When Is Aggregate Peer Information More Informative?*

Kanying Xu†

This Draft: August 2, 2026

Abstract

I show that the informativeness of aggregate peer signals depends on the information span of the peer set. To proxy information span empirically, I construct Peer Information Dispersion (PID) from the distribution of analyst-linked peer exposure across industries. I document a non-monotonic pattern: peer signals are more informative about future profitability when peers are moderately dispersed across industries. This informativeness arises from aggregating individually weak but economically relevant signals beyond historical common components. At moderate dispersion, peer signals are associated with stronger short-term price responses, little subsequent return drift, and greater convergence in analyst beliefs.

Keywords: Peer Signals, Non-Monotonicity, Information Span, Prices, Analyst Beliefs

JEL Classification: G10, D83, G12, G40

*I thank Christian Schlag, Cecilia Parlatore, Michael Weber, Falko Fecht, and Daniel Streitz, as well as participants at the 3rd Frankfurt Summer School, for helpful comments and suggestions. I also thank Matteo Vacca, Mikko Leppämäki, and participants at the 2026 GSF Summer Workshop in Finland for valuable discussions and feedback. All remaining errors are my own.

†University of Vaasa, Wolffintie 32, 65200 Vaasa, Finland. Email: kanying.xu@uwasa.fi.

1 Introduction

Information from economically related firms can reveal economic conditions relevant to a focal firm and consequently affect its stock price (Foster, 1981; Ramnath, 2002; Hou, 2007; Cohen and Frazzini, 2008; Honkanen and Schmidt, 2022). Managers may also learn from peers' stock prices when making corporate investment decisions (Foucault and Frésard, 2014). Market participants often interpret signals from a set of related firms jointly. These signals may contain overlapping information, complement one another when combined, or provide incremental information beyond the other signals. Their informational value may therefore depend on how they relate to one another.

However, the existing literature provides less insight into the informational value of the constructed aggregate and of the component signals it contains. Much of this literature takes a constructed aggregate measure of peer information as its starting point[1] and examines how market participants incorporate it into expectations and prices (Ramnath, 2002; Hou, 2007; Cohen and Frazzini, 2008; Cohen and Lou, 2012; Ali and Hirshleifer, 2020). These studies focus on why peer information is incorporated gradually and offer different explanations, including limited attention, investor specialization, and market segmentation (Hong and Stein, 1999; Hong et al., 2000; Peng and Xiong, 2006; Menzly and Ozbas, 2010). By contrast, I unpack the aggregate peer signal and examine how its informativeness varies with the information span of the underlying peer set.

The information span of the peer set may matter because it captures a distinct property of the underlying peer information set. When peer signals are drawn from a narrow set of economic sources, the aggregate may contain more overlapping information

[1]In this paper, an aggregate measure of peer information refers to a measure that combines multiple public signals from related firms into a single signal about a focal firm. In the classic market information aggregation literature, aggregation instead refers to agents' dispersed private signals about a common payoff-relevant state being aggregated through market interactions and reflected in equilibrium prices (Grossman, 1976; Grossman and Stiglitz, 1980; Hellwig, 1980; Diamond and Verrecchia, 1981; Vives, 2008).

about conditions common to those sources. Drawing signals from a broader set of sources may expand the information span of the aggregate by adding information about economic conditions not captured by the existing signals.[2] Beyond some point, however, additional sources may contribute signals that are less relevant to the focal firm's future fundamentals. An important question is therefore: How does the information span of a firm's peer set affect the informativeness of the aggregate peer signal?

To examine this question, I construct Peer Information Dispersion (PID), a firm-month measure of the breadth of economic sources represented in a firm's peer set. Importantly, PID captures information span rather than signal precision: it measures how dispersed the firm's exposure to shared-analyst peers is across industries.[3] I construct PID by combining shared-analyst coverage with industry classifications. I use shared-analyst coverage to define the peer set because it captures multiple dimensions of economic relatedness, including similarities in business models (Kaustia and Rantala, 2021), as well as industry, geographic, customer–supplier, business-segment, and technology links (Ali and Hirshleifer, 2020).[4] Industry classifications provide a natural way to group the identified peers into economically meaningful source categories, as firms within an industry tend to share broad demand conditions, input costs, technologies, and valuation benchmarks. Thus, shared-analyst links determine which peer signals enter the

[2]This intuition is informed by theoretical work on multidimensional and correlated information (Lou et al., 2019; Siga and Mihm, 2021). I apply it to public signals from economically related firms.

[3]I draw on the distinction between information span and signal precision in He et al. (2024). In their credit-market framework, information span captures the breadth of payoff-relevant fundamental states covered by hard information, whereas signal precision captures the accuracy with which a given set of states is assessed. In my setting, I apply an analogous notion of span to the peer information set and proxy for it using the cross-industry dispersion of analyst-linked peer exposure.

[4]Bartov et al. (2026) show that the detection of earnings-announcement spillovers depends on peer identification, the directional heterogeneity of peer responses, and sophisticated investor ownership. Their analysis focuses on identifying economically related peers and measuring announcement-period spillovers. In contrast, I use a consistent peer-identification rule based on shared-analyst coverage and examine how the cross-industry distribution of multiple signals within each resulting peer set shapes the informativeness of their aggregate for the focal firm's future fundamentals.

information set, while the cross-industry dispersion of those links captures the breadth of economic sources represented in that set.

To formalize how PID can affect the informativeness of the aggregate peer signal, I develop a simple Bayesian signal-extraction framework that links the information span of a firm's peer set to the informativeness of the aggregate peer signal. Each peer signal contains a component related to the focal firm's future fundamentals and a nonfundamental component that may be correlated across peers. Distributing peer information more broadly across industries can increase the precision of the aggregate signal by reducing the covariance among its nonfundamental components. Greater dispersion, however, may place more weight on industries that are less economically related to the focal firm, reducing the aggregate signal's loading on the firm's future fundamentals. The framework shows that aggregate peer information can be more informative at an interior level of PID when the first effect dominates at low PID and the second dominates at high PID.

I test this prediction using a panel of U.S. public firms from 2000 to 2024. The empirical analysis delivers three main findings. First, the aggregate peer signal, constructed as a weighted average of peer earnings surprises,[5] contains incremental information about the focal firm's future fundamentals beyond its own public information. Its predictive content follows an inverted U-shaped pattern: it increases as PID rises from low to moderate levels and declines as PID rises further. To assess whether this pattern extends beyond in-sample fit, I conduct recursive out-of-sample tests. These tests produce the same pattern, with the largest reduction in forecast errors occurring at moderate levels of PID. At moderate levels of PID, a one-standard-deviation increase in the aggregate peer signal predicts a 6-basis-point increase in average future profitability change. This estimate is 2.5 times the corresponding full-sample estimate and represents

[5]Peer weights are proportional to the number of analysts each peer shares with the focal firm and sum to one within each firm-month.

approximately 11% of the median absolute profitability change in the estimation sample.[6] The magnitude is economically meaningful given that the aggregate peer signal is constructed from peer earnings surprises and predicts the focal firm's future fundamentals beyond its own public information.

Second, I show that, when peers are moderately dispersed across industries, the informativeness of the aggregate signal reflects the aggregation of multiple economically relevant peer signals that are not fully captured by historical common components. I provide four pieces of evidence. First, I separate the dominant peer[7] from the aggregate of the remaining peers. The dominant peer has little information content, while the aggregate of the remaining peers remains informative. Second, I use a rolling principal component analysis to separate historical patterns of co-movement from the peer signals.[8] At moderate levels of dispersion, the leading historical components have little information content, while the remaining component remains informative. Thus, stable historical co-movement does not account for the result. Third, I randomly sample peer subsets of increasing size. Individual peer signals have little information content, but the aggregate becomes more informative as additional peer signals are included and is most informative when the full peer set is used. This pattern is absent when peers are highly concentrated or highly dispersed across industries. Finally, at moderate levels of dispersion, the aggregate signal becomes more informative as the economic relevance of the peers to the focal firm increases. Together, these results show that the greater informativeness of the aggregate peer signal at moderate PID arises from combining multiple individually weak but economically relevant peer signals and extends beyond the variation captured by historical common components.

[6]The median absolute profitability change in the estimation sample is 0.554 percentage points; thus, $0.060/0.554 \approx 0.11$.

[7]The dominant peer is the peer with the largest shared-analyst weight.

[8]For each month, the principal components are estimated from peer earnings surprises over the preceding 60 months. Peer signals are then separated into components associated with these historical patterns and a remaining component.

The economic intuition behind the findings is straightforward. Consider a retailer whose peers are concentrated among other retailers. Their earnings signals may reflect the same industry conditions and therefore contain overlapping information. Their nonfundamental components may also be correlated, limiting the reduction in noise from aggregation. Each peer signal may still be informative about the focal retailer, but combining similar signals adds relatively little information. Now consider a peer set distributed across several industries that remain economically relevant to the retailer, such as consumer-goods production, logistics, and payment services. Signals from these industries provide information about different aspects of the retailer's business, including product demand, inventory flows, distribution costs, and transaction activity. Aggregating these signals can increase informativeness because they contain distinct relevant information and less correlated nonfundamental variation. As the peer set extends into more economically distant industries, however, the additional signals may become less relevant to the retailer. The loss of relevance can then offset the benefit from combining signals across industries.

Third, I examine the market relevance of aggregate peer information using stock-price responses and subsequent changes in analyst disagreement. Previous research relates announcement-period price changes to the information conveyed by an announcement and analyst forecast dispersion to uncertainty and consensus (Kim and Verrecchia, 1991; Barron et al., 1998). I show that, when peers are moderately dispersed across industries, a one-standard-deviation increase in the aggregate peer signal is associated with abnormal returns of 0.084 percentage points through trading day 5 and 0.091 percentage points through trading day 10, with little subsequent return drift. The corresponding price responses are smaller when peers are highly concentrated or highly dispersed. I also show that, at moderate dispersion, a larger weighted average of absolute earnings surprises across peers is followed by lower analyst disagreement, particularly three to six months

after the news becomes available. Disagreement rises at low dispersion and changes little at high dispersion. Thus, at moderate dispersion, aggregate peer information has greater market relevance: prices respond more strongly to the aggregate peer signal, and larger absolute earnings surprises are followed by greater convergence in analyst beliefs.

The contributions of the paper are threefold. First, the paper contributes to the literature on peer effects and information diffusion. A common approach in this literature is to identify peers through an economic network, aggregate their returns or earnings signals, and examine how the resulting signal is incorporated into expectations and prices (Hou, 2007; Cohen and Frazzini, 2008; Menzly and Ozbas, 2010; Aobdia et al., 2014; Ali and Hirshleifer, 2020; Avramov et al., 2026). Rather than treating the constructed aggregate as a homogeneous signal, I unpack its underlying peer information set and introduce PID, an economically interpretable measure of the breadth of economic sources represented in that set.

Second, I distinguish among the mechanisms underlying the informativeness of aggregate peer information. Previous studies document the predictive content of signals constructed from multiple economically related firms (Cohen and Lou, 2012; Ali and Hirshleifer, 2020), but this evidence does not identify how the information content of the aggregate is generated. It may come primarily from a dominant peer, reflect common variation across peers, or accumulate across multiple component signals. I distinguish among these explanations and provide evidence consistent with information accumulating across multiple economically relevant peer signals.

Finally, the paper provides an integrated assessment of the market relevance of aggregate peer information. I examine price responses and analyst disagreement to assess how peer information is reflected in prices and analyst beliefs. These tests connect the fundamental informativeness of aggregate peer information to its broader market relevance.

2 A Framework for Aggregate Peer Information

This section develops a parsimonious framework for learning from multiple peer signals. The framework shows that the precision of the aggregate peer signal depends jointly on its loading on the focal firm's future fundamentals and the variance of its nonfundamental component. It then characterizes sufficient conditions under which precision is maximized at an intermediate level of PID and provides an illustrative relevance-decay environment in which such an interior optimum can arise. Proofs are provided in Appendix B.

