Turnitin Originality Report

Processed on: 25-Apr-2025 12:07 IST
ID: 2656401903
Word Count: 3684
Submitted: 1

MTP Report 20AG3FP30 By Gaurav Patidar

Similarity Index	Similarity by Source	
5%	Internet Sources:	5%
	Publications:	1%
	Student Papers:	3%

1% match (Internet from 13-May-2020)
https://www.coursehero.com/file/46959142/mainpdf/

1% match (Internet from 01-Mar-2022)
https://www.coursehero.com/file/96346170/1501CE24pdf/

1% match (Internet from 06-Feb-2022)
http://docplayer.net/42053049-Comparison-of-option-price-from-black-scholes-model-to-actual-values.html

1% match (student papers from 18-Feb-2025)
Submitted to DY Patil International School on 2025-02-18

1% match (Internet from 15-Nov-2023)
https://theses.ubn.ru.nl/server/api/core/bitstreams/2d084cf6-d899-484a-908c-0a32a4c3723c/content

< 1% match (Internet from 08-Jul-2024)
https://ideas.repec.org/a/gam/jjrfmx/v14y2021i9p402-d621475.html

< 1% match (student papers from 20-Apr-2025)
Submitted to University of Cincinnati on 2025-04-20

< 1% match (Internet from 16-Dec-2022)
https://gupea.ub.gu.se/bitstream/handle/2077/37435/gupea_2077_37435_1.pdf?sequence=1

< 1% match (Internet from 17-May-2016)
http://www.inderscience.com/storage/f793211810612514.pdf

< 1% match (Internet from 12-Mar-2014)
http://ro.uow.edu.au/cgi/viewcontent.cgi?article=2936&context=theses

< 1% match (Internet from 01-Mar-2021)
http://www1.iitkgp.ac.in/downloads/sm_1516_mdp.pdf

A Data-Driven Approach to Developing Novel Multi-Leg Trading Strategies Presented To the Faculty of Vinod Gupta School Of Management, IIT KHARAGPUR MTech thesis project Report Supervised By: Prof. Arun Kumar Misra Written By: Gaurav Patidar 20AG3FP30 Evaluators Name Grade Table 1: Grades 1 VINOD GUPTA SCHOOL OF MANAGEMENT INDIAN INSTITUTE OF TECHNOLOGY KHARAGPUR – 721 302 CERTIFICATE This is to certify that the project report entitled “A Data-Driven Ap- proach to Developing

