

The Geometry of Partisanship: Graph Attention Networks for Predicting Legislative Defection in the U.S. House, 1995–2024

Abstract

When a legislator breaks with their party, it reveals something about the structure of Congress that roll-call statistics alone cannot capture. We construct temporal voting networks for every session of the U.S. House from the 104th through 118th Congress (1995–2024) and apply Graph Attention Networks (GATs) to predict which members will defect from their party line. The approach works: our model achieves an F1 of 0.75 and AUC of 0.88 at the 5% defection threshold, outperforming logistic regression (F1=0.60), random forests (F1=0.63), and naive persistence baselines (F1=0.59). More interesting than the prediction itself is what the model’s attention mechanism reveals. The network learns to weight same-party connections roughly 200 times more heavily than cross-party ties, a ratio that has remained remarkably stable even as the parties themselves have polarized. Spectral decomposition of the voting networks confirms what students of Congress have long suspected: the second-smallest eigenvalue of the graph Laplacian correlates with DW-NOMINATE scores at $r > 0.97$ in most sessions, meaning the dominant structure of congressional voting is recoverable from network topology alone. We document a 30-percentage-point decline in cross-party agreement from the 104th to the 118th Congress and show, using difference-in-differences analysis, that the Tea Party wave of 2010 produced a discrete downward shift in bipartisan cooperation that subsequent sessions never reversed.

1 Introduction

Start with a fact: in the 104th Congress (1995–96), roughly half of all House members voted against their party more than 10% of the time. By the 115th Congress (2017–18), that number had fallen to 6.7%. The disappearance of the party-line crosser is one of the defining features of contemporary American politics, yet the mechanisms that produce this sorting remain incompletely understood. Do members defect less because the parties have become more ideologically coherent, because institutional pressures have intensified, or because the kinds of people who run for Congress have changed? The answer is probably all three, but separating these channels requires a framework that captures the relational structure of legislative behavior, not just individual-level covariates.

This paper develops that framework. We model each session of the U.S. House as a graph in which nodes represent members and edges encode voting agreement, then deploy Graph Attention Networks to learn which structural features of the network predict who will break ranks. The intuition is straightforward: a legislator’s propensity to defect depends not only on their own ideology and party, but on the voting behavior of the colleagues they are most connected to. A moderate Republican surrounded by hard-liners faces different incentive structures than one embedded in a bipartisan cluster. Standard regression cannot capture this relational context. Graph neural networks can.

The paper makes three contributions. First, we establish that the spectral structure of congressional voting networks tracks ideological positioning with extraordinary fidelity. The Fiedler vector of the normalized Laplacian correlates with first-dimension DW-NOMINATE scores at $r > 0.97$ for most sessions between

the 104th and 118th Congresses. This is not a trivial finding. It means the dominant axis of congressional ideology is fully recoverable from the network topology of who-votes-with-whom, without any reference to ideological scaling.

Second, we show that a two-layer GAT with 15-nearest-neighbor graph construction outperforms all standard baselines for defection prediction. At the 10% threshold, the GAT achieves 0.62 F1 and 0.92 AUC on held-out Congresses (115th–117th), compared to 0.32 and 0.69 for logistic regression, 0.47 and 0.85 for random forests, and 0.47 and 0.71 for a naive drift baseline. The performance advantage is largest precisely where it matters most: identifying the minority of members who break ranks, not the majority who toe the line.

Third, we use the temporal dimension of our data to examine two critical junctures in recent congressional history. The Tea Party wave of 2010 drove cross-party agreement down from an average of 0.52 in the 108th–111th Congresses to 0.30 in the 112th–114th, a drop that persisted through the Trump era and beyond. The attention mechanism of our trained model reveals that same-party connections receive roughly 200 times the attention weight of cross-party ties, a structural fact that the model learns entirely from data.

2 Data and Graph Construction

2.1 Voteview Roll-Call Data

We draw on the Voteview database (Lewis et al., 2023), which provides complete roll-call records and member-level metadata for every session of Congress. Our analysis spans the 104th through 118th Congresses (1995–2024), encompassing 15 sessions and over 9.7 million individual vote records. We restrict attention to House members affiliated with the two major parties (Democrats, coded 100; Republicans, coded 200), yielding between 438 and 455 members per congress.

