Divining the Future with Option Flows: A Volatility Model Extension of Gamma Exposure and the Net Options Pricing Effect

Jonas Paul Schmitten
Noah Julian Angara
Moritz Leonhardt Köhler
Giovanni Magagnin

May 2021

1 Introduction

Since the 1970s, derivatives have steadily increased their importance in financial markets. Be it as a risk management tool or for speculation purposes, the use of derivatives, especially equity derivatives, is as important as ever. So important have derivatives become that literature has emerged documenting the effect of market makers' hedging behavior through derivatives on the price of their underlying (see Barbon & Buraschi, 2020). More recently, the Gamestop saga has shown how option Greeks such as gamma can lead to large swings in the price of the underlying due to market maker hedging (a so-called gamma-squeeze).

To structure the discussion around options and their effect on the underlying one needs a pricing model. With their seminal paper, Black & Scholes (1973) introduced a formal framework for pricing options. The only unobservable input in their pricing formula is the measure of volatility. This makes obtaining an accurate forecast of volatility paramount in correctly pricing an option and obtaining correct measures of the option Greeks. A popular estimate is implied volatility, which results from solving the Black-Scholes equation for the volatility using current option prices. However, we argue that using forecasted volatilities from conditional heteroscedastic models such as GARCH leads to more accurate option Greeks. Hence, the goal of this paper is to show how forecasted volatilities can be used to better document the predictive power of Greeks on the price of the underlying asset.

We contribute to the literature in several ways. First, by forecasting financial volatility, we argue to get more precise volatility estimates compared to implied volatility. We use the predicted volatility to get accurate option prices and Greeks. Finally, we show that the predicted Greeks can consistently predict the openings of the underlying asset via thresholds in the Net Option Pricing Effect (NOPE) and Gamma Exposure (GEX). To the best of our knowledge, we are the first to employ GARCH-type volatility prediction models to complement and improve models which account for the effect delta-gamma hedging has on the price of the underlying.

This paper proceeds as follows. First, we provide an overview of the existing literature on volatility models, option pricing, and on how market makers' hedging requirements affect price movements in the underlying. In particular, we show how we can bridge the gap between different strands in the literature. Second, the utilized data is described and the necessary pre-processing steps are explained. Third, we focus on the mathematical foundations and derive the volatility, and delta-gamma hedging models. We emphasize how NOPE and GEX can be improved through our volatility forecasting techniques. Lastly, we present the results and discuss potential limitations as well as future applications.

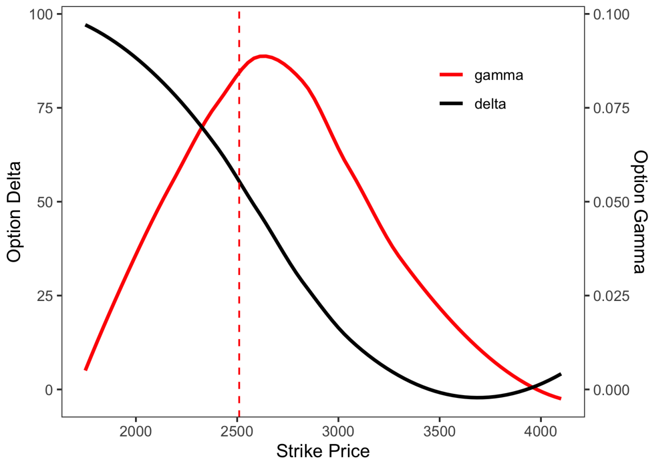
2 Literature Review

Volatility is the single most important factor in the pricing of derivative securities. In order to price options, the volatility of the underlying assets is needed, for the whole duration of the contract. Given that volatility is not directly observable, accurately estimating and forecasting volatility is crucial in correctly pricing an option. To have accurate volatility forecasts, it is fundamental to understand the peculiarities of financial time series. The literature on financial volatility gives a clearer idea of the specific features of financial time series, and which volatility models can be used to address them (see Poon & Granger (2003) for a comprehensive review). The most salient features of financial time series are the following: fat-tailed distributions of the returns, volatility clustering, leverage effects, and seasonality. The literature proposes three types of models to describe the evolution of volatility. The first category assumes a parametric form governing the time dynamics of volatility, while the second employs stochastic equations as an additional building block to describe volatility. The third category relies on realized volatility implied by current option prices. This paper focuses on the first category. In his seminal paper, Engle (1982) laid the groundwork for modeling the time-varying dynamics of the conditional variance with his Autoregressive Conditional Heteroscedasticity (ARCH) model, followed by the introduction of a generalized ARCH (GARCH) by Bollerslev (1986). The inability to capture some of the stylized facts seen in the data has led to the proliferation of extensions of the standard GARCH model such as the E-GARCH, GJR-GARCH, T-GARCH and MS-GARCH models. For example, E-GARCH specifies the conditional variance in its logarithmic form which allows to capture the stylized facts of a higher conditional variance in periods subsequent to negative shocks. These models are often the most popular structure for forecasting financial volatility. Relevant application of forecasted volatility can be found in option trading, and market making.

