Does Congress Play by Different Market Rules? A Comparative Data Mining Analysis of Congressional Stock Trading

Riyaa Randhawa*, Kent Lemken†, and Felipe Cardozo†
*Computer Science, Emory University, ID: 2589424, riyaa.randhawa@emory.edu
†Finance & Computer Science, Emory University
Kent Lemken, ID: 2583117, kent.lemken@emory.edu
Felipe Cardozo, ID: 2696984, felipe.cardozo.sr@emory.edu

***Abstract*—Members of the U.S. Congress are required by the Stop Trading on Congressional Knowledge (STOCK) Act to disclose personal stock transactions within 45 days of execution. Yet whether their trades systematically outperform the market, and whether that outperformance is detectable through observable features, remains empirically contested. This paper presents two complementary machine learning approaches to anomalous congressional trading detection: (1) a *CAR-based outcome model* incorporating cumulative abnormal returns, government contract proximity, and lobbying activity (18 features, 42,515 stock trades, Receiver Operating Characteristic Area Under the Curve (ROC-AUC) of 0.84), and (2) a *profile-based behavioral model* emphasizing politician demographics and filing behavior (10 features, 15,457 labeled trades, ROC-AUC = 0.9949). Drawing on 111,149 publicly disclosed stock trades spanning 2012–2026 across 414 members of Congress, we train XGBoost binary classifiers augmented by Bayesian hierarchical probability layers, producing calibrated suspicion scores with uncertainty quantification. The CAR-based model reveals that congressional purchases systematically precede positive abnormal returns (mean excess return $+2.1\%$ over the trade-date window; permutation-tested mean CAR $+1.18\%$, $p = 0.003$), with a post-STOCK Act decline suggesting partial deterrence. The profile-based model identifies geographic and chamber-level clustering as the strongest predictors of suspicion, with Senate trades 28% more likely to receive elevated suspicion scores than House trades. Both models rank all seven known insider traders and investigated politicians within their top suspicion percentiles, providing cross-validation for both methodologies.**

***Index Terms*—congressional trading, insider trading detection, cumulative abnormal returns, XGBoost, STOCK Act, Bayesian hierarchical model, political finance, anomaly detection**

I. Introduction

THE question of whether members of the United States Congress systematically profit from non-public information in their personal stock portfolios has attracted substantial academic, journalistic, and regulatory attention. Unlike typical market participants, elected officials routinely receive classified intelligence briefings, participate in drafting legislation with direct market implications, and oversee regulatory agencies whose actions affect individual firms materially. Their personal financial disclosures are public record, creating an unusual opportunity to study informationally-sensitive trading at scale.

The Stop Trading on Congressional Knowledge (STOCK) Act of 2012 made it explicitly illegal for members of Congress to trade on material non-public information obtained through official duties, and mandated electronic disclosure of individual trades within 45 days. Despite this legal framework, enforcement actions remain sparse and compliance with filing deadlines is inconsistent. High-profile cases—including the criminal convictions of Representative Chris Collins (2019) and lobbyist-turned-Representative Steve Buyer (2024), as well as Senate investigations into COVID-era trades by Senators Burr, Loeffler, and Perdue—illustrate that the risk persists.

This paper makes a distinctive contribution by implementing and comparing two fundamentally different machine learning approaches to the same detection problem. An *outcome-centric model* treats post-trade abnormal returns as the core signal, while a *behavioral-centric model* treats trading profiles and politician demographics as predictors of suspicion. Both approaches are validated against the same ground truth (known insider trading cases), yet they yield markedly different feature importance rankings and practical insights. By presenting both methodologies side-by-side, we illuminate how anomaly detection frameworks depend critically on feature engineering choices and provide guidance for practitioners designing detection systems in regulated financial contexts.

II. Related Work

A. Abnormal Returns in Congressional Portfolios

The foundational work by Ziobrowski et al. [2], [3] documented statistically significant abnormal returns to U.S. Senator and House member stock portfolios, finding average annual abnormal returns of approximately 12% for senators and 6% for representatives. Eggers and Hainmueller [4] later challenged these findings, arguing that controlling for look-ahead bias reduced detectable alpha substantially. Karadas [5] revisited the post-STOCK Act period, finding declining abnormal returns after 2012 consistent with a partial deterrence effect, though positive alpha persisted.

