Linköping University | Department of Management and Engineering
Master's Thesis in Economics
Master's Programme in Business and Economics
Spring 2025|LIU-IEI-FIL-A--25/04901--SE

Convexity in Motion

Leveraging Gamma Exposure to Predict Equity Market Returns and Improve Predictive Modeling

Gustav Jonsson
Theo Nyberg

Supervisor: Axel Hedström

Linköping University
SE-581 83 Linköping, Sweden
+46 013 28 10 00, www.liu.se

Acknowledgements

We would like to express our sincere gratitude to our supervisor Axel Hedström, PhD for his continuous support throughout the writing process and for valuable discussions.

We also wish to extend our thanks to Gustav Fyring, our supervisor at the Third Swedish National Pension Fund (AP3), for his guidance and feedback. Furthermore, we are grateful to AP3 and CIO Jonas Thulin for making this collaboration possible and for their engagement and support.

Finally, we would like to thank one another for a constructive and rewarding collaboration. We are grateful for the enriching discussions, mutual learning, and complementary contributions that have shaped the development of this thesis.

Linköping, 26 May 2025

Gustav Jonsson

Theo Nyberg

Abstract

This thesis investigates whether aggregate gamma exposure (GEX) in the S&P 500 index options market contains predictive information about future equity returns and whether its inclusion can enhance short-term forecasting models. Motivated by the rapid growth of the options market and the increasing influence of delta-hedging flows on equity price dynamics, the study examines the relationship between changes in GEX and subsequent movements in the S&P 500 index. Using daily data from 2011 to 2025, we estimate an Autoregressive Distributed Lag (ARDL) model to capture dynamic short- and long-run effects and construct two predictive models: one including GEX and one excluding it. These models are evaluated against each other and a random walk benchmark using out-of-sample, rolling-window forecasts and Diebold-Mariano tests. Our results show that changes in GEX are significantly and positively associated with S&P 500 returns across multiple time horizons. Moreover, incorporating GEX significantly improves forecast accuracy compared to both the GEX-excluding model and the benchmark. These findings suggest that gamma-related hedging activity by options market makers introduces mechanically induced flows with directional implications for short-term equity market returns. The study contributes to the literature by quantifying this relationship at the index level for the following day(s) and offers practical insights for asset managers seeking to integrate market microstructure variables into predictive frameworks.

Table of Contents

1. Introduction

Over the past few years, the size of the U.S. options market has expanded significantly, with 2024 marking the fifth consecutive year of strong growth. In 2019, approximately 4.9 billion options contracts we traded (OCC, 2020), and by 2024, the volume had exceeded 12 billion (OCC, 2025). Based on these numbers, the options market size has grown by nearly 150% since 2019. At the same time, according to SIFMA (2024), the U.S. equity market measured by average trading volume has grown by approximately 74%. Given the rapid growth of the options market relative to the equity market, our primary focus is to evaluate the potential impact of increased options trading on the stock market. In particular, we examine how option gamma, the rate of change in an option's price as delta shifts, influences short-term movements in the stock market.

Before we do so, understanding the key drivers behind the rapid expansion of the options market is essential. Improved infrastructure for options trading have made the U.S. options market more accessible to a broader range of participants. Specialized market-making firms along with a Payment for Order Flow (PFOF) structure have contributed to the rise of retail investment platforms (Ingrassia, 2021). By simplifying access to trading, these platforms have enabled millions of new investors to participate in both the equity and options market. Bryzgalova, Pavlova and Sikorskaya (2023) highlight that the growth of the options market has been particularly driven by retail investors who now account for over 60% of total options trading. This increase coincides with improvements in market infrastructure where commission-fee trading and PFOF have facilitated greater participation.

New regulatory frameworks such as Solvency II and Dodd-Frank have further increased incentives for institutions such as insurance companies and pension funds to utilize derivatives for risk management purposes. Barbon and Buraschi (2021) highlight the growing participation of the insurance sector in options trading, arguing that this development has introduced structural flows into the options market that did not previously exist. As the options market expands, driven by improved infrastructure, increased participation and the emergence of structural flows, liquidity in the options market improves. Moreover, access to U.S. options markets has become more accessible for international investors where Cboe (2021) played a key role in this development by introducing "Global Trading Hours" during 2021.

Additionally, a driving force of the industry's rapid growth in recent years has been the increased interest in short-term options contracts known as Zero Days to Expiration (0DTE) contracts. According to Cboe (Xu, 2023), the share of 0DTE contracts on the S&P 500 (SPX) has risen from 5% in 2016 to over 40% in 2023, and at times, these contracts have accounted for more than half of SPX options trading volume. A similar trend has been observed by the New York Stock Exchange (Poser, 2023) for both overall options trading and the increasing dominance of short-term options contracts. Due to their extremely short time to expiry, 0DTE options tend to be concentrated near at-the-money strikes, where gamma is at its highest.

Alongside the rapid growth of the options market, its role in driving stock price movements had gained increasing attention, particularly in the context of “gamma squeezes” where large options positions amplify price swings in the equity market. The most notable example occurred in 2021 when the share price of GameStop increased by over 2.800% in a short period of time. While part of this increase can be attributed to a “short squeeze’’, it is notable that a similar price movement, a second surge, occurred the following month in the same stock despite significantly lower short interest compared to the initial spike. On February 25, 2021, the trading volume in GameStop amounted to 184.7 million shares traded which was four times the number of outstanding shares. Over a three-day period, total volume reached 419.0 million shares traded, almost ten times the public float (Calhoun, 2021). In both instances, call option volume spiked to extreme levels, and in the second rally, a short squeeze was not the primary driver behind the sharp price increase. In an article for Forbes, Calhoun (2021) analyzed what he termed a ”gamma swarm”, illustrating minute-by-minute order flows that clearly shows how massive inflows into call options were followed within minutes by significant buying pressure in the underlying stock.

This study examines the impact of option trading flows, the increasing volume of options activity, and its influence on the stock market. Despite the rapid expansion of the options market and its documented effects on the equity markets, such as during option expirations, academic research on gamma’s impact remains limited. Much of the existing literature focuses on gamma’s role in market volatility or the mechanics of gamma and delta hedging. However, there is a lack of comprehensive systematic research that quantifies how gamma exposure influences market movements over the following day(s). By examining the relationship between changes in gamma exposure and subsequent stock price movements, and by employing GEX as a proxy for aggregate S&P 500 gamma exposure, this study contributes to the existing literature while offering practical insights. While several option greeks[1] play key roles in options pricing and risk management, gamma stands out as particularly relevant when analyzing how option market dynamics may influence the underlying asset. Gamma captures the convexity of an option’s value in relation to changes in the price of the underlying asset, and more importantly, it determines how frequently and aggressively delta must be adjusted in response to price movements. Since options market makers (OMMs) typically seek to remain delta-neutral, a position with high gamma necessitates continuous rebalancing of the underlying asset as prices fluctuate. Delta is a measure of directional exposure and drive this kind of reactive hedging.

[1] Option Greeks are partial derivatives of an option’s value with respect to key underlying variables, such as the price of the underlying asset, time, volatility and others. They quantify how sensitive the option’s price is to small changes in these inputs.

Thus, gamma is of particular interest because it directly links the positioning of OMMs to mechanically induced trading flows in the underlying equity market. These flows may not only influence short-term price dynamics but also serve as signals that shape the expectations and behaviors of other market participants, illustrating a reflexive feedback loop between derivative markets and stock prices.

As markets become increasingly driven by capital flows such as passive investments and quantitative strategies (CNBC, 2017; Morningstar, 2024), fueled in part by the expanding options market, new market dynamics may emerge. Traditional indicators may lose their explanatory power while new phenomena take shape. For professional asset managers and financial market participants it is crucial to adapt to this evolving market structure, develop new ways of analyzing market behavior and identify actionable signals to generate alpha. Our study's contribution to this field is to examine whether aggregated gamma exposure can help explain future stock market movements and by doing so, identify actionable market signals or modeling setups that can be utilized within a more tactical capital allocation framework.

Based on these considerations, this study sets out to answer the following research questions:

I. How does aggregated gamma exposure influence stock market movements before and after the options market expansion, and are there observable differences in these effects?

II. Does the inclusion of gamma exposure in predictive models improve the forecast accuracy of S&P 500 returns?

To examine how gamma exposure (GEX) may influence market movements and direction, we use the S&P 500 as a proxy for the stock market and an ARDL framework for the analysis. Further, we construct two predictive models based upon the ARDL structure which differs in the inclusions of GEX, which are compared against one another and to a benchmark random walk model. Our analysis focuses on the index level rather than individual stocks to identify actionable market signals or modeling frameworks that can be applied within active capital allocation strategies. Since GEX is specifically calculated for the S&P 500, it serves as a natural and relevant foundation for our analysis. Moreover, the S&P 500 is one of the most comprehensive and liquid indices, widely used in both academic research and professional investing as a measure of market dynamics, ensuring that our findings can be compared to previous studies and have broader applicability.

Our results clearly indicate that variations in the derivative of GEX have a statistically significant relationship with subsequent returns on the S&P 500. This effect is robust across both the pre and post options market expansion subperiods, albeit with somewhat diminished strength in the latter. Based on out-of-sample testing we find that incorporating GEX significantly enhances forecasting performance relative to both a GEX-excluding specification and a random walk benchmark.

2. Theoretical Framework

2.1 Delta and Gamma

To understand the natural relationship between options and their underlying asset price, we must first comprehend the most relevant option greeks. Delta represents how much the price of an option changes in relation to a price change in the underlying asset, i.e., how sensitive the option's price is to fluctuations in the underlying asset. Delta (Δ) is thus the first derivative of the option price (V) with respect to the price of the underlying asset (S):

$$\Delta = \frac{\partial V}{\partial S} \tag{1}$$

Gamma on the other hand, is usually explained as "the delta of delta," meaning how delta changes in response to a price movement in the underlying asset. In other words, gamma represents the rate of change. If you are driving a car at 30 km/h and increase your speed to 40 km/h, your speed has increased by 10 km/h. If we think of speed as delta, then gamma is the change in speed, i.e., the acceleration (CME Group, 2025). Gamma (Γ) is therefore the second derivative of the option price (V) with respect to a price change in the underlying asset (S):

$$\Gamma = \frac{\partial^2 V}{\partial S^2} = \frac{\partial \Delta}{\partial S} \tag{2}$$

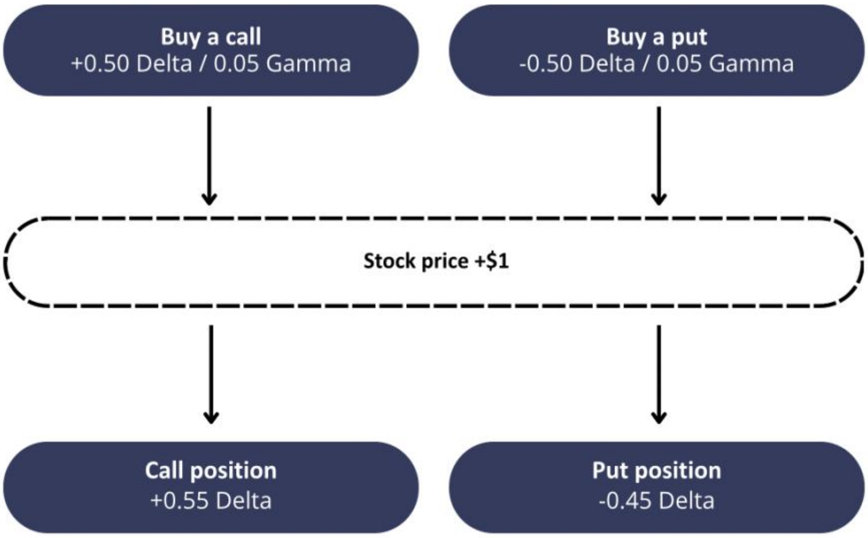
When trading options, one must think in three dimensions: the underlying price can and will change (Δ), time will pass (θ)[2], and volatility will fluctuate over the life of the options contract (ν)[3]. All these factors influence the value of an option. This study focuses on the price factor, i.e., the delta and gamma of an option, to understand directional exposure and related flows from OMMs. When OMMs quotes prices and buy or sell an options contract, they automatically take the opposite side of the trade from the buyer's or seller's perspective. This exposes OMMs to risks related to both the instrument that was bought or sold and the underlying asset. OMMs must therefore hedge in response to market movements to neutralize the risks that arise in their portfolios.

---

[2] Theta (θ) measures the rate at which an option loses value as time passes, holding all else constant.
[3] Vega (ν) measures how much an options value changes in response to a change in the implied volatility of the underlying asset, holding all else constant.

Gamma and delta hedging are therefore risk management strategies aimed at neutralizing the delta risk. This involves dynamically adjusting the delta hedge to manage the nonlinear price movements of an options portfolio. Fundamentally, gamma have implications on delta hedging, where an options trader continuously adjusts their position to manage changes in the option's delta. Understanding this interaction between delta and gamma is essential to grasp how option market positioning can mechanically influence stock price dynamics. This is further explained and visualized in Figure 1.

Figure 1: Impact of gamma on delta



Note: Illustration of how delta changes when the underlying stock price increases by $1, given an initial gamma of 0.05. A purchased option gains delta, increasing from +0.50 to +0.55 for the call option, and increasing from −0.50 to −0.45 for the put option. The figure highlights the convexity effect (gamma) embedded in options.

Delta hedging entails selling or buying a specific amount of the underlying asset to compensate for changes in the option's value. An OMM who is long a call option with a delta of 0.6 is effectively exposed to 0.6 units of the underlying stock. To remain delta neutral, the OMM must sell 0.6 shares of the stock per option contract, since the stock itself always has a delta of 1. However, the problem with delta hedging alone is that delta is not constant, it changes as the price of the underlying asset moves. To address this, gamma hedging is employed, supplementing the delta hedge by using other derivative contracts (often options) to neutralize gamma exposure. This, however, creates a need for delta hedging elsewhere, while at the same time, gamma hedging reduces the necessity for frequent delta hedge adjustments.

2.2 Gamma Exposure (GEX)

SqueezeMetrics (2017) examines how the dynamics of the options market can influence the stock market and its volatility, introducing the GEX index to track aggregate gamma exposure on the S&P 500. This index provides insight into how market participants, particularly large options traders and OMMs, may need to adjust their positions based on their gamma exposure along with price movements in the underlying asset.

As mentioned above, the option delta represents the rate of change in an option's price relative to movements in the underlying asset's price. A portfolio is considered delta-neutral when the total sum of its positive and negative delta positions equals zero. Additionally, gamma measures the rate at which delta shifts in response to changes in the underlying asset's price. OMMs adjust their hedges based on gamma values to maintain delta neutrality. If a portfolio is gamma-neutral, variations in the underlying asset's price will not impact its delta.

For a specific options chain, gamma exposure (GEX) is determined by multiplying gamma (Γ) and open interest (OI) at each strike price (*k*), as described in Equation 3. If the option is a put, the value is multiplied by negative one, assuming that OMMs typically take the selling side of put options, leading to a net short gamma position. Finally, since each option contract represents 100 shares of the underlying asset, the resulting sum is multiplied by 100 to reflect this standard lot size.

$$GEX = 100 \sum_{k=1}^{n} OI_k \Gamma_k 1(O) \tag{3}$$

$$Where \;\; 1(O) = \begin{cases} -1 \; if \; Put \\ 1 \; if \; Call \end{cases}$$

In this way, GEX is obtained for SPX, representing the sum of gamma exposure at each strike price for all outstanding options contracts. The aggregated gamma exposure reflects the extent to which option dealers and market participants need to adjust their hedging positions as the underlying asset moves or as open interest changes to manage delta and gamma risks in their books. The underlying dynamics can be summarized as follows: when GEX is positive (OMMs are net long gamma), the delta of the option portfolio increase following an increase in the underlying asset price, and they need to sell the underlying asset to maintain delta neutral. Reversely, if the underlying asset's price falls, they must buy to rebalance their exposure. This results in a dampening effect on the volatility of the underlying asset. Conversely, when GEX is negative, the effect is reversed, and OMMs hedging activities instead amplify the volatility of the underlying asset.

