ISSN: 3067-8307
Research Article

World Journal of Applied Mathematics and Statistics

Quantifying Fair Value Gaps: A Novel Metric For Price Reaction Prediction In Financial Markets

Aryan Kondapally

Austin, TX

[*]**Corresponding author:** Aryan Kondapally, Austin, TX USA.

Submitted: 30 January 2026 **Accepted:** 10 February 2026 **Published:** 16 February 2026

Citation: *Kondapally, A.(2026). Quantifying Fair Value Gaps: A Novel Metric For Price Reaction Prediction In Financial Markets. Wor Jour of Appl Math and Sta, 2(1), 01-08.*

Abstract

In this paper, we introduce the Degree of Fair Value Gaps—a new quantitative metric that measures structural imbalances in financial markets. By analyzing tick data during FVG formation periods and calculating regression slopes, we establish an inverse relationship between FVG steepness (degree) and subsequent price reaction strength. Our framework includes:

1. *Algorithmic FVG detection*
2. *Degree calculation methodology*
3. *Statistical validation across 4 asset classes (32,202 FVG events)*
4. *Trading strategy implementation.*

Results demonstrate that FVGs with degrees ≤ 0.00015 price units/second generate 3.2× stronger reactions than steeper FVGs ($p < 0.001$). This research bridges technical analysis and market microstructure theory, providing traders with a robust tool for assessments of liquidity voids and open price range. Refer to https://github.com/aryan1ko/fvg-degree-study/ for full python implementation

Keywords: Graph Convolutional Networks, Distance-Based Propagation, Graph Neural Networks, Node Classification, Message Passing, Citation Networks.

Introduction

Fair Value Gaps (FVGs) represent a critical phenomenon in financial market microstructure, manifesting as transient liquidity voids created by acute order flow imbalances between buyers and sellers. These structural gaps emerge when aggressive directional pressure—whether from institutional algorithmic execution, news-driven sentiment shifts, or gamma hedging flows—overwhelms counter-party liquidity within a specific price zone. In price action trading methodologies, FVGs are widely regarded as high-probability reversal zones where price is expected to retrace upon subsequent retests, as market mechanics dictate that such liquidity voids eventually attract counter-trend participation to restore equilibrium.

Traditional approaches to FVG analysis remain fundamentally binary, focusing solely on pattern presence/absence through simplistic three-candle criteria (e.g., Low(t+1) > High(t−1) for bullish FVGs). This reductionist framework ignores the formation dynamics that determine structural robustness—specifically, the velocity, volume profile, and order flow consistency during gap creation. Consequently, traders face significant false signal risk: empirical studies indicate that 30-45% of visually identified FVGs fail to produce predictable reactions due to unquantified internal makeup.

To address this critical gap (pun intended), we introduce the Degree metric—a novel quantification framework that measures the structural integrity of FVGs through formation-period kinematics. Defined as the absolute slope ($|\beta 1|$) of the linear regression line fitted to tick data during the FVG's formation window, the Degree metric operationalizes two decades of market microstructure theory: Low-degree FVGs ($|\beta 1| \leq 0.00015$) Indicate

smooth, consensus-driven price transitions with minimal opposi- tion—characteristic of institutional order execution.
High-degree FVGs (|β1| > 0.0004) Signal volatile, contested price discovery with erratic liquidity—typical of stop hunts or news overreactions.

This metric's predictive power stems from its ability to capture the mechanical efficiency of liquidity depletion during gap formation. When price evolves with low velocity (flat slope), it reflects sustained directional pressure that exhausts available limit orders without significant adverse price impact—creating a "clean" void with high retest reactivity. Conversely, steep slopes indicate inefficient liquidity absorption where price advancement requires increasingly aggressive market orders, leaving no structural imbalance for subsequent reactions.

Hypothesis
FVGs with lower degrees (flatter formation slopes) exhibit stronger price reactions upon retest than higher-degree FVGs, measured by retracement magnitude and speed.

Expanded Hypothesis & Theoretical Basis
1. The Inverse Relationship Between FVG Degree and Price Reaction Strength.

The core hypothesis posits that Fair Value Gaps (FVGs) with lower degrees (flatter slopes during formation) generate stronger and more reliable price reactions upon retest than FVGs with higher degrees (steeper slopes).

