

CONGRESSIONAL TRADING REVIEW

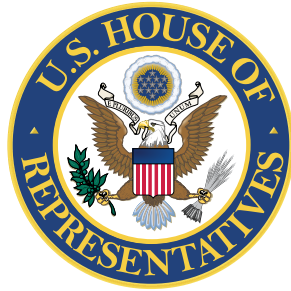
TECHNICAL METHODOLOGY

Insider Trading Pattern Tests for the U.S. Congress

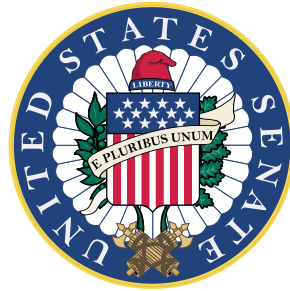
Technical Methodology: Definitions, Estimators, Confidence Intervals, Algorithms and
KPIs for All House and Senate Members

October 25, 2026

Public records, open methods, stated limits.



U.S. HOUSE OF REPRESENTATIVES



U.S. SENATE

Joint Mission Statement

To serve as the voice of the people by enacting laws, funding government programs, and providing democratic oversight to protect the general welfare and constitutional rights of the nation.

The mission of Congress rests on Article I of the Constitution of the United States, whose first section reads: "All legislative Powers herein granted shall be vested in a Congress of the United States, which shall consist of a Senate and House of Representatives."

U.S. Const. art. I, § 1.

The seals of the House of Representatives and the Senate are shown to identify the institutions whose members' trading this document examines. This methodology is an independent analysis of public records. It was not produced, sponsored or endorsed by the House of Representatives, the Senate or any member.

Contents

1	Purpose and Scope	4
2	Legal Definition and Detection Literature	5
2.1	What is being detected	5
2.2	How informed trading is detected	5
2.3	What the congressional literature finds	6
3	Data Model and Notation	6
3.1	Disclosure records	6
3.2	Interval censored amounts	6
3.3	Market data and benchmarks	7
3.4	Event calendars	7
3.5	Classification and jurisdiction maps	8
4	Statistical Framework	9
4.1	Hypotheses	9
4.2	From statistics to readings: equivalence and superiority	9
4.3	Bootstrap	10
4.4	Permutation tests for timing	10
4.5	Power and minimum detectable effects	11
5	Screening 535 Members: What Changes at Scale	11
5.1	Two level multiplicity control	11
5.2	Peer benchmarks	12
5.3	Empirical Bayes pooling	12
5.4	Coverage and <i>not assessed</i>	12
6	The Tests	12
6.1	T1: Trade size	12
6.2	T2: Breadth and concentration	14
6.3	T3: Cadence and event proximity	15
6.4	T4: Event timing at the trade level	16
6.4.1	Trade level event study	16

6.4.2	Calendar time portfolio	18
6.5	T5: Direction before market wide shocks	19
6.6	T6: Control and independence from the member's information	19
6.7	T7: Disclosure timeliness	20
6.8	T8: Committee overlap	21
6.9	T9: Policy topic overlap	21
6.10	T10: Behavior after scrutiny	22
6.11	T11: Performance relative to Congress	23
6.12	T12: Anticipation of large stock moves	24
6.13	T13: Cotrading with other members	24
6.14	T14: Family account trades	25
6.15	T15: Short swing round trips	25
6.16	T16: Options and leverage	26
7	KPIs	26
8	Aggregation and Reporting	26
8.1	The scorecard	26
8.2	Robustness	28
9	Implementation and Validation	28
9.1	Code	28
9.2	Validation on synthetic data with known answers	29
10	Application Across Congress	29
11	Limitations	30
A	Formula Reference	31
B	Committee Jurisdiction Map	31

1 Purpose and Scope

This document gives the full technical specification of the tests we apply to the public stock trading records of every member of the U.S. House of Representatives and Senate, applied as an automated screen of all 535 members. For each test it sets out the trading signature the test looks for and the literature behind it, a formal definition in common notation, the statistical estimator, its confidence interval or test distribution, an algorithm that computes it from disclosure and market data, and the decision rule that maps the result to one of four readings: *does not fit*, *open question*, *fits*, or *not assessed* when the data needed are missing.

Ten tests (T1 to T10) examine each member’s own trades, with decision rules built for use at scale, and six more (T11 to T16) that exploit what a population of traders offers and a single trader does not: a cross section to benchmark against, other members to cotrade with, and enough trades in family accounts, round trips and options to test them directly. Section 5 explains what changes when the same tests are run 535 times; Section 7 defines the member level and Congress level KPIs that the public dashboard reports; and Section 9 maps every estimator to the code that computes it.

What the readings mean. The tests are detection tests, not proof. A *fits* reading says the record carries a statistical signature that informed trading leaves, after correcting for the number of tests run on the member and the number of members screened. It is a reason to look closer, not a finding of wrongdoing. A *does not fit* reading says the record shows the opposite of, or none of, the signature with enough precision to rule out an economically meaningful effect. An *open question* says neither could be established, or that a compliance or conflict of interest concern exists that does not bear on information use. Results are published with their statistics, intervals, data coverage and sources so that every reading can be checked.

Insider trading under SEC Rule 10b5-1 turns on whether a trader was aware of material nonpublic information when trading [48], which only an investigation can establish. What trading records can show is whether a portfolio carries the statistical signatures that informed trading leaves in prosecuted cases and in research [34, 11, 51]. The framework below is built so that a clean result is informative (an equivalence test, not merely a failure to reject), so that the number of members screened does not manufacture false positives, and so that its limits are stated alongside every result.

2 Legal Definition and Detection Literature

2.1 What is being detected

Rule 10b5-1(a) treats as a manipulative and deceptive device the purchase or sale of a security on the basis of material nonpublic information in breach of a duty of trust or confidence, and Rule 10b5-1(b) provides that a trade is on the basis of such information if the trader was aware of it when trading [48]. Liability extends to tippees whose tipper breached a duty for a personal benefit [55] and to those who misappropriate information from its source [56]. Section 4 of the STOCK Act confirms that members of Congress owe such a duty with respect to information obtained through their positions, and the Act requires periodic transaction reports within 30 days of notice and no later than 45 days after a covered trade [54]. Rule 10b5-1(c) provides an affirmative defense for trades made under arrangements adopted before the trader became aware of the information, which the SEC tightened in 2022 [49]; its logic, that a trader who does not control timing cannot trade on information, underlies our control test. Section 16(b) of the Securities Exchange Act requires corporate insiders to disgorge profits from a purchase and sale within six months [50]; it does not apply to members of Congress, but its six month window is the natural horizon for our short swing test (T15).

2.2 How informed trading is detected

Theory predicts that an informed trader trades in the direction of private information and that prices move toward that information after the trade [29]. Empirical detection uses four families of evidence:

1. **Returns after the trade.** Abnormal returns concentrated in the days after the trade and before an announcement are the primary signature in SEC cases [34] and in studies of corporate insiders [23, 43, 30, 24]. Studies of members of Congress apply the same event study and calendar time methods [46, 47, 18, 19, 26, 4].
2. **Trade characteristics.** Informed trades are opportunistic rather than routine [11], larger than the trader’s normal activity, and concentrated in the names tied to the information [34, 2]. Informed traders favour leverage, including options, when they can [3].
3. **Network evidence.** Inside information spreads through social and professional networks, so informed trades cluster across connected traders [2].
4. **Market side anomalies.** Abnormal volume or order imbalance before announcements, including in options [3, 13]. The SEC’s Market Abuse Unit runs an Analysis and Detection Center that uses data analysis tools to detect suspicious trading patterns [51, 52].

Our tests use the first three families, which can be computed from members’ disclosures and public daily prices. The fourth requires order level market data and is outside the scope of a public record review.

2.3 What the congressional literature finds

Early studies found large abnormal returns for Senate [46] and House [47] portfolios in the 1990s. Later work using broader samples and calendar time methods found that the average congressional portfolio does not beat the market [18, 4], but that trades in companies connected to the member [19] and trades by members with more power or information access [26] perform better. The lesson for a screen is that a Congress wide average can be near zero while a minority of members or trades carry the signature, so tests must be run member by member with explicit control of false discoveries (Section 5), and benchmarked against the congressional cross section (T2, T8, T11, T13).

3 Data Model and Notation

3.1 Disclosure records

Each member $m = 1, \dots, 535$ files periodic transaction reports (PTRs) with the Clerk of the House or the Secretary of the Senate. Pooled across members they yield a trade table indexed by i , with $\mathcal{I}_m$ the trades of member m and $N_m = |\mathcal{I}_m|$. Each trade records:

- $m(i)$, the member (Bioguide identifier), with chamber, party and state;
- t_i , the transaction date, and f_i , the date the report was filed;
- $j(i)$, the issuer, mapped to a ticker, a GICS sector and a GICS subindustry;
- $d_i \in \{+1, -1, 0\}$, the direction (purchase = +1, full or partial sale = -1, exchange = 0);
- $[L_i, U_i]$, the disclosed value band (for example [1,001, 15,000]);
- o_i , the owner (self, spouse, joint, dependent) and a_i , the asset type (stock, option, fund, bond, other).

Let $\mathcal{J}_m$ be the set of issuers member m traded. Where no ambiguity arises we drop the member index and write $N, \mathcal{J}$ for a single member's record.

3.2 Interval censored amounts

The true dollar amount v_i is known only to lie in $[L_i, U_i]$. We treat this as interval censoring and carry it through every statistic in three ways.

Bounds (partial identification). For any statistic $S(\mathbf{v})$ that is monotone in each v_i , sharp bounds are obtained by evaluating S at the band endpoints. For nonmonotone statistics we compute

$$\underline{S} = \min_{\mathbf{v} \in \prod_i [L_i, U_i]} S(\mathbf{v}), \quad \overline{S} = \max_{\mathbf{v} \in \prod_i [L_i, U_i]} S(\mathbf{v}), \quad (3.1)$$

following the partial identification approach of [33]. Results are reported as identified sets $[\underline{S}, \overline{S}]$ whenever the conclusion depends on the amounts.

Point imputation. For descriptive figures we use the geometric midpoint $\tilde{v}_i = \sqrt{L_i U_i}$, which is the median of a log uniform distribution on the band and is less sensitive than the arithmetic midpoint to the long upper tail of the wide bands.

Stochastic imputation. For confidence intervals that must reflect amount uncertainty, we draw $v_i^{(b)} \sim \text{LogUniform}(L_i, U_i)$ independently across trades and replications $b = 1, \dots, B$:

$$v_i^{(b)} = \exp(\ln L_i + U_i^{(b)}(\ln U_i - \ln L_i)), \quad U_i^{(b)} \sim \text{U}(0, 1), \quad (3.2)$$

and nest this draw inside the bootstrap of Section 4.3.

Open ended bands. The top bands (“Over \$50,000,000”, and “Over \$1,000,000” for spouse assets) have no upper limit. We set $U_i = 2L_i$, which is conservative for bounds on shares because it limits how much weight a single trade can carry, and flag every statistic that uses such a trade.

3.3 Market data and benchmarks

For issuer j on trading day t , $R_{j,t}$ is the total return computed from adjusted closing prices, $R_{m,t}$ the market return (the SPDR S&P 500 ETF, SPY, as a total return proxy), $R_{f,t}$ the risk free rate, and $\mathbf{F}_t = (\text{MKT}_t, \text{SMB}_t, \text{HML}_t, \text{MOM}_t)$ the Fama French and momentum factors [20, 8].

