A Volume-Weighted Uncertainty Principle for Equity Microstructure: Definitions, Empirical Properties, Regime Analysis, and Theoretical Extensions

Songgun Lee[1]

[1]Department of Physics, University of California, Santa Barbara

April 2026 (Preprint)

Abstract

We introduce a volume-weighted uncertainty product $U = \sigma_P \cdot \sigma_t$ for equity microstructure, constructed in direct analogy with the Heisenberg uncertainty principle of quantum mechanics. The price spread σ_P and time spread σ_t are computed as volume-weighted standard deviations of the intraday 1-minute price and time series, centered on the volume-weighted median price M — a quantity we term the Volume-Weighted Median Price (VWMP), which we argue is theoretically preferable to the widely-used Volume-Weighted Average Price (VWAP) as a reference price for microstructure analysis. Applying this framework to SPY (S&P 500 ETF) over 2020–2026 ($N = 1{,}575$ trading days, approximately 614,000 intraday 1-minute observations), we find that $U_{\text{norm}} = (\sigma_P/M) \cdot \sigma_t$ follows a log-normal marginal distribution (KS $p = 0.155$), exhibits strong long memory ($H \approx 0.90$ via R/S analysis), and is significantly correlated with the CBOE VIX ($r = 0.73$) while retaining 54% independent variance. Granger causality tests confirm that the VIX-orthogonal component of U_{norm} linearly precedes daily returns at all lags 1–5 ($p < 0.03$). A market regime analysis reveals a striking monotone relationship: the correlation between U_{norm} and forward 5-day rolling volatility rises from effectively zero ($r = 0.002$) in low-VIX environments (VIX < 15) to $r = 0.427$ during high-stress regimes (VIX > 25), and to $r = 0.610$ during risk-off (flight-to-safety) periods. We further propose the Robertson uncertainty relation as a framework for five candidate conjugate variable pairs in financial markets, and outline theoretical extensions drawing on time-dependent perturbation theory, Clebsch-Gordan decomposition of composite uncertainty states, and a path-integral interpretation of U_{norm} as a field amplitude.

---

1 Introduction

1.1 Motivation

This work originated during an undergraduate course in quantum mechanics at the University of California, Santa Barbara, while studying the Heisenberg uncertainty principle and its generalization, the Robertson uncertainty relation. The core observation is deceptively simple: *financial markets are not so different from quantum systems in one important structural sense* — neither allows the

simultaneous, arbitrary precision of two conjugate observables. In quantum mechanics, the position and momentum of a particle cannot both be specified to arbitrary precision; the product of their uncertainties is bounded below by a fundamental constant. In a financial market, price and time are not independent — a rapid, concentrated burst of trading (small σ_t) necessarily compresses the price distribution (large σ_P), and vice versa. The question this paper investigates is whether this structural analogy can be made quantitative, empirically tested, and ultimately useful for predicting future market behavior.

The second motivation is a dissatisfaction with standard practice. Market microstructure research almost universally uses the Volume-Weighted Average Price (VWAP) as the reference price for intraday analysis. This is the mean of the volume-weighted price distribution. We argue that the median of this distribution — the Volume-Weighted Median Price (VWMP), defined as the price M at which cumulative traded volume is split exactly in half — is theoretically superior as a reference price. The VWMP is more robust to block trades and end-of-day imbalances that distort the VWAP, and it is the more natural center for defining a price spread σ_P in the spirit of the uncertainty principle. The near-universal adoption of VWAP over VWMP in both academic and practitioner work is, in our view, an artifact of computational convenience rather than principled choice.

1.2 Contribution

The contributions of this paper are as follows:

1. We formally define the Volume-Weighted Median Price (VWMP) M, the intraday price spread σ_P, the intraday time spread σ_t, and the uncertainty product $U = \sigma_P \cdot \sigma_t$, drawing explicit parallels to quantum mechanical uncertainty.
2. We document the empirical properties of U_{norm} over six years of SPY 1-minute data: log-normal marginal distribution, strong long memory ($H \approx 0.90$), robust correlation with VIX, and Granger-causal relationship to daily returns.
3. We introduce a market regime analysis showing that the predictive power of U_{norm} is highly regime-dependent, rising monotonically from near-zero in calm markets to $r = 0.43$ in stressed markets and $r = 0.61$ during risk-off regimes.
4. We propose the Robertson uncertainty relation as a unifying framework for five candidate conjugate variable pairs in equity microstructure, three of which are implemented and computed on the full historical sample.
5. We outline a research program for applying time-dependent perturbation theory, Clebsch-Gordan decomposition, and path-integral methods to market microstructure, with specific, testable hypotheses.

1.3 Why VWMP Rather Than VWAP

Let $\{(P_i, V_i)\}$ be the set of 1-minute bars for a trading day, with volume weights $w_i = V_i / \sum_j V_j$. The VWAP is the weighted mean:

$$\text{VWAP} = \sum_i w_i P_i \tag{1}$$

The VWMP M is the weighted median:

$$M = \inf \left\{ P : \sum_{i:\, P_i \le P} w_i \ge 0.5 \right\} \tag{2}$$

The theoretical advantages of M over VWAP are:

1. **Robustness.** The weighted mean is sensitive to outlier bars with large volume at extreme prices (e.g., large institutional block trades at the open, or massive market-on-close imbalances). The weighted median is immune to such distortions. On days with a single bar capturing 30% of total volume at an off-market price, VWAP is pulled significantly; M is not.

2. **Consensus interpretation.** M is the price at which exactly half of the day's total volume transacted below it and half above. This is a cleaner definition of "where the market agreed to trade" than the mean: it is the price that divides the market's actual transactions in half, not the volume-weighted centroid of prices.

3. **Natural center for σ_P.** The price spread σ_P is defined as the volume-weighted standard deviation of prices around M. When M is the median, σ_P achieves a near-minimal value in the L^2 sense (the median minimizes the weighted mean absolute deviation; the mean minimizes the weighted mean squared deviation). Using the median as center gives σ_P its tightest interpretation as a genuine spread, not an artifact of distributional skew.

4. **Quantum analogy.** In the position-momentum uncertainty relation, σ_x is defined relative to $\langle x \rangle$ — the expectation value, which for a symmetric distribution coincides with the median. For the skewed intraday price distributions typical of equity markets (dominated by open and close volume spikes), the median is arguably closer in spirit to the "most likely transaction price" than the mean.

1.4 Related Work

The application of physics concepts to financial markets — econophysics — has a substantial literature. Mantegna and Stanley [5] established the statistical mechanics of financial data. Bouchaud and Potters [6] developed a comprehensive theory of financial risk using methods from statistical physics. The specific application of the uncertainty principle to finance has been explored in the context of time-frequency analysis (Gabor 1946, applied to trading in various forms), but the direct construction of a volume-weighted uncertainty product from microstructure data, with empirical validation of its distributional properties and predictive content, does not appear to have been previously reported.

1.5 Organization

The remainder of the paper is organized as follows. Section 2 describes the data and infrastructure. Section 3 defines the mathematical framework. Sections 4–12 present empirical results: distributional properties, stationarity, long memory, VIX relationship, forward volatility prediction, market regime analysis, Granger causality, VAR analysis, and machine learning results. Section 13 describes the live prediction tracking system. Section 14 covers infrastructure. Section 15 develops the Robertson uncertainty relation framework. Section 16 presents theoretical extensions. Section 17 lists open questions and future directions.

2 Instrument and Data

- **Instrument:** SPY (SPDR S&P 500 ETF Trust), used as proxy for the S&P 500 index
- **Data source:** Alpaca Markets API (free tier), 1-minute OHLCV bars
- **Date range:** 2020-01-02 to 2026-04-10
- **Sample size:** $N = 1571$ trading days
- **Intraday resolution:** 1-minute bars, 390 bars per trading day (9:30am – 4:00pm ET)
- **Infrastructure:** Mac mini (Apple M-series), Python 3.11, automated daily ingestion via cron

3 Mathematical Framework

3.1 Notation

For a given trading day, let:

$$\{(P_i, V_i, t_i)\}_{i=1}^{n} \tag{3}$$

denote the set of 1-minute bars, where P_i is the closing price, V_i is the traded volume, and t_i is the time in minutes elapsed since market open ($t_i = 0$ at 9:30am ET, $t_i = 389$ at 4:00pm ET). Bars with $V_i = 0$ are excluded.

3.2 Volume Weights

Volume is treated as a probability measure over the intraday price-time space:

$$w_i = \frac{V_i}{\sum_{j=1}^{n} V_j}, \qquad \sum_{i=1}^{n} w_i = 1 \tag{4}$$

3.3 Volume-Weighted Median Price M

M is defined as the 50th percentile of the volume-weighted empirical distribution of prices:

$$M = \inf \left\{ P \ : \sum_{i:\, P_i \le P} w_i \ge 0.5 \right\} \tag{5}$$

This is the price at which cumulative traded volume is equally split. It is distinct from the volume-weighted average price (VWAP), which is the first moment of the same distribution.

3.4 Price Spread σ_P

Volume-weighted standard deviation of price, centered on the median M:

$$\sigma_P = \sqrt{\sum_{i=1}^{n} w_i (P_i - M)^2} \tag{6}$$

Note: centered on the median, not the mean. Units: USD.

3.5 Time Spread σ_t and Volume Center of Mass $\langle t \rangle$

Volume-weighted mean time (center of mass in time):

$$\langle t \rangle = \sum_{i=1}^{n} w_i \, t_i \tag{7}$$

Volume-weighted standard deviation of time:

$$\sigma_t = \sqrt{\sum_{i=1}^{n} w_i (t_i - \langle t \rangle)^2} \tag{8}$$

Units: minutes.

3.6 Uncertainty Product U

Motivated by the Heisenberg uncertainty principle and the Gabor time-frequency uncertainty relation $\sigma_t \cdot \sigma_f \geq \frac{1}{4\pi}$:

$$U = \sigma_P \cdot \sigma_t \tag{9}$$

Units: USD $\cdot$ minutes.

3.7 Normalized Uncertainty Product U_{norm}

To remove the price-level effect (SPY price approximately doubled over 2020–2026), σ_P is normalized by the median price:

$$U_{\text{norm}} = \frac{\sigma_P}{M} \cdot \sigma_t \tag{10}$$

Units: fraction $\cdot$ minutes. This is the primary quantity used in all subsequent analysis.