Novel Multi-Leg Trading Strategies" submit- ted by Gaurav Patidar (Roll No . 20AG3FP30) to Indian Institute of Technology Kharagpur towards partial fulfilment of requirements for the award of degree in Agricultural and Food Engineering and Financial Engineering is a record of bona fide work carried out by him under my supervision and guidance during Spring Semester, 2024-25. Date: 25 April 2025 Professor Arun Kumar Misra Vinod Gupta School of Management Indian Institute of Technology Kharagpur Kharagpur-721302, India 2 DECLARATION I, Gaurav Patidar, Roll No. 20AG3FP30, a student of Indian Institute of Technology Kharagpur, hereby declare that the project report entitled "A Data-Driven Approach to Developing Novel Multi-Leg Trading Strategies" submitted to the faculty of Vinod Gupta School of Man- agement, IIT Kharagpur, is my original work. I further declare that this project report: • Has not been submitted elsewhere in full or in part for the award of any other degree, diploma, or certification. • Is free from plagiarism and follows all academic integrity guidelines set forth by the institute. • Has been prepared based on the data and information obtained through legitimate means and research, with appropriate acknowledgments given to any references or resources used. Signature: Date: Place: Kharagpur Name: Gaurav Patidar Roll No: 20AG3FP30 3 Abstract The general option spreads try to hedge delta or harvest theta but rarely ad- dress the behavior of all Greeks, as real-world markets diverge from Black–Scholes assumptions—implied volatilities cluster, liquidity varies by strike, and deep I/OTM quotes often carry placeholder vols—systematic mispricings arise in Vega and higher-order sensitivities. In this study we approach with statis- tical analysis on past options data and try to find their relation with the Greeks. This project explores which Greek plays important role via statistical anal- ysis and then we try to create a trading strategy that gives us neutrality corresponding to that Greek, leveraging the insights we perform simulations and come up with the strategy and then back-test by applying this strat- egy on past data to check whether our portfolio that is made after applying strategy gives better Greek neutrality than naive approach or not. The re- sult talks about factor-neutral option strategies, demonstrating that careful Greek diagnostics can unlock risk-controlled portfolio. 4 Contents Abstract 4 1 Introduction 8 1.1 Problem Statement . . . . 8 1.2 Scope . . . . 8 1.3 Outcomes/Objectives . . . . 9 1.4 Methodology . . . . 9 1.4.1 Data Collection and Greek Computation . . . . 9 1.4.2 Statistical Analysis . . . . 10 1.4.3 Strategy Design . . . . 10 1.4.4 Back-Testing and Visual Assessment . . . . 10 2 Background and Review of Literature 11 3 Data and Methods 13 3.1 Data Description . . . . 13 3.1.1 Engineered Option Dataset: extracted option data.csv 13 3.1.2 Spot Price History: spy spot history.csv . . . . . . 14 3.2 Statistical Analysis . . . . 14 3.1.3 Full Option Chain: spy option chain.csv . . . . 14 3.2.1 Regression . . . . 14 3.3 Simulation . . . . 17 3.4 Generalisation of Strategies . . . . 19 4 Back-Testing of Strategy and Results 21 4.1 Framework . . . . 21 5 Managerial Implications and Future Work 24 5.1 Managerial Implications . . . . 24 5.2 Future Work . . . . 24 5 List of Tables 1 Grades............................ 1 6 List of Figures 2.1 Comparison of Option Price from Black-Scholes Model to Ac- tual Values . . . . 12 3.1 OLS regression summary for extracted SPY option data . . . 15 3.2 Correlation matrix . . . . 16 3.3 Net premium vs. volatility for Strategy 3.1 . . . . 17 3.4 Net premium vs. volatility for Strategy 3.2 . . . . 18 3.5 Net premium vs. volatility for Strategy 3.3 . . . . 19 4.1 Per-contract Vega for each leg . . . . 23 4.2 Position-sized Vega contribution of Strategy 3.3 legs . . . . . . 23 4.3 Net Vega: Strategy