As mentioned above, volatility is unobservable and hence forecasts are needed as an input in option pricing formulae, the most famous one being the Black-Scholes option pricing formula. There are two main approaches in computing volatility forecasts for option pricing. The first relies on calculating realized volatility from historical price data, while the second approach uses the volatility implied by backsolving the Black-Scholes equation with current prices. Embedded in the latter approach is the assumption that the market's expectation of future volatility is informationally efficient. While this assumption is intuitively appealing, the consensus in the literature is not yet clear on the predictive power of implied volatility. Canina & Figlewski (1993) find that implied volatility is a poor predictor for future volatility on S&P100 index options. On the other hand, Szakmary, Ors, Kim, & Davidson III (2003), Christensen & Prabhala (1998) and Jorion (1995) find implied volatility to have superior predictive power compared to econometric methods employing historical volatility. Further examples of using implied volatility for predictive purposes are shown by Ahoniemi (2006), who finds that an ARIMA model can be used to forecast the directional change in the CBOE Volatility Index (VIX). Similarly, Bartels & Lu (2000) evaluated competing volatility forecasting methods for trading strategies on the DAX with four approaches: implied volatility, historical volatility, GARCH and E-GARCH. Bartels & Lu (2000) find that the GARCH and E-GARCH models have the best performances when evaluated in terms of a trading strategy. More recently, the formulation of a Markov Switching GARCH model has shown promising results vis-à-vis standard GARCH models, particularly in their ability to capture different volatility regimes. For example, Bauwens, Preminger, & Rombouts (2010) show that an MS-GARCH can fit S&P500 returns better than standard GARCH models, while la Torre-Torres, Oscar, Galeana-Figueroa, & Álvarez-García (2020) give similar results when analyzing oil prices.

Parallel to the literature on forecasting volatility for option pricing, work has emerged on the effect options have on the price of the underlying. In particular, they highlight the channel between market makers' hedging behavior due to their derivative exposure and the returns of the underlying. The goal of market makers who write option contracts to market participants (take the short position in puts and calls) is to profit from the bid-ask spread, while remaining directionally agnostic with respect to the price of the underlying. Market makers do not make any bets on price directions and the bid-ask spread can be thought of as a reward for providing liquidity, with usually larger spreads in less liquid tickers. In its most simple form, the market maker has to hedge their position by buying or selling the underlying asset depending on what type of contract they sell. For example, if a market maker writes a call option, they have a negative delta on their position. This is a result of the fact that a call option becomes worth more as the price of the underlying increases, which would move against the market maker. Hence, if the price of the underlying increases, the value of their position decreases. To be delta-neutral, i.e. to not be exposed to those price movements in the underlying, the market maker buys the underlying. A market maker would hedge the other way around if they sell a put option, resulting in increasingly shorting the stock if the price goes down and buying it – hence reducing their short position – if the price goes up. This is referred to as delta-gamma hedging and occurs on a dynamic basis as convexity cannot be hedged away. The gamma changes delta when the price of the underlying changes, with gamma being the lowest for far out-the-money and in-the-money options and highest for at-the-money (figure 2). Most market makers have specific delta mandates at the end of the day.

Figure 1: SPX Delta and Gamma Across Strike Prices



Notes: This figure shows the variation of gammas and deltas of SPX options according to different strike prices. The data covers call options on 02-01-2019 with an expiration date on 17-01-2020. In the figure the dashed red line marks the closing price of the S&P500 on 02-01-2019 to show which call options are in vs. out of the money.

The two measures we focus on which seek to quantify this hedging behavior are the NOPE and GEX. Lily (2020) shows that there is a monotonic decrease of the proportion of close-to-close returns greater than zero with increasing NOPE. Furthermore, the author shows that the predictive effect mainly results from the magnitude of NOPE itself instead of deviations of the NOPE relative to the 30-day median. While unitless, a positive NOPE would imply that the market maker is overweight on calls and vice versa. Zambito (2016) puts forward the GEX which measures the second order price sensitivity of an option or portfolio to changes in the price of an underlying security. It essentially quantifies how much market makers have to buy for hedging purposes with zero being the turning point. A positive GEX implies that market makers hedge in a non-directional way, i.e. stifling volatility, while a negative GEX implies the opposite. A GEX approaching zero indicates that markets can move naturally without market makers' re-hedging trades affecting the price of the underlying. Barbon & Buraschi (2020) show similar results referring to it as gamma imbalances. Both approaches exhibit predictive power in terms of the underlying asset price dynamics.

3 Data

Our empirical analysis rests on two main data sets: data on the S&P500 pulled from Yahoo Finance and options data provided by IvyDB OptionMetrics which cover European-style SPX options. The option data only provides the best bid and ask offers, from which we take the mid-point to be the "true price". In addition to the price, the data set provides information on the option type (put or call), the expiration date, the option Greeks, the volume and the open interest. As the SPX does not consider dividends, dividend payments do not enter our computations.

On top of these two data sets we retrieved the LIBOR as of May 21, 2021 as the risk-free rate input into the Black-Scholes option pricing formula. We scale it in proportion to the time to maturity of the option, in accordance with the first step of the interest rate calculation in the IvyDB US Reference Manual v3.0. While we acknowledge that this leads to only an approximation of the option price, it is sufficient for the purpose of this paper. The options data set contains

Figure 2: Daily Log Returns SPX

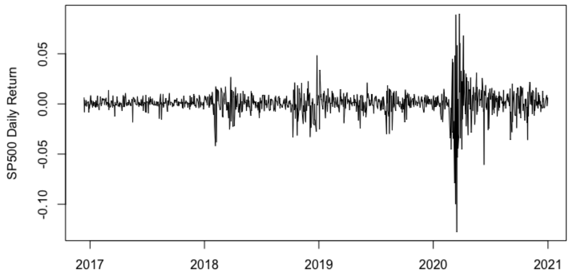
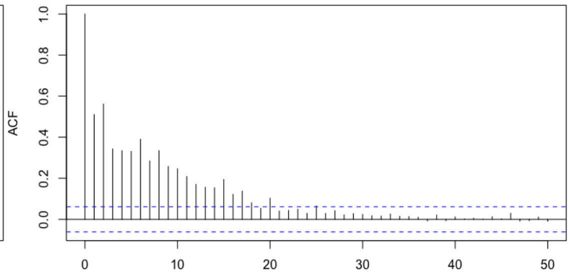


