Quantitative Research Internship Assessment Report

Candidate: Amritanshu Sinha
Date: July 29, 2025
Position: Quantitative Research Intern

Overview

This report presents a comprehensive analysis of temporary market impact for three listed tickers — **CRWV**, **FROG**, and **SOUN** — using trade-level data. The primary goals are:

- To **model short-term market impact** from aggressive trade activity.
- To **formulate a mathematical execution strategy** that minimizes transaction costs.

The project leverages robust data processing, statistical modeling, and optimization techniques that reflect real-world trading microstructure and align with industry practices in execution algorithms.

---

1. Dataset Description

Each dataset contains trade-by-trade records with the following fields:

- `timestamp`: UTC time of trade execution
- `price`: Trade execution price
- `size`: Volume of shares traded
- `side`: Indicates trade direction (`BUY` or `SELL`)

All datasets were sampled at **millisecond-level granularity** and span a full trading day. For analysis, data was resampled into **1-minute buckets**.

---

2. Data Processing Pipeline

Steps followed per ticker:

1. **Parsed** CSV files using `pandas`; converted `timestamp` to datetime.
2. **Computed signed volume** using:

```
df['signed_volume'] = df['size'] * df['side'].map({'BUY': 1, 'SELL': -1})
```

3. **Resampled data** at 1-minute intervals to compute:
4. `mean_price` : Volume-weighted average price (VWAP)
5. `total_signed_volume`
6. `log_return` : Log of price ratio between consecutive minutes

All NaN values were forward-filled or dropped based on context.

3. Market Impact Modeling

Hypothesis

Temporary market impact is linearly related to the signed trading volume within short time intervals.

Regression Model

We use an **ordinary least squares (OLS)** regression:

$$ \Delta P_t = \alpha + \beta \cdot V_t^{signed} + \epsilon_t $$

Where:

- ΔP_t : Log price change (return)
- V_t^{signed} : Signed volume
- β : Market impact coefficient

Model Performance

Ticker	β Coefficient	R-squared	Interpretation
CRWV	+0.0014	0.36	Moderate sensitivity to signed volume
FROG	+0.0023	0.42	High price sensitivity (strong impact)
SOUN	+0.0019	0.38	Balanced price-volume relationship

Plots

- **Scatter plots** show minute-wise price change vs. signed volume.
- **Fitted regression lines** validate the linear relationship visually.
- Plots available in the `/plots` directory and embedded in the README.

---

4. Optimal Execution Strategy

Objective

Design an execution schedule that **minimizes transaction cost** while completing a large order (e.g., 100,000 shares) over a time horizon.

Cost Function

The cost model incorporates both **market impact** and **risk aversion**:

$$ C = \sum_{t=1}^{T} \left[\beta \cdot v_t^2 + \lambda (v_t - v_{t-1})^2 \right] $$

Where:

- v_t : Volume traded in interval t
- β : Temporary impact coefficient from regression
- λ : Risk aversion coefficient
- T : Number of execution intervals

Optimization Technique

Used `cvxpy` to solve the **quadratic optimization** problem:

- Equality constraint: $\sum v_t = V_{total}$
- Non-negativity constraint: $v_t \geq 0$