3.3 vs. naïve long-call benchmark . . . . . 23 7 Chapter 1 Intro duction 1.1 Problem Statement The listed option market is highly dynamic, prices react simultaneously to underlying moves, term-structure shifts, order-book liquidity, and macro- volatility events. Many Investors properly monitor the the five primary Greeks, but many academic and reatil literature stills see them in isolation during making spreads and strategies mostly focuses on neutrlising delta without seeing interaction about other greeks and how they impact or how real-world deviations from Black–Scholes distort those relationships. This gap leaves traders with a hidden risk. This project aims to address that gap by applying modern statistical tech- niques correlation analysis, OLS regressions. It tries to quantify which Greeks most strongly explain price variance and then try to come up with a strategy or portfolio which minimize that specific Greek. 1.2 Scope The scope of this research centres on building a data-driven framework for designing and validating Vega-neutral, multi-leg options strategies. Specifi- cally, the project focuses on: • Applying correlation analysis, ordinary-least-squares regression, to iden- tify which Greeks dominate option-price variance under real-market conditions. • The system provides traders with a fully specified, immediately imple- mentable template that methodically hedges portfolios against unfore- 8 seen spikes or collapses in implied volatility, while still enabling them to hold on to, express, and gain from precisely calibrated directional opinions in the underlying market. The results are meant to show that there is a reproducible process for trans- forming statistical Greek diagnostics into operational, factor-insensitive op- tion portfolios, and that they offer a template for future studies on data- driven risk-controlled strategies. 1.3 Outcomes/Ob jectives The primary objectives of this research are as follows: 1. Uncover Key Influences in Option Pricing: Use regression and correlation analysis to determine which Greeks—Delta, Gamma, Theta, Vega, and Rho—have the largest impact on observed option premiums by strike and maturity. 2. Develop a Novel Multi-Leg Trading Strategy: Develop an inno- vative options portfolio that goes beyond vanilla spreads like straddles or butterflies specifically intended to take advantage of the prevalent Greek exposures revealed in Objective 1 and take advantage of the resultant mispricings. 3. Validate Strategy Performance: Back-test the proposed multi-leg strategy on historical SPY option data (2023–2025) and benchmark its risk profile and profitability against a naïve long-call baseline, demon- strating improved Greek neutrality—especially in Vega—and superior P&L stability. 1.4 Methodology The research proceeded through four sequential stages, each informed by the insights of the previous step. 1.4.1 Data Collection and Greek Computation A complete SPY option chain was downloaded via yfinance, yielding the fields Date, Option Type, Spot Price, Strike Price, Option Price, Implied Volatility, and Time-to-Maturity. Black–Scholes closed-form expressions were applied to generate the derived Greeks Delta, Theta, Vega, and Rho for every contract. 9 1.4.2 Statistical Analysis Ordinary-least-squares regressions and a full correlation matrix were run on the engineered dataset. Both techniques showed that price movements are explained most strongly by Vega, motivating a focus on volatility exposure. 1.4.3 Strategy Design Guided by the statistical findings, we iterated through candidate multi-leg combinations in a targeted simulation environment until a portfolio structure that drives net Vega toward zero was obtained. The resulting configuration, referred to as the novel multi-leg strategy, was retained for empirical testing. 1.4.4 Back-Testing and Visual Assessment Historical spot and option data (spy spot history.csv and spy option chain.csv) for July 2024–April 2025 were used to back-test the new strategy. For each weekly trade date the strategy was priced, Greeks were tallied, and profit- and-loss was recorded. The same procedure was repeated for a naïve five- contract long-call benchmark. Comparative plots of individual-leg Vega, to- tal portfolio Vega, and cumulative P&L visually demonstrate that the pro- posed strategy achieves effective Vega-neutrality while the naïve approach remains highly volatility-sensitive. 10