2.1 Peer Signals and Information Precision

Let $\mathcal{F}_i$ denote the information available before the arrival of peer signals, and let Y_i denote firm i's subsequent fundamental outcome. Define

$$u_i \equiv Y_i - E[Y_i \mid \mathcal{F}_i],$$

the component of the outcome not predictable from existing information. Conditional on $\mathcal{F}_i$,

$$u_i \sim N(0, \tau_u^{-1}), \qquad \tau_u > 0.$$

Firm i is associated with N_i peers whose signals satisfy

$$\boldsymbol{y}_i = \boldsymbol{b}_i u_i + \boldsymbol{\eta}_i, \qquad \begin{pmatrix} u_i \\ \boldsymbol{\eta}_i \end{pmatrix} \middle| \, \mathcal{F}_i \sim N\left[\begin{pmatrix} 0 \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \tau_u^{-1} & \mathbf{0}' \\ \mathbf{0} & \Sigma_{\eta,i} \end{pmatrix} \right], \tag{1}$$

where $\Sigma_{\eta,i}$ is positive definite. The vector $\boldsymbol{b}_i$ contains the loadings of the peer signals on u_i, while $\boldsymbol{\eta}_i$ contains variation unrelated to u_i. The elements of $\boldsymbol{\eta}_i$ may be correlated

because peer signals can reflect common underlying sources. This multivariate Gaussian structure is standard in models of information aggregation (e.g., Vives, 2008).

Let

$$s_i = \boldsymbol{a}_i' \boldsymbol{y}_i$$

denote an aggregate of the peer signals, where

$$a_{ij} \geq 0, \qquad \sum_{j=1}^{N_i} a_{ij} = 1.$$

The vector $\boldsymbol{a}_i$ is predetermined conditional on $\mathcal{F}_i$ and is not selected within the model to maximize signal precision. The aggregate signal can be written as

$$s_i = B_i u_i + \varepsilon_i, \qquad B_i \equiv \boldsymbol{a}_i' \boldsymbol{b}_i, \qquad \varepsilon_i \equiv \boldsymbol{a}_i' \boldsymbol{\eta}_i, \tag{2}$$

where

$$\Omega_i \equiv \mathrm{Var}(\varepsilon_i \mid \mathcal{F}_i) = \boldsymbol{a}_i' \Sigma_{\eta,i} \boldsymbol{a}_i.$$

For $B_i \neq 0$, normalize the aggregate signal as

$$m_i \equiv \frac{s_i}{B_i} = u_i + \omega_i, \qquad \omega_i \equiv \frac{\varepsilon_i}{B_i}.$$

Because

$$\mathrm{Var}(\omega_i \mid \mathcal{F}_i) = \frac{\Omega_i}{B_i^2},$$

the precision of m_i about u_i is

$$\kappa_i = \frac{B_i^2}{\Omega_i} = \frac{(\boldsymbol{a}_i' \boldsymbol{b}_i)^2}{\boldsymbol{a}_i' \Sigma_{\eta,i} \boldsymbol{a}_i}. \tag{3}$$

Thus, signal precision depends on the aggregate loading on u_i and on the covariance structure of the remaining variation.

2.2 Aggregation Across Peer Signals

To illustrate the role of aggregation, consider a symmetric special case of the general signal structure in Section 2.1. Each peer signal has the same loading on u_i:

$$y_{ij} = b_i u_i + \eta_{ij},$$

where

$$\mathrm{Var}(\eta_{ij} \mid \mathcal{F}_i) = \sigma^2_{\eta,i}, \qquad \mathrm{Corr}(\eta_{ij}, \eta_{ik} \mid \mathcal{F}_i) = \rho_i \in [0, 1)$$

for $j \neq k$. Consider the equal-weight aggregate

$$\bar{y}_i(N) = \frac{1}{N} \sum_{j=1}^{N} y_{ij} = b_i u_i + \bar{\eta}_i(N),$$

where

$$\mathrm{Var}(\bar{\eta}_i(N) \mid \mathcal{F}_i) = \sigma^2_{\eta,i} \left[\rho_i + \frac{1 - \rho_i}{N} \right].$$

Its precision about u_i is

$$\kappa_i(N) = \frac{b_i^2}{\sigma^2_{\eta,i} \left[\rho_i + \frac{1-\rho_i}{N} \right]}. \tag{4}$$

Proposition 1. *If $b_i \neq 0$, $\sigma^2_{\eta,i} > 0$, and $\rho_i < 1$, then*

$$\kappa_i(N+1) > \kappa_i(N).$$

Aggregation leaves the loading on u_i unchanged while reducing the variance of the aggregate nonfundamental component. The reduction is greater when the underlying

components are less correlated. Correlation across peer signals therefore limits the precision gain from aggregation.

Proposition 1 holds because every peer signal has the same loading b_i on u_i. Once peer signals differ in their relevance to the focal firm, adding a more distant source can reduce the aggregate loading as well as the nonfundamental variance. The net effect is then ambiguous.

2.3 Peer Information Dispersion

Suppose the peers of firm i are partitioned into G_i underlying sources. Let $\mathcal{G}_{ig}$ denote the set of peers associated with source g, and define

$$\pi_{ig} \equiv \sum_{j \in \mathcal{G}_{ig}} a_{ij}.$$

Thus, π_{ig} is the share of aggregate peer weight associated with source g, where

$$\pi_{ig} \geq 0, \qquad \sum_{g=1}^{G_i} \pi_{ig} = 1.$$

Peer information dispersion is defined as

$$PID_i = 1 - \sum_{g=1}^{G_i} \pi_{ig}^2 \in \left[0, 1 - \frac{1}{G_i}\right]. \tag{5}$$

PID is low when peer information is concentrated among a small number of sources and high when it is distributed more evenly across sources.

PID is not a sufficient statistic for signal precision: information sets with the same PID may differ in their loadings and covariance structures. Consider a differentiable path indexed by p, satisfying

$$1 - \sum_g \pi_{ig}(p)^2 = p.$$

Along this path,

$$\kappa_i(p) = \frac{B_i(p)^2}{\Omega_i(p)},$$

where $B_i(p)$ and $\Omega_i(p)$ denote the aggregate loading and nonfundamental variance.

An increase in PID may reduce the nonfundamental variance of the aggregate signal by distributing information across less correlated sources. It may also reduce the aggregate loading on u_i. Define

$$D_i(p) \equiv -\frac{d \log \Omega_i(p)}{dp}, \qquad L_i(p) \equiv -\frac{d \log B_i(p)^2}{dp}. \tag{6}$$

It follows that

$$\frac{d \log \kappa_i(p)}{dp} = D_i(p) - L_i(p). \tag{7}$$

Proposition 2 (Sufficient conditions for a unique interior optimum)**.** *Let*

$$\mathcal{P}_i = [\underline{p}_i, \overline{p}_i] \subseteq \left[0, 1 - \frac{1}{G_i}\right],$$

with $B_i(p) \neq 0$ and $\Omega_i(p) > 0$ throughout $\mathcal{P}_i$. Suppose that $D_i(p)$ is continuous and strictly decreasing and that $L_i(p)$ is continuous and strictly increasing. If

$$D_i(\underline{p}_i) > L_i(\underline{p}_i) \qquad \textit{and} \qquad D_i(\overline{p}_i) < L_i(\overline{p}_i),$$

then there exists a unique $p_i^* \in (\underline{p}_i, \overline{p}_i)$ *such that*

$$D_i(p_i^*) = L_i(p_i^*).$$

Signal precision increases for $p < p_i^*$ *and decreases for* $p > p_i^*$.

Under these sufficient conditions, the reduction in nonfundamental variance dominates at low PID, whereas the decline in the loading on u_i dominates at high PID. Signal precision is therefore maximized at an interior level of PID. Proposition 2 is a conditional characterization: the monotonicity and endpoint conditions describe the information environment rather than follow solely from the definition of PID in equation (5).

The next subsection provides an explicit information environment in which the two forces arise from differences in peer relevance and common exposure across economic sources.

2.4 Relevance Decay Across Sources

Order the G_i sources of firm i by their economic distance from the focal firm, with $g = 1$ denoting the closest source. Suppose that the loading and common-factor exposure of source g satisfy

$$b_g = b_0 x^{g-1}, \qquad c_g = \sqrt{\rho_0}\, y^{g-1}, \qquad x, y \in (0,1), \quad \rho_0 \in (0,1). \tag{8}$$

The parameter x governs how quickly the relevance of peer information declines with economic distance, while y governs how quickly common exposure declines. Values of x and y close to one correspond to slow decay.

The nonfundamental component of source g follows

$$\eta_g = c_g f + \sqrt{1 - c_g^2}\, e_g, \tag{9}$$

where f and $e_1, \ldots, e_{G_i}$ are mutually independent and each has variance σ^2. Hence,

$$\mathrm{Var}(\eta_g) = \sigma^2$$

and, for $g \neq h$,

$$\mathrm{Corr}(\eta_g, \eta_h) = c_g c_h.$$

The covariance matrix is therefore positive definite by construction.

Consider an equal-weight aggregate of the k sources closest to the focal firm, where $k \in \{1, \ldots, G_i\}$. Because each included source receives weight $1/k$, equation (5) gives

$$PID(k) = 1 - \frac{1}{k}.$$

The aggregate loading and nonfundamental variance are

$$B(k) = \frac{b_0}{k} \frac{1 - x^k}{1 - x}, \tag{10}$$

and

$$\Omega(k) = \frac{\sigma^2}{k} + \frac{\sigma^2 \rho_0}{k^2} \left[\frac{(1 - y^k)^2}{(1 - y)^2} - \frac{1 - y^{2k}}{1 - y^2} \right]. \tag{11}$$

Signal precision is

$$\kappa(k) = \frac{B(k)^2}{\Omega(k)}.$$

Proposition 3 (Relevance decay and the value of dispersion)**.** *Under equations* (8)–(9)*, the following results hold for every* $x, y \in (0, 1)$ *and* $\rho_0 \in (0, 1)$*:*

(i) As the number of available sources grows,

$$\lim_{k\to\infty} \kappa(k) = 0.$$

(ii) Adding a second source raises signal precision if and only if

$$\kappa(2) > \kappa(1) \quad \Longleftrightarrow \quad x^2 + 2x - 1 > 2\rho_0 y.$$

(iii) If the inequality in part (ii) holds, then there exists a finite $\bar{k} \geq 3$ such that

$$\kappa(k) < \kappa(2) \qquad \textit{for all } k \geq \bar{k}.$$

Consequently, whenever $G_i \geq \bar{k}$, signal precision attains its maximum over $k \in \{1, \ldots, G_i\}$ at some

$$k^* \in \{2, \ldots, \bar{k} - 1\}.$$

Part (i) shows that spreading information across an increasing number of progressively less relevant sources eventually eliminates signal precision. Although diversification continues to reduce source-specific noise, the aggregate loading on the focal firm's fundamental innovation becomes too small.

Part (ii) characterizes when the initial diversification gain is positive. Adding a second source is beneficial when its incremental relevance, governed by x, is sufficiently large relative to the remaining common exposure, governed by y and ρ_0. Unlike the symmetric benchmark in Proposition 1, adding another source therefore need not raise precision.

Part (iii) combines these two results. If moving from one source to two sources raises precision, but precision eventually approaches zero, then neither minimal nor

maximal dispersion is optimal when the available source set is sufficiently rich. The model therefore produces an interior optimum from an economically interpretable relevance-decay structure, together with the threshold condition in part (ii) and a sufficiently large set of available sources. The proposition does not require the direct monotonicity assumptions on $D_i(p)$ and $L_i(p)$ imposed in Proposition 2, although it also does not imply uniqueness of the discrete maximizer.

2.5 Learning from the Aggregate Signal

Given the normalized signal m_i, Gaussian updating implies

$$E[u_i \mid m_i, \mathcal{F}_i] = \lambda_i m_i, \qquad \lambda_i \equiv \frac{\kappa_i}{\tau_u + \kappa_i}, \tag{12}$$

and

$$\mathrm{Var}(u_i \mid m_i, \mathcal{F}_i) = \frac{1}{\tau_u + \kappa_i}. \tag{13}$$

The coefficient λ_i also equals the proportional reduction in posterior variance and the squared conditional correlation between u_i and the aggregate signal:

$$\lambda_i = 1 - \frac{\mathrm{Var}(u_i \mid m_i, \mathcal{F}_i)}{\mathrm{Var}(u_i \mid \mathcal{F}_i)} = \mathrm{Corr}^2(u_i, s_i \mid \mathcal{F}_i). \tag{14}$$

Thus, λ_i measures both the weight placed on the normalized signal and its explanatory power for the component of subsequent fundamentals not predictable from $\mathcal{F}_i$. Since

$$\frac{d\lambda_i}{d\kappa_i} = \frac{\tau_u}{(\tau_u + \kappa_i)^2} > 0,$$

the reduction in uncertainty and explanatory power increase with signal precision. Under Proposition 2, or under the threshold and source-richness conditions in Proposition 3,

both are maximized at an interior level of PID. Neither proposition ranks their magnitudes at low and high PID. Empirically, I examine whether this non-monotonic implication is supported by the out-of-sample forecast improvement from aggregate peer information across Low, Mid, and High PID.