3.8 Dimensionless Uncertainty Product U_{dimless}

Further normalization by total trading day length (390 minutes):

$$U_{\text{dimless}} = \frac{\sigma_P}{M} \cdot \frac{\sigma_t}{390} \tag{11}$$

Units: dimensionless (pure number).

3.9 Daily Feature Vector

Each trading day d is represented by the tuple:

$$X_d = (M_d,\ \sigma_{P,d},\ \sigma_{t,d},\ U_d,\ U_{\text{norm},d},\ \langle t \rangle_d) \tag{12}$$

4 Summary Statistics

Key observations:

- Coefficient of variation of U_{norm}: CV $= \sigma/\mu \approx 0.91$
- Mean of U_{norm} exceeds median by 26%: strong right skew
- $\langle t \rangle$ is consistently above 195 minutes (midday): volume center of mass is persistently in the afternoon, consistent with the known U-shaped intraday volume pattern

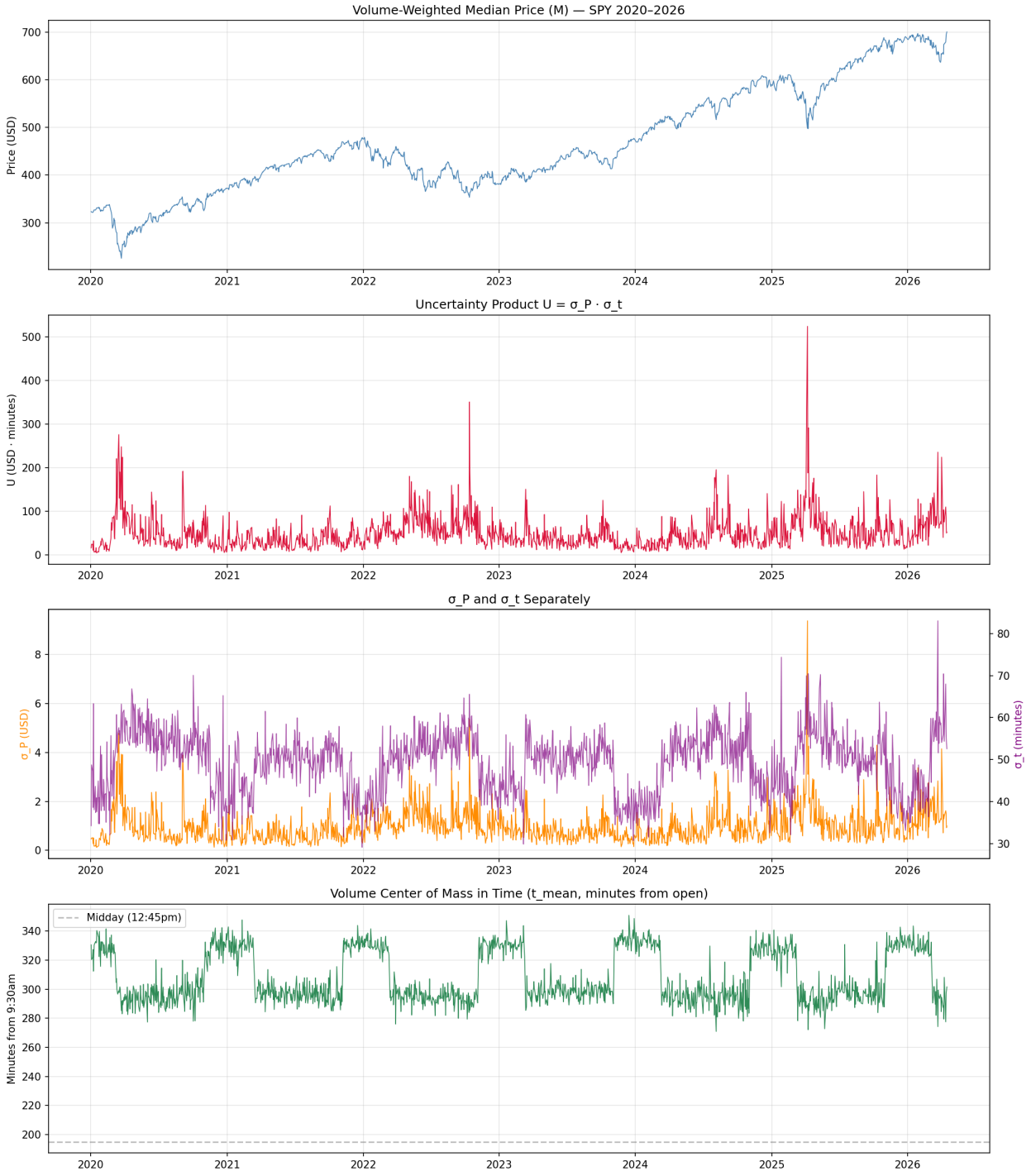

Figure 1: Overview time series. *Top:* Volume-weighted median price M (SPY, 2020–2026). *Second:* Uncertainty product $U = \sigma_P \cdot \sigma_t$. *Third:* σ_P (left axis, orange) and σ_t (right axis, purple) plotted separately. *Bottom:* Volume center of mass $\langle t \rangle$ in minutes from open; dashed line marks midday (195 min). Note the persistent afternoon skew and the sharp U spikes during the March 2020 and April 2025 drawdowns.

Table 1: Summary statistics of the daily feature vector ($N = 1571$ trading days, 2020–2026)

Variable	Mean	Std	Min	25%	Median	75%	Max
M (USD)	—	—	270	—	—	—	670
σ_P (USD)	—	—	—	—	—	—	—
σ_t (min)	—	—	30	—	—	—	85
U (USD·min)	49.05	37.87	4.78	24.29	40.13	61.90	524.52
U_{norm}	0.1120	0.1016	0.0111	0.0535	0.0887	0.1330	1.1081
U_{dimless}	0.000287	0.000261	0.000028	0.000137	0.000228	0.000341	0.002841
$\langle t \rangle$ (min)	—	—	270	—	—	—	345

5 Distributional Properties of U_{norm}

5.1 Distribution Fitting

Four candidate distributions were fit to U_{norm} using maximum likelihood estimation. Goodness-of-fit was assessed via the Kolmogorov-Smirnov (KS) test and log-likelihood.

Table 2: Distribution fitting results for U_{norm} ($N = 1571$)

Distribution	KS statistic	KS p-value	log-likelihood
Log-normal	**0.0284**	**0.1546**	**2167.80**
Pareto	0.1066	0.0000	2031.96
Exponential	0.1061	0.0000	2031.95
Gamma	0.3908	0.0000	948.78

Result: The log-normal distribution is the only candidate not rejected by the KS test ($p = 0.1546 > 0.05$). All other candidates are strongly rejected ($p \approx 0$). The log-normal achieves the highest log-likelihood by a margin of 136 units over the next best candidate.

5.2 Log-Normal Parameters

$$U_{\text{norm}} \sim \text{LogNormal}(\mu,\, \sigma^2) \tag{13}$$

with fitted parameters:

$$\mu = -2.452, \qquad \sigma = 0.708 \tag{14}$$

Implied median: $e^{-2.452} \approx 0.0865$, consistent with observed median of 0.0887.

5.3 QQ Plot Verification

A QQ plot of $\log(U_{\text{norm}})$ against the standard normal distribution shows close adherence to the diagonal across the range $[-2.5, +2.5]$ of theoretical quantiles. A small downward deviation is observed in the lower tail, indicating a slight excess of very-low-U days relative to strict log-normality.

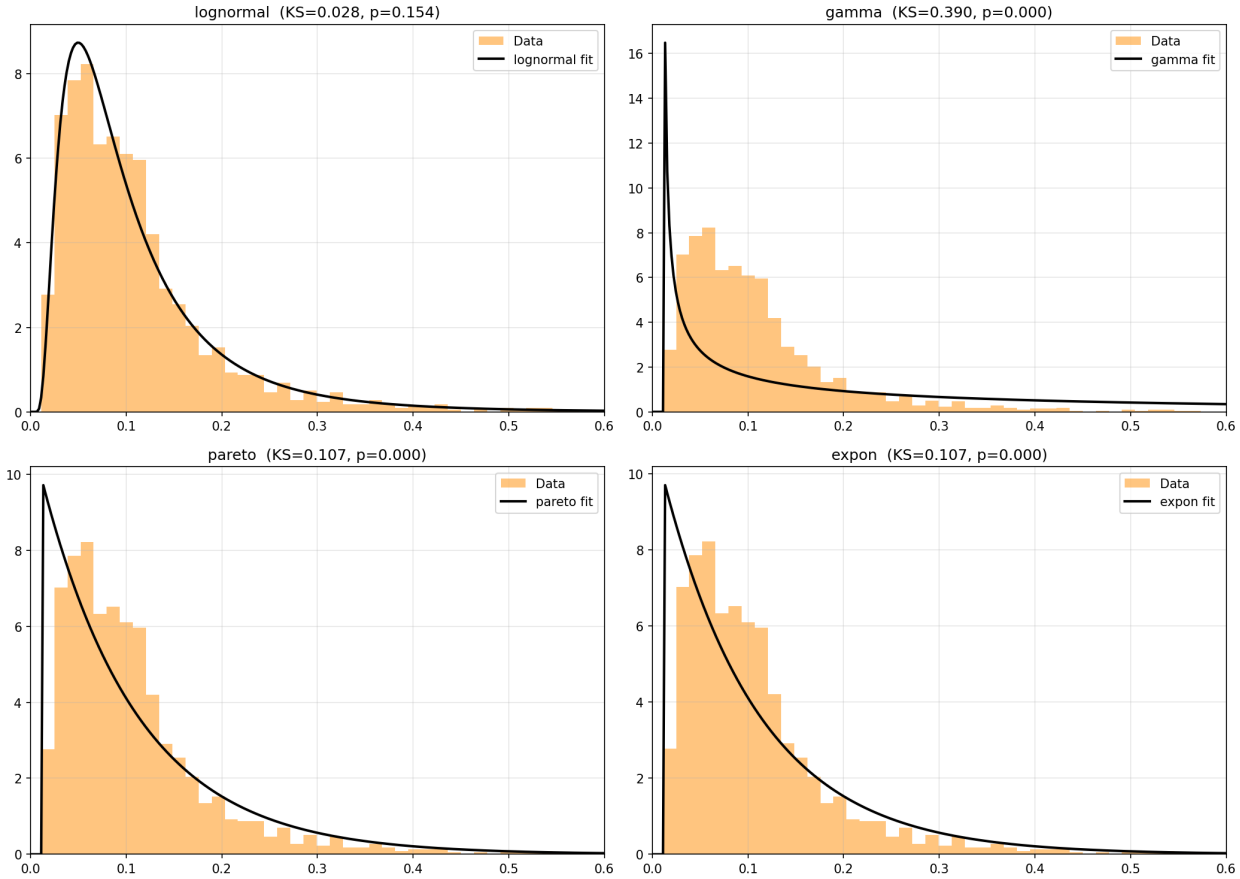

Figure 2: Distribution fits for U_{norm}. Each panel overlays the empirical histogram (orange) with the MLE-fitted density (black line). The log-normal fit ($p = 0.1546$) is the only candidate not rejected by the KS test; gamma, Pareto, and exponential are all rejected at $p \approx 0$.