Figure 3: Correlogram Squared SPX Returns



Notes: Figure 2 shows the daily log return of the SPX. Figure 3 shows the autocorrelation of the squared SPX log return series with lags ranging from 0 to 50. The dashed blue lines indicate 95% confidence bounds. The data covers the period 13-12-2016 until 30-12-2020.

daily observations covering the period from the 2nd of January 2019 until the 31st of December 2020. Several option contracts showed missing values (NA) for the option Greeks. As these are critical components in our calculations for the NOPE and GEX we removed them from the data set. Data on the S&P500 covers a larger time frame, namely since December 13, 2016. We need more historical data on the S&P500 in order to initialize our rolling volatility forecast to the day on which our options data starts.

Table 1: Descriptive Statistics of the S&P500 Log Return Series

	N	NA	Min	Max	Mean	Variance	Stdev	Skew	Kurtosis
SPX	$1,019$	0	-0.128	0.090	0.0005	0.0002	0.013	-1.148	21.162

Notes: This table shows descriptive statistics for the daily S&P500 log return series. The data covers the period of 13-12-2016 until 30-12-2020.

For our volatility model specification we compute the daily log returns of the S&P500. An initial inspection of these confirms some of the stylized facts of financial time series. The descriptive statistics in table 1 show that the log returns have significantly fatter tails than a normal distribution as indicated by the excess kurtosis. We confirmed the deviation of the return series from a normal distribution by conducting a Jarque-Bera test, which led to the rejection of the null hypothesis. Figure 2 gives evidence of volatility clustering in the series, i.e. large swings in the returns are also followed by large swings in either direction. A qualitative assessment of the correlogram in figure 3 suggests that there is information flow to be captured in the squared log return series. Using the Ljung-Box Test, we find that the SPX does indeed exhibit serial correlation in both returns and squared returns. Furthermore, we confirmed with the Augmented Dickey-Fuller Test that the returns data is stationary. To test for structural breaks in the returns series, i.e. whether multiple volatility regimes exist, we use the Quandt Likelihood Ratio which extends the classical Chow Test because it does not require to pick a break point. We find evidence of structural breaks in the SPX returns, which seems unsurprising given that the COVID-19 recession is included in our sample period.

4 Methodology

To predict volatility, we contrast three different approaches. First, we use the standard GARCH (1,1) as a benchmark model given by:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{1}$$

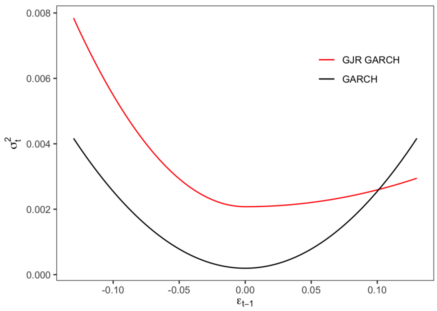
where σ_t^2 is the conditional variance given $\mathcal{F}_{t-1}$, $\varepsilon_{t-1}^2 = \sigma_{t-1}^2 z_{t-1}^2$ with z_{t-1}^2 being an iid centered noise process with $\mathcal{N}(0,1)$, and α_0, α_1, and β_1 being constants with specific restrictions. Even though it does not capture most of the stylized facts of returns, we still think it is valuable to have a widely used benchmark. Second, the GJR-GARCH with non-Gaussian error distribution is used to both incorporate the leverage effect, i.e., that volatility increases more after price declines, and the high kurtosis of the return series. The GJR-GARCH (1,1) is given by:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + \eta_1 \varepsilon_{t-1}^2 \mathrm{I}_{t-1}, \text{ where} \tag{2}$$

$$\mathrm{I}_t = \begin{cases} 1 & \text{if } \varepsilon_t < 0 \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

with the leverage effect incorporated via the dummy variable and z_{t-1}^2 following a Student's t distribution with $t(0,1)$. Presence of the leverage effect can be confirmed ex-post via a significant non-zero η_1. Figure 4 shows that a negative innovation has a larger effect on the volatility estimate for the GJR-GARCH (red line). The normal GARCH does not account for asymmetries and both negative and positive returns enter equally (black line).

Figure 4: News Impact Curve for the GARCH and GJR-GARCH



Notes: This figure shows the news impact curves for the GARCH and GJR-GARCH models on the S&P500 log return series.

Lastly, we factor in structural breaks in the volatility patterns via a Markov Switching GARCH, which allows us to incorporate different volatility regimes in our prediction. This is particularly interesting since our sample period includes the COVID-19 recession from early 2020 on. The Markov Switching GARCH is a generalization of the standard GARCH model that permits regime switching in the parameters and accommodates volatility clustering. This is achieved by allowing for different persistence in the conditional variance of each regime. The volatility depends on the state variable s_t, which evolves according to a first-order Markov chain with a transition matrix and is independent by definition of the returns. A Markov chain is a stochastic model in which the transition probability of each event depends only on the realization of the state in the previous period. We assume two states and it is assumed that both regimes are captured by a GJR-GARCH with non-Gaussian errors. The standard Markov Switching GARCH is given by:

$$\sigma_t^2 = \alpha_{0,s_t} + \alpha_{1,s_t}\varepsilon_{t-1}^2 + \beta_{1,s_t}\sigma_{t-1}^2 \tag{4}$$

We compute one-step ahead predictions with rolling windows, keeping the sample size constant. Hence, for each forecast the furthest observation is dropped, the most recent one is added and then the model is re-estimated. While Ahoniemi (2006) and Bartels & Lu (2000) use longer windows, we choose a window size of 505 observations for the sake of computational feasibility.