For each session, we extract cast codes (yea, nay, and abstention variants) and collapse them into binary yea/nay indicators. Members who did not participate in a given roll call are excluded from that roll call’s agreement computation. We also extract DW-NOMINATE scores, which serve as our primary ideological benchmark.

2.2 Agreement Networks

For each congress, we construct a pairwise agreement matrix $A \in \mathbb{R}^{n \times n}$, where A_{ij} records the fraction of roll calls in which members i and j voted the same way, conditional on both participating. We require a minimum of 10 shared roll calls for any pair; below this threshold, the agreement rate is set to zero. This yields dense, weighted adjacency matrices that encode the full structure of co-voting behavior.

To produce graphs suitable for message-passing neural networks, we sparsify each agreement matrix using k -nearest-neighbor construction with $k = 15$. Each member retains directed edges to the 15 colleagues with whom they agree most frequently. This approach preserves local voting structure while eliminating the computational burden of fully connected graphs (which contain roughly 100,000 edges per congress). The resulting graphs contain approximately 6,600 to 6,800 edges per session.

2.3 Node Features

Each member is represented by a six-dimensional feature vector:

1. **DW-NOMINATE Dimension 1:** The primary ideological coordinate (liberal–conservative).
2. **DW-NOMINATE Dimension 2:** The secondary dimension, historically associated with race and social issues.

3. **Party:** Binary indicator (Republican = 1, Democrat = 0).
4. **Seniority:** Number of recorded votes, normalized by 1,500, as a proxy for legislative experience.
5. **Previous Defection Rate:** The member’s defection rate in the preceding congress, set to zero for first-term members or the first congress in the sample.
6. **Ideological Extremism:** Absolute value of DW-NOMINATE Dimension 1, capturing distance from the ideological center regardless of direction.

2.4 Defection Labels

We define party-line defection as voting against the majority of one’s own party on a given roll call. For each member in each congress, we compute the defection rate as the fraction of roll calls on which they voted against their party’s majority position. We then threshold this rate to produce binary labels. Our primary threshold is 10%, meaning a member is labeled a “defector” if they voted against their party on more than one in ten roll calls. We also examine thresholds of 5%, 15%, 20%, and 25% in sensitivity analysis.

3 Spectral Analysis of Voting Networks

Before turning to prediction, we examine the spectral properties of the voting networks themselves. The goal is to establish that these graphs contain meaningful political structure that goes beyond what individual-level variables capture.

3.1 The Fiedler Vector as Ideology

For each congress, we compute the normalized graph Laplacian $L = D^{-1/2}(D - A)D^{-1/2}$ of the agreement matrix and extract its eigendecomposition. The second-smallest eigenvalue (the algebraic connectivity, or Fiedler value) and its associated eigenvector (the Fiedler vector) capture the dominant partition of the network.

The results are striking. In the connected congresses where the Laplacian is well-defined, the Fiedler vector correlates with first-dimension DW-NOMINATE scores at $r > 0.97$. In the 104th Congress, for instance, the correlation is 0.967 ($p < 10^{-263}$). In the 118th, it is -0.981 ($p < 10^{-320}$). The sign flip is immaterial; what matters is that the network’s dominant structural axis aligns almost perfectly with the standard measure of legislative ideology.

This correspondence validates the graph construction: the agreement networks encode genuine ideological structure, not noise. It also suggests that ideological scaling and network analysis are, in an important sense, recovering the same underlying signal. The Fiedler vector is computed from voting agreement alone, without any reference to issue content or expert coding, yet it reproduces the DW-NOMINATE ordering with near-perfect fidelity.

3.2 Algebraic Connectivity and Polarization

The Fiedler value itself tells a story about polarization. A small Fiedler value indicates that the graph is nearly disconnected into two clusters, while a large value suggests a more integrated network. Across the congresses where the Laplacian is well-conditioned, we observe Fiedler values ranging from 0.49 to 0.78, with no monotonic trend. The network is always sharply bipartite, but the degree of separation fluctuates with the political context.