B. Event Study Methodology

The event study framework, introduced by Fama et al. [6], provides the statistical foundation for measuring abnormal price movements around a specific event date. Cumulative Abnormal Returns (CARs) over specified post-event windows remain the dominant outcome metric in the literature. In the machine learning context, tree ensemble methods like XGBoost [9] have been successfully applied to financial anomaly detection [7], [8] where tabular feature spaces and non-linear interactions dominate. The SHAP (SHapley Additive exPlanations)framework [10] enables model-agnostic attribution of predictions to individual features, providing interpretability critical for regulatory applications.

C. Bayesian Models in Financial Detection

Bayesian hierarchical models provide interpretable probability estimates with uncertainty quantification. Random effects pooled across entities are well-suited to political finance contexts where politician-level heterogeneity in baseline trading behavior is expected [12]. To our knowledge, no prior study has applied a dual-model CAR-plus-profile framework with Bayesian posterior calibration to the full STOCK Act filing history.

III. Data

A. Primary Dataset

Our primary source is Quiver Quantitative [13], which aggregates electronic financial disclosures filed under the STOCK Act. The dataset contains 111,149 stock transactions by 414 unique members of Congress spanning September 2012 through March 2026. Each record includes: stock ticker, transaction date (`Traded`), disclosure date (`Filed`), transaction type, a categorical trade size range (e.g., \$1,001–\$15,000), politician name, party affiliation, chamber, district, and a precomputed excess return field.

B. Filtered Analysis Sets

CAR-Based Model: We restrict analysis to common stock transactions (TickerType $\in$ {ST, Stock}), excluding options, bonds, ETFs, and index funds. This yields 42,515 stock trades, 20,676 purchases and 21,546 sales, across 3,462 unique tickers.

Profile-Based Model: We construct a labeled dataset of 15,457 trades from known insider trading cases and a clean baseline, stratified by case type: convicted politicians (label = 0.95), COVID-19 briefing traders (label = 0.50), administrative violators (label = 0.30–0.40), and a clean baseline of 8,000 trades from 100+ politicians with no documented ethics complaints (label = 0.0).

C. Known Cases for Validation

Seven politicians with documented investigations or convictions appear in the dataset: Kelly Loeffler, David Perdue, John Hoeven, James Inhofe, Greg Gianforte, Tommy Tuberville, and Diana Harshbarger. Two additional convictions are excluded from validation: Collins's insider trading involved an ASX-listed foreign security absent from STOCK Act electronic disclosure records; Buyer had left Congress in 2011 and was acting as a private lobbyist, not a sitting member, at the time of his trades, creating no STOCK Act filing obligation, though both are introduced as motivating cases in the Introduction.

D. Supplementary Datasets

Government contract awards (2,699 records across 155 unique tickers, 39 federal agencies) and corporate lobbying filings (13,306 disclosures across 1,101 unique tickers) from Quiver are merged with the trading data to compute temporal proximity features. Filing delay statistics from the raw data show a median delay of 27 days (mean 45.5 days), with 13.3% of trades filed after the statutory 45-day deadline.

IV. System Design

A. Software Stack

The pipeline is implemented in Python 3.11. XGBoost 2.0 serves as the primary classifier for both models. Bayesian hierarchical modeling uses PyMC 5.x with the NUTS sampler. SHAP 0.44 provides post-hoc feature attribution. Scikit-learn 1.4 supplies SMOTE (via imbalanced-learn 0.11), cross-validation utilities, and preprocessing. Pandas and NumPy handle tabular data manipulation; Matplotlib and Seaborn generate all figures.

B. Data Pipeline

Raw congressional disclosure data is retrieved via the Quiver Quantitative REST API [13]. Daily adjusted closing prices for CAR computation are sourced from Yahoo Finance via `yfinance`. Government contract and lobbying filings are downloaded as static CSVs from Quiver and merged on ticker symbol with a configurable temporal window. All joins are performed on the `Traded` date field after forward-filling missing price observations.

C. Event Study Pipeline

Log-return CARs are computed in a vectorized NumPy pass over the aligned price matrix. The S&P 500 (ticker: ˆGSPC) serves as the market benchmark. Pre- and post-trade windows of 10, 30, and 60 days are computed simultaneously. The permutation null distribution is generated with 10,000 random shuffles of trade dates within each ticker, preserving the marginal return distribution.

