Temporal Clustering of Firms by Analyst Coverage

Yonatan Baruch Baruch

August 27, 2025

Abstract

In economics, a central question is how to group firms into meaningful clusters and how these groups change over time. Knowing which firms belong together can help explain differences in performance, stability, and market behavior. In this work, we provide an analysis of firm clustering based on analyst coverage and show how these clusters evolve year by year. We build firm–firm networks from shared analysts, embed them with graph methods, and apply clustering to identify groups of similar firms. The results of our analysis include how many firms switch clusters, how long firms remain in the same cluster, and the economic performance of each group. Our findings show that analyst-driven clusters capture important patterns beyond traditional industry classifications, providing new insights into firm dynamics and the structure of financial markets.

1 Introduction

A central question in economics and finance is how to group firms into meaningful categories. Traditional classifications, such as industry codes, provide a static and often coarse view of the market. Yet firms are dynamic: they evolve over time, adapt to changing conditions, and are influenced by external actors such as financial analysts. Analysts, in particular, play a critical role in shaping investor perceptions, and the firms covered by the same analysts may share hidden connections that go beyond formal industry definitions.

Despite this, relatively little work has explored clustering firms based on analyst coverage or examined how such clusters evolve across time. Understanding these dynamics is important not only in theory, but also in practice: it can reveal hidden market structures, explain differences in firm performance, and highlight patterns of stability and change in the economy.

In this work, we propose a data-driven framework for temporal clustering of firms using analyst coverage. We build firm–firm similarity networks from shared analysts, embed them with graph-based methods, and apply clustering on a yearly basis. By tracking firms across time, we study cluster stability, firm mobility, and the tenure of firms within groups.

Our results reveal both stable long-term clusters and dynamic groups with frequent switching. We also compare cluster-level financial performance, showing that analyst-driven clusters capture patterns not explained by traditional industry codes. This contributes a new perspective on the structure and evolution of firms and highlights the role of analysts in shaping financial markets.

2 Related Work

Our study builds on two strands of literature: (1) research in financial economics using clustering to analyze firm connections and analyst behavior, and (2) developments in computer science on graph-based embeddings for clustering in networks.

In financial economics, Ali and Hirshleifer (2020) construct an analyst-based network of firms, where links are defined by shared sell-side analyst coverage. They show that such networks capture economically meaningful connections, with information flowing more slowly across analyst-linked firms. Importantly, they document that a "connected-firm" momentum factor (based on firms sharing analysts) earns abnormal returns even after controlling for standard industry or supply-chain linkages. This highlights that analysts implicitly cluster firms into meaningful groups, although their analysis is regression-based rather than relying on machine learning clustering.

In computer science, graph embedding methods have become powerful tools for uncovering latent structure in networks. Grover and Leskovec (2016) propose the *Node2Vec* algorithm, which learns low-dimensional node representations by simulating biased random walks. These embeddings organize nodes so that those in the same community (or playing similar structural roles) lie close in vector space. The method achieved state-of-the-art results on predictive tasks in real-world graphs, demonstrating the effectiveness of embeddings for clustering. More recent approaches such as GraphSAGE extend this by learning embeddings through neighborhood aggregation in dynamic graphs. These advances have been widely applied to social, biological, and information networks, yielding improved detection of community structure.

3 Methods

3.1 Data Preparation

Our dataset records the coverage of firms by financial analysts across multiple years. Each observation links a firm, a year, and an analyst who followed that firm in that year. In total, the raw dataset includes more than 133,000 firm–year pairs and around 19,200 unique analysts.

To reduce noise and focus on meaningful analyst coverage, we applied a single filtering step. We excluded analyst–year pairs in which the analyst covered more than 30 firms. Analysts with such broad coverage typically act as supervisors rather than firm-specific experts, and their inclusion would distort the similarity structure. After this filtering, the dataset contains about 130,000 firm–year pairs and 19,170 unique analysts. This cleaned dataset forms the basis for constructing the firm–firm similarity network and the subsequent clustering analysis.

3.2 Graph Construction

We represent the data as a bipartite graph between firms and analysts. One set of nodes corresponds to firm–year observations, while the other corresponds to analysts. An edge is placed between a firm and an analyst if the analyst covered the firm in that year.