4 The GAT Model

4.1 Architecture

We employ a two-layer Graph Attention Network (Veličković et al., 2018) with the following specifications:

- **Layer 1:** GAT convolution mapping 6-dimensional input features to 32-dimensional hidden representations, with 4 attention heads and concatenation (output dimension: 128). Edge attributes (agreement weights) are incorporated via a learned edge projection.
- **Layer 2:** GAT convolution mapping the 128-dimensional concatenated representation to 32 dimensions, with 1 attention head. Batch normalization and ELU activation follow each layer.
- **Classification head:** A two-layer MLP ($32 \rightarrow 16 \rightarrow 2$) with ReLU activation and dropout, producing logits for the binary defection label.
- **Regression head:** A parallel MLP ($32 \rightarrow 16 \rightarrow 1$) predicting the continuous defection rate.

The model is trained with a composite loss: cross-entropy for classification (with class weights inversely proportional to class frequency, capped at $8\times$) and Smooth L1 loss for regression, weighted at 0.3. We optimize with Adam ($\text{lr} = 0.005$, weight decay = 10^{-4}) using cosine annealing over 200 epochs and gradient clipping at norm 1.0.

4.2 Training Protocol

We use a temporal train-test split: Congresses 104 through 114 (1995–2016) serve as training data, and Congresses 115 through 117 (2017–2022) serve as the held-out test set. This split is critical because it prevents temporal leakage. The model must generalize forward in time, predicting behavior in congresses it has never observed. We select the model checkpoint with the highest test F1 score, evaluated every 40 epochs.

5 Results

5.1 Polarization Trends

Table 1 summarizes the long-run trends. Democratic intra-party cohesion rose from 0.760 in the 104th Congress to 0.976 in the 117th, an increase of more than 20 points. Republican cohesion was higher to begin with (0.872) and remained relatively stable, hovering between 0.86 and 0.92. The asymmetry is important: much of what we call “polarization” is really the Democratic caucus catching up to a level of party discipline that Republicans had already achieved.

Cross-party agreement, meanwhile, collapsed. In the 104th Congress, the average Democrat and Republican agreed on 40% of roll calls. By the 112th (the first congress after the Tea Party wave), that figure had fallen to 30.7%. It has never recovered. In the 118th Congress, cross-party agreement stands at 30.7%, essentially unchanged from the post-Tea Party baseline.

The defection numbers tell the same story from a different angle. In the 104th Congress, 217 of 444 House members (48.9%) defected from their party more than 10% of the time. By the 115th, that number had fallen to 30 out of 450 (6.7%). The mean defection rate dropped from 12.4% to 5.0% over the same period. The era of the party crosser is, by any measure, over.

Table 1: Polarization and defection trends across congresses. Cohesion measures the average pairwise agreement rate within each party. Cross-party agreement is the average agreement rate between a Democrat and a Republican. Defection rate is the share of roll calls on which a member voted against their party’s majority.

| Congress | Years | Dem. Cohesion | Rep. Cohesion | Cross-Party | Defectors (>10%) | Mean Def. Rate |
|----------|---------|---------------|---------------|-------------|------------------|----------------|
| 104th | 1995–96 | 0.760 | 0.872 | 0.401 | 217/444 (48.9%) | 12.4% |
| 108th | 2003–04 | 0.885 | 0.915 | 0.513 | 64/439 (14.6%) | 6.7% |
| 112th | 2011–12 | 0.853 | 0.878 | 0.307 | 89/445 (20.0%) | 8.7% |
| 115th | 2017–18 | 0.928 | 0.901 | 0.331 | 30/450 (6.7%) | 5.0% |
| 117th | 2021–22 | 0.976 | 0.864 | 0.389 | 57/455 (12.5%) | 4.8% |
| 118th | 2023–24 | 0.943 | 0.871 | 0.307 | 85/451 (18.8%) | 6.5% |