Factor data. Daily factors come from the Kenneth French Data Library (the five factor 2×3 daily file and the daily momentum file), converted from percent to decimal returns and stored in `data/market/ff_factors_daily.csv`; the current file covers 3 January 2012 to 31 August 2026. The file also carries RMW and CMA, which are not used. Its RF column supplies $R_{f,t}$. Fama French factors enter only the calendar time portfolio regression of T4 (Section 6.4) and the Congress level calendar time alpha KPI. Trade level abnormal returns, and every test built on them, use the single index market model instead (see the note in Section 6.4). If the factor file is missing or covers less than 80% of trading days, the calendar time regression falls back to the market excess return alone. Trading days are indexed in exchange calendar time defined by the benchmark; a trade on a nontrading day is assigned to the next trading day. A missing price leaves a missing return rather than a multiday jump.

3.4 Event calendars

Five calendars are used. The first two are built automatically from prices for every member, which is what allows the timing tests to run Congress wide; the last three are curated and used where available.

1. $\mathcal{E}^{\text{mkt}}$: market wide shocks, the benchmark days on which $|R_{m,t}| > 3\hat{\sigma}_{m,t}$, with $\hat{\sigma}_{m,t}$ the trailing 250 day standard deviation (used by T5);
2. $\mathcal{E}_j^{\text{firm}}$: firm news for issuer j , proxied by the days on which the market adjusted return $R_{j,t} - R_{m,t}$ exceeds three trailing 250 day standard deviations in absolute value. This captures earnings surprises, guidance changes, mergers, regulatory decisions and contract awards whether or not a curated calendar lists them (used by T1, T2, T3);
3. $\mathcal{E}^{\text{leg}}$: legislative events (floor votes, committee markups, hearings), each tagged with a committee and the affected industries or policy topic (used by T9);

4. $\mathcal{P}_m$: the member’s presence calendar, with leave or absence periods (used by T6);
5. $\mathcal{S}_m$: scrutiny and arrangement dates (press reports, ethics complaints, blind trust adoption) and documented episodes (used by T5, T6 and T10).

3.5 Classification and jurisdiction maps

Each issuer carries a GICS sector and subindustry from a security master (S&P 500 constituents plus the dataset’s own security table). A jurisdiction map κ assigns each committee c a set of GICS sectors and subindustries over which it has legislative jurisdiction, built from House Rule X and Senate Rule XXV. Table 6 in Appendix B summarizes it; the full map, with the rule paragraph that justifies each row, is the file `analysis/reference/jurisdiction.csv`. Committee assignments come from the `unitedstates/congress-legislators` project and are matched at the committee level (subcommittees inherit their parent’s jurisdiction). Appropriations, Budget, Rules, House Administration and Ethics have jurisdiction too broad or too internal to map to industries and are excluded from T8.

Algorithm 1 Disclosure ingestion and normalization (all members)

Require: PTR filings $\mathcal{F}$ for all members; member roster; security master; exchange calendar

Ensure: Trade table $\{(m_i, t_i, f_i, j(i), d_i, L_i, U_i, o_i, a_i)\}$

- 1: **for** each filing $F \in \mathcal{F}$ **do**
 - 2: Match filer to a Bioguide identifier; flag unmatched filers
 - 3: **for** each transaction line $\ell \in F$ **do**
 - 4: Parse date, asset name, ticker, type (P, S, S (partial), E), owner, asset type and amount band
 - 5: Map asset to issuer j by ticker, then CUSIP, then fuzzy name match; flag unmatched
 - 6: $d \leftarrow +1$ if type = P; $d \leftarrow -1$ if type $\in \{S, S \text{ (partial)}\}$; $d \leftarrow 0$ for exchanges
 - 7: $[L, U] \leftarrow$ band endpoints, $U \leftarrow 2L$ for open ended bands; $t \leftarrow$ next trading day on or after the transaction date
 - 8: Append $(m, t, f_F, j, d, L, U, o, a)$
 - 9: **end for**
 - 10: **end for**
 - 11: Mark rows superseded by later amendments as duplicates and drop them
 - 12: Keep equities and options on equities for the tests; funds, bonds and other assets enter activity counts only
 - 13: **return** trade table
-

4 Statistical Framework

4.1 Hypotheses

Each test k is a contrast between two data generating processes for a member:

- H_0 (*uninformed account*): trades are chosen by a process independent of nonpublic information available to the member, so that signed post trade abnormal returns have mean zero and trade timing is independent of the member’s information calendar;
- H_1 (*informed trading*): some trades are timed or chosen using such information, so that signed abnormal returns are positive and trades cluster before information events.

4.2 From statistics to readings: equivalence and superiority

A failure to reject H_0 is weak evidence when a test has little power. We therefore pair each superiority test with an equivalence test in the two one sided tests (TOST) form [42]. Let θ_k be the parameter of test k , oriented so that larger values indicate informed trading, and let $\delta_k > 0$ be the smallest effect that would be economically meaningful (a margin fixed before seeing the data). With a $(1 - 2\alpha)$ confidence interval $[\hat{\theta}_k^-, \hat{\theta}_k^+]$ (90% at $\alpha = 0.05$), and adjusted p values p_k^{Holm} and q_k^{BH} from Section 5:

$$\text{Reading}_k = \begin{cases} \text{fits} & \text{if } p_k^{\text{Holm}} < \alpha, q_k^{\text{BH}} < q^* \text{ and } \hat{\theta}_k^- > 0, \\ \text{does not fit} & \text{if } \hat{\theta}_k^+ < \delta_k, \\ \text{open question} & \text{otherwise,} \\ \text{not assessed} & \text{if the inputs are missing or } N < N_{\min}. \end{cases} \quad (4.1)$$

Tests whose evidence is documentary rather than statistical (control, blind trust status) use the decision rules stated in their own sections. An *open question* is also assigned when a test reveals a compliance or conflict of interest concern that does not bear on whether information was used, such as late filing or overlap between holdings and committee jurisdiction.

Power limited open questions. When an open question arises only because the interval is wide, with the point estimate itself short of the margin ($\hat{\theta}_k < \delta_k$), the reading is marked *power limited*. The dashboard shows it as “inconclusive (too few trades)” and does not count it among a member’s flagged tests. This keeps members with a handful of trades from appearing flagged merely because their data cannot rule anything out.

Screening signals. A *fits* reading must survive both the within member Holm and the cross member Benjamini Hochberg corrections (Section 5). To surface leads for further review, an open question whose *unadjusted* one sided p value is below 0.05 is also marked a *screening signal*. It is not corrected for multiplicity: if no member traded on information, about 5% of all such tests would still show one. The dashboard therefore reports the number expected by chance next to the number observed, lists members with two or more signals first, and never presents a screening signal as a fit.

Minimum sample. Statistical readings require $N_{\min} = 10$ directional trades (or, for tests on a subset, at least three trades in each compared group). Below that the test is *not assessed* and only descriptive counts are reported.

4.3 Bootstrap

Unless an exact distribution is available, intervals are bootstrap intervals [14, 17]. Because trades in a batch share a date and therefore a market shock, we resample *dates*, not trades (a cluster bootstrap): draw D trading dates with replacement from the D distinct trade dates, take all trades on each drawn date, recompute the statistic, and repeat B times. Intervals use the bias corrected and accelerated (BCa) method [15], with the acceleration estimated by a date level jackknife (on a random subset of 200 dates when $D > 200$). Amount imputation (Section 3.2) is nested inside each replication. The default is $B = 499$ per member.

Algorithm 2 Date clustered bootstrap with amount imputation

Require: Trade table; statistic $S(\cdot)$; replications B ; level α

- 1: $\mathcal{D} \leftarrow$ distinct trade dates; $D \leftarrow |\mathcal{D}|$
 - 2: $\hat{S} \leftarrow S(\text{trades with } v_i = \tilde{v}_i)$
 - 3: **for** $b = 1$ to B **do**
 - 4: Draw D dates from $\mathcal{D}$ with replacement; collect all trades on those dates
 - 5: Draw $v_i^{(b)} \sim \text{LogUniform}(L_i, U_i)$ for each collected trade
 - 6: $S^{(b)} \leftarrow S(\text{collected trades with amounts } v^{(b)})$
 - 7: **end for**
 - 8: Compute bias correction $\hat{z}_0 = \Phi^{-1}(\#\{S^{(b)} < \hat{S}\}/B)$
 - 9: Compute acceleration $\hat{a}$ from the date level jackknife
 - 10: **return** BCa interval from the adjusted quantiles of $\{S^{(b)}\}$
-

4.4 Permutation tests for timing

Tests T1, T3 and T13 compare an observed statistic with its distribution when the member's trades are shifted in time relative to an event calendar or to other members' trades. Circular shifts by a single common offset preserve both the spacing of events and the batching of trades, which a naive shuffle would destroy.

Algorithm 3 Circular shift permutation test for event proximity

Require: Trades $\{(t_i, j(i))\}$ in trading day positions; event calendar $\mathcal{E}$; window w ; sample span $[T_0, T_1]$; replications B

- 1: $S_{\text{obs}} \leftarrow S(\text{trades}, \mathcal{E})$ $\triangleright$ e.g., proximity share $\bar{W}$
 - 2: **for** $b = 1$ to B **do**
 - 3: Draw offset $\omega_b \sim \text{U}\{1, \dots, T_1 - T_0\}$
 - 4: $t_i^{(b)} \leftarrow T_0 + ((t_i - T_0 + \omega_b) \bmod (T_1 - T_0))$ for every trade
 - 5: $S^{(b)} \leftarrow S(\text{trades at } t^{(b)}, \mathcal{E})$
 - 6: **end for**
 - 7: $p \leftarrow (1 + \#\{b : S^{(b)} \geq S_{\text{obs}}\})/(B + 1)$
 - 8: **return** S_{obs} , p , the mean and the 95th percentile of $\{S^{(b)}\}$
-

Shifting trades against fixed events is equivalent to shifting events against fixed trades, and it is the form that extends to T13, where the “events” are the other members' trades.

4.5 Power and minimum detectable effects

For a mean signed abnormal return $\bar{\theta}$ estimated from N trades with idiosyncratic H day return volatility σ_H and average pairwise correlation ρ among trades (induced by same date batches), the standard error is

$$\text{SE}(\bar{\theta}) = \frac{\sigma_H}{\sqrt{N}} \sqrt{1 + (N_{\text{cl}} - 1)\rho}, \quad (4.2)$$

where N_{cl} is the average cluster size. The minimum detectable effect at size α (two sided) and power $1 - \beta$ is

$$\text{MDE} = (z_{1-\alpha/2} + z_{1-\beta}) \text{SE}(\bar{\theta}). \quad (4.3)$$

Worked example 4.1 (illustrative power calculation). Assume a daily idiosyncratic volatility of 2% for large capitalization stocks, so that $\sigma_{21} \approx 2\%\sqrt{21} \approx 9.2\%$ over 21 days, and ignore clustering ($\rho = 0$). With $N = 648$ trades the MDE at $\alpha = 0.05$ and 80% power is about 1.0% per trade; with $N = 50$ trades it is about 3.6%. The 50 basis point margin δ_4 is therefore attainable as an equivalence result only for members with several thousand trades or low volatility holdings, and most members' T4 readings will be power limited unless the signature is strong. These are assumptions for illustration, not estimates; clustering raises both figures.

5 Screening 535 Members: What Changes at Scale

Running sixteen tests on each of 535 members is 8,560 hypothesis tests. At $\alpha = 0.05$ with no correction, roughly 27 members would fail T4 alone by chance, and nearly every active trader would fail at least one test. The screen therefore adds a second layer of multiplicity control, replaces fixed benchmarks with congressional peer benchmarks where possible, and pools information across members where single member estimates are too noisy.

5.1 Two level multiplicity control

Within a member we control the family wise error rate over the tests that produce a superiority p value with Holm's step down procedure [22]: order the K_m p values $p_{(1)} \leq \dots \leq p_{(K_m)}$ and reject $H_{(k)}$ for $k < k^*$, where k^* is the first index with $p_{(k)} > \alpha/(K_m - k + 1)$. This answers: given that we looked at this member sixteen ways, is any one result surprising?