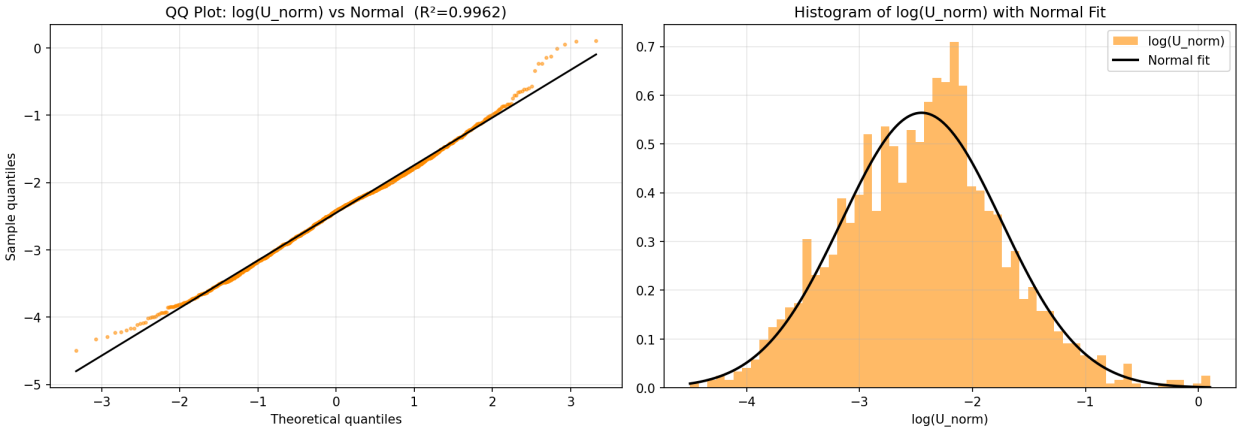

Figure 3: *Left:* QQ plot of $\log(U_{\text{norm}})$ against the standard normal. The close adherence to the diagonal confirms log-normality of the marginal distribution. *Right:* Histogram of $\log(U_{\text{norm}})$ with normal fit overlaid.

5.4 Caveat

The KS test assumes i.i.d. observations. U_{norm} exhibits strong autocorrelation (lag-1 ≈ 0.62), which violates this assumption and makes the reported p-value optimistic. The distributional claim should be interpreted as applying to the marginal distribution, not the joint process.

6 Stationarity Analysis

Table 3: Stationarity tests. ADF null: unit root present. KPSS null: series is stationary.

Series	ADF stat	ADF p	ADF verdict	KPSS stat	KPSS verdict
U_{norm}	-6.025	0.000	stationary	0.512	non-stationary ($p = 0.039$)
$\log(U_{\text{norm}})$	-6.266	0.000	stationary	0.457	stationary ($p > 0.05$)
$\Delta \log(U_{\text{norm}})$	-15.89	0.000	stationary	0.027	stationary ($p > 0.10$)

Conclusion: U_{norm} gives conflicting test results (likely due to structural breaks / regime changes). $\log(U_{\text{norm}})$ passes both tests and is the appropriate series for long-memory analysis. $\Delta \log(U_{\text{norm}})$ is over-differenced for Hurst estimation purposes.

7 Long Memory Analysis

7.1 Autocorrelation Function

- Lag-1 autocorrelation of U_{norm}: ≈ 0.62
- Autocorrelation remains statistically significant above the 95% confidence band until approximately lag 30 (6 trading weeks)
- Lag-1 autocorrelation of U_{residual} (VIX-orthogonal component): ≈ 0.25, significant to lag ≈ 20

7.2 Hurst Exponent

Two estimators were applied to $\log(U_{\text{norm}})$:

Table 4: Hurst exponent estimates for $\log(U_{\text{norm}})$

Method	Description	H
R/S analysis	Rescaled range	0.8962
DFA	Detrended Fluctuation Analysis	0.9955
Difference		0.0993

Interpretation: Both estimators give $H \gg 0.5$, indicating strong long memory. $H = 0.5$ corresponds to a random walk (no memory); $H \to 1$ corresponds to strong persistence. The disagreement of 0.099 is within acceptable range; the conservative estimate is $H \approx 0.90 \pm 0.05$.

Caveats:

- R/S analysis is known to overestimate H in the presence of strong autocorrelation

Figure 4: *Top:* Raw uncertainty product U (dollar-minutes). *Middle:* Normalized $U_{\text{norm}} = (\sigma_P/M) \cdot \sigma_t$, which removes the price-level trend visible in U. *Bottom:* Autocorrelation function of U_{norm}; the ACF remains significant well past lag 30, indicating strong persistence.

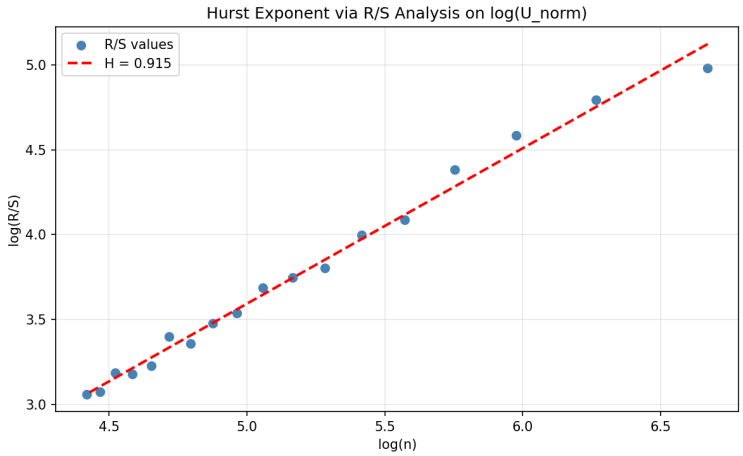

Figure 5: R/S analysis for $\log(U_{\text{norm}})$. The log-log plot of rescaled range R/S against window size n yields slope $H \approx 0.896$, well above the $H = 0.5$ random-walk baseline.

- DFA value of 0.9955 approaches the boundary $H = 1.0$ (Brownian motion), suggesting possible residual non-stationarity
- Both estimators agree qualitatively: H is well above 0.5

8 Relationship to VIX

8.1 Correlation

$$\text{Corr}(U_{\text{norm}}, \text{VIX}) = 0.7305 \tag{15}$$

$$\text{Corr}(\log U_{\text{norm}}, \log \text{VIX}) = 0.6792 \tag{16}$$

8.2 Log-Log Regression

$$\log(U_{\text{norm}}) = 1.5629 \cdot \log(\text{VIX}) - 7.1288 \tag{17}$$

$$R^2 = 0.4614 \tag{18}$$

Interpretation: $U_{\text{norm}} \propto \text{VIX}^{1.56}$. The superlinear exponent (> 1) means U_{norm} amplifies VIX signals: when VIX doubles, U_{norm} nearly triples ($2^{1.56} \approx 2.95$). The $R^2 = 0.46$ implies 54% of $\log(U_{\text{norm}})$ variance is statistically independent of VIX.

8.3 VIX-Orthogonal Residual

Define:

$$U_{\text{residual}} = \log(U_{\text{norm}}) - \widehat{\log(U_{\text{norm}}) \mid \log(\text{VIX})} \tag{19}$$

- U_{residual} has mean ≈ 0, std $= 0.52$
- Residual correlation with log(VIX): 0 (by construction)

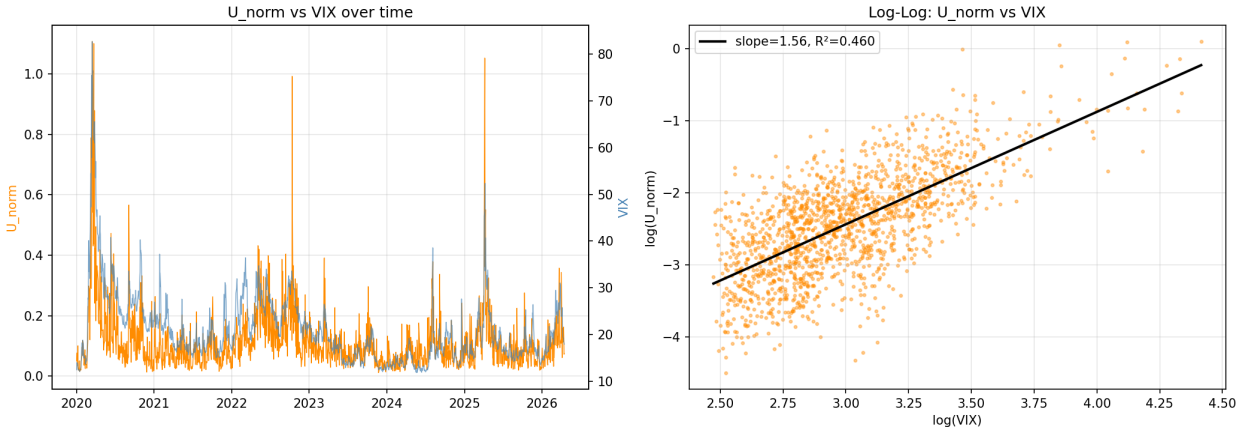

Figure 6: *Left:* U_{norm} (orange, left axis) and VIX (blue, right axis) over time; spikes co-occur during high-volatility episodes. *Right:* Log-log scatter of U_{norm} vs VIX with fitted regression line (slope $= 1.56$, $R^2 = 0.46$).

