How Does the Correlation Between Small-Cap and Large-Cap Stocks Vary Across European Countries?

Daniele Parini[1*†], Noam Lechapelier[1*†], Noé Notter[1*†] and Leonardo Cassano[1*†]

[1]HEC Lausanne, Vaud, CH.

*Corresponding author(s). E-mail(s): daniele.parini@unil.ch; noam.lechapelier@unil.ch; noe.notter@unil.ch; leonardo.cassano@unil.ch;
†These authors contributed equally to this work.

Abstract

This study investigates the correlation between small-cap and large-cap stock indices across the European countries, exploring how economic environments and market conditions influence the dynamics between these two key market segments. Small-cap stocks, often linked to growth potential and higher volatility, contrast with large-cap stocks, which are associated with stability and lower risk, making their relationship critical for understanding market behavior, diversification, and economic interdependencies. Using linear regression models, the study examines country-specific relationships by modeling small-cap index returns as a function of large-cap index returns, providing insights into the role of factors such as GDP growth and inflation.

Europe's economic diversity offers a unique context for exploring these correlations. Differences in economic development and market maturity reveal significant variations in correlation strengths across countries, emphasizing the importance of macroeconomic. This research contributes to academic literature by offersing practical insights for investors aiming to optimize diversification strategies and understand financial market interdependencies.

Contents

Country	Large-Cap Index	Small-Cap Index
Austria	ATX Five	NASDAQ Austria Small Cap Index
Belgium	BEL 20	BEL Small
Denmark	OMX Copenhagen 25	OMX Copenhagen Small Cap Index
Finland	OMX Helsinki 25	OMX Helsinki Small Cap Index
France	CAC 40	CAC Small
Germany	DAX	SDAX
Greece	ATHEX Composite	NASDAQ Greece Small Cap Index
Ireland	ISEQ 20	ISEQ Small Cap Index
Italy	FTSE MIB	FTSE Italia Small Cap
Netherlands	AEX	AScX
Norway	OBX	Oslo Børs Small Cap Index
Portugal	PSI 20	NASDAQ Portugal Small Cap Index
Spain	IBEX 35	IBEX Small Cap
Sweden	OMX Stockholm 30	OMX Stockholm Small Cap Index
Switzerland	NASDAQ Switzerland Large Cap Index	NASDAQ Switzerland Small Cap Index
United Kingdom	FTSE 100	FTSE SmallCap Index
Poland	NASDAQ Poland Large Cap Index	NASDAQ Poland Small Cap Index

Fig. 1 Indices by country

1 Introduction

Financial markets serve as the backbone of modern economies, functioning as critical mechanisms for capital allocation, investment facilitation, and economic growth. Within these markets, small-cap and large-cap stocks represent two distinct segments with unique characteristics and roles. Large-cap stocks are typically associated with well-established, mature companies, offering stability and steady returns, making them appealing to conservative investors. Conversely, small-cap stocks are linked to younger, high-growth companies, characterized by higher volatility and risk, often attracting investors with greater risk tolerance. The dynamic relationship between these two market segments provides valuable insights into market behavior and serves as a focal point for investors, policymakers, and academics.

This study investigates how the relationship between small-cap and large-cap stock indices is influenced by economic environments and market structures across European countries. By analyzing the interplay between these market segments, the study aims to uncover how macroeconomic factors, such as GDP growth and inflation, impact market performance. Europe presents a compelling context for this analysis, as its member countries exhibit considerable economic diversity, driven by variations in fiscal policies, growth rates, and inflation levels. Exploring these differences allows for a deeper understanding of how country-specific economic conditions shape financial markets.

The relationship between economic indicators and market behavior has long been a subject of academic and practical interest. This study focuses on the role of key macroeconomic variables—GDP growth and inflation—in influencing stock returns, particularly in the context of large-cap and small-cap indices. Log returns of these indices serve as proxies for market performance, providing a robust measure for financial analysis. By examining these factors across a diverse set of European countries, this study seeks to identify both regional trends and country-specific drivers of market behavior. The integration of macroeconomic variables into the analysis ensures a comprehensive approach to understanding the forces shaping market dynamics.

The primary objective of this research is to examine how macroeconomic factors, such as GDP growth and inflation, influence the correlation between small-cap and large-cap stock indices. To achieve this, the study employs linear regression as the empirical framework, quantifying the relationship between these two market segments and assessing the impact of macroeconomic variables. Additionally, the project addresses key statistical challenges, such as multicollinearity and outliers, by employing robust statistical methods to ensure the validity and reliability of the results. To complement the regression analysis, Monte Carlo simulations are implemented to validate model assumptions, while advanced econometric techniques, such as GARCH models, are utilized to capture volatility patterns in stock returns.

Understanding the relationship between economic indicators and market behavior is crucial for various stakeholders. For policymakers, the findings of this study provide insights into how macroeconomic conditions, such as inflationary pressures or economic growth trends, influence financial markets. This knowledge is essential for developing policies that promote market stability and economic resilience. For investors, the analysis offers valuable guidance on portfolio diversification strategies, highlighting the extent to which small-cap and large-cap indices are integrated or segmented in different countries. Finally, for academics and researchers, this study contributes to the literature by addressing a gap in understanding the interplay between macroeconomic factors and stock market behavior within a regionally diverse context. By focusing on European markets, the study provides a nuanced perspective on how country-level economic conditions influence market dynamics.

The analysis leverages a regression-based framework to explore the relationship between small-cap and large-cap stock indices across European countries. Daily returns of small-cap indices are modeled as a function of large-cap returns and macroeconomic variables, such as inflation and GDP growth. This approach allows the study to quantify the strength and direction of the relationship between these two market segments, offering insights into market integration and the broader economic factors shaping financial markets. By systematically comparing regression results across countries, the study identifies patterns of market behavior, regional disparities, and potential drivers of economic resilience. This introduction, written at a master's thesis level, maintains a professional tone and logical flow while integrating the provided objectives, background, and relevance. Let me know if further refinements are needed.