Using the Black-Scholes Pricing formula, we calculate the theoretical option prices of the SPX. The value of a European call option is given by:

$$C = e^{-\delta m} S\Phi(d_1) - e^{-ym} K\Phi(d_2), \quad \text{where} \tag{5}$$

$$d_1 = \frac{\ln(S/K) + (y - \delta + \sigma^2/2)m}{\sigma\sqrt{m}} \quad \text{and} \quad d_2 = d_1 - \sigma\sqrt{m} \tag{6}$$

where S is the SPX closing price, K is the option strike price, δ is the dividend, y is the risk-free rate, σ is the volatility of the underlying asset, and m the time to expiration of the option in years. $\Phi(d)$ is the value of the normal distribution with $\mathcal{N}(0,1)$ at d. In order to calculate the value of a European put option, we use the put-call parity formula:

$$P = e^{-ym} K\Phi(-d_2) - e^{-\delta m} S\Phi(-d_1) \tag{7}$$

which arises as a result of the law of one price and no-arbitrage assumption. Constructing a portfolio of a put and call option on the same underlying, strike, and expiration date is equivalent to a forward contract at the same strike and expiry. First, we calculate the theoretical option prices with the original data. To do so, we plug in the implied volatility and use the closing price of the SPX. Since the SPX does not pay dividends, we set $\delta = 0$ and y is the scaled LIBOR. Then we use the predicted scaled volatility of the three models to calculate new option prices. It is important to note here, that the predicted volatility is the value in $t+1$ but we need the volatility over the option's lifespan, i.e., the time to expiration. Even though Diebold et al. (1998) show that it overestimates the variability of long-horizon volatility, we multiply $\hat{\sigma}$ by $\sqrt{m}$ for sake of simplicity.

Next, we calculate the Greeks from the Black-Scholes model. The focus is here on Δ, which measures the responsiveness of the option's value w.r.t. to changes in the price on the underlying, and Γ, which measures the rate of change of Δ w.r.t. to changes in the underlying. Δ and Γ for a European call option are given by:

$$\Delta = \frac{\partial C}{\partial S} = e^{-\delta m}\Phi(d_1)\ , \quad \Gamma = \frac{\partial^2 C}{\partial S^2} = \frac{e^{-\delta m}\phi(d_1)}{S\sigma\sqrt{m}} \tag{8}$$

where all variables are defined as previously. $\phi(d)$ is the standard normal probability density function, i.e. the derivative of $\Phi(d)$. We calculate the Greeks on the predicted option prices with the volatility predictions from our GARCH models. It is important to note, that all Δ and Γ are multiplied by 100. This is done because one contract consists of 100 options.

Δ is used as the main component of the Net Options Pricing Effect. This is calculated by summing up the Δ weighted by the volume of every call c on all expirations q and option chains ω. This gives us the total call Δ and total put Δ:

$$\varphi_t = \sum_{\omega\in\Omega}\sum_{\mathrm{q}\in\mathrm{Q}}\sum_{\mathrm{c}\in\mathrm{C}} \mathrm{V}_{c_{\omega,q}} \cdot \Delta_{c_{\omega,q}} \ , \quad \psi_t = \sum_{\omega\in\Omega}\sum_{\mathrm{q}\in\mathrm{Q}}\sum_{\mathrm{p}\in\mathrm{P}} \mathrm{V}_{p_{\omega,q}} \cdot \Delta_{p_{\omega,q}} \tag{9}$$

The Net Options Pricing Effect is then given by the difference divided by the total shares traded in the underlying V_t. Note here, the total put Δ is negative and therefore added:

$$\mathrm{NOPE} = \frac{\varphi_t + \psi_t}{\mathrm{V}_t} \tag{10}$$

A positive NOPE would indicate an overweight of call options in the market and vice versa. Alternatively, we model the gamma exposure as follows:

$$\vartheta_t = \sum_{\omega\in\Omega}\sum_{\mathrm{q}\in\mathrm{Q}}\sum_{\mathrm{c}\in\mathrm{C}} \Gamma_{\mathrm{C}_{\omega,q}} \cdot \mathrm{OI}_{\mathrm{C}_{\omega,q}} \ , \quad \zeta_t = \sum_{\omega\in\Omega}\sum_{\mathrm{q}\in\mathrm{Q}}\sum_{\mathrm{p}\in\mathrm{P}} \Gamma_{\mathrm{P}_{\omega,q}} \cdot \mathrm{OI}_{\mathrm{P}_{\omega,q}} \tag{11}$$

where OI is the open interest, i.e. the total number of contracts at a given strike and exercise day which have not been liquidated yet. The summation is similar to NOPE. Open interest is a measure of liquidity and usually not updated during the trading day. The GEX is calculated as the sum (difference) between the GEX of put and call options:

$$\mathrm{GEX} = \vartheta_t + \zeta_t \tag{12}$$

which is denominated in USD for the SPX.

