Insider Signals: To Trade or Not To Trade

Team: Bears. Bulls. Battlestar Galactica

Abhinav Chhabra, Ananya Bajaj

November 2025

Abstract

This report examines whether corporate insider transactions continue to convey predictive information about short- and medium-term stock returns in today's evolved regulatory, technological, and market environment. We utilize a comprehensive dataset of Form 3, 4, and 5 filings from 2006–2025, integrated with price data, Fama–French risk factors, trading volume metrics, and insider-specific characteristics.

We develop a systematic data pipeline to extract, clean, and classify insider transactions, removing noise from routine or pre-planned activity and organizing trades into economically interpretable behavioral clusters using unsupervised learning techniques. These clusters capture heterogeneity in insider seniority, transaction size, prior performance, and market attention.

To assess the informational content of insider activity, we estimate factor-adjusted regressions incorporating firm and time fixed effects, and construct both naïve and refined trading strategies based on rolling z-scores of net insider transactions and a weighted insider seniority index. Performance is evaluated through realistic backtesting procedures that minimize look-ahead bias.

Our results indicate that insider trading activity retains statistically significant predictive power at short horizons (7–30 days), particularly when aggregating buy and sell transactions into a unified signal and assigning higher importance to senior executives' actions. While the refined signal improves upon the naïve approach, practical profitability remains constrained by reporting delays, market adaptation, and the inherently low signal-to-noise ratio of insider data. These findings highlight both the continued relevance and the limitations of insider trading signals in increasingly efficient financial markets.

1 Introduction

Insider transactions - purchases or sales of company stock made by corporate officers, directors, and major shareholders - are often viewed as "informed" signals in financial markets. Insiders possess privileged knowledge about their firms' operations, strategic decisions, and future prospects, making their trading behaviour a potentially valuable source of information. A large body of prior research documents that insider purchases are typically followed by positive abnormal returns, while insider sales can predict short-term underperformance.

However, the informational environment surrounding insider trading has evolved significantly over the last decade:

- Regulatory scrutiny has intensified, particularly around Rule 10b5-1 trading plans.
- Many firms now encourage or require insiders to pre-schedule trades, reducing opportunistic timing.
- Faster information diffusion, algorithmic trading, and higher market efficiency may dilute the predictive power of traditional insider signals.

These changes raise a natural concern: whether insider trading still provides economically meaningful and statistically reliable signals in modern financial markets.

To address this, we re-evaluate insider trading activity using a comprehensive empirical and machine-learning–assisted framework. In particular, this study is structured around the following research questions:

1. Are insider trading signals still as strong and informative as they were in earlier decades?
2. Can a refined, data-driven trading strategy outperform naïve insider-based strategies?
3. How can the low signal-to-noise ratio of raw insider activity be mitigated to improve reliability and robustness?
4. Does a causal relationship exist between qualitative attributes of the stock, the insider, or the transaction and the resulting buy/sell signal?

By combining insider transaction data from 2006–2025 with market, volume, and firm-level metadata, this study provides a modern reassessment of the predictive power of insider trading under today's regulatory and technological conditions.

2 Motivation

2.1 Why This Problem Matters

- Insider trades represent real capital being deployed by informed corporate decision-makers.
- Prior literature finds insider activity to be a robust predictor of abnormal returns, especially when trades occur in clusters.
- While sparse, insider signals are often high-conviction compared to many noisy and overfit market indicators.

2.2 Why Study It Now?

- The informational content of insider trading may have changed due to tighter enforcement, the rise of Rule 10b5-1 trading plans, and faster incorporation of information into prices.
- It is unclear whether insider signals have weakened structurally or merely become more subtle and conditional.
- Advances in data availability and computation allow for a cleaner and more granular re-examination of insider behavior.