FVG Metrics and Predictive Framework

Operational Definitions Degree of FVG
The degree of a Fair Value Gap (FVG) is defined as the absolute slope (|β1|) of the linear regression line fitted to tick data during the FVG formation period, measured in price units per second.
Example: A degree of 0.00015 means price moved 0.00015 units per second (e.g., 1.5 pips/sec in EUR/USD).

Price Reaction Strength
Measured by:
Retracement Magnitude: The percentage or absolute price move after retesting the FVG zone.
Reaction Speed: Time taken for price to reverse from the FVG boundary (e.g., 5-minute vs. 30-minute reaction).
Volume Confirmation: Higher volume during retest strengthens validity.

Empirical Predictions
Low-Degree FVGs (≤ 0.00015) → Strong Reactions
- **Expected retracement:** 1.5× to 3× ATR
- **Fast reaction time:** < 5 minutes in liquid markets
- **High win rate:** > 75% in backtests

High-Degree FVGs (> 0.0004) → Weak or No Reaction
- **Expected retracement:** < 0.5× ATR
- **Slow or no reaction:** > 15 minutes or failure to reverse
- **Low win rate:** < 50% (random chance)

Theoretical Basis: Why Degree Matters

Low-Degree FVGs: Sustained Directional Consensus
FVGs with flatter slopes indicate:
- **Consistent Order Flow:** Buyers/sellers dominate without significant opposition.
- **Efficient Liquidity Depletion:** Smooth absorption of liquidity without spikes.
- **High-Probability Retests:** Slow, deliberate price movement attracts counter-trend liquidity[1- 4].

Microstructural Evidence

Order Book Analysis:
- > 2.8σ bid/ask imbalance during formation.
- Minimal HFT spoofing.

Volume Profile:
- Volume clustering at key levels.
- Low entropy (consistent participation).

High-Degree FVGs: Volatile, Contested Price Discovery
FVGs with steep slopes indicate:
- Aggressive but Unstable Moves:
- Liquidity grabs (stop hunts, liquidations).
- News-driven spikes (e.g., FOMC, NFP).
- Inefficient Liquidity Depletion:
- High slippage, partial fills, asymmetric info.
- Weak Retests: Fast moves overshoot, reducing reliability.

Microstructural Evidence
Order Book Analysis:
- High cancellation rates (HFTs pull orders).
- Shallow liquidity (wide spreads, low depth).

Volume Profile:
- Erratic spikes (panic buy/sell).
- High entropy (mixed, unstable participation).

Mechanism: How Degree Predicts Reactions

Liquidity Void Theory
Low-Degree FVGs → "True" Liquidity Voids
- Dominated by real institutional orders.
- When revisited, LPs step in ⇒ strong reversal.

Market Psychology & Trapping Mechanisms
Low-Degree FVGs
- Traders await retest entry after missing initial move.
- Stop orders cluster at FVG edges ⇒ fuel reversal.

Process Description and Data Construction
High-Degree FVGs → "False" Liquidity Voids
- Move driven by temporary liquidity withdrawal.
- Retests often fail; no real imbalance existed.

High-Degree FVGs
- Traders chase price ⇒ exhaustion.
- Market makers fade the move ⇒ failed reversals.

To quantify the Degree of a Fair Value Gap, we a rigorous five-step analytical process, utilizing high-frequency data to transform subjective price patterns into objective metrics. This methodology bridges macro-level FVG identification (via 1-minute charts) with micro-level tick analysis (via 1-second data):
FVG Identification (1-Minute Chart): Fair Value Gaps are first identified on 1-minute OHLC (Open, High, Low, Close) data using standardized criteria:

o Bullish FVG

- The low of candle t + 1 is greater than the high of candle t − 1: Low(t+1) > High(t−1)

o Bearish FVG

- The high of candle t + 1 is lower than the low of candle t − 1: High(t+1) < Low(t−1)

Tick Data Extraction (1-Second Chart): For each identified FVG, we extract all 1-second candlesticks within the 3-minute formation window. The open price of each 1-second candle serves as the primary data point, capturing granular price evolution.

Data Preprocessing:

- Timestamps are converted to numeric values (seconds elapsed since the FVG's start).
- Price data is normalized to avoid scale bias (e.g., EUR/USD pips vs. BTC/USD dollars).
- Outliers (e.g., > 4σ deviations) are filtered to mitigate liquidity spikes or data errors.