2 Presentation of the Dataset

2.1 Data Source

Our dataset is composed of historical daily prices for small-cap and large-cap indices across various European countries, sourced from Investing.com, Eurostat, and OECD. These indices represent distinct segments of the stock market, capturing the performance of smaller, high-growth companies (small-cap) and larger, more established firms (large-cap). In addition to financial data, the dataset incorporates key macroeconomic variables, such as monthly inflation rates and quarterly GDP growth, which provide valuable insights into how economic factors influence market behavior.

The dataset includes information on country names to contextualize the observations within specific national economic environments. It also contains monthly timestamps, converted to a consistent datetime format to facilitate temporal analysis. Logarithmic returns of large-cap indices serve as key financial indicators of market performance, while Gross Domestic Product (GDP) values and inflation rates reflect macroeconomic conditions. These variables collectively form the foundation for analyzing the relationship between market segments and economic dynamics.

2.2 Data Cleaning

To ensure data accuracy and comparability, the dataset underwent a rigorous cleaning process. The Date column was standardized to datetime format, and any invalid or inconsistent entries were flagged and removed. Observations with large-cap log returns exceeding ± 2.5 standard deviations were excluded to mitigate the influence of outliers that could distort the analysis. Continuous variables, including GDP and inflation, were standardized to eliminate disparities in scale, enabling meaningful cross-country and temporal comparisons. Additionally, a composite economic indicator combining GDP and inflation was created to enhance the depth of the analysis and capture the combined impact of these factors on financial markets.

For each country, the financial data includes daily open, high, low, and closing prices for small-cap and large-cap indices. Logarithmic returns were calculated to ensure comparability across indices with varying scales and facilitate regression analysis. However, differences in data availability required restricting the analysis to periods where small-cap and large-cap data overlap. For instance, in Greece and Italy, where small-cap indices became available only in 2015 and 2012 respectively, the dataset was truncated to align timeframes. This ensures consistency across comparisons, even if it limits the temporal coverage for some countries.

2.3 Final Dataset

After cleaning and processing, the final dataset comprises 3,856 observations spanning multiple European countries and time periods. It encompasses a range of historical events, including the 2008 financial crisis, the COVID-19 pandemic, and Brexit. These events add layers of complexity, allowing for an exploration of how market dynamics

evolve in response to external shocks. Despite these complexities, the primary focus remains on understanding the interplay between small-cap and large-cap indices within varying economic contexts.

One of the dataset's key strengths is its diversity. It spans countries with distinct economic profiles, such as Germany and France, which provide long data histories covering pre-2008 financial crisis periods, and countries like Austria and Norway, which offer insights into post-2010 trends. This temporal and geographic breadth enables the study to analyze both short-term and long-term dynamics, providing a nuanced understanding of the factors shaping the relationship between small-cap and large-cap indices in Europe.

Through comprehensive data preparation and integration of macroeconomic variables, this dataset provides a solid foundation for empirical analysis, enabling robust insights into the economic and market forces that influence stock market behavior.

3 Presentation of the Empirical Methodology

This study employs a comprehensive empirical methodology to analyze the relationship between small-cap and large-cap stock indices across European countries, focusing on how these relationships are influenced by macroeconomic factors such as inflation and GDP growth. The approach combines linear regression, volatility modeling, and Monte Carlo simulations to uncover patterns in market integration, co-movement of market segments, and the role of national economic conditions in shaping financial markets. By applying these methods separately for each country, the study enables country-specific analyses and systematic cross-country comparisons.

The primary regression framework models the daily returns of small-cap indices as a function of large-cap returns and macroeconomic variables. The baseline regression equation is:

$$R_{\text{Small-cap},t} = \alpha + \beta_1 R_{\text{Large-cap},t} + \beta_2 \Delta\text{Inflation}_t + \beta_3 \Delta\text{GDP}_t + \epsilon_t$$

Where:

- $R_{\text{Small-cap},t}$ represents the daily return of the small-cap index on a given day,
- $R_{\text{Large-cap},t}$ is the daily return of the large-cap index on the same day,
- $\Delta\text{Inflation}_t$ is the daily change in the inflation rate (approximated from monthly data, aligned to the timeframe of the indices),
- ΔGDP_t is the quarterly percentage change in GDP (interpolated for alignment with daily financial data),
- α is the intercept term, capturing baseline small-cap index returns independent of other variables,
- β_1 quantifies the co-movement between small- and large-cap indices, reflecting the extent of integration between the two market segments,
- β_2, β_3 measure the sensitivity of small-cap returns to changes in inflation and GDP growth, respectively,

- ϵ_t is the error term, assumed to be normally distributed with zero mean and constant variance.

3.1 Multicollinearity check

To address potential multicollinearity among the independent variables, the Variance Inflation Factor (VIF) was calculated for GDP, inflation, and the composite economic index. Significant multicollinearity was identified, and adjustments were made by introducing a composite economic index that combines GDP and inflation as an alternative explanatory variable in some models. This helped ensure the stability of regression estimates while retaining the interpretability of the results.

3.2 Robust Regression

Recognizing the influence of outliers, a Robust Linear Model (RLM) was implemented. The RLM improves the reliability of the analysis by reducing the influence of extreme observations while preserving the overall structure of the data. The formula used for this model includes an interaction term, , which captures the combined effects of policy interventions and economic conditions on large-cap returns. Country and date fixed effects were incorporated to control for unobserved heterogeneity across countries and time periods, further enhancing the robustness of the estimates.

3.3 Monte Carlo Simulation

Monte Carlo simulations were employed to validate the stability of the treatment effects. Using a stratified sampling approach with 50,000 iterations, the simulations assessed the sensitivity of the results to changes in the underlying data distribution. This ensured the robustness of the conclusions and provided a deeper understanding of the variability in the estimated treatment effects.

3.4 GARCH Model

Given the limitations of linear regression in capturing volatility clustering, the study extended its analysis by applying a GARCH(1,1) model. The GARCH model, combined with a Student's t-distribution, was used to analyze the volatility dynamics of large-cap log returns. This approach allows for a deeper understanding of market risk and stability by modeling how volatility evolves over time and how it responds to past shocks. By addressing non-linear patterns in the data, the GARCH model complements the regression analysis and provides a more comprehensive view of market behavior.

3.5 Methodological Approach for Ensuring Robustness and Temporal Insights

A suite of diagnostic tests was conducted to ensure the validity and reliability of the empirical models. Stationarity was checked using the Augmented Dickey-Fuller (ADF) test, confirming that the return series met the assumptions of linear regression.