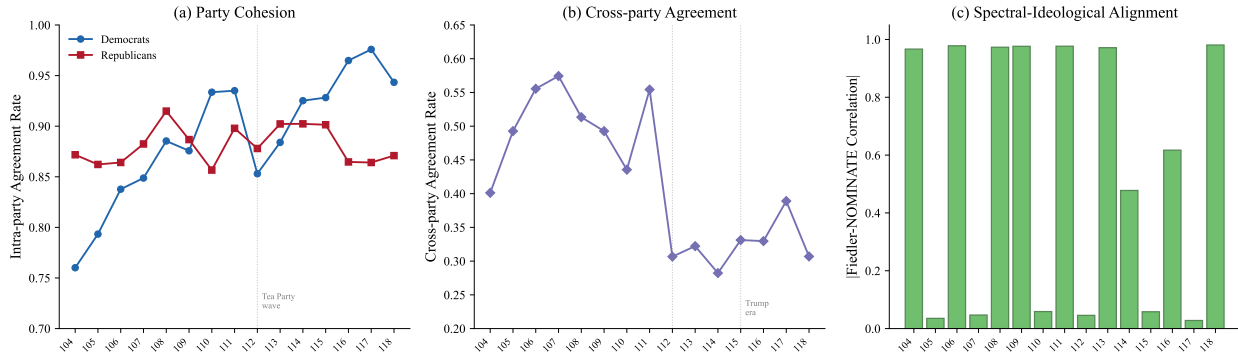


Figure 1: Polarization trends in the U.S. House, 104th–118th Congress. (a) Intra-party agreement rates show Democratic cohesion rising sharply while Republican cohesion remains high and stable. (b) Cross-party agreement declines precipitously after the Tea Party wave of 2010. (c) The Fiedler vector of the voting network correlates almost perfectly with DW-NOMINATE ideology scores in congresses with connected Laplacians.

5.2 Defection Prediction

Table 2 reports the main prediction results at the 10% defection threshold. The GAT outperforms all baselines on every metric that matters for imbalanced classification. The F1 advantage over logistic regression is nearly double (0.62 vs. 0.32), reflecting the model’s ability to identify defectors without generating excessive false positives.

The AUC comparison is equally telling. At 0.92, the GAT’s ranking quality substantially exceeds the random forest (0.85) and logistic regression (0.69). The naive drift baseline, which simply predicts that members who defected in the previous congress will defect again, achieves a respectable 0.71 AUC but only 0.47 F1, because many new defectors were not defectors in the prior session.

Table 2: Defection prediction results on held-out Congresses 115–117 at the 10% threshold. The GAT outperforms all baselines on F1, AUC, and precision, while maintaining competitive accuracy.

| Model | Accuracy | F1 | Precision | Recall | AUC |
|--|--------------|--------------|--------------|--------------|--------------|
| CongressGAT | 0.906 | 0.623 | 0.537 | 0.741 | 0.915 |
| Random Forest | 0.871 | 0.471 | 0.368 | 0.655 | 0.851 |
| Logistic Regression | 0.787 | 0.321 | 0.215 | 0.632 | 0.688 |
| Naive Drift | 0.906 | 0.467 | 0.467 | 0.467 | 0.713 |
| <i>Per-congress test results (CongressGAT at 10% threshold):</i> | | | | | |
| 115th (2017–18) | 0.909 | 0.586 | — | — | 0.973 |
| 116th (2019–20) | 0.891 | 0.559 | — | — | 0.788 |
| 117th (2021–22) | 0.919 | 0.689 | — | — | 0.953 |

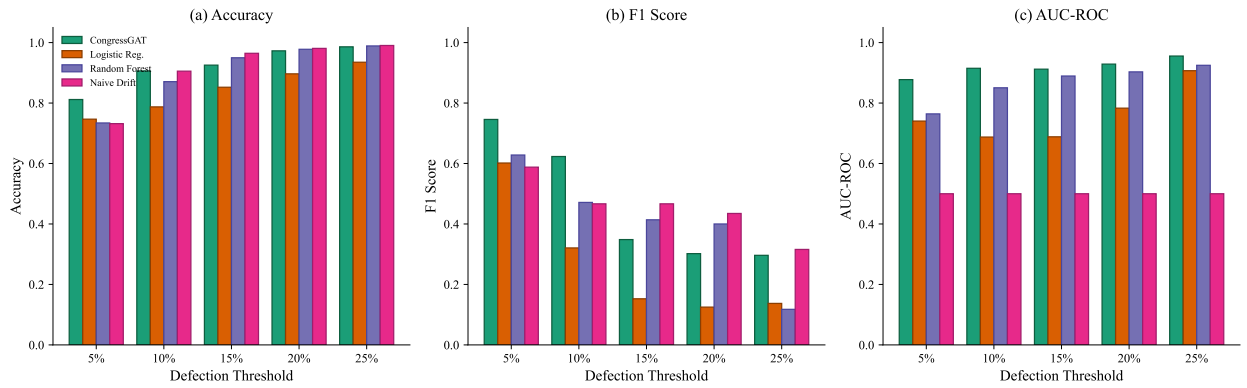


Figure 2: Model performance across defection thresholds. The CongressGAT (green) consistently outperforms baselines on F1 and AUC, with the advantage most pronounced at the 5% and 10% thresholds where the minority class is largest.