The GEX index is built upon a set of simplifying assumptions intended to make the estimation of aggregate hedging flows feasible. It assumes that all traded options are facilitated by delta-hedging participants, typically OMMs, who manage option portfolios in a risk-neutral manner. Furthermore, it presumes that call options are sold by investors and bought by OMMs, consistent with strategies such as call overwriting, while puts are mostly purchased by investors for downside protection and sold by OMMs. Finally, GEX assumes that OMMs hedge precisely to the option delta, although in practice, they apply hedging bands to balance execution costs with risk and thereby translating changes in gamma into predictable rebalancing flows in the underlying asset. What is important to remember about the GEX computation is that it is based on open interest and sum the gamma over all strike prices and expirations dates. Given that short-dated options like 0DTE options often have negligible open interest, especially early in the trading day, their gamma exposure is likely underrepresented in GEX calculations. Even though 0DTE options can account for a large portion of the trading volume, they do not leave a mark on the open interest at the end of the day since all positions are closed. Meaning, the gamma exposure as calculated by GEX might underestimate the actual gamma levels in the market.

2.3 Gamma Hedging and Delta Neutrality

As explained by Leoni (2014, pp. 86–87), gamma captures the convexity of an option's value in relation to changes in the price of the underlying asset. To better understand this relationship, a Taylor expansion can be used to approximate how the value of an option position, denoted $\pi(S)$, changes with respect to the underlying price S. The first-order approximation including only the delta (Δ) is illustrated in equation (4):

$$\pi(S_0 + dS) \simeq \pi(S_0) + \Delta(S_0) \cdot dS \tag{4}$$

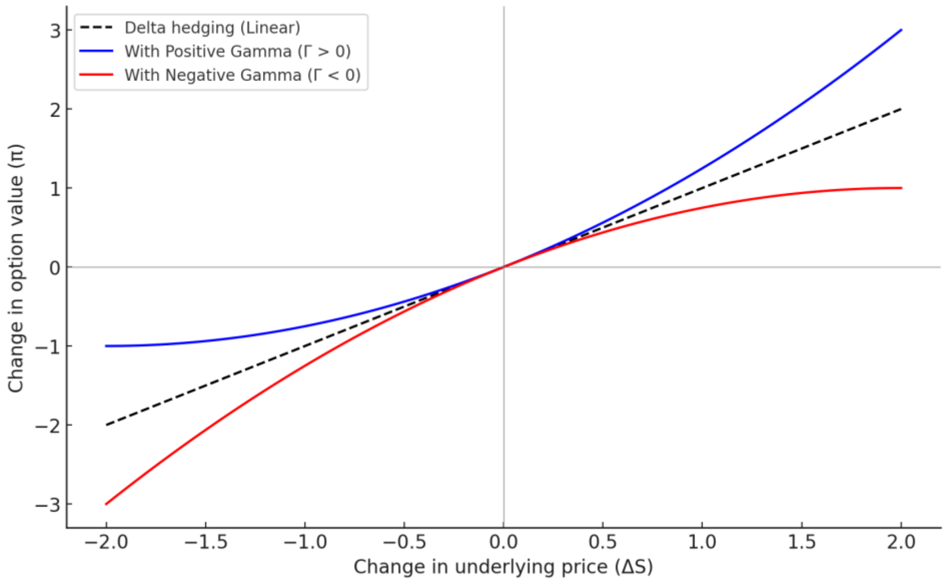
This equation provides a simple estimate of how the option's value changes in response to small movements in the underlying asset. However, it does not account for the non-linear behavior of options, especially when price movements are large, or gamma exposure is significant. To improve the approximation, a second-order term is added that includes gamma (Γ). This leads to the parabolic approximation:

$$\pi(S_0 + dS) \simeq \pi(S_0) + \Delta(S_0) \cdot dS + \frac{1}{2}\Gamma(S_0) \cdot (dS)^2 \tag{5}$$

The key insight from equation (5) is that the correction term $[\frac{1}{2}\Gamma(S_0) \cdot (dS)^2]$, depends entirely on the sign of gamma. While $(\Delta S)^2$ is always positive, the impact of this term on the option's value hinges on whether gamma is positive or negative:

When $\Gamma > 0$ (positive gamma), the correction is positive. This means the option value increases more than what the delta-only approximation predicts. This typically occurs when a trader is long options, such as having bought calls or puts. When $\Gamma < 0$ (negative gamma), the correction is negative. In this case, the option value increases less than the delta would suggest. This is common when a trader is short options, such as having sold calls or puts to e.g., collect option premiums or hedge positions.

Figure 2: Gamma convexity



Note: Illustration of the curvature in option value changes depending on the sign of gamma.

This relationship is illustrated in Figure 2, where the dashed line represents the linear approximation based on delta hedging alone (Equation 4), while the curved lines reflect the full approximation including gamma (Equation 5). The blue curve (positive gamma) bends upward, indicating increased value gain for larger price moves. The red curve (negative gamma) bends downwards, reflecting reduced gains or larger losses for the same price move. In practice, this dynamic is crucial for option traders and OMMs. A position with negative gamma is vulnerable to large moves in either direction, requiring frequent adjustments in line with market movements to maintain neutral delta. This "chasing the market" behavior leads to buying in rallies and selling in declines, amplifying volatility. On the other hand, positive gamma positions benefit from volatility and naturally offset price swings, dampening market moves. To visualize the interplay between delta and gamma, it can be explained as driving at night with only your low beams on. You can see a bit ahead, but not the full road. That's equivalent to relying solely on delta: it shows you the current direction, but not how quickly it might change. Gamma is your high beam, it reveals how the road curves further ahead. Thus, gamma is essential to manage risk exposure beyond what delta hedging alone can manage, and the illustration provides a visual representation of how the value of an option diverges from the delta-predicted path as gamma changes.

2.4 Causality and simultaneity

Establishing causal relationships within financial time series is inherently challenging due to the presence of endogeneity, where explanatory variables may be influenced by the same factors driving the dependent variable. In the context of our study, the relationship between GEX and S&P 500 returns may be subject to simultaneity, as price movements can influence option positioning, while GEX-induced hedging flows may in turn affect price dynamics. This circular feedback loop reflects a broader principle of reflexivity in financial markets, where perception and behavior are mutually reinforcing. As a result, our empirical approach focuses on identifying statistically robust associations, while acknowledging the limitations in fully disentangling cause and effect.

Building on the theoretical relationship between delta hedging and movements in the underlying asset, we expect that a positive GEX environment is associated with OMMs adopting hedging behavior that dampens price fluctuations. This dynamic contributes to lower equity market volatility and supports a risk-on environment, typically characterized by positive returns. Conversely, a decline in GEX, particularly into negative territory, suggests that hedging flows may amplify market movements, leading to heightened volatility and a risk-off regime. However, while these relationships are consistent with economic intuition and widely referenced in market literature, it is important to acknowledge that GEX captures potential hedging pressure rather than realized flows.

3. Literature Review

Research on gamma exposure and its impact on financial markets has largely been divided into distinct areas of focus. A significant portion of the literature examines how gamma exposure influences intraday price movements in the S&P 500 or its behavior around options expiration dates, where large-scale hedging adjustments by OMMs are known to drive short-term market dynamics. Another stream of research investigates the relationship between gamma exposure and volatility, analyzing how changes in option gamma contribute to periods of market stabilization or heightened price fluctuations. A third category focuses on the impact of gamma exposure on individual stocks, assessing how hedging flows related to single-stock options influence their price trajectories.

While previous studies have provided valuable insights into the mechanics of gamma exposure, much of the research has concentrated on short-term price movements or volatility effects rather than aggregated and more long-lasting effects which could translate into actionable trading signals for capital allocation. This study aims to bridge that gap by exploring how GEX influences the future returns of the S&P 500 over the following day(s), with the objective of identifying persistent effects that possibly could inform tactical investment decisions. By examining the relationship between GEX and subsequent market movements, this research seeks to contribute to the existing literature while offering practical implications for portfolio management and systematic trading strategies.

A fundamental question that should be addressed in this context is the Efficient Market Hypothesis (EMH), as formulated by Fama (1970). He argued that all available and publicly accessible information is already fully reflected in asset prices, implying that no systematic excess returns can be achieved by analyzing past price movements or publicly known data. Lo and MacKinlay (1988) challenged this hypothesis early on by demonstrating that stock market prices do not follow a random walk, suggesting that markets may not always be fully efficient. Our study directly but unintentionally challenges the EMH by examining whether gamma exposure provides predictive insights into future market movements, which if confirmed, could suggest that the market is not fully efficient. Whether the market is efficient or not is not directly addressed in this study but given that we assume the market is inefficient in terms of information processing, it is relevant to highlight the long-standing debate we are entering.

Chen, Chung & Lien (2016, p. 448) utilize regression models based on Permanent-Transitory models and Information Shares models, which in turn rely on Vector Error Correction Models, to analyze the role of index options in price discovery for the S&P 500. Their findings suggest that, relative to other derivative instruments, index options lack explanatory power in this regard. Several other studies challenge this conclusion by placing greater emphasis on the fundamental relationship between OMMs delta hedging, volatility and price movements in underlying assets. Pearson, Poteshman & White (2007), for instance, find a statistically and economically significant negative relationship between the volatility of a stock and the net gamma exposure or options positions held by market participants likely engaged in delta hedging. Their study shows that approximately 12% of the daily absolute returns of stocks with active options trading can be attributed to options market participants rebalancing their hedging positions in the underlying asset. This relationship has been demonstrated across multiple financial markets. Anderegg, Ulmann and Sornette (2022) examine the impact of option hedging, specifically delta hedging, on FX spot market volatility. They develop a theoretical model that quantifies the feedback effects induced by hedging activities and empirically validate their framework using DTCC data. Their findings reveal that a net short gamma exposure among OMMs is significantly associated with increased spot market volatility.

While the volatility aspect is highly relevant to this topic, our study focuses on the directional impact of delta hedging resulting from the aggregate gamma exposure of OMMs. Egebjerg and Kokholm (2024) investigate this relationship, but on an intraday basis, particularly examining SPX price movements during the last 30 minutes of the day. Their findings confirm that OMMs hedging activities significantly impact SPX price dynamics. Using high-frequency regression analysis, they quantify the effect of hedging flows on SPX futures and demonstrate that changes in OMMs delta positions are a significant predictor of SPX future returns at the end of the trading session. They also find that these hedging effects intensify during extreme market movements. Their results align with those of Baltussen, Da, Lammers and Martens (2021), who identify a similar effect in the final 30 minutes of the trading day. Their study empirically documents OMMs net gamma positions and their impact on volatility across multiple asset classes (gamma effects). Barbon and Buraschi (2021) further support the relationship between OMMs intraday gamma-related activity and its effects on both price and volatility of the underlying asset. Focusing on a related aspect, Dubois (2022) also establishes a connection between GEX and the S& P500 around option expiration periods. All these studies primarily examine short-term time horizons, such as intraday returns and volatility, whereas our study extends the analysis to longer timeframes up to subsequent trading day(s). Building upon Barbon and Buraschi (2021), Soebhag (2023) employs cross-sectional and panel regressions to show a statistically significant negative relationship between a stock's net gamma exposure and its future returns. His study concludes that net gamma exposure predicts realized volatility for the following month, primarily driven by hedging activities among OMMs rather than information-based options trading. What distinguishes Soebhag (2023) from prior research is its confirmation of this relationship over a longer time horizon, demonstrating that net gamma exposure for individual stocks predicts returns not only for the next trading day but also up to the following month.

Following the findings presented above, we assume that capital flows derived from options trading influence both volatility and price information in the underlying asset. Previous studies have confirmed this relationship, which aligns with the intuitive link between OMMs hedging activity and the price of the underlying asset, especially given the growing significance of the options market and the increasing volume of capital flows into the equity market. However, we remain critical of the idea that prices should not be significantly affected in a sufficiently liquid market, where market participants should, in theory, be able to absorb and offset these flows. We do know that flow-driven and price-insensitive trading has the potential to amplify and reinforce market movements. This suggests that gamma exposure may still play a crucial role in predicting stock market fluctuations by driving flows and price information. Like many of the studies mentioned above, our research aims to examine and quantify the effect of OMMs hedging activity on the S&P 500. Unlike many previous studies we use the publicly available GEX index as an independent variable to represent aggregated gamma exposure on SPX.

Previous studies have largely focused on the gamma effect, which describes how OMMs must delta hedge their existing positions to manage risk. Egebjerg and Kokholm (2024) developed a model, without using GEX, that demonstrates how OMMs delta hedging impacts the underlying stock market through both gamma and inventory effects. Their study finds that the inventory effect is highly significant, particularly in high-frequency data when assessing the impact of OMMs delta hedging on the price of the underlying asset. The inventory effect arises when OMMs adjust their exposure by entering new options trades, thereby introducing additional risk to their inventory. The gamma effect suggests that OMMs hedge their delta- and gamma because of fluctuations in already held positions. While GEX is primarily suited to capturing the gamma effect, any inventory-driven shifts in total gamma exposure will also be reflected in GEX.

The contribution of this study lies in complementing previous research by examining the directional movement of S&P 500 as a consequence of OMMs hedging activities stemming from shifts in their aggregated gamma exposure. While individual components of this approach, such as the use of gamma, directional effects, or index-level analysis over the following days have been explored in isolation, there are no prior studies that jointly considers all of these elements. Specifically, we examine aggregated gamma exposure at the index level, focus on directional price movements, and assess predictive effects over multiple forward-looking time horizons. This integrated approach aims to bridge the gap between academic research and practical financial modeling implications for tactical capital allocation.

4. Data

4.1 Data collection

This study relies on daily data to examine whether and how Gamma Exposure (GEX) influences market movements, with focus on the S&P 500 index. The dataset consists of GEX values, S&P 500 daily closing prices, and complementary financial indicators acting as control variables that help assess a robust model, able to isolate the effect of GEX on future equity price movements.

The source for GEX data is SqueezeMetrics, which provides daily estimates of gamma positioning in the options market. Since GEX is derived from options positioning and reflects dealer hedging behavior, it serves as a proxy for how OMMs rebalancing flows may impact underlying asset prices. The primary reason for choosing GEX is its widespread use in the financial industry as a tool for tracking dealer activities and the directional exposure of the options market. Even if GEX is not an optimal measure of SPX gamma exposure due to its underlying assumptions, there is value in using an indicator that market participants actively monitor. Financial metrics with high visibility and frequent usage often become self-fulfilling prophecies if a significant portion of market participants incorporate them into their trading decisions. GEX is readily accessible and simplifies modeling, and if market signals can be identified using this approach, they can also be seamlessly integrated into the market analysis of active asset managers.

Equity market data, including daily closing prices for the S&P 500 index, is obtained from LSEG/Refinitiv. The S&P 500 is chosen as the benchmark due to its natural link to the GEX index. Also, it is highly suitable due to its liquidity, broad market representation, and its central role in institutional investment strategies. To further contextualize the impact of GEX on market dynamics, additional financial indicators such as the Cboe Volatility Index (VIX), put and call option open interest, Dark Pool Index (DIX), economic policy uncertainty index (EPU) and SPY trading volume are incorporated into the dataset. These variables, sourced from Bloomberg, St. Louis FRED and SqueezeMetrics, or provided by AP3, help capture broader market sentiment and trading activity, which may influence or interact with the effects of GEX.