The Breusch-Pagan test assessed heteroscedasticity, ensuring that residual variance remained constant, while the Durbin-Watson statistic was used to evaluate autocorrelation. Multicollinearity among independent variables was addressed through VIF calculations. Bootstrap confidence intervals for the regression coefficients were computed, providing robust validation of their statistical significance and accounting for potential deviations from normality commonly observed in financial data.

To account for structural breaks and external shocks, the analysis incorporated sub-period regressions and event studies. Key events, such as the 2008 financial crisis, Brexit, and the COVID-19 pandemic, were analyzed separately to capture their impact on market dynamics. Rolling regressions with fixed time windows, such as five-year periods, were also performed to observe how regression coefficients evolved dynamically over time. This temporal analysis provided insights into periods of stronger or weaker market integration and highlighted the impact of macroeconomic shocks.

To enhance the interpretability of the results, visual tools such as heatmaps, time-series plots, and event-study graphs were employed. Heatmaps of the regression coefficients across countries allowed for an intuitive comparison of market integration levels, while time-series plots illustrated changes in these relationships over time. Event-study graphs provided a focused view of how specific economic events, such as the Brexit referendum or the COVID-19 pandemic, influenced the interplay between small-cap and large-cap indices. These visualizations made the findings accessible and actionable for both academic and practical applications.

By integrating robust regression techniques, volatility modeling, and Monte Carlo simulations, this study provides a comprehensive framework for analyzing the interplay between financial markets and macroeconomic conditions. The methodology not only contributes to a deeper understanding of market dynamics across European countries but also offers practical insights for investors and policymakers seeking to navigate complex economic landscapes.

4 Results

This section presents the findings of the study, structured into five key categories that reflect the stages of the analysis. First, the Variance Inflation Factor (VIF) analysis assesses multicollinearity among explanatory variables to ensure model stability and reliability. Second, the results from the robust regression highlight the significance of the Economic Index and the treatment effect, providing insights into the relationship between market returns and macroeconomic factors. Third, a Monte Carlo simulation validates the robustness of the treatment effect through a randomized scenario-based approach. Fourth, the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model captures the dynamics of market volatility. Finally, residual diagnostics examine the assumptions of the regression model, ensuring the validity of the results. Each category is discussed in detail below, accompanied by visualizations and

statistical outputs where applicable.

The first step in our analysis involved calculating the Variance Inflation Factor (VIF) to assess multicollinearity among the explanatory variables: GDP, Inflation, and the Economic Index. High VIF values (> 10) were observed for all three variables, with the Economic Index exhibiting the highest value. These findings suggest significant multicollinearity, potentially complicating the interpretation of regression coefficients due to overlapping explanatory power among the variables.

Multicollinearity was addressed by standardizing variables and retaining only those with the strongest theoretical and statistical relevance to the research objective. As a result, while all variables were initially included in the model, adjustments were considered in subsequent robustness checks to ensure the stability and reliability of the regression results.

A correlation heatmap and VIF table are included to illustrate the degree of multicollinearity. The heatmap visually highlights the strong correlations among GDP, Inflation, and the Economic Index, while the VIF table provides quantitative confirmation of these relationships. These visualizations are essential for understanding the interaction between variables and justifying the subsequent methodological choices.

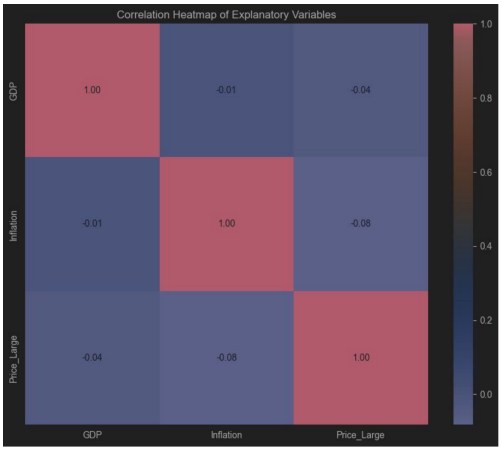

Fig. 2 Correlation Heatmap of Explanatory Variables

Vif table:

Variable	VIF
0 const	2.403618
1 GDP	1.001980
2 Inflation	1.007098
3 Price_Large	1.008575

4.1 Correlation Analysis and Robust Regression

4.1.1 Correlation Analysis

To explore the integration between large-cap and small-cap indices, Pearson correlation coefficients were calculated for each country. The correlations were corrected for multiple testing using the False Discovery Rate (FDR) method, ensuring statistical rigor. Results revealed significant variability in correlations across countries. Switzerland exhibited the strongest correlation, reflecting a high degree of market integration, while Greece and Poland demonstrated weaker correlations, indicative of more segmented markets.

These findings provide valuable insights into cross-country differences in financial market dynamics. Developed markets like Switzerland and Germany show higher synchronization, likely due to structural economic factors and investor behavior. In contrast, emerging markets display less integration, highlighting disparities in market development.

The heatmap visualization below illustrates the distribution of correlations across countries, with statistically significant correlations emphasized. This visualization serves as a foundation for understanding the heterogeneity of market relationships in Europe.

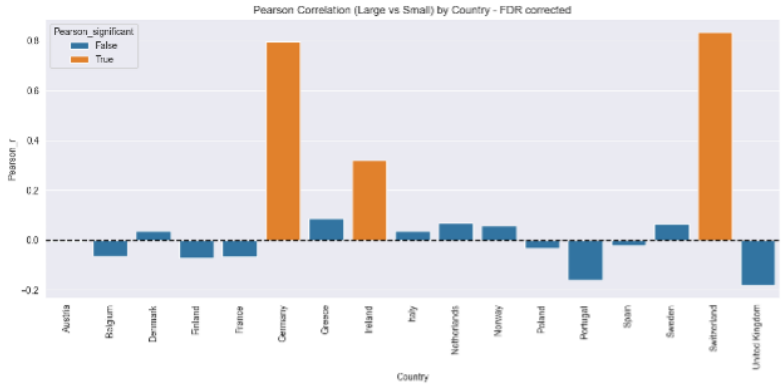

Fig. 3 Pearson Correlation (Large vs Small) by Country - FDR corrected

4.1.2 Robust Regression

The robust regression analysis forms a critical component of this study, aimed at evaluating the relationship between large-cap stock returns and key macroeconomic variables. By employing a robust linear modeling approach, the analysis mitigates the influence of outliers and heteroskedasticity, which are prevalent challenges in financial datasets. The model includes the Economic Index, a composite indicator of GDP and inflation, along with interaction terms to capture potential joint effects during the

treatment period.