2.3 Historical Signal Strength

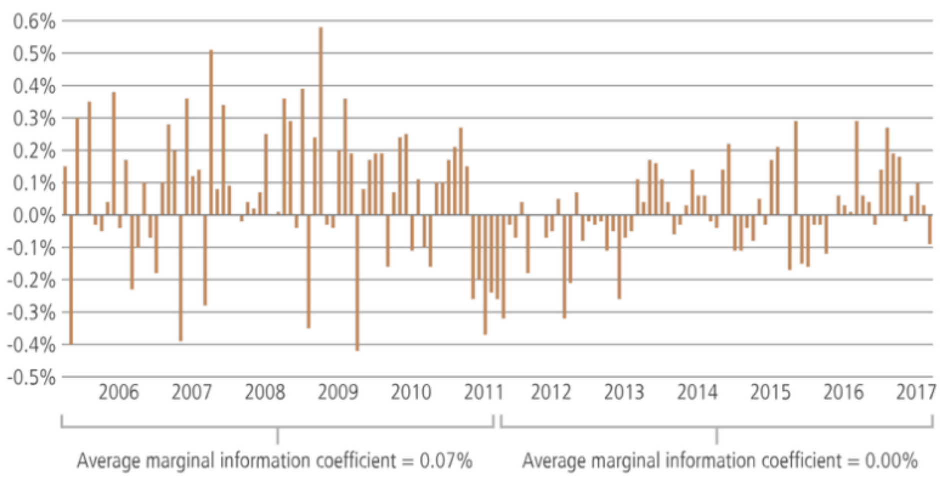

Figure 1: Marginal information coefficient from using the "Insider Boost", 2006–2017. Source: Bloomberg, S&P Capital IQ, Neuberger Berman. The chart shows the additional information coefficient gained from applying an "Insider Boost"-a binary adjustment that increases a stock's ranking by 30 percentage points when top insiders (excluding CEOs) purchase at least $100,000 worth of shares within the past 90 days and the firm is classified as high-value. The universe consists of the top 1,500 stocks by market capitalization in the Russell 3000 Index. Trades are executed seven days after the transaction date.

The figure above presents the marginal information coefficient from incorporating an "Insider Boost" metric over the period 2006-2017, as documented in Neuberger Berman's [Vukadinovic, 2019] report. The early years exhibit stronger predictive power (around 0.07%), which fades to near zero after 2012, highlighting the potential erosion of traditional insider signals and motivating a fresh empirical reassessment.

These results suggest that while insider activity historically improved predictability, its marginal contribution has largely vanished in recent years. This deterioration may stem from trading restrictions, non-informational motivations behind trades, or adaptive market behavior as the signal became widely known. Understanding whether modern, more structured approaches can recover useful information from insider activity is therefore an open and important research question.

2.4 Preliminary Analysis

2.5 Example of Variable Components

We created three primary variables. They provide contextual information that helps explain and structure insider behavior:

- **NET_TRANSACTION**: Net insider buying or selling activity. Although theoretically informative, its raw form shows minimal predictive power on its own.
- **relative_return_r_to_r5**: The stock's relative return over the previous 5 trading days, capturing short-term momentum or recent price pressure.
- **relative_return_r_to_r21**: The stock's relative return over the previous 21 trading days, reflecting short-term to medium-term performance trends.

To demonstrate the limitations of using raw insider activity alone as a trading signal, Table 1 presents a simplified correlation analysis between the above mentioned variables.

Table 1: Example of Variable Components

Variable	**Correlation**
NET_TRANSACTION	1.00
relative_return_r_to_r5	0.01
relative_return_r_to_r21	0.01

One striking observation is that while `NET_TRANSACTION` captures the direction and magnitude of insider activity, it exhibits very weak standalone correlation with short- and medium-term excess returns. This illustrates the poor signal-to-noise ratio of raw insider data when used in isolation.

This highlight that insider trades frequently occur under very different market conditions and are not directly aligned with recent price movements. This helps explain why simple, unadjusted insider trading strategies tend to perform poorly, and motivates our use of clustering and feature engineering to extract more structured and economically meaningful signals from the raw data.

3 Data Collection and Preparation

3.1 Sources

Our dataset is constructed from the following sources:

- **SEC Form 3/4/5 filings** from 2006–2025.
- **Daily stock prices, trading volumes, and risk factors** for return computation.
- **Firm characteristics**: sector, industry classification, and S&P 500 membership.

