O Bubble, (Where) Art Thou?

Market Concentration in an Era of Flux

Abstract—The increasing concentration of a few dominant stocks in the US stock market has significant implications for portfolio construction and risk management. This paper analyzes the drivers of this concentration and its impact on market dynamics. We decompose market sensitivities, assess the effects of concentration on volatility, and apply a Log-Periodic Power Law Singularity (LPPLS) model to detect speculative behavior. Additionally, we examine the market response to tail risk events, particularly the behavior of the Magnificent Seven during the COVID-19 crash. Based on these findings, we propose a trading strategy that refines the traditional long small-cap, short large-cap approach by incorporating effects from concentration trends and volatility spillovers. Our results suggest that increasing dominance by a few firms alters systemic risk transmission, with implications for both passive investing as well as active portfolio management.

1 Introduction

Stock market concentration has increased significantly in recent years. The top 10 stocks in the S&P 500 accounted for 28% of total market capitalization in early 2024, a stark increase from 14% in 2014. Much of this growth has been driven by the 'Magnificent Seven' (Apple, Microsoft, Alphabet, Amazon, Meta, Nvidia, and Tesla), which have collectively exerted extreme influence over market returns. This development has been further pronounced in market-capitalization-weighted indices such as the S&P 500, and naturally raises concerns regarding diversification, systematic risk and potential speculative bubbles.

Past instances of market concentration have been followed by corrections and deconcentration, a cyclicality that has associated concentrated regimes with speculative excess and the existence of bubbles. The present market condition is indeed highly concentrated in relation to the previous decade, but some argue that this time, we are observing something structurally different. What appears to differentiate this occasion from prior cycles — such as the late 1990s dot-com boom — the large valuations that drive the concentration today are largely underpinned by strong fundamentals [Michael Mabousinn, 2024]. Today's market leaders not only command a significant share of capitalization, but also contribute disproportionately to fundamentals, such as profit and return on invested capital (ROIC). This is, in general, achieved through a combination of economies of scale, network effects, and intangible assets such as software, data, and intellectual property, which reinforce their dominance.

A further characteristic that distinguishes our era is the rapid pace of concentration growth, which is the fastest to be recorded since the 1950s. This, of course, raises questions about this regime's long-term sustainability, and has broader implications for financial markets. Some argue that this surge in fundamentals is not merely cyclical, but rather indicative of a broader economic regime shift, where "superstar firms" capture an increasing share of economic value due to their structural advantages in technology, scale, and productivity [David Autor, 2020].

This paper investigates the drivers and implications of this concentration from several perspectives. First, we quantify the sensitivity of the S&P 500 to the Magnificent Seven by decomposing its market sensitivity, in terms of beta. We then examine how their additional growth affects volatility, and if their dominance has introduced systematic risk. Finally, we attempt to detect the potential existence of speculative dynamics (bubbles), and the extent of volatility spillovers, together with their implications regarding market stability in tail risk events.

To address these questions, we apply empirical methods such as beta decomposition, realized volatility estimation, and a DCC-GARCH model to analyze dynamic correlations [Engle, 2002]. We utilize a Log-Periodic Power Law Singularity (LPPLS) model [Riza Demirer, 2018] to assess the possibility of a potential speculative bubble regime. Finally, we leverage these insights towards a trading strategy that aims to capitalize on the potential existence of distinct market regimes modifying the traditional long small-cap, short large-cap investing framework.

The remainder of this paper is structured as follows. Section 2 reviews the literature on market concentration, risk decomposition, volatility clustering, and speculative cycles. Section 3 presents the methodology, including beta decomposition, volatility modeling, and the LPPLS model. Section 4 discusses and analyzes our results, followed by a discussion which informs a Trading Strategy in Section 5. and a conclusion in Section 6.

2 Literature Review

The increasing concentration of equity markets has been widely studied in financial economics, particularly with regard to its implications for systematic risk, market efficiency, and investment strategies. While market cap-weighted indices naturally assign higher weights to larger firms, the recent surge in concentration has reached levels not seen since the early 20th century. Empirical studies have examined whether this phenomenon is driven by fundamentals or speculative excess, whether it introduces systemic risk, and how it affects volatility transmission, beta dynamics, and trading strategies. This section reviews key contributions in these areas, situating our study within the broader literature.

2.1 Market Concentration and Systematic Risk

Historically, episodes of rising concentration have often been associated with financial instability, as dominant firms are often viewed as amplifiers of systemic shocks. Research has identified that past peaks in concentration, such as those observed in the 1900s, 1960s, and the late 1990s, were often followed by de-concentrations in periods of market corrections. Studies suggest that concentration increases both idiosyncratic risk (as investors become more exposed to a few firms) and systematic risk (as these firms account for a greater share of macroeconomic fluctuations) [Michael Mabousinn, 2024].

However, some researchers argue that the current circumstances are structurally different. Unlike previous episodes, in which dominant firms frequently lost their edge during downturns, today's market leaders exhibit profoundly strong fundamentals. Broad evidence suggests that the Magnificent Seven generate disproportionate returns on invested capital (ROIC) relative to smaller firms, showcasing a fundamentals-backed concentration rather than a speculative bubble [Michael Mabousinn, 2024]. The 'Superstar Firm Hypothesis', highly relevant to this work, posits that globalization, economies of scale, and the increasing role of intangible assets allow the most dominant firms to capture an ever-growing share of economic output. Some argue that this hints towards a novel economic regime rather than a cyclical overvaluation episode, where leading firms not only maintain dominance but structurally reshape market returns [David Autor, 2020].

At the same time, concerns remain regarding the fragility of a market where a small group of firms dictate broader index movements. The proliferation of passive investing has reinforced this concentration by directing capital flows toward the largest firms, creating a self-reinforcing cycle. Studies on the rise of mega-firms have even argued that passive fund inflows have made index weightings increasingly inelastic, rendering price movements now more reflective of flows rather than traditional fundamentals [Hao Jiang, 2024].