3 Data and Variable Definitions

I construct a monthly firm-level panel by merging I/B/E/S analyst forecasts and earnings surprises with CRSP stock-market data and Compustat accounting and business-segment data. The sample spans January 2000 through December 2024 and includes NYSE-, AMEX-, and NASDAQ-listed common stocks with CRSP share codes 10 and 11. I require an absolute CRSP share price of at least \$5 at the end of the preceding month.[9]

3.1 Constructing Analyst-Linked Peers

Following Ali and Hirshleifer (2020), I identify analyst-linked peers using the I/B/E/S Detail History file. Two firms are linked in a given month if they are covered by at least one common analyst, and the number of common analysts measures the strength of the link. I update the peer network monthly to match the timing of the disagreement and return variables. A daily network would be sparse because forecast revisions are infrequent and asynchronous across firms, whereas a quarterly network would obscure the timing between peer earnings news and subsequent changes in beliefs and prices.[10]

[9]The \$5 cutoff is commonly used to mitigate microstructure concerns associated with low-priced stocks (Stambaugh and Yuan, 2017; Müller and Schmickler, 2025). CRSP records prices based on bid-ask averages with a negative sign, so I use absolute prices.

[10]More than 90% of firm-quarters with at least two monthly observations exhibit within-quarter variation in analyst disagreement.

I use FY1 and FY2 EPS forecasts for U.S. firms.[11] To avoid look-ahead bias, forecasts issued through month $t-1$ determine the peer network in month t. Analyst a covers firm i in month t if the analyst issued at least one FY1 or FY2 EPS forecast for the firm during the preceding twelve months. Each analyst-firm pair is counted at most once per month, regardless of how many forecasts the analyst issued for the firm.

For $i \neq j$, define $SC_{ij,t}$ as the number of analysts covering both firms i and j in month t:

$$SC_{ij,t} = \sum_{a} \mathbf{1}\{a \text{ covers } i \text{ in } t\}\mathbf{1}\{a \text{ covers } j \text{ in } t\}. \tag{15}$$

Firm j is an analyst-linked peer of firm i when $SC_{ij,t} > 0$. In constructing the aggregate earnings signal, I assign each peer a weight proportional to $SC_{ij,t}$.

[Table 1]

Table 1 summarizes the analyst-linked peers in the sample. Firms with at least one analyst-linked peer constitute 85% of the sample and represent 99% of sample market capitalization. Peer information is therefore available for firms accounting for nearly all of the market capitalization in the sample.

Firms are also typically connected to many peers. Among firms with at least one peer, the median firm has 60 peers, with an interquartile range of 37 to 88. The aggregate earnings signal consequently incorporates information from multiple firms, while PID captures how this information is distributed across economic sources. The number of shared analysts also differs across firm pairs. Among linked firm pairs, the median number of shared analysts is one, and the 75th percentile is two. Since peer weights are proportional to $SC_{ij,t}$, peers sharing more analysts with the focal firm contribute more to its aggregate earnings signal.

[11] Including FY2 prevents coverage from being understated near fiscal-year ends, when FY1 forecasts approach realization.

Firms with analyst-linked peers tend to be larger, with a median market capitalization of $1,076.68 million, compared with $767.71 million for the full sample. This difference reflects the greater availability of analyst-linked peers among larger firms.

3.2 Measuring Aggregate Peer Signals

I construct an aggregate peer earnings signal from peers' earnings surprises. Previous research shows that earnings announcements can convey information to economically related firms (Foster, 1981; Ramnath, 2002; Thomas and Zhang, 2008; Pandit et al., 2011). I use the number of shared analysts to measure the strength of each peer link when aggregating these earnings surprises.

For each quarterly earnings announcement, I obtain the standardized unexpected EPS score from the I/B/E/S Surprise History file, denoted SUE.[12] Positive values indicate earnings above analyst expectations, and negative values indicate earnings below expectations. I assign each surprise to its announcement month and winsorize it by month at the 1st and 99th percentiles.

I define $PES_{i,t}$ using peer earnings information available as of the beginning of month t. It aggregates the quarterly earnings surprises announced by firm i's analyst-linked peers during month $t-1$. This timing ensures that the earnings surprises used to construct the signal are public before the future outcomes examined in the baseline regressions. Let $\widetilde{SUE}_{j,t-1}$ denote peer j's winsorized earnings surprise announced in month $t-1$, and let $\mathcal{P}^*_{i,t-1}$ denote the subset of analyst-linked peers with a valid earnings surprise in that month. The signal is

$$PES_{i,t} = \sum_{j \in \mathcal{P}^*_{i,t-1}} \omega_{ij,t-1} \widetilde{SUE}_{j,t-1}, \tag{16}$$

[12] I/B/E/S constructs `SUESCORE` from reported EPS and the corresponding consensus forecast.

where, following Ali and Hirshleifer (2020),

$$\omega_{ij,t-1} = \frac{SC_{ij,t-1}}{\sum_{k \in \mathcal{P}^*_{i,t-1}} SC_{ik,t-1}}. \tag{17}$$

Among peers with a valid earnings surprise in month $t-1$, those sharing more analysts with the focal firm receive greater weight. The weights sum to one within $\mathcal{P}^*_{i,t-1}$ for each firm-month. I set $PES_{i,t}$ to missing when none of firm i's analyst-linked peers has a valid earnings surprise in month $t-1$.[13] Thus, $PES_{i,t}$ is a realized monthly signal generated by the subset of analyst-linked peers that announce earnings in month $t-1$.

3.3 Measuring Peer Information Dispersion (PID)

PES is a realized monthly signal constructed from the earnings surprises of analyst-linked peers that announced earnings in the preceding month. PID instead serves as an empirical proxy for the information span of the focal firm's full analyst-linked peer set, capturing how its peer exposure is distributed across economic sources. In the empirical implementation, I define economic sources using peers' two-digit SIC industries.

For firm i in month t, let $\mathcal{P}_{i,t}$ denote its set of analyst-linked peers. I group these peers by two-digit SIC industry because industry membership separates signals associated with different areas of economic activity, consistent with evidence that earnings information is transmitted across firms in the same industry (Foster, 1981; Ramnath, 2002; Piotroski and Roulstone, 2004).[14]

[13] The absence of a peer earnings announcement is distinct from an announced earnings surprise of zero.

[14] One-digit SIC sectors are too broad to separate many economically different peer firms, whereas three- and four-digit SIC classifications often leave too few firms to construct reliable monthly industry earnings signals. Two-digit SIC industries provide an intermediate level of aggregation that preserves economically meaningful differences while maintaining adequate monthly coverage. Because industry-classification schemes differ in how well they capture economic similarity (Bhojraj et al., 2003), Internet Appendix IA.1.2 reconstructs PID using GICS industries while holding the analyst-linked peer set and shared-analyst weights fixed. The main results are preserved under this alternative classification.

Let $\mathcal{S}_{i,t}$ denote the set of SIC2 industries represented among firm i's analyst-linked peers. The share of its shared-analyst exposure attributable to peer industry s is

$$w_{i,s,t} = \frac{\sum_{j \in \mathcal{P}_{i,t}} SC_{ij,t} \mathbf{1}\{j \in s\}}{\sum_{j \in \mathcal{P}_{i,t}} SC_{ij,t}}. \tag{18}$$

Here, $SC_{ij,t}$ is the number of analysts covering both firms. I define peer information dispersion as

$$PID_{i,t} = 1 - \sum_{s \in \mathcal{S}_{i,t}} w_{i,s,t}^2. \tag{19}$$

PID equals zero when all peer exposure comes from one industry and increases as exposure becomes more evenly distributed across industries. If exposure is equally distributed across G industries, then $PID_{i,t} = 1 - 1/G$.[15]

PID is constructed from all analyst-linked peers, including peers that do not announce earnings in month $t-1$. By contrast, $PES_{i,t}$ is generated by the subset of analyst-linked peers that announce earnings in month $t-1$. Using the full peer network makes PID an ex ante characteristic of the peer information environment and prevents it from changing mechanically with firms' fiscal calendars and earnings-announcement timing.

[Table 2]

Table 2 describes how firms differ across the contemporaneous monthly PID distribution. For this descriptive analysis, firms are assigned each month to low-, mid-, and high-PID terciles based on their position in that month's PID distribution.[16]

[15] PID is one minus the Herfindahl concentration of peer-industry exposure shares (Hirschman, 1945; Herfindahl, 1950). The same functional form is used in network science as the participation coefficient, which measures how a node's links are distributed across network modules (Guimerà and Amaral, 2005). Here, peer industries correspond to modules and shared-analyst exposure corresponds to weighted links.

[16] The tercile cutoffs are calculated separately each month. These descriptive groups are based on contemporaneous monthly PID.

Panel A verifies that the PID groups differ in the industry distribution of analyst-linked peers. Peer exposure is concentrated in the focal firm's industry in the low-PID group and extends across more industries as PID increases. This difference does not appear to reflect the number of peers alone. The low- and mid-PID groups have similar average peer counts but substantially different levels of dispersion. PID therefore captures how peer exposure is distributed across industries in addition to the size of the peer set.

Panels B and C examine the extent to which PID is associated with observable firm characteristics and broad differences across industries and time. Panel B shows that the PID groups differ in several firm characteristics, raising the possibility that these characteristics account for some of the variation in PID. Panel C evaluates their joint explanatory power. The included firm characteristics explain 6% of the variation in monthly PID rank. Adding month and SIC2 fixed effects raises the adjusted R^2 to 54%. Thus, the included firm characteristics, industry fixed effects, and month fixed effects account for slightly more than half of the variation in monthly PID rank but do not subsume it. Substantial within-industry and within-month variation in PID remains after accounting for the included firm characteristics.

Panel D relates PID to two dimensions of the peer information environment using measures of historical industry earnings correlation. The first compares the earnings signal of each peer industry with that of the focal firm's industry and measures how closely the peer industries are related to the focal firm's economic environment. The second compares the earnings signals of different peer industries with one another and

measures the common historical variation among the industries represented in the peer information set. Both measures use shared-analyst exposure weights.[17]

The group means of both correlation measures decline as PID rises. Peer industries in higher-PID sets are, on average, less closely related to the economic activity of the focal firm and share less common historical earnings variation with one another. Panel D therefore shows that PID is associated with both the connection of peer industries to the focal firm and the common variation among the peer industries. These descriptive relations provide economic content for PID and motivate the subsequent informativeness tests; they do not by themselves establish that an intermediate level of PID is optimal.

Figure 1 provides a visual example using analyst-linked peers in December 2024. Each panel contains a focal firm, shown in black, and its analyst-linked peers, colored by two-digit SIC industry. The examples have similar numbers of peers, facilitating a comparison of how peer exposure is distributed across industries. Low PID corresponds to exposure concentrated in a small number of industries, whereas high PID corresponds to exposure distributed across many industries.

[Figure 1]

Figure 2 summarizes the distribution and time-series pattern of PID in the regression sample. Panel (a) shows substantial cross-sectional variation, including firms whose peer exposure is concentrated in one industry and firms whose peer exposure is spread across many industries. Panel (b) shows that the average level of PID also changes over time,

[17] Let $g(i)$ denote the SIC2 industry of firm i, and define $\rho^{+}_{a,b,t} = \max\{\rho_{a,b,t}, 0\}$, where $\rho_{a,b,t}$ is the historical correlation between the earnings signals of industries a and b. The two measures are

$$\frac{\sum_{s\neq g(i)} w_{i,s,t-1}\rho^{+}_{g(i),s,t}}{\sum_{s\neq g(i)} w_{i,s,t-1}} \quad \text{and} \quad \frac{\sum_{s\neq r} w_{i,s,t-1}w_{i,r,t-1}\rho^{+}_{s,r,t}}{\sum_{s\neq r} w_{i,s,t-1}w_{i,r,t-1}}.$$

The first excludes the focal firm's industry, and the second excludes correlations of an industry with itself. The weights are renormalized after these exclusions. Correlations are estimated over the preceding 60 months, require at least 24 valid observations, and use exposure weights from month $t-1$. A measure is set to missing when its denominator is zero.

especially after 2020. This time-series movement motivates the use of monthly PID ranks and monthly terciles in the main tests. The main specifications therefore compare firms within the same calendar month, while robustness tests consider alternative PID group definitions and alternative sample periods. In particular, Table IA.6 reports estimates for both a pre-2020 sample and a sample excluding 2020–2021.