The studied period spans from 2 May 2011 (when GEX was first introduced) to 17 January 2025, ensuring a sufficiently long sample to capture different market conditions. The sample period also includes periods of high option market growth and the structural shift during 2020. The data series in full contain 3451 observations when adjusted for non-common dates (such as weekends for EPU when the S&P 500 is not trading). By making this adjustment we ensure that the dataset is accurate and reliable for examination in our econometric model. Adjustments for stock splits and index rebalancing are included in the equity time series, ensuring that price and return calculations remain accurate. Prior to empirical analysis, the data undergoes a series of preprocessing steps to ensure robustness.

Out of the full data set we created two sub-data frames: one before 2020 and one from 2020 onward, containing 2182 and 1269 observations respectively. This was done to highlight potential differences following the expansion of the options market from around 2020 onward, in line with the first research question and as discussed in the introduction. For the construction of the predictive models, we trained each model on the first 95% of the full data sample and generated forecasts for the remaining 5%. This approach aligns with the short-term focus of our analysis, which emphasizes forecasting over the immediate days ahead. Robustness is ensured by the relatively large dataset, which provides a sufficient number of observations for both the training and testing periods.

The choice of a daily frequency is motivated by 1) the short-term nature of GEX fluctuations, as dealer hedging flows can change rapidly and potentially affect market sentiment within short time horizons, and 2) To complement this field of research by evaluating the effect of gamma exposure over the following day(s) on an aggregated level, and 3) The necessity of actionable insights to include in active asset allocation. A shorter data frequency (hours, minutes) might show more clear relationships, as stated in previous studies in the literature review, but for the relationships to be actionable for an asset allocator and not a trader, the time frame needs to be set accordingly.

All variables in the model except for GEX are expressed in daily changes, where the rate of change from one day to another is calculated. Due to GEX's extreme value fluctuations and frequent shifts between positive and negative territory, it is not well-suited for computing logarithmic or standard daily returns. Instead, we compute its first derivative as the daily changes. To estimate the first derivative of GEX we applied the Generalized Orthogonal Local Derivative (GOLD) method for derivative approximation (Deboeck, 2010). This approach is particularly well-suited for time series data, as it estimates derivatives using local polynomial approximations based on a defined number of surrounding data points (i.e., embedding). In our case, we set the embedding parameter to 2 and the derivative order n=1 to compute the first derivative of GEX. The derivative was calculated based on the temporal differences between successive observations of GEX and its associated time vector. The derivative of GEX is therefore the difference in GEX from one day to another divided by the number of days between them, making the change in GEX an approximation of its derivative. By focusing on the first derivative rather than the absolute level of GEX, we aim to capture how changes in gamma positioning affect short-run market movements, consistent with theoretical expectations of delta-hedging flows responding to shifts in gamma exposure.

4.2 Control Variables

A. Volatility Index (VIX)

The Cboe Volatility Index (VIX), often referred to as the "fear index," measures the market's expected volatility over the next 30 days, derived from S&P 500 index options. The VIX index is derived from real-time S&P 500 option prices by taking the midpoints between bid and ask quotes across a range of strike prices and maturities. It incorporates both standard and weekly S&P 500 options traded on the Cboe Options Exchange (Cboe, 2025).The methodology relies on risk-neutral pricing principles, where implied volatility is extracted from option premiums. VIX is constructed as an estimate of the risk-neutral expected variance, weighting option prices to approximate the market's expectation of future volatility. As a result, VIX serves as a key indicator of market uncertainty, with higher values reflecting increased risk aversion and potential turbulence in financial markets. The generalized formula used in the VIX calculation is:

$$\sigma^2 = \frac{2}{T} \Sigma_i \frac{\Delta K_i}{K_i^2} e^{rt} Q(K_i) - \frac{1}{T}\left[\frac{F}{K_0} - 1\right]^2 \tag{6}$$

Time to option expiration (T), expressed in years, ensures that volatility is measured over a consistent period. K_i denotes the strike price of the i-th option, while difference between consecutive strike prices (ΔK_i), is essential for weighing the options in a way that reflects market expectations. The risk-free interest rate (r) is used to discount option values, capturing the time value of money in derivative pricing. To estimate implied volatility, the midpoint of bid-ask prices for options with strike price K_i, denoted as $Q(K_i)$, is used to minimize distortions from bid-ask spreads. Additionally, the forward price of the S&P 500 (F), plays a central role, as it is derived from put-call parity and is used to identify K_0, the strike price just below F, which serves as a reference point in the calculation (Cboe, 2025). By combining these components, VIX provides an option market-based estimate of expected volatility, making it a key indicator for investors seeking to assess market risk and potential turbulence.

The inclusion of VIX allows us to control for return variations driven by changes in implied volatility, as captured by the option greek Vega. Its inclusion is particularly relevant given its structural similarity to GEX, as both measures are derived from S&P 500 options and reflect option-implied market dynamics. Also, the VIX Index is a widely used measure in financial modeling, serving as a benchmark for market-implied volatility and risk assessment. It is often used as an explanatory variable in empirical asset pricing studies, risk management frameworks, and portfolio allocation strategies, making it a well-established tool for analyzing market sentiment and expected fluctuations in equity returns.

Figure 3: Volatility Index (VIX)

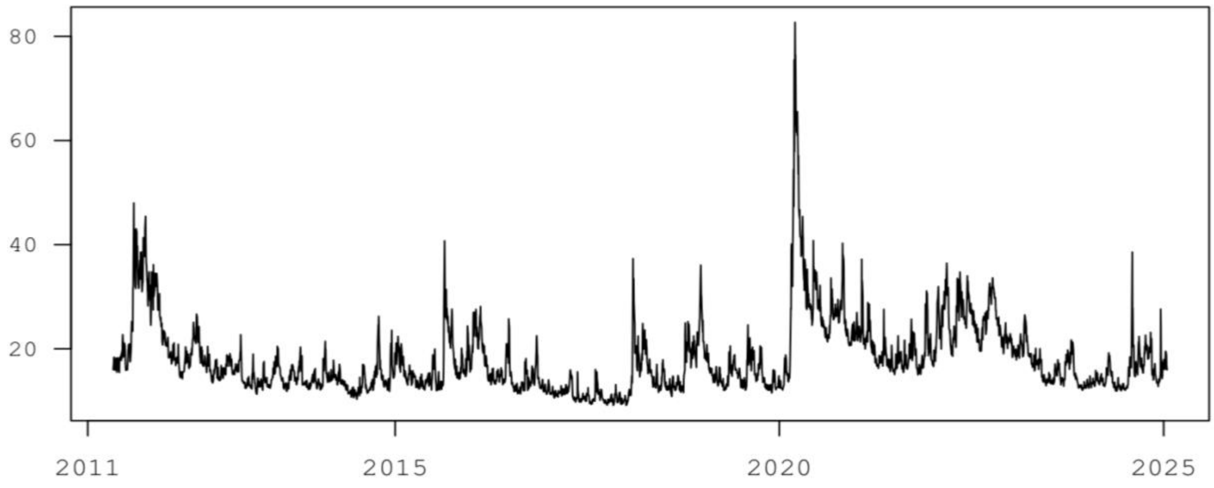
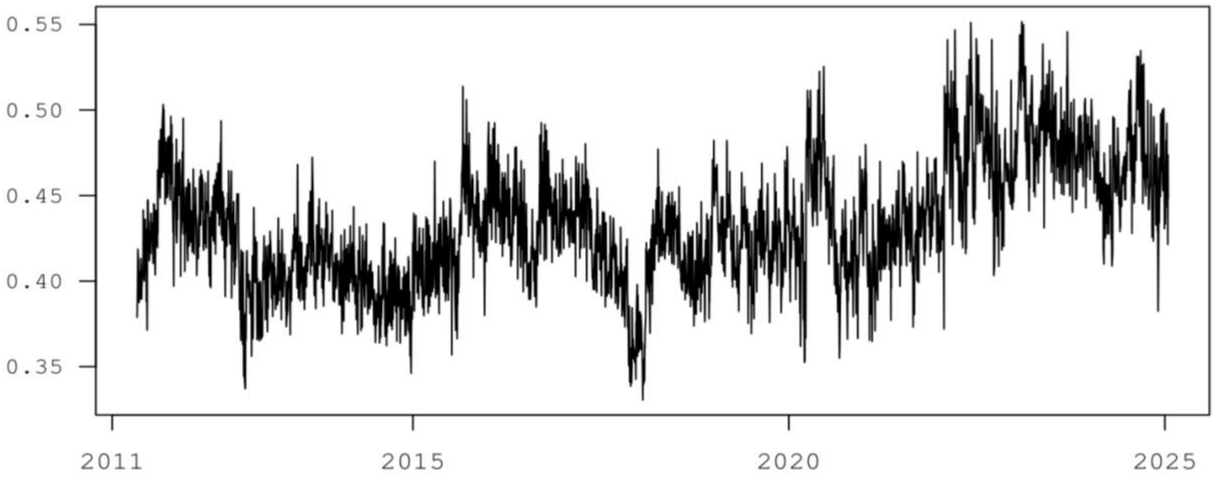


Figure 3 illustrates the development of VIX throughout our studied period. The spikes in volatility occurs in times of uncertainty in the financial markets, e.g., in 2011 during the eurozone debt crisis and U.S. debt ceiling standoff, and in 2020 during the initial phase of the covid pandemic. The plot highlights cyclical fluctuations in market implied volatility with elevated levels during periods of financial stress.

B. Dark Index (DIX)

The Dark Index is developed by SqueezeMetrics (2018) and captures institutional sentiment by measuring net buying activity in dark pools. Dark pools are private, non-transparent trading venues primarily used by large institutional investors. These off-exchange markets lack a visible order book, allowing institutions to execute large orders with minimal market impact. A high DIX value typically signals strong net buying interest from institutional investors, whereas a low DIX reflects net selling pressure. Since institutional trading flows can significantly influence price dynamics in the broader market, DIX is included in the model to control for changes in institutional positioning, helping to isolate the explanatory power of GEX on S&P 500.

Figure 4: Dark Pool Index (DIX)

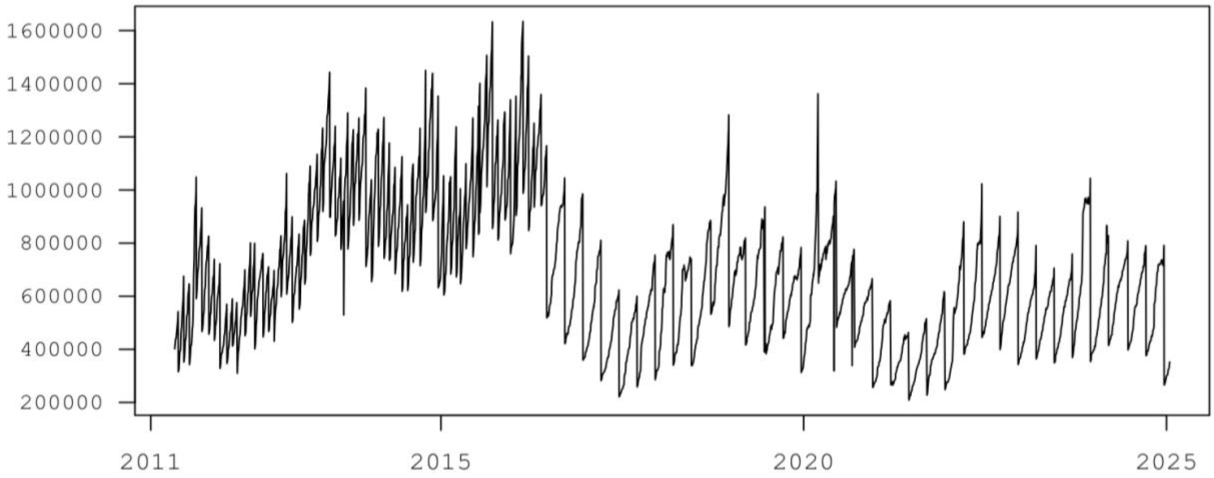


The DIX plot suggests an increasing dark pool activity throughout our studied period. The DIX shows relatively stable institutional sentiment from 2011 to 2015, followed by a gradual upward trend. A sharp drop in 2020 reflects the COVID-19 shock, after which sentiment rises to historically high levels. Post-2020, the index remains elevated indicating sustained institutional activity.

C. Aggregate Call Open Interest

Aggregate Call Open Interest (Call OI) refers to the total number of outstanding call option contracts on the S&P 500, reflecting the market's cumulative positioning for potential upside. A high level of Call OI typically indicates increased demand for bullish exposure, as investors seek to benefit from potential price increases in the underlying index. Conversely, lower Call OI levels may suggest fading optimism and reduced participation in upward speculation. By including Call OI as a control variable, we account for prevailing sentiment and positioning in the options market that could influence return dynamics in the S&P 500. This helps isolate the unique contribution of GEX by controlling for broader option-driven market pressures.

Figure 5: Call Open Interest

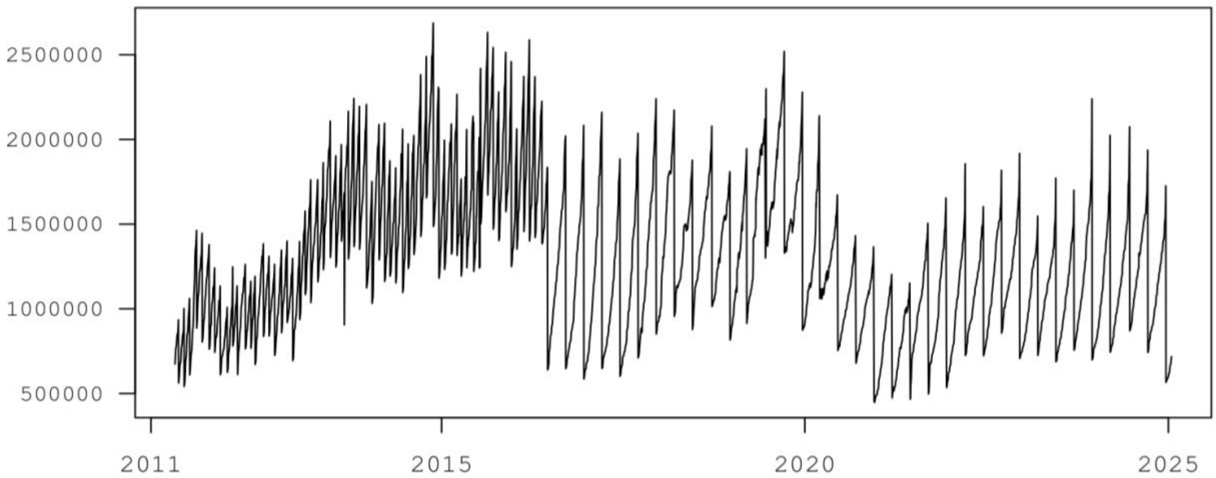


The aggregate call open interest plot exhibits a downward trend from the first third of our sample period into the later parts of the period. This could be explained by a higher (retail) interest for short dated options like 0DTE options, implying that the OI doesn't aggregate at the same pace as before. Another reason could be the shift from classic SPX standard options to other instruments like e.g., the mini-SPX (XSP) options or SPY ETF options. A third explanation could be a shift to trading unique stock options, as seen in the interest for e.g., Tesla or Nvidia options. The noise in the data is clearly diminishing from 2016 and forward, and the pattern exhibits a clear structural shift. This could possibly be explained by changes in open interest reporting, more standardized maturities or changes in data sources. Overall, the pattern in the data suggests systemic expiry or roll-over behavior likely linked to monthly or quarterly option cycles.

D. Aggregate Put Open Interest

Aggregate Put Open Interest (Put OI) captures the total number of outstanding put option contracts on the S&P 500. This metric reflects market demand for downside protection, often linked to increased hedging activity by institutional investors or elevated risk aversion among market participants. High levels of Put OI may indicate a broader expectation of volatility or negative market developments, while lower levels suggest reduced concern over potential drawdowns. Put OI is included in the model alongside Call OI to control for the overall positioning in the options market. By accounting for prevailing hedging activity and bearish sentiment, the model can more accurately isolate the impact of GEX on equity returns.