2.2 β Decomposition and Market Sensitivity

One of the fundamental tools used to understand concentration risk is beta decomposition, which examines how different market segments contribute to the overall sensitivity of the market. Traditionally, the Capital Asset Pricing Model (CAPM) tacitly assumes that beta remains stable over time [Sharpe, 1964]. Recent empirical findings have come to challenge this assumption, particularly in the context of highly concentrated indices.

The Fama French multifactor model [Fama and French, 1993] expands on CAPM by identifying additional risk factors, including size (SMB) and value (HML). Their framework has been instrumental in explaining excess returns in different market regimes. However, recent findings suggest that these traditional factors are becoming less explanatory in an era of increasing concentration. As the largest stocks dominate market returns, the market beta appears to be increasingly skewed toward the performance of a few firms rather than broader economic factors [Michael Mabousinn, 2024].

Empirical research has found that the S&P 500's sensitivity to the Magnificent Seven has increased significantly, with these stocks now accounting for a dominant share of index movements despite comprising less than 2% of the total firms in the index. This phenomenon has led to a growing divergence between top-heavy beta (driven by large firms) and broad-market beta (which includes smaller firms). Understanding this decomposition is critical for risk management and trading strategies, as periods of high concentration can lead to misleading signals about market risk [Michael Mabousinn, 2024].

2.3 Volatility Spillovers and Market Stability

Another crucial dimension of market concentration is its effect on volatility transmission and spillover effects. Historically, research on volatility clustering has shown that large firms disproportionately impact market volatility, particularly during macro-downturns. The Dynamic Conditional Correlation GARCH (DCC-GARCH) model has been widely used to analyze time-varying volatility relationships, demonstrating that concentration amplifies systematic risk through higher correlation regimes [Engle, 2002].

Empirical studies have found that the volatility of the Magnificent Seven has become increasingly correlated with overall market volatility, suggesting that these firms have effectively become a market risk factor in their own right. This is a departure from historical norms, where smaller stocks contributed to index diversification by exhibiting uncorrelated or lower-beta movements [Hao Jiang, 2024].

Moreover, research on volatility spillovers suggests that shocks originating from dominant firms propagate more broadly than in past market cycles. Large firms exhibit higher volatility persistence, meaning that price shocks in these stocks have longer-lasting effects on the broader market. This is particularly relevant for the pricing of index options, as volatility feedback loops become stronger in concentrated markets. The VIX, often referred to as the market's "fear gauge," is increasingly influenced by movements in a handful of stocks rather than broad market sentiment [Hao Jiang, 2024].

2.4 Speculative Dynamics and Bubble Detection

Although many argue that the current concentration trend is fundamentals-driven, historical precedent suggests that excessive concentration can introduce speculative dynamics. One widely used framework to detect bubbles is the Log-Periodic Power Law Singularity (LPPLS) model, which identifies unsustainable price growth through log-periodic oscillations [V. Filimonov, 2013].

Past research applying LPPLS to financial bubbles has successfully detected pre-crash conditions in the dot-com bubble (1999), the housing crisis (2007), and the cryptocurrency boom (2017) [Riza Demirer, 2018]. While market participants argue that today's concentration is different, early LPPLS signals suggest that certain characteristics of unsustainable growth may be present in highly concentrated stocks.

Importantly, while traditional bubble detection methods rely on valuation multiples, the LPPLS approach focuses on the underlying structure of price movements. This allows researchers to differentiate between fundamentally justified growth and excessive speculation. If certain stocks exhibit self-reinforcing price acceleration beyond fundamental value, it could indicate an increased probability of mean reversion in the future.

2.5 Implications for Trading Strategies

The implications of these findings extend to trading strategies, particularly in the context of the Betting Against Beta (BaB) framework [Andrea Frazzini, 2013]. Historically, low-beta stocks have outperformed high-beta stocks on a risk-adjusted basis, leading to the popularity of long-small-cap, short-large-cap trades. However, recent research suggests that this strategy has become less effective in the current market regime, as mega-cap stocks, showcasing high betas, increasingly dominate index returns [Andrea Frazzini, 2013].

3 Methodology

3.1 Beta Decomposition

To assess the influence of stock concentration on market risk, we decompose the beta of the S&P 500 index by isolating the contributions of its largest constituents. This allows us to quantify the extent to which the most dominant firms drive market movements and evaluate whether their growing market share has structurally altered the systematic risk profile of the index.

A crucial first step in this analysis is the construction of a historical approximation of the S&P 500 index, which enables us to study its evolving composition over time. We aggregate historical index constituents using data from the CRSP database [Center for Research in Security Prices (CRSP), 2024], which provides a comprehensive and survivorship-bias-free dataset of U.S. equities. The dataset includes changes in index membership, stock splits, mergers, and other corporate actions, allowing for an accurate reconstruction of past index compositions.

To validate the accuracy of our reconstructed index, we compute its correlation with the official historical S&P 500 index. The results indicate a 99.8% correlation with index levels and a 99.7% correlation with historical index returns. This high degree of alignment confirms that our reconstructed index reliably replicates the behavior of the S&P 500, ensuring a robust foundation for further decomposition.

To quantify the role of large-cap firms in driving index performance, we segment the S&P 500's returns into two components. The first group consists of the seven largest firms by market capitalization at each point in time. While this group overlaps significantly with the Magnificent Seven in recent years, its composition may have varied in earlier periods. The second group includes all remaining index constituents.

We express total market returns as a linear combination of the returns from these two groups, allowing us to determine how much of the index's overall performance is attributable to each segment. This decomposition is formalized using a regression model of the form [SPACING]

$$r_{\mathrm{SP500},t} = \alpha + \beta_1 r_{\mathrm{MAG7},t} + \beta_2 r_{\mathrm{ExMAG7},t} + \varepsilon_t \qquad (1)$$

where $r_{\mathrm{MAG7},t}$ represents the returns of the seven largest firms, $r_{\mathrm{ExMAG7},t}$ represents the returns of the remaining index constituents, and ε_t is the residual term.