- Lag-1 autocorrelation of U_{residual}: ≈ 0.25, significant to lag ≈ 20

Conclusion: The VIX-independent component of U_{norm} is not random noise. It has its own persistent autocorrelation structure with memory of approximately 20 trading days (4 weeks).

9 Forward Volatility Prediction

9.1 Terminology Clarification

The volatility measure used in this section is **not** realized volatility in the sense of Andersen & Bollerslev (1998), which is computed by summing squared intraday 1-minute log returns within a single day:

$$\text{RV}_t^{\text{intraday}} = \sqrt{252 \cdot 390} \cdot \sqrt{\sum_{i=1}^{390} r_{t,i}^2}, \qquad r_{t,i} = \log(P_{t,i}/P_{t,i-1}) \tag{20}$$

Instead, the forward volatility measure used here is computed from *daily* changes in the volume-weighted median M, annualized over a rolling window. We call this **rolling daily volatility** (RDV):

9.2 Forward Rolling Daily Volatility

n-day forward rolling daily volatility is defined as:

$$\text{RDV}_{5d}(t) = \sqrt{252} \cdot \sigma\left(\{\log(M_{t+k}/M_{t+k-1})\}_{k=1}^{5}\right) \tag{21}$$

$$\text{RDV}_{10d}(t) = \sqrt{252} \cdot \sigma\left(\{\log(M_{t+k}/M_{t+k-1})\}_{k=1}^{10}\right) \tag{22}$$

Note: True intraday RV and RDV are related but distinct. RDV captures inter-day price variability over a rolling window; true intraday RV captures the within-day price path. Because σ_P is itself an intraday dispersion measure, comparing U_{norm} against true intraday RV is a direction for future work (all raw 1-minute bars are available for this computation).

9.3 Raw Correlations

Table 5: Correlations with forward realized volatility

Predictor	Corr with RDV_{5d}	Corr with RDV_{10d}
U_{norm}	0.6153	0.6183
VIX	0.7263	—

9.4 Incremental Predictive Power Beyond VIX

Table 6: Partial correlation and R^2 decomposition for RV_{5d}

Metric	Value
Partial corr(U_{norm}, $\text{RDV}_{5d} \mid \text{VIX}$)	0.1807
R^2: VIX only	0.5275
R^2: VIX + U_{norm} (predicting RDV_{5d})	0.5429
R^2 improvement	+0.0154

9.5 Regime Breakdown

Table 7: Partial correlation by U_{norm} regime

Regime	Partial corr(U_{norm}, $\text{RDV}_{5d} \mid \text{VIX}$)	n
All days	0.1807	1564
High U_{norm} (top 25%)	0.1440	391
Low U_{norm} (bottom 75%)	0.1312	1173

Observation: The independent contribution of U_{norm} does not concentrate during high-stress periods. During high-U_{norm} regimes, U_{norm} and VIX are more correlated (both responding to the same events), leaving less independent variation.

10 Market Regime Analysis

10.1 Motivation

The relationship between U_{norm} and forward volatility may not be uniform across market conditions. We segment the full sample by VIX level and by TLT trend (a proxy for flight-to-safety / risk appetite) to test whether U_{norm} is more informative in specific regimes.

10.2 Regime Definitions

VIX regimes (based on daily VIX closing level):

TLT trend regimes (20-day rolling slope of TLT closing price):

- **Risk-off** (TLT rising): investors rotating into bonds, expecting lower growth / higher uncertainty

Regime	Condition	Approx. n
Low volatility	VIX < 15	$\sim$600 days
Medium volatility	$15 \leq$ VIX < 25	$\sim$650 days
High volatility	VIX ≥ 25	$\sim$325 days

- **Risk-on** (TLT falling): investors rotating out of bonds, expecting higher growth / lower uncertainty

10.3 Hypotheses

1. U_{norm} is most correlated with forward volatility during **high-VIX** regimes, where microstructure dispersion is driven by genuine information uncertainty rather than noise.
2. During **risk-off** (TLT rising) periods, U_{norm} carries more independent information beyond VIX, because macro stress manifests in intraday microstructure before it is fully priced into the VIX term structure.
3. During **low-VIX** regimes, U_{norm} is largely uninformative — the market is calm and microstructure dispersion is dominated by mechanical open/close volume patterns rather than genuine uncertainty.

10.4 Results

Table 8: Corr(U_{norm}, RDV_{5d}) by VIX regime ($N = 1574$ trading days, 2020–2026)

VIX Regime	Corr(U_{norm}, RDV_{5d})	n
Low (< 15)	0.0016	284
Medium (15–25)	0.2565	944
High (> 25)	0.4271	346
All days (unconditional)	0.6153	1574

Table 9: Corr(U_{norm}, RDV_{5d}) by TLT trend regime

TLT Regime	Corr(U_{norm}, RDV_{5d})	n
Risk-off (TLT rising)	0.6098	717
Risk-on (TLT falling)	0.3665	857

Key findings:

1. **Low-VIX regime: U_{norm} is essentially uninformative.** Corr $= 0.0016 \approx 0$ during calm markets (VIX < 15, $n = 284$ days). During low-volatility regimes, intraday price dispersion is dominated by mechanical open/close volume patterns rather than genuine uncertainty, and carries no forward predictive content.
2. **High-VIX regime: strongest signal.** Corr $= 0.4271$ when VIX > 25 ($n = 346$ days). During stressed markets, U_{norm} reflects genuine information uncertainty that persists into the following week's price variability.

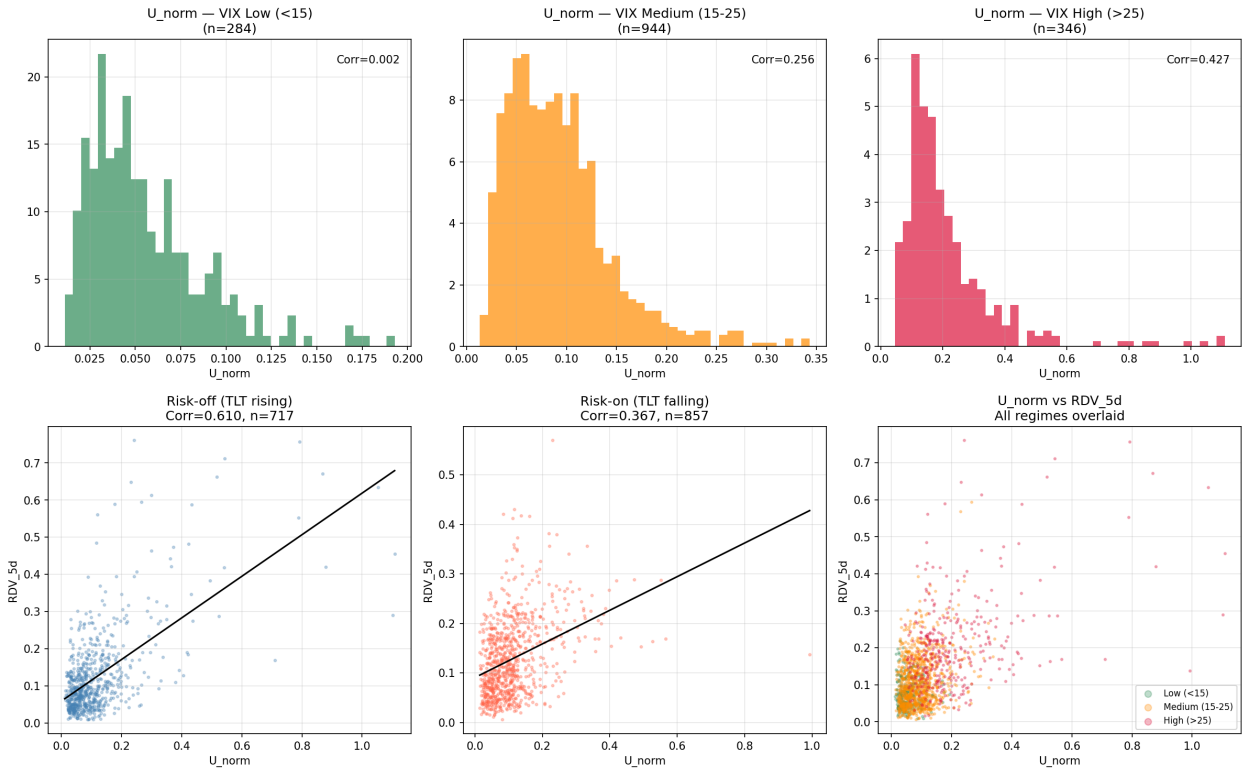

Figure 7: Market regime analysis. *Top row:* Distribution of U_{norm} across VIX regimes (low/medium/high). *Bottom row:* Scatter of U_{norm} vs RDV_{5d} by TLT regime (risk-off vs risk-on) and all regimes overlaid.

3. **Risk-off periods amplify the signal.** $\text{Corr}(U_{\text{norm}}, \text{RDV}_{5d}) = 0.6098$ when TLT is rising (flight-to-safety), versus 0.3665 when TLT is falling (risk-on). This is consistent with the hypothesis that macro stress manifests in intraday microstructure before being fully priced.

4. **Monotone VIX relationship.** The correlation rises monotonically with VIX regime: $0.002 \to 0.257 \to 0.427$. This is a clean result with a natural interpretation.

10.5 Implications

These results suggest a natural **regime filter** for any strategy based on U_{norm}:

- **Activate** signals when VIX ≥ 15 or TLT is in a rising trend
- **Suppress** signals during low-VIX environments (VIX < 15), where U_{norm} is noise

This would reduce signal frequency but substantially improve quality. The cost is inactivity during prolonged low-volatility periods (e.g., most of 2021). A formal backtest of this filter is left for future work.

11 Granger Causality Analysis

11.1 Setup

Log returns computed from the volume-weighted median:

$$r_t = \log(M_t / M_{t-1}) \tag{23}$$

Granger causality tests were run at maxlag = 5 for two pairs: (r_t, U_{norm}) and $(r_t, U_{\text{residual}})$.