[Figure 2]

3.4 Control Variables

The control variables follow prior research on peer effects and information diffusion (Hou, 2007; Menzly and Ozbas, 2010; Cohen and Lou, 2012; Hameed et al., 2015). In the empirical analyses, I control for size, book-to-market, turnover, return volatility, business concentration, and analyst coverage to account for differences in firm characteristics, trading activity, business complexity, and the information environment. Where relevant, I also control for profitability, past returns, analyst disagreement, analyst consensus, and the focal firm's own earnings surprise to separate peer information from information about the focal firm. The exact set of controls is reported in each table. Detailed variable definitions are provided in Appendix A.

3.5 Summary Statistics

Table 3 reports summary statistics for the main variables used in the empirical analysis.[18]

[Table 3]

[18]The number of observations varies because the construction of the variables requires different data sources. PES is observed when at least one peer announces earnings in the preceding month, while PID requires information on the industry distribution of the peer set.

The aggregate peer earnings signal, PES, and peer information dispersion, PID, both exhibit substantial variation across firm-months. The standard deviation of PES is 2.09, while the interquartile range of PID extends from 0.23 to 0.79. The sample therefore provides variation in both the aggregate signal and the distribution of its component signals across industries. The remaining variables summarize the outcomes and firm characteristics used in the empirical specifications.

Table 4 next examines how the two peer information measures relate to each other, to the focal firm's earnings surprise, and to the other variables used in the analysis.

[Table 4]

The correlation between PES and PID is 0.091, indicating that the realized peer earnings signal is only weakly associated with the cross-industry dispersion of the focal firm's analyst-linked peer environment. PES measures the signed aggregate signal generated by peers that announced earnings in the preceding month, whereas PID measures how the focal firm's full analyst-linked peer exposure is distributed across industries. This distinction allows the empirical tests to examine whether the informativeness of the realized peer signal varies with the breadth of the underlying peer information environment.

PES has a positive but modest correlation with the focal firm's earnings surprise. The empirical specifications control for the focal firm's earnings surprise to isolate the incremental information contained in PES. They also control for profitability and other firm characteristics that may be related to PID.

4 Empirical Results

In this section, I present four sets of empirical tests. I examine whether the aggregate peer earnings signal predicts the focal firm's subsequent profitability changes

and whether its informativeness varies non-monotonically with PID (Section 4.1); whether the same non-monotonic pattern holds out of sample (Section 4.2); whether this informativeness arises from aggregating information across multiple economically relevant peers (Section 4.3); and whether aggregate peer information is reflected in stock prices and analyst disagreement (Section 4.4).

4.1 Non-monotonic Informativeness of Aggregate Peer Signal

In this section, I examine whether the aggregate peer earnings signal predicts the focal firm's subsequent profitability changes and whether its informativeness varies non-monotonically with PID. I measure subsequent profitability changes using the average change from quarter q to quarters $q+1$, $q+2$, $q+3$, and $q+4$. Averaging across these horizons captures fundamental information that may be realized at different points over the following year and reduces the influence of transitory quarterly fluctuations. I also estimate the specifications separately for each horizon to examine the timing and persistence of the relation. Table IA.9 reports separate estimates for horizons $q+1$ through $q+4$.

I build the control structure sequentially. The first set controls for the fundamental conditions of the focal firm and its industry. The second set controls for the focal firm's information environment. The third set further controls for firm characteristics, market characteristics, and characteristics of the peer set. All specifications include firm fixed effects and month fixed effects. The coefficients therefore measure the incremental explanatory power of the aggregate peer earnings signal for subsequent profitability changes after controlling for the variables described above, time-invariant firm characteristics, and common month shocks.

For the heterogeneity analysis, I measure PID using information available before the regression month. For each firm-month, I average monthly PID over the preceding

twelve months and require at least nine nonmissing observations. I then calculate its percentile rank within each month and assign firms to Low, Mid, and High PID groups based on monthly terciles.[19]

Table 5 reports regressions of future profitability changes on the aggregate peer earnings signal. Panel A first reports the pooled PES specification. PES continues to predict future profitability changes as fundamental controls, information controls, and firm and market controls are progressively added. This result establishes the baseline finding of this section: the aggregate peer earnings signal contains incremental information about future fundamentals beyond the included public information controls.

[Table 5]

Panel B examines whether this information content varies across the PID distribution. For firms in the Low PID group, PES is insignificant across control specifications, indicating limited incremental information when peer information is less dispersed across economic sources. For firms in the High PID group, PES has some predictive power when controls are sparse, but the estimate weakens as information controls and firm and market controls are added. Thus, greater dispersion of peer information does not necessarily produce a more informative signal about the focal firm.

The aggregate peer earnings signal is more informative for firms in the Mid PID group. Under full controls, a one-standard-deviation increase in PES is associated with a 6-basis-point increase in average future profitability change. This magnitude is economically meaningful, representing approximately 5.3% of the interquartile range and 2.3% of the standard deviation of the dependent variable. Under the same specification, the corresponding coefficients are only 0.008 percentage points for the Low PID group and 0.014 percentage points for the High PID group, and neither is statistically

[19]The nine-month requirement ensures that the measure is based on at least 75% of the preceding twelve months without requiring a complete twelve-month history.

significant. Wald tests reject equality between the Mid coefficient and each of the Low and High coefficients. Thus, the economic significance of the aggregate peer earnings signal depends importantly on how peer information is distributed across economic sources: the signal has meaningful predictive content at intermediate levels of PID, but little incremental content at either extreme.

Panel C examines whether the greater informativeness observed in the Mid PID group depends on the tercile classification. Discrete sorting can assign observations with similar PID values to different groups and mask variation within each group. The continuous quadratic specification instead allows the informativeness of PES to vary smoothly across the PID distribution. The estimates are consistent with an effect of PES that first increases and then decreases with PID. Under full controls, the estimated effect is largest at approximately the 51.5th percentile of the PID distribution, with a 95% confidence interval from the 40.8th to the 62.1st percentile. These estimates provide continuous-specification evidence consistent with the non-monotonic pattern in Panel B.

Overall, Table 5 shows that the aggregate peer earnings signal contains incremental information about subsequent profitability changes and that its informativeness varies non-monotonically with PID and is greatest near the middle of the PID distribution.[20]

4.2 Out-of-Sample Predictive Evidence

Table 5 shows that the informativeness of PES varies across PID within the estimation sample. I next examine whether this pattern improves predictions for observations not used in estimation. In each test period, I compare a benchmark based

[20]The estimated location near the median is an empirical result. The framework predicts a possible interior optimum but does not restrict that optimum to the midpoint of the PID distribution.

on public information with an augmented model that also includes the aggregate peer earnings signal.[21]

Forecasts are generated monthly using expanding-window OLS. Predictors are standardized using only observations in the estimation window, and a strict 12-month gap separates the estimation and test observations to avoid overlap in future profitability outcomes. The estimation window must contain at least 60 months. I evaluate forecast performance using both monthly and quarterly evaluation periods. The monthly evaluation organizes forecast errors by forecast month, preserving the timing with which PES, PID, and the controls become available. Each month therefore represents a distinct forecast origin with an updated information set, even when the profitability outcome is measured at the quarterly frequency. For the quarterly evaluation, the same firm-quarter forecasts are grouped by calendar quarter, matching the reporting frequency of profitability. Because each firm-quarter contributes a single forecast, the two evaluations use the same forecast errors but differ in the time aggregation used for statistical inference.[22]

[Table 6]

Table 6 reports the results. Panel A evaluates the forecast contribution of PES separately for firms in the Low, Mid, and High PID groups. Adding PES produces the strongest and most statistically reliable improvement in forecast accuracy for the Mid PID group. It does not consistently improve forecasts for the Low PID group and produces a smaller improvement for the High PID group. Pairwise tests confirm that the improvement for the Mid PID group exceeds the corresponding improvements for the

[21]The public information controls are the full set of controls used in Table 5, including the focal firm's fundamentals, industry conditions, analyst information, firm and market characteristics, and characteristics of the peer set.

[22]The dependent variable is the future profitability measure used in Table 5.

Low and High PID groups. This pattern is consistent across the monthly and quarterly evaluations.

Panel B examines whether this evidence depends on the use of PID-tercile cutoffs. Instead of assigning firms to discrete groups, the quadratic specification allows the predictive contribution of PES to vary smoothly across the PID distribution.[23] The augmented model improves forecast accuracy under both monthly and quarterly evaluations. The tercile and continuous specifications therefore provide consistent out-of-sample evidence that the aggregate peer earnings signal is most useful for predicting the focal firm's future profitability at intermediate levels of PID.

4.3 Mechanism: Aggregating Weak but Relevant Peer Signals

The preceding analyses show that the predictive content of the aggregate peer earnings signal varies systematically with PID, both in sample and out of sample. This section investigates whether the greater informativeness observed at intermediate levels of PID arises from aggregating information across multiple economically relevant peers. To distinguish this mechanism from alternative explanations, I examine whether the predictive content is concentrated in a single influential peer, whether it can be explained by historical common variation across peers, whether it increases as signals from additional peers are combined, and whether it is stronger when the aggregated peers are more economically relevant to the focal firm.

I first decompose PES into the component attributable to the peer receiving the largest analyst weight and the component attributable to all remaining peers. This decomposition examines whether the predictive content of PES is concentrated in one influential peer or remains present in the broader peer set.

[23] The benchmark includes the public information controls, PID, and PID^2. The augmented model additionally includes PES, $PES \times PID$, and $PES \times PID^2$.

[Table 7]

Table 7 shows that the component attributable to the peer receiving the largest weight has a small and statistically insignificant coefficient after controlling for the component attributable to the remaining peers. In contrast, the remaining-peer component is positive and statistically significant for firms in the Mid-PID group. Wald tests further show that its coefficient is significantly larger in the Mid-PID group than in the Low- and High-PID groups. Thus, the predictive content of PES for Mid-PID firms is not driven solely by the peer receiving the largest weight and remains present in the broader peer set.

The predictive power of the aggregate signal from the remaining peers does not necessarily imply that it combines distinct information from multiple peers. Individual peer signals may reflect the same historical common variation and therefore contain overlapping information. I use a rolling PCA decomposition to separate the variation captured by the first five historical principal components from the residual variation in peer earnings surprises. This decomposition examines whether the predictive relation remains after removing the variation captured by these historical components.[24]

[Table 8]

Table 8 shows that, for Mid-PID firms, the coefficients on the components explained by the first principal component and by the second through fifth principal components are small and statistically insignificant, whereas the coefficient on the residual component is positive and statistically significant. The residual-component coefficients are statistically insignificant for Low- and High-PID firms, although the coefficient on the component

[24]For each month, I estimate the first five PCA loading vectors using peer earnings surprises observed over the preceding 60 months. I then use these loading vectors to decompose the peer earnings surprises entering PES into the variation explained by the five historical principal components and the residual variation. Each constructed component retains the analyst weights used to construct PES. The PCA loadings are estimated using only information available before the month in which the signal is formed.

explained by the second through fifth principal components is weakly significant for High-PID firms.

I next examine whether the predictive content of the residual signal is concentrated in a few peers or increases as additional signals are combined. I construct residual signals from one, two, three, five, and ten randomly selected peers and compare them with the signal from all available peers. The sample is held fixed, the analyst weights are renormalized within each peer set, and the estimates for each peer count are averaged across 100 random selections.

[Figure 3]

Figure 3 traces the predictive coefficient as additional peers are included in the residual signal. For firms in the Mid-PID group, the coefficient generally increases with the number of peers and reaches its highest value when all available peers are included. The corresponding patterns for firms in the Low- and High-PID groups are considerably flatter. This coefficient pattern suggests that the predictive content of the Mid-PID residual signal is not concentrated in a single randomly selected peer and becomes stronger as information from additional peers is combined. The following table provides the corresponding estimates and formal tests of the gain from aggregation.

[Table 9]

Table 9 presents the corresponding estimates and formal tests. For firms in the Mid-PID group, the all-peer coefficient exceeds the mean one-peer coefficient by 0.026, with a bootstrap p-value of 0.007. The corresponding all-peer-minus-one-peer differences are not statistically significant in the Low- or High-PID groups.