Across members we control the false discovery rate for each test with the Benjamini Hochberg procedure [5, 16]: for test k , order the raw p values of the M_k members for which it was computed and set

$$q_{(r)}^{\text{BH}} = \min_{s \geq r} \left\{ \frac{M_k}{s} p_{(s)} \right\}. \quad (5.1)$$

This answers: of all the members we flag on this test, what share do we expect to be false alarms? A *fits* reading requires both $p^{\text{Holm}} < \alpha = 0.05$ and $q^{\text{BH}} < q^* = 0.10$. FDR control is the right criterion across members because the screen's purpose is to direct attention, and a list of flagged members in which at most one in ten is expected to be a false alarm is interpretable, whereas family wise control across 8,560 tests would make the screen powerless.

5.2 Peer benchmarks

Besides comparing a member with a market index, the tests also compare the member with Congress. T2 benchmarks concentration against random portfolios drawn from all congressional trades; T8 benchmarks committee industry exposure against the larger of the S&P 500 share and the share of all congressional trades in the same industries, so that a member is not flagged for holding sectors every member holds; T11 ranks the member’s returns within the cross section; and T13 asks whether a member’s trades coincide with other members’ trades more than chance predicts.

5.3 Empirical Bayes pooling

A member with 20 trades and a mean signed CAR of 4% is less informative than one with 2,000 trades and 1%. T11 makes the comparison explicit with normal normal empirical Bayes shrinkage [36, 16]: each member’s estimate is pulled toward the congressional mean in proportion to its noise (Section 6.11). Rankings published on the dashboard use the shrunk estimates.

5.4 Coverage and *not assessed*

Tests degrade gracefully. Each declares its inputs (Table 1); when an input is missing for a member the test reports *not assessed* with evidence type *not computed*, and where a weaker descriptive form exists it is reported with evidence type *descriptive*. Return based tests need daily prices for the traded issuer; documentary tests (T6, T10) need a curated record that exists for few members; and T9 needs a legislative calendar. The dashboard shows coverage alongside every reading.

6 The Tests

Each subsection follows the same order: signature, definitions, estimator, inference, algorithm, decision rule. Table 1 summarizes them; T1 to T10 examine the member’s own trades and T11 to T16 compare the member with the rest of Congress.

6.1 T1: Trade size

Signature. Informed trades are large relative to the trader’s normal activity, because the expected profit from an informational edge scales with position size [29, 34]. SEC actions against members of Congress involved positions of hundreds of thousands of dollars in a single name ahead of a specific event [53].

Null hypothesis. Trade sizes are drawn from the account’s ordinary size distribution, unrelated to information events.

Primary statistic. The share of trades in the lowest disclosure band, and the size gap between event linked and other trades.

Definitions. Let the disclosure bands be $b = 1, \dots, K$ and n_1 the number of trades in the lowest band (\$1,001 \$15,000). The lowest band share is $\hat{p}_1 = n_1/N$. For trade i define the log geometric

Table 1: Overview of the tests.

#	Test	Primary statistic	Inference	Inputs
T1	Trade size	Lowest band share; event linked size gap Δ_z	Wilson; within issuer permutation	Trades, prices
T2	Breadth and concentration	HHI identified set, N_{eff} , sector entropy	Exact bounds; peer Monte Carlo; bootstrap	Trades, sectors
T3	Cadence and proximity	Batch share, routine share, proximity share $\bar{W}$	χ^2 dispersion; circular shift	Trades, prices
T4	Event timing	Mean signed CAR; calendar time alpha	KP adjusted BMP, Patell, Corrado; Newey West	Trades, prices
T5	Direction before shocks	Hit rate before market wide shocks	Exact binomial; Clopper Pearson; sign flip	Trades, prices
T6	Control	Documentary record; absence rate ratio	Exact conditional binomial	Curated record
T7	Disclosure timeliness	Late share; lag quantiles; late minus timely CAR	Wilson; bootstrap	Filing dates
T8	Committee overlap	Jurisdiction excess share; return premium γ_1	Bootstrap; clustered regression	Committees, sectors, prices
T9	Policy topic overlap	Pre event signed CAR; pre trade drift	Event study	Legislative calendar
T10	Behavior after scrutiny	Rate ratio at a known date; change point	Exact rate ratio; permutation LR	Scrutiny dates
T11	Relative to Congress	Shrunk signed CAR; percentile	Empirical Bayes	T4 for all members
T12	Large move anticipation	Hit rate before 3σ stock moves	Exact binomial	Trades, prices
T13	Cotrading	Cotraded share vs. circular shift null	Permutation; KP test	All members' trades
T14	Family accounts	Family signed CAR; family minus self	KP test; bootstrap	Owner, prices
T15	Short swing round trips	Opening trade signed CAR of round trips	KP test	Trades, prices
T16	Options and leverage	Option share; option signed CAR	KP test	Asset type, prices

midpoint $x_i = \ln \tilde{v}_i$ and the robust size score

$$z_i = \frac{x_i - \text{med}(x)}{1.4826 \text{ MAD}(x)}, \quad (6.1)$$

where MAD is the median absolute deviation and the constant makes the score comparable to a standard normal deviate under normality.

Estimator and interval. $\hat{p}_1$ is a binomial proportion with a Wilson score interval [45]:

$$\hat{p}_1^\pm = \frac{\hat{p}_1 + \frac{z^2}{2N} \pm z \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{N} + \frac{z^2}{4N^2}}}{1 + \frac{z^2}{N}}, \quad z = z_{1-\alpha/2}. \quad (6.2)$$

Band membership is observed exactly, so no imputation is needed.

Event linked size. Let $E_i = \mathbf{1}\{\exists e \in \mathcal{E}_{j(i)}^{\text{firm}} : 0 < e - t_i \leq 5\}$ indicate a trade within five trading days before a firm event. The size gap $\Delta_z = \bar{z}_{E=1} - \bar{z}_{E=0}$ is compared with its distribution when E is permuted *within issuer*, which holds fixed how often each issuer has news and so removes the confound that volatile issuers are both newsworthy and traded in different sizes. The one sided permutation p value enters the multiplicity layer.

Decision rule. *Fits* if $\Delta_z > 0$ with $p^{\text{Holm}} < \alpha$ and $q^{\text{BH}} < q^*$. *Does not fit* if $\hat{p}_1^- > 0.5$ and no event linked trade has $z_i > 3$ in a band of \$50,001 or more. Otherwise *open question*.

Design note: a single event linked outlier no longer suffices for *fits*. Across 535 members, some large trades will precede news by chance; the outlier now prevents a *does not fit* reading but only the permutation test can produce *fits*.

6.2 T2: Breadth and concentration

Signature. Informed trading concentrates in the few issuers tied to the information [34, 2]; a diversified account spreads activity across many issuers and sectors.

Null hypothesis. The issuer and sector mix resembles that of a random draw of the same number of congressional trades.

Primary statistic. The dollar weighted Herfindahl Hirschman index of issuer concentration, its identified set, and the effective number of issuers.

Definitions. Dollar shares are $s_j(\mathbf{v}) = V_j/V$ with $V_j = \sum_{i:j(i)=j} v_i$ and $V = \sum_j V_j$. The concentration index [41] and effective number of issuers are

$$\text{HHI}(\mathbf{v}) = \sum_{j \in \mathcal{J}} s_j(\mathbf{v})^2, \quad N_{\text{eff}} = 1/\text{HHI}. \quad (6.3)$$

Sector diversity uses Shannon entropy [44] over the $G = 11$ GICS sectors with dollar shares q_g , normalized by its maximum: $\tilde{H} = -\sum_g q_g \ln q_g / \ln G$.

Exact bounds under interval censoring. HHI depends on $\mathbf{v}$ only through the issuer totals V_j , which range over $[a_j, b_j] = [\sum_{i:j(i)=j} L_i, \sum_{i:j(i)=j} U_i]$, so both bounds are problems in $|\mathcal{J}|$ rather than N dimensions.

- *Lower bound.* For a fixed total $V = T$, $\sum_j V_j^2$ is minimized subject to the box by water filling, $V_j = \text{clip}(\lambda, a_j, b_j)$ with λ chosen so that the totals sum to T . Since $\text{HHI} = \sum_j V_j^2 / T^2$, the global minimum is $\underline{\text{HHI}} = \min_{\lambda} \text{HHI}(\text{clip}(\lambda, \mathbf{a}, \mathbf{b}))$, a one dimensional search over λ that we solve exactly over the breakpoints $\{a_j, b_j\}$ and refine with a bounded scalar minimization.
- *Upper bound.* The partial derivative $\partial \text{HHI} / \partial V_j = \frac{2}{V^2} (V_j - V \cdot \text{HHI})$ is positive exactly when $s_j > \text{HHI}$. At a maximum every issuer with $s_j > \text{HHI}$ sits at b_j and every other at a_j , so $\overline{\text{HHI}}$ is attained at a vertex. We run vertex coordinate ascent from each prefix of issuers sorted by b_j and keep the largest value.

A narrow identified set means the conclusion does not depend on the unknown amounts.

Peer benchmark. The trade count concentration $\text{HHI}^\# = \sum_j (N_j / N)^2$ is observed exactly. Its null distribution is obtained by drawing N trades at random from the pooled congressional trade table (Monte Carlo, $B = 499$), which reflects the issuers members actually trade rather than index weights that no member holds.

Inference. The interval for $\tilde{H}$ uses the date clustered bootstrap with nested amount imputation (Algorithm 2).

Decision rule. *Does not fit* if $\text{HHI}^\# \leq \max(0.10, 95\text{th percentile of the peer benchmark})$, that is, ten or more effective issuers or no more concentrated than random congressional trades, and $\tilde{H}^- > 0.7$ (or sectors are unmapped). The floor of 0.10 is needed because real accounts trade their holdings repeatedly, so random draws from the pooled trade table are far more dispersed than any actual portfolio: in the first House run, records with 30 or more effective issuers still exceeded the random draw benchmark. Otherwise *open question*. This test has no p value and does not enter the multiplicity layer.

Design note: T2 can no longer produce *fits*. The earlier rule (top three issuers above half of dollar volume, one of them traded before a firm event) has no statistical test behind it, and in the first House run it was met by members who hold a single family or employer stock for years and trade around it routinely. Concentration describes a portfolio, not information use, so it can raise an open question but the informed trading signature must come from the return and timing tests.

6.3 T3: Cadence and event proximity

Signature. Opportunistic trades carry information and routine trades do not [11]. Informed traders trade sporadically around events; managed accounts trade in rebalancing batches across many unrelated names.

Null hypothesis. Trade timing is unrelated to the timing of firm news.

Primary statistic. The proximity share of trades to firm events, with the batch share, the index of dispersion and the routine share as context.

Batch share. A batch is a date with at least $k = 3$ trades in at least three distinct issuers. The batch day share is the number of batch dates over the number of trade dates $|\mathcal{D}|$, with a Wilson interval at the date level to respect clustering.

Dispersion. Over the T business days in the member's trading span, the index of dispersion is $I = s_n^2 / \bar{n}$ for daily counts n_t . Under a homogeneous Poisson process, $(T - 1)I \sim \chi_{T-1}^2$ [10].

Routine share. Following [11], a trade is routine if the same issuer was traded in the same direction in the same calendar month in each of the prior two years. Because tenures are often short, a trade is also routine if it falls on a batch date whose issuer set overlaps an earlier batch by at least 50% (Jaccard index).

Event proximity. With $W_i = \mathbf{1}\{\exists e \in \mathcal{E}_{j(i)}^{\text{firm}} : 0 < e - t_i \leq 5\}$, the proximity share is $\bar{W} = N^{-1} \sum_i W_i$. Its null distribution is obtained by circular shifts of the member's trade dates (Algorithm 3).