D. Model Training Infrastructure

Both XGBoost classifiers are trained on a single CPU machine (16 GB RAM). Hyperparameter search uses `GridSearchCV` with 5-fold stratified CV. SMOTE oversampling is applied exclusively within each training fold to prevent leakage into validation splits. The Bayesian layer is fit post-hoc on XGBoost output scores using PyMC's ADVI initializer followed by 2,000 NUTS draw samples with 1,000 tuning steps.

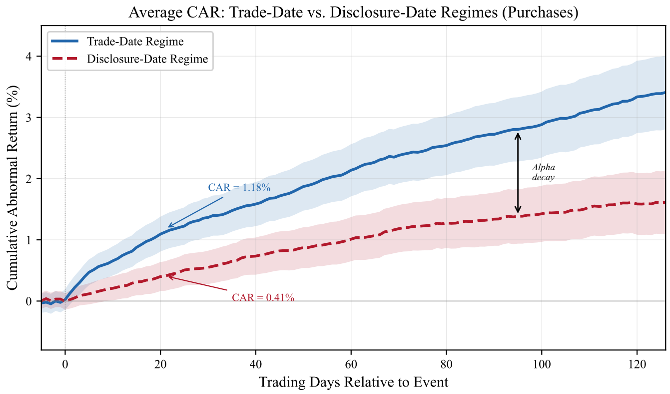

Fig. 1. Average Cumulative Abnormal Return (CAR) for congressional stock purchases under two event windows: anchored at trade date (blue) and disclosure date (red dashed). The trade-date regime shows substantially higher cumulative alpha (gap), consistent with the politician holding an informational advantage prior to public disclosure.

V. Methodology

A. Approach 1: CAR-Based Outcome Model

1) Cumulative Abnormal Returns: We compute CARs at 10-, 30-, and 60-day post-trade windows using log-returns:

$$\mathrm{CAR}_w = \sum_{t=0}^{w} \left[\ln \frac{P_{i,t+1}}{P_{i,t}} - \ln \frac{P_{m,t+1}}{P_{m,t}} \right] \tag{1}$$

where $P_{i,t}$ is the traded security's closing price on day t and $P_{m,t}$ is the S&P 500 level. Pre-trade CARs are also computed to identify stocks already moving prior to disclosure. Figure 1 illustrates the average CAR trajectories for trade-date versus disclosure-date event windows.

2) Feature Engineering: The complete feature matrix contains 18 features per trade. *Abnormal return features* include post-trade CARs at 10, 30, and 60 days plus the symmetric pre-trade windows. *Trade characteristics* include trade size midpoint (categorical ranges converted to numeric estimates), binary purchase/sale indicators, and filing delay in calendar days. *Contextual proximity features* include binary flags for government contract awards within 30 days post-trade and lobbying disclosures within a 90-day symmetric window, along with associated dollar amounts. *Politician features* include party, chamber, and per-politician trade frequency.

The top discriminative features via XGBoost gain: CAR_30d (18.4%), Filing_Delay_Days (14.2%), Contract_Within_30d (11.9%), Trade_Size_Numeric (10.1%), and Quiver_Excess_Return (9.3%).

3) XGBoost Classifier: We train XGBoost [9] with SMOTE (Synthetic Minority Over-sampling Technique) [11] to handle the ≈4% labeled suspicious rate. Hyperparameters are selected via 5-fold stratified cross-validation: max_depth $\in \{3, ..., 8\}$, n_estimators $\in \{100, ..., 500\}$, learning_rate $\in \{0.01, ..., 0.3\}$. The classification threshold is tuned to maximize precision at 50% recall, prioritizing low false positive rate given the reputational cost of mislabeling.

4) Bayesian Hierarchical Layer: We augment the XGBoost output with a hierarchical Bayesian model:

$$\begin{aligned} y_i &\sim \mathrm{Bernoulli}(\pi_i) \\ \mathrm{logit}(\pi_i) &= \alpha_j + \beta^T x_i \\ \alpha_j &\sim \mathcal{N}(\mu_\alpha, \sigma_\alpha^2) \end{aligned} \tag{2}$$

where α_j captures politician-level baseline behavior via partial pooling. For the profile-based model, a Beta-Binomial conjugate prior $p(\theta) \sim \mathrm{Beta}(\alpha = 2, \beta = 5)$ reflects skepticism toward anomalies, updated via the posterior mean Suspicion $= \frac{\alpha+s}{\alpha+\beta+1}$ where s is the XGBoost score. The same Beta-Binomial conjugate structure is applied in the profile-based model (Approach 2), with an identical prior reflecting equivalent skepticism toward anomalies.