Because the peer weights are renormalized within each selected peer set, coefficient comparisons across peer counts may partly reflect changes in signal scale. I therefore standardize each peer-count signal within month and compare its incremental explanatory

power using the common sample. In the Mid-PID group, the partial R^2 increases from 0.096 basis points for one randomly selected peer to 0.426 basis points for ten peers and 1.069 basis points for the complete peer set. Although these values are small in absolute terms, their cross-peer-count pattern shows that the increase is also present in a scale-invariant measure of incremental explanatory power. No comparable increase appears in the Low- or High-PID groups.

The preceding tests show that the predictive content of the Mid-PID residual signal increases as information from additional peers is combined. This pattern establishes a gain from broader aggregation, but it does not show whether that gain depends on the economic relevance of the information being combined. The improvement could otherwise reflect the mechanical averaging of a larger number of signals, including signals with limited relevance to the focal firm. I therefore examine whether the predictive relation between PES and future profitability strengthens with peer relevance within the Mid-PID group.

[Table 10]

Table 10 shows that the PES-by-relevance interaction is positive and statistically significant for Mid-PID firms. The implied PES coefficient within this group increases from 0.035 at the 25th percentile of peer relevance to 0.042 at the median and 0.046 at the 75th percentile. Thus, within the Mid-PID group, the aggregate peer signal becomes more informative as the economic relevance of the underlying peers increases. This result complements the peer-count analysis by showing that the aggregation mechanism depends not only on combining a broader set of peer signals but also on the economic relevance of the information being combined.

Taken together, the tests provide a sequence of results supporting the proposed aggregation mechanism: for Mid-PID firms, PES predicts future profitability by

combining residual earnings information distributed across multiple economically relevant peers. The dominant-peer decomposition shows that this predictive content is not driven solely by the peer receiving the largest weight, while the PCA decomposition shows that it remains after removing the variation captured by the first five historical principal components. The random-subset analysis then shows that the residual signal becomes more informative as additional peers are aggregated, and the relevance analysis shows that the predictive relation is stronger when the underlying peers are more economically relevant to the focal firm.

4.4 Market Responses to Aggregate Peer Information

To assess the market relevance of aggregate peer information, I examine price responses and analyst disagreement. Earnings announcements can convey price-relevant information (Ball and Brown, 1968; Beaver, 1968), while delayed incorporation may generate subsequent return continuation (Kothari, 2001). If peer earnings news is informative, it should be associated with a price response when it becomes public; little subsequent drift would be consistent with timely incorporation. Analyst forecast dispersion has been linked to uncertainty, consensus, and differences of opinion (Barron et al., 1998; Diether et al., 2002). Informative news that provides analysts with a common basis for updating may therefore reduce disagreement, whereas redundant or diffuse news may leave greater scope for different interpretations.

[Table 11]

Table 11 examines whether the aggregate peer signal is incorporated into prices when the underlying peer earnings signals become available. Because a target firm-month may contain several linked peer earnings announcements, I measure the contemporaneous response as the link-weighted average of the target firm's announcement-window

abnormal returns, with weights based on shared-analyst coverage. This construction mirrors the monthly aggregation of PES while preserving the link between the aggregate signal and the announcements from which it is formed. Columns (3) and (4) test for delayed incorporation using the target firm's abnormal return in month $m+1$. Column (3) uses firm and month fixed effects, while column (4) uses monthly Fama–MacBeth estimates (Fama and MacBeth, 1973) with Newey–West standard errors (Newey and West, 1987).

The aggregate peer signal is reflected primarily in announcement-window returns. In the full sample, a one-standard-deviation increase in PES is associated with abnormal returns of 5.7 basis points over $CAR[0,5]$ and 7.2 basis points over $CAR[0,10]$. The next-month fixed-effects estimate is small and statistically insignificant.

The price response varies systematically with PID. At intermediate levels of PID, a one-standard-deviation increase in PES is associated with abnormal returns of 8.4 basis points over $CAR[0,5]$ and 9.1 basis points over $CAR[0,10]$. At low levels of PID, announcement-window responses are smaller, with weak evidence of next-month drift in the Fama–MacBeth specification. At high levels of PID, contemporaneous responses are weakest and no reliable drift follows. Thus, the aggregate peer signal has the largest immediate valuation effect when peer information is sufficiently broad to contain distinct information but remains economically relevant to the focal firm.

[Table 12]

Price responses reveal whether the aggregate signal affects valuations. Analyst disagreement provides complementary evidence on whether it leads to a more common assessment of the focal firm's prospects. Table 12 examines changes in analyst

disagreement after the signal becomes available.[25] The dependent variable is the change in analyst disagreement from month m to month $m + k$.

The estimates are small over the first two months. Beginning three months after signal formation, a nonlinear PID pattern emerges: larger absolute peer signals are followed by greater disagreement near the tails of the PID distribution but lower disagreement near its middle. At the six-month horizon, the coefficients on the absolute peer signal and its linear and quadratic interactions with PID are all statistically significant. The delayed decline in disagreement is consistent with the signal gradually providing analysts with a common basis for assessing the focal firm's fundamentals.

Taken together, the two outcomes show that the predictive content of aggregate peer information has market consequences. At intermediate levels of PID, the signal generates the strongest contemporaneous price response and the largest subsequent decline in analyst disagreement, with little evidence of next-month return drift. Aggregate peer information is therefore reflected most clearly in valuations and beliefs when the peer set combines distinct but still economically relevant information.

5 Alternative Explanations and Robustness

I conduct additional tests of the non-monotonic relation between PID and the informativeness of PES and of the aggregation mechanism at moderate PID. The detailed results are reported in the Internet Appendix.

The main pattern is preserved under alternative variable constructions. Internet Appendix Table IA.1 uses log-transformed and equal-weighted versions of PES, while Table IA.2 excludes forecasts issued by analysts who also cover the focal firm. Table IA.3

[25]The analysis uses the shared-analyst-weighted average of the absolute earnings surprises of analyst-linked peers. This measure captures the magnitude of peer earnings news without allowing positive and negative surprises to offset one another. Both positive and negative surprises may provide information that helps analysts update their assessments.

reconstructs PID using six-digit GICS industries. The predictive content of PES continues to peak near the middle of the PID distribution, indicating that the non-monotonic relation is not specific to the baseline signal, shared-analyst forecasts, or the SIC industry classification.

Firm and peer-network characteristics do not account for the greater informativeness of peer earnings information at moderate PID. Internet Appendix Table IA.5 allows the PES slope to vary with network complexity, analyst coverage, firm size, and business concentration, and the predictive relation at moderate PID remains positive. Table IA.7 further examines whether the residual signal reflects contemporaneous variation shared by firms in the same source industry or information assigned through similarly structured artificial networks. The signal remains informative at moderate PID after removing its SIC2-by-month component, whereas randomly reassigned signals do not reproduce the actual-network evidence. These results indicate that the predictive content is tied to information conveyed through the analyst-linked peer network rather than broad source-industry variation or network structure alone.

The residual aggregation evidence is also robust to how historical common variation is estimated and how the residual peer signals are aggregated. Internet Appendix Table IA.4 varies the PCA estimation window, the aggregation weights, and the treatment of the strongest peer link. Across these specifications, the residual peer signal remains informative at moderate PID.

Internet Appendix Table IA.8 clarifies the negative Low-PID coefficient in the peer-relevance analysis. Reestimating the baseline specification on the restricted relevance sample produces a similar negative coefficient before the relevance terms are introduced. The estimate therefore reflects the composition of the restricted sample rather than the relevance interaction and does not alter the evidence that economically relevant peer signals are more informative at moderate PID.

Finally, the predictive relation is not confined to a particular period or forecast horizon. Internet Appendix Table IA.6 shows that peer earnings information remains predictive at moderate PID after excluding the financial-crisis and pandemic periods and within the pre-2020 sample. Table IA.9 shows that the relation remains positive when future profitability is examined separately over each of the next four quarters.

Overall, the non-monotonic relation and the aggregation evidence at moderate PID are robust to alternative variable constructions, firm and network characteristics, historical and contemporaneous common variation, the composition and weighting of the peer set, sample periods, and forecast horizons.

6 Conclusion

The evidence indicates that the value of aggregating peer information depends on the composition of the underlying information set. Greater dispersion can broaden the information represented in the aggregate, but it can also introduce signals that are less relevant to the focal firm. Aggregate peer signals are therefore more informative when the peer set spans multiple, yet still economically relevant, sources. In this region, individually weak peer signals acquire predictive content through aggregation rather than through a dominant peer or historical common variation. The corresponding responses in prices and analyst beliefs further indicate that this information is relevant to market participants.

The paper contributes to the peer-effects and information-diffusion literature by identifying the information span of the peer set as a determinant of when aggregation produces an informative signal. More broadly, the results show that the value of peer information is determined not simply by the availability or number of peer signals, but by how those signals are distributed across economic sources.

References

Ali, U. and D. Hirshleifer (2020). "Shared Analyst Coverage: Unifying Momentum Spillover Effects". *Journal of Financial Economics* 136.3, pp. 649–675.

Aobdia, D., J. Caskey, and N. B. Ozel (2014). "Inter-Industry Network Structure and the Cross-Predictability of Earnings and Stock Returns". *Review of Accounting Studies* 19.3, pp. 1191–1224.

Avramov, D., S. Ge, S. Li, and O. Linton (2026). "Dual Peer Effects and Cross-Stock Predictability". *Journal of Financial Economics* 180, 104274.

Ball, R. and P. Brown (1968). "An Empirical Evaluation of Accounting Income Numbers". *Journal of Accounting Research* 6.2, pp. 159–178.

Barron, O. E., O. Kim, S. C. Lim, and D. E. Stevens (1998). "Using Analysts' Forecasts to Measure Properties of Analysts' Information Environment". *The Accounting Review* 73.4, pp. 421–433.

Bartov, E., G. Clinch, and W. Li (2026). "Unpacking Earnings Information Spillovers: Shared Analyst Coverage, Peer Heterogeneity, and Investor Sophistication". Working paper, revised June 28, 2026.

Beaver, W. H. (1968). "The Information Content of Annual Earnings Announcements". *Journal of Accounting Research* 6: *Empirical Research in Accounting: Selected Studies 1968*, pp. 67–92.

Bhojraj, S., C. M. C. Lee, and D. K. Oler (2003). "What's My Line? A Comparison of Industry Classification Schemes for Capital Market Research". *Journal of Accounting Research* 41.5, pp. 745–774.

Carhart, M. M. (1997). "On Persistence in Mutual Fund Performance". *The Journal of Finance* 52.1, pp. 57–82.

Clark, T. E. and K. D. West (2007). "Approximately Normal Tests for Equal Predictive Accuracy in Nested Models". *Journal of Econometrics* 138.1, pp. 291–311.

Cohen, L. and A. Frazzini (2008). "Economic Links and Predictable Returns". *The Journal of Finance* 63.4, pp. 1977–2011.

Cohen, L. and D. Lou (2012). "Complicated Firms". *Journal of Financial Economics* 104.2, pp. 383–400.

Diamond, D. W. and R. E. Verrecchia (1981). "Information Aggregation in a Noisy Rational Expectations Economy". *Journal of Financial Economics* 9.3, pp. 221–235.

Diether, K. B., C. J. Malloy, and A. Scherbina (2002). "Differences of Opinion and the Cross Section of Stock Returns". *The Journal of Finance* 57.5, pp. 2113–2141.

Fama, E. F. and K. R. French (1993). "Common Risk Factors in the Returns on Stocks and Bonds". *Journal of Financial Economics* 33.1, pp. 3–56.

— (2015). "A Five-Factor Asset Pricing Model". *Journal of Financial Economics* 116.1, pp. 1–22.

Fama, E. F. and J. D. MacBeth (1973). "Risk, Return, and Equilibrium: Empirical Tests". *Journal of Political Economy* 81.3, pp. 607–636.

Foster, G. (1981). "Intra-industry Information Transfers Associated with Earnings Releases". *Journal of Accounting and Economics* 3.3, pp. 201–232.

Foucault, T. and L. Frésard (2014). "Learning from Peers' Stock Prices and Corporate Investment". *Journal of Financial Economics* 111.3, pp. 554–577.

Grossman, S. J. (1976). "On the Efficiency of Competitive Stock Markets Where Trades Have Diverse Information". *The Journal of Finance* 31.2, pp. 573–585.

Grossman, S. J. and J. E. Stiglitz (1980). "On the Impossibility of Informationally Efficient Markets". *American Economic Review* 70.3, pp. 393–408.