3.2 Cleaning and Filtering

To improve data quality and interpretability, we impose the following filters:

- Retain only pure buy and sell transaction codes: `P`, `M`, `S`, `D`.
- Remove trades executed under Rule 10b5-1 pre-planned trading programs.
- Exclude transactions with:
  - filing–transaction date gaps greater than 7 days,
  - missing share counts, transaction prices, or amended filings,
  - aggregate transaction values below $1,000,000 for a given firm on a given day.

3.3 Constructed Variables

For each transaction, we construct the following variables:

- **Net transaction value:** positive for net buys and negative for net sells.
- **Pre-trade returns:** 7-day and 30-day excess returns relative to the market.
- **Volume attention:** 7-day and 30-day trading volume z-scores.
- **60-day beta:** estimated using rolling regressions to capture market sensitivity.
- **Insider seniority counts:** number of transactions by Level 1 (executives), Level 2, and Level 3 insiders within a given filing.

4 Clustering Insider Behavior

Raw insider transaction data suffers from a very low signal-to-noise ratio: many trades are routine, liquidity-driven, or mechanically scheduled, and only a small subset reflects meaningful private information. To address this, we apply clustering to group transactions into economically interpretable behavioral patterns. The goal is not purely dimensionality reduction, but rather to identify structured insider behaviors that better correlate with subsequent stock performance and improve signal separability.

The following features are used for clustering:

- **Insider seniority:** counts of Level 1 (executives), Level 2, and Level 3 insiders.
- **Transaction magnitude:** net share or dollar value of insider trading (buy/sell intensity).
- **Market context:** 7-day and 30-day volume z-scores, and estimated beta.
- **Pre-trade performance:** 7-day and 30-day excess returns.

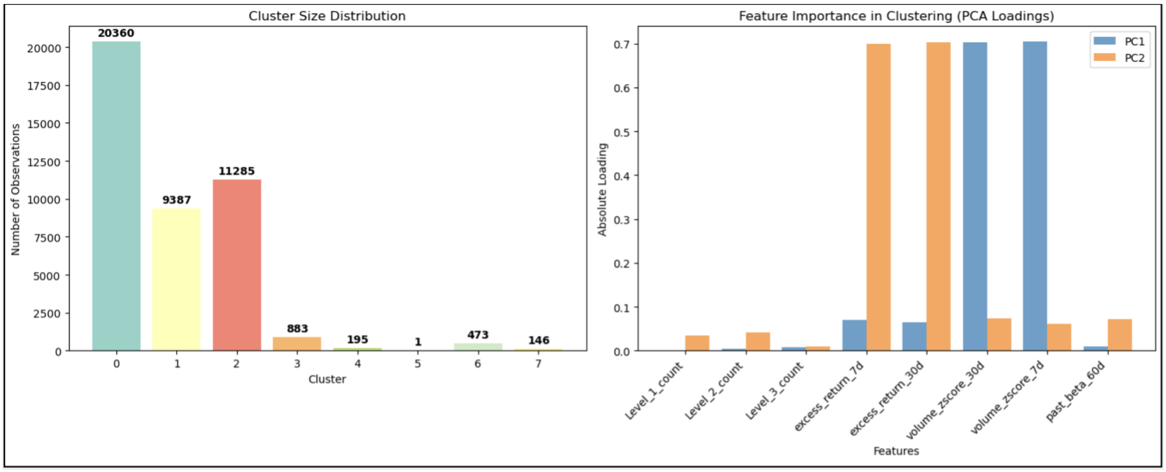

Figure 2: Insider Transaction Clustering based on Economic Variables

We initially generated eight clusters using KMeans clustering, but the raw clusters were not always economically interpretable. Therefore, we applied economic intuition to aggregate these into broader, conceptually meaningful groups. In particular, we focused on three dominant cluster archetypes:

1. **Buy-dominant clusters**: characterized by moderate to large positive net transactions, often concentrated among senior insiders, signaling confidence or perceived undervaluation.

2. **Sell-dominant clusters**: featuring large negative net transactions, commonly associated with insider caution, negative private information, or valuation concerns.