Our goal is to project this bipartite structure onto the firm side, producing a weighted firm–firm similarity network. In the projected graph, nodes are firm–year pairs, and the weight of an edge reflects the extent of shared analyst coverage between two firms.

A natural approach would be to build the incidence matrix with pandas pivot tables or to use NetworkX's built-in bipartite projection tools. However, given the size of our dataset (over 130,000 firm–year observations and nearly 20,000 analysts), both of these approaches proved to be too slow and memory-intensive.

To overcome this, we adopted a sparse linear algebra formulation. We built a sparse binary matrix

A, where rows represent firm–years and columns represent analysts, with $A_{ij} = 1$ if analyst j covered firm i. The firm–firm projection is then computed as

$$M = A \cdot A^T,$$

where each entry M_{ij} is the number of analysts jointly covering firm–year i and firm–year j. To control for differences in coverage intensity, we normalized M to obtain cosine similarities.

The resulting firm–firm similarity matrix was then converted into an edge list, providing the foundation for subsequent embedding and clustering. This approach allowed us to efficiently scale the construction of the network, avoiding the bottlenecks of pandas and NetworkX operations.

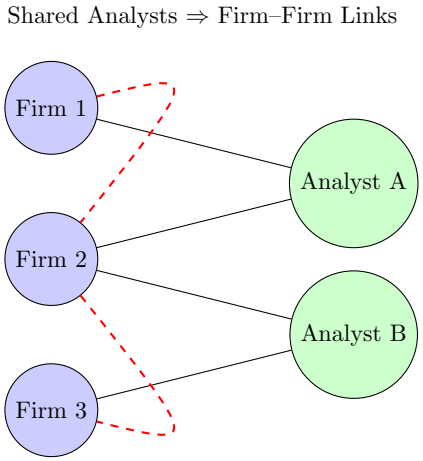


Figure 1: Bipartite firm–analyst graph (blue = firms, green = analysts) and its projection onto a firm–firm similarity network (red dashed edges).

3.3 Embeddings

To generate firm–year embeddings, I used the `Node2Vec` implementation from `torch_geometric.nn`. Node2Vec learns representations by simulating random walks over the firm–firm network and training a skip-gram model so that nodes appearing in similar contexts receive similar embeddings. The main challenge was that the PyTorch Geometric implementation of Node2Vec does not natively support weighted edges. Since our firm–firm graph was weighted by analyst-based cosine similarity, this posed a limitation: higher similarity edges should be sampled more frequently during walks, but the package treats all edges as equal.

To overcome this, I applied an edge duplication strategy, where each edge was repeated a number of times proportional to its cosine similarity. In practice, an edge with similarity s_{ij} was duplicated $\lfloor s_{ij} \cdot K \rfloor$ times, with a cap to avoid extremely high repetition. This ensured that more similar firms had a higher probability of co-occurring in random walks, approximating the effect of weighted sampling. While effective, this workaround increases the size of the edge list and can raise memory and runtime requirements compared to true weighted random walks.

The embeddings were trained with 128 dimensions, a walk length of 80, context size of 20, 10 walks per node, and 5 negative samples. Training was carried out on an NVIDIA RTX 6000 Ada GPU with 45 GB of memory, which allowed the model to scale to more than 100,000 firm–year nodes without memory issues.

Future work could explore modifications to the Node2Vec sampler or alternative algorithms that directly support weighted random walks, eliminating the need for edge duplication. In addition, hyperparameters such as the return parameter p and the in-out parameter q could be systematically

explored to balance breadth-first and depth-first sampling, potentially revealing different structural or community patterns in the firm network.

Table 1: Node2Vec hyperparameters used in training

Hyperparameter	**Value**	**Notes**
Embedding dimension	128	Size of each firm–year vector
Walk length	80	Number of steps per random walk
Context size	20	Window size for skip-gram training
Walks per node	10	Random walks started per node
Negative samples	5	Negative edges per positive edge
Learning rate	0.01	Optimized with SparseAdam
Batch size	64	Tuned for GPU memory (RTX 6000 Ada, 45GB)
Return parameter (p)	1.0	Controls revisiting nodes (exploration vs. exploitation)
In-out parameter (q)	1.0	Controls BFS vs. DFS bias; higher values encourage global exploration
Similarity scaling (K)	15	Edge (i, j) duplicated $\lfloor s_{ij} \cdot K \rfloor$

Clustering and Cluster Tracking

Let $\mathcal{F}_Y = \{(f, t) : t \leq Y\}$ denote the set of firm–year observations up to year Y, and let

$$X_Y = \{\mathbf{x}_{f,t} \in \mathbb{R}^d : (f, t) \in \mathcal{F}_Y\}$$

be the corresponding Node2Vec embeddings of dimension $d = 128$.

