

# Sefirot

Financial Research

## Econometric indicators

**Authors:**  
Mallia Corrado  
Croci Mattia

---

May, 2025

# Index

1	Introduction	1
2	Predictability of Returns	1
3	The CAPM	2
4	ARMA Representation and Stocks	3
5	Heteroskedasticity (GARCH)	4

# 1 Introduction

Within this paper we report an in-depth study of the calculation of econometric indicators on selected stocks. In this case we were asked to identify three of the largest Italian listed companies and based on their stock performance calculate the required econometric parameters.

In the first part of the paper we focused on the predictability of returns while the second part will be vertical on the application from CAPM (Capital Asset Pricing Models). Then the representation of values with the ARMA (Autoregressive moving-average model) will be deepened while in the last part insights will be made on the GARCH heteroschedasticity (Generalized AutoRegressive Conditional Heteroskedasticity) of the returns obtained from the selected stocks.

[Click here to find the database and matlab code.](#)

The analysis of our research takes as a sample of study the ones of the biggest stocks in the Italian market: Generali (GNR), Enel (ENE), Stellantis (STL).

## 2 Predictability of Returns

Compare the simple and logarithmic returns on the whole period with the sum of the daily returns.

From the close price series, we calculate the simple and logarithmic return on the whole period and we compare it with the sum of the daily returns. The simple returns on the whole period that we obtain is different from the sum of the daily simple returns, but the log-return on the whole period is equal to the sum of the daily logreturns. This is verified for all the three stocks.

For example, the sum of the log-returns of ENE stock is -0.4550 and its log-return on the whole period is -0.4550, but the simple return on the whole period is 0.5762.

Estimate the regression coefficients of the returns over their first and second lags. Discuss the results.

Analyzing the results obtained for Generali with the first lag we can observe that alpha and beta are not statistically significant, as they have a p-value greater than 0.05.

As for the results obtained with the second lag, alpha and beta are not statistically significant. In the second case the beta remained negative but very close to zero and in fact R2 is very low, suggesting that the model does not explain well the variability of y.

For Enel, in both models, the intercept is not statistically significant, while the betas are negative and statistically significant, indicating an inverse relationship with y. However, the R2 values are very low, therefore, despite the existence of a significant correlation, the effect is weak.

Analyzing the results obtained for Stellantis we can observe in the model with the first lag that the intercept (alpha) is statistically significant whereas the beta is not. The model with the second lag instead, reports an alpha and a beta statistically significant. R2 values are very low in both models.

In all stocks and models the beta shows a value that is very close to zero which means that the regression draws a line almost parallel to the x axis.

For the three stock series, working with daily arithmetic returns on the whole sample perform the Jarque-Bera test on the series of daily returns and discuss.

Jarque-Bera test measures the deviation from normality, while the p-value assesses the statistical significance of that deviation. In this case, the result is 1 in all the three

series of daily returns and it implies that the null hypothesis of normality is rejected. Rejecting the null hypothesis ( $h=1$ ) suggests that the data significantly deviate from a normal distribution.

For the three stock series, working with daily arithmetic returns on the whole sample compute the cross-sectional covariance matrix and discuss the results.

The cross-sectional covariance matrix is composed of the variances of the individual variables along the diagonal and the covariances between pairs of variables in the out-diagonal positions.

Correlation refers to a statistical relationship between two or more stocks. It measures the degree to which changes in one stock are associated with changes in another stock.

There is not any important correlation between the stocks because all the values are very close to 0. We can note the STL returns have a larger variance than the others. It means that there is more volatility in the returns of this stock as we can observe from the Figure 1. In carrying out this exercise we used the European Market Factor as a market proxy.

### 3 The CAPM

The standard Sharpe-Linter CAPM assumes the existence of a constant risk-free rate and is written as

$$E(r_i) = r_f + \beta_i(E(r_m - r_f)).$$

For each stock, estimate the betas on the Jan 2003 - Dec 2022 period, using the European Market Factor as a market proxy.

Stellantis beta is greater than 1 and it indicates that the activity is more volatile than the market.