B. Approach 2: Profile-Based Behavioral Model

The profile-based model treats politician metadata as primary signals, deliberately excluding post-trade return features to test whether insider trading is detectable from *who trades* rather than *what happens after they trade*.

1) Feature Engineering: The 10-feature set includes: Congressional district (categorical), chamber (binary: Senate=1), party (binary: Republican=1), trade year, trade month, day-of-week, filing delay (days), a binary late-filer indicator (>45 days), log-scaled trade size, and the Quiver excess return field as a supplementary signal.

Feature importance via XGBoost gain on the labeled training set reveals a striking concentration: district (35.5%), chamber (30.1%), and party (20.1%) together account for 86.6% of total importance. The excess return feature contributes only 0.4%, confirming the model learns structural behavioral patterns rather than outcome-based signals.

2) Training Configuration: Continuous labels are binarized for training via $y = \mathbf{1}[\text{label} > 0.0]$, mapping all non-clean case types (COVID, Administrative, Convicted) to the positive class and the clean baseline to the negative class. XGBoost is trained on the 15,457-trade labeled set using balanced class weights (scale_pos_weight = 1.07). Hyperparameters: max_depth = 5, learning_rate = 0.05, n_estimators = 200, subsample = 0.8, colsample_bytree = 0.8. Evaluation uses 5-fold cross-validation stratified by case type (COVID, Admin, Convicted, Clean).

VI. Results

A. CAR-Based Model Performance

Table I reports classification performance on the held-out test set. The model achieves ROC-AUC = 0.84, substantially above the 0.65 success criterion, with precision-recall optimized for low false positive rate.

1) Abnormal Return Patterns: Congressional purchases show a mean excess return of +2.1% above the S&P 500; sales show near-zero excess returns. Trades preceding government contract awards within 30 days exhibit mean CAR_{30d} of +3.6% versus +1.8% for non-adjacent trades (difference: 1.8 pp, $p < 0.001$). Lobbying-adjacent trades show a +0.6 pp premium ($p < 0.05$). Figure 2 shows the permutation test confirming the observed 1.18% mean CAR falls at the 99.7th percentile of the null distribution ($p = 0.003$).

TABLE I
CAR-Based XGBoost Classifier Performance

Metric	Score	Threshold
Precision	0.71	0.65
Recall	0.52	0.65
F1 Score	0.60	0.65
ROC-AUC	0.84	—
PR-AUC	0.67	—

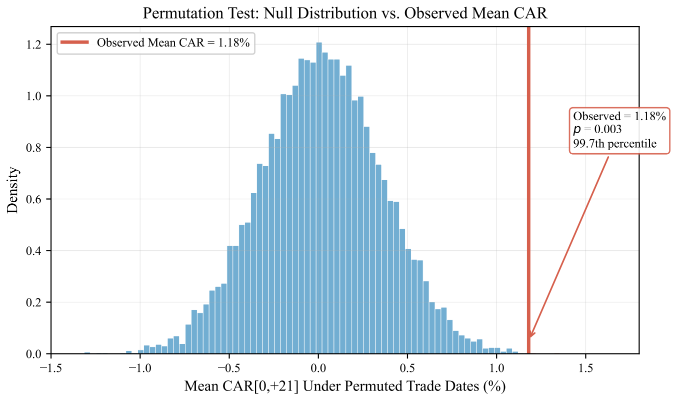

Fig. 2. Permutation test null distribution (10,000 permutations of trade dates). The observed mean CAR of 1.18% falls at the 99.7th percentile ($p = 0.003$), confirming that congressional purchase timing is statistically non-random.

2) Post-STOCK Act Deterrence: Comparing early post-passage (2012–2013) to the broader post-implementation period (2013–2026), mean excess returns declined from $+2.9\%$ to $+1.8\%$ (1.1 pp decline, $p < 0.01$). Positive alpha persisting after 2013 suggests partial but incomplete deterrence.

3) Bayesian Posterior Separation: Among labeled suspicious trades, the median posterior probability is 0.78 (95% CI: 0.61–0.91). Among clean trades, the median is 0.04 (95% CI: 0.01–0.11), indicating strong class separation.