Chapter 2 Background and Review of Literature The option–pricing literature is full of analytic models and empirical tests, but even much of this work still focuses on single-Greek hedges or makes Black–Scholes realism. Paper Analysis of Option Butterfly Portfolio Mod- els by Xiangyu Ge extends the traditional butterfly portfolio by suggesting a single–index non-parametric state-price density (SPD) estimator. Apply- ing local polynomial regression, the author demonstrates that the proposed estimator produces lower mean-squared error compared to kernel–based ap- proaches in calculating butterfly returns at minimum strike intervals, thus providing practitioners with a more accurate method to calibrate multi-leg positions. While our research does not directly estimate SPDs, the paper inspires the notion that well-selected leg combinations can take advantage of pricing inefficiencies that traditional spreads miss. An independent empirical strand examines to what extent Black–Scholes prices match traded premiums. Paper Comparison of Option Price from Black-Scholes Model to Actual Values by Matthew J. Krznaric compares theoretical call prices—under different volatilities and risk-free rates—with real S P 500 option quotes. The research finds systematic under-valuation ex- cept when volatility or interest-rate inputs are unrealistically exaggerated, supporting the perception that real markets incorporate risk premia and volatility dynamics the Black–Scholes model excludes. This result supports our choice to use historical regressions and correlation analysis, instead of closed-form calibration, when determining which Greeks actually move op- tion prices. 11 More generally, core texts like Hull and Sankarshan's Options and derivatives present the derivative-pricing equations and Greek definitions used through- out this work. Hull's discussion of Delta—, Vega– and Theta-neutral spreads provides the theoretical reference point by which we assess our strategy, and Sankarshan's examples of worked problems explain Greek interactions under discrete hedging. Figure 2.1: Comparison of Option Price from Black-Scholes Model to Actual Values 12 Chapter 3 Data and Metho ds 3.1 Data Description Three distinct CSV datasets underpin the empirical work: an engineered option–Greek panel for exploratory analysis, and two raw feeds for out-of- sample back-testing. 3.1.1 Engineered Option Dataset: extracted option data.csv This file contains the initial SPY option snapshot enriched with calculated Greeks. Key columns are: • Date – calendar date of the quote. • Option Type – call or put. • Underlying Price (Spot) – SPY close on the quote date. • Strike Price – contractual strike. • Option Price (Premium) – mid-quote used as transaction cost. • Implied Volatility (IV) – market-implied annualised volatility. • Time to Maturity (Days) – days until expiration. • Delta, Gamma, Theta, Vega, Rho – first-order Greeks derived via Black–Scholes; these variables form the basis of the regression and correlation analysis. 13 3.1.2 Spot Price History: spy spot history.csv Downloaded through yfinance, this file provides the reference price path for back-testing. • Date – trading day. • Open, High, Low, Close – daily OHLC quotes; the Close column is used as the spot price S0 in Greek and payoff calculations. • Volume – shares traded; used only to sanity-check data quality. 3.1.3 Full Option Chain: spy option chain.csv A comprehensive dump of every listed SPY option on the extraction date. • contractSymbol – unique OCC code (identifies strike, expiry, type). • lastTradeDate – timestamp of most recent trade. • strike – strike price (numeric). • bid, ask – best bid and ask; their midpoint is the modelled entry price. • volume, openInterest – liquidity indicators. • inTheMoney – Boolean flag used to segment ATM, ITM, and OTM contracts. • optionType – call or put. • expiration – option expiry (day–month–year format). The engineered dataset supports the initial Greek influence study, while the spot–history and full-chain files feed the back-test that compares the naïve long-call benchmark with the Vega-neutral multi-leg strategy. 3.2 Statistical Analysis 3.2.1 Regression Ordinary-least-squares (OLS) was run on extracted option data.csv, us- ing the option price as the dependent variable and the computed Greeks (Delta, Gamma, Theta, Vega, Rho) plus