Decision rule. *Fits* if $\bar{W}$ exceeds its null with $p^{\text{Holm}} < \alpha$ and $q^{\text{BH}} < q^*$, and the proximity share of nonroutine trades exceeds that of routine trades. *Does not fit* if the proximity p value is at least α . Otherwise *open question*. When prices are unavailable, the descriptive form reads *does not fit* if $\hat{\pi}_B^- > 0.5$ (most trading days are rebalancing batches) and *not assessed* otherwise.

Design note: the batching condition is no longer required for *does not fit*. At scale the permutation test carries the inference, and requiring batches would leave every self directed but uninformed trader as an open question.

6.4 T4: Event timing at the trade level

Signature. Informed purchases are followed by positive abnormal returns and informed sales by negative abnormal returns, concentrated in the days before announcements [34, 23, 43].

Null hypothesis. Signed post trade abnormal returns have mean zero.

Primary statistic. The mean signed cumulative abnormal return over $H = 21$ trading days, and the calendar time alpha of a long purchases, short sales portfolio.

6.4.1 Trade level event study

Normal returns. For each trade, estimate the market model on the estimation window $[\tau_1, \tau_2] = [-250, -30]$ trading days relative to the trade [32, 7]:

$$R_{j,t} = \alpha_j + \beta_j R_{m,t} + \varepsilon_{j,t}, \quad \varepsilon_{j,t} \sim (0, \sigma_j^2), \quad (6.4)$$

by ordinary least squares, requiring $M \geq 120$ valid observations. Estimates are cached by (issuer, trade date) because many members trade the same stock on the same day.

Why not a factor model here. Trade level abnormal returns use the market model, not a Fama French model. Over a 21 day horizon the size, value and momentum loadings of a single stock add little explained variance relative to the market factor [32], while estimating four or five loadings from at most 220 daily observations adds estimation noise to every CAR. Factor exposure that persists across a member's trades, such as a tilt toward small or value stocks, is instead absorbed by the calendar time portfolio regression below, which uses the Fama French three factors plus momentum. So every test that uses trade level abnormal returns (T4, T7, T8, T9, T12, T13, T14, T15 and T16) measures them relative to the market model; the firm news calendar used by T1 to T3 uses market adjusted returns $R_{j,t} - R_{m,t}$; T5 uses the benchmark return itself; and only the calendar time alpha is factor adjusted.

Abnormal and cumulative abnormal returns. For event day τ relative to the trade,

$$AR_{i,\tau} = R_{j(i),t_i+\tau} - \hat{\alpha}_{j(i)} - \hat{\beta}_{j(i)} R_{m,t_i+\tau}, \quad CAR_i(1, H) = \sum_{\tau=1}^H AR_{i,\tau}. \quad (6.5)$$

The window starts at $\tau = 1$ because the disclosed date is the trade date, not the execution time, and the trade day return partly reflects execution. We report $H \in \{5, 21, 63\}$ and use $H = 21$ for the reading.

Signed returns. The informed trading parameter is the mean signed CAR,

$$\theta_4 = \mathbb{E}[d_i CAR_i(1, H)], \quad \hat{\theta}_4 = \frac{1}{N} \sum_{i=1}^N d_i CAR_i(1, H), \quad (6.6)$$

and its dollar weighted analogue $\hat{\theta}_4^s(\mathbf{v}) = \sum_i v_i d_i CAR_i / \sum_i v_i$, reported as an identified set. Because $\hat{\theta}_4^s$ is linear fractional in $\mathbf{v}$, its maximum is attained at the vertex that sets $v_i = U_i$ for the trades with the largest $d_i CAR_i$ and $v_i = L_i$ for the rest; sorting trades on $d_i CAR_i$ and scanning the $N + 1$ thresholds finds it exactly, and the minimum follows symmetrically.

Variance of the CAR. Under the market model the prediction error variance of $CAR_i(1, H)$ is

$$\sigma_{CAR,i}^2 = \hat{\sigma}_j^2 \left(H + \frac{H^2}{M} + \frac{(\sum_{\tau=1}^H (R_{m,t_i+\tau} - \bar{R}_m))^2}{\sum_{t=\tau_1}^{\tau_2} (R_{m,t} - \bar{R}_m)^2} \right), \quad (6.7)$$

where $\bar{R}_m$ is the estimation window mean market return [32]. Standardized CARs are $SCAR_i = d_i CAR_i / \sigma_{CAR,i}$.

Test statistics. We report five statistics, because each fails under different conditions:

1. *Cross sectional t* [7]: $t_{CS} = \hat{\theta}_4 / (s_{dCAR} / \sqrt{N})$.
2. *Patell* [40]: $Z_P = \sum_i SCAR_i / \sqrt{\sum_i \frac{M_i - 2}{M_i - 4}}$.
3. *Boehmer Musumeci Poulsen* [6]: $t_{BMP} = \overline{SCAR} \sqrt{N} / s_{SCAR}$, which remains valid when the event itself raises return variance.
4. *Kolari Pynnönen adjustment* [27]: with average correlation $\bar{r}$ of estimation window residuals across the member's trades (estimated on a random sample of up to 60 trades, over overlapping estimation days, truncated to $[0, 0.99]$),

$$t_{KP} = t_{BMP} \sqrt{\frac{1 - \bar{r}}{1 + (N - 1)\bar{r}}}. \quad (6.8)$$

This is the primary statistic because batches are the norm in managed accounts.

5. *Rank test* [12, 28]: scaled ranks $U_{i,\tau} = K_{i,\tau} / (n_i + 1) - \frac{1}{2}$ of the abnormal returns within each trade's combined estimation window and 63 day event window of n_i valid days; the statistic is $Z_{\text{rank}} = \sum_i d_i \sum_{\tau=1}^H U_{i,\tau} / \sqrt{\sum_i H \widehat{\text{Var}}(U_{i,\cdot})}$, a nonparametric check robust to fat tails.

The $(1 - 2\alpha)$ confidence interval for θ_4 is

$$\hat{\theta}_4 \pm t_{N-1, 1-\alpha} \widehat{SE}_{KP}, \quad \widehat{SE}_{KP} = \widehat{SE}_{CS} \sqrt{\frac{1 + (N - 1)\bar{r}}{1 - \bar{r}}}. \quad (6.9)$$

Algorithm 4 Trade level signed event study (all members)

Require: Trade table; daily returns $R_{j,t}$, $R_{m,t}$; horizons $\mathcal{H}$; window $[\tau_1, \tau_2]$
Ensure: Per trade $\text{CAR}_i(1, H)$, SCAR_i , ranks, pre trade drift, large move flags

```

1: for each distinct (issuer  $j$ , trade day  $t$ ) in the trade table do
2:   Fit  $R_{j,s} = \alpha + \beta R_{m,s} + \varepsilon$  on  $s \in [t + \tau_1, t + \tau_2]$ ; skip if  $M < 120$ 
3:    $\text{AR}_\tau \leftarrow R_{j,t+\tau} - \hat{\alpha} - \hat{\beta} R_{m,t+\tau}$  for  $\tau = -10, \dots, 63$ 
4:   Store  $\text{CAR}(1, H)$  and  $\sigma_{\text{CAR}}$  (Eq. 6.7) for  $H \in \mathcal{H}$ ; ranks;  $\text{CAR}(-10, -1)$ ; first  $|\text{AR}_\tau| > 3\hat{\sigma}$ 
     for  $\tau \leq 10$ 
5: end for
6: for each member  $m$  do
7:    $\hat{\theta}_4 \leftarrow N^{-1} \sum_i d_i \text{CAR}_i$ ;  $t_{\text{BMP}} \leftarrow \overline{\text{SCAR}} \sqrt{N} / s_{\text{SCAR}}$ 
8:    $\bar{r} \leftarrow$  mean pairwise residual correlation;  $t_{\text{KP}} \leftarrow t_{\text{BMP}} \sqrt{(1 - \bar{r}) / (1 + (N - 1)\bar{r})}$ 
9:   Compute  $Z_P$ ,  $Z_{\text{rank}}$ , dollar weighted bounds and the calendar time alpha
10: end for
    
```

6.4.2 Calendar time portfolio

Event time tests overstate significance when events overlap in calendar time [35, 31]. Following the congressional and insider literature [46, 18, 24], we form a daily portfolio that holds each purchased stock and shorts each sold stock for $H = 21$ days after the trade, equal weighted across open positions. Its excess return is regressed on the Carhart four factor model, the Fama French market, size and value factors plus momentum [20, 8] (or on the market excess return alone when the factor file is unavailable):

$$R_{p,t} - R_{f,t} = a + \mathbf{b}^\top \mathbf{F}_t + u_t. \quad (6.10)$$

The intercept a is the calendar time alpha. Standard errors are Newey West with $\lfloor 4(T/100)^{2/9} \rfloor$ lags [37]. At least 60 portfolio days are required.

Decision rule. With margins $\delta_4 = 50$ basis points per trade over 21 days and 2% per year in calendar time, *fits* if the Kolari Pynnönen lower bound and the calendar time lower bound both exceed zero with $p^{\text{Holm}} < \alpha$ and $q^{\text{BH}} < q^*$, where the p value is the larger of the two one sided p values (an intersection union test); *does not fit* if both upper bounds fall below their margins; otherwise *open question*.

Robustness variants. Horizons of 5 and 63 days (margins scaled by $H/21$); excluding the largest batch date; self and joint accounts only; and dollar weighting at both amount bounds. The calendar time half of the rule is already factor adjusted, so there is no separate factor model variant. A reading is marked robust if every variant gives the same reading.

6.5 T5: Direction before market wide shocks

Signature. When a market moving event is coming, an informed trader's trades just before it point in the direction of the coming move.

Null hypothesis. The net direction of trades before a shock is independent of the sign of the shock.

Primary statistic. The hit rate $\hat{h}$: the share of episodes in which pre event net trading points in the direction of the subsequent move.

Episodes. Episodes are built systematically for every member: each day in $\mathcal{E}^{\text{mkt}}$ with at least one of the member's trades in the five trading days before it is an episode, with move Δ_e equal to the benchmark return that day. Documented episodes from $\mathcal{S}_m$ (with their affected issuers and moves) are added where curated. For each episode, $\bar{d}_e = \text{sign}(\sum_{i \in e} \tilde{v}_i d_i)$; the sign is reported as robust if it is the same at both amount bounds. The episode is a hit if $\bar{d}_e \text{sign}(\Delta_e) = +1$. Trades in the five days after the shock are counted as post news and do not enter the statistic.

Estimator and exact inference. With n episodes and X hits, $\hat{h} = X/n$ and, under H_0 , $X \sim \text{Binomial}(n, 1/2)$. The one sided p value is $\Pr(\text{Bin}(n, 1/2) \geq X)$ and the Clopper Pearson interval [9] is $[B^{-1}(\frac{\alpha}{2}; X, n - X + 1), B^{-1}(1 - \frac{\alpha}{2}; X + 1, n - X)]$. The magnitude weighted statistic $S_5 = \sum_e \bar{d}_e \Delta_e$ is compared with its sign flip permutation distribution.

Worked example 6.1 (two wrong way episodes). Suppose a member sold a batch of stocks the day before the benchmark rose sharply, and later bought a stock shortly before it fell sharply on company news. Both episodes are wrong way: $X = 0$, $n = 2$, one sided p value 1, and the 95% Clopper Pearson interval for h is $[0, 0.842]$. Under the systematic construction, a market shock day is in $\mathcal{E}^{\text{mkt}}$ for every member, so every member who traded in the five trading days before it contributes that episode.