Guimerà, R. and L. A. N. Amaral (2005). "Functional Cartography of Complex Metabolic Networks". *Nature* 433.7028, pp. 895–900.

Hameed, A., R. Morck, J. Shen, and B. Yeung (2015). "Information, Analysts, and Stock Return Comovement". *The Review of Financial Studies* 28.11, pp. 3153–3187.

He, Z., J. Huang, and C. Parlatore (2024). *Information Span in Credit Market Competition*. NBER Working Paper 33141. Revised September 2025. National Bureau of Economic Research.

Hellwig, M. F. (1980). "On the Aggregation of Information in Competitive Markets". *Journal of Economic Theory* 22.3, pp. 477–498.

Herfindahl, O. C. (1950). "Concentration in the U.S. Steel Industry". Ph.D. dissertation. New York, NY: Columbia University.

Hirschman, A. O. (1945). *National Power and the Structure of Foreign Trade*. Berkeley, CA: University of California Press.

Hong, H., T. Lim, and J. C. Stein (2000). "Bad News Travels Slowly: Size, Analyst Coverage, and the Profitability of Momentum Strategies". *The Journal of Finance* 55.1, pp. 265–295.

Hong, H. and J. C. Stein (1999). "A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets". *The Journal of Finance* 54.6, pp. 2143–2184.

Honkanen, P. and D. Schmidt (2022). "Learning from Noise? Price and Liquidity Spillovers around Mutual Fund Fire Sales". *The Review of Asset Pricing Studies* 12.2, pp. 593–637.

Hou, K. (2007). "Industry Information Diffusion and the Lead–Lag Effect in Stock Returns". *The Review of Financial Studies* 20.4, pp. 1113–1138.

Kaustia, M. and V. Rantala (2021). "Common Analysts: Method for Defining Peer Firms". *Journal of Financial and Quantitative Analysis* 56.5, pp. 1505–1536.

Kim, O. and R. E. Verrecchia (1991). "Trading Volume and Price Reactions to Public Announcements". *Journal of Accounting Research* 29.2, pp. 302–321.

Kothari, S. P. (2001). "Capital Markets Research in Accounting". *Journal of Accounting and Economics* 31.1–3, pp. 105–231.

Lou, Y., S. Parsa, D. Ray, D. Li, and S. Wang (2019). "Information Aggregation in a Financial Market with General Signal Structure". *Journal of Economic Theory* 183, pp. 594–624.

Menzly, L. and O. Ozbas (2010). "Market Segmentation and Cross-Predictability of Returns". *The Journal of Finance* 65.4, pp. 1555–1580.

Müller, K. and S. N. M. Schmickler (2025). "Interacting Anomalies". *The Review of Asset Pricing Studies* 15.2, pp. 162–216.

Newey, W. K. and K. D. West (1987). "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix". *Econometrica* 55.3, pp. 703–708.

Pandit, S., C. E. Wasley, and T. Zach (2011). "Information Externalities along the Supply Chain: The Economic Determinants of Suppliers' Stock Price Reaction to Their Customers' Earnings Announcements". *Contemporary Accounting Research* 28.4, pp. 1304–1343.

Peng, L. and W. Xiong (2006). "Investor Attention, Overconfidence and Category Learning". *Journal of Financial Economics* 80.3, pp. 563–602.

Piotroski, J. D. and D. T. Roulstone (2004). "The Influence of Analysts, Institutional Investors, and Insiders on the Incorporation of Market, Industry, and Firm-Specific Information into Stock Prices". *The Accounting Review* 79.4, pp. 1119–1151.

Ramnath, S. (2002). "Investor and Analyst Reactions to Earnings Announcements of Related Firms: An Empirical Analysis". *Journal of Accounting Research* 40.5, pp. 1351–1376.

Siga, L. and M. Mihm (2021). "Information Aggregation in Competitive Markets". *Theoretical Economics* 16.1, pp. 161–196.

Stambaugh, R. F. and Y. Yuan (2017). "Mispricing Factors". *The Review of Financial Studies* 30.4, pp. 1270–1315.

Thomas, J. and F. Zhang (2008). "Overreaction to Intra-industry Information Transfers?" *Journal of Accounting Research* 46.4, pp. 909–940.

Vives, X. (2008). *Information and Learning in Markets: The Impact of Market Microstructure*. Princeton, NJ: Princeton University Press.

Table 1.

Summary statistics: analyst-linked peers

Variable	Mean	SD	p25	Median	p75
# Peers per firm-month	72	71.04	37	60	88
# Shared analysts per firm-pair	2	2.59	1	1	2
Market cap: all firms	6,654.06	41,080.41	222.53	767.71	2,920.99
Market cap: firms with peers	7,796.41	44,549.09	350.88	1,076.68	3,699.40
Share of firms with peers	0.85	0.05	0.82	0.87	0.89
Share of market cap with peers	0.99	0.01	0.99	0.99	0.99
# firms per month	3,352	526.78	2,968	3,126	3,821

Notes: Market capitalization is reported in millions of dollars. The number of peers per firm-month is summarized for observations with at least one peer. The number of shared analysts is summarized across linked firm pairs in each month. Market capitalization is summarized separately for all eligible observations and for observations with at least one peer. The share of firms with peers, the share of market capitalization represented by firms with peers, and the number of firms are calculated each month and summarized over the 300-month sample period.

Return to text.

Table 2.

Peer information dispersion and firm characteristics

Variable	Low PID	Mid PID	High PID
A. Distribution of analyst-linked peers across industries			
Mean PID	0.13	0.59	0.83
Peer industry share (SIC2)	0.89	0.47	0.20
Peer sector share (SIC1)	0.93	0.62	0.45
Number of peer industries (SIC2)	4.44	11.68	19.27
Number of peer sectors (SIC1)	2.82	5.32	6.58
Number of analyst-linked peers	69.35	73.57	94.02
B. Characteristics of focal firms			
Size	6.88	7.34	7.14
Book-to-market	-0.82	-0.95	-0.94
Analyst coverage	7.36	9.18	7.98
Analyst disagreement	0.10	0.11	0.09
ROA	-1.06%	0.52%	1.08%
Business concentration	0.33	0.30	0.29

C. Association between PID and firm characteristics

Specification	Adj. R^2	Unexplained share
Firm characteristics only[26]	0.06	0.94
Firm characteristics, month FE, and SIC2 FE	0.54	0.46

D. Positive industry earnings correlations

Variable	Low PID	Mid PID	High PID
Correlation between peer industries and the focal industry	0.22	0.20	0.18
Observations	184,695	203,552	195,607
Correlation among peer industries	0.21	0.20	0.17
Observations	197,483	254,005	255,352

Notes: Firm-month observations are assigned to monthly PID terciles. Panels A and B report group means. Peer industry share is the share of shared-analyst exposure originating from the focal firm's two-digit SIC industry; peer sector share is defined analogously using one-digit SIC sectors. Panel C reports adjusted R^2 from regressions of monthly PID rank; the unexplained share equals one minus adjusted R^2. Panel D reports the positive industry earnings correlations defined in the text. Its observation counts are the firm-month observations available for each measure. Continuous firm characteristics are winsorized by month at the 1st and 99th percentiles.

Return to text.

[26]"Firm characteristics only" refers to a single joint regression of monthly PID rank on size, book-to-market, the logarithm of one plus analyst coverage, analyst disagreement, ROA, and business concentration. The second specification additionally includes month and SIC2 fixed effects.

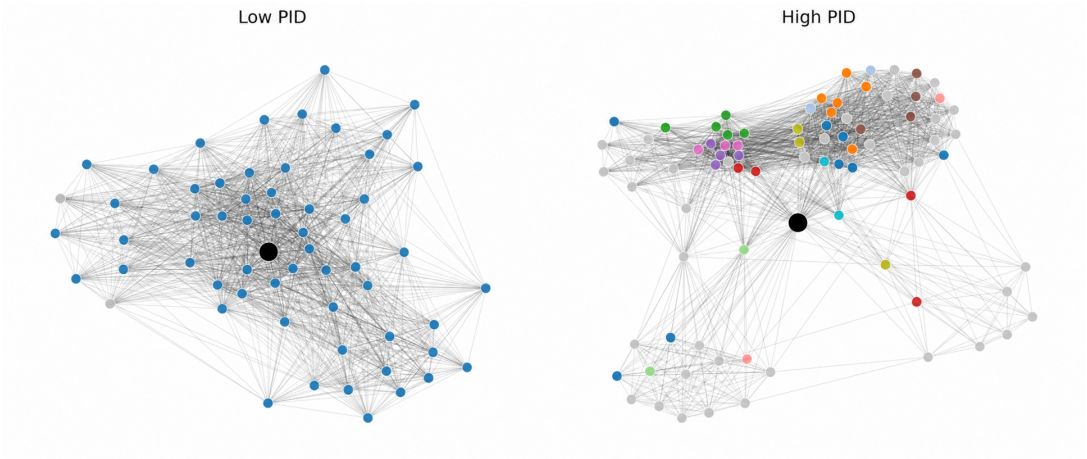


Figure 1: Peer Information Dispersion: Real Data Illustration

Alt text. Two peer networks from December 2024 illustrate low and high PID. Black nodes identify the focal firms, other colors identify peers' two-digit SIC industries, and gray lines represent shared-analyst links. Peer industries are concentrated in the low-PID network, corresponding to a narrower information span, and more dispersed in the high-PID network, corresponding to a broader information span.

[Return to text]

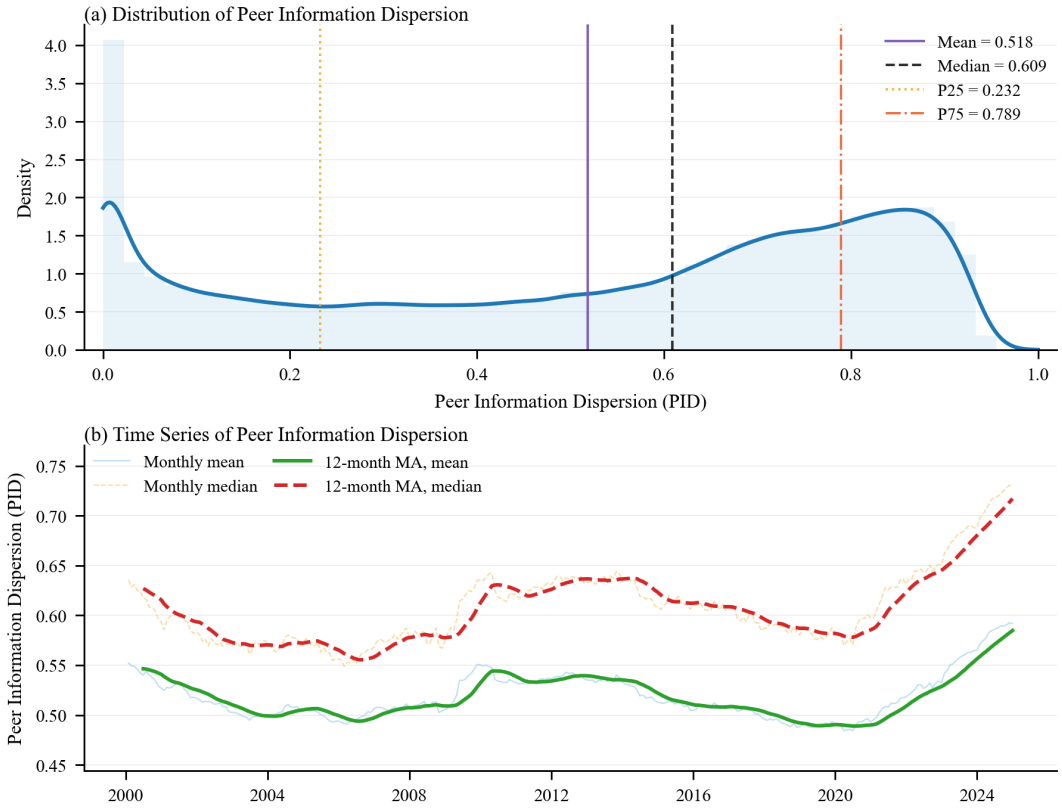


Figure 2: Peer Information Dispersion: Distribution and Time-Series Patterns

Alt text. Panel (a) shows the distribution of PID over 2000–2024, with vertical lines marking the mean, median, and 25th and 75th percentiles. The distribution exhibits substantial cross-sectional variation from narrower to broader peer information spans. Panel (b) shows the monthly mean and median PID, together with their 12-month moving averages. The median generally remains above the mean, and both measures vary over time and rise markedly after 2020, indicating an increase in the information span of peer sets in recent years.