Time-to-Maturity as predictors. Figure 3.1 displays the regression output table. 14 Figure 3.1: OLS regression summary for extracted SPY option data Key findings. • Model fit. The R2 value is below 0.30, indicating that a simple lin- ear specification explains less than one-third of the variation in option premiums. • Joint significance. A very large F -statistic (4.40 × 103) with a near- zero p-value shows that, taken together, the predictors do contribute statistically to price formation. • Dominant factor. Coefficient magnitudes and subsequent Greek- ranking analysis point to Vega as the single most influential explana- tory variable. Designing any hedging portfolio should therefore priori- tise neutralising Vega exposure. Correlation Matrix A heat-map of pairwise correlations among the Greeks, option price, and Time-to-Maturity is provided in Figure 3.2. 15 Figure 3.2: Correlation matrix Conclusion from correlation analysis. • Time-to-Maturity is strongly correlated with Theta, Vega, and Rho, confirming that longer-dated contracts are more sensitive to volatility and interest-rate changes. • Option Price shows its highest correlation with Delta, followed by Vega, reinforcing the regression result that these two Greeks dominate price behaviour. • Because a small subset of Greeks captures most of the variation, fo- cusing on Delta and especially Vega allows the hedging strategy to be simplified without sacrificing effectiveness. 16 3.3 Simulation To visualise how each successive design improves Vega-neutrality, we simulate the net portfolio premium as volatility (σ) varies from 15 % to 30 %. Figures for Strategies 3.1, 3.2, and 3.3 are inserted below; each uses Black–Scholes prices and the leg quantities specified earlier. Strategy 3.1: 10 Long Calls / 15 Short Puts • Each ATM call contributes roughly +3 Vega; each OTM put about −2 Vega. Ten calls and fifteen puts offset to a net Vega of 0. • Net premium is flat across the volatility range, showing that gains in the long-call leg are cancelled by losses in the short-put leg. • Conclusion: Vega hardly fluctuates, confirming effective neutrality. Figure 3.3: Net premium vs. volatility for Strategy 3.1 Strategy 3.2: 5 Long 6-Month Calls / 5 Short 1-Month Straddles • Long-dated calls supply high positive Vega; short-dated straddles sup- ply nearly equal negative Vega. 17 • The simulated premium curve is close to horizontal, indicating minimal sensitivity to σ. • Conclusion: Volatility changes have little impact, again demonstrat- ing Vega-neutral behaviour. Figure 3.4: Net premium vs. volatility for Strategy 3.2 Strategy 3.3: +5 Long 6-Month Calls, −7 Short 1-Month Calls, −3 Short 3-Month Puts • Positive Vega from the long 6-month calls is offset by negative Vega from the two short legs. • The resulting premium line is the flattest of all three strategies, con- firming the strongest Vega neutralisation. • Conclusion: Strategy 3.3 delivers the most stable portfolio value across the tested volatility range. 18 Figure 3.5: Net premium vs. volatility for Strategy 3.3 3.4 Generalisation of Strategies Strategy 3.2 — Mixed-Maturity Call & Short Straddle 1. Long-Volatility Leg (Positive Vega) Select near-the-money calls with a long expiration (typically 6–12 months). • Why? Vega rises with time-to-maturity and is largest when the option is at– or slightly in–the–money. • Quantity rule: Start with nlCong = N contracts. 2. Short-Volatility Leg (Negative Vega) Sell a short-dated straddle at the same strike: 1 ATM call + 1 ATM put, expiry 1–2 months. • Why? Short-dated, at-the-money options carry high negative Vega per dollar of premium, allowing efficient offset. • Quantity rule: Set the number of straddles nstraddle so that nlCong×Vegalong call ≈ nstraddle×(Vegashort call + Vegashort put). 3. Net Effect With the positive Vega from the long-dated calls offset by the negative Vega of the short straddles, the portfolio’s value remains nearly constant when volatility shifts, while still allowing directional or theta views. 19 Strategy 3.3 — Three-Maturity Five–Seven–Three Port- folio 1. Leg A (+5 Long Calls, High Vega) Buy +5 ATM calls expiring in roughly 6 months. • Generates the bulk of positive Vega exposure and provides upside Delta. 2. Leg B (−7 Short Calls, Strong Negative Vega) Sell −7 ATM calls expiring in about 1 month. • Short maturity produces a large negative Vega per contract, effi- ciently offsetting Leg A. • Because strike matches spot, the Delta of this leg partially hedges the Delta of Leg A. 3. Leg C (−3 Short Puts, Supplemental Negative Vega) Sell −3 out-of-the-money puts expiring in roughly 3 months. • Adds additional negative Vega and a modest positive Theta credit. • Slightly OTM strike limits downside Delta until the market falls materially. 4. Balancing Rule Compute each leg's Vega using current IVs; adjust the 5–7–3 ratio so that 5×Vegalong 6M call − 7×Vegashort 1M call − 3×Vegashort 3M put ≈ 0. 5. Resulting Profile • Net Vega near zero across a wide volatility band. • Limited net Delta (long-biased when spot is near strike, but cush- ioned by the short calls). • Positive Theta from the two short legs compensates for long-call time decay. Both templates let a practitioner substitute another underlying, shift strikes, or fine-tune quantities while following the same Vega-neutral construction logic. 