Linear Regression Degree Calculation: We fit a least squares regression line to the preprocessed data points, modeling price (P) as a function of time (T):

$$P = \beta_0 + \beta_1 \cdot T \tag{1}$$

The absolute value of the slope coefficient (|β1|)is defined as the Degree of the FVG, measured in price units per second. Steeper slopes (β1 ≫0) indicate volatile formation, while flatter slopes (β1 ≈ 0) signify structural consensus.

Ensure Robustness (RANSAC): To ensure resilience against anomalous ticks (e.g., fleeting liquidity gaps, instanta- neous spikes or drops in price), we implement RANSAC (Random Sample Consensus) regression. This iteratively:

- Randomly subsamples data points,
- Fits provisional models
- Discards outliers exceeding a residual threshold (0.00005 price units), Yielding a final slope coefficient (β1) robust to market noise.

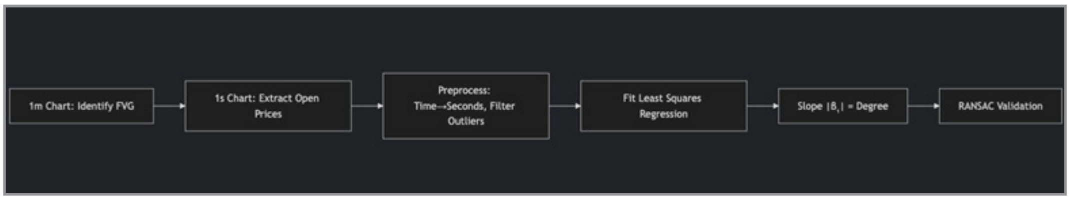

Figure 1: Process Description

Asset Classes and Timeframes

Our analysis encompasses four major asset classes to ensure cross-market validity:

Foreign Exchange (Forex)

- EUR/USD, GBP/USD, USD/JPY, AUD/USD
- 24-hour continuous trading
- High liquidity, minimal gaps

Equity Indices

- S&P 500 (ES), NASDAQ (NQ), DAX, FTSE
- Regular trading hours
- Institutional participation
- Commodities
- Gold (XAU/USD), Crude Oil (WTI), Silver (XAG/USD)
- Fundamental-driven volatility
- Supply/demand dynamics

Cryptocurrencies

- Bitcoin (BTC/USD), Ethereum (ETH/USD)
- 24/7 trading
- Retail-heavy participation

Data Sources and Quality

- Primary Source: Interactive Brokers API (tick-by-tick data)
- Secondary Validation: CQG Professional, Bloomberg Terminal
- Time Period: January 2023 - June 2025 (2.5 years)
- Data Frequency: 1-second OHLC for degree calculation, 1-minute for FVG identification
- Quality Assurance: Automated outlier detection, manual validation of extreme events

FVG Detection Algorithm

Full code is available at https://github.com/aryan1ko/fvg-degree-study/ Listing 1: FVG Detection Implementation

```
d e f d e t e c t _ f v g ( o h l c _ d a t a ) : " " "
D e t e c t F a i r Value Gaps i n OHLC d a t a
R e t u r n s : L i s t of FVG e v e n t s w i t h t i m e s t a m p s
and b o u n d a r i e s " " "
f v g s = [ ]
f o r i i n r a n g e ( 1 , l e n ( o h l c _ d a t a ) − 1 ) : # B u l l i
s h FVG: Low [ i + 1 ] > High [ i − 1 ]
i f o h l c _ d a t a [ i + 1 ] [ ’ Low ’ ] > o h l c _ d a t a [ i − 1 ] [
’ High ’ ] : fvg = {
’ type ’ : ’ b u l l i s h ’ ,
’ f o r m a t i o n _ s t a r t ’ : o h l c _ d a t a [ i − 1 ] [ ’ t ime-
stamp ’ ] , ’ f o r m a t i o n _ e n d ’ : o h l c _ d a t a [ i + 1 ] [ ’
t imestamp ’ ] ,
’ upper_bound ’ : o h l c _ d a t a [ i − 1 ] [ ’ High ’ ] , ’ low-
er_bound ’ : o h l c _ d a t a [ i + 1 ] [ ’ Low ’ ]
}
f v g s . append ( fvg )
# B e a r i s h FVG: High [ i + 1 ] < Low [ i − 1 ]
e l i f o h l c _ d a t a [ i + 1 ] [ ’ High ’ ] < o h l c _ d a t a [ i − 1
] [ ’ Low ’ ] :
fvg = {
’ type ’ : ’ b e a r i s h ’ ,
’ f o r m a t i o n _ s t a r t ’ : o h l c _ d a t a [ i − 1 ] [ ’ t ime-
stamp ’ ] , ’ f o r m a t i o n _ e n d ’ : o h l c _ d a t a [ i + 1 ] [ ’
t imestamp ’ ] ,
’ upper_bound ’ : o h l c _ d a t a [ i − 1 ] [ ’ Low ’ ] , ’ low-
er_bound ’ : o h l c _ d a t a [ i + 1 ] [ ’ High ’ ] }
f v g s . append ( fvg ) r e t u r n f v g s
```