Decision rule. Fits if $\hat{h} > 1/2$ with $p^{\text{Holm}} < \alpha$ and $q^{\text{BH}} < q^*$. Does not fit if $\hat{h} < 1/2$ and T4 does not fit. Otherwise *open question*; *not assessed* if no episode exists.

6.6 T6: Control and independence from the member's information

Signature. The person holding the information directs the trade. Rule 10b5-1(c) recognizes that a trader who does not control timing cannot trade on information [48, 49].

Null hypothesis. Trading activity is independent of the member's access to legislative information.

Primary statistic. Documentary evidence of discretion, and the rate ratio of trading intensity during periods of reduced access versus normal access.

Documentary component. The test records the stated arrangement from primary documents: who has discretion, whether the member is consulted, whether a qualified blind trust exists. These are supplied in a curated file (`data/events/arrangements.csv`); members without a record are *not assessed*.

Statistical component. With absence days $\mathcal{A}$ (from $\mathcal{P}_m$), business day exposures T_A , T_{A^c} and trade counts Y_A , Y_{A^c} , the Poisson rate ratio is $\hat{\rho}_6 = (Y_A/T_A)/(Y_{A^c}/T_{A^c})$. Conditional on $Y =$

$Y_A + Y_{A^c}$, $Y_A \sim \text{Binomial}(Y, \pi)$ with $\pi = \rho_6 T_A / (\rho_6 T_A + T_{A^c})$, and the Clopper Pearson limits $\pi^\pm$ map to an exact interval [21, 1]

$$\rho_6^\pm = \frac{T_{A^c}}{T_A} \cdot \frac{\pi^\pm}{1 - \pi^\pm}. \quad (6.11)$$

The same ratio is computed on trading days (days with any trade) as a robustness check against within batch overdispersion.

Absence periods. House roll call records (clerk.house.gov) give, for each member, the voting days on which every roll call was missed; runs of at least two consecutive such voting days form $\mathcal{A}$. Missing votes is a coarse proxy for reduced access to information: members campaign or travel while trading through a phone. Without an arrangement record the rate ratio is therefore reported as context and the test stays *not assessed*.

Decision rule. *Does not fit* if the record assigns discretion to others and either no absence period exists or the rate ratio interval includes one and excludes values below 1/2. *Fits* if the member is shown to direct trades and T4 or T5 fits. Otherwise *open question*.

6.7 T7: Disclosure timeliness

Signature. Delayed or missing disclosure can conceal well timed trades. The STOCK Act requires a report within 30 days of notice and no later than 45 days after the trade [54].

Null hypothesis. Filing lags are within the statutory limit.

Primary statistic. The late share $\hat{\pi}_L$, the distribution of filing lags, and the late minus timely signed CAR.

Definitions. The filing lag is $\ell_i = f_i - t_i$ in calendar days. Trade i is late if $\ell_i > 45$, unless the report states a notification date $n_i \geq t_i$ and was filed within 30 days of it ($f_i - n_i \leq 30$). That is the statute's clock for trades by a spouse, dependent or adviser that the member learns of after the fact. Notification dates are self reported (and read by OCR on scanned filings), so the trade date only rule $\ell_i > 45$ is reported as a robustness variant. The late share $\hat{\pi}_L$ carries a Wilson interval and is computed at the filing level as well as the trade level, because a single late filing can contain many trades. The median and 90th percentile lag summarize the distribution. Lines from amended reports are excluded from the late share, because an amendment's filing date is when the correction was filed, not when the trade was first disclosed; in the House data 84% of amendment lines would otherwise count as late. Their lags are right censored, and the Kaplan Meier estimator [25] of time to disclosure handles them where the original report can be linked.

Linking timeliness to returns. Concealment matters only if late trades are informed. We estimate $\theta_4^{\text{late}} - \theta_4^{\text{timely}}$ with the date clustered bootstrap; its one sided p value enters the multiplicity layer.

Worked example 6.2 (lower bound from partial records). If public records document at least eight late trades and the member made 189 trades in office, the late share is at least $8/189 = 4.2\%$, with a 95% Wilson interval of $[2.2\%, 8.1\%]$.

Decision rule. *Does not fit* if no trade was late. *Open question* if any trade was late. *Fits* only if late trades earn significantly more than timely trades after adjustment.

6.8 T8: Committee overlap

Signature. Members may trade on information about industries their committees oversee [46, 47, 19].

Null hypothesis. Trading in committee jurisdiction industries is no larger than a benchmark implies, and in jurisdiction trades earn no more than other trades.

Primary statistic. The jurisdiction excess share Δ_8 and the in minus out of jurisdiction signed CAR premium γ_1 .

Definitions. Let $J_i = 1$ if trade i 's issuer has a GICS sector or subindustry in $\kappa(c)$ for some committee c on which member m sits (Section 3.5). Among trades with a mapped industry, the dollar share in jurisdiction is $s_J(\mathbf{v}) = \sum_i v_i J_i / \sum_i v_i$, with sharp bounds obtained by setting in jurisdiction trades to U_i and others to L_i (upper) and the reverse (lower). The benchmark is

$$s_J^* = \max(s_J^{\text{S\&P}}, s_J^{\text{Congress}}), \quad (6.12)$$

the larger of the share of S&P 500 constituents and the share of all congressional trades that fall in the same jurisdiction, so that a member is not flagged for holding industries every member holds. The excess share is $\Delta_8 = s_J - s_J^*$ with a date clustered bootstrap interval.

Return premium. The regression $d_i \text{CAR}_i(1, 21) = \gamma_0 + \gamma_1 J_i + \boldsymbol{\eta}^\top \mathbf{x}_i + e_i$ with sector and month fixed effects $\mathbf{x}_i$ and standard errors clustered by trade date estimates the connected versus unconnected premium of [19]. Month effects are dropped when there are fewer than three trades per month.

Decision rule. *Fits* if $\hat{\gamma}_1 > 0$ with $p^{\text{Holm}} < \alpha$ and $q^{\text{BH}} < q^*$. *Does not fit* if the excess share interval does not lie above zero (no tilt) and $\gamma_1^+ < \delta_4$. *Open question* otherwise, including whenever holdings are tilted toward jurisdiction but the return test is inconclusive. Without prices, the descriptive form reads *open question* if the excess share is significantly positive and *does not fit* otherwise.

6.9 T9: Policy topic overlap

Signature. Members may trade on knowledge of legislation or policy they are shaping, such as a bill's content before release or a vote outcome before it is public.

Null hypothesis. Trades in topic stocks are unrelated to the timing of the member's policy actions.

Primary statistic. The pre event signed CAR of topic stock trades, and the pre trade drift that separates trading on news from trading ahead of it.

Definitions. A legislative event $e \in \mathcal{E}^{\text{leg}}$ carries a date, a committee and either a list of affected sectors or a policy topic φ with issuers $\mathcal{J}_\varphi$. The member's events are those of their committees or those the member sponsored. Topic trades are trades in affected issuers within 30 calendar days before the event; each trade counts once however many events follow it. The calendar is built from public records (python -m ingest.legislative): GovInfo bill status files give, for every bill and joint resolution since the 113th Congress, the committees of referral, committee markup and report dates, chamber passage and enactment. Each milestone is an event of every committee the bill was referred to, and its affected issuers are the sectors and subindustries of that committee in the

jurisdiction map of T8. This is coarser than a bill level topic map: a large committee generates many events, so T9 then approaches a timed version of T8.

Pre event abnormal return and drift. For each topic trade, $\theta_{9,i} = d_i \text{CAR}_i(1, 21)$, and $\hat{\theta}_9$ is their mean. Because a member with general skill earns positive signed CARs everywhere, the test statistic is the timing contrast $\hat{\Delta}_9 = \hat{\theta}_9 - \bar{\theta}_{-9}$ against the member's other priced trades, with a date clustered BCa bootstrap interval and one sided bootstrap p value; at least ten trades are required in each group, and the pre event trades must fall on at least ten distinct dates, since a bootstrap over two or three batch dates understates the uncertainty. As a placebo, the same contrast is computed for trades in the affected industries in the 30 days *after* an event: if those also beat the member's other trades, the premium belongs to the industry rather than to the timing.

Robustness. Windows of 15 and 60 days and the exclusion of the largest batch date are reported next to the main reading. The pre trade drift $\psi_i = d_i \sum_{\tau=-10}^{-1} \text{AR}_{i,\tau}$ is positive when a sale follows a price drop (or a purchase follows a rise); a large positive ψ with $\theta_9 \approx 0$ is the pattern of reacting to public news.

Worked example 6.3 (a sale after public news). Suppose a member sells a stock two days after the company suspended guidance and its shares fell about 16% in one day, and a week before a related floor vote. The pre trade drift ψ_i is large and positive (a sale after a public drop), and the drop is also an entry in $\mathcal{E}_j^{\text{firm}}$, so T3 classifies the trade as post news rather than pre news.

Decision rule. *Fits* if $\hat{\theta}_9 > 0$ and $\hat{\Delta}_9 > 0$ with $p^{\text{Holm}} < \alpha$ and $q^{\text{BH}} < q^*$, the placebo does not fit, and mean ψ is not significantly positive (one sided t test at α), so the gains are not explained by trading after public moves. *Does not fit* if $\hat{\Delta}_9^+ < \delta_4$. Otherwise *open question*; *not assessed* without a legislative calendar, with fewer than ten trades in either group, or with pre event trades on fewer than ten dates.

6.10 T10: Behavior after scrutiny

Signature. A member who trades on information has an incentive to continue trading and to resist arrangements that remove control; a member who does not has little cost in stopping or in adopting a blind trust.

Null hypothesis. Trading intensity is unchanged by scrutiny or by changes in arrangements.

Primary statistic. The change in trading intensity at a known date, and an estimated change point.

Known date test. For a scrutiny date τ^* from $\mathcal{S}_m$, the rate ratio of trades after versus before, ρ_{10} , uses the exact conditional binomial interval of T6.

Unknown change point. With daily trade day indicators n_t over the member's span, the Poisson likelihood ratio for a single change at τ is

$$\Lambda(\tau) = 2 \left[Y_{\leq \tau} \ln \frac{Y_{\leq \tau}}{\tau} + Y_{> \tau} \ln \frac{Y_{> \tau}}{T - \tau} - Y \ln \frac{Y}{T} \right], \quad (6.13)$$

$\hat{\tau} = \arg \max_{\tau} \Lambda(\tau)$, and the significance of $\max_{\tau} \Lambda(\tau)$ is assessed by permuting days [38]. The change point is reported for every member with ten or more trades, so the dashboard can show when trading stopped or started.

Scrutiny record. Scrutiny dates come from a curated file (`data/events/scrutiny_events.csv`): ethics complaints (for example by the Campaign Legal Center), Office of Congressional Ethics referrals, and press investigations from nonpartisan outlets or reports of them, each with a source link. The earliest documented date is used; it can lag the first public report, and a tally that lists a member is dated by the tally. The comparison needs at least 60 business days of the member’s trading history on each side of the date; otherwise the test is *not assessed*. Members’ public statements about who directs their trades (“a financial adviser makes the trades”) are recorded in `data/events/arrangements.csv` for T6; they are statements, not audited arrangements.

Decision rule. *Does not fit* if trading fell sharply after scrutiny ($\rho_{10}^+ < 1/2$) or discretion was transferred to a qualified blind trust. *Open question* if trading continued without a blind trust. *Not assessed* without a documented scrutiny date or arrangement. This test cannot produce *fits*.

6.11 T11: Performance relative to Congress

Signature. A member with an information edge stands out from the cross section of congressional traders, not only from zero [26, 19].

Null hypothesis. The member’s signed abnormal return is drawn from the same distribution as other members’.

Primary statistic. The empirical Bayes posterior mean of the member’s θ_4 and its percentile among members.