3. **Neutral clusters**: consisting of small, dispersed, or mixed trades that do not convey a clear directional signal and are more likely to represent routine or liquidity-driven activity.

This economically guided aggregation allows us to preserve the structure uncovered by the clustering algorithm while improving interpretability and relevance for trading strategy design. The resulting clusters significantly reduce noise and help isolate higher-conviction insider behaviors.

5 Causal Relationships

To assess whether insider activity contains incremental predictive information about future stock performance, we estimate a sequence of increasingly rich regression models. Each specification links forward excess returns r_{t+h} to our constructed insider signal (binary variable representing a net buy/sell by all executives of a company in a given time window) while progressively controlling for risk factors, firm heterogeneity, and time-specific shocks.

5.1 Basic Model

This baseline specification tests the raw association between insider activity and future returns:

$$r_{t+h} = \beta_0 + \beta_1 \cdot \text{InsiderSignal} + \varepsilon.$$

While simple, this model provides a useful first pass at detecting directional predictability before introducing additional controls.

5.2 Fama–French 3-Factor Model

To adjust for known sources of systematic return variation, we augment the regression with the market, size, and value factors:

$$r_{t+h} = \beta_0 + \beta_1 MKT + \beta_2 SMB + \beta_3 HML + \beta_4 \cdot \text{InsiderSignal} + \varepsilon.$$

A significant β_4 in this setting indicates that insider activity captures information beyond standard risk premia.

5.3 Fama-French and Firm Fixed Effects

We further incorporate firm fixed effects α_i, which absorb time-invariant differences across firms such as disclosure culture, insider compensation structure, and governance quality. This allows identification to come from within-firm variation in insider activity over time.

5.4 Fama-French, Firm, and Year Fixed Effects

Finally, adding year fixed effects γ_t controls for market-wide shocks, regulatory changes (e.g., shifts in enforcement of Rule 10b5-1), and macroeconomic conditions:

$$r_{t+h} = \ldots + \alpha_i + \gamma_t + \varepsilon.$$

This specification isolates the "pure" informational content of insider activity net of firm identity and year-specific noise.

5.5 Findings

Across all specifications, several patterns emerge:

- **Short-horizon excess returns (7–30 days) display statistically significant predictability.** This supports the hypothesis that markets incorporate insider information gradually rather than instantaneously.
- **Senior executives (Level 1 insiders) generate materially stronger signals.** Their trades are more closely tied to strategic decisions and private information, consistent with the literature on managerial information advantage.
- **Insider selling shows mild but non-negligible predictability.** Even after controlling for risk factors, sales retain a small negative coefficient, suggesting that at least a subset reflects adverse private information rather than liquidity needs alone.

- **Combining buys and sells into a unified net-transactions-based signal yields the highest explanatory power.** This aggregation helps mitigate noise from heterogeneous motives and reporting anomalies.

These findings collectively indicate that insider activity continues to contain economically meaningful information, provided one conditions appropriately on insider seniority and aggregates across transaction types.

Table 2: Summary of Significant Regression Results (7–30 Day Horizons)

Model / Variable	**Horizon**	**Coeff.**	**P-value**
Short-Horizon Predictability			
FF3_FE (Neutral Data)	7 Day	0.00058	0.0211
FF3_FE (Level 1 Executives)	7 Day	0.00101	0.0037
FF3_FE_Time (All Buys & Sells)	7 Day	0.00729	0.00060
Level 1 Executives (Stronger Signals)			
FF3_FE (Level 1 Exec Data)	7 Day	0.00101	0.0037
FF3_FE_Time (Level 1 Exec Data)	7 Day	0.00111	0.00160
Insider Selling (Mild Predictability)			
Basic (Sell Data)	30 Day	-0.00284	0.00030
FF3 (Sell Data)	30 Day	-0.00160	0.0170
Combined Buy + Sell Signal (Stronger Predictive Power)			
FF3_FE (All Buys & Sells)	7 Day	0.00065	0.00180
FF3_FE_Time (All Buys & Sells)	7 Day	0.00729	0.00060
FF3_FE (All Buys & Sells)	30 Day	0.00086	0.0465
FF3_FE_Time (All Buys & Sells)	30 Day	0.00093	0.0319