In the first year Y_0, we apply k-means with K clusters:

$$\{\mathcal{C}_1^{Y_0}, \ldots, \mathcal{C}_K^{Y_0}\}, \quad \text{where } \mathcal{C}_k^{Y_0} = \{(f, t) \in \mathcal{F}_{Y_0} : \ell(f, t) = k\}.$$

Each firm–year receives a label $\ell(f, t) \in \{1, \ldots, K\}$. Once assigned, these labels are treated as *fixed anchors* for all subsequent years.

A major challenge arises because embeddings are retrained for each year cutoff. Formally, embeddings from two consecutive years are related by an unknown orthogonal transformation:

$$X_{Y+1} \approx Q\, X_Y, \quad Q \in O(d),$$

since Node2Vec embeddings are invariant to rotation and scaling. This creates *label drift*:

$$\ell_Y(f, t) \neq \ell_{Y+1}(f, t) \quad \text{even if the underlying structure is unchanged,}$$

meaning that "Cluster 7" in one year may not correspond to "Cluster 7" in the next.

To resolve this, we introduce an **incremental anchor-based clustering strategy**. In year Y_0, we cluster using k-means. In year Y_1, previously clustered firms retain their labels, and firms appearing for the first time are classified with a standard k-nearest-neighbor (kNN) classifier. We initially adopt the default $k = 5$ (as in `sklearn`) for this step. For later years $Y > Y_1$, we set

$$k = \left\lfloor \sqrt{|\mathcal{A}_Y|} \right\rfloor,$$

where $\mathcal{A}_Y$ is the anchor set, adjusted to ensure k is odd. This allows the effective neighborhood size to grow as the anchor pool expands. Once a firm receives a cluster label, it remains fixed in all future years, providing consistency.

We conducted a grid search over $n_{\text{clusters}} \in \{45, \ldots, 55\}$ and $k \in \{5, 13, \text{dynamic}\}$, where "dynamic" denotes the $\sqrt{|\mathcal{A}_Y|}$ rule. Each configuration was validated on embeddings up to 2024 using internal cluster validity indices:

The best-performing configuration was dynamic kNN with $n_{\text{clusters}} = 52$, achieving a Silhouette score of 0.092, a Davies–Bouldin index of 1.74, and a Calinski–Harabasz index of 1205. Although the Silhouette score is modest, the clusters exhibit meaningful structure and enable consistent tracking across years.