[Return to text]

Table 3.

Summary Statistics: Main Firm-Month Variables

Variable	N	Mean	SD	p25	Median	p75
Return	1,005,307	1.56	13.26	-5.06	0.81	7.09
Excess return	1,005,307	1.40	13.27	-5.22	0.66	6.95
Abnormal return	988,040	0.68	11.12	-5.08	0.06	5.46
Analyst disagreement	748,804	0.10	0.19	0.02	0.04	0.09
Profitability	998,570	0.00	0.04	0.00	0.01	0.02
Average future profitability change	175,600	-0.19	2.63	-0.67	-0.03	0.45
Absolute future profitability change	175,600	1.35	2.26	0.18	0.55	1.46
PID	851,794	0.52	0.31	0.23	0.61	0.79
Peer earnings signal (PES)	761,934	1.05	2.09	0.20	1.05	1.96
Focal earnings surprise (SUE)	784,553	1.10	4.38	-0.54	0.76	2.51
Size	1,005,582	6.75	1.88	5.41	6.64	7.98
Analyst coverage	838,419	8.12	6.85	3.00	6.00	11.00
Return volatility	901,778	0.11	0.07	0.07	0.10	0.14
Turnover	1,005,582	0.18	0.20	0.06	0.12	0.22
Book-to-market	960,123	-0.78	0.87	-1.25	-0.65	-0.18
Business concentration	879,597	0.31	0.16	0.20	0.33	0.33

Notes: This table reports summary statistics for the main firm-month variables in the 2000–2024 sample. Variables are reported after monthly winsorization at the 1st and 99th percentiles where applicable. Average future profitability change is the average change in quarterly profitability from quarter q to quarters $q+1$ through $q+4$, winsorized by availability month within the final regression sample and reported in percentage points. Absolute future profitability change is its absolute value. Variable construction details are provided in Appendix A.

Return to text.

Table 4.

Pooled correlations of peer signal, PID, disagreement, fundamentals, and returns

	PES	PID	Disagr.	Profit.	Return	Abret.	SUE	Coverage	Size	BM
PID	0.091	1.000								
	(0.00)									
Disagr.	−0.036	−0.038	1.000							
	(0.00)	(0.00)								
Profit.	0.039	0.255	−0.175	1.000						
	(0.00)	(0.00)	(0.00)							
Return	−0.001	0.001	0.026	−0.032	1.000					
	(0.28)	(0.67)	(0.00)	(0.00)						
Abret.	−0.003	−0.015	0.017	−0.036	0.800	1.000				
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)					
SUE	0.128	0.035	−0.107	0.160	−0.004	−0.005	1.000			
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)				
Coverage	0.073	0.024	−0.101	0.124	−0.018	−0.022	0.060	1.000		
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)			
Size	0.111	0.061	−0.164	0.222	0.020	0.013	0.107	0.707	1.000	
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)		
BM	−0.146	−0.032	0.082	0.024	0.044	0.012	−0.123	−0.192	−0.269	1.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	

Return to text.

Table 5.
Non-monotonic Informativeness of Aggregate Peer Earnings Signal

	DV: future profitability change $\overline{\Delta Profitability}_{i,\{q+1,q+2,q+3,q+4\}}$								
	Baseline PES			PID-tercile			Quadratic PID		
	Fund.	+ Info env.	+ Firm/mkt	Fund.	+ Info env.	+ Firm/mkt	Fund.	+ Info env.	+ Firm/mkt
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. Baseline PES									
PES	0.040***	0.036***	0.024**						
	(3.79)	(3.49)	(2.54)						
Panel B. PID-tercile									
PES× Low PID				0.018	0.016	0.008			
				(1.11)	(1.00)	(0.56)			
PES× Mid PID				0.086***	0.081***	0.060***			
				(4.25)	(4.03)	(3.24)			
PES× High PID				0.034**	0.028*	0.014			
				(2.26)	(1.86)	(0.98)			
Wald p-value: Mid = Low				0.010	0.015	0.045			
Wald p-value: Mid = High				0.011	0.011	0.021			
Panel C. Quadratic PID									
PES							-0.031	-0.030	-0.020
							(-1.43)	(-1.38)	(-0.92)
PES × PID							0.438***	0.414***	0.276**
							(3.32)	(3.15)	(2.20)
PES × PID^2							-0.416***	-0.398***	-0.269**
							(-3.18)	(-3.06)	(-2.17)
Implied effect at PID p10							0.009	0.008	0.005
							(0.64)	(0.56)	(0.39)
Implied effect at PID p50							0.085***	0.078***	0.051***
							(4.20)	(3.94)	(2.87)
Implied effect at PID p90							0.027	0.021	0.012
							(1.63)	(1.28)	(0.75)
PID percentile at which the PES effect is largest [95% CI]							0.527 [0.452, 0.601]	0.521 [0.444, 0.597]	0.515 [0.408, 0.621]
Firm FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Time FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	175,600	175,600	175,600	175,600	175,600	175,600	175,600	175,600	175,600
Within adj. R^2	0.356	0.358	0.395	0.356	0.358	0.395	0.356	0.359	0.395

Notes: The dependent variable, $Y_{i,q}$, is the average of the changes in quarterly profitability from quarter q to each of quarters $q+1$ through $q+4$, winsorized by availability month and reported in percentage points. Here, $t(q)$ denotes the Compustat availability month associated with fiscal quarter q. Panel A reports the relation between the peer earnings signal (PES) and future profitability changes. Panel B allows the coefficient on PES to vary across PID terciles.

$$Y_{i,q} = \sum_{g\in\{\text{Low,Mid,High}\}} \beta_g \text{PES}_{i,t(q)} \times \mathbf{1}\Big\{\text{PID}_{i,t(q)} \in g\Big\} + \Gamma' \text{Controls}_{i,t(q)} + \alpha_i + \delta_{t(q)} + \varepsilon_{i,q}.$$

Panel C uses a quadratic PID specification.

$$\begin{aligned} Y_{i,q} = {} & \beta_0 \text{PES}_{i,t(q)} + \beta_1 \text{PES}_{i,t(q)} \times \text{PID}_{i,t(q)} + \beta_2 \text{PES}_{i,t(q)} \times \text{PID}^2_{i,t(q)} \\ & + \theta_1 \text{PID}_{i,t(q)} + \theta_2 \text{PID}^2_{i,t(q)} + \Gamma' \text{Controls}_{i,t(q)} + \alpha_i + \delta_{t(q)} + \varepsilon_{i,q}. \end{aligned}$$

PES is constructed from analyst-linked peer firms' earnings surprises using shared-analyst weights and is standardized within month. PID is the monthly percentile rank of mean PID over the preceding 12 months, requiring at least nine nonmissing monthly observations. The PID terciles are formed using the same measure. All interaction specifications include the corresponding main effects. The final row of Panel C reports the PID percentile rank, scaled from zero to one, at which the estimated effect of PES is largest, together with its 95% confidence interval. Fundamental controls include current and lagged focal-firm profitability and industry earnings signals. Information environment controls include analyst consensus, the focal firm's earnings surprise, its absolute value, analyst coverage, and analyst disagreement. Firm and market controls include past return, size, book-to-market, return volatility, turnover, business concentration, the numbers of analyst-linked peers and peer industries, and total shared-analyst link weight. All regressions include firm and month fixed effects. Standard errors are two-way clustered by firm and month; t-statistics are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Return to text.

Table 6.

Out-of-Sample Forecast Gains from the Aggregate Peer Earnings Signal

Forecast comparison	Monthly evaluation		Quarterly evaluation	
	OOS R^2 (%)	CW-adjusted $\Delta MSFE$	OOS R^2 (%)	CW-adjusted $\Delta MSFE$
Panel A. PID-tercile PES				
$PES\times$ Low PID	-0.267	0.009	-0.267	-0.002
$PES\times$ Mid PID	0.194	0.042***	0.194	0.018***
$PES\times$ High PID	0.047	0.010*	0.047	0.006*
p-value: Mid > Low		0.012		0.001
p-value: Mid > High		< 0.001		< 0.001
Panel B. Quadratic PID				
PES, $PES \times PID$, $PES \times PID^2$	0.067	0.022***	0.067	0.010***
Observations	129,951		129,951	
Test periods	205 months		69 quarters	

Notes: This table reports out-of-sample tests for the future-profitability outcome in Table 5. The benchmark uses public information, and the augmented model adds PES. Forecasts are generated monthly using expanding-window OLS with a 12-month gap. Following Clark and West (2007), the adjusted loss differential is $d_{i,t}^{CW} = (e_{i,t}^{B})^2 - (e_{i,t}^{A})^2 + (\widehat{Y}_{i,t}^{B} - \widehat{Y}_{i,t}^{A})^2$. Differentials are averaged by forecast month or calendar quarter and evaluated using one-sided Newey–West tests with 12 or 4 lags, respectively. Equal weighting of periods allows the monthly and quarterly CW means to differ. Both evaluations use the same forecasts, so their $R^2_{OOS} = 1 - MSFE_{aug}/MSFE_{bench}$ values are identical. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Return to text.

Table 7.

Aggregate Peer Signals Versus the Dominant Peer

	$\overline{\Delta Profitability}_{i,\{q+1,q+2,q+3,q+4\}}$		
	Low PID	Mid PID	High PID
Aggregate Peer Earnings Signal (PES)	0.009	0.063***	0.014
	(0.46)	(3.39)	(0.97)
Component of PES attributable to the dominant peer	0.004	0.014	-0.004
	(0.38)	(1.19)	(-0.43)
Component of PES attributable to the remaining peers	0.004	0.060***	0.014
	(0.23)	(3.47)	(1.03)
Wald p-value: Remaining-peer component, Mid = Low		0.033	
Wald p-value: Remaining-peer component, Mid = High		0.012	
Public controls	Y	Y	Y
Firm FE	Y	Y	Y
Time FE	Y	Y	Y
Observations		171,691	
Firms		6,393	
Months		277	

Notes: This table decomposes PES into the component attributable to the peer receiving the largest shared-analyst weight and the component attributable to all remaining peers. The first row estimates the total PES coefficient using the same decomposition sample. The two components enter the decomposition regressions jointly and use the weights employed in constructing PES. All specifications include the public controls used in Table 5, PID-group indicators, firm fixed effects, and month fixed effects. Standard errors are two-way clustered by firm and month. t-statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Return to text.

Table 8.

PCA Decomposition of Peer Earnings Surprises

	$\overline{\Delta Profitability}_{i,\{q+1,q+2,q+3,q+4\}}$		
	Low PID	Mid PID	High PID
PES^{PC1}	0.024	0.005	0.006
	(1.31)	(0.38)	(0.50)
PES^{PC2-5}	0.005	0.000	0.021*
	(0.33)	(0.03)	(1.90)
$PES^{Residual}$	0.003	0.032***	0.000
	(0.20)	(2.60)	(0.00)
Controls	Y	Y	Y
Firm FE	Y	Y	Y
Time FE	Y	Y	Y
Observations		139,332	

Notes: PCA loadings are estimated each month using peer earnings surprises observed over the preceding 60 months. PES^{PC1} is the component explained by the first principal component, PES^{PC2-5} is the component explained by the second through fifth principal components, and $PES^{Residual}$ is the remaining component. All three components enter the regression jointly. All specifications include the public controls used in Table 5, firm fixed effects, and month fixed effects. Standard errors are two-way clustered by firm and month. t-statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Return to text.

Table 9.

Aggregation Gains across Random Peer Subsets

	Low PID	Mid PID	High PID
Panel A. Average Coefficient by Number of Peers			
1 peer	-0.001	0.005	-0.002
2 peers	0.002	0.008	0.000
3 peers	0.001	0.011	0.001
5 peers	0.002	0.015	0.001
10 peers	0.001	0.020	0.001
All peers	0.000	0.031	0.003
Aggregation gain: All peers minus 1 peer	0.001	0.026	0.005
Bootstrap p-value	0.953	0.007	0.556
Panel B. Standardized Partial R^2 (Basis Points)			
1 peer	0.120	0.096	0.083
10 peers	0.081	0.426	0.041
All peers	0.000	1.069	0.009
Observations		114,078	
Random draws		100	

Notes: The sample is held fixed across peer counts and requires at least 10 peers with valid earnings surprises. For each finite peer count, peers are randomly selected and their shared-analyst weights are renormalized to sum to one. Panel A reports coefficients averaged across 100 random draws; "All peers" uses the complete eligible peer set. Aggregation gain is the all-peer coefficient minus the mean one-peer coefficient, with a p-value based on 499 month-clustered bootstrap replications. Panel B reports partial R^2 values in basis points after each signal is standardized within month, also averaged across 100 random draws. All specifications include the public controls used in Table 5, firm fixed effects, and month fixed effects.