Model. Let $\hat{\theta}_m$ be member m ’s T4 estimate with Kolari Pynnönen standard error s_m , and suppose $\hat{\theta}_m \mid \theta_m \sim N(\theta_m, s_m^2)$ and $\theta_m \sim N(\mu, \tau^2)$ across members. The precision weighted mean and the method of moments between member variance are

$$\hat{\mu} = \frac{\sum_m s_m^{-2} \hat{\theta}_m}{\sum_m s_m^{-2}}, \quad \hat{\tau}^2 = \max\left(0, \widehat{\text{Var}}(\hat{\theta}_m) - \overline{s_m^2}\right), \quad (6.14)$$

and the posterior for each member [36, 16] is

$$\tilde{\theta}_m = B_m \hat{\mu} + (1 - B_m) \hat{\theta}_m, \quad \tilde{s}_m^2 = (1 - B_m) s_m^2, \quad B_m = \frac{s_m^2}{s_m^2 + \hat{\tau}^2}. \quad (6.15)$$

When $\hat{\tau}^2 = 0$ every member is shrunk fully to $\hat{\mu}$, which is the honest result when the cross section shows no more dispersion than noise.

Decision rule. *Fits* if $\tilde{\theta}_m - 1.96 \tilde{s}_m > 0$. *Does not fit* if $\tilde{\theta}_m + 1.645 \tilde{s}_m < \delta_4$. Otherwise *open question*. Because shrinkage already accounts for the cross section, T11 does not enter the Holm or BH layers. The dashboard’s “top performers” ranking uses $\tilde{\theta}_m$, not $\hat{\theta}_m$, so that members with few lucky trades do not top the list.

6.12 T12: Anticipation of large stock moves

Signature. Buying shortly before a large jump or selling shortly before a large drop, the pattern behind most SEC insider trading actions [34, 53].

Null hypothesis. Among trades followed by a large abnormal move, the direction of the trade is independent of the direction of the move.

Primary statistic. The hit rate among trades followed within ten trading days by an abnormal move beyond three estimation window standard deviations.

Definitions. For trade i let $\tau_i^* = \min\{\tau \in \{1, \dots, 10\} : |\text{AR}_{i,\tau}| > 3\hat{\sigma}_{j(i)}\}$ when it exists. The qualifying set is $\mathcal{Q} = \{i : \tau_i^* \text{ exists}\}$ with $n = |\mathcal{Q}|$, and the hit count is $X = \sum_{i \in \mathcal{Q}} \mathbf{1}\{d_i \text{AR}_{i,\tau_i^*} > 0\}$.

Inference. Under H_0 , $X \sim \text{Binomial}(n, 1/2)$; the one sided exact p value and the Clopper Pearson interval follow as in T5. Unlike T4, which averages over all trades and can be diluted by many routine ones, T12 concentrates on the few trades where an informational edge would pay most. Unlike T3, which asks whether trades come before news more often than chance, T12 asks whether they are on the right side of it.

Decision rule. *Fits* if $\hat{h} > 1/2$ with $p^{\text{Holm}} < \alpha$ and $q^{\text{BH}} < q^*$. *Does not fit* if $\hat{h}^+ < 1/2 + \delta_{12}$ with $\delta_{12} = 0.10$. Otherwise *open question*; *not assessed* when $n = 0$.

6.13 T13: Cotrading with other members

Signature. Inside information travels through networks and shows up as clustered trades by connected traders [2]. Members share committees, briefings and staff.

Null hypothesis. The timing of the member's trades in each issuer is independent of other members' trades in that issuer.

Primary statistic. The cotraded share c_m : the share of the member's trades for which another member traded the same issuer in the same direction within five business days.

Definitions. With business day positions b_i , trade i of member m is cotraded if

$$C_i = \mathbf{1}\{\exists i' : m(i') \neq m, j(i') = j(i), d_{i'} = d_i, |b_{i'} - b_i| \leq 5\}, \quad c_m = \frac{1}{N_m} \sum_{i \in \mathcal{I}_m} C_i. \quad (6.16)$$

Inference. Popular stocks are traded by many members, so a high c_m alone means little. The null distribution is obtained by circularly shifting member m 's trade dates (Algorithm 3) while holding every other member's trades fixed, which preserves the member's issuer mix and batching and the popularity of each issuer. The excess p value is p^{ex} . Where prices are available, the signed CAR of the cotraded trades is tested as in T4, and its p value enters the multiplicity layer.

Decision rule. *Fits* if cotrading is in excess ($p^{\text{ex}} < \alpha$) and cotraded trades earn significantly positive signed returns after Holm and BH adjustment. *Open question* if cotrading is in excess but returns are not significant. *Does not fit* otherwise. Excess cotrading is a pattern, not a channel: it is also what common reactions to public news, shared financial advisers or shared briefings would produce, and the dashboard lists the cotrading members and issuers so the pattern can be examined.

6.14 T14: Family account trades

Signature. Routing informed trades through a spouse's or dependent's account, which members may believe draws less attention.

Null hypothesis. Signed abnormal returns in family accounts have mean zero and do not exceed those in the member's own accounts.

Primary statistic. The mean signed CAR of family account trades, and its difference from self account trades.

Definitions. Family trades are those with $o_i \in \{\text{spouse, dependent}\}$; joint accounts are treated as the member's own. The family parameter θ_{14} is θ_4 computed on family trades with the Kolari Pynnönen interval, and the difference $\Delta_{14} = \theta_4^{\text{family}} - \theta_4^{\text{self}}$ is estimated with the date clustered bootstrap.

Decision rule. *Fits* if $\theta_{14}^- > 0$ with $p^{\text{Holm}} < \alpha$ and $q^{\text{BH}} < q^*$, and the difference is significantly positive. *Does not fit* if $\theta_{14}^+ < \delta_4$. Otherwise *open question*; *not assessed* with fewer than ten priced family trades. Many spouses are investment professionals trading on their own account, which the reading does not distinguish.

6.15 T15: Short swing round trips

Signature. A temporary informational edge is captured by opening and closing a position before the information becomes public; Section 16(b) of the Exchange Act treats six months as the horizon for such profits by corporate insiders [50].

Null hypothesis. Round trip returns in excess of the market have mean zero.

Primary statistic. The mean market adjusted return between the opening and closing trade of each round trip.

Matching. Trades in the same issuer are sorted by date and matched greedily first in: each unmatched trade is paired with the earliest later unmatched trade in the opposite direction within 183 calendar days. A purchase followed by a sale is a long round trip; a sale followed by a repurchase is a short round trip. Bands make exact share matching impossible, so a partial sale may close part of a position; the statistic is per round trip, not per dollar.

Estimator. The natural statistic, the market adjusted return realized between the opening and closing trade, is biased upward: the closing date is chosen after seeing the price path, and investors close winning positions sooner than losing ones [39], so round trips that close within six months are disproportionately winners even when the opening was uninformed. In the first House run the realized statistic flagged seven members, with mean realized returns of 1% to 28% per round trip, none of whose opening trades showed a significant edge at a fixed horizon. The test therefore uses the fixed horizon signed CAR of the opening trades,

$$\hat{\theta}_{15} = \frac{1}{K} \sum_k d_{o,k} \text{CAR}_{o,k}(1, 21), \quad (6.17)$$

with the Kolari Pynnönen interval of T4. It asks whether the trades that the member went on to reverse quickly were well timed when made, which is the short swing signature, without conditioning

on the exit. The realized return $g_k = d_{o,k} \sum_{t=t_{o,k}+1}^{t_{c,k}} (R_{j,t} - R_{m,t})$ and the share of winning round trips are reported for context only.

Decision rule. *Fits* if $\hat{\theta}_{15} > 0$ with $p^{\text{Holm}} < \alpha$ and $q^{\text{BH}} < q^*$. *Does not fit* if there are no round trips or $\hat{\theta}_{15}^+ < \delta_4$. Otherwise *open question*.

6.16 T16: Options and leverage

Signature. Informed traders favour options because leverage raises the payoff to private information, and informed options trading before takeovers is well documented [3].

Null hypothesis. Option trades earn no signed abnormal return on the underlying.

Primary statistic. The option share of trades and the mean signed 21 day CAR of the underlying after option trades.

Definitions. Option trades are those with $a_i = \text{option}$; the underlying ticker is used for returns, and the direction is the disclosed transaction type. Disclosures rarely state whether an option was bought or written, or whether it was a call or a put, so the direction is approximate and the test is conservative: misclassified directions bias $\hat{\theta}_{16}$ toward zero.

Decision rule. *Does not fit* if the member reported no option trades or $\theta_{16}^+ < \delta_4$. *Fits* if $\theta_{16}^- > 0$ with $p^{\text{Holm}} < \alpha$ and $q^{\text{BH}} < q^*$. Otherwise *open question*. Without prices, any option use is an *open question* with descriptive evidence.

7 KPIs

The dashboard reports two layers of indicators. Member KPIs summarize one record; Congress KPIs summarize the population. Every KPI is computed by the pipeline in Section 9 and is published with its interval where one exists.

The Congress signed CAR uses the date clustered bootstrap, which allows for correlation among trades on the same day but not across nearby days; the calendar time alpha is the estimate that is robust to overlapping windows and is the one to cite.

8 Aggregation and Reporting

8.1 The scorecard

The sixteen readings are reported side by side with their statistics, intervals, adjusted p values, coverage and sources. We do not combine them into a single probability of insider trading, for three reasons. First, the tests are not independent: T1, T2 and T3 describe the same portfolio from different angles, and T11 is a transformation of T4. Second, a single number would hide which evidence drives the result. Third, the readings answer a question about resemblance to known patterns, and a posterior probability would require a prior on misconduct that the data cannot supply. The dashboard’s member headline is generated mechanically from the readings: the tests that fit, or else the tests that are open questions, or else a statement that no test found the signature.

Table 2: Member level KPIs.

KPI	Definition	Source
Trades, issuers, volume	$N_m, \mathcal{J}_m , [\sum L_i, \sum U_i]$	Dataset
Readings profile	Counts of fits, open question (excluding power limited), does not fit, not assessed	All tests
Flagged tests	Tests reading <i>fits</i> , or <i>open question</i> that is not power limited	All tests
Mean signed 21 day CAR	$\hat{\theta}_4$ with KP interval; dollar weighted identified set	T4
Shrunk performance and percentile	$\tilde{\theta}_m, \tilde{s}_m$, percentile among members	T11
Calendar time alpha	Annualized $\hat{a}$ with Newey West interval	T4
Large move hit rate	X/n with Clopper Pearson interval	T12
Late disclosure share	$\hat{\pi}_L$ (trades and filings), median and 90th percentile lag	T7
Committee excess share	Δ_8 with bootstrap interval and identified set	T8
Concentration	N_{eff} with identified set, sector entropy	T2
Cotraded share	c_m against its null mean	T13
Family account share and return	Share of trades; θ_{14}	T14
Round trips	Count, share of trades, $\hat{\theta}_{15}$, realized $\bar{g}$	T15
Option share	Share of trades in options; θ_{16}	T16
Activity change point	$\hat{\tau}$ and rates before and after	T10

Table 3: Congress level KPIs.

KPI	Definition	Source
Coverage	Members total, members trading, trades, volume bounds	Dataset
Congress signed CAR	Mean $d_i \text{CAR}_i(1, H)$ over all priced trades, $H \in \{5, 21, 63\}$, date clustered 95% interval	T4
Congress calendar time alpha	Alpha of the pooled long purchases, short sales portfolio	T4
Late filing share	Share of all trades disclosed after 45 days	T7
Readings by test	Members per reading for each test, with the power limited count	All tests
Members flagged	Members with at least one flagged test; members with at least one <i>fits</i>	All tests
By party and chamber	Members trading, trades, flagged members, members with <i>fits</i>	All tests
Monthly flow	Purchases and sales per month	Dataset
Rankings	Top 20 members (10+ trades) by $\hat{\theta}_4, \tilde{\theta}_m, \hat{\pi}_L$ and c_m , with intervals	T4, T11, T7, T13

8.2 Robustness

Every return based reading is rerun under the variants listed for T4. A reading is reported as robust only if it is unchanged across them; the variants and their readings are published with each test.