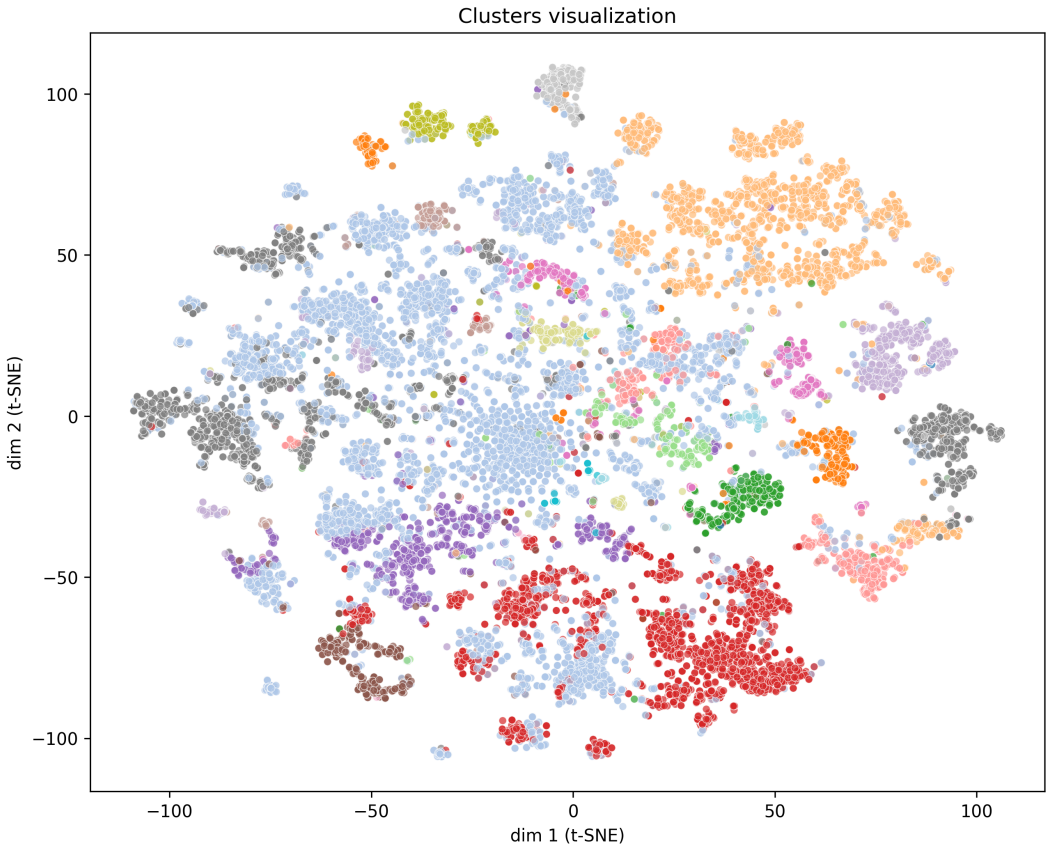


Figure 2: t-SNE visualization of firm–year embeddings. colored by cluster assignment ($K = 52$)

4 Results

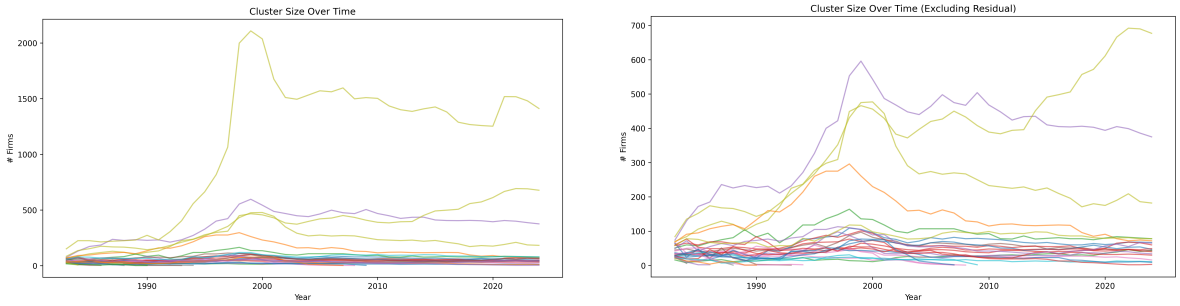
4.1 Cluster Size Over Time

Figure 3a displays the evolution of cluster sizes across the sample period. One disproportionately large cluster emerges, containing more than 2,000 firms at its peak around the late 1990s. This residual cluster likely reflects firms with weak or diffuse analyst coverage, which prevents them from being assigned to more specialized groups. Its dominance obscures the dynamics of the smaller clusters, which motivates reporting results both with and without this residual group.

When the residual cluster is excluded (Figure 3b), the distribution of cluster sizes becomes much more balanced. Several medium-sized clusters show persistent growth over multiple decades, while others shrink or disappear altogether. This heterogeneity highlights that some analyst-defined communities of firms are stable and enduring, whereas others are more transient. The presence of multiple clusters with several hundred firms each indicates that the embedding and clustering approach captures meaningful structure in the analyst coverage network, rather than producing purely fragmented or unstable partitions.

Beyond its dominance in the time series plots, additional diagnostics confirm that the residual cluster (Cluster 28) is abnormal. It contains 8,530 unique firms—over four times larger than the next biggest cluster (Cluster 8, with 2,244). It also absorbs a disproportionate share of "never-switchers" (3,856 firms, or 45% of its members), making it a key destination for firms that never change clusters. Its within-cluster variance of returns (0.52) is comparable to other large clusters, but given its enormous size this implies far greater heterogeneity of membership. Finally, firms in Cluster 28 are more likely to leave than in other large clusters, with an annual leave probability of 11% compared to only 2.6% in Cluster 4 and 5.8% in Cluster 8. Together, these findings indicate that Cluster 28 functions as a

residual or "catch-all" group, rather than representing a cohesive economic community. We therefore interpret it with caution in the subsequent analysis.