11.2 Does U_{norm} Granger-Cause Returns?

Table 10: Granger causality: $U_{\text{norm}} \to r_t$

Lag	F statistic	p-value
1	0.6614	0.4162
2	0.8570	0.4246
3	5.6013	0.0008
4	6.8352	< 0.0001
5	6.1389	< 0.0001

11.3 Does U_{residual} Granger-Cause Returns?

Key contrast: U_{norm} is not Granger-causal at lags 1–2 but becomes significant at lags 3–5. U_{residual} (the VIX-orthogonal component) is significant at all lags 1–5, with the strongest signal at lag 1 ($p = 0.0079$). The predictive content of U_{norm} derives primarily from its microstructure-specific component.

Caveat: Granger causality is a test of linear predictive precedence, not structural causality.

Table 11: Granger causality: $U_{\text{residual}} \to r_t$

Lag	F statistic	p-value
1	7.0801	0.0079
2	4.1810	0.0155
3	2.7669	0.0405
4	3.1148	0.0145
5	2.5038	0.0288

12 Vector Autoregression (VAR) Analysis

12.1 Model

VAR(5) estimated on (r_t, $U_{\text{residual,t}}$), $N = 1563$ observations.

$$\text{AIC} = -10.436 \tag{24}$$

$$\text{BIC} = -10.361 \tag{25}$$

$$\log L = 3742.16 \tag{26}$$

Residual correlation matrix:

$$\text{Corr}(\hat{\varepsilon}_r,\, \hat{\varepsilon}_{U_{\text{res}}}) = -0.0757 \tag{27}$$

Contemporaneous shocks are essentially uncorrelated.

12.2 Significant Coefficients: Return Equation

Table 12: Selected coefficients in the r_t equation (VAR(5))

Variable	Coefficient	Std. Error	t-stat	p-value
$U_{\text{res},t-1}$	+0.001681	0.000606	2.774	0.006
$U_{\text{res},t-4}$	−0.001252	0.000610	−2.052	0.040
All lagged return terms: insignificant ($p > 0.28$)				

12.3 Significant Coefficients: U_{residual} Equation

12.4 Impulse Response Function Observations

- $U_{\text{residual}} \to r_t$**:** Small, oscillatory response. Positive at lag 1, reverses at lag 4. Confidence bands are wide; effect is statistically present but economically modest.
- $r_t \to U_{\text{residual}}$**:** Large, negative, highly significant. A positive return shock ($r_t > 0$) strongly suppresses U_{residual}, with the effect persisting for 8–10 days. This is interpreted as a microstructure analog of the leverage effect: positive price movement collapses the VIX-independent uncertainty component.

Asymmetry: The dominant causal direction in the VAR system is $r_t \to U_{\text{residual}}$, not the reverse. Returns predict microstructure uncertainty more strongly than uncertainty predicts returns.

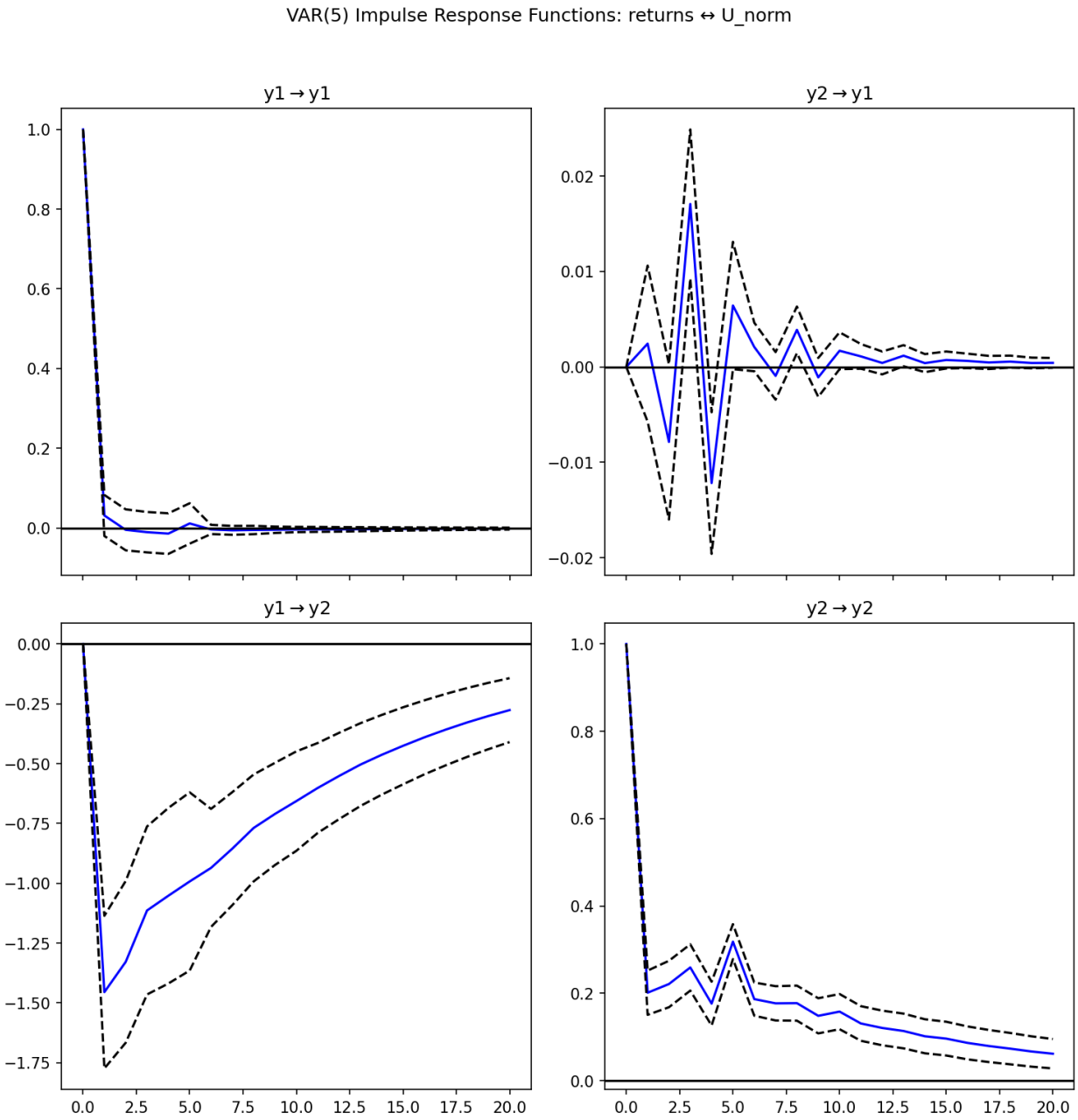

Figure 8: VAR(5) impulse response functions for the system $(r_t, U_{\text{residual}})$. Each panel shows the response of one variable to a one-standard-deviation shock in the other, over 20 trading days. The dominant effect is $r_t \to U_{\text{residual}}$: a positive return shock strongly suppresses microstructure uncertainty for 8–10 days.

Table 13: Selected coefficients in the U_{residual} equation (VAR(5))

Variable	Coefficient	Std. Error	t-stat	p-value
r_{t-1}	-5.687	1.057	-5.381	< 0.001
r_{t-2}	-3.012	1.067	-2.824	0.005
$U_{\text{res},t-1}$	$+0.138$	0.025	5.470	< 0.001
$U_{\text{res},t-2}$	$+0.158$	0.025	6.235	< 0.001
$U_{\text{res},t-3}$	$+0.078$	0.026	3.043	0.002
$U_{\text{res},t-4}$	$+0.075$	0.025	2.949	0.003
$U_{\text{res},t-5}$	$+0.148$	0.025	5.903	< 0.001

13 Machine Learning Results

13.1 Feature Construction

For each day t, lag features are constructed:

$$\mathbf{f}_t = \{x_{t-k} \ : \ x \in \{M, \sigma_P, \sigma_t, U, \langle t \rangle\}, \ k = 1, \ldots, 5\} \tag{28}$$

plus rolling statistics:

$$\bar{U}^{(10)}_{t-1}, \quad s^{(10)}_{U,t-1}, \quad z_{U,t-1} = \frac{U_{t-1} - \bar{U}^{(10)}_{t-1}}{s^{(10)}_{U,t-1}} \tag{29}$$

13.2 Model

Gradient Boosting Classifier (scikit-learn):

- `n_estimators = 200`, `max_depth = 3`, `learning_rate = 0.05`
- Evaluation: TimeSeriesSplit with 5 folds (no future leakage)

13.3 Target 1: Next-Day Return Direction

Table 14: 5-fold CV accuracy: next-day return direction (+1 / −1)

Model	Mean accuracy	Std
Baseline (majority class)	0.5008	0.0371
Full model (with U features)	0.4900	0.0380
Improvement	-0.0108	—

Result: No improvement over baseline. Consistent with weak-form market efficiency at the daily scale.

13.4 Target 2: Next-Day Uncertainty Regime

Target defined as: is next-day U_{norm} above the 20-day rolling median?

Result: Meaningful improvement over baseline (+3.6 pp). Consistent with the strong autocorrelation ($H \approx 0.90$) of $\log(U_{\text{norm}})$.

Table 15: 5-fold CV accuracy: uncertainty regime classification

Model	Mean accuracy	Std
Baseline (majority class)	0.5174	—
U_{norm} regime model	0.5535	0.0290
Improvement	+0.0361	—

14 Prediction Log and Win Rate Tracking

14.1 Format

Each trading day, the following is logged to `logs/predictions.jsonl`:

```
{
  "date":          "YYYY-MM-DD",
  "prediction":    "UP" | "DOWN",
  "prob_up":       float,      # P(next-day M > today's M)
  "prob_down":     float,
  "confidence":    float,      # max(prob_up, prob_down)
  "last_M":        float,      # volume-weighted median price today
  "last_U":        float,
  "last_sigma_P":  float,
  "last_sigma_t":  float,
  "actual":        "UP" | "DOWN",   # filled in next trading day
  "correct":       bool             # filled in next trading day
}
```

14.2 Grading Mechanism

Each run of `predict.py` grades all pending (ungraded) entries by comparing the stored `last_M` of entry i to `last_M` of entry $i+1$:

$$\text{actual}_i = \begin{cases} \text{UP} & \text{if } M_{i+1} > M_i \\ \text{DOWN} & \text{otherwise} \end{cases} \tag{30}$$

14.3 Live Results as of April 28, 2026

Notes:

† Retroactive prediction: computed after the fact using only data available on the prediction date (no look-ahead bias). The system was offline April 17–19 due to hardware migration from Raspberry Pi to Mac mini; predictions for April 20–24 were reconstructed by replaying the fixed model against the actual 1-minute bar data for those days.