5 Results

According to the AIC and BIC, we find that the Markov Switching GARCH performs the best. Table 2 shows the average AIC and BIC for the different models used in this paper, with the lowest AIC and BIC indicating the best model. It is important to note, that there is no formal decision criterion if both AIC and BIC diverge but Medel & Salgado (2013) show that the BIC is more robust in forecasting. However, we also argue that the MS-GARCH is preferred from a qualitative angle since it incorporates structural breaks in addition to the stylized facts captured by the GJR-GARCH. Therefore, we use the MS-GARCH as our main approach from here onwards. Unsurprisingly, the standard GARCH underperforms.

Table 2: AIC and BIC

	GARCH	GJR-GARCH	MS-GARCH
AIC	$-3,386.792$	$-3,459.195$	$-3,510.513$
BIC	$-3,365.893$	$-3,438.297$	$-3,459.818$

Notes: This table shows the average Akaike Information Criterion, and Bayesian Information Criterion for the different models in our study. The lowest values indicate the preferred model.

To evaluate how well our volatility models estimate the second moment, we run Mincer Zarnowitz regressions. While the second moment is not directly observable, we use the squared daily SPX returns as a proxy. Alternatively, and arguably more accurate, one could have used scaled high-frequency returns via a HAR model. However, due to data restrictions we decided to go with the former. According to Mincer & Zarnowitz (1969), we regress the estimated volatility on the squared SPX returns proxy where the constant would equal zero and the coefficient one under the Null Hypothesis. While we do not expect that the GARCH-type models reject the Alternative Hypothesis, we still consider this to be an important step for the sake of completeness. Table 3 shows the results for the GARCH, GJR-GARCH, and MS-GARCH.

Table 3: Mincer Zarnowitz Regressions of Volatility Forecasts

	Dependent variable:		
	r_t^2 Proxy		
	(1)	(2)	(3)
MS-GARCH	0.059***		
	(0.003)		
GJR-GARCH		0.052***	
		(0.003)	
GARCH			0.057***
			(0.004)
Constant	-0.0004***	-0.0004***	-0.0004***
	(0.0001)	(0.0001)	(0.0001)
Observations	505	505	505
R^2	0.384	0.334	0.304
Adjusted R^2	0.383	0.333	0.302
Residual Std. Error	0.001	0.001	0.001
F Statistic	313.526***	252.260***	219.396***

Notes: This table shows the output of the Mincer Zarnowitz regressions of the respective volatility forecasts. Statistical significance is given by *p<0.1; **p<0.05; ***p<0.01. Under H_0, $\alpha = 0$ and $\beta = 1$.

Using the volatility models to estimate the Greeks, we can replicate the NOPE well. Figure 5 and figure 6 plot the NOPE for implied volatility and MS-GARCH volatility vis-à-vis the 1-day SPX close-to-close returns. We can clearly see that there is higher volatility in the returns in a negative NOPE environment and lower volatility in positive NOPE territory, with a distinct switching point at zero. Similarly, we find promising results for the GEX measure using the estimated volatility. Figures 7 and 8 show the GEX calculated with implied and estimated volatility, respectively. Interestingly, the predicted volatility does show the volatility stifling behavior we expected in a positive gamma environment and vice versa.

Figure 5: NOPE with Implied Volatility

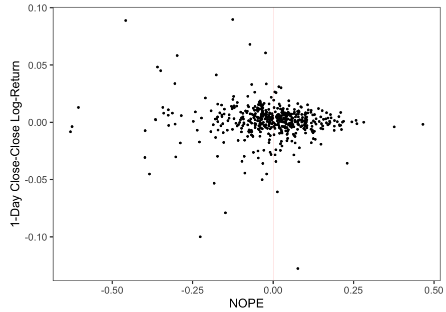
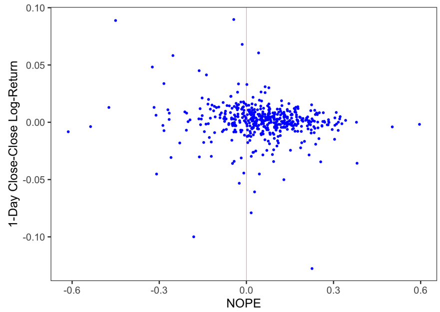


Figure 6: NOPE with MS-GARCH Volatility



Notes: Figure 5 shows the Net Options Pricing Effect calculated from deltas with the implied volatility from OptionMetrics versus the 1-day close-to-close return of the SPX. Figure 6 shows the Net Options Pricing Effect calculated from deltas with the estimated volatility from the Markov-Switching GARCH versus the 1-day close-to-close return of the SPX. The vertical line indicates the NOPE zero point, i.e., where we switch from overweight puts to overweight calls. The option data covers the period from 2nd January 2019 until the 31st December 2020.

Figure 7: GEX with Implied Volatility

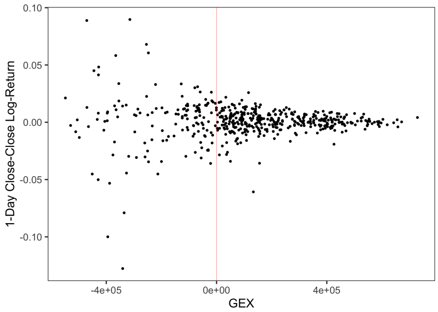
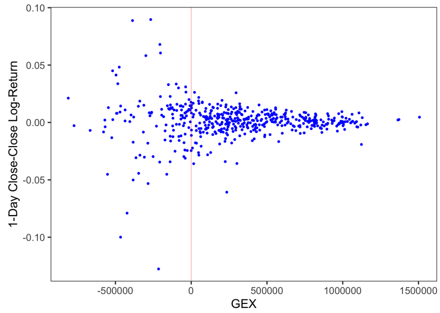


Figure 8: GEX with MS-GARCH Volatility



Notes: Figure 7 shows the Gamma Exposure from gammas calculated with the implied volatility from OptionMetrics. Figure 8 shows the Gamma Exposure from gammas calculated using the forecasted volatility with an MS-GARCH specification. The data covers the period from 2nd January 2019 until the 31st December 2020.