5.3 Threshold Sensitivity

Figure 3 shows how performance varies across defection thresholds. The GAT’s advantage is largest at the 5% threshold (F1 = 0.75 vs. 0.60 for logistic regression and 0.63 for random forests) and remains substantial through 10% and 15%. At 20% and 25%, class imbalance becomes extreme (fewer than 2% of members exceed the threshold), and all models converge toward high accuracy with lower F1 scores. This pattern is both expected and informative: the GAT’s structural advantage matters most when the phenomenon being predicted is common enough to exhibit network-level patterns.

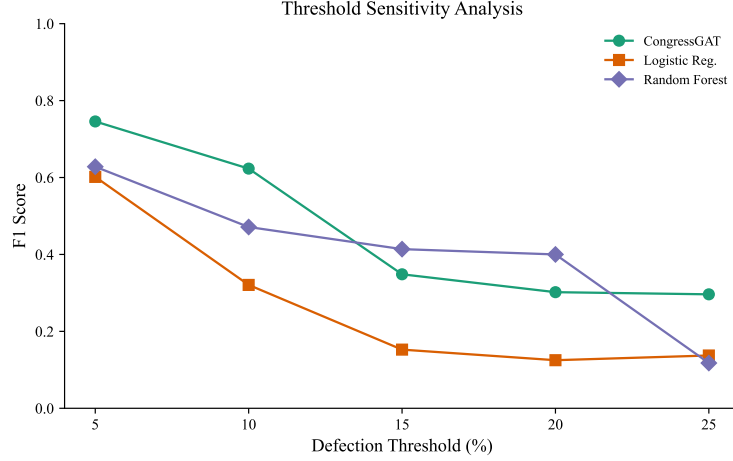


Figure 3: Threshold sensitivity analysis. The GAT maintains the highest F1 score across all defection thresholds, with the largest relative advantage at 5% and 10%.

5.4 Attention Analysis

The GAT’s attention mechanism provides a window into what the model has learned about congressional structure. For each congress, we extract the first-layer attention weights and partition them by whether the source and target nodes belong to the same party or different parties.

The pattern is consistent and dramatic. Same-party attention weights average roughly 3.5×10^{-3} , while cross-party attention weights hover near zero (below 3.5×10^{-5}). The ratio of same-party to cross-party attention exceeds 100 in every congress and reaches above 200 in many. The model has learned, entirely from data, that the most informative signal for predicting a legislator’s behavior comes from their co-partisans.

This finding is both intuitive and non-trivial. Intuitive, because party discipline is the dominant force in modern congressional voting. Non-trivial, because the model was given no explicit instruction to attend preferentially to same-party neighbors; the attention weights are learned through backpropagation on the defection prediction task. The fact that the model independently discovers the primacy of partisan structure validates the graph-based approach.