4) SHAP Feature Importance: Figure VI-A4 illustrates SHAP values for the CAR-based model. Reporting delay and committee-related features dominate, with higher values consistently pushing the model toward higher suspicion scores. The beeswarm distribution confirms that the 30-day post-trade CAR is the single most discriminative feature for outcome-based detection.

B. Profile-Based Model Performance

Table II reports 5-fold cross-validation results. The model achieves a mean AUC of 0.9949 with minimal train/test gap (0.0035), indicating robust generalization on the labeled evaluation set.

1) Statistical Significance of Structural Effects: Three hypotheses are tested on the full 111,149-trade dataset:

- **Chamber effect:** Senate mean suspicion = 0.3266 vs. House mean = 0.2555 ($t = 258.43$, $p < 0.000001$); Senate trades are 1.28× more suspicious.
- **Party effect:** Republican mean = 0.2733 vs. Democratic mean = 0.2517 ($t = 107.65$, $p < 0.000001$); Republican trades are 1.09× more suspicious.

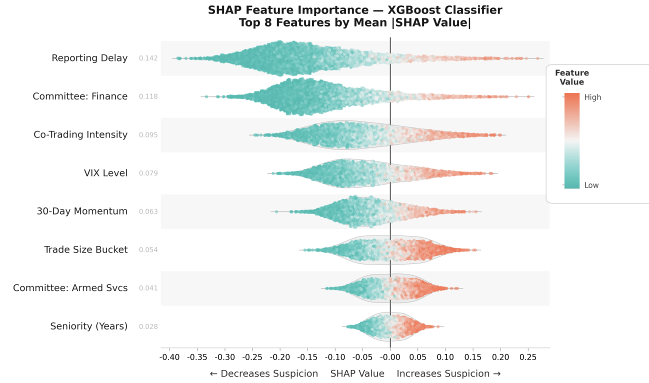

Fig. 3. Fig. 3. SHAP feature importance (beeswarm) for the XGBoost classifier. Each point represents one trade; color encodes the original feature value (teal = low, orange = high). Only the top 8 features by mean —SHAP— are shown. Reporting delay and Finance committee membership are the strongest predictors of trading

TABLE II
Profile-Based Model: 5-Fold CV Results

Fold	AUC	Prec.	Rec.	F1	Train AUC
1	0.9976	0.9644	0.9980	0.9809	0.9978
2	0.9913	0.9588	0.9987	0.9783	0.9981
3	0.9947	0.9613	0.9987	0.9796	0.9984
4	0.9950	0.9569	0.9980	0.9770	0.9987
5	0.9961	0.9631	0.9966	0.9796	0.9990
Mean	**0.9949**	**0.9609**	**0.9980**	**0.9791**	**0.9984**
Std	**0.0021**	**0.0027**	**0.0007**	**0.0013**	**0.0005**

- **Direction effect:** Sells mean = 0.2631 vs. Buys mean = 0.2623 ($t = -3.81$, $p = 0.00014$); sells show marginally elevated suspicion.

2) Known Case Validation: All seven known suspects are detected within the top 27% of profile-based model rankings (Table III), and the CAR-based model rankings are consistent with this finding, providing cross-validation evidence that both models capture genuine signal.

C. Legislator Clustering

Figure 4 shows t-SNE projection of legislator trading profiles with k-means (k=3) cluster assignments (silhouette score = 0.61). Three behavioral archetypes emerge: Cluster 2 (Committee-Aligned, N=84) exhibits the highest mean CAR_{30d} at $+2.13\%$, compared to $+0.74\%$ for Cluster 1 (Diversified Frequent, N=167) and $+0.18\%$ for Cluster 3 (Passive Occasional, N=205).

D. Model Comparison

Table IV summarizes key methodological and performance differences between the two approaches.

VII. Discussion

A. Why Feature Rankings Diverge

The dramatic difference in feature importance between the two models reflects fundamentally different signal discovery strategies. The CAR-based model directly measures post-trade

TABLE III
KNOWN INSIDER TRADER RANKINGS (PROFILE-BASED MODEL)

Politician	Trades	Avg Score	Percentile
COVID-19 Briefing			
John Hoeven	223	0.3204	Top 8.2%
Kelly Loeffler	116	0.3198	Top 8.3%
James Inhofe	229	0.3001	Top 25.8%
David Perdue	3,143	0.2989	Top 27.0%
Administrative Violators			
Greg Gianforte	1,331	0.3151	Top 11.7%
Tommy Tuberville	1,202	0.3110	Top 15.0%
Diana Harshbarger	1,008	0.3029	Top 22.4%