6 Signal Construction

6.1 Naïve Insider Signal

Our baseline signal relies solely on the intensity of insider trading. For each stock, we compute a 52-week rolling z-score of net insider transaction value:

$$z_i = \frac{x_i - \mu_i}{\sigma_i},$$

where x_i is the net dollar value of insider trades for stock i during the formation window, and (μ_i, σ_i) are the historical mean and standard deviation over the prior 52 weeks. This normalization allows comparison across firms with different trading frequencies and firm sizes.

Portfolio formation rules:

- Go **long** the top 10% of stocks (largest positive z-scores: concentrated insider buying).
- Go **short** the bottom 10% of stocks (most negative z-scores: concentrated insider selling).

While intuitive, this naïve signal treats all insiders equally and does not differentiate between routine trades, liquidity-driven sales, and high-conviction actions by senior executives.

6.2 Refined Insider Score (Public Score)

To address the heterogeneity among insiders, we construct a composite score that integrates both trading intensity and insider seniority. Since executives at higher organizational levels are more likely to possess private information about strategic decisions, their trades arguably carry greater informational value.

The refined score is defined as:

$$\text{PublicScore} = z(\text{NetTransaction}) + z(0.5\,L_1) + z(0.3\,L_2) + z(0.2\,L_3),$$

where L_1, L_2, and L_3 represent the counts of trades by senior executives, mid-level insiders, and general insiders, respectively.

Rationale for Weights. The weights (0.5, 0.3, 0.2) reflect the empirical intuition that:

- Level 1 insiders (executives) have the strongest strategic insight and historically produce the most informative trades.
- Level 2 insiders (mid-level officers) may react to operational information but with less precision.
- Level 3 insiders (general employees/ownership filers) provide diffuse and noisier signals.

These weights also align with patterns observed in our clustering analysis: clusters dominated by Level 1 insiders showed the clearest directional signals.

The trading strategy mirrors that of the naïve approach, but ranks stocks using the refined PublicScore, effectively giving more weight to high-conviction insider clusters.

7 Backtesting Framework

7.1 Formation and Holding Periods

Our backtesting procedure is designed to mimic a realistic implementation environment with weekly refresh cycles:

- Portfolios are **formed every Monday**.
- Only filings **publicly available during the previous week** are used to compute the signal.
- Positions are held for either:
  - **1 week**: captures high-frequency informational leakage, or
  - **1 month**: reduces turnover and emphasizes medium-term drift.

7.2 Avoiding Look-Ahead Bias

A critical component of the backtesting design is ensuring that trading decisions only use information that would have been available in real time. To prevent inadvertent look-ahead bias, we impose the following timing rules:

- All signals for week t are computed using data available *only up to the end of week t*; no future filings or returns are incorporated.
- Portfolios formed on the basis of these week-t signals are implemented at the **open of week $t+1$**.

This structure mirrors the operational reality that Form 4 insider filings become actionable only after they are disclosed, parsed, and incorporated into the dataset. By lagging trades one full week, we eliminate any artificial inflation of strategy performance caused by inaccessible information.

8 Backtesting Results

8.1 Performance Summary

Across both weekly and monthly rebalancing frequencies, several consistent patterns emerge:

- The naïve strategy yields small but generally positive average excess returns, albeit with unstable drawdowns and low Sharpe ratios.
- The refined Public Score strategy delivers higher Sharpe ratios, lower drawdowns, and more stable cumulative return trajectories.
- Weekly rebalancing captures short-lived insider-related price reactions but exhibits higher volatility.
- Monthly rebalancing smooths portfolio paths and reduces noise at the cost of slower signal responsiveness.

8.2 Interpretation

Although insider activity retains statistically significant short-horizon predictive power, it is not sufficiently strong to generate large standalone excess returns in a modern and efficient market setting. Noise in filings, reporting delays, and heterogeneous insider motives dampen raw signal strength. These results suggest that insider signals function best as a complementary factor—particularly when combined with momentum, valuation, or volatility measures—rather than as a standalone strategy.