Figure 6: Put Open Interest



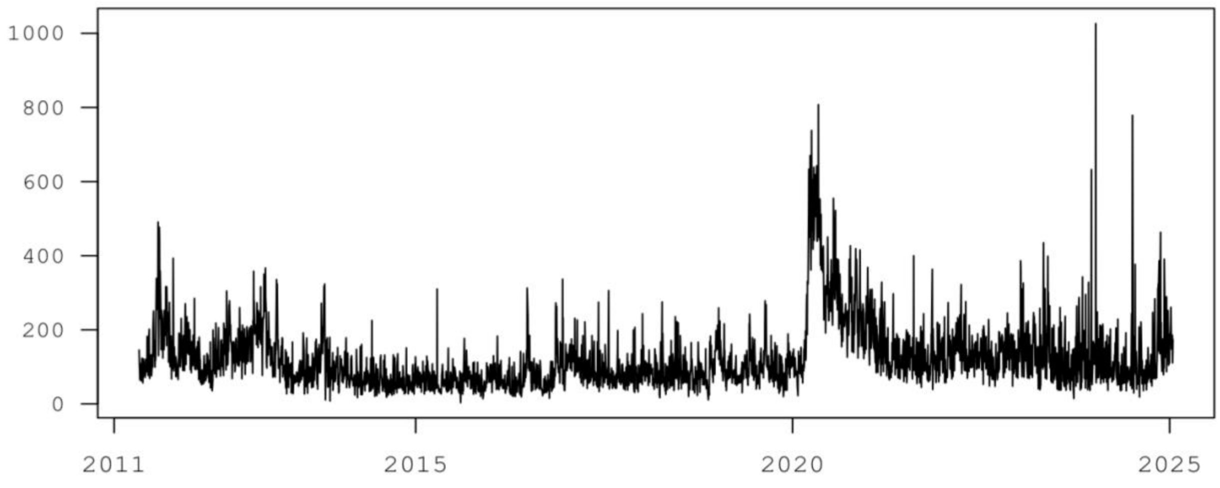
For the aggregate put open interest, the same theme of declining open interest and patterns linked to noise and roll-over behavior as for the aggregated call OI can be observed. The downtrend in the aggregated open interest for put options is not to the same extent as for call options, which *could* indicate that SPX options is more frequently used for protection rather than speculation. As for the Call OI the pattern exhibits a clear structural shift. This could possibly be explained by changes in open interest reporting, more standardized maturities or changes in data sources.

E. U.S. Economic Policy Uncertainty (EPU) Index

To account for the impact of macroeconomic and policy-related uncertainty on financial markets, we include the U.S. Economic Policy Uncertainty (EPU) Index as a control variable in our model. Developed by Baker, Bloom, and Davis (2016), the EPU Index quantifies economic policy uncertainty by combining the frequency of newspaper articles referencing terms related to the economy, uncertainty, and policy, with measures of tax code expirations and forecaster disagreement.

Incorporating the EPU Index is relevant in our context as heightened policy uncertainty may plausibly influence investor behavior, OMMs risk appetite, and overall market liquidity, factors that can affect both Gamma Exposure (GEX) and S&P 500 returns. By controlling for fluctuations in economic policy uncertainty, we aim to isolate the marginal explanatory power of GEX and ensure that our results are not confounded by broader macroeconomic uncertainty shocks. The EPU Index thus serves as a meaningful proxy for external uncertainty that may otherwise bias our estimates if omitted.

Figure 7: U.S. Economic Policy Uncertainty Index (EPU)

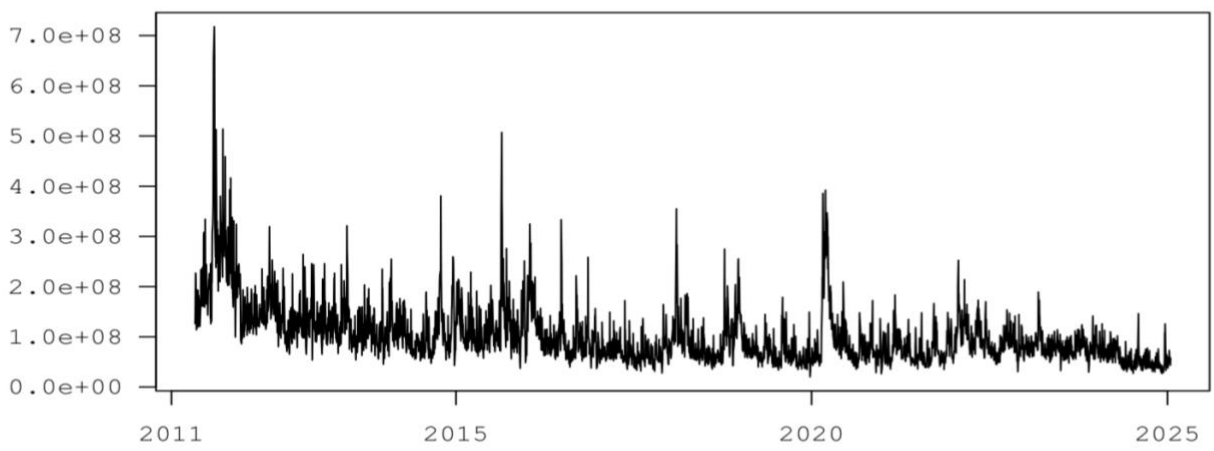


The development of the EPU index clearly coincides with uncertainty in the economy and the financial markets. This is clear when evaluating the years from 2020 throughout 2024. During this period, we have seen a pandemic, rising inflation and recession fears, the Russian invasion of Ukraine and an inverted U.S. yield curve - all of which is reflected in heightened EPU-levels.

F. SPY US Equity ETF Trading volume

SPY is one of the most heavily traded exchange-traded funds (ETFs) globally and is designed to track the performance of the S&P 500 index. It offers investors a highly liquid and efficient way to gain broad exposure to the U.S. equity market. Because of the cash settled nature of S&P 500 options, instruments like the SPY ETF enables OMMs to hedge their delta position since there is no trading in the actual index. Given its strong correlation with the underlying index and its central role in institutional trading, the daily trading volume of SPY is often viewed as a real-time indicator of investor activity and sentiment. SPY trading volume is included as a control variable to account for variations in liquidity and trading intensity that could independently influence movements in the S&P 500. High trading volumes may signal elevated investor attention, increased hedging or speculative behavior, all of which can affect short-term price dynamics. By controlling for SPY volume, we aim to isolate the effect of GEX from broader shifts in market participation and trading flows.

Figure 8: SPY US Equity ETF Trading volume



In recent years, trading volume in the SPDR S&P 500 ETF (SPY) has declined despite a strong performance in the underlying S&P 500 index. Several structural factors may explain this trend. First, SPY faces increasing competition from lower-fee alternatives such as IVV and VOO, fragmenting trading volume across multiple S&P 500 tracking products. Second, market participants have increasingly utilized alternative instruments including futures, sector ETFs, and short-dated options for tactical exposures, further reducing SPY's role as the primary vehicle for index-level trading activity. Despite this, the SPY ETF makes up a feasible proxy for our purpose since we are interested in the short-term effects of changes in volume rather than long term structural trends.

4.3 Descriptive statistics

This chapter provides an overview of the key variables used in our empirical analysis through a combination of descriptive statistics and graphical representations. By summarizing the data both numerically and visually, we aim to offer the reader a clear understanding of the structure, dynamics, and relationships inherent in the dataset.

Table 1: Descriptive statistics (Daily changes), May 2011 - December 2019

	Mean	Median	Std. Dev	Skewness	Kurtosis
S&P 500	0.04%	0.05%	0.91%	-0.46	7.93
GEX	8.54	35.9	740	-0.39	9.34
DIX	0.11%	0.06%	4.66%	0.054	3.43
VIX	0.31%	-0.47%	8.32%	2.37	24.01
Put OI	0.63%	1.94%	9.33%	-3.40	32.69
Call OI	0.58%	1.74%	9.38%	-2.12	45.74
EPU	15.39%	-1.00%	87.08%	11.34	268.61
SPY Volume	5.69%	-0.18%	37.18%	1.43	6.83

Note: GEX represents its derivative and is expressed in million units.

Table 1 summarizes the descriptive statistics of all daily changes prior to 2020. The S&P 500 displays relatively modest daily fluctuations, in line with expectations for a broad equity index. GEX shows substantial variability, indicating frequent and sometimes sharp shifts in the data, affecting dealer hedging behavior. The average daily change in GEX is clearly positive, suggesting a general upward drift in gamma positioning over time.

DIX and VIX exhibit more pronounced day-to-day variation, capturing shifts in investor sentiment and changes in implied volatility. Both put- and call open interest changes are sizable, reflecting dynamic positioning in the options market. EPU, changes in economic policy uncertainty, stands out with large swings across the sample. Finally, trading volume in SPY also varies significantly, reflecting periods of elevated market activity.

Table 2 presents the descriptive statistics of daily changes from January 2020 to January 2025. Compared to the pre-2020 period, the S&P 500 exhibits a clear rise in daily volatility, with its standard deviation increasing from 0.91% to 1.34%, indicating more unstable market conditions in recent years.

Table 2: Descriptive statistics (Daily changes), January 2020 - January 2025

	Mean	Median	Std. Dev	Skewness	Kurtosis
S&P 500	0.06%	0.09%	1.34%	-0.51	16.08
GEX	-61.5	47.1	2425	-0.86	11.06
DIX	0.12%	0.09%	4.87%	0.15	3.93
VIX	0.33%	-0.81%	8.38%	2.35	16.79
Put OI	0.52%	1.21%	7.85%	-7.38	58.91
Call OI	0.59%	1.1%	10.03%	-3.98	148.41
EPU	17.55%	-1.41%	85.36%	5.01	50.06
SPY Volume	4.43%	-1.24%	32.19%	1.27	6.03

Note: GEX represents its derivative and is expressed in million units.

The most significant shift is observed in the GEX derivative. Its average value has turned negative, moving from 8.5 million to -61.5 million, accompanied by a substantial increase in standard deviation from 740 million to 2.425 billion. The data implies a fundamental change in the dynamics of gamma exposure during the post-2020 period.

Table 3: Pearson Correlation Coefficients

	S&P 500	GEX	DIX	VIX	Put OI	Call OI	EPU	SPY Volume
S&P 500	1.00	0.39	-0.18	-0.73	0.01	0.03	0.00	-0.25
GEX	0.39	1.00	-0.11	-0.35	0.03	0.03	0.01	-0.18
DIX	-0.18	-0.11	1.00	0.13	-0.04	-0.06	-0.01	0.13
VIX	-0.73	-0.35	0.13	1.00	0.06	0.04	0.03	0.36
Put OI	0.01	0.03	-0.04	0.06	1.00	0.87	-0.01	0.00
Call OI	0.03	0.03	-0.06	0.04	0.87	1.00	-0.01	-0.01
EPU	0.00	0.01	-0.01	0.03	-0.01	-0.01	1.00	0.00
SPY Volume	-0.25	-0.18	0.13	0.36	0.00	-0.01	0.00	1.00

Note: The table reports Pearson correlation coefficients, which measure the strength and direction of the linear relationship between two continuous variables. Coefficients range from −1 to +1, with values closer to ±1 indicating stronger linear association. The measure is sensitive to outliers and assumes normally distributed data.

The correlation matrix is based on daily changes rather than levels which aligns with the structure of the ARDL model used in this study. The most relevant observation is the positive correlation between the derivative of GEX and S&P 500 returns, suggesting that increases in gamma exposure tend to coincide with rising equity prices. This aligns with the theoretical expectations around delta-hedging behavior. The strong negative correlation between S&P 500 returns and the change in VIX confirms that periods of higher volatility typically coincide with falling equity prices, a well-known dynamic in financial markets. Interestingly, SPY trading volume shows a moderate negative correlation with the S&P 500 and with GEX, potentially indicating that increased trading volume coincides with weaker market conditions and lower gamma positioning. The relatively weak correlations between GEX and the other control variables suggest that multicollinearity is unlikely to pose a significant issue in the regression framework.

Table 4: Spearman Correlation Coefficients

	S&P 500	GEX	DIX	VIX	Put OI	Call OI	EPU	SPY Volume
S&P 500	1.00	0.68	-0.24	-0.78	-0.09	-0.03	0.00	-0.26
GEX	0.68	1.00	-0.18	-0.56	-0.09	-0.04	-0.02	-0.23
DIX	-0.24	-0.18	1.00	0.17	0.00	-0.04	0.01	0.13
VIX	-0.78	-0.56	0.17	1.00	0.15	0.06	0.02	0.29
Put OI	-0.09	-0.09	0.00	0.15	1.00	0.50	0.01	0.14
Call OI	-0.03	-0.04	-0.04	0.06	0.50	1.00	-0.02	0.09
EPU	0.00	-0.02	0.01	0.02	0.01	-0.02	1.00	-0.02
SPY Volume	-0.26	-0.23	0.13	0.29	0.14	0.09	-0.02	1.00

Note: The table reports Spearman correlation coefficients, which assess monotonic relationships based on ranked values. Unlike Pearson correlation, Spearman does not assume normality and is less sensitive to outliers, making it suitable for non-linear or ordinal data. Coefficients range from –1 to +1.

The correlation between the GEX derivative and S&P 500 increases when you adjust for the scale of the variables and only account for the direction. As seen in VIX, the Pearson and Spearman correlation with S&P 500 does not change significantly, but for the GEX derivative the correlation increases from 0.39 to 0.68. This notable jump implies that while the absolute magnitudes of changes in GEX and S&P 500 may not always align, their directional movement is consistent. In other words, when the gamma exposure derivative increases, equity returns tend to rise with high regularity, even if the changes vary in size. This directional alignment may reinforce the relevance of gamma dynamics as a potential driver of market movements.

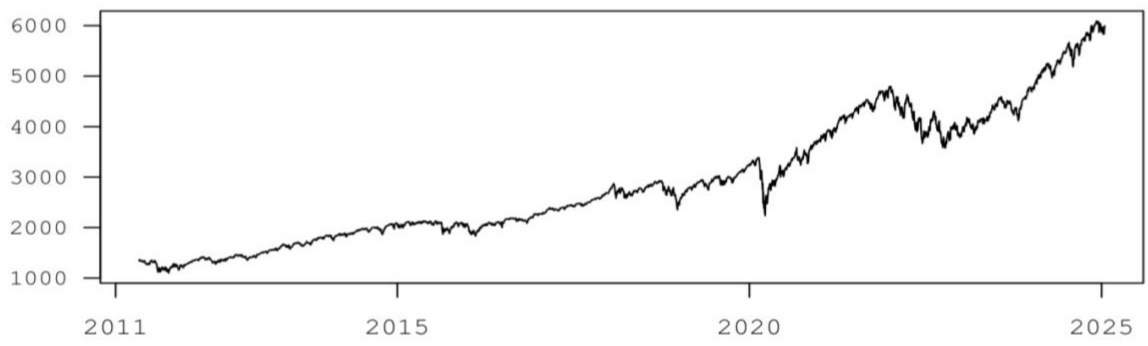
The Spearman matrix also confirms the strong negative relationship between S&P 500 and VIX, as well as the limited associations between GEX and the remaining control variables. Together with the Pearson results, this further supports the notion that GEX adds unique explanatory value with minimal multicollinearity risk.

To test for collinearity, we conducted a variance inflation factor (VIF) test to measure the collinearity for each factor in our regression. An indication of collinearity is when there is correlation between the regressors which can affect conclusions made about the significance of effects and model practicality (Craney & Surles, 2002). There are no formal criteria but values of 5 and 10 are frequently used as rules of thumb in econometric studies, which is why we used 5 and 10 as cutoffs for our test. The VIF-test was conducted for all coefficients in the model over the full period, resulting in values ranging from approximately 1 to 2,5 with a few reaching above 4. Of particular interest is the interpretation of multicollinearity in GEX as it is the explanatory variable whose coefficients are the focus of the analysis. The VIF values for the GEX lags are just above 1,5 which is well below the commonly accepted threshold and therefore pose no issues for interpreting the GEX coefficients, a key priority in this study. It should also be noted that no VIF values exceed 5, indicating that multicollinearity is not a problem in our model.