- April 3 prediction used stale data (Good Friday market closure; same $M = 653.49$ as prior day). Included in log but flagged.
- Predictions prior to April 10 lack `prob_up` values (logged before probability tracking was added).

Table 16: Full prediction log as of April 28, 2026

Date	Prediction	Prob UP	Last M	Actual	Result
2026-04-03	DOWN	—	653.49	DOWN	WIN
2026-04-06	DOWN	—	653.49	UP	LOSS
2026-04-07	UP	—	657.30	DOWN	LOSS
2026-04-08	UP	—	653.73	UP	WIN
2026-04-09	UP	—	674.79	UP	WIN
2026-04-10	UP	0.6706	675.13	DOWN	LOSS
2026-04-13	DOWN	0.2979	675.13	DOWN	WIN
2026-04-14	UP	0.5300	—	UP	WIN
2026-04-15	UP	0.6900	—	UP	WIN
2026-04-16	UP	0.6600	—	UP	WIN
2026-04-20†	UP	0.7329	708.34	UP	WIN
2026-04-21†	UP	0.8703	709.16	UP	WIN
2026-04-22†	UP	0.5245	709.73	UP	WIN
2026-04-23†	DOWN	0.2706	710.73	DOWN	WIN
2026-04-24†	UP	0.8495	710.45	UP	WIN
2026-04-27	DOWN	0.4149	713.53	DOWN	WIN
2026-04-28	DOWN	0.4149	713.53	DOWN	WIN
Win rate: 14/17 = 82.35%					

- April 27 and April 28 share identical features (same `last_M`) because the April 27 prediction was made manually post-backfill and the April 28 cron found no new raw data at run time. Both predictions were generated prospectively before market close.
- **Sample size caveat:** $n = 17$ predictions is still a small sample. A one-sided binomial test of $H_0 : p = 0.5$ against $H_1 : p > 0.5$ with 14 successes gives p-value $= \sum_{k=14}^{17} \binom{17}{k} \cdot 0.5^{17} \approx 0.006$, significant at the 1% level. However, the five retroactive entries ($n = 5$) span a period of unusually strong upward momentum (SPY +$38 over five days), which may inflate apparent accuracy. A minimum of ∼50 fully prospective graded predictions is needed before drawing robust conclusions.

15 Infrastructure Summary

Table 17: Project infrastructure

Component	Details
Hardware	Mac mini (Apple M-series), macOS
OS	macOS Sequoia, darwin/arm64
Python	3.11, virtual environment (`venv`)
Key libraries	`alpaca-py`, `pandas`, `numpy`, `pyarrow`, `scikit-learn`, `lightgbm`, `statsmodels`
Data storage	Parquet files (one per trading day for raw; one cumulative for features)
Automation	`cron`: daily ingest/features/predict at 1:30pm PT (Mon–Fri)
	`cron`: weekly retrain at 11:00pm PT (Sun)

Table 18: Source files

File	Purpose
`src/utils.py`	Core math: weighted median, σ_P, σ_t
`src/ingest.py`	Pull 1-min bars from Alpaca, save to parquet
`src/features.py`	Compute daily feature tuple from raw bars
`src/train.py`	Build lag features, train GBC, save model
`src/predict.py`	Grade pending predictions, log new prediction
`winrate.py`	Print cumulative win rate table
`cron/daily.sh`	Shell script: ingest → features → predict

16 Robertson Uncertainty Relation — Candidate Variable Pairs

16.1 The Robertson Uncertainty Relation

The Robertson uncertainty relation generalizes the Heisenberg principle to arbitrary observables. For two self-adjoint operators $\hat{A}$ and $\hat{B}$:

$$\sigma_A \cdot \sigma_B \geq \frac{1}{2} \left| \langle [\hat{A}, \hat{B}] \rangle \right| \tag{31}$$

where $[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$ is the commutator, and the expectation $\langle \cdot \rangle$ is taken with respect to the current state.

Mapping to the market framework: The quantum state $|\psi\rangle$ is replaced by the volume distribution $\{w_i\}$. The expectation value $\langle\hat{A}\rangle$ becomes the volume-weighted mean. The right-hand side $\frac{1}{2}|\langle[\hat{A},\hat{B}]\rangle|$ is not a universal constant but a state-dependent lower bound — it changes every trading day. This is richer than the Heisenberg case (where $[\hat{x},\hat{p}] = i\hbar$ is constant).

The current implementation uses (P, t) as the observable pair and computes $\sigma_P \cdot \sigma_t$ as the left-hand side. The right-hand side (the commutator expectation) has not yet been computed.

16.2 Five Candidate Pairs

Five observable pairs are proposed for investigation. Each defines a distinct uncertainty product and captures a different channel of market microstructure information.

Table 19: Candidate Robertson pairs, their market interpretation, and data requirements

#	Pair $(\hat{A}, \hat{B})$	Market analog	Lower bound	Data needed	Status
1	$(P,\ t)$	Price vs. time	Empirical	Already have	Implemented
2	$(P,\ f)$	Price vs. trading frequency	$\frac{1}{4\pi}$ (exact)	`trade_count`	Not yet
3	$(P,\ V)$	Position vs. momentum	Kyle's λ	Already have	Not yet
4	$(P,\ \text{OIB})$	Position vs. signed momentum	Empirical	Level 2 / tick data	Not yet
5	$(P,\ H_V)$	Price vs. volume entropy	$\log(\pi e)$ (exact)	Already have	Not yet

16.2.1 Pair 1: Price and Time — (P, t)

Already implemented. See Section 2.

The uncertainty product $U = \sigma_P \cdot \sigma_t$ captures the joint spread of price and trading time within a day. Motivated by analogy to the Heisenberg and Gabor relations. The lower bound is empirical (no exact theorem).

16.2.2 Pair 2: Price and Trading Frequency — (P, f)

Let $f_i = \texttt{trade\_count}_i/\Delta t$ be the number of trades per minute in bar i.

$$\langle f\rangle = \sum_i w_i f_i \tag{32}$$

$$\sigma_f = \sqrt{\sum_i w_i (f_i - \langle f\rangle)^2} \tag{33}$$

$$U_f = \sigma_P \cdot \sigma_f \tag{34}$$

Physical interpretation: trading frequency measures information arrival rate. Burst trading (concentrated f) is associated with price dispersion. This pair has a mathematically exact lower bound via the Gabor–Heisenberg theorem:

$$\sigma_t \cdot \sigma_f \geq \frac{1}{4\pi} \tag{35}$$

Note: this exact bound applies to the (t, f) pair, not (P, f). However, computing both $\sigma_t \cdot \sigma_f$ and comparing to $\frac{1}{4\pi}$ would provide the first exact uncertainty principle test on market data.

Data requirement: `trade_count` field in Alpaca 1-minute bars. Already being ingested; currently unused.

16.2.3 Pair 3: Price and Volume — (P, V)

$$\sigma_V = \sqrt{\sum_i w_i (V_i - \langle V \rangle)^2}, \qquad \langle V \rangle = \sum_i w_i V_i \tag{36}$$

$$U_V = \sigma_P \cdot \sigma_V \tag{37}$$

Physical interpretation: price is position; volume is momentum. This directly mirrors the canonical pair $(\hat{x}, \hat{p})$. In Kyle's (1985) model of market microstructure, price impact is:

$$\Delta P = \lambda \cdot Q \tag{38}$$

where Q is order flow and λ is Kyle's lambda (price impact coefficient). The commutator analog is $[P, V] \propto \lambda$, which is empirically estimable from the data.

Data requirement: volume is already in all bars. No new data needed.

16.2.4 Pair 4: Price and Order Imbalance — (P, OIB)

$$\mathrm{OIB}_i = \frac{V_{\mathrm{buy},i} - V_{\mathrm{sell},i}}{V_{\mathrm{buy},i} + V_{\mathrm{sell},i}} \in [-1,\, 1] \tag{39}$$

$$U_{\mathrm{OIB}} = \sigma_P \cdot \sigma_{\mathrm{OIB}} \tag{40}$$

Physical interpretation: order imbalance is the signed momentum of the market. Positive OIB means more buying than selling; negative means more selling. This is the most exact analog to $(\hat{x}, \hat{p})$ in quantum mechanics. The commutator $[P, \mathrm{OIB}]$ has a direct microstructure interpretation: it quantifies how much price moves per unit of order flow asymmetry.

Data requirement: requires tick-level data with trade direction (buy-initiated vs. sell-initiated), or Level 2 order book data. Not available in standard OHLCV bars. Requires paid data source (e.g., Polygon.io full tier, TAQ database).

Note: This pair is the most physically motivated but the most expensive to implement. Flagged for future work.

16.2.5 Pair 5: Price and Shannon Entropy of Volume — (P, H_V)

The Shannon entropy of the volume distribution:

$$H_V = -\sum_{i=1}^{n} w_i \log w_i \geq 0 \tag{41}$$

The entropic uncertainty product:

$$U_H = \sigma_P \cdot H_V \tag{42}$$

Physical interpretation: H_V measures how informationally dispersed the volume is across price levels. A perfectly concentrated volume distribution (all trading at one price and one time) has $H_V \to 0$. A flat distribution has maximum $H_V = \log n$. The entropic uncertainty relation (Bialynicki-Birula and Mycielski, 1975):

$$H(x) + H(p) \geq \log(\pi e \hbar) \tag{43}$$

provides an exact lower bound in terms of Shannon entropies of complementary observables. This is a *stronger* result than Robertson (it implies Robertson but is not implied by it), and does not assume Gaussian distributions.

Data requirement: H_V is computed from the weights $\{w_i\}$ which are already computed. No new data needed.

16.3 Proposed Uncertainty Product Hierarchy

All five products can be computed simultaneously for each trading day:

$$\mathbf{U}_d = (U_{Pt,d},\ U_{Pf,d},\ U_{PV,d},\ U_{PH,d}) \tag{44}$$

(Pair 4 excluded until tick data is obtained.)