8.3 Weekly Strategy Performance

Figure 3 presents the weekly (5-day horizon) strategy results constructed with proper avoidance of look-ahead bias. For comparison, Figure 4 reports performance using the same methodology but without bias correction, illustrating the sensitivity of insider-based strategies to timing assumptions.

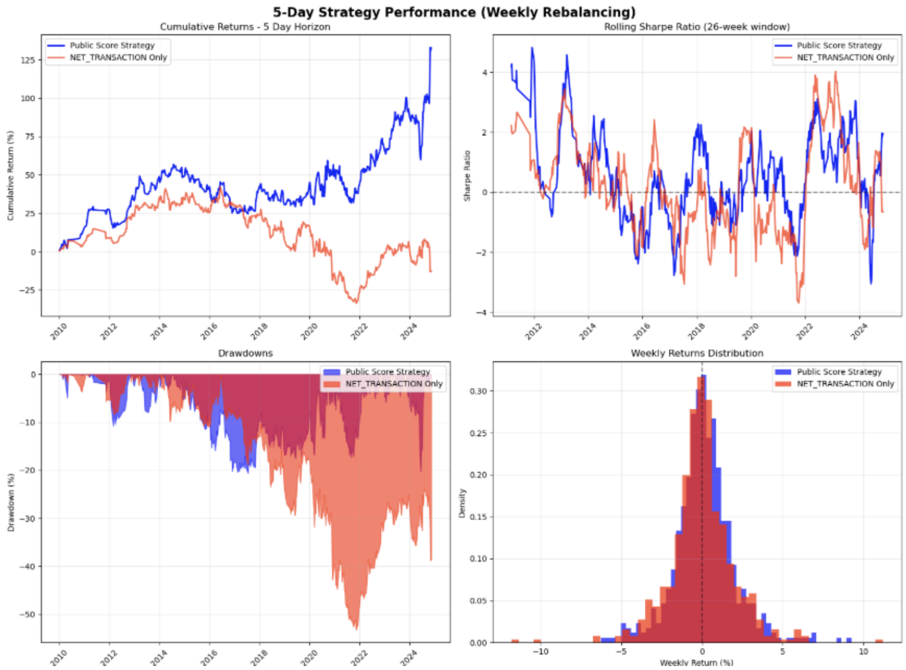

Figure 3: Weekly Strategy Performance (With Bias)

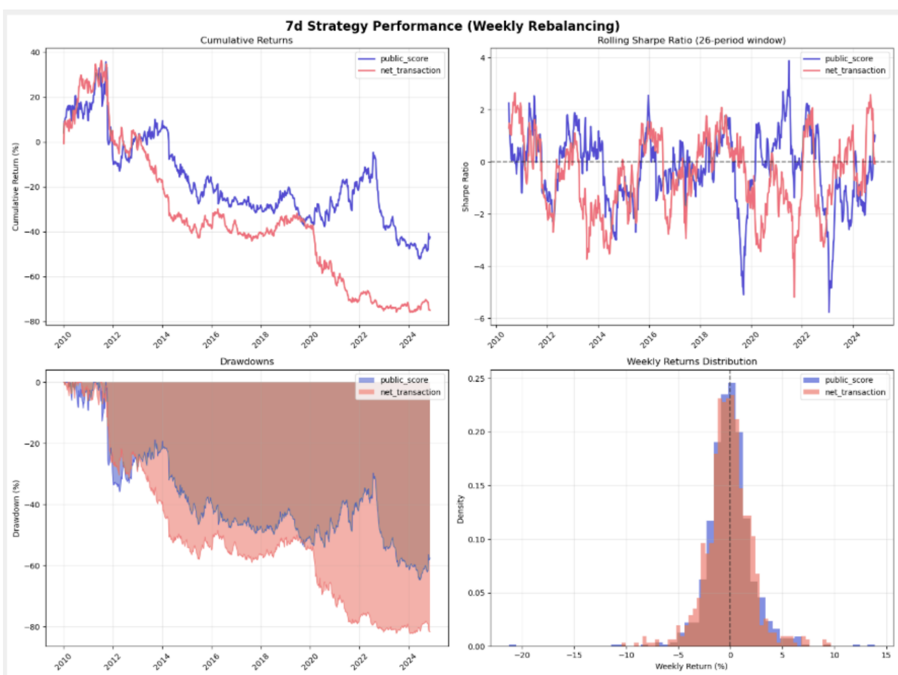

Figure 4: Weekly Strategy Performance (Without Bias)