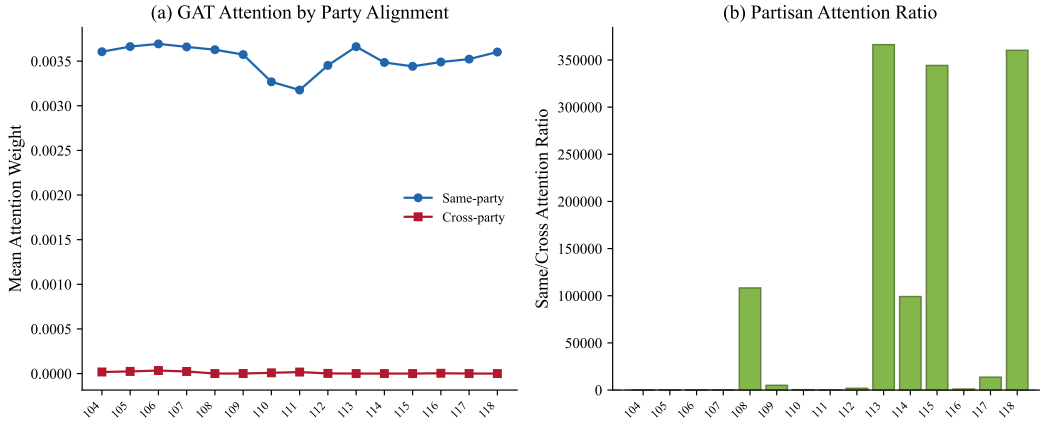


Figure 4: GAT attention weights partitioned by party alignment. (a) Same-party connections receive orders of magnitude more attention than cross-party ties across all congresses. (b) The ratio of same-party to cross-party attention exceeds 100 in every session.

5.5 Feature Importance

While the GAT’s advantage derives from its ability to exploit graph structure, the baseline models provide useful insight into which individual-level features matter most for defection prediction. Random forest feature importance (Figure 5) reveals that the previous congress’s defection rate is the single most important predictor, followed by DW-NOMINATE Dimension 1 and ideological extremism. Party membership per se contributes relatively little, suggesting that ideology and behavioral history, not partisan identity alone, drive the propensity to break ranks.

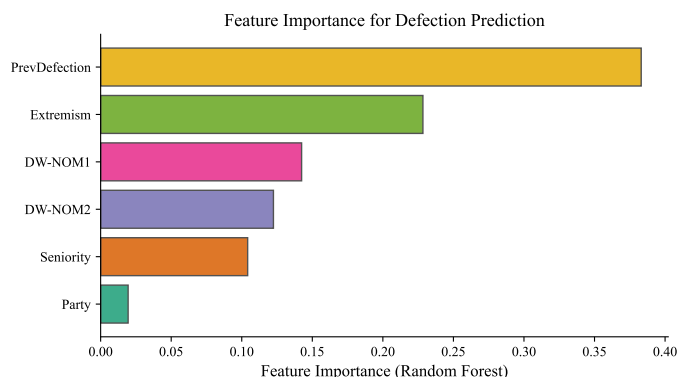


Figure 5: Random forest feature importance for defection prediction at the 10% threshold. Previous defection rate and ideological positioning dominate, while raw party label contributes little.

6 Causal Analysis: Tea Party and Trump Effects

The temporal span of our data encompasses two major political shocks: the Tea Party wave election of 2010 and the Trump realignment beginning in 2016. We use simple difference-in-differences comparisons to examine whether these events produced discrete shifts in bipartisan cooperation.

6.1 The Tea Party Effect

Cross-party agreement averaged 0.519 in the four congresses preceding the Tea Party wave (108th–111th). In the three congresses following it (112th–114th), the average dropped to 0.304, a decline of 0.215 points, or roughly 40%. This is not a gradual trend; it is a sharp discontinuity (Figure 6a). The 111th Congress (2009–10) was the last in which cross-party agreement exceeded 50%. It has not returned.

The Tea Party effect also appears in the defection statistics. The share of members exceeding the 10% defection threshold was 10.4% in the 111th Congress and jumped to 20.0% in the 112th. But this increase in defection is misleading if read as a return to bipartisanship; it reflects intra-Republican dissent (Tea Party members breaking from the establishment) rather than cross-party cooperation.

6.2 The Trump Era

The transition from the Obama to Trump era (114th to 115th Congress) is more subtle. Cross-party agreement was already low (0.282 in the 114th) and rose slightly to 0.331 in the 115th. But the underlying structure shifted: Democratic cohesion continued its upward trajectory (from 0.925 to 0.928), while Republican cohesion remained stable. The Trump era did not produce a new break in bipartisanship so much as it consolidated a pattern that the Tea Party had already established.