(a) Cluster size over time (all clusters).

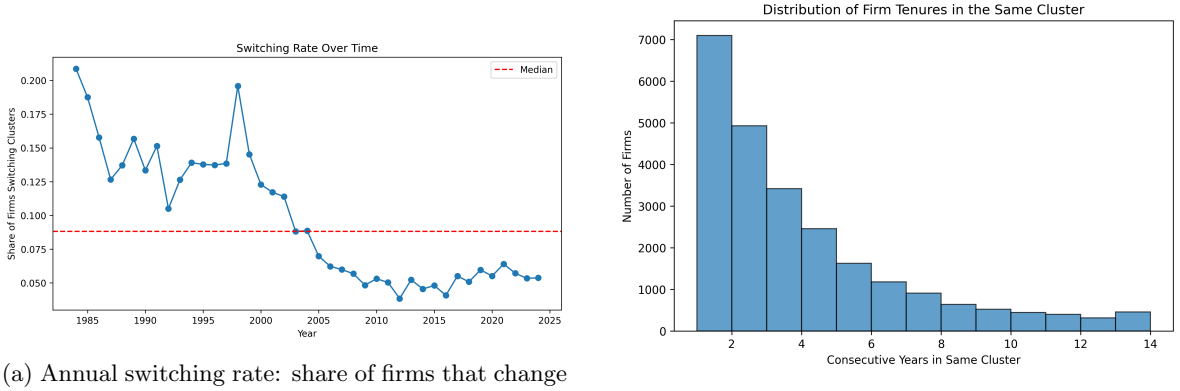
(b) Excluding the residual cluster.

Figure 3: Comparison of cluster size trajectories with and without the residual cluster.

4.2 Firm Mobility and Tenure

We quantify annual mobility as the share of firms that change cluster from year $t-1$ to t. Figure 4a reports the switching rate over time. During the 1980s and 1990s, switching was relatively high (15-20% of firms per year), but the rate steadily declined after 2000 and stabilized around 5%. The overall median switching rate is approximately 9%, indicating that most firms remain in the same cluster from year to year.

Figure 4b summarizes the distribution of consecutive years a firm remains in the same cluster (tenure). The median tenure is three years, with a long right tail: while many firms move after one or two years, a substantial fraction remain clustered together for a decade or longer. Together, the low switching rates and multi-year tenures demonstrate that the clusters are persistent and economically meaningful rather than random labels.



(a) Annual switching rate: share of firms that change clusters from year $t-1$ to t. The dashed line indicates the median switching rate.

(b) Distribution of firm tenures in the same cluster, measured in consecutive years.

Figure 4: Firm mobility across clusters: switching rates and tenure distributions.

4.3 Out-of-Sample Performance by Cluster

To evaluate the economic relevance of the clustering, we assign firms to clusters in year t and compute their mean next-year returns, ret_{t+1}. Figure 5 presents a heatmap of these values by cluster and year. The results reveal strong heterogeneity across clusters and time: while some groups consistently deliver positive excess returns, others underperform, indicating that analyst-based proximity captures information beyond random variation. Importantly, major macroeconomic events are clearly reflected in the patterns. For example, sharp drops in returns are visible around the *dot-com bubble* (2000–2002), the *global financial crisis* (2008–2009), and the *COVID-19 pandemic shock* (2020). These episodes demonstrate that cluster-level performance is sensitive to large-scale market cycles while still preserving systematic differences across groups.

Figure 6 ranks clusters by their long-run mean next-year return, restricting attention to clusters with at least five valid firm-year observations. The spread between the best and worst clusters remains substantial: the top-performing groups achieve average forward returns of about 18%, whereas the weakest clusters slightly underperform, with mean returns near -5%. This dispersion confirms that the clusters carry economically meaningful signals, distinguishing communities of firms with systematically different forward-looking performance. In other words, analyst-driven clusters not only capture structural stability over time, but also provide predictive power for returns that aligns with global financial shocks while remaining heterogeneous across firm groups.