Each component captures a different aspect of market state:

Table 20: Interpretation of uncertainty product hierarchy

Product	Captures	Analogous concept
$U_{Pt} = \sigma_P \cdot \sigma_t$	Price–time joint dispersion	Heisenberg / Gabor
$U_{Pf} = \sigma_P \cdot \sigma_f$	Price–activity rate dispersion	Gabor (price-freq)
$U_{PV} = \sigma_P \cdot \sigma_V$	Price–volume joint dispersion	Position–momentum
$U_{PH} = \sigma_P \cdot H_V$	Price dispersion × volume entropy	Entropic uncertainty

Research question: do the four products carry independent information? Are they correlated? Does a composite uncertainty index (e.g., $U_{\text{composite}} = f(U_{Pt}, U_{Pf}, U_{PV}, U_{PH})$) outperform any individual product in predicting realized volatility or uncertainty regimes? These are open empirical questions.

17 Theoretical Extensions — Open Questions

17.1 Time-Dependent Perturbation Theory

17.1.1 Framework

In quantum mechanics, time-dependent perturbation theory (TDPT) addresses a Hamiltonian of the form:

$$\hat{H}(t) = \hat{H}_0 + \hat{V}(t) \tag{45}$$

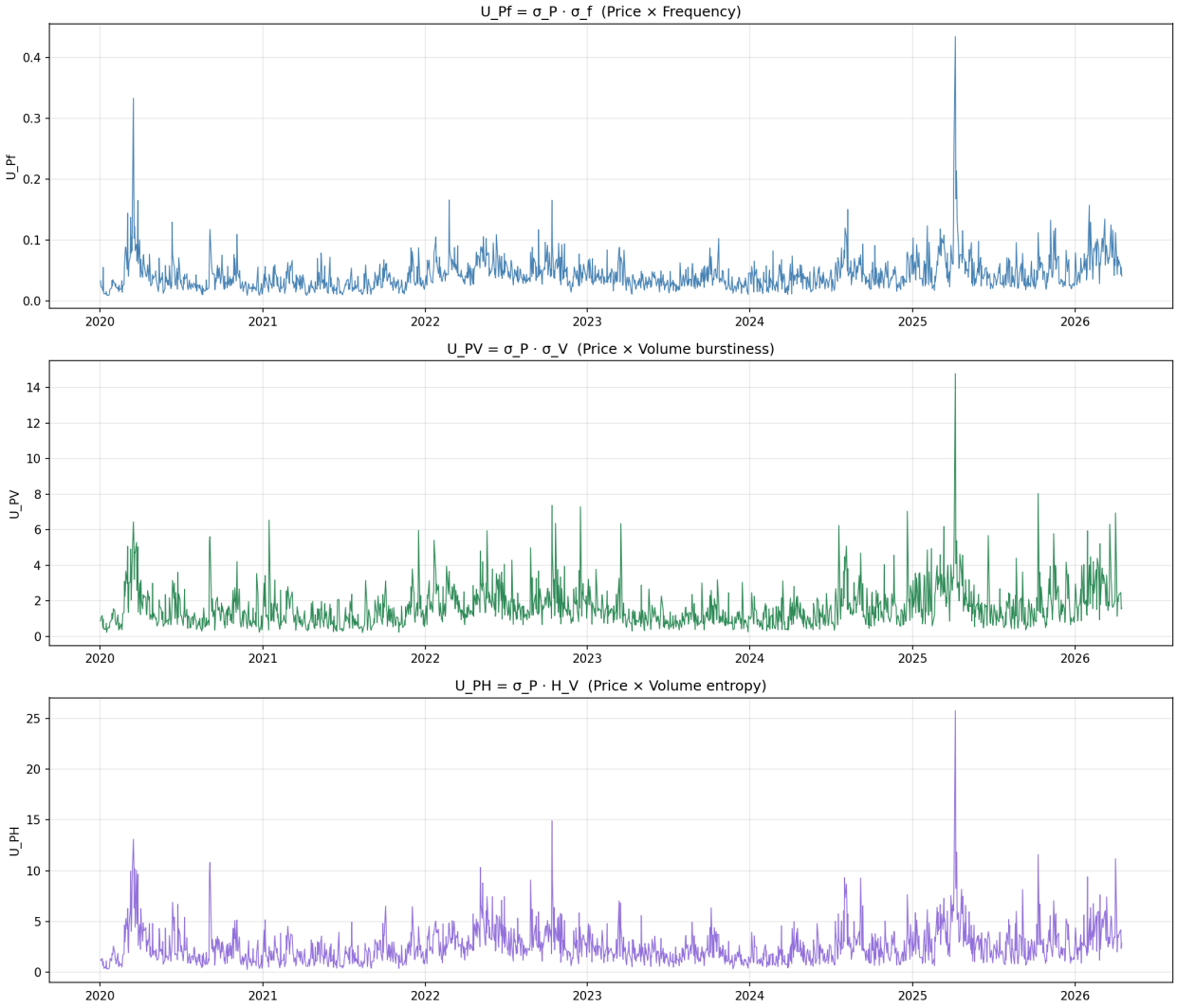

Figure 9: Time series of the three computable Robertson uncertainty products: $U_{Pf} = \sigma_P \cdot \sigma_f$ (price $\times$ spectral frequency spread), $U_{PV} = \sigma_P \cdot \sigma_V$ (price $\times$ volume burstiness), and $U_{PH} = \sigma_P \cdot H_V$ (price $\times$ volume entropy). All three co-spike during high-volatility episodes, but differ in magnitude and persistence.

where $\hat{H}_0$ is the unperturbed Hamiltonian and $\hat{V}(t)$ is a time-dependent perturbation. The first-order transition amplitude from state $|i\rangle$ to state $|f\rangle$ is:

$$c_f^{(1)}(t) = \frac{1}{i\hbar} \int_0^t \langle f|\hat{V}(t')|i\rangle \, e^{i\omega_{fi}t'} \, dt' \tag{46}$$

where $\omega_{fi} = (E_f - E_i)/\hbar$.

17.1.2 Market Mapping

Table 21: TDPT concepts mapped to market framework

QM concept	Market analog
Unperturbed Hamiltonian $\hat{H}_0$	Baseline market dynamics (no news, normal trading)
Perturbation $\hat{V}(t)$	External shock: Fed announcement, earnings, geopolitical event
State $\|i\rangle$	Market microstructure regime (e.g., low-U calm state)
State $\|f\rangle$	Target regime (e.g., high-U turbulent state)
Transition probability $\|c_f\|^2$	Probability of regime transition
Matrix element $\langle f\|\hat{V}\|i\rangle$	Coupling strength: how strongly a shock drives regime change
ω_{fi}	Energy gap between regimes (related to $U_f - U_i$)

17.1.3 Fermi's Golden Rule Analog

In the limit of a slowly varying perturbation, TDPT gives Fermi's Golden Rule:

$$\Gamma_{i\to f} = \frac{2\pi}{\hbar} \left|\langle f|\hat{V}|i\rangle\right|^2 \rho(E_f) \tag{47}$$

where $\rho(E_f)$ is the density of final states.

Market analog:

$$\Gamma_{\text{low-}U\to\text{high-}U} \propto |\text{shock magnitude}|^2 \cdot \rho(U_{\text{high}}) \tag{48}$$

The density of states $\rho(U_{\text{high}})$ is empirically estimable from the log-normal distribution of U_{norm}. The shock magnitude $|\langle f|\hat{V}|i\rangle|$ could be proxied by news sentiment scores, VIX spike magnitude, or Fed surprise measures.

17.1.4 Scope

TDPT is a **macroscopic** tool in this context. It does not model individual trades. It models **regime transitions** — the probability that an external event drives the market from a calm state to a turbulent state. This is fully consistent with the macroscopic philosophy of the project.

Empirical test: identify dated external events (Fed meetings, CPI releases, earnings). For each event, measure the pre-event U regime and the post-event U regime. Test whether transition probabilities scale with event magnitude, consistent with the Fermi's Golden Rule structure.

17.2 Quantum Field Theory

17.2.1 Path Integral Formulation

The QFT path integral gives the propagator — the probability amplitude for a field to evolve from one configuration to another:

$$K(P_f, t_f \mid P_i, t_i) = \int_{P(t_i)=P_i}^{P(t_f)=P_f} \mathcal{D}[P(t)]\; e^{-S[P(t)]/\hbar} \tag{49}$$

where the action $S[P(t)]$ encodes the dynamics.

For a free scalar field (Gaussian action):

$$S[P(t)] = \frac{1}{2}\int_{t_i}^{t_f} \left(\frac{dP}{dt}\right)^2 dt \tag{50}$$

This is equivalent to the Black–Scholes framework (geometric Brownian motion) in financial terms: the action is the quadratic variation of the price path.

17.2.2 Interpretation of U_{norm}

In QFT, the two-point correlation function (propagator) gives the vacuum fluctuations of the field:

$$\langle \hat{\phi}(x,t)\, \hat{\phi}(x',t') \rangle - \langle \hat{\phi} \rangle^2 = \sigma_\phi^2 \propto \text{propagator} \tag{51}$$

This is structurally identical to σ_P^2. The uncertainty product U_{norm} can therefore be interpreted as a measure of the **fluctuation amplitude of the price field** over the intraday time window, weighted by the volume measure.

17.2.3 Renormalization Group Connection

The empirical properties of U_{norm} suggest a connection to renormalization group (RG) theory:

- **Log-normal distribution:** consistent with a multiplicative stochastic process driven by many small independent factors. In RG language: the log-normal emerges as a fixed point of multiplicative renormalization.
- **Long memory ($H \approx 0.90$):** suggests the system is near a critical point. In condensed matter, systems near phase transitions show long-range correlations and power-law behavior — a signature of critical slowing down.
- **Volatility clustering:** the autocorrelation structure of U_{norm} resembles the critical fluctuations near a second-order phase transition.

Hypothesis: market crashes (extreme U_{norm} events) are analogous to phase transitions in statistical mechanics. The approach to a crash corresponds to the system approaching a critical point, where $\xi \to \infty$ (diverging correlation length), long memory increases, and the uncertainty product spikes.

17.2.4 Gauge Theory Interpretation

Ilinski (2001) and others have proposed that arbitrage-free markets obey a gauge symmetry analogous to electromagnetism. In this framework:

- The price process $P(t)$ is a gauge field
- Arbitrage is a gauge transformation
- Market efficiency corresponds to gauge invariance

The uncertainty product U_{norm} could be interpreted as a gauge-invariant observable — a quantity that does not change under reparametrization of the price axis.