To further assess the sensitivity of these groups to systematic risk, we estimate their respective betas using the Capital Asset Pricing Model (CAPM) [?]. The beta of a given segment is calculated as:

$$\beta_X = \frac{\mathrm{Cov}(r_X, r_F)}{\mathrm{Var}(r_F)} \tag{2}$$

where r_X represents the returns of the segment under consideration, and r_F denotes the returns of the full S&P 500.

Given that market concentration is dynamic, we estimate rolling-window betas over a 36-month period to capture potential time variation in market sensitivity. This approach allows us to track whether the beta of large firms has increased over time and whether their influence on index risk characteristics has intensified.

We compute three sets of beta estimates:

- The beta of the seven largest firms relative to the S&P 500, assessing their contribution to systematic risk.
- The beta of the top seven firms by market capitalization at each point in time, capturing the influence of any temporary shifts in market leadership.
- The beta of the remaining S&P 500 constituents, which serves as a benchmark for how the broader market responds to systematic risk.

3.2 Volatility Analysis

To analyze the time-varying relationship between the volatility of the Magnificent Seven (MAG7) and the broader S&P 500, we employ the Dynamic Conditional Correlation GARCH (DCC-GARCH) model [Engle, 2002]. Unlike static correlation measures, which assume constant relationships between assets, DCC-GARCH allows for dynamic adjustments in correlations over time, making it particularly suited for studying volatility spillovers and market interdependence.

The model consists of two components. First, each return series is modeled using a univariate GARCH(1,1) process:

$$h_{i,t} = \omega_i + \alpha_i \varepsilon_{i,t-1}^2 + \beta_i h_{i,t-1} \tag{3}$$

where $h_{i,t}$ represents the conditional variance, ω_i is a constant, α_i captures the reaction to past shocks, and β_i represents the persistence of volatility.

Second, the model estimates the time-varying correlation structure using a dynamic covariance matrix Q_t:

$$Q_t = (1-\theta_1-\theta_2)\bar{Q} + \theta_1(\varepsilon_{t-1}\varepsilon_{t-1}^T) + \theta_2 Q_{t-1} \tag{4}$$

where $\bar{Q}$ is the unconditional correlation matrix, and θ_1 and θ_2 govern the responsiveness of correlations to new information. The final correlation matrix is obtained as:

$$R_t = (Q_t^*)^{-1} Q_t (Q_t^*)^{-1} \tag{5}$$

This model allows us to capture how fluctuations in MAG7 volatility influence the broader index and whether increased stock concentration has altered systemic risk transmission. By examining the evolution of correlations, we assess the extent to which volatility spillovers from the largest firms affect overall market stability.

3.3 LPPLS Model Application

To identify potential speculative bubbles in the market, we apply the Log-Periodic Power Law Singularity (LPPLS) model [Riza Demirer, 2018].

3.3.1 Foundation

The Log-Periodic Power Law Singularity (LPPLS) model is used to identify and predict financial bubbles by capturing their faster-than-exponential growth. It is based on the assumption that speculative behavior in financial markets leads to positive feedback loops, where price increases attract more investors, further accelerating the rise. The model builds on the Johansen-Ledoit-Sornette (JLS) framework [Johansen et al., 1999], which classifies market participants into fundamental investors, who trade based on intrinsic value, and noise traders, who exhibit herding behaviour and reinforce price trends.

This dynamic results in an increasingly unstable market, leading to a critical transition point t_c where the probability of a crash peaks.

$$\frac{dp}{p} = \mu(t)dt + \sigma(t)dW - kdj \tag{6}$$

where $\mu(t)$ is the expected return, $\sigma(t)$ is the volatility, dW is a standard Wiener process representing random fluctuations, dj represents a discontinuous jump (a crash event), and k is the magnitude of the crash.

The crash hazard rate, which represents the probability per unit time of a market correction, is given by

$$h(t) = \alpha(t_c - t)^{m-1}(1 + \beta\cos(\omega\ln(t_c - t) - \phi)) \tag{7}$$

where t_c is the critical time when the bubble bursts, m governs the power-law acceleration of prices, ω is the log-periodic angular frequency capturing oscillatory behavior, and β and ϕ modulate the oscillations.

The hazard rate $h(t)$ reflects the interplay of herding behavior and speculative dynamics, which amplify price movements leading up to a crash.

3.3.2 Price Dynamics and the LPPLS formula

The expected log-price trajectory of an asset undergoing a bubble phase follows the log-periodic power law equation

$$\begin{aligned}\ln p(t) = A + B(t_c - t)^m \\ + C(t_c - t)^m \cos(\omega \ln(t_c - t) - \phi)\end{aligned} \tag{8}$$

where A represents a baseline trend, B and C are coefficients that control bubble intensity, the term $(t_c - t)^m$ accounts for super-exponential growth in asset prices, and the log-periodic oscillations arise from investor feedback mechanisms.

The presence of $(t_c - t)^m$ with $0 < m < 1$ ensures that prices accelerate faster than an exponential function, differentiating bubbles from normal growth. The oscillatory term reflects waves of optimism and panic, as traders respond to price movements. The estimated value of t_c signals the likely time of market collapse.

In empirical applications, the LPPLS model successfully detects past bubbles, including the Black Monday crash of 1987, the dot-com bubble of 2000, and the global financial crisis of 2008.

3.3.3 Model Calibration

Fitting the LPPLS model to financial data requires estimating its parameters using statistical techniques. The standard approach is nonlinear least squares fitting, but calibration can be challenging due to parameter sensitivity.