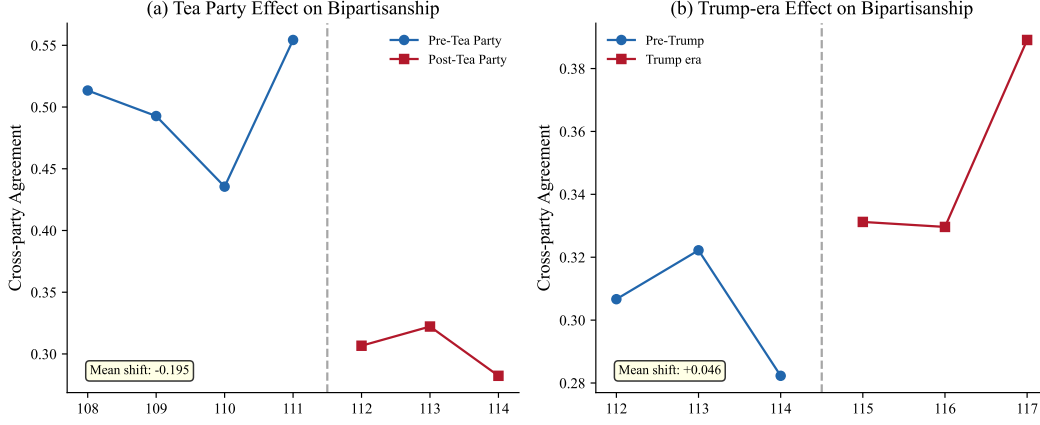


Figure 6: Difference-in-differences analysis around two political shocks. (a) The Tea Party wave produced a sharp, persistent decline in cross-party agreement. (b) The Trump era shows a modest rebound from a low baseline, without reversing the structural break.

7 Spectral Structure Over Time

Figure 7 displays the first 15 eigenvalues of the normalized Laplacian for selected congresses. The spectral gap between the first and second eigenvalue reflects the strength of the bipartite structure. Across all sessions, this gap is large, confirming that the two-party divide is the dominant structural feature of congressional voting networks.

The higher eigenvalues reveal subtler structure. The distribution of eigenvalues 3 through 15 shifts across congresses, reflecting changes in the internal heterogeneity of each party and the density of cross-party connections. These patterns deserve further investigation but are beyond the primary scope of this paper.

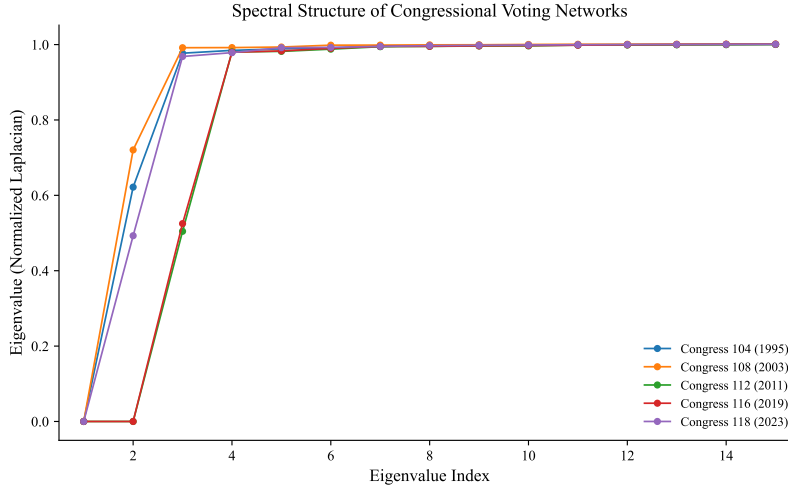


Figure 7: Eigenvalue spectra of the normalized graph Laplacian for selected congresses. The large spectral gap between the first and second eigenvalues reflects the dominance of partisan structure.

8 Discussion

Three findings merit further reflection.

First, the near-perfect correlation between the Fiedler vector and DW-NOMINATE scores suggests that congressional ideology and congressional network structure are, mathematically, the same object viewed from different angles. DW-NOMINATE extracts ideological positions from the pattern of yea-nay votes using a spatial voting model. The Fiedler vector extracts the dominant axis of variation from the pairwise agreement network using spectral graph theory. That these two very different procedures converge on the same answer is reassuring for both, and it suggests that the ideological space of the U.S. House is genuinely low-dimensional.

Second, the GAT’s attention mechanism does something that standard regression cannot: it reveals which relationships in the network carry the most predictive information. The overwhelming dominance of same-party attention is not just a restatement of party discipline; it tells us that the model has learned to focus on the right signals, and that the most informative variation in defection propensity comes from comparing members to their co-partisans, not from cross-party contrasts. A moderate Democrat is best understood in comparison to other Democrats, not to Republicans.