4.4 Data visualization

In addition to basic summary statistics such as means, standard deviations, and distributions, we include time-series plots and scatter plots to illustrate how the main variables of interest evolve over time and interact with one another. Particular emphasis is placed on visualizing the problem under investigation: whether and how GEX affects equity market returns. This visual foundation serves to contextualize our methodological approach and provides preliminary insights that inform the formal econometric analysis in subsequent sections.

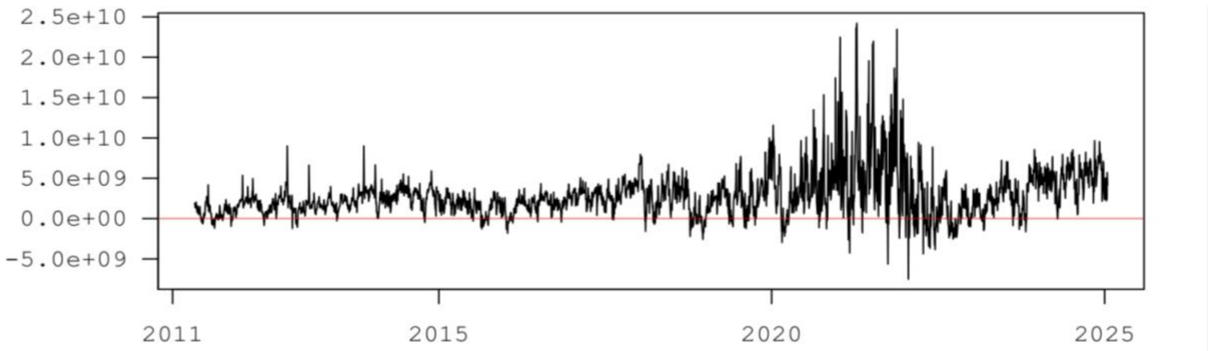
Figure 9: S&P 500 (SPX)



Above, Figure 9 illustrates the development of the S&P 500 index over the period examined in this study, and it shows that the S&P 500 have had some strong years performance wise resulting in a net gain of +341% during our observed period. Even though we are not interested in the long-term development of the index, this visualization provides a contextual backdrop for our analysis, including both upward trends and periods of heightened volatility. The index's performance reflects a combination of macroeconomic conditions, monetary policy regimes, and investor sentiment, which may all interact with the dynamics of option-based positioning.

Figure 10 displays the evolution of aggregated gamma exposure for the S&P 500 index over the sample period. A few key patterns emerge from the visualization. First, GEX values are predominantly positive, suggesting that option OMMs are, on average, net long gamma. This typically results from the buying of at-the-money call options where gamma exposure is at its highest, combined with the selling of out-of-the-money put options, which contribute relatively little gamma. As shown in Figures 6 and 7, open interest is generally higher for put options. This indicates that the positive GEX levels are not driven by open interest per se, but rather by the gamma profile of the options. This positioning pattern, together with the assumptions of GEX that OMMs buy calls and sell puts, contributes to a consistently positive gamma environment, clearly illustrated in Figure 10.

Figure 10: Gamma Exposure (GEX)



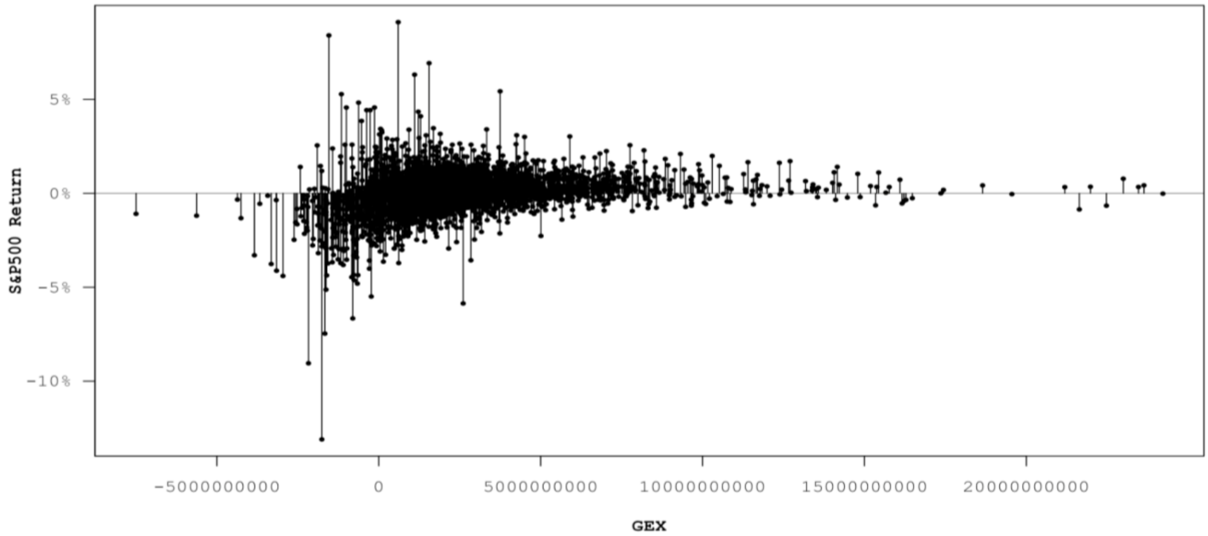
Note: The red line highlights the zero mark on the y-axis, visualizing the fluctuations between positive and negative values.

This positioning implies that OMMs are expected to dampen market volatility through their hedging activity, buying the underlying asset when prices fall and selling when prices rise. Conversely, periods where GEX falls below zero, though less frequent, suggest that hedging flows may instead amplify market moves. Second, GEX exhibits a notable increase in magnitude and variability over time, especially following the onset of the option market expansion in 2020. This likely reflects a combination of elevated option activity and shifts in market structure (e.g., growth in retail option trading and volatility-targeting strategies). This is also evident in the data, where the daily change in GEX exceeded 1 billion (in absolute terms) on 13% of the days prior to 2020, compared to 32% of the days after 2020.

Importantly, GEX values frequently reach into the tens of billions of USD, which has methodological implications for our econometric analysis. Due to the scale of these figures, coefficient estimates involving GEX should be interpreted with appropriate context, as small changes in the dependent variable may are reflected from large nominal shifts in GEX. However, this does not present an interpretational issue, GEX is known to fluctuate substantially from day to day, often moving by several billion USD. These dynamics justify the use of the original scale and support the practical relevance of the resulting coefficient magnitudes, despite the apparent mismatch in numerical scale between the variables.

In Figure 11, S&P 500 returns are illustrated across different GEX levels where the dispersion confirm that GEX levels tend to be mostly positive as earlier stated. The figure clearly shows how the dispersion in S&P 500 returns varies at different levels of GEX, with a distinct difference between positive and negative GEX values. This relationship aligns with the economic explanation of delta hedging previously outlined. When GEX is negative, the return distribution is much wider, reflecting heightened volatility caused by destabilizing hedging flows. In contrast, as GEX increases and moves into positive territory, the return dispersion becomes increasingly compressed, suggesting that price swings are mechanically suppressed by stabilizing delta-hedging activity. The return behavior between negative and positive GEX regimes underscores the predictive value of gamma positioning. The market appears significantly more volatile and reactive in low or negative GEX environments.

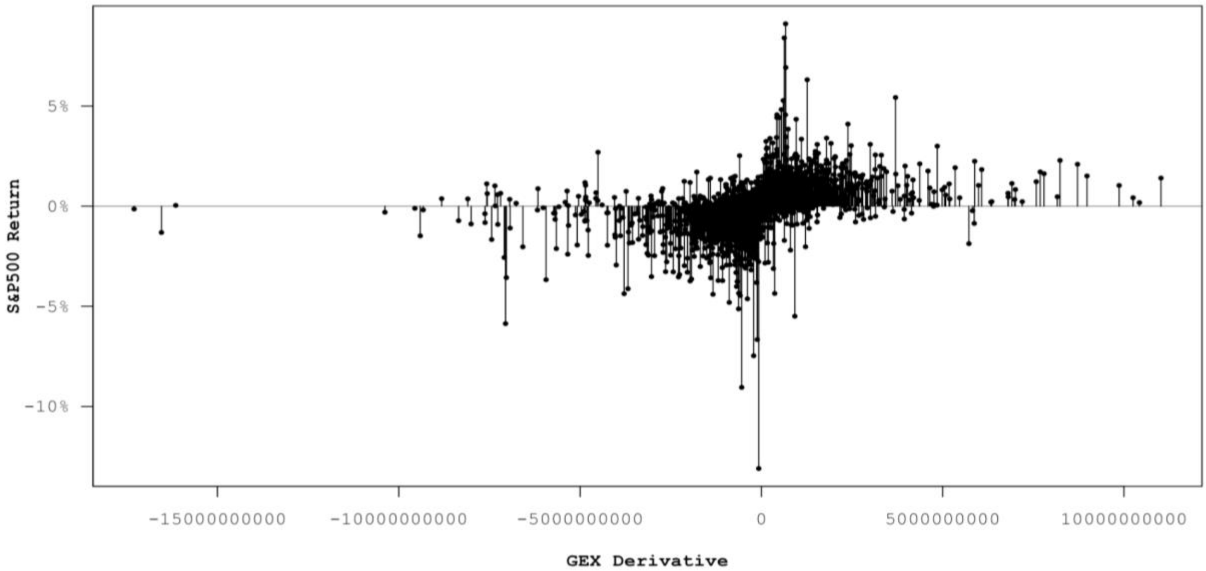
Figure 11: GEX vs S&P 500 Returns



Note: The dispersion of points indicates how S&P 500 return volatility varies with GEX values.

The distributional plot of S&P 500 returns against the daily GEX derivative, seen in Figure 12, reveals even more distinct structural patterns in the data. Most notably is that return dispersion is highest when the GEX derivative is near zero. This indicates that market behavior becomes more unpredictable when aggregate gamma exposure is undergoing significant change. One explanation is that when the GEX derivative approaches zero, the market may be transitioning between gamma regimes, from positive to negative or vice versa. These structural shifts create uncertainty, increasing the influence of non-hedging-related order flow, speculative trading and broader macro factors which can drive volatility. When the GEX derivative is strongly positive or negative, return dispersion appears lower. This suggests that during periods of strong directional shifts in gamma exposure, OMMs engage more consistently in delta hedging activity. While such hedging can dampen or amplify volatility depending on the gamma regime, the presence of clear positioning seems to reduce the role of uncertainty and discretionary trading in the market.

Figure 12: GEX Derivative vs S&P 500 Returns

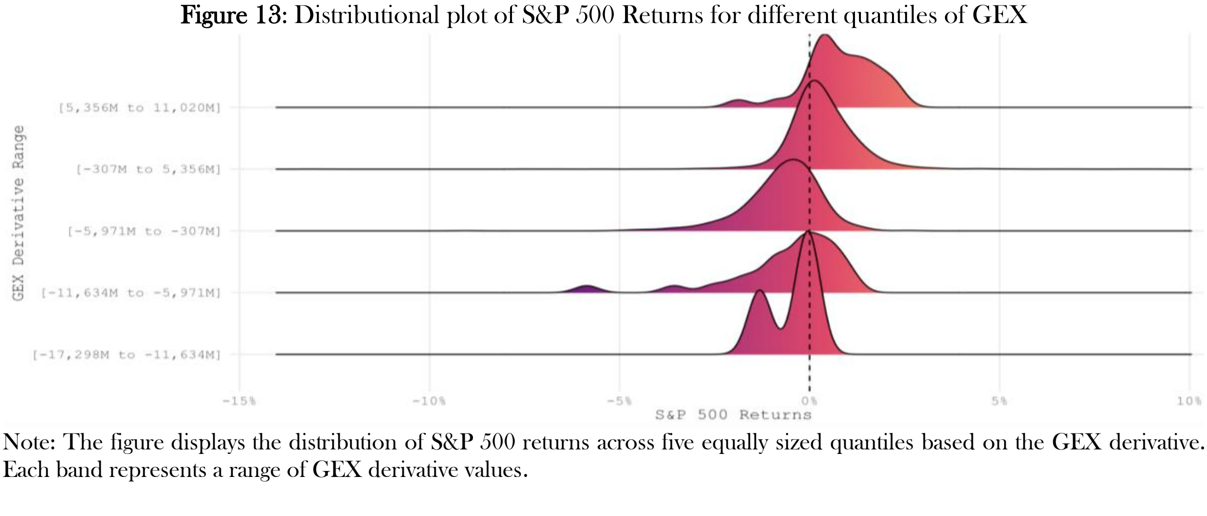


Note: Illustration of how S&P 500 returns dispersion varies across levels of the GEX derivative.

In addition to these patterns, we find that the S&P 500 return and the GEX derivative share the same sign in most cases: when GEX is above average and the derivative is positive, the S&P 500 returns are positive in 82.4% of instances, and when GEX is below average with a negative derivative, the S&P 500 declines in 85.3% of cases. This directional alignment further supports the hypothesis that changes in gamma exposure, and the hedging they induce, have a meaningful relationship with price movements in the underlying equity market.

Figure 13 illustrates a distributional view of S&P 500 returns across different quantiles of the GEX derivative. A clear pattern emerges: the more positive GEX derivative, the more right skew of the S&P 500 returns. Meaning, the conditional return's distribution is dependent on the sign of the GEX derivative. This pattern clearly outlines the subject of this study. The distribution for S&P 500 returns is centered around zero (as seen in Table 1 and 2), and when the returns are conditional on the GEX derivative the skewness and center of the distribution shifts slightly (as seen in Figure 13).

Figure 13: Distributional plot of S&P 500 Returns for different quantiles of GEX



Note: The figure displays the distribution of S&P 500 returns across five equally sized quantiles based on the GEX derivative. Each band represents a range of GEX derivative values.

5. Method

5.1 Research and modeling framework

Our study conducts a quantitative approach examining the relationship between GEX and S&P 500 returns in a systematic, objectively and measurable way. A quantitative approach is required since the study is based on time series data and aims to identify statistical relationships using econometric methods. By using a quantitative approach, we can make conclusions based on empirical data rather than subjective interpretations, which increases the reliability and allows for future replication. The foundation of the study's analysis is data gathering, data processing, descriptive statistics and econometric modelling. Embracing this method, our study can measure and analyze the relationship between our variables in a way that makes it possible to conclude statistical relationships. This approach enables handling large quantities of data which is necessary when analyzing constantly moving financial markets. By quantifying these relationships, we can make more objective conclusions compared to if we were to conduct a more qualitative method.

To ensure that our econometric model accurately reflect the structural conditions of the market, we estimate the effects separately for two distinct subperiods: pre-2020 and post-2020. This is necessary to highlight potential differences in the relationship between GEX and S&P 500 following the expansion of the options market. While a full-sample analysis could provide more statistical power due to a larger number of observations and a broader view of the overall relationship, it relies on the assumption that the underlying dynamics remain stable over time. This assumption is not justified given the structural changes in the options market. Dividing the data into sub-periods enables an accurate comparison of the effect pre and post 2020, while also maintaining the statistical power to enable exploration of the relationship. Applying a unified model across these different regimes could lead to biased or misleading coefficient estimates. By estimating the effects separately for both periods, we allow for time-varying dynamics and capture differences in how GEX interacts with the equity market. This approach allows us to directly assess whether the informational value of GEX has changed over time.

5.2 ARDL Model

The analysis follows a multivariate framework and employs an *Autoregressive Distributed Lag (ARDL)* model specification as explained in Natsiopoulos and Tzeremes (2022). An ARDL model is an econometric model used to analyze the relationship between a dependent variable and one or more independent variables, incorporating both current and past values of these variables. The autoregressive (AR) part incorporates lags of the dependent variable to capture its own past influence, and the distributed lag (DL) part includes current and lagged values of independent variables to assess their impact over time.