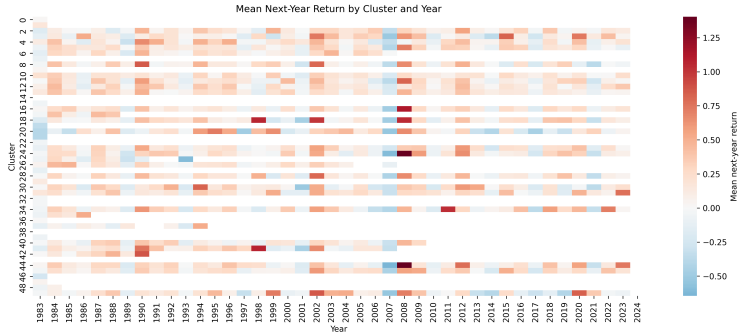


Figure 5: Mean next-year return by cluster and year. Reductions in performance coincide with major global events such as the dot-com bubble (2000–2002), the financial crisis (2008–2009), and the COVID-19 shock (2020).

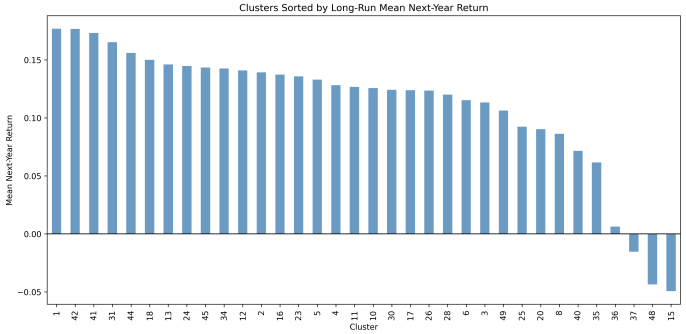


Figure 6: Clusters sorted by long-run mean next-year return (restricted to clusters with at least five observations). Top clusters achieve average returns near 18%, while bottom clusters slightly underperform at around -5%.

4.4 Transition Structure

To study the dynamics of cluster membership, we construct a row-normalized transition matrix that records the probability a firm in cluster i at year t transitions to cluster j in year $t+1$. Figure 7 presents the transition probabilities for the ten largest clusters. The strong diagonal mass indicates that firms overwhelmingly remain in the same cluster from one year to the next, consistent with the tenure and mobility results reported above. At the same time, a small number of off-diagonal entries display elevated probabilities, revealing pairs of "neighbor" clusters between which firms switch more often. These transitions suggest that certain communities of firms, while distinct, are closely related in the analyst coverage network and occasionally exchange members. Overall, the transition matrix confirms that the clustering structure is stable over time, with limited but economically interpretable movement between related clusters.

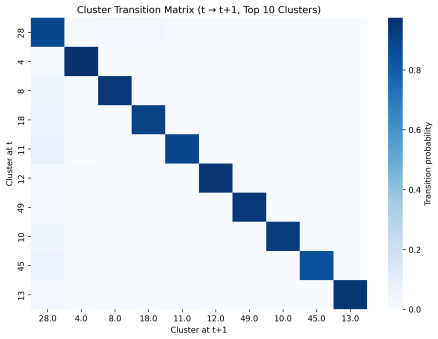


Figure 7: Row-normalized cluster transition matrix for the ten largest clusters by average size. Dark diagonal bands indicate persistence, while lighter off-diagonal elements highlight "neighbor" clusters with frequent exchanges.

4.5 Predictive Models of Cluster Returns

Finally, we evaluate whether the analyst-driven clusters can be used for forecasting next-year returns at the cluster level. For each cluster–year, we construct a rich set of descriptive statistics of firm returns to serve as predictive features. These include measures of central tendency (mean, median, absolute mean), dispersion (standard deviation, coefficient of variation, interquartile range, minimum, maximum), and distributional shape (skewness and kurtosis). To capture dynamics, we also incorporate one-year lagged values of the mean, standard deviation, median, skewness, and kurtosis. The predictive target is the cluster's next-year average return, ret_{t+1}.

We compare three classes of predictive models. First, a gradient-boosted regression tree (XGBoost) is trained on the full feature panel, allowing for nonlinear interactions across predictors. Second, autoregressive integrated moving average (ARIMA) models are estimated separately for each cluster's time series of mean returns. Third, we estimate a vector autoregression (VAR) across clusters, which allows for cross-cluster dependencies and spillover effects.