Algorithm 5 Scorecard construction for all members

Require: Trade table; market data; event calendars; documentary record; margins $\{\delta_k\}$; levels α , q^*

- 1: Run Algorithm 1; attach sectors and jurisdiction flags; run Algorithm 4
 - 2: Build peer quantities: pooled issuer distribution, sector shares, cotrading index
 - 3: **for** each member m **do**
 - 4: **for** $k = 1$ to 16, $k \neq 11$ **do**
 - 5: Compute statistic, interval and raw superiority p value p_{mk} (Section 6); record coverage
 - 6: **end for**
 - 7: **end for**
 - 8: T11: empirical Bayes over all members with a T4 estimate
 - 9: Holm across k within each m ; Benjamini Hochberg across m within each k
 - 10: **for** each member m and $k = 1$ to 16 **do**
 - 11: Assign the reading by the decision rule of test k ; mark power limited open questions
 - 12: Rerun robustness variants; flag the reading as robust or sensitive
 - 13: **end for**
 - 14: Compute Congress KPIs; write scorecards, KPIs and summaries (Section 9)
-

9 Implementation and Validation

9.1 Code

The methodology is implemented in the Python package `analysis/citests` in the project repository. Table 4 maps each part of this document to its code.

Table 4: Where each estimator is computed.

Section	Module and function
Data model, Algorithm 1	<code>data.py</code> : <code>load</code> , <code>normalize_trades</code> , <code>attach_sectors</code> , <code>in_jurisdiction</code>
Intervals, multiplicity, bootstrap, permutation, shrinkage	<code>stats.py</code> : <code>wilson</code> , <code>clopper_pearson</code> , <code>rate_ratio_exact</code> , <code>holm</code> , <code>bh</code> , <code>cluster_bootstrap</code> , <code>empirical_bayes_normal</code>
Event study, calendar time alpha, Algorithm 4	<code>market.py</code> : <code>event_study</code> , <code>mean_residual_corr</code> , <code>calendar_time_alpha</code> , <code>market_events</code>
Tests T1 to T16	<code>battery.py</code> : one function per test, <code>t1_size</code> to <code>t16_options</code>
Firm events, Algorithm 5, KPIs	<code>pipeline.py</code> : <code>FirmEvents</code> , <code>run</code> , <code>cross_member_t11</code> , <code>adjust_and_decide</code> , <code>congress_summary</code>
Test descriptions for the dashboard	<code>catalog.py</code> (written to <code>data/site/tests.json</code>)

The pipeline runs with `python -m citests.pipeline -root .` from the `analysis` directory and

writes: `analysis/output/results.json` (every statistic for every member and test), `kpis.csv` (one row per member), `trade_metrics.csv` (per trade CARs, sectors and jurisdiction flags) and `summary.json`; and, for the dashboard, `data/site/tests.json`, `summary.json`, and the readings merged into `members.json` and `members/<id>.json` following the site data contract. Random seeds are fixed so that results are reproducible.

9.2 Validation on synthetic data with known answers

The test suite (`analysis/tests`) checks each primitive against reference values (including the worked examples in this document: the Wilson interval [2.2%, 8.1%] for 8 of 189, and the Clopper Pearson interval [0, 0.842] for 0 of 2), checks that the HHI and dollar weighted bounds bracket every feasible amount vector, and runs the full pipeline on a synthetic Congress with known ground truth: an informed member who trades ahead of 21 day abnormal moves, an adviser run account that rebalances randomly in monthly batches, a small random trader, and a member whose spouse’s trades are informed and own trades are not. The suite requires that the informed member *fits* T4 and T11, that the family account member *fits* T14, and that neither uninformed member *fits* any test.

10 Application Across Congress

Table 5 states, for each test, the form in which it is applied Congress wide and which inputs determine coverage. “Statistical” means the estimator and its interval are computed for every member with enough trades; “curated” means the test depends on a documentary record that is assembled member by member.

Table 5: Form of application across all members.

Test	Form	Coverage
T1 Size	Statistical	Members with 10+ trades; event gap needs prices
T2 Breadth	Statistical	Members with 10+ trades
T3 Cadence	Statistical	Members with 10+ trades and prices
T4 Timing	Statistical	Members with 10+ priced trades
T5 Shocks	Statistical	Members trading before market wide shocks
T6 Control	Curated	Members with documented arrangements or absences
T7 Timeliness	Statistical	All members with filing dates
T8 Committee	Statistical	Members with committees and 10+ mapped trades
T9 Topic	Curated	Members with legislative events in the calendar
T10 Scrutiny	Curated	Members with a documented scrutiny date
T11 Relative	Statistical	All members with a T4 estimate
T12 Large moves	Statistical	Members with priced trades followed by large moves
T13 Cotrading	Statistical	Members with 10+ trades
T14 Family	Statistical	Members with 10+ priced family trades
T15 Round trips	Statistical	Members with round trips and prices
T16 Options	Statistical	All members; returns for option trades with prices

Reading the published scorecards. At Congress scale most members trade too little for the return based tests to reach either *fits* or *does not fit*, and their open questions are marked power limited. A *fits* reading has survived correction for sixteen tests on the member and for 535 members on the test; with $q^* = 0.10$, about one in ten members flagged as *fits* on a given test is expected to be a false discovery. Every reading is a statement about patterns in public records, not a finding about any member's knowledge or intent.

11 Limitations

- **Amounts are censored.** Bands as wide as \$1,001 to \$15,000 limit dollar weighted statistics to identified sets; conclusions that depend on amounts are reported as bounds.
- **Dates, not times.** Disclosures give the transaction date, not the execution time or price, so returns are measured from the next close; same day information cannot be tested.
- **Low power for small samples.** Most members trade too little for return tests to reach a conclusion (Section 4.5); a clean result for a small set is weak evidence, which is why the equivalence criterion is required for *does not fit* and why power limited open questions are separated from substantive ones.
- **Multiplicity is controlled, not eliminated.** The FDR target allows about one false *fits* in ten per test; the tests are correlated, so the expected number of members with at least one false *fits* is larger than for any one test.
- **Information is unobserved.** The tests detect patterns; they cannot show what a member knew. A member who traded on information rarely and in small size could pass every test.
- **Event calendars are proxies or incomplete.** Firm events are proxied by large abnormal moves; legislative information arrives through meetings and briefings that leave no public date.
- **Committee assignments are current.** The public roster records current assignments; a member's earlier committees are not in the map, which biases T8 toward finding no overlap for trades made before an assignment changed.
- **The jurisdiction map is coarse.** It works at GICS sector or subindustry level; jurisdiction over specific programs (for example a defense contract) is narrower than an industry.
- **Coverage of securities.** Funds, bonds, private companies and unlisted assets are excluded from return tests; issuers without a usable price history are excluded from T4 and the tests that depend on it.
- **Stated arrangements are not verified.** Discretionary management described in letters to the Ethics Committee is a representation by counsel, not an audited fact.
- **Market side evidence is excluded.** Order level data, options activity and broker records used by the SEC [51, 3] are not available in a public record review.

A Formula Reference

Wilson interval for x successes in n trials at level $1 - \alpha$, $z = z_{1-\alpha/2}$:

$$\frac{\hat{p} + \frac{z^2}{2n} \pm z \sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{z^2}{4n^2}}}{1 + \frac{z^2}{n}}, \quad \hat{p} = x/n.$$

Clopper Pearson interval: $[B^{-1}(\alpha/2; x, n - x + 1), B^{-1}(1 - \alpha/2; x + 1, n - x)]$.

Exact Poisson rate ratio: with Y_A events in exposure T_A and Y_{A^c} in T_{A^c} , $\pi^\pm$ the Clopper Pearson limits for Y_A out of $Y_A + Y_{A^c}$, the limits are $\frac{T_{A^c}}{T_A} \frac{\pi^\pm}{1 - \pi^\pm}$.

Market model CAR variance over H event days with M estimation days:

$$\sigma_{\text{CAR}}^2 = \hat{\sigma}^2 \left(H + \frac{H^2}{M} + \frac{(\sum_{\tau=1}^H (R_{m,\tau} - \bar{R}_m))^2}{\sum_t (R_{m,t} - \bar{R}_m)^2} \right).$$

BMP statistic: $t_{\text{BMP}} = \frac{N^{-1} \sum_i \text{SCAR}_i}{\sqrt{\frac{1}{N(N-1)} \sum_i (\text{SCAR}_i - \overline{\text{SCAR}})^2}}.$

Kolari Pynnönen adjustment: $t_{\text{KP}} = t_{\text{BMP}} \sqrt{\frac{1 - \bar{r}}{1 + (N-1)\bar{r}}}.$

Newey West variance of the calendar time alpha with lag q : $\hat{\Omega} = \hat{\Gamma}_0 + \sum_{l=1}^q (1 - \frac{l}{q+1})(\hat{\Gamma}_l + \hat{\Gamma}_l^\top)$, $\hat{\Gamma}_l = T^{-1} \sum_t \hat{u}_t \hat{u}_{t-l}^\top \mathbf{z}_t \mathbf{z}_{t-l}^\top$, $\mathbf{z}_t = (1, \mathbf{F}_t^\top)^\top$.

Holm procedure: reject $H_{(1)}, \dots, H_{(k^*-1)}$ where $k^* = \min\{k : p_{(k)} > \alpha/(m - k + 1)\}$.

Benjamini Hochberg q value: $q_{(r)} = \min_{s \geq r} \{Mp_{(s)}/s\}$, capped at one.

TOST equivalence: conclude $\theta < \delta$ at level α if the upper limit of the $(1 - 2\alpha)$ interval is below δ [42].

Empirical Bayes posterior: $\tilde{\theta}_m = B_m \hat{\mu} + (1 - B_m) \hat{\theta}_m$, $\tilde{s}_m^2 = (1 - B_m) s_m^2$, $B_m = s_m^2 / (s_m^2 + \hat{\tau}^2)$.

HHI bounds: $\underline{\text{HHI}} = \min_\lambda \text{HHI}(\text{clip}(\lambda, \mathbf{a}, \mathbf{b}))$; $\overline{\text{HHI}}$ at the vertex with $V_j = b_j$ for $s_j > \text{HHI}$ and $V_j = a_j$ otherwise.

B Committee Jurisdiction Map

Table 6 summarizes the map used by T8. Each row of the full file cites the paragraph of House Rule X or Senate Rule XXV (or the select committee's charter) that grants jurisdiction.

Table 6: Committees and the industries mapped to them
(GICS sector in bold, otherwise subindustry).