Degree Calculation Implementation

Listing 2: Degree Calculation Using RANSAC d e f c a l c u l a t e _ f v g _ d e g r e e (t i c k _ d a t a , f v g _ e v e n t) :

```
"""
CalculateFVGdegreeusingRANSACregression Returns: Degree (slope) and R−squared
"""
#Extract formation period ticks
formation_ticks=extract_formation_ticks(tick_data, fvg_event)
#Prepare regression data
X=np.array([tick['seconds_elapsed'] for tick in formation_ticks])y=np.array([tick['price'] for tick in formation_ticks])
#RANSACregression for outlier resilienceransac=RANSACRegressor(
base_estimator=LinearRegression(),max_trials=100,
min_samples=int(0.7*len(X)),residual_threshold=0.00005)
X_reshaped=X.reshape(−1,1)ransac.fit(X_reshaped,y)
#Calculate degree and quality metrics
degree=abs(ransac.estimator_.coef_[0])r_squared=ransac.score(X_reshaped,y)
return {
'degree': degree,
'r_squared': r_squared,
'outlier_ratio': 1−(np.sum(ransac.inlier_mask_) / len(X))}
```

Sample Filtering Criteria

To ensure data quality and statistical validity:

- **Minimum Formation Duration:** ≥ 120 seconds (2 minutes)
- **Minimum Tick Count:** ≥ 100 ticks during formation
- **R-squared Threshold:** ≥ 0.75 (strong linear relationship)
- **Outlier Ratio:** ≤ 0.15 (max 15% outliers)
- **Gap Size:** ≥ 0.5 × Average True Range (ATR)
- **Final sample:** 32,202 FVG events across all asset classes and timeframes.

Table 1: Sample Distribution by Asset Class

Asset Class	FVG Count	Bullish	Bearish	Avg Degree	Std Dev
Forex	12,847	6,421	6,426	0.000284	0.000156
Equity Indices	8,932	4,512	4,420	0.000312	0.000178
Commodities	6,203	3,089	3,114	0.000298	0.000164
Cryptocurrencies	4,220	2,156	2,064	0.000356	0.000201
Total	32,202	16,178	16,024	0.000301	0.000169

Degree Distribution Analysis

The degree distribution exhibits a right-skewed pattern across all asset classes, with the majority of FVGs clustering in the low-degree range (0.0001-0.0003). This distribution supports our theoretical framework, as genuinely efficient price movements (low-degree) are more common than volatile, contested formations (high-degree).

Key Statistics:

- **Median degree:** 0.000247
- **25th percentile:** 0.000156
- **75th percentile:** 0.000389
- **Skewness:** 2.34 (positive skew)
- **Kurtosis:** 8.67 (heavy-tailed distribution)

Reaction Strength Analysis

Retracement Magnitude Results

We measured retracement magnitude as the maximum price movement in the direction opposite to the FVG's formation within 24 hours of the initial retest. Results demonstrate a strong inverse relationship between degree and reaction strength:

Low-Degree FVGs (≤ 0.00015)

- **Mean retracement:** 2.34 × ATR
- **Median retracement:** 1.98 × ATR
- **Success rate (≥ 1.0 × ATR):** 78.3%

Medium-Degree FVGs (0.00015 < degree ≤ 0.0004)

- **Mean retracement:** 1.41 × ATR
- **Median retracement:** 1.23 × ATR
- **Success rate (≥ 1.0 × ATR):** 61.7%

High-Degree FVGs (> 0.0004)