8.4 Monthly (21-Day) Strategy Performance

Figure 5 displays the monthly (21-day horizon) strategy results using strict bias-free signal formation. Figure 6 provides the corresponding performance without bias handling, again highlighting the degree to which insider-based strategies may appear overstated when timing constraints are not properly enforced.

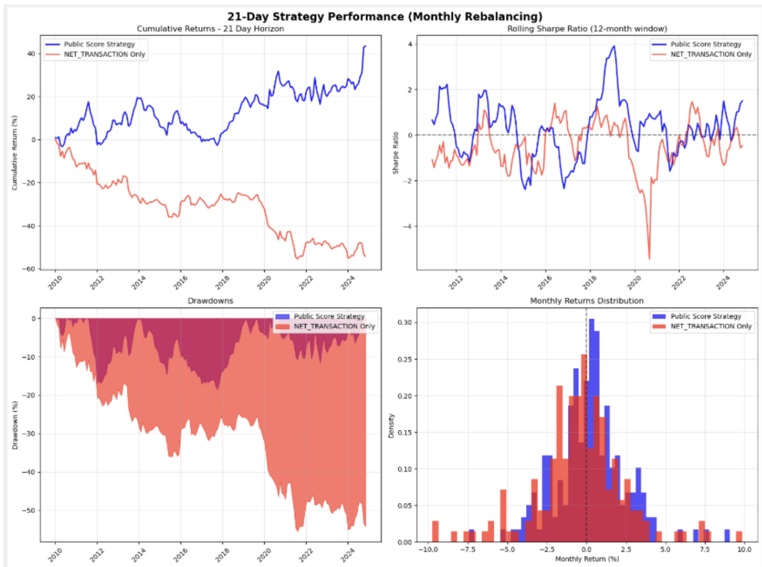

Figure 5: Monthly Strategy Performance (Without Bias)

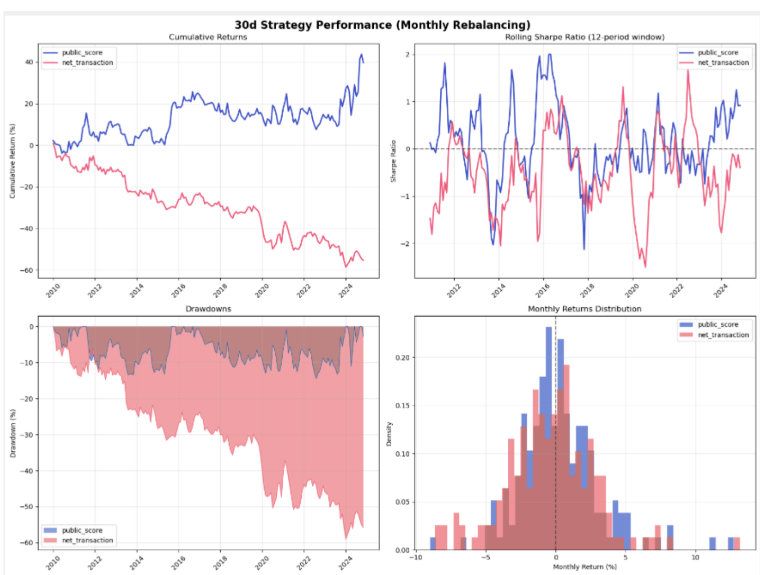

Figure 6: Monthly Strategy Performance (Without Bias)

9 Conclusion

This study provides a contemporary reassessment of the informational content embedded in insider trading activity, using a comprehensive dataset from 2006-2025 and a methodology that combines econometric analysis, clustering, and systematic portfolio construction. Our results demonstrate that insider transactions continue to convey measurable predictive power, but only under specific conditions and over short investment horizons.