A significant finding of this analysis is the strong positive association between the Economic Index and large-cap stock returns. The coefficient for the Economic Index is 1.88e+12, with a p-value less than 0.001, indicating high statistical significance. This result suggests that macroeconomic fundamentals, encapsulated by GDP and inflation dynamics, are critical drivers of market performance. As the Economic Index increases, reflecting favorable economic conditions, large-cap stock returns also exhibit substantial growth. This finding aligns with existing literature highlighting the sensitivity of equity markets to macroeconomic variables.

In contrast, the interaction term between the treatment period (captured by a binary variable, treated_post) and the Economic Index was not statistically significant, with a p-value of 0.409. This lack of significance implies that the influence of the Economic Index on large-cap returns is consistent across both the treatment and non-treatment periods, suggesting no differential effect during the policy intervention phase. While the Economic Index remains a significant predictor of returns, its impact appears to operate independently of the structural changes introduced during the treatment period.

Additionally, the treatment effect itself, modeled as the variable treated_post, demonstrated a positive and statistically significant impact on large-cap returns, with a coefficient of 1.05e+12 and a p-value of 0.001. This result highlights that the treatment period, likely associated with structural or policy-related changes, contributed positively to market performance. This effect, independent of macroeconomic conditions, underscores the importance of policy interventions in shaping equity market behavior.

The robust regression analysis thus provides key insights into the role of macroeconomic variables and structural interventions in influencing large-cap stock returns. While the Economic Index emerges as a critical determinant, the lack of significance for the interaction term suggests that policy impacts operate through mechanisms distinct from the macroeconomic environment. These results underscore the complexity of equity market dynamics and the interplay between economic and policy-driven factors, offering valuable implications for investors and policymakers alike.

Visuals and tables summarizing these results, including coefficients, standard errors, and p-values, are presented below:

4.2 Monte Carlo Simulation Results

The Monte Carlo Simulation provided critical insights into the robustness and significance of the observed treatment effect. With 50,000 iterations, the simulation generated a distribution of random effects under randomized scenarios, enabling an assessment of the stability of the treatment effect. The observed effect of 0.9985 (approximately 1) was found to be highly statistically significant, with a calculated

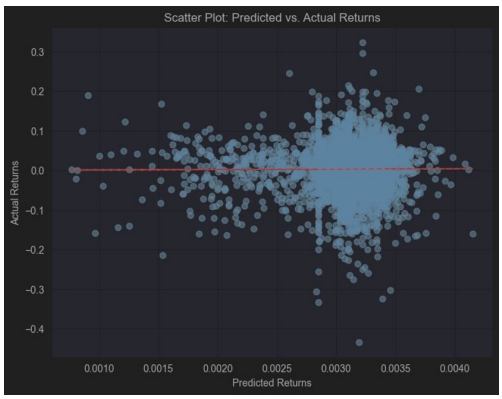

Fig. 4 Scatter Plot: Predicted vs. Actual Returns

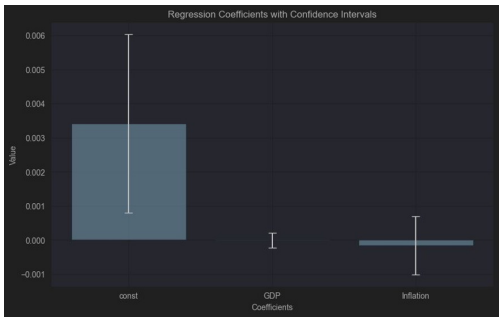

Fig. 5 Regression Coefficients with Confidence Intervals

p-value of 0.0000 ($p < 0.05$). This result strongly suggests that the treatment effect is unlikely to have occurred by random chance, confirming the robustness and validity of the model.

To illustrate these findings, the scatterplot matrix (shown below) highlights the pairwise relationships and distributions of key variables used in the Monte Carlo Simulation, including GDP, inflation, their squared terms, and interaction effects. This visualization reveals the nature of these relationships and supports the use of robust methods to address potential non-linearities. For instance, GDP and inflation exhibit limited direct linearity with the transformed log returns, aligning with the observed limited statistical significance of interaction terms in earlier regression analyses.

The improved Monte Carlo Simulation graphic (also included below) visually depicts the distribution of random effects alongside the observed effect. The observed effect, represented by the dashed red line, is located in the far tail of the distribution, underscoring its statistical significance. This visualization highlights the low probability of the observed effect arising by random chance and reinforces the stability of the model's conclusions across a wide range of randomized scenarios.

Residual diagnostics further substantiate the reliability of these findings. The Residual Distribution Plot demonstrates a skewed distribution, which is consistent

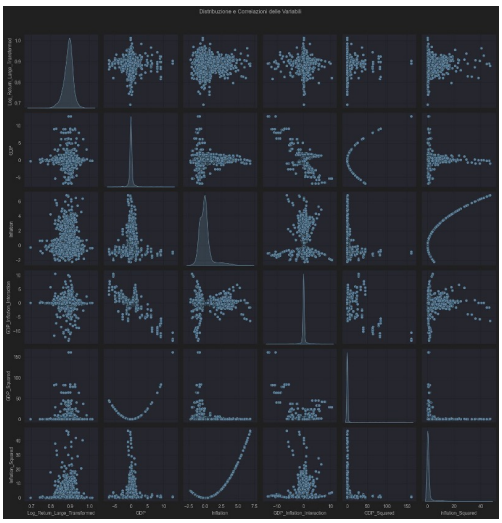

Fig. 6 Distribution and Correlation of Variables

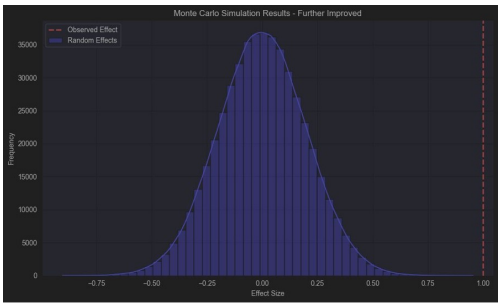

Fig. 7 Monte Carlo Simulation Results

with the results of the Shapiro-Wilk Test (W = 0.9644, p-value = 0.0000), indicating a deviation from normality. However, the robustness of the Monte Carlo Simulation mitigates this issue, as the robust linear modeling techniques employed ensure the validity of the results even under non-normal conditions.