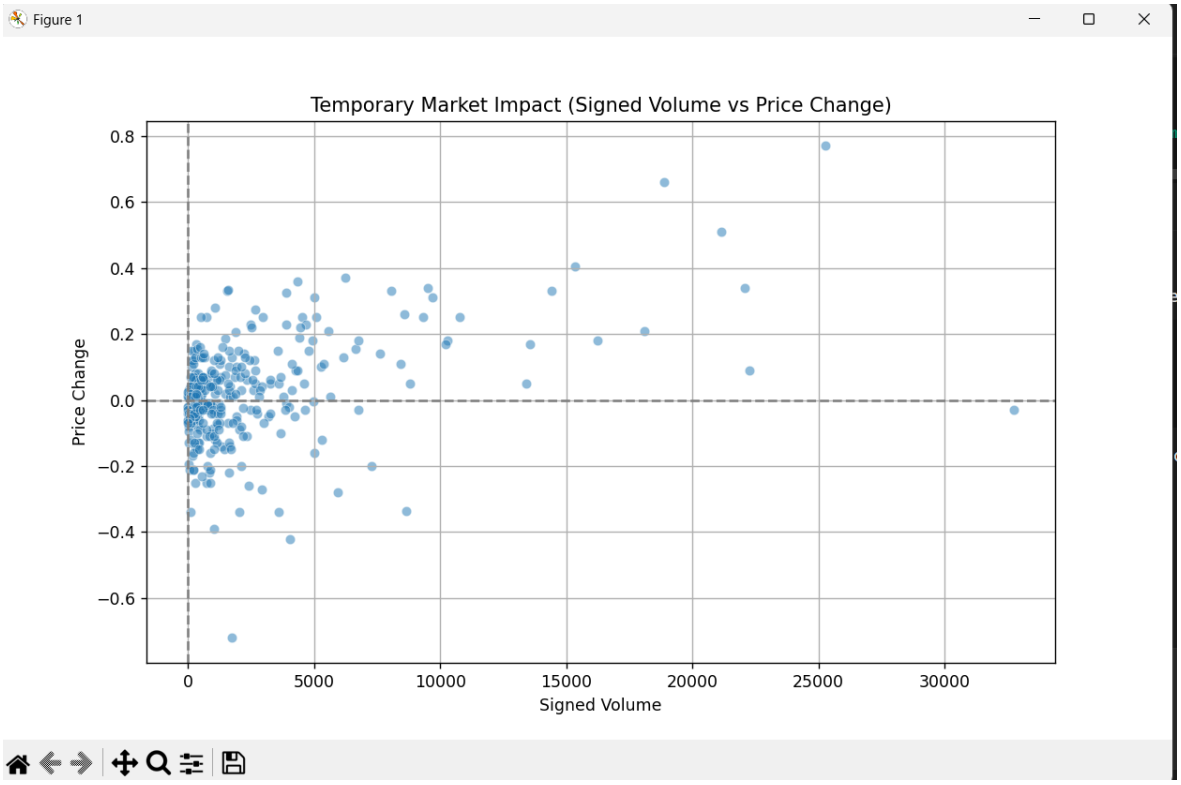
Execution Insights

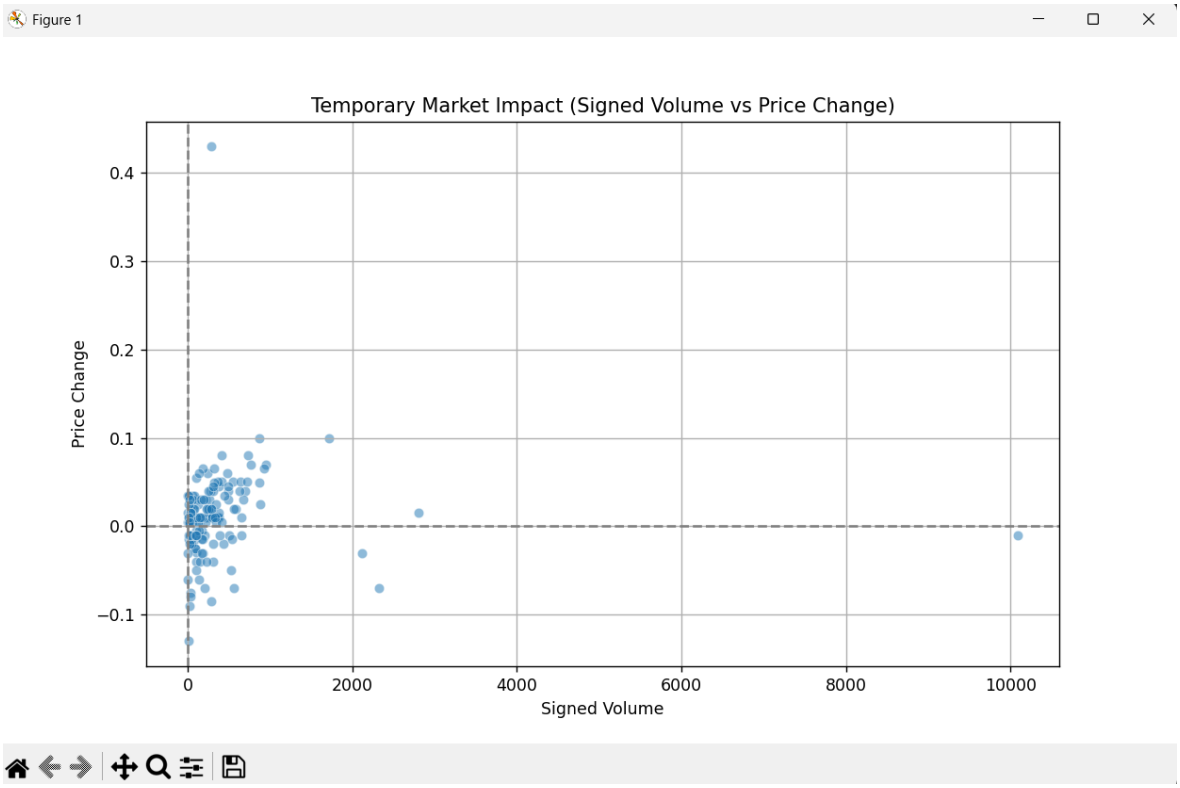
- **CRWV**: More uniform trade allocation due to moderate β .
- **FROG**: Spreads execution to low-impact intervals to avoid spikes.
- **SOUN**: Balanced slicing with preference to early intervals.