For an ARDL (p,q_1, ..., q_k) model with k independent variables, the general equation is:

$$y_t = c_0 + \sum_{i=1}^{p} \psi_{y,i}\, y_{t-1} + \sum_{j=1}^{k} \sum_{l=0}^{q_j} \beta_{j,l}\, x_{j,t-1} + \epsilon_t \tag{7}$$

where c_0 is a constant term, $\psi_{y,i}$ are coefficients for the lagged dependent variable y_{t-1}, $\beta_{j,l}$ are coefficients for the current and lagged independent variables $x_{j,t-1}$ and ϵ_t denotes the error term.

The optimal lag order was determined by testing all possible lag structures, with a maximum restriction of 10 lags per variable and selecting the specification with the lowest Akaike Information Criterion (AIC). AIC is a metric used for model selection in statistical and econometric analysis, balancing model fit and complexity to identify the most efficient model. It is calculated as: $[AIC = -2ln(L) + 2k]$ where L represents how well the model fits the data and k is the number of parameters in the model. A lower AIC indicates a better balance between fit and complexity. By using AIC, we can compare different ARDL model specifications and select the optimal one based on the dataset. Minimizing AIC ensures that the model is efficient, robust and not overfitted. The AIC was chosen over BIC given the relatively large sample size of the dataset and AIC's suitability for time series models where residuals may exhibit autocorrelation. To ensure a robust examination between the pre- and post 2020 period, the optimal lag structure was calculated based on AIC for the full dataset and applied on both sub-periods. The optimal lag structure for the model is as follows:

Table 5: Optimal lag structure for the ARDL model

	Current Value	Optimal lag order
S&P 500	N/A	$S\&P500_{t-9}$
GEX	INCLUDED	GEX_{t-9}
DIX	INCLUDED	DIX_{t-8}
VIX	INCLUDED	VIX_{t-9}
Put OI	INCLUDED	$PUT\ OI_{t-6}$
Call OI	INCLUDED	$CALL\ OI_{t-6}$
EPU	INCLUDED	N/A
SPY Volume	INCLUDED	$SPY\ Volume_{t-5}$

Note: Overview of the ARDL modeling framework, showing the lag structure applied to each explanatory variable. The contemporaneous value of the dependent variable is excluded, as it is the outcome being explained. Current values are retained for the explanatory variables to allow for contemporaneous relationships in the analysis.

In the model specification, all current values (except for the dependent variable) and lags up until the optimal lag is included, where the lags included for e.g., GEX is GEX_{t-1} to GEX_{t-9}. Although a lag structure is absent for EPU, it is still included in the model to control for the variation in the other variables. Dynamic multipliers were calculated to illustrate whether a time-varying effect over and beyond the lag structure exists, how that effect evolves over multiple periods and whether it accumulates or fades. This provides complementary insights beyond the standard coefficients obtained from the ARDL model.

Furthermore, the inclusion of multiple control variables such as VIX, open interest, and trading volume within the ARDL framework enables us to isolate the marginal contribution of GEX to return dynamics, while accounting for other known drivers of market behavior. Financial markets are inherently complex and influenced by a multitude of interacting factors, which motivates the inclusion of several control variables in our model. This enhances the credibility of our empirical findings and supports the strength of our claim regarding GEX as a meaningful predictor of short-term market movements.

Among the merits of using the ARDL framework is its ability to incorporate a flexible lag structure, which helps mitigate issues related to simultaneity. By relying on lagged values of both the dependent and independent variables, the model reduces the risk of contemporaneous feedback effects that can obscure causal interpretation. In the context of our study, this is particularly relevant given the potential two-way interaction between GEX and S&P 500 returns. While current returns may influence option positioning, and thereby GEX, the use of lagged GEX values ensures that we isolate the predictive component of GEX that precedes the observed market outcome. This temporal separation strengthens the empirical identification of directionality and allows us to explore dynamic relationships without relying on strict exogeneity assumptions. Although the ARDL framework allows for the inclusion of variables integrated of both I(0) and I(1) (Kripfganz & Schneider, 2023), the Augmented Dickey-Fuller test confirms that all variables in our model are stationary. We opt for the ARDL framework because of its ability to capture lagged effects, and the fact that it is a well-established approach for dynamic modelling. The ARDL methodology offers both theoretical and practical advantages that align well with the structure of our data and the objectives of our study.

Further advantages of the ARDL model are that we can estimate several key outputs in a multi-dimensional perspective relevant to our research objective: standard regression coefficients for each variable and its respective lags, as well as both short-run and long-run multipliers. The ARDL approach thus provides a dynamic structure for identifying temporal relationships between variables, allowing us to evaluate short-term fluctuations alongside long-term cumulative effects. We acknowledge, however, that the ARDL model imposes certain assumptions, most notably linearity and unidirectional causality, from the independent variables to the dependent variable. These assumptions implicitly rule out asymmetric effects or feedback dynamics, which may limit the model's ability to capture more complex market mechanisms.

The short-run multiplier (SRM) effects as derived from an ARDL:

$$\frac{\partial y_t}{\partial x_{j,t}} = b_{j,0} \quad j \in \{1, \dots, k\} \tag{8}$$

The SRM is equal to the lag zero coefficient at time *t* and shows the immediate effect of a change in an independent variable on the dependent variable. The rationale behind the inclusions of SRM is to compare the effects with the long-run multipliers and evaluate whether the effects are persistent, and if so, how they change over time.

The long-run multiplier (LRM) effects as derived from an ARDL:

$$\frac{\partial y_{t+\infty}}{\partial x_{j,t}} = \theta_j = \frac{\Sigma_{l=0}^{q_j} b_{j,l}}{1 - \Sigma_{i=1}^{p} b_{y,i}} \quad j \in \{1, \dots, k\} \tag{9}$$

The LRM is the total accumulated effect on the S&P 500 returns from a change in the explanatory variables when all the short run effects have played out and the system has stabilized. It is computed by summing the short-run coefficients for the explanatory variable and scaling this by the persistence of the dependent variable. The long-run multiplier θ_j indicates how much the equilibrium level of y changes in response to a one-unit increase in x_j, given the internal dynamics of the model. It therefore captures the permanent effect of x_j on y. As for the SRM, the LRM is presented in a single coefficient which can be evaluated on the basis of statistical significance.

The model's robustness was subsequently tested. Based on a Breusch-Godfrey test, the resulting p-value of 0.3633 indicates that there is no statistically significant autocorrelation in the model's residuals (> 0.05). This suggests that the model does not suffer from autocorrelation, which is a positive sign and implies that the estimated coefficients are more reliable. To account for potential heteroskedasticity, robust standard errors were estimated using the White heteroskedasticity consistent covariance matrix (HC0). This ensures that the inference (t- and p-values) remains valid even if the error variance is not constant.

While our main analysis employs an Autoregressive Distributed Lag (ARDL) framework, we carefully considered alternative modeling strategies prior to final model selection. Vector Autoregressive (VAR) models were considered, as they are commonly applied in financial time series analysis and could impose valuable tools to evaluate the effects in our system. Although, a VAR model assumes a system of jointly endogenous variables, making it less suitable for our aim of modeling the unidirectional impact of an independent variable on the S&P 500 index. Having ruled out multidirectional models, a Nonlinear ARDL (NARDL) model could in theory capture potential asymmetries in the effect of gamma exposure by estimating separate effects for positive and negative changes. The NARDL specification requires an ex-ante decomposition of each explanatory variable into positive and negative changes, effectively doubling the number of parameters for each variable. Given our relatively limited sample size, particularly after splitting the data into pre- and post-2020 subperiods, we opted for a more parsimonious model structure to preserve statistical power and avoid overfitting.

5.3 Predictive Models

To evaluate the informational value of the GEX variable, we construct two separate forecasting models based on the previously established ARDL framework. The first model includes lagged values for GEX as predictors, while the second model excludes GEX entirely. This setup allows for a direct comparison of forecast performance with and without the inclusion of GEX, thereby isolating its contribution to the model. In addition, we introduce a naïve random walk model as a baseline benchmark. It is important to emphasize that the primary objective of these forecasting models is *not* to predict stock market movements per se. Rather, the forecasts are used strictly to assess whether GEX adds explanatory power in a statistical and predictive sense. The models are deliberately designed to mirror the structure of our earlier ARDL specifications with adjustments to enhance practical application to ensure that any differences in forecast accuracy can be attributed specifically to the inclusion (or exclusion) of GEX.

To construct the predictive modeling setup, the ARDL model is based on 95% of the full dataset, which resulted in minor changes in the lag structure. There are two major differences from our base model. First, all current values at time t are excluded to mitigate simultaneity and exclude nonrealistic prediction inputs. Second, our objective is to evaluate whether GEX adds informational value within a predictive framework, rather than to compare predictive performance across different time periods. Therefore, there is no analytical benefit in dividing the data into sub-samples for these models. Utilizing the full sample also allows us to maximize the amount of training data available, which enhances model robustness and improves the reliability of out-of-sample forecast evaluation.

Two multivariate linear regression models were constructed using the optimal lag structure from the ARDL framework, with the main difference being the inclusion or exclusion of GEX. The linear regression models are based on the following generalized formula:

$$y_t = \beta_0 + \sum_{i=1}^{p} \beta_i x_{i,t-l_i} + \epsilon_t \tag{10}$$

Where p equals all predictors in the model and l represent all lags for each predictor. This follows the coefficient estimation of the ARDL model, but with all current values excluded in order to predict more realistically. By using this framework, we can: 1) leverage the already established and robustness tested multivariate framework, and 2) isolate the effect of GEX in a predictive model. The predictive models are based on the following structure:

Table 6: Lag structure base for predictive models

	Current Value	Optimal lag order
S&P 500	N/A	$S\&P500_{t-9}$
GEX	N/A	GEX_{t-9}
DIX	N/A	DIX_{t-8}
VIX	N/A	VIX_{t-9}
Put OI	N/A	$PUT\ OI_{t-5}$
Call OI	N/A	$CALL\ OI_{t-7}$
EPU	N/A	*N/A*
SPY Volume	N/A	$SPY\ Volume_{t-5}$

Note: Lags for put- and call open interest changed from 6 lags each to 5 and 7 lags respectively and current values are excluded. Apart from that, the optimal lag structure is identical to our base ARDL model.

The predictions are based on a three-day rolling window forecast structure, where each new round of predictions builds on the latest available historical observations. This framework mimics a realistic out-of-sample forecasting scenario, where the model does not have access to future information but relies on the most recent available data. We chose a three-day forecast horizon to strike a balance between short-term reactivity and medium-term relevance. A single-day forecast may be too sensitive to noise and lack practical significance, while longer horizons may dilute the forecasting relevance. The three-day horizon thus provides a meaningful compromise that allows for timely portfolio rebalancing or signal generation, without overfitting to daily fluctuations. Hence, a three-day rolling window is appropriate for tradable insights but not too far-fetched model wise. The rolling window approach ensures that our model continuously adapts to new data as it becomes available. Rather than retraining the model from scratch for each prediction, we use a fixed training sample and generate multi-step forecasts in overlapping sequences. After each three-day prediction block, the window advances forward by three days, taking the new available data into consideration.

By forecasting in intervals and incorporating all available historical data into future inputs, we aim to simulate a more realistic setting and better evaluate the predictive performance of the model under conditions that resemble real-time investment decision-making. But once again, the primary object of the predictive models is *not* to predict stock market movements per se. Rather, the forecasts are used strictly to assess whether GEX adds explanatory power in a statistical and predictive sense.

5.4 Model comparison and constructing a horse race

Beyond comparing the two gamma including/excluding predictive models, a random-walk model was constructed as a benchmark for evaluating out-of-sample predictive accuracy. The random walk is a widely accepted baseline in financial time series forecasting, particularly for stock returns, due to its simplicity and the efficient market hypothesis. By comparing our models to a random walk, we are able to construct a horse race and assess whether our models provide meaningful predictive improvements over a naïve model that assumes the best forecast for tomorrow's price level is simply the value of today adjusted by a normally distributed error term. To allow for statistical testing and comparison between our models, the random walk model needs to be constructed with all independent variables being expressed in returns, in line with the other models. The random walk forecast is estimated in price level and then converted into returns to enable comparison. The setup follows the general formula for asset pricing with a random walk, expressed in returns:

$$y_t = y_{t-1} + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma^2) \tag{11}$$

To enable a robust comparison between the random walk and the other models the exact same predictive structure is used. The random walk model is, just as the ARDL-derived linear regression models, based on a three-day rolling window forecast structure, where each new round of predictions builds on all available historical observations. To statistically evaluate the difference in forecast accuracy between competing models, we employ the Diebold-Mariano (DM) test as proposed by Harvey, Leybourne and Newbold (1997). The DM test assesses whether one model is superior to another based on the predictive errors from the two models, using the Mean Squared Error measure. The null hypothesis is that two methods have the same forecast accuracy. For string "less" the alternative hypothesis is that the model without GEX is less accurate than the model containing lagged GEX values. For string "two.sided" the alternative hypothesis is that the models including/excluding lagged GEX values have different levels of accuracy. Depending on the string ("less" or "two.sided"), a significant result for "less" indicates that the model including lagged GEX values provides superior forecasts compared to the other beyond what could be attributed to random variation, and for "two.sided" that the two models are statistically different from another in terms of predictive accuracy. These test results are presented in Table 11 in the result section.

The DM results are particularly important to state if there is a statistically significant difference between the models. Besides the DM test, we focus on three widely used performance metrics for forecast evaluation. These metrics capture complementary aspects of forecast accuracy: *Root Mean Squared Error,* penalizes large forecast errors more heavily, making it useful when outliers or volatility are of particular concern. *Mean Absolute Error,* provides a robust measure of average error without disproportionately amplifying larger deviations. *Mean Absolute Percentage Error,* expresses errors in percentage terms, making it easily interpretable and scale-independent, which is helpful when comparing accuracy across time periods. Together, these three metrics provide a comprehensive and interpretable evaluation framework. Although the metrics value the performance of the models, what's important for our research question is not at which accuracy the models predict, but if and how they differ in accuracy.

Root Mean Squared Error:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(Predicted_i - Actual_i)^2}{N}} \tag{12}$$

Mean Absolute Error:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Predicted_i - x| \tag{13}$$

Mean Absolute Percentage Error:

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{Predicted_i - Actual_i}{Actual_i}\right| \tag{14}$$

5.5 Research ethics

The thesis adheres to good research practice in accordance with the principles outlined by the Swedish Research Council (Vetenskapsrådet, 2024). In this study, artificial intelligence (AI) has been used exclusively as a language support tool to improve grammar, sentence structure, and stylistic clarity. The content, analysis, and conclusions have been entirely formulated by the authors. ChatGPT has served as an aid for rephrasing sentences, identifying potential linguistic issues, and suggesting structural improvements. All AI-generated suggestions have been manually reviewed and edited, in accordance with ethical guidelines for academic AI usage (Lim, Gottipati and Cheong, 2023). This approach ensures that the content, reasoning, and research contributions of the thesis remain the sole work of the authors.

6. Results

The interpretation for the GEX coefficients will, to enhance readability, be presented as the coefficients from the result tables followed by a parenthesis containing the equivalent value if the change in the GEX derivative was a one-billion-unit change. Example: a one-unit change in the GEX derivative leads to an increase of approximately 2.44e-12 in the S&P 500. This is equivalent to a one billion unit increase in the GEX derivative resulting in a gain of 0.24% in the S&P 500. This will be presented as 2.44e-12 (0.24%) in the text below. During the full period of data, the GEX derivative has changed with one billion units or more for 846 days (≈25% of observations), which makes it a fully reasonable adjustment for interpretation.