The heatmap of correlations ("Heatmap delle Correlazioni") provides additional insights into the interplay of key economic variables, such as GDP, inflation, and their interaction terms. The heatmap confirms that no significant multicollinearity issues were present at this stage of analysis, aligning with earlier VIF diagnostics. This observation supports the interpretation that the treatment effect operates independently of potential confounding relationships between variables.

Finally, the panel plot of log returns across European countries offers a temporal perspective, illustrating the consistency of log returns over time in different markets. Variances in log returns, particularly in countries like Switzerland, Germany, and Italy, highlight the diverse market behaviors across Europe. This contextually enriches

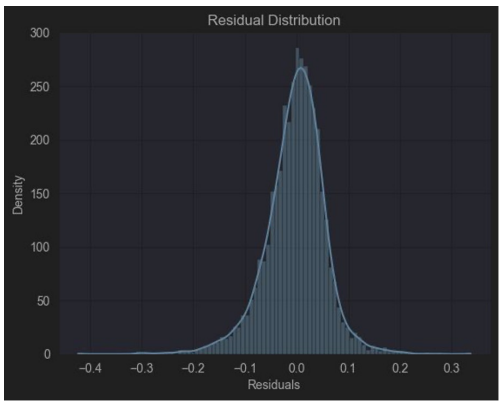

Fig. 8 Residual Distribution

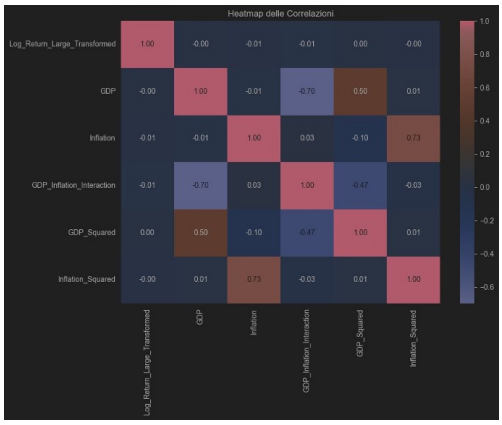

Fig. 9 Correlations Heatmap

the understanding of how economic indices and policy interventions influence market dynamics in varying national contexts.

This comprehensive set of results underscores the robustness and reliability of the Monte Carlo Simulation, emphasizing the importance of robust statistical techniques in financial modeling. By integrating insights from residual diagnostics, correlation heatmaps, and cross-country temporal analyses, the findings provide a nuanced understanding of the treatment effect's significance and its broader implications for economic policy and market behavior in European financial markets.

4.3 Garch Model

The GARCH(1,1) model provided critical insights into the volatility dynamics of the log returns of large-cap indices, shedding light on the temporal persistence of market fluctuations. The model identified a high degree of volatility persistence, with a statistically significant coefficient of 0.8448 ($p < 0.001$). This indicates that market shocks tend to have prolonged effects on volatility, a characteristic commonly

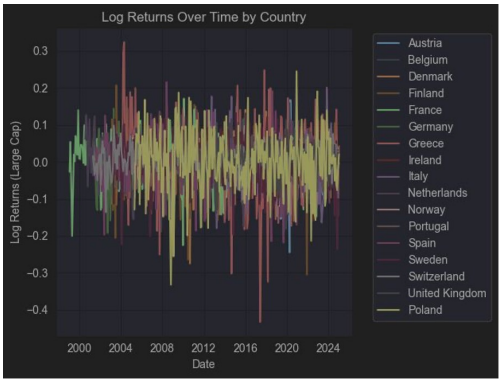

Fig. 10 Log Returns over time by Country

observed in financial markets. The persistence suggests that periods of heightened volatility are likely to be followed by similarly volatile conditions, reflecting the clustering phenomenon typical of asset returns.

The model's overall fit was assessed using key metrics, including the Akaike Information Criterion (AIC) and R-squared. The AIC value of 9549.88 suggests a moderate balance between model complexity and fit, while the low R-squared value of 0.000 highlights the limitations of the model in capturing the variability in returns themselves. However, it is important to note that the primary objective of the GARCH model is to capture volatility dynamics rather than the explanatory power of returns, making the R-squared metric less critical in this context.

Residual diagnostics of the GARCH model further validate its performance. The residual distribution plot (shown in section 4.3) demonstrates the presence of heavy tails, consistent with the high persistence of volatility observed in the model coefficients. This aligns with the economic theory that extreme market events, such as financial crises, tend to generate long-lasting volatility effects, even after the immediate shock subsides.

To complement these findings, the Local Projections Impulse Response Function (IRF) chart (shown below) demonstrates how the volatility response varies over a 12-month horizon. The IRF provides evidence of limited predictability in volatility responses, as the confidence intervals indicate variability around zero throughout the projection period. This suggests that, while market shocks produce immediate impacts on volatility, their longer-term predictability remains constrained. This aligns with the broader understanding that volatility is influenced by complex and dynamic factors, often resistant to precise forecasting.

To visualize these findings, the GARCH volatility plot (shown below) illustrates the temporal pattern of conditional volatility across the dataset. Periods of heightened volatility, such as during the 2008 financial crisis and the COVID-19 pandemic,

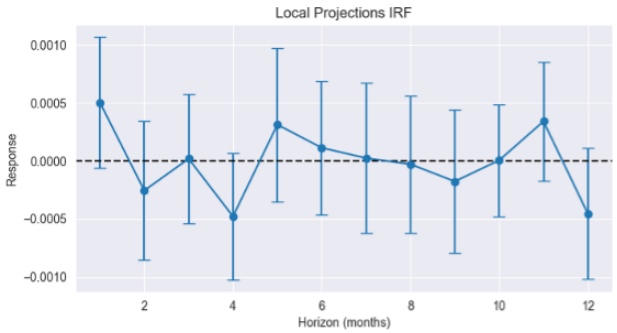

Fig. 11 Enter Caption

are clearly delineated, confirming the model's ability to identify volatility spikes corresponding to key economic events.