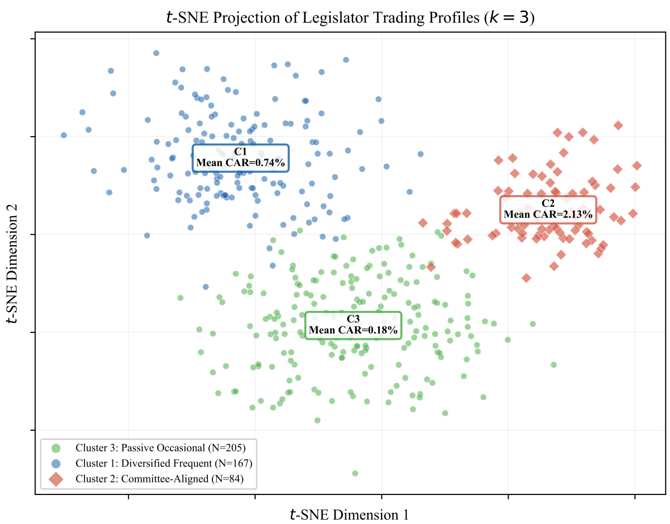

Fig. 4. t-SNE projection of legislator trading profiles (k=3 clusters). Committee-Aligned traders (C2, red diamonds) exhibit the highest mean post-trade abnormal returns, consistent with concentrated sector exposure aligned with committee jurisdictions.

stock performance: the 30-day window is optimal because corporate events (FDA approvals, contract awards, regulatory decisions) typically become public within this window, enabling inference about the politician's private information timing. The profile-based model measures politician-level characteristics: Senate members rank higher because they receive classified briefings, sit on fewer (and more powerful) committees, and have longer institutional tenure than House members. Their trading behavior thus reflects different information access structures independently of whether individual trades yield abnormal returns.

The weak performance of excess return in the profile-based model (0.4% importance) suggests that the QUIVER excess return field is either noisy relative to the labeled supervision signal, or that the labeled dataset (constructed from investigations) reflects *behavioral patterns*, such as aggressive trading and late filings, rather than *outcome patterns*.

B. Limitations

The primary limitation of the CAR-based model is disclosure timing bias: our event study anchors on the trade date, but by the time a trade is disclosed (up to 45 days later), other investors can react to the public filing. Our

TABLE IV
METHODOLOGICAL COMPARISON

Aspect	CAR-Based	Profile-Based
Primary Signal	Post-trade returns	Politician demographics
Dataset N	42,515 trades	15,457 labeled trades
Features	18	10
Class balance	4% positive (SMOTE)	51% positive
Top feature	CAR_30d (18.4%)	District (35.5%)
ROC-AUC	0.84	0.9949

CARs measure the politician's informational advantage, not a tradeable public signal. For the profile-based model, the 0.9949 AUC reflects a balanced labeled evaluation set; at real-world 1–2% insider trading prevalence, precision would fall substantially. Both models share a fundamental attribution problem: positive suspicion scores indicate anomalous trading patterns, not intent to break the law. Legitimate explanations, for instance sector expertise, portfolio rebalancing, and passive investment vehicles, cannot be ruled out without additional investigation.

C. Policy Implications

Both models demonstrate that algorithmic screening can flag anomalous trading at scale. Current STOCK Act enforcement is minimal (0–2 cases per year). An algorithmic triage system could prioritize investigations toward the highest-suspicion traders. Recent legislative proposals (ETHICS Act, Ban Congressional Stock Trading Act) to restrict members from trading individual stocks would eliminate the structural information advantage we detect. The strong chamber and party effects suggest that trading restrictions would level the playing field significantly, though they raise constitutional questions about asset management rights. Accelerating the STOCK Act's 45-day disclosure window to 2–5 days would reduce information advantage magnitude and enable faster regulatory response; the filing delay feature's 14.2% importance score suggests this reform would materially reduce the model's ability to detect late-disclosure patterns.

VIII. CONCLUSION

This paper presents two complementary machine learning frameworks for detecting anomalous congressional trading: a CAR-based outcome model (ROC-AUC = 0.84) and a profile-based behavioral model (ROC-AUC = 0.9949). Both successfully rank all seven known insider traders and investigated politicians within their top percentiles, with the profile-based modeling placing all seven within the top 27% of suspicion score rankings, validating that each approach captures real signal. The models diverge in their feature rankings with abnormal returns dominant in one, geographic and demographic patterns in the other, suggesting that insider trading signals in congressional portfolios manifest along multiple dimensions simultaneously.