To improve estimation accuracy, the LPPLS equation is reformulated by introducing new variables

$$C_1 = C\cos\phi, \quad C_2 = C\sin\phi \tag{9}$$

This transforms the model into

$$\begin{aligned}\ln p(t) = A + B(t_c - t)^m + C_1(t_c - t)^m \cos(\omega \ln(t_c - t)) \\ + C_2(t_c - t)^m \sin(\omega \ln(t_c - t))\end{aligned} \tag{10}$$

where the phase parameter ϕ is eliminated, reducing complexity. The number of nonlinear parameters decreases from four to three, making optimization more stable.

To ensure meaningful estimates, parameter values are constrained based on empirical observations

$$\begin{aligned} 0.01 \leq m \leq 0.99, \\ 6 \leq \omega \leq 13, \\ B < 0 \end{aligned} \tag{11}$$

where $m < 1$ ensures faster-than-exponential growth, ω is restricted to avoid overfitting to noise, and $B < 0$ enforces bubble behavior.

The estimation process consists of two steps. First, linear parameter estimation is performed by solving for A, B, C_1, C_2 using ordinary least squares. Second, nonlinear optimization is used to estimate t_c, m, ω using global search algorithms such as genetic algorithms, taboo search, and Markov switching models for identifying bubble phases.

To improve predictive power, multi-scale indicators are constructed by applying the LPPLS model to different time windows. The short-term indicator considers 30 to 90 days, the medium-term indicator spans 91 to 300 days, and the long-term indicator covers 301 to 750 days. The formulation for the calculation of the indicators is

$$I(t) = \frac{1}{N} \sum_{i=1}^{N} 1(t_c^{(i)} - t \leq \Delta t) \tag{12}$$

where $I(t)$ is the multi-scale LPPLS confidence indicator, representing the probability that a bubble is present at time t. The summation runs over N different rolling windows, each indexed by i. The term $t_c^{(i)}$ is the estimated critical time from the LPPLS model within the i-th rolling window, indicating the point where a bubble is expected to burst. The difference $t_c^{(i)} - t$ measures the time remaining until the estimated bubble collapse. The parameter Δt is a predefined forecast horizon, determining how far into the future the model expects a potential crash. The indicator function $1(\cdot)$ equals 1 if $t_c^{(i)} - t \leq \Delta t$, meaning the estimated critical time falls within the forecast horizon, and 0 otherwise. The fraction aggregates these signals across all rolling windows, providing a confidence level for the presence of a speculative bubble. This approach allows detection of bubbles at multiple time scales, making it useful for market monitoring.

In our analysis, we used daily adjusted closing prices for the S&P 500 index (ĜSPC), covering a period ranging from January 1, 1990, to February 25, 2025. The dataset is preprocessed by computing the natural logarithm of prices to capture exponential growth trends. To detect bubble regimes, we employed a multi-scale rolling estimation approach, setting a 120-day window size with a 60-day minimum window constraint. The model parameters were estimated using an iterative optimization procedure, with a maximum of 25 independent searches to ensure robustness. The LPPLS confidence indicator was subsequently computed to evaluate the probability of an ongoing bubble phase, providing a quantitative measure of speculative pressure within the index.

3.4 Market Crash Response Estimation: The Mag 7 in the time of COVID-19

To analyze the response of the Magnificent Seven to the COVID-19 market crash, we employ a combination of return analysis, correlation dynamics, and time-series modeling to quantify their behavior under extreme market conditions.

Daily adjusted closing prices for each of the seven firms, along with the S&P 500 index, are obtained for the period spanning January 2020 to June 2020. This window is selected to capture both the pre-crash period, the market drawdown between February and March, and the initial recovery phase following March 23, 2020. To examine return patterns, we computed daily logarithmic returns and rolling pairwise correlation coefficients between each of the Magnificent Seven and the S&P 500, using a 250-day rolling window:

$$\rho_{i,m}(t) = \frac{\mathrm{Cov}(r_{i,t}, r_{m,t})}{\sigma_{i,t}\sigma_{m,t}} \tag{13}$$

where $r_{i,t}$ and $r_{m,t}$ are the daily returns of stock i and the market index, respectively, and σ represents their rolling standard deviations. The evolution of these correlations provides insight into whether the Magnificent Seven became more or less synchronized with the broader market during the crisis.

To analyze the propagation of shocks, a vector autoregression (VAR) model [Sims, 1980] is estimated on the return series. The VAR framework allows for the examination of dynamic interactions between the Magnificent Seven and the S&P 500. The model takes the form:

$$R_t = A + \sum_{j=1}^{p} B_j R_{t-j} + \varepsilon_t \tag{14}$$

where R_t is the vector of stock returns at time t, A is a vector of intercepts, B_j represents the lagged coefficient matrices up to lag order p, and ε_t is the error term. The lag length p is selected based on the Akaike Information Criterion (AIC) [Akaike, 1973] to optimize model fit versus over-determination.

To evaluate spillover effects from the Magnificent Seven, impulse response functions (IRFs) from the estimated VAR model measure how shocks to one stock influence others over time. The magnitude and duration of these responses reveal which firms exert the greatest systemic influence during crises.

Finally, a cumulative return analysis is performed to compare the recovery speeds of different stocks.

The combination of these methods provides a rigorous framework to quantify the behavior of the Magnificent Seven during the COVID-19 market crash, allowing for an empirical assessment of their role in shock absorption, systemic influence, and market leadership during the recovery phase.