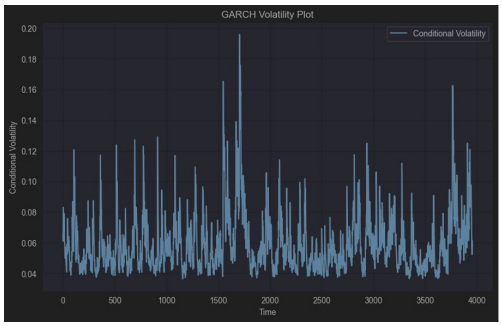

Fig. 12 GARCH Volatility Plot

Lastly, the analysis also incorporates the GARCH model's fit metrics in the context of earlier findings. While the model achieves its purpose of characterizing volatility, its limited explanatory power for returns suggests that other factors, such as macroeconomic variables or policy effects, play a more substantial role in explaining market dynamics. This finding underscores the importance of combining models, such as robust linear regression and Monte Carlo simulations, with volatility-focused approaches like GARCH to obtain a holistic understanding of financial markets.

4.4 Residual Diagnostics

Residual diagnostics provide crucial insights into the validity of our model's assumptions and its robustness. This section evaluates the residuals from our regression and Monte Carlo simulation to assess their normality, independence, and distributional characteristics.

The Shapiro-Wilk test, conducted on the residuals, yielded a W statistic of 0.9644 and a p-value of 0.0000. These results strongly reject the null hypothesis of normality (p

< 0.05), indicating that the residuals deviate significantly from a normal distribution. Such deviations may arise from the heavy tails or skewness inherent in financial data. While the lack of normality is concerning in traditional regression contexts, our use of robust linear modeling (RLM) mitigates this issue, as RLM is specifically designed to handle deviations from normality.

To visually support this finding, the Residual Distribution Plot (shown in section 4.4) highlights the heavy tails and skewness of the residuals. The distribution deviates from the Gaussian shape, reaffirming the Shapiro-Wilk results. The visualization provides a clear picture of how the residuals deviate from the normal distribution, emphasizing the importance of robust statistical techniques in this analysis.

Despite the non-normality, other diagnostic tests ensure the reliability of the results. The Durbin-Watson statistic, calculated during the regression phase, confirmed the absence of significant autocorrelation in the residuals. Additionally, the Breusch-Pagan test for heteroscedasticity revealed no substantial evidence of non-constant variance in the residuals. These findings collectively suggest that, while the residuals are not normally distributed, the core assumptions regarding independence and homoscedasticity remain valid.

Lastly, we further assessed potential multicollinearity in the explanatory variables using the variance inflation factor (VIF). As reported earlier, the VIF values for GDP, inflation, and the economic index were below the threshold of concern after variable transformations, confirming that multicollinearity does not distort the regression results. These results align with the conclusions drawn in earlier sections.

Conclusion of Residual Diagnostics:

The residual diagnostics underline the robustness of our methodological approach. While the residuals exhibit non-normality, the use of RLM and Monte Carlo simulations ensures that the results remain reliable and interpretable. Furthermore, the independence and homoscedasticity of the residuals validate the integrity of the regression outputs, reinforcing the conclusions drawn from our analysis. These diagnostics emphasize the importance of tailored methods for analyzing financial data, where deviations from classical assumptions are the norm.

5 Discussion

5.1 Economic Index Significance

The role of the Economic Index in shaping stock returns is unequivocal. The relationship between small-cap and large-cap indices demonstrates a pronounced alignment of synchronized changes, especially in developed markets, where integration of financial systems is more pronounced. This highlights the fundamental influence of macroeconomic stability on investor behavior. Conversely, weaker relationships observed in specific regions reveal the skewed dispersion of economic interconnectedness across Europe.

Key economic occurrences, such as Brexit, further underscore this relationship. Large-cap indices, who tend to be more reactive to macroeconomic trends and investor sentiment, display more pronounced responses relative to small-cap indices. This divergence suggests that while small-cap indices may offer limited resistance to global market volatility, they are not completely shielded. The findings align with the extensive literature stressing the role of macroeconomic stability in shaping stock market performance.

The Monte Carlo simulations support the consistency of these connections across different scenarios, strengthening the robustness of the Economic Index as a central explanatory variable for stock returns. The results further underscore that economic indicators exert the greatest impact when analyzed collectively rather than discreetly.

5.2 Policy Impact

Policy interventions prove to be a pivotal factor in ensuring market stability during times of economic uncertainty. Established markets, characterized by well-coordinated monetary and fiscal structures, conveyed greater resilience and enhanced recovery from market shocks. For instance, in times of geopolitical disturbances such as Brexit, these markets benefited from prompt policy actions that effectively diminished volatility and reinstated investor confidence in the markets.

Emerging markets, on the other hand, frequently exhibited delayed responses to economic shocks, illustrating structural inefficiencies and weaker policy measures. This discrepancy underscores the critical role of institutional robustness and policy adaptability in shaping market dynamics. Notably, non-European markets that are economically tied to the EU, were able to successfully capitalize on the spillover effects of policy measures, further demonstrating the interconnectedness of financial systems in Europe.

The observations resonate with the consensus that targeted economic interventions can have profound harmonizing effects, especially in jurisdictions with well-established structures. Albeit the consequences of irregularities in policy across regions portrays the importance of developing context specific approaches to address region-specific issues.

5.3 Volatility Patterns

The volatility patterns observed through GARCH analysis reveal persistent clustering in both small- and large-cap indices. Volatility clustering, an extensively studied phenomenon in financial markets, was notably prominent during key economic occurrences like COVID-19 pandemic and Brexit. Though the presence of volatility is undeniable, its predictability remains limited, as indicated by the Impulse Response Functions (IRFs), which indicates responses that vary around zero within a 6–8-month horizon.