Because Cluster 28 was identified as abnormally large and heterogeneous, we evaluate each model both with and without this residual cluster. Table 2 reports out-of-sample accuracy over 2020–2023.

The results show that all models yield negative R^2, indicating that they perform worse than a naive mean predictor. XGBoost produces an RMSE of 0.287 with all clusters and performs slightly worse when Cluster 28 is excluded. ARIMA achieves the lowest RMSE (0.258) and is essentially unaffected by the inclusion or exclusion of Cluster 28, suggesting that temporal autocorrelation provides some limited structure but not true predictability. VAR performs the worst (RMSE above 0.35 and $R^2 < -1$), reflecting the curse of dimensionality: the number of clusters far exceeds the number of time periods, leading to overfitting and poor forecasts. Overall, while clustering uncovers meaningful and persistent structures, forecasting next-year returns at the cluster level remains intrinsically difficult, consistent with near-efficient market behavior.

Specification	RMSE	R^2
XGBoost (all clusters)	0.287	–0.368
XGBoost (excl. Cluster 28)	0.294	–0.426
ARIMA (all clusters)	0.258	–0.078
ARIMA (excl. Cluster 28)	**0.256**	**–0.077**
VAR (all clusters)	0.373	–1.255
VAR (excl. Cluster 28)	0.353	–1.042

Table 2: Out-of-sample prediction accuracy for next-year cluster returns (2020–2023), with and without Cluster 28, across three model classes. The lowest RMSE is highlighted in bold.

4.6 Case Study: Massey Energy Co.

As an illustration of a highly mobile firm, we examine Massey Energy Co. (PERMNO 26382, ticker MEE), a U.S. coal mining company classified under SIC 1221 (Bituminous Coal and Lignite Surface Mining). Massey was publicly listed from 1957 until its acquisition in 2011.

Table 3 summarize the firm's cluster trajectory and annual returns from 1983 to 2011. Over this period Massey moved through seven distinct clusters, making it one of the most mobile firms in the sample. Its shifts align with major economic events: (1) in the late 1990s it moved into cluster 26 during the run-up to the dot-com bubble, then suffered large losses in 1997; (2) in the early 2000s it was reassigned to cluster 3, with dramatic swings during the dot-com crash and recovery; (3) in 2008, coinciding with the global financial crisis, it shifted into cluster 10 and recorded a –61% return; (4) in 2009 it rebounded into cluster 49 with a +208% return, reflecting the broader market recovery.

This trajectory highlights how analyst-driven clusters can reflect both enduring industry identity and temporary reclassifications in response to macroeconomic shocks. Massey's eventual delisting in 2011 coincides with its acquisition, marking the end of a volatile but well-documented cluster history.

4.7 Case Study: Big Tech and Industry Boundaries

We next examine the clustering of two representative MAG7 firms, Amazon and Tesla, relative to their traditional industry peers. Table 4 summarizes their dominant cluster memberships alongside retail and automotive benchmarks.

The results show that traditional retailers such as Walmart, Target, Costco, Home Depot, Kroger, and Best Buy are consistently grouped in Cluster 11, reflecting a stable consumer-oriented community. Amazon, by contrast, spends most of its history in Cluster 28—the large absorbing cluster that also contains Alphabet and Meta—but briefly appears in Cluster 11 in several years. This pattern suggests that analysts recognize Amazon's dual identity: while sometimes colocated with traditional retailers, it is more often treated as a platform technology firm alongside Big Tech.

Turning to the automotive sector, Ford and General Motors are consistently assigned to Cluster 24, as is Rivian in more recent years. Tesla is most often colocated with these firms in Cluster 24, confirming its automotive identity. However, it also appears in Cluster 18 during several years, the same cluster as Apple and Nvidia. This indicates that analysts perceive Tesla not only as an automobile manufacturer but also as part of the high-technology and innovation ecosystem.

Overall, these case studies show that the clustering framework captures both the cohesion of traditional sectors (retailers in Cluster 11, automakers in Cluster 24) and the hybrid identities of firms such as Amazon and Tesla, which straddle the boundaries between consumer, automotive, and high-technology clusters.