By comparing the Pearson correlations between NOPE and GEX with the SPX Close-to-Close return, we can further quantify the relationships shown in figures 5 through 8 and demonstrate which forecasting model best captures this relation. Table 4 depicts correlations according to the method that has been used to estimate NOPE and GEX. Surprisingly, the model that returns the highest absolute value of correlation between NOPE and SPX returns is the standard GARCH. In terms of GEX, the model that returns the highest absolute correlation with the SPX returns is the GJR-GARCH. The MS-GARCH performs relatively well in terms of NOPE Pearson correlation, but does not give remarkable results in terms of GEX Pearson correlation. The signs for the OptionMetrics GEX and the GEX calculated with the implied volatility are unexpected. One possible reason for the unexpected signs may be due to the fact that the scale of GEX is in dollars. However, we acknowledge that this finding requires further research. Even though some of the results are unexpected, we do find initial empirical evidence that our models with forecasted volatility have a stronger relationship with SPX returns than models using implied volatility. Nonetheless, we also acknowledge the various limitations of simple correlations.

Table 4: Pearson Correlation for NOPE and GEX, with SPX Close-to-Close Return

	NOPE	GEX
OptionMetrics	−0.086	0.003
GARCH	−0.146	−0.006
GJR-GARCH	−0.145	−0.008
MS-GARCH	−0.138	−0.001
Implied Volatility	−0.086	0.001

Notes: This table shows the Pearson correlation between NOPE/GEX, and the 1-day SPX close-to-close return. OptionMetrics refers to the NOPE and GEX calculated from the given Greeks. The data covers the period from 2nd January 2019 until the 31st December 2020.

6 Discussion

The initial evidence seems promising in that it shows that we can use predicted volatility to accurately quantify the delta-gamma hedging behavior of market makers. In addition, it shows a slight improvement relative to the models calculated with the Greeks provided by OptionMetrics.

Interestingly, we find that the MS-GARCH has difficulties estimating deltas and gammas for options with short times to expiration. Figures 9 and 10 show the call deltas for different strike prices at different times to expiration for the implied and MS-GARCH volatility, respectively. Deltas in general behave differently at shorter times to expiration, with a much steeper delta curve. Hence, we conjecture that the MS-GARCH seems to have difficulties to estimate short-term deltas which may be the result of a large noise component close to expiration. We argue that volatility seems to be less important for option prices with only a few days to expiration due to a rise of speculation given the tilted risk-reward ratio. In line with this, figures 11 and 12 plot the call gammas. Unsurprisingly, given the steep delta in the MS-GARCH setting, the gamma in the MS-GARCH is narrowly centered around high values. For both deltas and gammas, the MS-GARCH performs better with an increasing time to expiration which supports our speculation narrative. Therefore, we argue that we can better quantify the hedging behavior and predict directionality of the underlying when there is an overweight of options with longer time to expiration.

Figure 9: Delta with Implied Volatility

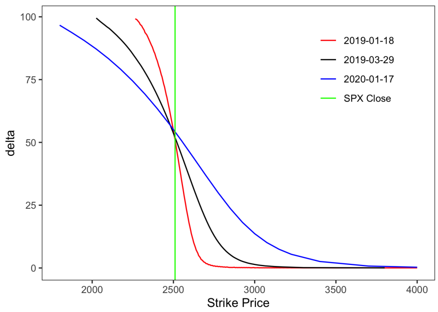
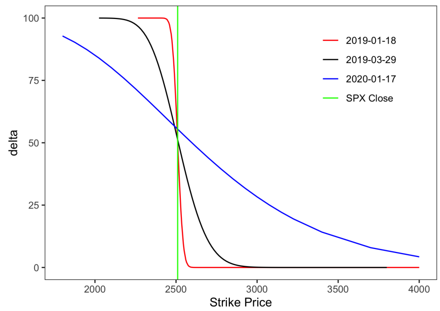


Figure 10: Delta with MS-GARCH Volatility



Notes: Figure 9 shows the SPX call deltas calculated with the implied volatility from OptionMetrics for various strike prices. Figure 10 shows the SPX call deltas calculated with the estimated volatility from the Markov-Switching GARCH for various strike prices. The option data covers the period from 2nd January 2019 until the 31st December 2020. The vertical green line shows the closing price of the SPX on the 2nd January 2019 to indicate which call options are in vs. out of the money.

Figure 11: Gamma with Implied Volatility

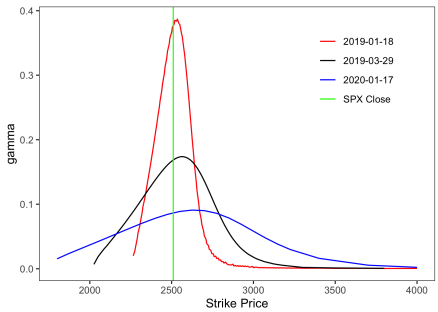
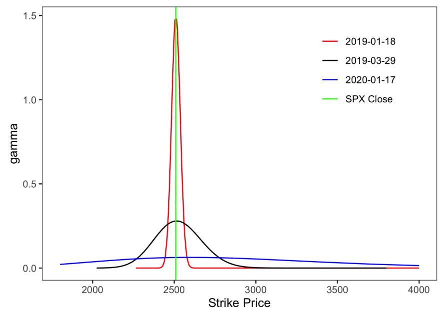


Figure 12: Gamma with MS-GARCH Volatility



Notes: Figure 11 shows the SPX call gammas calculated with the implied volatility from OptionMetrics for various strike prices. Figure 12 shows the SPX call gammas calculated with the estimated volatility from the Markov-Switching GARCH for various strike prices. The option data covers the period from 2nd January 2019 until the 31st December 2020. The vertical green line shows the closing price of the SPX on the 2nd January 2019 to indicate which call options are in vs. out of the money.