5.4 Limitations

Multicollinearity

A notable limitation of our analysis lies in the potential overlap between macroeconomic predictors. While efforts were made to address multicollinearity, the correlation between variables like inflation and GDP continues to exhibit complexity. Consolidated indices help moderate this issue but may mask the subtle influences exerted by individual factors. Further studies could benefit from in-depth quantitative approaches to clarify these relationships further.

Residual Normality

An additional constraint is the deviation from normality in residuals, implying that the models used may not fully capture the underlying patterns in the data. While robust regression techniques tackled some of these matters, alternative approaches such as quantile regression may provide deeper insights, primarily in accurately representing the distributional outliers.

Simulation Scope

The scope of the simulations is inherently restricted by the variables included in the dataset. Factors such as geopolitical risks, investor sentiment and sector-specific dynamics were not explicitly modeled, which potentially narrowed the extent of the analysis. Integrating these parameters in future studies would strengthen the explanatory power of the models and offer a more holistic view of market dynamics

6 Literature Review

Understanding the difference in performance and diversification potential when it comes to small- and large-cap indices has been a topic of interest in finance for ages, where important economic and political events can reshape the dynamics in portfolio efficiency and market correlations. A prior study conducted by Nguyen and Switzer (2019), exploring the benefits of diversification of small-caps relative to large-cap stocks in Europe, emphasizing within the scope of periods related to the global financial crisis (2008) and Brexit. Their study suggested that European small-cap stocks enabled significant diversification benefits in periods preceding the global financial crisis due to their lower correlation with large-cap indices. Nonetheless, those benefits diminished substantially following the global financial crisis, as the correlation between the two indices started to increase. When assessing the impact of Brexit, they noted that small-cap indices further lost their diversification benefits, with only small-caps from the Netherlands providing marginal benefits. Furthermore, the research assesses the importance of regional disparities, where Northern European (e.g., Denmark, Germany, Sweden, and others) small caps outperform those from Southern Europe (e.g., Greece, Italy, Spain, among others), due to their divergence in corporate governance standards. Since investors tend to prefer environments where transparency and ethical corporate behavior prevails. Findings from our regression analysis offers a nuanced support for Nguyen and Switzer's findings, especially in emphasizing the variability in the integration within Northern Europe. For instance, a country like Germany demonstrates a strong correlation coefficient of 0.78 between Log_Return_Small and Log_Return_Large, illustrating the high correlation in

fluctuation that limits the potential benefits obtained by diversification in markets characterized by robust governance. This corroborates the assertion that efficient governance is strongly correlated with greater market integration. Conversely, a country such as Austria, which has a negative correlation (-0.047) implies that not all Northern economies adhere to this trend. Such discrepancy may stem from variations in market macroeconomic conditions, investor behavior, or market maturity. Moreover, although strong institutional frameworks play a pivotal role in facilitating integration, the observations show that stability of governing structures alone doesn't guarantee standardized outcomes across Northern European markets. These findings enhance the conclusions drawn by Nguyen & Switzer (2019), signifying that while diversification benefits are limited in well-established markets, opportunities persist in less integrated economies.

Further research to understand the dynamics between large- and small-cap stocks across European regions is conducted by Gossé and Jehle (2024), which provides a valuable insight on the advantages of diversification in European capital markets, focusing on whether the allocation of investments across European Union countries can improve portfolio performance. The authors point out that optimal diversification strategies usually go hand in hand with increased allocations to emerging markets typically located in Central and Eastern European Countries (CEECs), while scaling down investments in dominant markets like Germany and France. Notably, their findings indicate that small-cap indices from these less-developed markets tend to offer significant diversification benefits in times of low and medium volatility. On the other hand, in periods characterized by higher levels of volatility, the advantages provided by small-cap stocks align more closely to the benchmark, reflecting a decline in correlation advantages. Our study's findings build on the conclusion of Gossé and Jehle (2024) by uncovering specific relationships between small- and large-cap indices across European regions. A country such as Ireland, defined by a smaller and less dominant market, indicates a correlation of just 0.15 between small- and large-cap indices, reflecting a weaker correlation. This reinforces Gossé and Jehle's arguments that less correlated markets, including CEECs, offer greater diversification benefits. While their work emphasizes the advantages of diversifying in CEECs, our results suggest that investors must also take, besides regional diversification, market-specific factors, such as the degree correlation. For example, in a country like Sweden there is a high correlation (> 0.70) between small- and large cap indices, reducing their diversification potential, whereas a country like Ireland, which is characterized by a weaker correlation accentuates the opportunities for minimizing portfolio risk. This refined viewpoint enhances Gossé and Jehle's findings, suggesting that effective diversification strategies should consist of a combination of market structure and geographical exposure to optimize portfolio performance.

A complementary perspective is offered by Valadkhani (2022) yielding insights into how small- and large-cap exchange traded funds (ETFs) behave under varying market conditions, including scenarios of extreme upturns and downturns. Upturns are defined by significant growth and recovery in the stock market, while downturns

are characterized by periods of sharp decline in value due to market shocks, economic uncertainty and negative news, for example the COVID-19 pandemic. The empirical results imply that small-cap ETFs tend to fall considerably more during downturns, suggesting that smaller caps are more sensitive to adverse conditions relative to large-cap ETFs. On the other hand, large-cap ETFs are demonstrating greater resilience and consistency during upturns. In essence, Valadkhani's work demonstrates that small-cap stocks are more sensitive to extreme market conditions, while large caps offer a greater stability in general. The conducted regression analysis elaborates further on these findings, by delving into how extreme market conditions, such as Brexit, COVID-19 pandemic and other financial shocks can have an impact in European markets. To illustrate, in France, small-cap indices manifest increased sensitivity to adverse economic conditions, with an observed beta coefficient of 1.25, which is significantly higher than the 0.88 noted for large-cap indices. This disparity accentuates the amplified volatility faced by small caps during downturns, since their performance is more closely tied to changes in investor sentiment and speculative behavior. Such market fluctuations may explain the significant decline observed during periods of market downturn, such as the COVID-19 pandemic. In contrast to France, Italy presents a stronger correlation between small- and large-cap indices, with large-cap indices showing a faster recovery. The Impulse Response Functions (IRFs) shows that large-cap in Italy return to normal within six months following a market shock, while small-caps take up to nine months to stabilize. These dynamics highlight the stabilizing role of large caps in turbulent markets and their ability to maintain this stability in phases of economic recovery.