Committee	Industries
House Agriculture (HSAG)	Agricultural Products & Services; Fertilizers & Agricultural Chemicals; Agricultural & Farm Machinery; Packaged Foods & Meats; Food Distributors; Timber REITs; Financial Exchanges & Data
House Armed Services (HSAS)	Aerospace & Defense
House Financial Services (HSBA)	Financials; Real Estate
House Energy and Commerce (HSIF)	Health Care; Energy; Utilities; Communication Services; Automobile Manufacturers; Interactive Media & Services; Tobacco
House Natural Resources (HSII)	Oil & Gas Exploration & Production; Integrated Oil & Gas; Gold; Copper; Timber REITs
House Transportation and Infrastructure (HSPW)	Passenger Airlines; Air Freight & Logistics; Rail Transportation; Cargo Ground Transportation; Passenger Ground Transportation; Construction & Engineering; Construction Materials; Construction Machinery & Heavy Transportation Equipment; Water Utilities; Environmental & Facilities Services; Hotels, Resorts & Cruise Lines
House Science Space and Technology (HSSY)	Aerospace & Defense; Semiconductors; Semiconductor Materials & Equipment; Biotechnology
House Homeland Security (HSHM)	Aerospace & Defense; Systems Software; Application Software
House Intelligence (HLIG)	Aerospace & Defense; Systems Software; IT Consulting & Other Services
House Ways and Means (HSWM)	Health Care; Tobacco; Distillers & Vintners; Life & Health Insurance
House Veterans' Affairs (HSVR)	Health Care Facilities; Health Care Services
House Education and Workforce (HSED)	Education Services; Human Resource & Employment Services
House Foreign Affairs (HSFA)	Aerospace & Defense; Semiconductors; Semiconductor Materials & Equipment
House Judiciary (HSJU)	Interactive Media & Services; Broadline Retail; Pharmaceuticals
House Small Business (HSSM)	Regional Banks
House Select Committee on the CCP (HSZS)	Semiconductors; Semiconductor Materials & Equipment; Technology Hardware, Storage & Peripherals
Senate Agriculture (SSAF)	Agricultural Products & Services; Fertilizers & Agricultural Chemicals; Agricultural & Farm Machinery; Packaged Foods & Meats; Food Distributors; Timber REITs; Financial Exchanges & Data

Committee	Industries
Senate Armed Services (SSAS)	Aerospace & Defense
Senate Banking (SSBK)	Financials; Real Estate; Aerospace & Defense
Senate Commerce (SSCM)	Communication Services; Passenger Airlines; Air Freight & Logistics; Rail Transportation; Cargo Ground Transportation; Automobile Manufacturers; Interactive Media & Services; Semiconductors; Aerospace & Defense
Senate Energy and Natural Resources (SSEG)	Energy; Utilities; Gold; Copper; Timber REITs
Senate Environment and Public Works (SSEV)	Construction & Engineering; Construction Materials; Environmental & Facilities Services; Water Utilities; Electric Utilities; Commodity Chemicals
Senate Finance (SSFI)	Health Care; Life & Health Insurance; Tobacco; Distillers & Vintners
Senate HELP (SSHR)	Health Care; Human Resource & Employment Services
Senate Foreign Relations (SSFR)	Aerospace & Defense
Senate Homeland Security and Governmental Affairs (SSGA)	Aerospace & Defense; Systems Software; IT Consulting & Other Services
Senate Intelligence (SLIN)	Aerospace & Defense; Systems Software
Senate Judiciary (SSJU)	Interactive Media & Services; Broadline Retail; Pharmaceuticals
Senate Veterans' Affairs (SSVA)	Health Care Facilities; Health Care Services
Senate Aging (SPAG)	Managed Health Care
Senate Small Business (SSSB)	Regional Banks

References

- [1] A. Agresti. *Categorical Data Analysis*, 3rd ed. Wiley, 2013.
- [2] K. R. Ahern. “Information Networks: Evidence from Illegal Insider Trading Tips.” *Journal of Financial Economics* 125(1):26–47, 2017.
- [3] P. Augustin, M. Brenner and M. G. Subrahmanyam. “Informed Options Trading Prior to Takeover Announcements: Insider Trading?” *Management Science* 65(12):5697–5720, 2019.
- [4] W. Belmont, B. Sacerdote, R. Sehgal and I. Van Hoek. “Do Senators and House Members Beat the Stock Market? Evidence from the STOCK Act.” *Journal of Public Economics* 207:104602, 2022.
- [5] Y. Benjamini and Y. Hochberg. “Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing.” *Journal of the Royal Statistical Society, Series B* 57(1):289–300, 1995.
- [6] E. Boehmer, J. Musumeci and A. B. Poulsen. “Event-Study Methodology under Conditions of Event-Induced Variance.” *Journal of Financial Economics* 30(2):253–272, 1991.
- [7] S. J. Brown and J. B. Warner. “Using Daily Stock Returns: The Case of Event Studies.” *Journal of Financial Economics* 14(1):3–31, 1985.

- [8] M. M. Carhart. “On Persistence in Mutual Fund Performance.” *Journal of Finance* 52(1):57–82, 1997.
- [9] C. J. Clopper and E. S. Pearson. “The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial.” *Biometrika* 26(4):404–413, 1934.
- [10] W. G. Cochran. “Some Methods for Strengthening the Common χ^2 Tests.” *Biometrics* 10(4):417–451, 1954.
- [11] L. Cohen, C. Malloy and L. Pomorski. “Decoding Inside Information.” *Journal of Finance* 67(3):1009–1043, 2012.
- [12] C. J. Corrado. “A Nonparametric Test for Abnormal Security-Price Performance in Event Studies.” *Journal of Financial Economics* 23(2):385–395, 1989.
- [13] D. Easley, N. M. Kiefer, M. O’Hara and J. B. Paperman. “Liquidity, Information, and Infrequently Traded Stocks.” *Journal of Finance* 51(4):1405–1436, 1996.
- [14] B. Efron. “Bootstrap Methods: Another Look at the Jackknife.” *Annals of Statistics* 7(1):1–26, 1979.
- [15] B. Efron. “Better Bootstrap Confidence Intervals.” *Journal of the American Statistical Association* 82(397):171–185, 1987.
- [16] B. Efron. *Large-Scale Inference: Empirical Bayes Methods for Estimation, Testing, and Prediction*. Cambridge University Press, 2010.
- [17] B. Efron and R. J. Tibshirani. *An Introduction to the Bootstrap*. Chapman & Hall, 1993.
- [18] A. C. Eggers and J. Hainmueller. “Capitol Losses: The Mediocre Performance of Congressional Stock Portfolios, 2004–2008.” *Journal of Politics* 75(2):535–551, 2013.
- [19] A. C. Eggers and J. Hainmueller. “Political Capital: Corporate Connections and Stock Investments in the U.S. Congress, 2004–2008.” *Quarterly Journal of Political Science* 9(2):169–202, 2014.
- [20] E. F. Fama and K. R. French. “Common Risk Factors in the Returns on Stocks and Bonds.” *Journal of Financial Economics* 33(1):3–56, 1993.
- [21] F. Garwood. “Fiducial Limits for the Poisson Distribution.” *Biometrika* 28(3/4):437–442, 1936.
- [22] S. Holm. “A Simple Sequentially Rejective Multiple Test Procedure.” *Scandinavian Journal of Statistics* 6(2):65–70, 1979.
- [23] J. F. Jaffe. “Special Information and Insider Trading.” *Journal of Business* 47(3):410–428, 1974.
- [24] L. A. Jeng, A. Metrick and R. Zeckhauser. “Estimating the Returns to Insider Trading: A Performance-Evaluation Perspective.” *Review of Economics and Statistics* 85(2):453–471, 2003.
- [25] E. L. Kaplan and P. Meier. “Nonparametric Estimation from Incomplete Observations.” *Journal of the American Statistical Association* 53(282):457–481, 1958.
- [26] S. Karadas. “Trading on Private Information: Evidence from Members of Congress.” *Financial Review* 54(1):85–131, 2019.
- [27] J. W. Kolari and S. Pynnönen. “Event Study Testing with Cross-Sectional Correlation of Abnormal Returns.” *Review of Financial Studies* 23(11):3996–4025, 2010.
- [28] J. W. Kolari and S. Pynnönen. “Nonparametric Rank Tests for Event Studies.” *Journal of Empirical Finance* 18(5):953–971, 2011.
- [29] A. S. Kyle. “Continuous Auctions and Insider Trading.” *Econometrica* 53(6):1315–1335, 1985.
- [30] J. Lakonishok and I. Lee. “Are Insider Trades Informative?” *Review of Financial Studies* 14(1):79–111, 2001.
- [31] J. D. Lyon, B. M. Barber and C.-L. Tsai. “Improved Methods for Tests of Long-Run Abnormal Stock Returns.” *Journal of Finance* 54(1):165–201, 1999.

- [32] A. C. MacKinlay. “Event Studies in Economics and Finance.” *Journal of Economic Literature* 35(1):13–39, 1997.
- [33] C. F. Manski. *Partial Identification of Probability Distributions*. Springer, 2003.
- [34] L. K. Meulbroek. “An Empirical Analysis of Illegal Insider Trading.” *Journal of Finance* 47(5):1661–1699, 1992.
- [35] M. L. Mitchell and E. Stafford. “Managerial Decisions and Long-Term Stock Price Performance.” *Journal of Business* 73(3):287–329, 2000.
- [36] C. N. Morris. “Parametric Empirical Bayes Inference: Theory and Applications.” *Journal of the American Statistical Association* 78(381):47–55, 1983.
- [37] W. K. Newey and K. D. West. “A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix.” *Econometrica* 55(3):703–708, 1987.
- [38] E. S. Page. “Continuous Inspection Schemes.” *Biometrika* 41(1/2):100–115, 1954.
- [39] T. Odean. “Are Investors Reluctant to Realize Their Losses?” *Journal of Finance* 53(5):1775–1798, 1998.
- [40] J. M. Patell. “Corporate Forecasts of Earnings per Share and Stock Price Behavior: Empirical Tests.” *Journal of Accounting Research* 14(2):246–276, 1976.
- [41] S. A. Rhoades. “The Herfindahl-Hirschman Index.” *Federal Reserve Bulletin* 79:188–189, March 1993.
- [42] D. J. Schuirmann. “A Comparison of the Two One-Sided Tests Procedure and the Power Approach for Assessing the Equivalence of Average Bioavailability.” *Journal of Pharmacokinetics and Biopharmaceutics* 15(6):657–680, 1987.
- [43] H. N. Seyhun. “Insiders’ Profits, Costs of Trading, and Market Efficiency.” *Journal of Financial Economics* 16(2):189–212, 1986.
- [44] C. E. Shannon. “A Mathematical Theory of Communication.” *Bell System Technical Journal* 27(3):379–423, 1948.
- [45] E. B. Wilson. “Probable Inference, the Law of Succession, and Statistical Inference.” *Journal of the American Statistical Association* 22(158):209–212, 1927.
- [46] 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* 39(4):661–676, 2004.
- [47] 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* 13(1), 2011. doi:10.2202/1469-3569.1308.
- [48] U.S. Securities and Exchange Commission. Rule 10b5-1, “Trading on the basis of material nonpublic information in insider trading cases,” 17 C.F.R. § 240.10b5-1.
<https://www.ecfr.gov/current/title-17/section-240.10b5-1>
- [49] U.S. Securities and Exchange Commission. *Insider Trading Arrangements and Related Disclosures*, Final Rule, Release No. 33-11138, 87 Fed. Reg. 80362 (December 29, 2022).
- [50] Securities Exchange Act of 1934, § 16(b), 15 U.S.C. § 78p(b).
- [51] U.S. Securities and Exchange Commission. “SEC Files Multiple Insider Trading Actions Originating from the Market Abuse Unit’s Analysis and Detection Center,” Press Release 2022-129, July 25, 2022.
<https://www.sec.gov/news/press-release/2022-129>
- [52] U.S. Securities and Exchange Commission. Press Release 2023-204, September 28, 2023.
<https://www.sec.gov/news/press-release/2023-204>

- [53] U.S. Securities and Exchange Commission. *SEC v. Stephen Buyer*, complaint filed July 25, 2022 (S.D.N.Y.). Press Release 2022-128. <https://www.sec.gov/news/press-release/2022-128>
- [54] Stop Trading on Congressional Knowledge (STOCK) Act of 2012, Pub. L. No. 112-105, 126 Stat. 291.
- [55] *Dirks v. SEC*, 463 U.S. 646 (1983).
- [56] *United States v. O'Hagan*, 521 U.S. 642 (1997).