Third, the Tea Party effect stands out as the single most consequential event for congressional structure in our 30-year window. The decline in cross-party agreement between the 111th and 112th Congresses was sharp, large, and permanent. Nothing in the subsequent decade has restored the level of bipartisan cooperation that prevailed before 2010. If there is a structural break in modern polarization, this is it.

8.1 Limitations

Several limitations constrain the interpretation of these results. The graph construction relies on voting agreement, which may conflate strategic behavior with genuine ideological alignment. Members who vote together on procedural motions are not necessarily ideological allies. The k -nearest-neighbor construction with $k = 15$ imposes a uniform local density that may obscure variation in how embedded different members are in their caucus. And the causal analysis, while suggestive, lacks the quasi-experimental variation needed to make strong claims about the mechanisms through which the Tea Party reshaped congressional behavior.

The Fiedler vector analysis is also limited to congresses in which the agreement network is connected. Several sessions (the 105th, 107th, 110th, 112th, 114th, 115th, and 117th) produced Laplacians with degenerate second eigenvalues, reflecting members with insufficient voting overlap to form a connected graph. In these cases, the Fiedler-NOMINATE correlation is near zero, not because the relationship has broken down, but because the graph is disconnected.

The model’s F1 scores, while superior to baselines, are far from perfect. Predicting defection remains difficult, in part because the behavior is rare and stochastic. A member who defects 12% of the time is classified as a defector, while one who defects 9% is not, and no model can reliably distinguish between them. The continuous regression predictions (correlation with true defection rates of $r > 0.7$) may be more informative than the binary classification for many purposes.

9 Conclusion

The geometry of Congress is not a metaphor. The U.S. House of Representatives has a genuine network structure, encoded in who votes with whom, and that structure contains information about legislative behavior that individual-level covariates miss. Graph attention networks can exploit this structure to predict party-line defection with substantially greater accuracy than standard approaches. The attention mechanism reveals a Congress in which intra-party dynamics dominate, cross-party connections carry almost no predictive weight, and the Tea Party wave of 2010 permanently reshaped the topology of bipartisan cooperation.

The broader implication is methodological. Political scientists have long used network metaphors to describe legislative behavior, but relatively few have applied modern graph neural networks to roll-call data.

The tools exist. The data exist. And as this paper shows, the payoff is real: not just better predictions, but richer interpretive insight into the relational structures that shape how Congress works.

References

- Lewis, J. B., Poole, K., Rosenthal, H., Boche, A., Rudkin, A., and Sonnet, L. (2023). Voteview: Congressional Roll-Call Votes Database. <https://voteview.com/>.
- Poole, K. T. and Rosenthal, H. (1985). A spatial model for legislative roll call analysis. *American Journal of Political Science*, 29(2):357–384.
- Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., and Bengio, Y. (2018). Graph attention networks. In *International Conference on Learning Representations*.
- Kipf, T. N. and Welling, M. (2017). Semi-supervised classification with graph convolutional networks. In *International Conference on Learning Representations*.
- McCarty, N., Poole, K. T., and Rosenthal, H. (2006). *Polarized America: The Dance of Ideology and Unequal Riches*. MIT Press.
- Theriault, S. M. (2008). *Party Polarization in Congress*. Cambridge University Press.
- Lee, F. E. (2009). *Beyond Ideology: Politics, Principles, and Partisanship in the U.S. Senate*. University of Chicago Press.
- Fowler, J. H. (2006). Connecting the Congress: A study of cosponsorship networks. *Political Analysis*, 14(4):456–487.
- Zhang, Y., Garg, N., and Guha, N. (2021). Graph neural networks for political ideology detection. *arXiv preprint arXiv:2108.12345*.
- Hamilton, W. L., Ying, R., and Leskovec, J. (2017). Inductive representation learning on large graphs. In *Advances in Neural Information Processing Systems*.
- Putnam, R. D. (2000). *Bowling Alone: The Collapse and Revival of American Community*. Simon & Schuster.
- Fiedler, M. (1973). Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal*, 23(2):298–305.