6.1. Statistical relationship (Pre 2020)

A. Granular view

Table 7 presents the results from our ARDL model for the pre-2020 period, with coefficients specified for each variable and the model's maximum number of lags. For all variables in the table, the coefficients are based on daily changes except for GEX, for which we use the first derivative.

Table 7: Coefficients from the ARDL-model (Pre 2020)

	t	t-1	t-2	t-3	t-4	t-5	t-6	t-7	t-8	t-9
S&P 500	N/A	-0.0352	-0.0145	-0.0696*	-0.008	-0.0541	-0.0161	-0.027	-0.0374	-0.0151
GEX	2.44e-12***	7.03e-13**	4.54e-13	7.59e-13**	2.29e-13	1.81e-13	1.77e-13	2.81e-13	3.45e-13	3.26e-13
DIX	-0.03***	-0.0144***	-0.012***	-0.0107**	-0.0045	-0.0085*	-0.0053	-0.0103**	-0.0038	N/A
VIX	-0.0744***	-0.0059	-0.0039	-0.0035	0.000268	0.001	-0.0045	-0.0042	-0.0026	-0.0012
Put OI	0.0053	-0.0036	5.32e-04	0.0044	0.0075	0.0016	1.53e-05	N/A	N/A	N/A
Call OI	-0.0034	0.004	-0.0017	-0.004	-0.0085	-0.0032	-0.0016	N/A	N/A	N/A
EPU	1.73e-04	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SPY Volume	0.0011*	-2.03e-05	2.36e-04	-4.28e-04	-6.06e-04	6.48e-04	N/A	N/A	N/A	N/A

Note: Columns represent lag structures from the current day (t) to nine days prior (t−9). "NA" indicates that the variable was not included at that specific lag, based on the optimal lag selection process. Statistical significance is denoted as follows: *** for $p < 0.001$, ** for $p < 0.01$, and * for $p < 0.05$.

For GEX, the effect on the S&P 500 returns is statistically significant on the same day *(t)*, the previous day *(t−1)*, and three days prior *(t−3)*. The second day prior *(t-2)* is significant at a 10% level, which we do not incorporate in our significance levels, but still could be of interest for the economic interpretation. The effects are positive, indicating that a positive change in the derivative of GEX contributes to positive returns in the S&P 500. The strongest relationship is observed on the same day, significant at the 0,1% level (***), where a one-unit change in GEX coincide with an increase of approximately 2.44e-12 (0.24%) in the S&P 500. GEX thus appears to have a strong immediate/contemporaneous effect on equity market returns, which supports our theoretical framework and prior studies on the subject presented earlier. However, it remains difficult to definitively establish a causal relationship between GEX and the S&P 500, which makes it particularly interesting that the effect is also confirmed at lagged values. The coefficient for lag one and lag three suggests that a one billion unit change in GEX would result in a gain of 0.07% and 0.008% on the S&P 500 respectively.

These results show that changes in GEX from one- and three-days prior are significantly and positively associated with S&P 500 returns. In practical terms, this suggests that an increases in the derivative of the gamma exposure tend to precede upward movements in the equity market. The findings provide evidence of a short-term transmission mechanism between shifts in GEX and stock market behavior, aligning with the underlying theory of delta-hedging feedback effects previously described. While focusing on the implications of changes in GEX, Dubois (2022) identified a connection between GEX and the S&P 500 around option expiration periods. Our study extends the analysis by showing that GEX is significantly associated with S&P 500 returns when taking all trading days in consideration, and over an extended time-period.

The statistically significant coefficients observed for the GEX derivative and SPY trading volume at time *t* suggest an immediate relationship between changes in option market positioning, equity market volume and S&P 500 returns. A plausible mechanism underlying this result begins with a shift in gamma exposure, which alters the hedging requirements of OMMs. As delta-hedging must be adjusted in response to changes in gamma, OMMs often rebalance their exposure in highly liquid instruments, such as the SPY ETF. Since SPX-options is cash-settled, and OMMs must use a proxy instrument for hedging when trading the underlying asset, increased volume in the SPY ETF could be a result of OMM hedging activities. This hedging activity can generate substantial trading volume in SPY, particularly given the scale and liquidity of the index options market. Increased buying or selling pressure in SPY mechanically propagate into the underlying basket of S&P 500 stocks. This flow-based transmission channel implies that option-driven hedging activity may indirectly affect the valuations of the index's components through passive rebalancing, thus exerting an immediate influence on S&P 500 returns. Hence, our findings support the flow-based transmission mechanism, highlighting our standpoint on how the interplay between the option market and equity flows can drive short-term price movements.

Both DIX and VIX are significant at the 0.1% level (***) on the same day, with VIX's negative relationship to equity market returns being generally accepted, something that is also confirmed in our study. DIX exhibits a very strong level of significance from time *t* through *t–3*, whereas the relationship between the S&P 500 and VIX is only statistically significant on the contemporaneous day. All significant coefficients for both DIX and VIX show a negative relationship with the S&P 500. Neither of the open interest metrics or the EPU variable is significant at any lag in our model.

B. Long and short-run multiplier effects

The short-run multiplier for GEX is approximately 2.44e-12 (0.24%) and statistically significant at the 0.1% level (***). The long-run multiplier for GEX is approximately 4.62e-12 (0.46%) and statistically significant at the 0.1% level (***), indicating a highly robust relationship between changes in gamma exposure and movements in the S&P 500. This suggests that the effect of changes in the GEX derivative persists and accumulates over time, as shown by the almost twice as large, long-run vs. short-run coefficient. The fact that the effect is increasing is not necessarily surprising since the long-run effects weigh in the effect when all the short run effects have worn off. It is particularly interesting that the long-run effect is statistically significant at the 0.1% level. This result indicates that gamma dynamics not only impact short-term price behavior through hedging flows, but also contribute to longer-term market trends, potentially via persistent positioning effects or regime shifts in market sentiment and liquidity.

Table 8: Multiplier coefficients (Pre 2020)

	Short run	**Long run**
(Intercept)	0.0009***	0.0007***
GEX	2.44e-12***	4.62e-12***
DIX	-0.03***	-0.078***
VIX	-0.074***	-0.077***
Put OI	0.0053	0.012
Call OI	-0.0034	-0.015
EPU	0.00017	0.00014
SPY Volume	0.0011**	-0.0003

Note: The short-run column reflects the immediate effect of changes in each explanatory variable on S&P 500 returns, while the long-run column captures the total accumulated effect after all lags have been accounted for. Statistical significance is denoted as follows: *** for $p < 0.001$, ** for $p < 0.01$, and * for $p < 0.05$.

Gamma-related hedging flows may act as a signal to the broader market by altering volatility and creating directional pressure (either buying or selling) that other market participants respond to and potentially amplify. A plausible explanation for the observed long-term persistence (over one lag to several days) is that the hedging activity of OMMs is not itself the primary driver of returns beyond the immediate horizon. Rather, these flows may influence market psychology, prompting follow-on behavior from other investors who interpret the initial movements as meaningful, thereby extending the impact through sustained trading activity.

The short run multipliers show statistically significant coefficients for DIX (***), VIX (***) and SPY Volume (**). The long run multiplier coefficients for DIX (***) and VIX (***) are still significant while SPY volume is no longer statistically significant.

6.2. Statistical relationship (Post 2020)

A. Granular view

For the post-2020 period, once again, we identified a strong contemporaneous relationship between GEX and the S&P 500 with high statistical significance, although the effect is weaker in this period. Specifically, the immediate effect of GEX has decreased from 2.44e-12 (0.24%) to 8.79e-13 (≈ 0.09%), suggesting a reduced sensitivity of the equity market to changes in the GEX derivative. One potential explanation is a shift in positioning dynamics post-2020, possibly linked to changes in options market participation and a higher usage of short-lived options such as 0DTE options.

Table 9: Coefficients from the ARDL-model (Post 2020)

	t	t-1	t-2	t-3	t-4	t-5	t-6	t-7	t-8	t-9
S&P 500	N/A	-0.1334	0.0398	-0.0282	-0.1455**	-0.0083	-0.1031	0.1184*	-0.0271	0.1353*
GEX	8.79e-13***	2.89e-13*	2.63e-13*	7.07e-14	6.64e-14	8.13e-14	2.77e-14	-4.48e-14	2.54e-14	2.98e-14
DIX	-0.0112	-0.0103	0.0169*	0.015	0.0074	-1.80e-04	-0.0108	-0.0054	-0.0135	N/A
VIX	-0.1019***	-0.019*	-0.0101	-0.0104	-0.0178**	-0.0084	-0.0103*	2.39e-04	0.0026	0.0029
Put OI	0.0047	-0.0011	0.004	0.0017	0.0056	0.0087*	0.0065	N/A	N/A	N/A
Call OI	0.0042	-0.0017	0.0033	0.0012	-0.0024	-0.0067**	-0.0068*	N/A	N/A	N/A
EPU	1.18e-04	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SPY Volume	0.0022*	0.0022*	0.0026*	0.002	8.07e-05	-5.25e-04	N/A	N/A	N/A	N/A

Note: Columns represent lag structures from the current day (t) to nine days prior (t−9). "NA" indicates that the variable was not included at that specific lag, based on the optimal lag selection process. Statistical significance is denoted as follows: *** for $p < 0.001$, ** for $p < 0.01$, and * for $p < 0.05$.

The lagged effect over the subsequent days is statistically significant, albeit at a lower level than pre-2020, and persist up to two days at the 5% significance threshold (*). The main difference compared to the first period is that lag two now is statistically significant while lag three is not. The coefficient for the prior day *t-1* decreases from 7.03e-13 (0.07%) to 2.89e-13 (≈ 0.03%). This implies that not only the immediate effect from GEX to the S&P 500 returns is reduced, but also the lagged effects. The coefficient for the second lag is now 2.89e-13 (0.03%) compared to 4.54e-13 (≈ 0.05%), a reduction in line with lag one and zero. Although, the second lag was only statistically significant on the 10% level in pre-period, why the specific evaluation should be done with caution.

Taken together, the results suggest a weakening of the relationship between GEX and S&P 500 returns in the latter period of the sample. While the effects remain statistically significant across several lags, their magnitude has decreased consistently at lag zero, one, and two. This reduction implies that the responsiveness of the equity market to changes in the GEX derivative has diminished since 2020, both in terms of immediate and short-term dynamics. Nevertheless, while the effect has weakened, the fact that it remains statistically detectable suggests that GEX still retains informational value, albeit to a lesser degree.

The relationship between DIX and the S&P 500, however, loses its statistical importance during this period as all lags, except from the second lag, becomes insignificant. At the same time, the negative relationship between VIX and the equity market remains strongly significant at time *t*, with a slightly larger coefficient than in the earlier period. Furthermore, the effect of VIX appears to be more persistent in the post-2020 period, with the first lag (*t*–*1*) also significant at the 5% level. This indicates that an increase in VIX now tend to have a more meaningful and prolonged impact on the S&P 500, typically resulting in negative returns over a two-day horizon. This supports the interpretation of rising volatility as a stronger and more reliable risk-off signal in recent years. Trading volume in the SPY ETF shows a statistically significant relationship on the 5% level (*) affecting S&P 500 returns for up to two days. In the earlier period, this effect was only observed in time *t* while it was not observed at all for lag one or two.

B. Long and short-run multiplier effects

The short-run multiplier for GEX in the second period is approximately 8.79e-13 (≈ 0.09%) and statistically significant at the 0.1% level (***). The long-run multiplier for GEX is approximately 1.47e-12 (0.15%) and statistically significant at the 1% level (**), which once again shows a meaningful relationship between changes in gamma exposure and movements in the S&P 500. Both coefficients (long and short-run) have decreased compared to the pre-2020 period, but just as before the long-run effect remains stronger.

Table 10: Multiplier coefficients (Post 2020)

	Short run	**Long run**
(Intercept)	0.0009**	0.0007**
GEX	8.79e-13***	1.47e-12**
DIX	-0.112*	-0.010
VIX	-0.102***	-0.15***
Put OI	0.0047	0.026
Call OI	0.0042	-0.008
EPU	0.00012	0.00010
SPY Volume	0.0022*	0.0074*

Note: The short-run column reflects the immediate effect of changes in each explanatory variable on S&P 500 returns, while the long-run column captures the total accumulated effect after all lags have been accounted for. Statistical significance is denoted as follows: *** for $p < 0.001$, ** for $p < 0.01$, and * for $p < 0.05$.

There are several possible explanations behind the decreasing effect when comparing the pre- and post-periods. First, GEX has had an unusual development since 2020 with more frequent extreme swings as shown in Figure 10. As previously shown by the data, GEX moved with at least one billion units in 13% of the days pre 2020 compared to 32% of the days in the post 2020 period. These swings are likely to be driven by the increased options trading activity which this study aims to explore, but the swings could also induce noise in the data which possibly affect the predictive power of GEX negatively. Furthermore, the increased swings in gamma exposure may dampen the implicit directional market signals generated by OMM hedging activity. This could help explain why the previously suggested follow-on behavior appears less pronounced after 2020 compared to the earlier period. Although this might affect the relationship between GEX and S&P 500 returns negatively, the results still imply a statistically significant relationship which motivates the decisions to evaluate GEX within a predictive modeling framework.

Both short and long-run multipliers show statistically significant coefficients for VIX (***) and SPY Volume (*).

These findings align with the results in Egebjerg and Kokholm (2024), but on a different time frame and when using a different quantitative approach. Their study demonstrated that changes in OMMs delta positions are a significant predictor of SPX future returns at the end of the trading session, while the result in our study implies that the effects last longer than so. Our results imply that the long-run multiplier effect is stronger than the short-run multiplier effect, even though the coefficient for specific lags tend to fade. Pearson et al. (2007) document a statistically and economically significant negative relationship between stock volatility and the net gamma exposure. Their findings suggest that approximately 12% of the daily absolute returns of stocks with active options trading can be explained by hedging-related rebalancing flows in the underlying asset. Our research builds upon this, stretching the results from single stocks and absolute returns to an aggregated equity benchmark and the directional impact of hedging flows. Several of the prior studies on this subject employ either a short-term time frame (intra-day), such as Baltussen et al. (2021), or a longer-term time frame (up to a month) as studied in Soebhag (2023).

6.3 Predictive Models Comparison

Below, the results from our forecast evaluation is presented, comparing the predictive performance of three models: (i) the model specification including GEX, (ii) the same model excluding GEX, and (iii) a random walk benchmark. The primary objective of this comparison is not to achieve accurate short-term return forecasts per se, but rather to assess whether the inclusion of GEX improves predictive performance in a statistically meaningful way. The evaluation is conducted using out-of-sample forecasts over the final 5% of the dataset, and performance is assessed through standard forecast error metrics and formal statistical testing.

When evaluating the forecast accuracy metrics for all our models, the results in Table 11 implies that the model including GEX have superior predictive accuracy compared to the other models. Although our models are not optimized for predictive accuracy, all the accuracy measures point out that including GEX into the predictive model enhances its accuracy.

Table 11: Predictive accuracy measures

	Model including GEX	**Model excluding GEX**	**Random Walk**
RMSE	0.008462	0.008499	0.01221
MAE	0.006133	0.006168	0.008764
MAPE	207.19%	211.83%	547.62%

Note: All models are constructed and evaluated utilizing the full data sample (2011-2025). Lower values indicate lower forecast errors. Each model should be compared to one another by each accuracy measure respectively.

In table 12, the model specification including GEX is compared with the other two models using the Diebold-Mariano test. The accuracy evaluation within the test is based on our chosen loss function, the Mean Squared Error. Since all forecasts are based on a rolling three-day window, we compare the models against each other on day *t+1, t+2* and *t+3*. By doing so, we can evaluate the models throughout the entire forecast horizon and ensure a more comprehensive comparison.