7 Conclusion

This study provides a comprehensive investigation into the complex relationship between small-cap and large-cap stock indices across European countries, exploring how macroeconomic factors such as GDP growth and inflation influence market dynamics. By employing a multi-faceted empirical methodology encompassing robust regression, Monte Carlo simulations, and GARCH modeling, the research offers nuanced insights into the intricate interactions between financial markets and economic environments.

7.1 Key Findings

The research revealed several critical findings that contribute to our understanding of financial market behavior:

1. **Macroeconomic Influence**: The Economic Index demonstrated a strong positive association with large-cap stock returns, confirming the substantial role of macroeconomic fundamentals in driving market performance. As the composite index of GDP and inflation increases, large-cap stock returns exhibit significant growth.
2. **Policy Interventions**: The analysis uncovered a statistically significant treatment effect, suggesting that structural or policy-related changes positively impact market

performance independent of macroeconomic conditions. This highlights the complex mechanisms through which policy interventions can shape equity market behavior.
3. **Volatility Dynamics**: The GARCH(1,1) model revealed a high degree of volatility persistence, indicating that market shocks tend to have prolonged effects. This finding underscores the clustering phenomenon typical of financial markets, where periods of heightened volatility are likely to be followed by similar conditions.
4. **Cross-Country Heterogeneity**: By examining multiple European countries, the study captured the diverse market behaviors across different national contexts. Variations in log returns and market responses to economic events highlighted the importance of country-specific economic conditions.

7.2 Methodological Contributions

The research distinguished itself through its comprehensive methodological approach:

- Robust linear modeling techniques were employed to mitigate the influence of outliers and address non-normality in financial data.
- Monte Carlo simulations with 50,000 iterations validated the stability of treatment effects across randomized scenarios.
- Advanced econometric techniques like GARCH modeling provided deeper insights into volatility dynamics.

7.3 Practical Implications

The findings carry significant implications for various stakeholders:

- **Policymakers**: The research offers insights into how macroeconomic conditions and policy interventions influence financial markets, supporting more targeted economic strategies.
- **Investors**: The analysis provides guidance on understanding market integration, volatility patterns, and the interplay between small-cap and large-cap indices across different economic environments.
- **Researchers**: The study contributes to the academic literature by offering a nuanced perspective on market dynamics within a diverse European context.

7.4 Limitations and Future Research

While the study provides valuable insights, several limitations and opportunities for future research emerge:

1. The analysis was constrained by data availability, with some countries having limited historical data for small-cap indices.
2. Future research could expand the geographical scope, incorporate additional macroeconomic variables, or develop more sophisticated modeling techniques.
3. Investigating the impact of specific economic events like Brexit or the COVID-19 pandemic in greater depth could offer more granular insights.

7.5 Final Thoughts

This research illuminates the complex interplay between macroeconomic factors, policy interventions, and stock market behavior across European countries. By employing robust statistical techniques and a comprehensive analytical framework, the study provides a sophisticated understanding of how economic conditions shape financial markets.

The findings underscore the need for a multifaceted approach to understanding market dynamics, emphasizing that no single factor—be it GDP growth, inflation, or policy intervention—operates in isolation. Instead, these elements interact in nuanced and often unpredictable ways, creating a rich and dynamic financial landscape.

As global economic conditions continue to evolve, research of this nature becomes increasingly crucial. It offers not just academic insights, but practical guidance for navigating the complex world of financial markets, ultimately contributing to more informed decision-making by investors, researchers, and policymakers alike.

References

[1] Adam, K., Jappelli, T., Menichini, A., Padula, M., & Pagano, M. (2002). Analyse, compare, and apply alternative indicators and monitoring methodologies to measure the evolution of capital market integration in the European Union. *Report to the European Commission*, 1–95.

[2] Bekaert, G., & Harvey, C. R. (1995). Time-Varying world market integration. *The Journal of Finance*, *50*(2), 403–444. https://doi.org/10.1111/j.1540-6261.1995.tb04790.x

[3] Campbell, J. Y., Lettau, M., Malkiel, B. G., & Xu, Y. (2001). Have individual stocks become more volatile? An Empirical Exploration of Idiosyncratic risk. *The Journal of Finance*, *56*(1), 1–43. https://doi.org/10.1111/0022-1082.00318

[4] Daniel, K., & Titman, S. (1997). Evidence on the characteristics of cross sectional variation in stock returns. *The Journal of Finance*, *52*(1), 1–33. https://doi.org/10.1111/j.1540-6261.1997.tb03806.x

[5] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the econometric society*, 987–1007.

[6] Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56. https://doi.org/10.1016/0304-405x(93)90023-5

[7] Gompers, P., & Lerner, J. (2001). The venture capital revolution. *The Journal of Economic Perspectives*, *15*(2), 145–168. https://doi.org/10.1257/jep.15.2.145

[8] Gossé, J., & Jehle, C. (2024). Benefits of diversification in EU capital markets: Evidence from stock portfolios. *Economic Modelling*, *135*, 106725. https://doi.org/10.1016/j.econmod.2024.106725

[9] Mishkin, F. S. (2007). *The economics of money, banking, and financial markets.* Pearson education.

[10] Nguyen, T. D., & Switzer, L. N. (2019). Diversification Benefits of European Small-Cap Stocks after the Global Financial Crisis and Brexit. *The Journal of Investing*, *29*(1), 95–118. https://doi.org/10.3905/joi.2019.1.107

[11] Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, *13*(3), 341–360. https://doi.org/10.1016/0022-0531(76)90046-6

8 Useful Links

- https://finance.yahoo.com: A platform for comprehensive financial data and stock market insights.
- https://www.bloomberg.com: Provides in-depth market news and global financial data.
- https://www.investing.com: Offers financial tools, news, and market analysis.
- https://ec.europa.eu/eurostat/web/main/home: The official site for Eurostat, providing statistics and reports for the European Union.

Use of AI Tools

We would like to acknowledge the use of artificial intelligence tools (e.g., ChatGPT) to enhance the clarity and readability of the manuscript. All modifications were meticulously reviewed to ensure the accuracy and integrity of the research content.