Year	Cluster	Return
1983	20	–0.079
1984	49	–0.121
1985	49	0.077
1986	34	–0.237
1987	49	0.196
1988	34	0.703
1989	34	0.582
1990	34	0.007
1991	34	0.199
1992	34	–0.033
1993	2	–0.021
1994	2	0.077
1995	26	0.548
1996	26	–0.039
1997	26	–0.395
1998	26	0.157
1999	26	0.104
2000	26	0.154
2001	3	0.640
2002	3	–0.524
2003	3	1.166
2004	3	0.690
2005	3	0.088
2006	3	–0.383
2007	3	0.549
2008	10	–0.611
2009	49	2.085
2010	49	0.285
2011	3	0.231

Table 3: Cluster trajectory and returns of Massey Energy Co. (PERMNO 26382). The firm moved through seven distinct clusters between 1983 and 2011, with transitions often coinciding with global economic shocks.

Firm(s)	**Dominant Cluster(s)**
Walmart, Target,Costco, Home Depot,Kroger, Best Buy	11
Amazon	28 (25 yrs), 11 (3 yrs)
Ford, GM, Rivian	24
Tesla	24 (11 yrs), 18 (4 yrs)

Table 4: Cluster memberships of Amazon and Tesla relative to traditional peers.

5 Discussion

This paper set out to evaluate whether clustering firms by analyst coverage can reveal meaningful economic structure and whether such structure can be leveraged to predict future returns. Our results provide a clear answer to the first question, but a more sobering one to the second.

On the structural side, the clustering reveals persistent and economically interpretable communities. Retailers consistently co-locate in Cluster 11, automakers in Cluster 24, and technology firms in Cluster 18. Hybrid identities are also captured: Amazon alternates between retail and Big Tech communities, while Tesla is primarily placed with automakers but occasionally joins the technology cluster alongside Apple and Nvidia. These findings suggest that analysts' coverage decisions embed an implicit

classification system that complements and refines standard industry codes such as SIC and NAICS.

At the same time, we identify a large residual cluster (Cluster 28), which acts as an absorbing group for thousands of firms with diffuse or unstable coverage. Its sheer size and heterogeneity make it less economically interpretable, and future work should explore methods to better handle such residual categories.

When turning to prediction, however, the results are much more limited. Across three modeling approaches—XGBoost using cross-sectional features, ARIMA on cluster-level time series, and VAR allowing for cross-cluster spillovers—all models deliver negative R^2 values in out-of-sample tests. ARIMA achieves the lowest error, but even here predictive power is weaker than a naive mean benchmark. Excluding the residual Cluster 28 does not improve performance.

These findings reinforce a central theme of financial economics: while analysts' coverage decisions uncover meaningful economic structure, forecasting future returns remains an inherently difficult mission. The efficient market hypothesis remains intact: structure can be described and interpreted, but it cannot easily be exploited for systematic prediction.

Future work could extend these results by exploring dynamic embedding methods that smooth cluster trajectories over time, incorporating analyst report text to add semantic context, or applying the methodology to international markets. Such extensions may provide further insights into the structure of analyst coverage, even if return predictability remains elusive.

6 Conclusions and Future Work

Future research could improve on these results in several ways. First, richer feature sets may prove more informative than returns alone: analyst coverage networks contain additional signals such as earnings forecast errors, forward guidance revisions, analyst reputation, and textual sentiment from reports. Second, alternative approaches to dimensionality reduction may extract more stable predictive structure, including temporal graph embeddings or dynamic node representations that evolve smoothly over time. Third, hybrid models that combine clustering with modern deep learning techniques (e.g., graph neural networks or attention mechanisms) may capture dependencies missed by static Node2Vec embeddings. Finally, expanding the analysis beyond U.S. firms to international markets would allow us to test whether the structural and predictive patterns identified here generalize more broadly.

In sum, clustering by analyst coverage uncovers clear economic structure and firm-level dynamics, but predicting financial outcomes based on this structure remains elusive. The approach is therefore best understood as a tool for classification and exploration, while future extensions incorporating richer features, temporal models, and alternative graph methodologies hold promise for deeper insights.

References

Usman Ali and David Hirshleifer. Shared analyst coverage: Unifying momentum spillover effects. *Journal of Financial Economics*, 136(3):649–675, 2020.

Aditya Grover and Jure Leskovec. node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 855–864. ACM, 2016.