4 Results and Analysis

4.1 Beta Decomposition and Correlation Results

Our decomposition results indicate that while the Magnificent Seven account for over 35% of the S&P 500's total market capitalization, their direct contribution to index returns is notably lower, at 28%, as shown in Fig 1.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:     sp500_total_return   R-squared:                       0.996
Model:                            OLS   Adj. R-squared:                  0.996
Method:                 Least Squares   F-statistic:                 6.672e+04
Date:                Tue, 25 Feb 2025   Prob (F-statistic):               0.00
Time:                        09:20:42   Log-Likelihood:                 3109.5
No. Observations:                 502   AIC:                            -6213.
Df Residuals:                     499   BIC:                            -6200.
Df Model:                           2
Covariance Type:            nonrobust
=======================================================================================
                          coef    std err          t      P>|t|      [0.025      0.975]
---------------------------------------------------------------------------------------
const                  -0.0003   2.24e-05    -15.588      0.000      -0.000      -0.000
mcw_return_top_7        0.2807      0.002    141.569      0.000       0.277       0.285
mcw_return_ex_top_7     0.7137      0.004    192.776      0.000       0.706       0.721
==============================================================================
Omnibus:                      176.023   Durbin-Watson:                   1.837
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              793.154
Skew:                          -1.504   Prob(JB):                    5.87e-173
Kurtosis:                       8.373   Cond. No.                         177.
==============================================================================
```

Figure 1: S&P500 Beta decomposition summary

At first glance, this may seem counterintuitive—given their outsized presence, one might expect a larger share of total returns to be driven by these stocks. However, it is important to recognize that the regression estimates the additional explanatory power of the Magnificent Seven beyond what is already captured by the rest of the index. This means that if these stocks move closely in line with the broader market, much of their influence is already reflected in the returns of other constituents, thereby reducing their standalone contribution in the regression. In other words, the model does not measure how much these stocks contribute mechanically due to their weight in the index but rather isolates the unique effect they have on overall returns.

Additionally, beta decomposition is influenced not only by market capitalization but also by the volatility and independent variation of each segment. If a stock or group of stocks exhibits high correlation with the rest of the market, its distinct contribution to index returns appears lower when assessed through regression analysis. The Magnificent 7 (MAG7) have a beta of 1.105, indicating higher market sensitivity. The Top 7 stocks have a slightly lower beta of 1.023, while the Ex-Top7 group, representing the rest of the S&P 500, has a beta of 0.9434, reflecting lower volatility. These values highlight the elevated market influence of mega-cap stocks.

We extend our analysis by computing and visualizing (Fig. 2) the quarterly rolling betas as a function of market capitalization, distinguishing between the top seven S&P 500 constituents and the remaining 493 companies. The results reveal a clear pattern: the large-cap group consistently exhibits higher betas compared to the rest of the index.

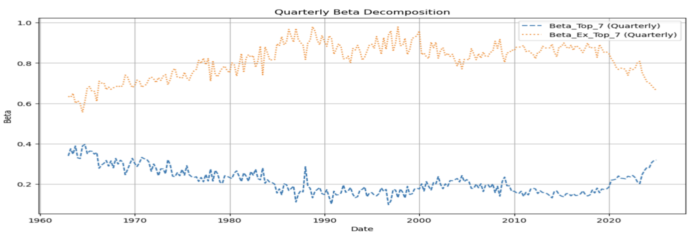


Figure 2: S&P 500 drivers decomposition

We proceed to calculate the correlation between the Top 7 group, the following 493, the Mag 7 and the Russel 2000 index, a proxy for small cap companies. The Top 7 stocks show a strong correlation (0.832) with the rest of the index, driving market performance. However, the Magnificent Seven exhibit a lower correlation of 0.526 with the broader market, suggesting that their movements are more independent, likely due to sectoral concentration and firm-specific factors.

4.2 Market Concentration and Volatility

Applying the Dynamic Conditional Correlation GARCH model confirms that the volatility of the Magnificent Seven is increasingly correlated with overall market volatility, indicating a shift in systemic risk contributions. The results of this analysis are shown in Fig. **??**:

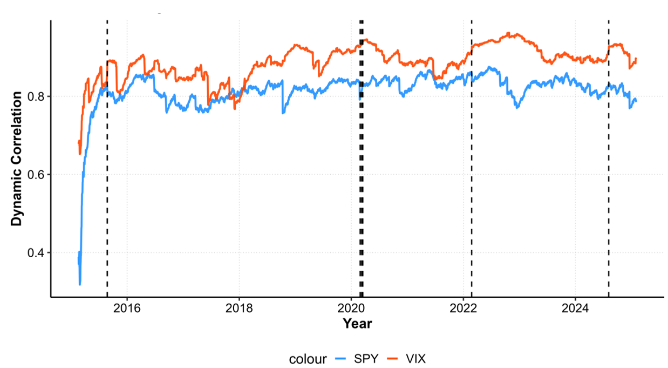


Figure 3: DCC between MAG7 and S&P 500

Now, we proceed to examine the potential of spillover effects, or transmission of volatility from one asset to another.