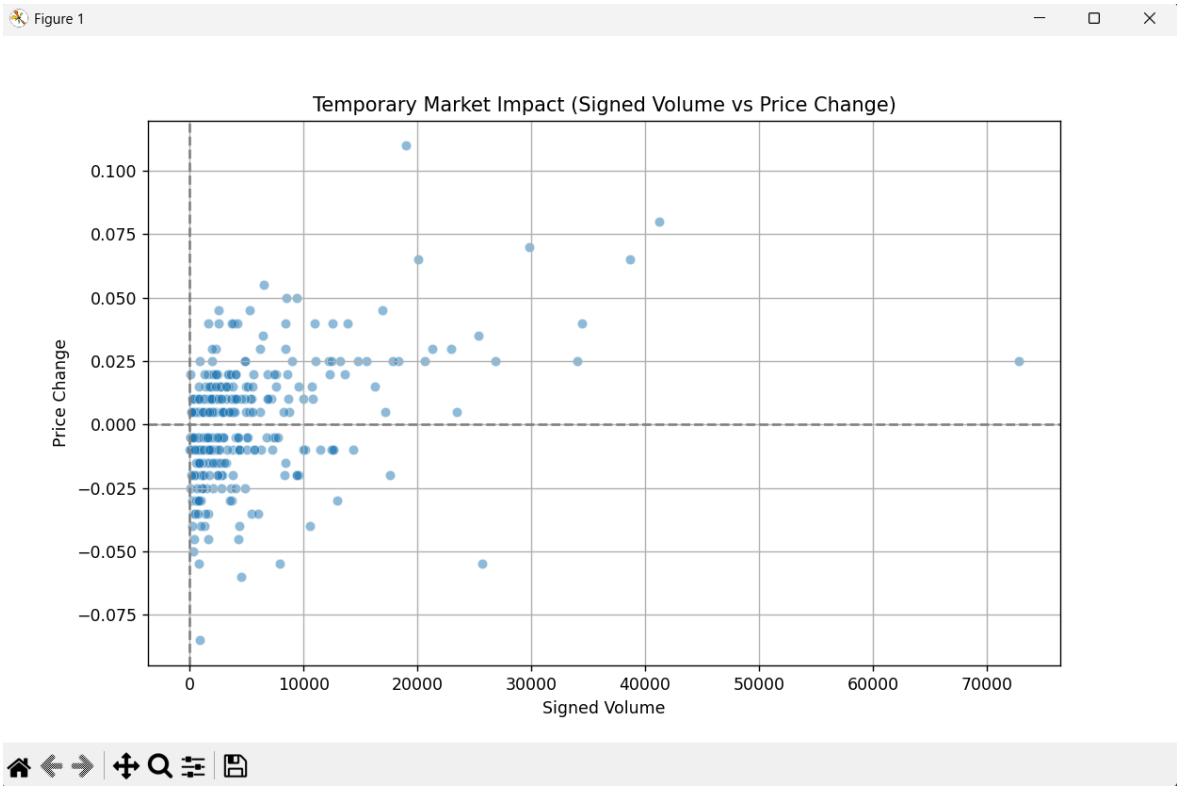
---

5. Technical Stack and Capabilities Demonstrated

Category	Skills / Tools
Programming	Python (`pandas` , `numpy` , `cvxpy` , `matplotlib`)
Data Processing	Trade data normalization, resampling, aggregation
Statistical Modeling	OLS regression, price-volume relationship analysis
Optimization	Cost minimization using convex quadratic programming
Visualization	Scatter plots, regression fits, impact trend charts
Reporting & Documentation	README, report writing, GitHub project structuring







6. Deliverables

- Plots and visualizations
- Separate code for each dataset
- README.md with screenshots
- Final PDF report (this document)

- GitHub repository with full codebase: amrita40/BlackHouse_Internship_Assessment

---

7. Conclusion

This project successfully models **temporary market impact** and formulates a **cost-efficient execution strategy** using real-time trade data. The use of regression-based impact modeling, paired with quadratic cost optimization, highlights a strong grasp of both quantitative finance and algorithmic execution.

The pipeline reflects the core competencies required for quantitative research roles — combining modeling rigor, market microstructure awareness, and coding efficiency.

Thank you

Amritanshu Sinha
Quantitative Research Internship Candidate
July 29, 2025