Generali and Enel betas are lower than 1 so these stocks are considered less risky and generally offer more stability but potentially lower returns than the Stellantis.

A first method to test the CAPM is based on the following timeseries regression:

$$z_{i,t} = \alpha_i + \beta_i z_{m,t} + \epsilon_{i,t},$$

where  $z_{i,t} = r_{i,t} - r_{f,t}$ , and  $\epsilon_{i,t}$  is a zero mean noise. The null hypothesis  $H_0 : \alpha_i = 0$  can be tested for each asset separately using standard t-statistics.

Using the whole sample, perform the naïve test of the CAPM and check how well the market explains individual stock returns.

All the alphas are significant but have low values, very close to zero, which suggests that the securities have had similar returns to those predicted by the CAPM, once adjusted for market risk. These results might indicate that the factors considered by the CAPM model influence the returns of the securities.

All the estimated beta coefficients are statistically significant, as indicated by the low p-values associated with the beta estimates.

A second approach for testing the Sharpe-Linter CAPM is based on a two-stage procedure:

1. Estimate the betas  $\beta_i$  using the time-series dimension.
2. For each period, run the following cross-sectional regression:

$$z_{i,t} = \phi_{0,t} + \phi_{1,t} \hat{\beta}_i + u_{i,t}, \quad i = 1, \dots, N$$

where  $u_{i,t}$  is the noise with mean zero.

3. Estimate  $\phi_0$  and  $\phi_1$  as the average of the cross-sectional estimates.
4. If the CAPM is valid, we should have  $H_0 : \phi_0 = 0$  and  $\phi_1 > 0$ .

$\phi_1 > 0$  should correspond to the market risk premium.

In our model  $p_0$  is different from 0, so the CAPM may not be an appropriate model to explain the returns of stock in the sample. In fact, a  $p_0$  different from zero suggests that there is an abnormal average yield not explained by the model.

## 4 ARMA Representation and Stocks

Use the three time-series of stock prices. After Plot and comment the sample ACF and PACF for each stocks. Comment the result obtained using log-prices.

ACF measures the correlation between time series values to different time lags (delays). In other words, it shows how much a value in the time series is correlated with a value before a certain lag.

The PACF measures the correlation between time series values at different lags, eliminating the effect of intermediate lag correlations.

We calculated the Autocorrelation Function (ACF) with simple returns and log returns and as we can see from the graphs, the values are very close to zero for all lags. This indicates that there is no significant correlation between your time series values and their delayed values.

Looking at the PACF graph, we notice that also in this case the values are very close to zero for all the lags considered. This indicates that there is no significant correlation between time series values and their specific delays, once the effect of intermediate correlations has been eliminated.

You can quantify the preceding qualitative checks for correlation with a formal hypothesis test. Using the Ljung-Box-Pierce Q-test, test the presence of correlation in the returns for up to 20 lags at the 0.05 level of significance.

The Ljung-Box-Pierce Q-test was performed on the returns of the time series models for Generali, Enel and Stellantis. For Stellantis and Enel, the test rejected the null hypothesis ( $h=1$ ), indicating the presence of significant autocorrelation in the returns. Although the ACF and PACF values are close to zero, the Ljung-Box Q test found relevant autocorrelations in the time series. This may indicate that there are weak but persistent autocorrelations that add up to be significant, or that there are higher lag autocorrelations that are not evident in the ACF and PACF charts.

For Generali, the test did not reject the null hypothesis ( $h = 0$ ). In this case, the values of Ljung Box Pierce Q-test are consistent with those of the ACF and PACF close to zero and we can conclude that the time series does not show significant autocorrelation.

Estimate the coefficients of an AR(p) for each stock by OLS.

We use linear regression in order to model the residues of the time series of the chosen stocks using a series of delayed regressors.

For Generali and Enel all the coefficients AR obtained are all lower than zero but very close to zero and it indicates a weak association between the independent variable and the dependent variable.

On the contrary, Stellantis has the coefficients AR(1) and AR(2) positive but still close to zero and only AR(3) is slightly negative.

Check that the residuals of your estimated model are white noise with Ljung-Box test.