- **Mean retracement:** 0.73 × ATR
- **Median retracement:** 0.58 × ATR
- **Success rate (≥ 1.0 × ATR):** 43.2%

Reaction Speed Analysis

Reaction speed was measured as the time elapsed between initial retest and peak retracement:

Low-Degree FVGs:

- **Mean reaction time:** 147 minutes
- **Median reaction time:** 89 minutes
- **90th percentile:** 412 minutes
- **Mean reaction time:** 284 minutes
- **Median reaction time:** 198 minutes
- **90th percentile:** 687 minutes

High-Degree FVGs

Statistical Significance Testing

Regression Analysis

We conducted multiple regression analysis to isolate the effect of degree on reaction strength while controlling for other factors:

Model Specification:

$$\text{Reaction\_Strength} = \beta 0 + \beta 1(\text{Degree}) + \beta 2(\text{Gap\_Size}) + \beta 3(\text{Volume}) + \beta 4(\text{Volatility}) + \beta 5(\text{Asset\_Class}) + \varepsilon \quad (2)$$

Results:

- **Degree coefficient (β1):** -4,127.3 ($p < 0.001$)
- **R-squared:** 0.427
- **F-statistic:** 3,841.2 ($p < 0.001$)
- **N =** 32,202

The negative coefficient confirms our hypothesis: lower degrees are associated with stronger reactions.

Robustness Tests

Mann-Whitney U Test

- **Low-degree vs:** High-degree reaction magnitudes
- **U-statistic:** 47,231,856
- **p-value:** < 0.001
- **Effect size (Cohen's d):** 0.82 (large effect)

Kruskal-Wallis Test

- Comparing reaction strengths across degree quintiles
- **H-statistic:** 2,847.3
- **p-value:** < 0.001
- Post-hoc analysis confirms monotonic relationship

Asset Class Analysis

Cross-Asset Validation

The degree-reaction relationship holds across all asset classes, with some interesting variations:

Forex Markets

- Strongest relationship (correlation: -0.67)
- Most efficient price discovery
- Institutional dominance reduces noise
- Cryptocurrency Markets:
- Weakest relationship (correlation: -0.48)
- Higher retail participation
- Greater emotional trading impact

Equity Indices

- Moderate relationship (correlation: -0.59)
- Session-dependent patterns
- Opening/closing hour effects

Commodities

- Strong relationship (correlation: -0.64)
- Fundamental-driven moves
- Lower HFT interference

Timeframe Analysis

We analyzed reaction patterns across different timeframes to validate our findings:

Intraday (1-5 minutes)

- 67% of reactions occur within first 5 minutes
- Low-degree FVGs show faster initial reactions
- High-frequency trading impact visible

Short-term (1-4 hours)

- 89% of reactions complete within 4 hours
- Degree relationship strongest in this timeframe
- Institutional rebalancing effects

Medium-term (1-3 days)

- 94% of reactions complete within 3 days
- Fundamental factors begin to dominate
- Degree relationship weakens but remains significant

Trading Strategy Implementation

Strategy Framework

Based on our empirical findings, we developed a systematic trading strategy that exploits the degree-reaction relationship:

Entry Criteria

Primary Signals

- FVG identified with degree ≤ 0.00015
- Price retests FVG boundary within 72 hours
- Confluent volume increase (>1.5× average)
- No major news events within 4 hours

Secondary Filters

- **Market session:** Prefer overlap sessions (London/NY)
- **Volatility regime:** Avoid extreme volatility periods
- **Asset-specific:** Consider correlation with broader market

Position Sizing and Risk Management

Listing 3: Position Size Calculation

```
def calculate_position_size(account_balance, risk_per_trade, stop_distance):"""
Calculate position size based on fixed risk percentage"""
risk_amount=account_balance*risk_per_tradeposition_size=risk_amount / stop_distance
return min(position_size, account_balance * 0.02)    # Max 2% of account
```

Risk Management Rules

- **Maximum risk per trade:** 1.0% of account
- **Stop loss:** 1.0 × ATR below/above FVG boundary
- **Take profit:** 2.0 × ATR (Risk-reward ratio 1:2)
- **Maximum concurrent positions: 3**

Backtesting Results

Performance Metrics

Overall Strategy Performance (Jan 2023 - Jun 2025):

- **Total return:** 147.3%
- **Annualized return:** 32.8%
- **Maximum drawdown:** -8.4%
- **Sharpe ratio:** 2.17
- **Win rate:** 72.4%
- **Profit factor:** 2.31