In Fig**??**, the upper subplot examines the impact of volatility shocks at the Magnificent 7 (M7) stocks on the VIX index. A clear positive correlation is observed between pre-shock and post-shock VIX volatility, as indicated by the clustering of data points above the reference dashed line. This represents the hypothetical scenario where post-shock volatility equals pre-shock volatility. The upward deviation of the majority of data points from this reference suggests that shocks originating from Magnificent 7 stocks contribute to heightened implied market volatility. This finding aligns with the notion that volatility in major market components can propagate into broader market expectations, as captured by the VIX.

In the lower subplot, we have attempted to capture the same dynamics between Magnificent 7 stock volatility shocks and the volatility of the SPDR S&P 500 ETF (SPY). Similar to the VIX, the scatter plot demonstrates a positive relationship between pre- and post-shock SPY volatility, albeit with a less pronounced slope compared to the VIX. This suggests that, while Magnificent 7 volatility shocks influence the overall market, the transmission of volatility to SPY is relatively weaker than to the VIX. Given that SPY reflects realized market volatility, whereas VIX captures forward-looking implied volatility, the findings imply that Magnificent 7 shocks have a more immediate and pronounced impact on market expectations than on realized price fluctuations.

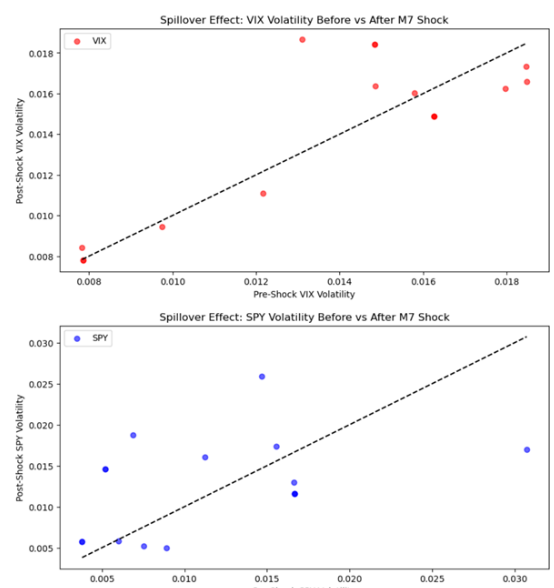


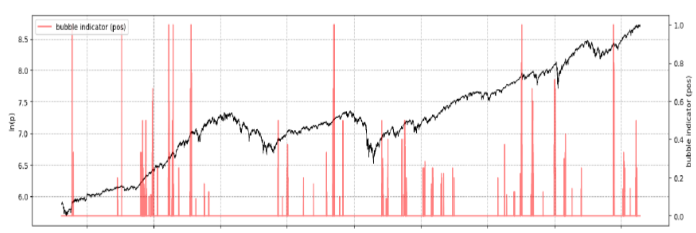
Figure 4: DCC between MAG7 and S&P 500

The patterns observed in our scatter plots resonate strongly with the findings of Jiang et al. (2024), reinforcing the idea that the largest firms no longer serve as stabilizers but rather as amplifiers of systemic volatility. The spillover effects captured in our charts, particularly the pronounced shift in VIX volatility post-shock, suggest that volatility originating in the Magnificent 7 stocks propagates unevenly, with a stronger transmission to implied market volatility than to realized price fluctuations in SPY.

This aligns with the argument that passive flows have reshaped market dynamics, concentrating risk in the largest firms [Hao Jiang, 2024], and that passive investing drives a feedback loop where inflows into index funds elevate the market capitalization of a few dominant firms, reinforcing their price momentum while simultaneously amplifying their volatility. Our results seem to support this conclusion—volatility spillovers that were once more evenly distributed now appear to be disproportionately driven by these mega-cap stocks, challenging the traditional assumption that firm size inversely correlates with volatility. The asymmetry between the VIX and SPY spillover effects further strengthens this narrative, indicating that investors' expectations of future risk are more reactive to Magnificent 7 volatility shocks than actual realized market movements. In essence, what we see in the data is a reflection of a structural shift in modern markets, where passive flows not only concentrate returns but also volatility, pushing the largest firms into a central role as key transmitters of systemic risk.

4.3 LPPLS Model Results

The Log-Periodic Power Law Singularity (LPPLS) model was applied to historical S&P 500 data from 1990 onward to assess its predictive accuracy in identifying speculative bubbles. By analyzing past market cycles, we evaluate whether the model effectively signals periods of unsustainable price growth that preceded major corrections. The corresponding results, signaling both positive and negative (aggressive correction) bubble formation, may be found below:



(a) LPPLS positive bubble indicator signals.

(b) LPPLS negative bubble indicator signals.

Figure 5: Bubble indicator signals based on LPPLS model. The positive indicator (a) highlights potential speculative periods, while the negative indicator (b) detects mean-reverting behavior.

Fig. 5 displays the LPPLS confidence indicators, which are measures of the existence of a speculative bubble probability as a function of time. The top panel (positive indicator) highlights historical periods where the model detected strong speculative dynamics, while the bottom panel (negative indicator) identifies phases of mean reversion. Notably, the model issued strong positive signals before major downturns, including the dot-com crash (2000-2002) and the global financial crisis (2008-2009). The recurrence of similar signals in recent years, particularly post-2020, suggests that the model captures structural patterns associated with speculative excess. This is a positive indicator towards the model's success in capturing them.

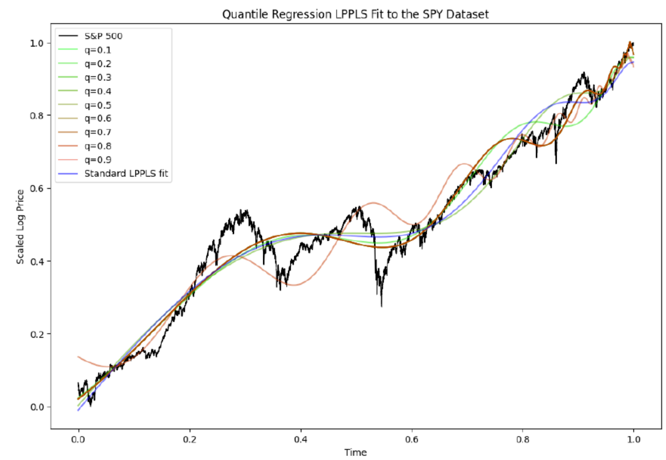


Figure 6: Quantile regression LPPLS fit to the S&P 500. The plot shows multiple quantile fits capturing price oscillations and instability over time.