For Generali and Enel slightly negative AR coefficients and  $h=0$  in the Ljung-BoxPierce test suggest that the AR model is adequate to model a time series with a tendency to marginal decline and residues are uncorrelated, indicating that the model could be a good representation of the data.

Testing the Ljung-Box-Pierce test for Stellantis, it returned a value of  $h = 1$ , suggesting that there is still significant autocorrelation in the model residue. This indicates that the AR model does not fully capture the time series autocorrelation structure and could be improved for more accurate modelling.

## 5 Heteroskedasticity (GARCH)

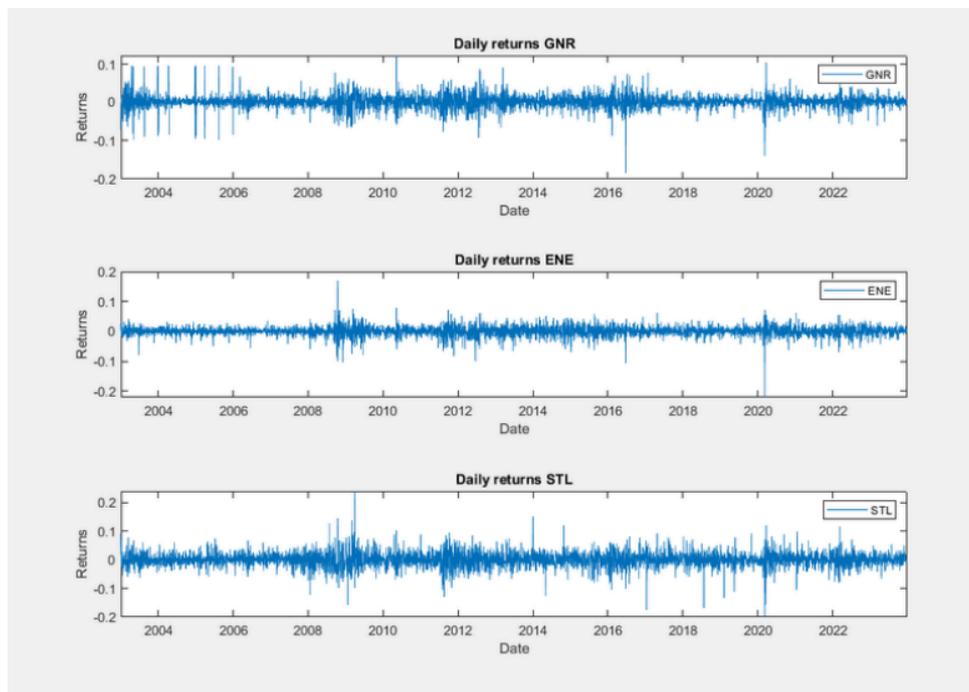


Figure 1: Plot the daily returns. Describe the series.

We gave a plot of the daily returns and observed the behaviour of the series at the time. There is volatility clustering: we can see periods of high volatility alternating with periods of low volatility. The code generates three plots, one for each series of daily returns (Generali, Enel and Stellantis datasets).

We can notice the tendency of volatility in financial markets to appear in bunches, because information, which drives price changes, occurs in bunches.

For example, these clusters are visible for the stocks selected in early 2020 due to the coronavirus pandemic or during the 2008 crisis.

We can also observe, as confirmed by empirical evidence, that Stellantis has a larger variance than the others.

The cause of this high volatility is to be attributed to the numerous merger and acquisition transactions during the period 2003-2023.

Estimate a GARCH(1,1) model.

Estimating the GARCH we quantified and modelled the changes of the volatility, we estimated that the GARCH model using the daily return series for each set, so that we

can observe the dynamics of the volatility in the data and then infer the conditional variances of the series. This approach is better than simply using an historical variance. Forecast the conditional variance of the return series 5 years into the future using the estimated GARCH model

Using the GARCH, we can also give a forecast of the conditional variances for the next 60 periods, and consequently we can estimate the volatility patterns for that future period. We forecast the next 5 years conditional variance and what we obtain is almost a constant line.