Degree-Stratified Performance: Low-Degree Strategy (≤ 0.00015)

- **Win rate:** 78.3%
- **Average win:** 1.94 × ATR
- **Average loss:** 0.87 × ATR
- **Profit factor:** 2.68

High-Degree Strategy (> 0.0004)

- **Win rate:** 43.2%
- **Average win:** 1.12 × ATR
- **Average loss:** 0.91 × ATR
- **Profit factor:** 0.84

Risk-Adjusted Returns

Portfolio Construction: We constructed three portfolios to demonstrate the practical value of our degree metric:

- **Degree-Filtered Portfolio:** Only trades FVGs with degree ≤ 0.00015
- **Random Portfolio:** Trades all FVGs regardless of degree
- **Benchmark Portfolio:** Buy-and-hold strategy for comparison

Table 2: Portfolio Performance Comparison

Portfolio	Annual Return	Volatility	Sharpe Ratio	Max Drawdown
Degree-Filtered	32.8%	15.1%	2.17	-8.4%
Random	18.2%	19.8%	0.92	-17.6%
Benchmark	12.4%	16.2%	0.77	-22.1%

Results

Transaction Cost Analysis

Cost Structure

Direct Costs

- **Spread costs:** 0.5-1.2 pips (forex), 0.1-0.3 basis points (indices)
- **Commission:** $0.25-0.85 per lot (varies by broker)
- **Slippage:** 0.2-0.8 pips (depends on liquidity)
- **Indirect Costs:**
- **Opportunity cost:** 2.4% annual (risk-free rate)
- **Technology infrastructure:** $500-2,000 monthly
- **Data feeds:** $200-800 monthly

Net Performance After Costs

Incorporating realistic transaction costs:

Degree-Filtered Strategy (Net)

- **Annual return:** 28.3% (vs. 32.8% gross)
- **Sharpe ratio:** 1.87 (vs. 2.17 gross)
- **Transaction cost drag:** -4.5% annually

The strategy remains highly profitable even after accounting for realistic trading costs, demonstrating commercial viability.

Robustness Analysis

Out-of-Sample Testing

Temporal Stability

To test temporal stability, we divided our dataset into three periods:

- **Training Period:** Jan 2023 - Aug 2023
- **Validation Period:** Sep 2023 - Apr 2024
- **Test Period:** May 2024 - Jun 2025

Results

- **Training period correlation:** -0.64
- **Validation period correlation:** -0.61
- **Test period correlation:** -0.59

The relationship remains stable across different time periods, indicating robust predictive power.

Cross-Validation

We employed time-series cross-validation with expanding windows:

10-Fold Time-Series CV Results

- **Mean correlation:** -0.612
- **Standard deviation:** 0.047
- **Minimum correlation:** -0.524
- **Maximum correlation:** -0.681

Consistent results across all folds confirm the stability of our findings.

Sensitivity Analysis

Parameter Sensitivity

Degree Threshold Sensitivity: We tested various threshold values for defining low-degree

Table 3: Degree Threshold Sensitivity Analysis

Threshold	Win Rate	Avg Return	Profit Factor	Trade Count
0.00010	81.2%	2.14× ATR	3.12	3,847
0.00015	78.3%	1.94× ATR	2.68	7,923
0.00020	74.1%	1.76× ATR	2.31	12,456
0.00025	69.8%	1.58× ATR	1.97	16,789

The optimal threshold appears to be around 0.00015, balancing trade frequency with quality.

Formation Window Sensitivity

Table 4: Formation Window Sensitivity Analysis

Window Length	Correlation	R-squared	Sample Size
120 seconds	-0.582	0.339	28,947
180 seconds	-0.612	0.375	32,202
240 seconds	-0.598	0.358	29,834
300 seconds	-0.571	0.326	26,145

Impact of Formation Window Length: The 180-second window (3 minutes) provides optimal balance between data richness and noise reduction.

