
- Adjusted R^2 — adjusts for model complexity.
- Durbin-Watson statistic: approximately 1.98 — suggests residuals are not autocorrelated.
- Information Criteria — AIC and Schwarz criteria for model fit assessment.

Diagnostic Summary

- All inverted AR and MA roots lie within the unit circle — indicating a stable, well-behaving model.
- Residual diagnostics (to be examined separately) should confirm white noise behavior and absence of autocorrelation.

Conclusion

The automatic ARIMA selection suggests that Spain's GDP requires a SARIMA$(2, 1, 1)(4, 0, 4)[4]$ model. This model accounts for both short-term dependencies and robust quarterly seasonal patterns — an excellent starting point for reliable forecasting.