Fig. 6 applies a quantile regression to the LPPLS model, capturing a range of possible price trajectories rather than a single expected path. By fitting multiple curves at different quantiles (q = 0.1 to q = 0.9), this approach accounts for variations in speculative intensity. Lower quantiles represent more conservative price paths, where speculative effects remain limited, while upper quantiles capture extreme log-periodic oscillations that precede market corrections.

The divergence between quantiles highlights increasing uncertainty in speculative phases, particularly in historical periods leading up to major downturns. The upper quantiles exhibit stronger oscillations before known market crashes, reinforcing the model's ability to identify speculative excess. In the current market, these high-quantile trajectories suggest that concentration and passive investing may be amplifying price instability, increasing the likelihood of a self-reinforcing feedback loop. While this does not imply an imminent correction, the pattern resembles conditions observed before past downturns, warranting closer attention to systemic risk. The model's success at predicting previous corrections, alongside its signals for the present, could be interpreted as a sign of caution.

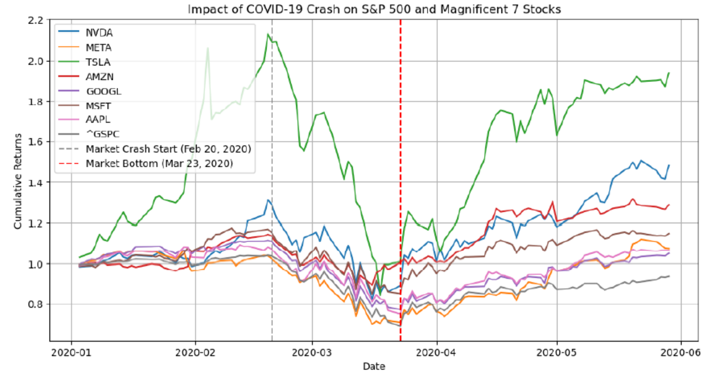
4.4 Response to shocks

After examining the possibility of being in a speculative bubble, we employ empirical results to identify the response of the Magnificent 7 to the COVID - 19 crash, which offers us a natural experiment to assess how these companies behave in times of market distress.

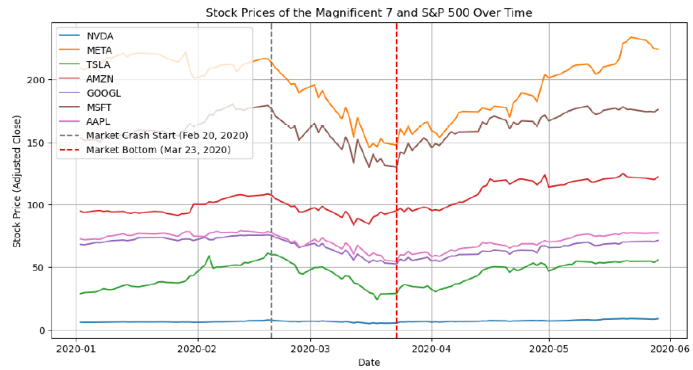
As shown in Figure 7a, the S&P 500 and the Magnificent 7 stocks experienced synchronized drawdowns during the February–March 2020 crash, confirming their role as dominant market drivers. Recovery dynamics diverged significantly, with Nvidia and Tesla rebounding at an accelerated pace, outpacing both the broader index and other tech leaders. This suggests that while all Magnificent 7 stocks are critical in downturns, some stocks exhibit higher sensitivity to liquidity-driven recoveries.

Amazon and Microsoft, in contrast, showed more stable recovery paths, likely due to their fundamental resilience amid increased reliance on e-commerce and cloud computing. This divergence indicates that while the Magnificent 7 collectively influence market-wide movements, the nature of their response varies based on business model and market positioning.

During the crisis, the correlation among Magnificent 7 stocks surged, reinforcing the hypothesis that market drawdowns are increasingly dictated by a small subset of firms. This is evident in Figure 7b, where the simultaneous decline in all Magnificent 7 stocks reflects a breakdown in diversification benefits. Interestingly, Tesla and Nvidia exhibited lower correlation with the rest of the group compared to Apple, Microsoft, and Alphabet. This suggests that while Tesla and Nvidia contribute to speculative volatility, systemic risk primarily flows through the largest index-weighted firms



(a) Cumulative returns of the Magnificent 7 and the S&P 500 during the COVID-19 market crash (February–April 2020).



(b) Stock prices of the Magnificent 7 and S&P 500 over time, illustrating long-term trends.

Figure 7: Market performance of the Magnificent 7 and the S&P 500, highlighting both short-term crisis response and long-term price evolution.

5 Trading Strategy and Regime-Based Allocation

Building on the insights derived from our research, we analyze the viability of a regime-based trading strategy that enhances the traditional long small-cap, short large-cap approach. Our results, in conjunction with prior work such as [Michael Mabousinn, 2024], suggest a cyclicality in market dominance, wherein large-cap stocks periodically outperform small-cap stocks before a reversion occurs. Given this observed cyclicality, we propose an enhanced strategy that incorporates market structure indicators to identify regime shifts and optimize trade allocations accordingly.

5.1 Regime Identification and Market Structure Dynamics

The conventional long small-cap, short large-cap strategy assumes persistent small-cap outperformance but is vulnerable to two primary risks: high small-cap volatility leading to drawdowns and periods of large-cap dominance causing underperformance on the short leg. The increasing concentration of the Magnificent Seven (M7) has amplified both risks, necessitating a more dynamic approach.

We define three distinct market regimes. Regime 0 represents market distress, where systemic risk is high, and investors shift to safety, historically favoring M7 stocks. Regime 1 occurs when large caps, particularly M7, drive market returns, outperforming small caps. Regime 2 is characterized by small-cap outperformance, where smaller stocks regain leadership.