Table 12: Model comparison using Diebold-Mariano test

	Model excluding GEX			**Random Walk**		
	t+1	t+2	t+3	t+1	t+2	t+3
P-value ("two.sided")	0.0525	0.02027*	0.0248*	<0.0000***	<0.0000***	<0.0000***
P-value ("less")	0.02625*	0.01013*	0.0124*	<0.0000***	<0.0000***	<0.0000***
DM Statistic	-1.9527	-2.3434	-2.2646	-4.117	-4.6555	-5.7675

Note: All models are constructed and evaluated utilizing the full data sample (2011-2025). The test "two.sided" evaluates if the models statistically differ from each other. The test "less" evaluates whether the model including GEX outperform the other models. The DM statistics are results from the hypothesis tests, where more negative values indicate stronger difference/outperformance. The significance codes used are: *** for $p < 0.001$, ** for $p < 0.01$, and * for $p < 0.05$.

The two-sided test version implies that the model including GEX is statistically different from the model without GEX in predictive accuracy at *t+2* and *t+3* at a 5% significance level. Also, for forecast horizon *t+1*, the model including GEX is statistically different at a 10% significance level (although very close to 5%). When comparing the model including GEX with the random walk-model, the two-sided test exhibits clear statistically significant difference in predictive accuracy at a 0.1% level for all forecast horizons. This means, including GEX into our predictive framework statistically differs the accuracy of predicting the S&P 500 returns compared to the other models. The one-sided test ("less") tests for predictive outperformance and allows for conclusions regarding whether one of the models is *better* than another. The model including GEX is statistically better than both the model excluding GEX and the random walk model in predictive accuracy. This is proven by the significant p-values at a 5% (although close to 1% for *t+2* and *t+3*) compared to the model excluding GEX, and 0.1% compared to the random walk model.

These tests combined show that including GEX into the models does not only differ the predictive accuracy, but statistically proves that the model gets better. All three Diebold-Mariano tests suggest that the model including GEX have significantly lower predicting errors compared to both other models over the entire forecast horizon. This implies that GEX adds valuable information to the model and enhances its capability to predict the S&P 500 returns.

The results from our analyses not only demonstrate that incorporating GEX into the predictive model improves forecasting accuracy, but they also provide robust statistical evidence that its inclusion significantly enhances the model's performance across all examined prediction horizons.

7. Conclusion

This study set out to answer two central research questions. First, we examined whether changes in aggregated gamma exposure influence future stock market movements. Our results clearly indicate that variations in the derivative of GEX have a statistically significant relationship with subsequent returns on the S&P 500. This effect is robust across both the pre- and post-2020 subperiods, albeit with somewhat diminished strength in the latter. These findings suggest that shifts in gamma positioning among OMMs, likely driven by delta-hedging dynamics, can generate price effects that persist beyond intraday horizons and into the following trading days.

Second, we assessed whether the inclusion of GEX in a predictive modeling framework improves the forecast accuracy of S&P 500 returns. Based on out-of-sample testing, including a model comparison using Diebold-Mariano tests, we find that incorporating GEX significantly enhances the model's forecasting performance relative to both a GEX-excluding specification and a random walk benchmark. This outcome reinforces the idea that GEX contains forward-looking informational value, which can be utilized to improve predictive modeling in equity markets.

Together, these results affirm that GEX is not only statistically linked to future equity returns but also adds meaningful predictive power within a structured econometric framework. In doing so, the study provides empirical evidence that GEX has the potential to serve as a valuable tool for tactical capital allocation.

The improved predictive performance of the model including GEX supports the notion that option-market-derived variables play a meaningful role in shaping short-term return dynamics. These findings are consistent with previous literature suggesting that option market flows, especially delta-hedging behavior by OMMS, generate mechanically induced feedback loops that can influence equity prices (e.g., Egebjerg and Kokholm, 2024; Pearson et al., 2007). Our results reinforce this view by demonstrating that GEX maintains explanatory power even in a model that controls for volatility (VIX), institutional flows (DIX), option open interest, and trading volume. Our research also extends the understanding of how gamma exposure affects the equity market. While Baltussen et al. (2021) and Barbon and Buraschi (2021) demonstrated that intraday hedging flows influence return dynamics in the S&P 500, our findings suggest that changes in gamma exposure provide both predictive and explanatory power over longer time horizons beyond the intraday level. Although the effect of changes in gamma exposure might be amplified over shorter time horizons or during specific events, such as option expiration dates as highlighted by Dubois (2022), our findings demonstrate that gamma exposure remains an important factor in explaining return dynamics across broader market conditions.

The fact that the forecast model including GEX outperform the random walk model is not surprising, although it is a well-established way of measuring and comparing predictive models. Since the ARDL-framework includes lagged values of S&P 500, it basically includes the main information that drives the random walk process (excluding the error term). The framework also adds relevant information into the forecast, i.e. the lagged values of the control variables, which helps in shaping more accurate forecasts. Although the model itself is not that accurate in predictive terms, it was expected to at least be better than the random walk. The main result is the statistically significant improvement of including GEX compared to the model excluding GEX, which suggests that aggregate gamma exposure provides informational value.

A particularly notable observation is that while the strength of the relationship between changes in GEX and S&P 500 returns appears to weaken post-2020, it remains statistically significant. This could reflect structural changes in the options market such as increased use of short-term contracts (e.g., 0DTE options), which may introduce gamma related noise outside the scope of the aggregated GEX index, thereby diluting the clarity of the signal. Alternatively, broader macroeconomic developments including heightened geopolitical tensions, persistent inflationary pressures, the COVID-19 pandemic, and recurring fears of recession, may have increased overall market uncertainty and introduced additional drivers of equity market returns. These forces could dampen the influence of market microstructure variables like gamma exposure, making it more challenging to isolate their effect in a noisier post-2020 environment. However, the persistence of statistical significance even under changing market environments and noise may itself be interpreted as a strength of the GEX signal, indicating its robustness across regimes.

At the same time, several limitations warrant discussion. First, GEX is constructed under simplifying assumptions, including the idea that OMMs always hedge to delta and that their positioning reflects aggregate net gamma. In reality, the market is more heterogeneous: not all option trades involve delta-hedging participants, and investor behavior may vary across regimes. Hedging behavior is not always symmetrical or continuous and the assumption that OMMs hedge exactly to delta ignores the use of wider hedging bands in practice. Second, the GEX measure does not account for intraday or 0DTE option dynamics, which as stated by Cboe (Xu, 2023), have grown significantly in recent years. This may contribute materially to short-term price action without affecting daily open interest. As options approach expiration, gamma becomes more sensitive and peaks when the option is at or near the money, leading to larger gamma-contribution into the market. Another measure of gamma exposure could be used when replicating this study to capture more detailed market information, such as the gamma exposure from short-live options. This study focused on GEX due to its availability, high usage and the self-fulfilling prophecy factor while being fully aware of the limitations.

From a methodological standpoint, this study exclusively employs an ARDL-based modeling framework. While ARDL is well-suited for capturing lagged and dynamic relationships, it may not fully accommodate potential non-linearities or regime shifts that characterize modern financial markets. Future research could explore alternative non-linear models and forecasting architectures, such as machine learning or regime-switching models.

For data-driven asset managers and tactical allocators, the practical implications of our findings suggest that derivatives-based signals such as GEX can serve as valuable components within a broader framework for managing risk-on and risk-off positioning. Furthermore, GEX can be used not only as a predictive indicator for anticipating short-term shifts in equity returns, but also as an explanatory variable that provides insight into market movements after they have occurred. Even once initial gamma-driven dynamics have played out, the changes in gamma exposure retains informational value that can support decision-making regarding portfolio allocation and risk adjustment. The statistically validated relationships between changes in GEX and S&P 500 returns, both in terms of explanatory power and predictive strength, contribute to a more evidence-based approach to portfolio management.

8. References

Anderegg, B., Ulmann, F. and Sornette, D., (2022). The impact of option hedging on the spot market volatility. *Journal of International Money and Finance*, 124, p.102627. https://doi.org/10.1016/j.jimonfin.2022.102627

Baker, S.R., Bloom, N. and Davis, S.J., (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, 131(4), pp.1593–1636. Available at: http://www.policyuncertainty.com/media/BakerBloomDavis.pdf [Accessed 7 Feb. 2025].

Baltussen, G., Da, Z., Lammers, S. and Martens, M., (2021). Hedging demand and market intraday momentum. *Journal of Financial Economics*, 142(1), pp.377–403. https://doi.org/10.1016/j.jfineco.2021.04.029

Barbon, A. and Buraschi, A. (2021). *Gamma Fragility*. University of St. Gallen, School of Finance Research Paper No. 2020/05. First published November 2020; last revised March 2021. https://doi.org/10.2139/ssrn.3725454

Bryzgalova, S., Pavlova, A. and Sikorskaya, T. (2023). Retail trading in options and the rise of the big three wholesalers. *Journal of Finance*, 78(6), pp.3465–3514. https://doi.org/10.1111/jofi.13285

Calhoun, G., (2021). GameStop, The Second Surge: Anatomy Of A 'Gamma Swarm'. *Forbes*, [online] 10 March. Available at: https://www.forbes.com/sites/georgecalhoun/2021/03/10/gamestop-the-second-surgeanatomy-of-a-gamma-swarm/ [Accessed 24 Feb. 2025].

Cboe Global Markets, (2025). *Cboe VIX FAQ*. [online] Available at: https://www.cboe.com/tradable_products/vix/faqs/ [Accessed 8 Mar. 2025].

Cboe Options Exchange, (2021). *Cboe Options Exchange to extend global trading hours*. [online] Available at: https://cdn.cboe.com/resources/release_notes/2021/Cboe-Options-Exchange-to-Extend-Global-Trading-Hours.pdf[Accessed 11 Feb. 2025].

Chen, W.-P., Chung, H. and Lien, D., (2016). Price discovery in the S&P 500 index derivatives markets. *International Review of Economics and Finance*, 45, pp.438-452. https://doi.org/10.1016/j.iref.2016.07.008

CME Group, (2025). Options Gamma – The Greeks. [online] CME Group. Available at: https://www.cmegroup.com/education/courses/option-greeks/options-gamma-the-greeks.html [Accessed 24 Feb. 2025].

CNBC, (2017). *Death of the human investor: Just 10 percent of trading is regular stock picking, J.P. Morgan estimates.* CNBC, [online] 13 Jun. Available at: https://www.cnbc.com/2017/06/13/death-of-the-human-investor-just-10-percent-of-trading-is-regular-stock-picking-jpmorgan-estimates.html [Accessed 24 Feb. 2025].

Craney, T.A. and Surles, J.G., (2002). Model-dependent variance inflation factor cutoff values. *Quality Engineering*, 14(3), pp.391–403. https://doi.org/10.1081/QEN-120001878

Dubois, A.J.G., (2022). *Predicting the options expiration effect using machine learning models trained with gamma exposure data.* Kandidatrapport. KTH Royal Institute of Technology. Available at: https://www.diva-portal.org/smash/get/diva2:1705054/FULLTEXT01.pdf [Accessed 10 Feb. 2025].

Deboeck, P.R., 2010. Estimating dynamical systems: Derivative estimation hints from Sir Ronald A. Fisher. *Multivariate Behavioral Research*, 45(5), pp.725-745. https://doi.org/10.1080/00273171.2010.498294

Egebjerg, S. and Kokholm, T., (2024). *A model for the hedging impact of option market makers.* Working paper. Aarhus University & Danish Finance Institute. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4936978 [Accessed 8 Feb. 2025].

Fama, E.F., (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25(2), pp.383–417. https://doi.org/10.1111/j.1540-6261.1970.tb00518.x

Harvey, D.I., Leybourne, S.J. and Newbold, P., (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, 13(2), pp.281–291. https://doi.org/10.1016/S0169-2070(96)00719-4

Ingrassia, P., (2021). *How Robinhood has revolutionized online trading and dramatically upended the traditional model for Payment for Order Flow (PFOF).* Virginia Law & Business Review, 16(1), pp.1–33. [online] Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3832424 [Accessed 22 May 2025].

Kripfganz, S. and Schneider, D.C., (2023). ardl: Estimating autoregressive distributed lag and equilibrium correction models. *The Stata Journal*, 23(4), pp.983–1019. https://doi.org/10.1177/1536867X231212434

Leoni, P., (2014). *The Greeks and Hedging Explained.* Financial Engineering Explained Series. Basingstoke: Palgrave Macmillan.

Lim, T., Gottipati, S. and Cheong, M.L.F., (2023). Ethical considerations for artificial intelligence in educational assessments. In: *Creative AI Tools and Ethical Implications in Teaching and Learning.* Hershey, PA: IGI Global, pp.32–79. Available at: https://www.researchgate.net/publication/376951624_Ethical_Considerations_for_Artificial_Intelligence_in_Educational_Assessments [Accessed 8 Mar. 2025].

Lo, A.W. and MacKinlay, A.C., (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. *The Review of Financial Studies*, 1(1), pp.41–66. Available at: https://www.jstor.org/stable/2962126 [Accessed 8 Mar. 2025].

Morningstar, (2024). *Recovery in US fund flows was weak in 2023.* Morningstar, [online] Available at: https://www.morningstar.com/funds/recovery-us-fund-flows-was-weak-2023 [Accessed 24 Feb. 2025].

Natsiopoulos, K. and Tzeremes, N., (2023). ARDL: ARDL, ECM and Bounds-Test for Cointegration. *R package version 0.2.4.* Available at: https://CRAN.R-project.org/package=ARDL [Accessed 6 Feb. 2025].

OCC, (2020). *OCC clears over 4.9 billion contracts in 2019. Options Clearing Corporation*, [online] 2 January. Available at: https://www.theocc.com/newsroom/press-releases/2020/occ-clears-over-4-9-billion-contracts-in-2019 [Accessed 8 Feb. 2025].

OCC, (2025). Annual 2024 and December 2024 volume. *Options Clearing Corporation,* [online] 3 January. Available at: https://www.theocc.com/newsroom/views/2025/01-03-occ-annual-2024-and-december-2024-volume [Accessed 5 May. 2025].

Pearson, N.D., Poteshman, A.M. and White, J., (2007). *Does option trading have a pervasive impact on underlying stock prices?* SSRN working paper. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=970592 [Accessed 8 Feb. 2025].

Poser, S.W., (2023). Trends in options trading. *New York Stock Exchange*, [online] Available at: https://www.nyse.com/data-insights/trends-in-options-trading [Accessed 24 Feb. 2025].

SIFMA, (2024). *SIFMA Insights: Equity Market Structure Compendium*. New York: Securities Industry and Financial Markets Association. Available at: https://www.sifma.org/wp-content/uploads/2024/03/SIFMA-Insights-Equity-Market-Structure-Compendium_as-of-2-26-1.pdf [Accessed 7 May 2025].

Soebhag, A., (2023). Option gamma and stock returns. *Journal of Empirical Finance*, 74, Article 101442. https://doi.org/10.1016/j.jempfin.2023.101442

SqueezeMetrics, (2017). *Gamma Exposure (GEX): Quantifying hedge rebalancing in SPX options*. Available at: https://squeezemetrics.com/monitor/download/pdf/white_paper.pdf [Accessed 17 Jan. 2025].

SqueezeMetrics, (2018). *SHORT IS LONG: Using dark pool short sales as a proxy for buying activity*. Available at: https://squeezemetrics.com/monitor/download/pdf/short_is_long.pdf [Accessed 17 Jan. 2025].

Vetenskapsrådet, (2024). *God forskningssed*. Stockholm: Vetenskapsrådet. Available at: https://www.vr.se/download/18.4c9f221a191e4edf9053a474/1727853946433/God%20forskningssed%20VR%202024.pdf [Accessed 28 Mar. 2025].

Xu, M., (2023). Volatility insights: Much ado about 0DTEs – Evaluating the market impact of SPX 0DTE options. *Cboe Global Markets*, [online] Available at: https://www.cboe.com/insights/posts/volatility-insights-evaluating-the-market-impact-of-spx-0-dte-options/ [Accessed 9 Feb. 2025].