Figures 13 and 14 plot the proportion of positive close-to-close returns for the next day against today's NOPE, calculated with the implied volatility and the MS-GARCH volatility, respectively. For both we observe an overall decline in the share of next day positive returns as today's NOPE increases. However, both graphs are not as monotonically decreasing as shown by Lily (2020). We hypothesize that this might be a result of the different sample sizes as our option data only covers

Figure 13: NOPE vs. Close-Close Return Implied Volatility

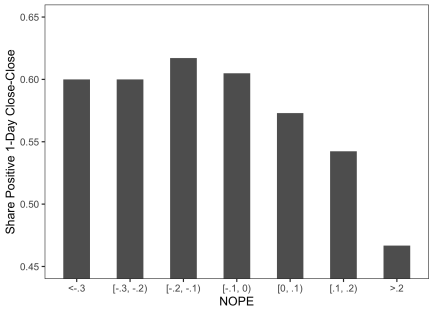
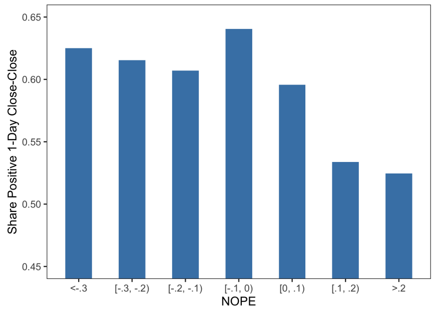


Figure 14: NOPE vs. Close-Close Return MS-GARCH Volatility



Notes: Figure 13 shows the NOPE calculated with the implied volatility from OptionMetrics versus the share of positive 1-day SPX close-to-close returns. Figure 14 shows the NOPE calculated with the estimated MS-GARCH volatility versus the share of positive 1-day SPX close-to-close returns. The bin size is 0.1.

two full trading years with 505 observations as opposed to almost 14 years in Lily's paper. As a result, the two bins for the negative NOPE on the left in figures 13 and 14 only consist of 30 or less observations combined. This limits the validity of the first two bars in the both graphs. The sample size is less an issue for the more positive NOPE where we can confirm the clear monotonic decline in positive next day returns found by Lily. It is important to note, that in contrast to Lily's results, our share of positive returns for positive NOPE are close to 50% over the plotted range for both volatility approaches. This limits the feasibility of potential trading strategies based on our results. However, we are confident that with both refined volatility models, i.e. take into account stochastic and/or implied volatility models, and richer data that we are able to build a robust trading strategy based on NOPE.

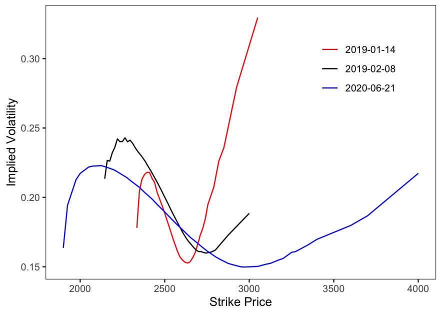
Table 5: Regression Results NOPE and GEX

	Dependent variable:	
	SPX 1-day Close-Close Return	
	(1)	(2)
NOPE Implied Volatility	-0.021^{***}	
	(0.008)	
GEX Implied Volatility	0.000^{*}	
	(0.000)	
NOPE MS-GARCH		-0.027^{***}
		(0.007)
GEX MS-GARCH		0.000^{**}
		(0.000)
Constant	-0.0002	0.001
	(0.001)	(0.001)
Observations	505	505
R^2	0.014	0.030
Adjusted R^2	0.010	0.027
Residual Std. Error	0.016	0.016
F Statistic	3.570^{**}	7.879^{***}

Notes: This table shows the regression results with the SPX close-to-close return as dependent variable and NOPE and GEX as covariates. Statistical significance is given by *p<0.1; **p<0.05; ***p<0.01. Under H_0, $\alpha = 0$ and $\beta = 1$.

While in figures 13 and 14 the comparison of implied volatility vs. MS-GARCH volatility is not as clear, table 5 gives a distinct indication. We estimated a multivariate regression of the 1-day close-to-close returns on the NOPE and the GEX for the two volatility models separately. When examining the coefficients for the NOPE, we find significance for both volatility models. For the MS-GARCH volatility, however, the coefficient is larger in absolute terms with a smaller standard error. This suggests that our MS-GARCH volatility model does a better job in predicting next day returns. With respect to the GEX coefficients, we observe close to zero coefficients in both regressions which is most likely as a result of universally lower gamma values relative to delta. Again, this simple regression framework is limited and only referred to as first evidence. Another drawback of our approach is that we estimate volatility purely on a daily basis. Therefore, we do not account for variation in volatility across strike prices and expiration dates. The standard Black-Scholes model also assumes that the underlying follows a Geometric Brownian Motion with constant volatility. However, as figure 15 shows, the volatility varies significantly according to both the strike price and the expiration date of the option, leading to volatility "smiles" and "smirks". Hence, adjusting our volatility estimates to account for heterogeneity in time to maturity and strike prices and using more advanced option pricing models may yield more conclusive estimates.