Regime transitions are driven by volatility trends, speculative excess, and liquidity cycles. To identify these shifts, we construct a feature set incorporating EGARCH-based volatility forecasts, composite momentum measures, risk-free rates, implied volatility (VIX), LPPLS-based bubble confidence, M7 concentration, and lagged small-cap versus large-cap returns. These inputs are used in regime detection models to classify market conditions in real time.

5.2 Trading Strategy Implementation

The strategy dynamically adjusts allocations based on prevailing market conditions. In Regime 0, characterized by systemic risk and market distress, historical data suggests that the M7 stocks provide relative stability.

During these periods, the strategy adopts a defensive stance by going long on the market cap-weighted returns of the M7 while shorting the rest of the S&P 500. This position is maintained until the speculative confidence indicator declines below a critical threshold, signaling reduced crash risk, at which point the strategy reverts to its standard regime-based allocation.

In Regime 1, where large-cap stocks dominate market performance, the strategy shifts to a long position in the S&P 500 total return while shorting small-cap equities, represented by the Russell 2000 ETF (IWM). This allocation capitalizes on large-cap strength while hedging against small-cap underperformance.

In Regime 2, where small caps regain leadership, the strategy implements the traditional long small-cap, short large-cap approach by taking a long position in small-cap equities while shorting the S&P 500 total return. This positioning remains in place until market indicators signal a shift back to large-cap dominance or heightened systemic risk.

5.3 Regime Detection and Model Performance

To classify market regimes, we employ Hidden Markov Models (HMMs), Bayesian Gaussian Mixture Models (BGMMs), and Long Short-Term Memory (LSTM) networks. Our approach involves segmenting historical market data and evaluating feature combinations based on their ability to maximize the Sharpe ratio and minimize drawdowns in backtesting.

Across all models, lagged small-cap vs. large-cap returns, top-seven concentration, VIX, and vuvvle confidence consistently emerged as key predictive variables. While these factors indicate regime shifts in hindsight, real-time classification proved unreliable.

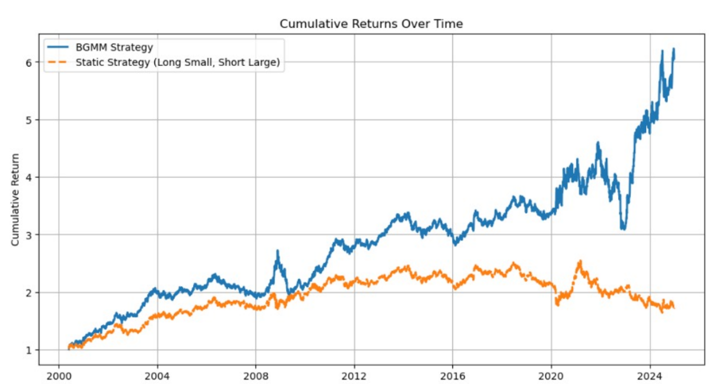


Figure 8: Cumulative returns over time for BGMM regime based strategy with optimal features

5.4 Limitations

A major limitation of this study is the uncertainty in determining whether the Magnificent Seven's dominance

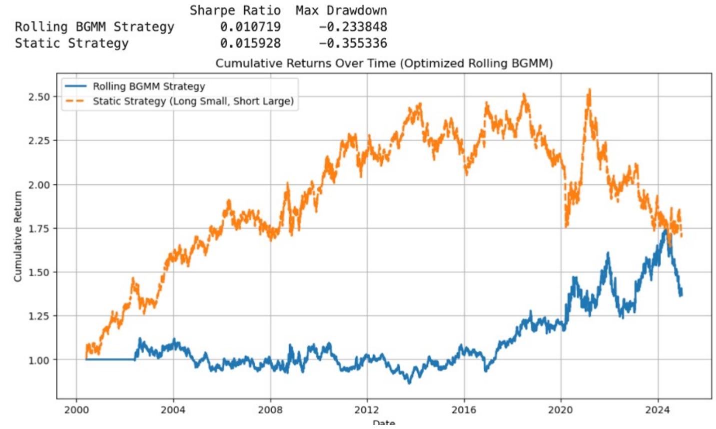


Figure 9: Cumulative returns over time and performance metrics for Rolling BGMM regime based strategy with the same features

is cyclical or a permanent structural shift. The LPPLS model, while effective in detecting speculative bubbles, is prone to false positives, leading to unnecessary portfolio rebalancing and increased transaction costs. Additionally, regime shifts are influenced by non-financial factors such as political events, regulatory changes, and technological disruptions, which are difficult to quantify. Structural changes in market dynamics, including passive investing and algorithmic trading, may also weaken historical patterns of small-cap versus large-cap dominance, reducing the reliability of past trends for future predictions

References

[Akaike, 1973] Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In Petrov, B. N. and Csaki, F., editors, *Proceedings of the 2nd International Symposium on Information Theory*, pages 267–281, Budapest. Akadémiai Kiadó.

[Andrea Frazzini, 2013] Andrea Frazzini, L. P. (2013). Betting Against Beta. *CDAR Berkley*.

[Center for Research in Security Prices (CRSP), 2024] Center for Research in Security Prices (CRSP) (2024). CRSP U.S. Stock Databases.

[David Autor, 2020] David Autor, David Dorn, L. K. C. P. J. R. (2020). The Fall of the Labor Share and the Rise of Superstar Firms. *Quarterly Jounal of Economics*.

[Engle, 2002] Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate garch models. *Journal of Business Economic Statistics*, 20(3):339–350.

[Fama and French, 1993] Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56.

[Hao Jiang, 2024] Hao Jiang, Dimitri Vayanos, L. Z. (2024). Passive investing and the rise of mega-frism. *SSRN*.

[Johansen et al., 1999] Johansen, A., Sornette, D., and Ledoit, O. (1999). Predicting financial crashes using discrete scale invariance. *Journal of Risk*, 1(4):5–32.

[Michael Mabousinn, 2024] Michael Mabousinn, D. C. (2024). Stock Market Concentration: How Much Is Too Much. *Morgan Stanley*.

[Riza Demirer, 2018] Riza Demirer, Guilherme Demos, R. G. D. S. (2018). On the predictability of stock market bubbles. *Routledge*.

[Sharpe, 1964] Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19(3):425–442.

[Sims, 1980] Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, 48(1):1–48.

[V. Filimonov, 2013] V. Filimonov, D. S. (2013). A stable and robust calibration scheme of the log-periodic power law model. *ElSevier*.