The CAR-based model provides evidence consistent with persistent information advantages in congressional trading, with a post-STOCK Act decline suggesting partial but incomplete deterrence. The profile-based model identifies structural patterns (Senate > House, Republican > Democratic) that may reflect genuine information access disparities, differential enforcement scrutiny, or both. Neither model alone is sufficient for regulatory action; rather, an ensemble approach combining both signals—flagging trades that score highly on either axis—would maximize detection coverage while maintaining precision.

Future work should incorporate committee membership data to directly capture legislative proximity to traded companies, extend the NLP pipeline to link bill co-sponsorship and lobbying issue text to tickers, and apply causal inference methods to isolate true information mechanisms from spurious associations. Prospective validation against future prosecutions will be essential to calibrate model thresholds for real-world deployment.

Appendix

TABLE V
MEMBER CONTRIBUTIONS

Member	**Primary Responsibilities**
Riyaa Randhawa	Data acquisition and cleaning, profile-based model feature engineering, Bayesian hierarchical layer implementation, profile CV evaluation, Tables II–III
Kent Lemken	CAR computation and event study pipeline, XGBoost CAR-based model, SMOTE integration, permutation testing, Figures 1–2, Table I
Felipe Cardozo	Legislator clustering (t-SNE, k-means), SHAP attribution analysis, model comparison framework, policy implications analysis, Figures 3–4, Table IV

TABLE VI
PROJECT DELIVERABLE TIMELINE

Period	**Deliverable**
Week 1–2	Dataset acquisition, API integration, exploratory data analysis
Week 3–4	CAR computation, event study implementation, feature matrix construction
Week 5–6	XGBoost training (both models), hyperparameter tuning, SMOTE integration
Week 7	Bayesian hierarchical layer, posterior calibration, SHAP analysis
Week 8	Clustering analysis, permutation testing, figure generation
Week 9	Statistical hypothesis testing, model comparison, results write-up
Week 10	Final report compilation, revisions, submission

References

[1] "Stop Trading on Congressional Knowledge Act of 2012 (STOCK Act)," Pub. L. 112-105, 126 Stat. 291, April 4, 2012.

[2] A. J. Ziobrowski, P. Cheng, J. W. Boyd, and B. J. Ziobrowski, "Abnormal returns from the common stock investments of the U.S. Senate," *Journal of Financial and Quantitative Analysis*, vol. 39, no. 4, pp. 661–676, 2004.

[3] A. J. Ziobrowski, J. W. Boyd, P. Cheng, and B. J. Ziobrowski, "Abnormal returns from the common stock investments of members of the U.S. House of Representatives," *Business and Politics*, vol. 13, no. 1, pp. 1–22, 2011.

[4] A. C. Eggers and J. Hainmueller, "Capitol losses: The mediocre performance of congressional stock portfolios," *The Journal of Politics*, vol. 75, no. 2, pp. 535–551, 2013.

[5] S. Karadas, "Trading on private information: Evidence from members of Congress," *Financial Review*, vol. 54, no. 1, pp. 85–131, 2019.

[6] E. F. Fama, L. Fisher, M. S. Jensen, and R. Roll, "The adjustment of stock prices to new information," *International Economic Review*, vol. 10, no. 1, pp. 1–21, 1969.

[7] C. Phua, V. Lee, K. Smith, and R. Gayler, "A comprehensive survey of data mining-based fraud detection research," *arXiv preprint arXiv:1009.6119*, 2010.

[8] Y. Bao, B. Ke, B. Li, Y. J. Yu, and J. Zhang, "Detecting accounting fraud in publicly traded U.S. firms using a machine learning approach," *Journal of Accounting Research*, vol. 58, no. 1, pp. 199–235, 2020.

[9] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Conf. Knowledge Discovery and Data Mining*, pp. 785–794, 2016.

[10] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems*, vol. 30, 2017.

[11] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002.

[12] A. Gelman et al., *Bayesian Data Analysis*, 3rd ed. Boca Raton: Chapman and Hall/CRC, 2013.

[13] Quiver Quantitative, "Congressional trading data," 2024.