20 Chapter 4 Back-Testing of Strategy and Results 4.1 Framework Data Sources Back-testing draws on two raw files obtained from yfinance: • spy spot history.csv – daily SPY open, high, low, close and volume. The Close price supplies the spot level S0 for each trade date. • spy option chain.csv – complete option chain for all listed expiries on the download date, including bid–ask quotes, strikes, expirations and option type (call/put). Trade-Date Selection • Window: every Friday from 2024-07-05 through 2025-04-25. • For each trade date t we identify three expiries: 1. the nearest contract at least 30 days away (short-dated), 2. the nearest contract at least 90 days away (intermediate), 3. the nearest contract at least 180 days away (long-dated). Leg Construction • Long leg (positive Vega): buy +5 at-the-money calls from the long- dated expiry. 21 • Short leg 1 (negative Vega): sell −7 at-the-money calls from the short-dated expiry. • Short leg 2 (negative Vega): sell −3 out-of-the-money puts from the intermediate expiry (strike immediately below S0). • Entry prices are the bid–ask mid-points quoted in spy option chain.csv. Benchmark A naïve baseline consisting of +5 long-dated at-the-money calls is priced in parallel for every trade date so that Vega exposure and P&L can be compared directly. Evaluation Metrics • Per-contract Vega for each leg, calculated from Black–Scholes on the trade date. • Position-sized Vega contribution (N ×Vega) for the three-leg port- folio. • Net portfolio Vega for Strategy 3.3 versus the naïve long-call posi- tion. Visualisation and Findings Three plots summarise the results: 1. Individual-leg Vega curves show how each contract's sensitivity evolves across trade dates. 2. Position-sized bar chart illustrates how the positive Vega from the long calls is offset by the two short legs. 3. Net Vega comparison demonstrates that Strategy 3.3 remains close to zero, whereas the naïve portfolio carries a persistent large positive Vega. Together, the charts confirm that Strategy 3.3 successfully neutralises volatil- ity risk while the naïve alternative remains highly sensitive to changes in implied volatility. 22 Figure 4.1: Per-contract Vega for each leg Figure 4.2: Position-sized Vega contribution of Strategy 3.3 legs Figure 4.3: Net Vega: Strategy 3.3 vs. naïve long-call benchmark 23 Chapter 5 Managerial Implications and Future Work 5.1 Managerial Implications • Volatility Risk Control. Strategy 3.3 returns a portfolio with net Vega centered around zero for the duration of the test interval. For risk managers this translates to a quantifiable decrease in P L swings when implied volatility surges, without the capital cost of delta–gamma rebalancing on a daily basis. • Capital and Margin Efficiency. Hedging long-dated positive Vega against short-dated negative Vega allows desks to reuse margin freed in the short legs, making a more collateral-efficient usage relative to having a large unhedged call position. • Flexible Delta Expression. Since the short-dated calls partially offset the Delta of the long-dated calls, managers can adjust directional exposure by simply adjusting the long-call position, without re-pricing the whole structures. 5.2 Future Work • Cross-Asset Generalization. Use the same Greek-ranking and multi- leg design methodology to other underlyings like equity indexes non- U.S., single-name liquid stocks, or even currency and commodity op- tions. 24 • Dynamic Rebalancing Rules. Back-test the longer term, adding intraday price feed and transaction-cost simulation, so that one can also estimate how frequently the strategy needs to be rebalanced to maintain Vega neutrality within an acceptable tolerance band. • Machine-Learning Enhancement. Replace the linear OLS step with tree-based or neural models in order to catch non-linear Greek interactions; periodically retrain such models so the ranking of the dominant-Greek adjusts to evolving market regimes. • Stress-Scenario Validation. Expose the strategy to historical volatil- ity shock periods (e.g., March 2020) in order to check how well the Vega-neutral characteristic performs when bid–ask spreads increase and liquidity dries up • Regulatory and Reporting Integration. Port the Vega-neutral measure into current risk dashboards to meet upcoming derivative- exposure requirements that focus on sensitivities in addition to Delta and Gamma. 25 Bibliography [1] Xiangyu Ge, Xia Zhu, Gang Bi, Hao Zheng, and Qing Li. Analysis of Op- tion Butterfly Portfolio Models Based on Non-parametric Estimation Deep Learning Method. Research Article, 2023. [2] Matthew J. Krznaric. Comparison of Option Price from Black-Scholes Model to Actual Values. Honours Thesis, University of Akron, Summer 2016. [3] G. Aït-Sahalia and M. Saglam. "Does Vega-Neutral Options Trading Con- tain Information?" Journal of Empirical Finance, vol. 66, 2021, pp. 77–94. :contentReference[oaicite:0]index=0 [4] M. Lundqvist. Delta Neutral and Gamma Neutral Hedging for Option-Based Books. Master's Thesis, KTH Royal Institute of Technology, 2024. :contentRe- ference[oaicite:1]index=1 [5] A. M. Matache and C. Reisinger. "Vega Risk and the Smile." Journal of Risk, vol. 4, no. 4, 2002, pp. 17–35. :contentReference[oaicite:2]index=2 [6] Merrill Edge. "Vega Explained: Understanding Options-Trading Greeks." Ed- ucational white paper, 2019. :contentReference[oaicite:3]index=3 [7] Corporate Finance Institute. "Vega Neutral: How It Works and How to Cre- ate One." Technical note, 2022. :contentReference[oaicite:4]index=4 [8] tastytrade. "What Is Vega in Options Trading and How Does It Work?" Online tutorial, 2023. :contentReference[oaicite:5]index=5 [9] J. Bryant. "Delta–Gamma Hedging: A 2025 Guide for Advanced Traders." The Trading Analyst, 2024. :contentReference[oaicite:6]index=6 [10] John C. Hull. Options, Futures, and Other Derivatives. 11th ed., Pearson, 2021. 26