First, we find that insider activity is most informative when aggregated across all filings within a period, rather than examining isolated trades. Second, trades executed by senior executives consistently exhibit stronger predictive relationships with future returns, reflecting their closer proximity to strategic decision-making and material non-public information. Third, the use of clustering techniques and contextual metadata substantially improves signal quality by filtering out routine, scheduled, or noise-driven trades. Finally, portfolio strategies based on our refined Public Score outperform naïve net-transaction approaches, yielding higher Sharpe ratios and shallower drawdowns, even after adjusting for potential look-ahead bias.

Taken together, these findings suggest that insider trading data remains economically relevant in modern markets, but extracting useful information requires careful preprocessing, contextualization, and aggregation. The signal is subtle, heterogeneous, and easily overwhelmed by noise - reinforcing that insider activity should be viewed as a complementary input rather than a standalone alpha source.

10 Future Work

Several avenues offer promising extensions to this research:

- **Expanding the universe:** Apply the methodology to small-cap firms, where information asymmetry is greater and insider signals may be stronger, as well as to non-U.S. markets with different disclosure rules.

- **Coordinated insider behaviour :** Develop models to identify clusters of insiders trading simultaneously within a firm, which may indicate higher conviction or collective information sharing.

- **Repeated trading patterns:** Investigate insider "track records," studying whether certain insiders historically demonstrate superior timing skill and whether their trades should be weighted more heavily.

- **Integration with machine learning:** Combine insider-derived features with predictive models (e.g., gradient boosting, sequence models, or latent factor embeddings) to construct multi-factor return forecasts.

- **Event-context modeling:** Explore how the informativeness of insider trades varies around earnings announcements, M&A activity, macro policy changes, or periods of elevated volatility.

- **Transaction-level microstructure effects:** Examine whether trade size relative to insider wealth or firm liquidity contains incremental predictive power.

By extending the analysis along these dimensions, future research may uncover deeper structure in insider trading behavior and potentially amplify the economic value of insider-based signals in multi-factor portfolio systems.

References

[Chan et al., 2012] Chan, K., Ikenberry, D. L., Lee, I., and Wang, Y. (2012). Informed traders: Linking legal insider trading and share repurchases. *Financial Analysts Journal*, 68(1). Available at SSRN: `https://ssrn.com/abstract=1990500`.

[Cosemans and Frehen, 2025] Cosemans, M. and Frehen, R. (2025). Strategic insider trading and its consequences for outsiders: Evidence from the eighteenth century. *Journal of Financial Economics*, 164:103974.

[Lakonishok and Lee, 2001] Lakonishok, J. and Lee, I. (2001). Are insider trades informative? *The Review of Financial Studies*, 14(1):79–111.

[Tikr, 2025] Tikr (2025). How to track insider trading in stocks. `https://www.tikr.com/blog/how-to-track-insider-trading-in-stocks`.

[University of Michigan Ross School of Business, 2024] University of Michigan Ross School of Business (2024). New study on insider trading discovers flaws in oversight and regulation. `https://michiganross.umich.edu/news/new-study-insider-trading-discovers-flaws-oversight-and-regulation`.

[U.S. Securities and Exchange Commission (SEC), 2025] U.S. Securities and Exchange Commission (SEC) (2025). Forms 3, 4, and 5. `https://www.sec.gov/files/forms-3-4-5.pdf`. Accessed November 22, 2025.

[Vukadinovic, 2019] Vukadinovic, A. (2019). Systematically speaking: Insider boost to a value-based stock assessment. `https://www.nb.com/en/global/insights/systematically-speaking-may-2019`.

[Wang et al., 2011] Wang, W., Shin, Y.-C., and Francis, B. B. (2011). Are CFOs' trades more informative than CEOs' trades? *Journal of Financial and Quantitative Analysis (JFQA)*. Forthcoming, Available at SSRN: `https://ssrn.com/abstract=1787482`.