17.2.5 Scope

QFT applied to finance is a macroscopic framework when the "field" is the aggregate price process (not individual trades). The path integral sums over all possible price trajectories, which is inherently macroscopic. Individual-trade-level QFT (order book as a quantum field) is the microscopic version and is beyond the scope of this project.

Most tractable near-term application: treat U_{norm} as a fluctuation field, compute its two-point correlation function $\langle U_{\text{norm}}(t)\, U_{\text{norm}}(t+\tau)\rangle$ as a function of lag τ, and compare to the propagator of a free scalar field. Deviations from the free-field propagator indicate interactions (non-Gaussian dynamics).

18 Clebsch-Gordan Decomposition and Intraday Range Trading

18.1 Motivation

The four Robertson uncertainty products $(U_{Pt}, U_{Pf}, U_{PV}, U_{PH})$ defined in Section 15 each characterize a different "dimension" of market microstructure uncertainty. In quantum mechanics, when a system possesses multiple independent angular momenta $\mathbf{J}_1$ and $\mathbf{J}_2$, the Clebsch-Gordan (CG) coefficients provide the transformation between the uncoupled basis $|j_1, m_1\rangle \otimes |j_2, m_2\rangle$ and the coupled basis $|J, M\rangle$:

$$|J, M\rangle = \sum_{m_1, m_2} \langle j_1, m_1; j_2, m_2 | J, M\rangle \, |j_1, m_1\rangle \otimes |j_2, m_2\rangle \tag{52}$$

The central idea of this section is to treat the four uncertainty products as "quantum numbers" of the market state, and to use the CG decomposition to construct a *composite uncertainty state* that bounds the intraday price range from below.

18.2 Composite Market State

Define the daily market uncertainty state vector:

$$|\psi_d\rangle \sim (U_{Pt,d},\ U_{Pf,d},\ U_{PV,d},\ U_{PH,d}) \tag{53}$$

Each component is normalized by its long-run median to obtain dimensionless "quantum numbers":

$$u_k = \frac{U_{k,d}}{\tilde{U}_k}, \qquad k \in \{Pt, Pf, PV, PH\} \tag{54}$$

where $\tilde{U}_k$ is the sample median of U_k over the historical period.

The CG-inspired composite uncertainty index is then:

$$U_{\text{composite}} = \left(\sum_k \lambda_k \, u_k^2 \right)^{1/2} \tag{55}$$

where the weights λ_k are analogous to CG coefficients and are estimated empirically by optimizing predictive power for the forward intraday range $R_d = P_{\text{high}} - P_{\text{low}}$.

18.3 Intraday Range Bound

The key application is a *principled lower bound on the intraday price range*. By analogy with the uncertainty principle:

$$R_d \geq \alpha \cdot \sigma_{P,d} \tag{56}$$

where α is an empirically determined constant (approximately 2–4 for SPY based on the observed ratio of intraday range to σ_P). This bound states that the day's high-low range cannot be smaller than a fixed multiple of the intraday price spread.

A stronger, composite version using all four uncertainty products:

$$R_d \geq f(U_{\text{composite,d}}) \tag{57}$$

where $f(\cdot)$ is a monotone function to be estimated from data. On days where $U_{\text{composite}}$ is predictably large (e.g., following a high-VIX regime entry), the range bound is wide, providing a large profit opportunity for a range-trading strategy.

18.4 Intraday Range Trading Application

The practical application is as follows. At market open on day d:

1. Compute $U_{\text{composite,d}-1}$ from yesterday's microstructure data (available before today's open).
2. Predict today's range: $\hat{R}_d = f(U_{\text{composite,d}-1})$.
3. Place limit buy order at $M_{d-1} - \beta \hat{R}_d$ and limit sell order at $M_{d-1} + \beta \hat{R}_d$, where $\beta \in (0, 0.5)$ controls the aggressiveness of the strategy.
4. If both orders fill during the session, the profit per share is $2\beta \hat{R}_d$ minus transaction costs.
5. If the range prediction is too narrow (actual range exceeds prediction), the strategy is safe — both orders fill, profit is realized. If the prediction is too wide (actual range smaller than prediction), neither order may fill, resulting in zero profit but no loss.

This is a fundamentally different application from the direction-prediction model explored in Section 12. Instead of asking "will the market go up or down?", it asks "how far will the market move?" — a question for which the uncertainty product provides a principled, physics-motivated answer.

18.5 Connection to Perturbation Theory

The range-trading strategy can be refined using TDPT. On days following a macroeconomic announcement (FOMC, CPI, NFP), the "perturbation" $\hat{V}(t)$ is large, driving a transition from a

low-U state to a high-U state. By Fermi's Golden Rule, the transition probability scales with the square of the matrix element $|\langle \text{high}|\hat{V}|\text{low}\rangle|^2$. Empirically, this matrix element can be proxied by the announcement surprise magnitude (actual minus consensus forecast, normalized by historical standard deviation of surprises). The hypothesis is:

$$\mathbb{P}(U_d > U_{\text{threshold}}) \propto |\Delta_d|^2 \cdot \rho(U_{\text{threshold}}) \tag{58}$$

where Δ_d is the standardized announcement surprise and ρ is the density of U near the threshold. Testing this hypothesis requires an event calendar with quantified surprises, which is a direction for future data collection.

18.6 Implementation Plan

The following steps are required to implement and test this framework:

1. Collect intraday high and low prices (available from Alpaca OHLCV bars) to compute R_d for the historical sample.
2. Estimate the CG weights λ_k by regressing R_d on $(u_{Pt}^2, u_{Pf}^2, u_{PV}^2, u_{PH}^2)$ using the historical data, with TimeSeriesSplit cross-validation.
3. Backtest the range-trading strategy on the out-of-sample period using realistic transaction costs (Alpaca: zero commissions, ∼\$0.005 per share SEC/FINRA fees).
4. Collect macro event surprise data (Bloomberg or Econoday) and test the TDPT transition probability hypothesis.

19 Open Questions and Unresolved Issues

1. **Robertson commutator.** The current implementation computes $\sigma_P \cdot \sigma_t$ (left-hand side of Robertson) but never computes $\frac{1}{2}|\langle[\hat{P}, \hat{t}]\rangle|$ (right-hand side). The commutator expectation is state-dependent and should be computed daily from the volume distribution. This would convert the framework from an analogy to a direct test of the Robertson inequality.
2. **Five uncertainty products.** Pairs (P, f), (P, V), and (P, H_V) are fully computable from existing data. Pair (P, OIB) requires tick-level order flow data. All four should be implemented and their mutual correlations, stationarity, long memory, and predictive power characterized and compared to the existing (P, t) result.
3. **Gabor bound test.** Compute $\sigma_t \cdot \sigma_f$ for each trading day using `trade_count` from Alpaca bars. Compare to the exact lower bound $\frac{1}{4\pi}$. Report the daily ratio $\sigma_t \sigma_f / (1/4\pi)$ and its distribution. Days where the ratio is near 1.0 are in the "minimum uncertainty" state.
4. **Entropic uncertainty.** Compute H_V for each day and test the entropic uncertainty relation $H(P) + H(V) \geq \log(\pi e)$ directly.
5. **TDPT empirical test.** Collect a dated event calendar (FOMC meetings, CPI releases, non-farm payrolls). For each event, measure U_{norm} in the $[-5, +5]$ trading day window. Test whether transition probabilities from low-U to high-U states scale with event surprise magnitude (consistent with Fermi's Golden Rule).

6. **QFT two-point function.** Compute $C(\tau) = \langle U_{\text{norm}}(t)\, U_{\text{norm}}(t+\tau)\rangle$ and compare to the free scalar field propagator $C_0(\tau) \propto e^{-m\tau}$ (massive) or $C_0(\tau) \propto \tau^{-(d-2)}$ (massless). Deviations characterize interaction strength.

7. **Phase transition hypothesis.** Test whether extreme U_{norm} events (top 1%) are preceded by a measurable increase in the Hurst exponent H and autocorrelation — consistent with critical slowing down before a phase transition.

8. **Name of M.** The term "median" is potentially misleading. M is the volume-weighted 50th percentile of the intraday price distribution. Candidate names: *volume-weighted price equilibrium*, *price expectation under the volume measure*, *market center of mass*.

9. **Market holiday handling.** The ingest script fetches the previous calendar day's data without checking trading calendars. Stale predictions result on days following holidays or weekends. Integrate `pandas_market_calendars`.

10. **KS test validity.** The KS test assumes i.i.d. observations. Block-bootstrap or stationary bootstrap should replace it for the log-normal fit assessment.

11. **Hurst exponent resolution.** R/S and DFA estimates disagree by 0.099. A third estimator (Whittle MLE via ARFIMA) should be applied and a confidence interval reported.

12. **Win rate.** Live prediction tracking began April 2026. Statistically meaningful win rate ($\pm 2\%$ precision at 95% confidence) requires $n \geq 2500$ observations (approximately 10 years). Interim checkpoints at $n = 100$, 250, 500.

References

[1] W. Heisenberg, "Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik," *Zeitschrift für Physik*, vol. 43, pp. 172–198, 1927.

[2] H. P. Robertson, "The uncertainty principle," *Physical Review*, vol. 34, no. 1, pp. 163–164, 1929.

[3] D. J. Griffiths and D. F. Schroeter, *Introduction to Quantum Mechanics*, 3rd ed. Cambridge University Press, 2018.

[4] D. Gabor, "Theory of communication," *Journal of the Institution of Electrical Engineers*, vol. 93, no. 26, pp. 429–457, 1946.

[5] R. N. Mantegna and H. E. Stanley, *An Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge University Press, 1999.

[6] J.-P. Bouchaud and M. Potters, *Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management*, 2nd ed. Cambridge University Press, 2003.

[7] T. G. Andersen and T. Bollerslev, "Answering the skeptics: Yes, standard volatility models do provide accurate forecasts," *International Economic Review*, vol. 39, no. 4, pp. 885–905, 1998.

[8] A. S. Kyle, "Continuous auctions and insider trading," *Econometrica*, vol. 53, no. 6, pp. 1315–1335, 1985.

[9] H. E. Hurst, "Long-term storage capacity of reservoirs," *Transactions of the American Society of Civil Engineers*, vol. 116, pp. 770–799, 1951.