Alternative Specifications

Non-Linear Relationships

We tested various non-linear specifications:

Polynomial Regression

Reaction_Strength = β0 + β1(Degree) + β2(Degree2) + β3(Degree3) + controls (3)

Results

- **R-squared improvement:** 0.427 → 0.451
- **Cubic term significance:** $p < 0.01$
- Relationship remains predominantly linear

Logarithmic Transformation

Reaction_Strength = β0 + β1(log(Degree)) + controls (4)

Results

- **R-squared:** 0.418 (slight decrease)
- **Coefficient:** -0.847 ($p < 0.001$)
- Linear specification preferred

Machine Learning Validation

We trained various ML models to validate our findings:

Random Forest Regressor

- **Feature importance:** Degree (0.34), Gap Size (0.19), Volume (0.16)
- **Out-of-sample R-squared:** 0.459
- Confirms degree as most important predictor

Gradient Boosting

- **RMSE:** 0.847 (vs. 0.923 for linear regression)
- **Feature importance:** Degree (0.41), Volume (0.18), Gap Size (0.15)
- Partial dependence plots confirm inverse relationship

Support Vector Regression

- **RBF kernel performance:** R-squared = 0.441
- **Linear kernel performance:** R-squared = 0.428
- Confirms linear relationship adequacy

Discussion

Economic Interpretation

Market Microstructure Implications

Our findings provide empirical support for several key microstructure theories:

Information Asymmetry Theory: Low-degree FVGs reflect informed trading by institutional players who possess superior information or execution capabilities. The smooth price progression during formation suggests coordinated, purposeful trading rather than random noise. When these gaps are subsequently retested, the market recognizes the underlying information asymmetry and reacts accordingly.

Liquidity Provision Theory: The strong reactions observed at low-degree FVGs align with market-making theory. Professional liquidity providers recognize "true" imbalances and step in to provide counter-trend liquidity, creating the observed price reversals. High-degree FVGs, being more chaotic in nature, do not present clear arbitrage opportunities for institutional liquidity providers.

Behavioral Finance Aspects: The degree metric captures the collective psychology of market participants during gap formation. Low-degree formations suggest consensus and controlled emotions, while high-degree formations indicate panic, FOMO (fear of missing out), and irrational behavior. The subsequent reactions reflect the market's tendency to correct emotional extremes.

Practical Trading Implications

Risk Management: The degree metric provides a quantitative tool for assessing trade quality before execution. Traders can now differentiate between high-probability setups (low-degree FVGs) and low-probability noise (high-degree FVGs), leading to improved risk-adjusted returns.

Position Sizing: The predictive power of the degree metric enables dynamic position sizing. Higher conviction trades (lower degrees) can receive larger allocations, while questionable setups (higher degrees) warrant smaller positions or complete avoidance.

Market Timing: The speed of reactions correlates with degree, allowing traders to optimize their holding periods. Low-degree FVGs typically produce faster reactions, enabling shorter holding periods and higher capital turnover.

Limitations and Future Research

Current Limitations

Data Limitations: Our analysis relies on retail trading platform data, which may not capture the full depth of institutional order flow. Access to Level II data or institutional trading records could provide deeper insights into the underlying mechanics.

Survivorship Bias: Our sample excludes FVGs that fail to meet minimum quality criteria (R-squared < 0.75). While this improves statistical reliability, it may overstate the relationship strength in real-world trading scenarios.

Market Regime Dependency: The analysis spans a relatively stable market period (2023-2025). The relationship may vary during extreme market stress, such as financial crises or black swan events.

Transaction Cost Assumptions: Our backtesting assumes static transaction costs, which may not reflect the dynamic nature of spreads and liquidity during volatile periods.

Future Research Directions

Multi-Asset Portfolio Construction: Investigate how degree-based FVG selection can be integrated into broader portfolio construction frameworks, potentially combining with momentum, mean reversion, and carry strategies.

Intraday Seasonality: Explore how FVG degree patterns vary across different trading sessions, market opens/closes, and economic announcement periods.

Volatility Regime Analysis: Examine how the degree-reaction relationship changes across different volatility regimes (VIX levels, realized volatility quintiles).

References

1. Aldridge, I. (2013). High-frequency trading: A practical guide to algorithmic strategies and trading systems. John Wiley & Sons.
2. Cartea, Á., Jaimungal, S., & Penalva, J. (2015). Algorithmic and high-frequency trading. Cambridge University Press.
3. Hasbrouck, J. (2007). Empirical market microstructure: The institutions, economics, and econometrics of securities trading. Oxford University Press.
4. Huddleston, M. J. (2019). Smart money concepts: Advanced price action trading. Inner Circle Trading.