Figure 15: Volatility Smile



Notes: This figure shows the variation of implied volatility across different strike prices and expiration dates on SPX call options. The data covers the period from 2nd January 2019 until the 31st December 2020.

7 Conclusion

In this paper we analyzed if we can improve on models which seek to capitalize on the effect delta-gamma hedging activities by market makers have on the price of the underlying asset. Predicting the hedging trades efficiently would allow the anticipation of a stock's demand and with it price movements. To do so, we used GARCH-type volatility models as inputs to calculate more accurate option prices, and more importantly, more accurate Greeks. Thus, this would allow us to better predict next day's price movements in the underlying and capitalize on this accordingly.

Given we are the first to do so, we provide exciting new evidence on the intersection of the effect of delta-gamma hedging on the price of the underlying, and the prediction of volatility. Using SPX options data from OptionMetrics and index-level data from Yahoo Finance, we show that GARCH-type models can to some extent improve models which predict next days' price movement based on market maker hedging behavior. First, we find that GARCH-type volatility models have difficulties estimating deltas and gammas of options with relatively short time to expiration. We hypothesize that this is a result of a significantly higher level of speculation given the tilted risk-reward rations of deep out-the-money options. In other words, it is the result of added noise from short-term speculators in the market. With an increasing time to expiration, however, the predicted volatility yields Greeks which are closer to realized ones. Furthermore, we confirm the findings of other studies, that in a situation of an overweight of put options – negative NOPE and GEX – the variance of the next day close-to-close returns increases significantly. This is also in line with the leverage effect property of return series which states that negative returns increase volatility with a larger magnitude than positive returns.

Even though our results suggest a lower probability of positive next day returns with an increasing NOPE, the magnitude of the decline does not seem large enough to open the possibility of capitalizing through an applied trading strategy with the limited data. Nevertheless, we generally observe a decrease in the proportion of 1-day close-to-close returns above zero with an increasing NOPE, which supports the findings from the literature that were based on a richer data set. We did find, however, that the NOPE derived with the best performing GARCH-type volatility model, the MS-GARCH, has a higher correlation with next day returns than the NOPE from implied volatility. Moreover, the regression results suggest that the magnitude of the effect is larger for the former. We also note that the merely relying on correlation and the multivariate regression may not be statistically rigorous, but argue that this yields interesting results regardless.

We do acknowledge that our results could be extended by increasing the sample size of our option data, which might show a clearer relationship between the NOPE and the share of positive 1-day SPX close-to-close returns. Furthermore, employing more refined volatility models which can account for variation across strike prices and expiration dates could give a more precise estimate of the NOPE and GEX. This would also allow us to make a more conclusive statement as to whether forecasted volatilities can beat implied volatilities when calculating option Greeks. Moreover, future research can focus on improving the delta-gamma hedging models by means of including open interest in NOPE and/or focus on Vega hedging strategies and third-order Greeks such as speed.

References

Ahoniemi, K. (2006). Modeling and forecasting implied volatility-an econometric analysis of the vix index.

Barbon, A., & Buraschi, A. (2020). Gamma fragility. *Available at SSRN 3725454*.

Bartels, H.-J., & Lu, J. (2000). Volatility forecasting and delta-neutral volatility trading for dtb options on the dax. *Proceedings AFIR*, 51–66.

Bauwens, L., Preminger, A., & Rombouts, J. V. (2010). Theory and inference for a markov switching garch model. *The Econometrics Journal*, *13*(2), 218–244.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, *81*(3), 637–654.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, *31*(3), 307–327.

Canina, L., & Figlewski, S. (1993). The informational content of implied volatility. *The Review of Financial Studies*, *6*(3), 659–681.

Christensen, B. J., & Prabhala, N. R. (1998). The relation between implied and realized volatility. *Journal of financial economics*, *50*(2), 125–150.

Diebold, F., Hickman, A., Inoue, A., & Schuermann, T. (1998). Scale models. *Risk*, 104–107.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the econometric society*, 987–1007.

Jorion, P. (1995). Predicting volatility in the foreign exchange market. *The Journal of Finance*, *50*(2), 507–528.

la Torre-Torres, D., Oscar, V., Galeana-Figueroa, E., & Álvarez-García, J. (2020). A test of using markov-switching garch models in oil and natural gas trading. *Energies*, *13*(1), 129.

Lily. (2020). Investigating in the impact of delta-gamma hedging in sp 500 returns from 2007-2020.

Medel, C. A., & Salgado, S. C. (2013). Does the bic estimate and forecast better than the aic? *Revista de Análisis Económico–Economic Analysis Review*, *28*(1), 47–64.

Mincer, J. A., & Zarnowitz, V. (1969). The evaluation of economic forecasts. In *Economic forecasts and expectations: Analysis of forecasting behavior and performance* (pp. 3–46). NBER.

Poon, S.-H., & Granger, C. W. (2003). Forecasting volatility in financial markets: A review. *Journal of economic literature*, *41*(2), 478–539.

Szakmary, A., Ors, E., Kim, J. K., & Davidson III, W. N. (2003). The predictive power of implied volatility: Evidence from 35 futures markets. *Journal of Banking & Finance*, *27*(11), 2151–2175.

Zambito, M. (2016). Gamma exposure (GEX), quantifying hedge rebalancing in SPX options. *squeezemetrics.com*.