Return to text.

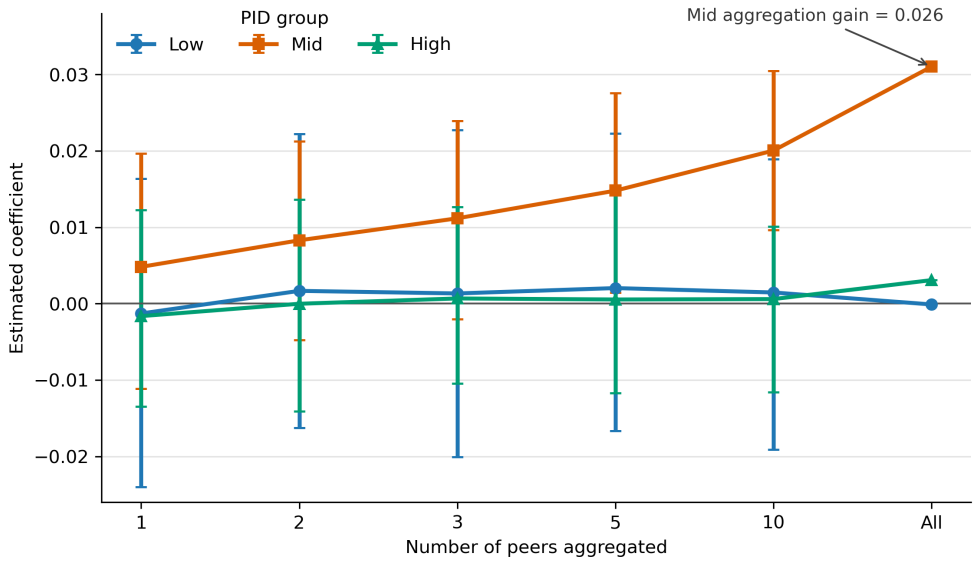


Figure 3

Predictive Gains from Residual Peer Aggregation

Notes: This figure plots the estimated coefficient on the aggregate residual peer earnings signal after removing the historical common components identified by the rolling PCA decomposition. The horizontal axis shows the number of peers included in the signal; "All" uses the full available peer set. For each finite peer count, the plotted coefficients are averages across 100 random selections, and the whiskers indicate the 5th and 95th percentiles of the randomization distribution. Analyst weights are rescaled to sum to one within each selection. All specifications include the public controls, firm fixed effects, and month fixed effects used in Table 5. The sample is held fixed across all peer counts.

Alt text: Line chart showing that the coefficient on the residual peer earnings signal rises with the number of peers for firms in the Mid-PID group and is highest when all peers are included. The coefficients remain comparatively flat for firms in the Low- and High-PID groups.

Return to text.

Table 10.

Peer Relevance and PES Informativeness

	Low PID	Mid PID	High PID
Panel A. Effect of Peer Relevance on PES Informativeness			
$PES\times$ Peer relevance	-0.016	0.028**	0.015
	(-0.80)	(2.31)	(1.46)
Panel B. Implied PES Coefficients			
Peer relevance at the 25th percentile	-0.038**	0.035***	0.002
	(-2.11)	(2.74)	(0.10)
Peer relevance at the 50th percentile	-0.042**	0.042***	0.005
	(-2.29)	(3.37)	(0.34)
Peer relevance at the 75th percentile	-0.045**	0.046***	0.008
	(-2.31)	(3.66)	(0.49)
Public controls	Y	Y	Y
Firm FE	Y	Y	Y
Month FE	Y	Y	Y
Observations		167,133	

Notes: This table examines how peer relevance relates to the informativeness of PES across PID groups. Panel A reports the interaction between PES and peer relevance within each PID group. Panel B reports the implied PES coefficients at the 25th, 50th, and 75th percentiles of peer relevance. The specification includes the public controls used in Table 5, peer-relevance level controls, the $PES\times$ peer-relevance interactions, and the peer-signal-redundancy terms used in the final specification. Internet Appendix Table IA.8 reports the stepwise transition from the baseline specification to this final specification. All regressions include firm and month fixed effects. Standard errors are two-way clustered by firm and month. t-statistics are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Return to text.

Table 11.
Peer Earnings Aggregation, Price Response, and Next-Month Drift

	Contemporaneous price response		Next-month abnormal return	
	$CAR[0,5]$	$CAR[0,10]$	FE	Fama–MacBeth
	(1)	(2)	(3)	(4)
Unconditional				
PES	0.057***	0.072***	0.032	0.068*
	(3.76)	(3.42)	(0.97)	(1.97)
PID terciles				
PES× Low PID	0.051*	0.074*	0.067	0.108*
	(1.82)	(1.94)	(1.25)	(1.94)
PES× Mid PID	0.084***	0.091***	0.040	0.074
	(4.13)	(3.12)	(0.87)	(1.45)
PES× High PID	0.037*	0.050*	-0.019	-0.043
	(1.86)	(1.84)	(-0.44)	(-0.81)
Controls	Y	Y	Y	Y
Firm FE	Y	Y	Y	
Event-month FE	Y	Y	Y	
Fama–MacBeth				Y
Observations	506,871	506,871	506,255	
Avg. monthly obs.				1,776
Fama–MacBeth months				285
Adj./Avg. R^2	0.010	0.011	0.006	0.038

Notes: Columns (1)–(2) report the link-weighted average of the target firm's announcement-window abnormal returns around linked peer earnings announcements in event month m. The link weights are based on shared-analyst coverage and are normalized within each target-firm-month. Columns (3)–(4) report the target firm's abnormal return in month $m+1$, estimated using the Fama–French five-factor model augmented with momentum. PES is the signed aggregate peer earnings signal, standardized within each calendar month. Controls are measured before event month m and include analyst disagreement, the focal firm's earnings surprise, size, book-to-market, analyst coverage, return volatility, turnover, business concentration, and profitability. Columns (1)–(3) include target-firm fixed effects and event-month fixed effects, where the event month is the calendar month containing the linked peer earnings announcement. Standard errors in columns (1)–(3) are two-way clustered by target firm and event month. Column (4) reports monthly Fama–MacBeth estimates with Newey–West standard errors computed using six monthly lags. t-statistics are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Return to text.

Table 12.
Aggregated Absolute Peer Earnings News and Belief Convergence

	Dependent variable: $\Delta Disagreement_{i,m+k}$ (percentage points)					
	$m+1$ (1)	$m+2$ (2)	$m+3$ (3)	$m+4$ (4)	$m+5$ (5)	$m+6$ (6)
PES^{Abs}	-0.016	0.043	0.146	0.237**	0.285**	0.361***
	(-0.23)	(0.53)	(1.44)	(2.01)	(2.22)	(2.76)
$\text{PES}^{Abs} \times PID$	-0.166	-0.365	-0.688*	-1.058**	-1.323***	-1.594***
	(-0.60)	(-1.14)	(-1.80)	(-2.38)	(-2.77)	(-3.27)
$\text{PES}^{Abs} \times PID^2$	0.251	0.424	0.645**	0.957**	1.193***	1.393***
	(1.02)	(1.50)	(1.97)	(2.53)	(2.96)	(3.36)
Current disagreement	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y
Time FE	Y	Y	Y	Y	Y	Y
Observations	543,679	534,258	525,600	516,921	508,836	501,337
Adj. R^2	0.114	0.187	0.249	0.300	0.341	0.377

Notes: This table tests whether the magnitude of peer earnings news is followed by convergence in analysts' beliefs and whether this relation varies nonlinearly with PID. The dependent variable is the change in monthly analyst forecast disagreement from month m to month $m+k$, expressed in percentage points. PES^{Abs} is the shared-analyst-weighted average of the absolute earnings surprises of the focal firm's analyst-linked peers. It is winsorized and standardized within each calendar month. The specifications include PID and PID^2 as main effects, together with the interactions of PES^{Abs} with PID and PID^2. They also control for current disagreement, profitability, the focal firm's earnings surprise and its absolute value, past returns, size, book-to-market, analyst coverage, return volatility, turnover, business concentration, the number of analyst-linked peers, the number of peer industries, and the total weight of shared-analyst links. All regressions include firm and month fixed effects. Standard errors are two-way clustered by firm and month; t-statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Return to text.

Appendix A Variable Construction

Table A1

Variable definitions

Variable	Definition
Analyst-linked peer	A firm covered by at least one analyst who also covers the focal firm.
Shared-analyst coverage	Number of analysts covering both the focal firm and peer firm.
Peer count	Number of distinct analyst-linked peers of the focal firm.
Peer earnings signal (PES)	Shared-analyst-weighted average of earnings surprises announced in month $t-1$ by analyst-linked peers: $$PES_{i,t} = \sum_{j \in \mathcal{P}^*_{i,t-1}} \omega_{ij,t-1} SUE_{j,t-1},$$ where $\mathcal{P}^*_{i,t-1}$ contains peers with valid earnings surprises announced in month $t-1$. The weights are proportional to shared-analyst coverage and sum to one within each firm-month.
Peer information dispersion (PID)	One minus the Herfindahl concentration of shared-analyst exposure across two-digit SIC industries: $$PID_{i,t} = 1 - \sum_{s \in \mathcal{S}_{i,t}} w^2_{i,s,t},$$ where $\mathcal{S}_{i,t}$ is the set of two-digit SIC industries represented in the peer set and $w_{i,s,t}$ is industry s's share of the focal firm's total shared-analyst exposure. PID uses all analyst-linked peers, including those that do not announce earnings in month $t-1$.
Peer relevance	Shared-analyst-exposure-weighted average of positive historical earnings-signal correlations between non-focal peer industries and the focal firm's industry.
Peer-signal redundancy	Weighted average of positive historical earnings-signal correlations across distinct peer industries, using pairwise products of shared-analyst exposure shares.
Industry earnings signal	Average earnings surprise announced in month m by firms in the focal firm's two-digit SIC industry, adjusted for the market-wide monthly mean and standardized by month.
Profitability	Quarterly net income divided by total assets, niq/atq.

(continued)

Table A1 – continued

Variable	Definition
Future profitability change	Average change in profitability from quarter q to quarters $q+1$ through $q+4$: $$\frac{1}{4}\sum_{h=1}^{4}\left(\text{Profitability}_{i,q+h} - \text{Profitability}_{i,q}\right).$$
Analyst consensus	Median EPS forecast across analysts covering the focal firm.
Analyst disagreement	Natural logarithm of one plus the cross-sectional standard deviation of analyst EPS forecasts scaled by the absolute median forecast: $$\text{Disagreement}_{i,t} = \log\left(1 + \frac{\text{SD}_a(\text{Forecast}_{a,i,t})}{\lvert\text{Median}_a(\text{Forecast}_{a,i,t})\rvert}\right).$$ The measure requires at least two analysts and a nonzero median forecast and is winsorized by month at the 1st and 99th percentiles (Diether et al., 2002).
Focal-firm earnings surprise	Most recently available earnings surprise of the focal firm.
Size	Natural logarithm of market capitalization.
Book-to-market	Natural logarithm of the ratio of book equity to market equity.
Analyst coverage	Number of analysts covering the focal firm.
Past return	Cumulative stock return from month $t-7$ through month $t-2$.
Return volatility	Stock-return volatility over the preceding 12 months.
Turnover	Monthly trading volume divided by shares outstanding.
Business concentration	Herfindahl index of segment sales.
Return	Monthly stock return.
Excess return	Monthly stock return minus the one-month risk-free rate.
Abnormal return	Stock return adjusted for the Fama–French five factors and momentum (Fama and French, 1993; Carhart, 1997; Fama and French, 2015).
Event cumulative abnormal return	Cumulative abnormal return of the target firm over the specified event window beginning on a linked peer firm's earnings-announcement date, such as CAR[0,5] or CAR[0,10].

Notes: Variables are measured at the firm-month level unless otherwise stated. Quarterly variables are aligned to the first available Compustat availability month. Event-level variables are measured around analyst-linked peer earnings announcement dates. Continuous variables are winsorized by month at the 1st and 99th percentiles where applicable and standardized as indicated in the corresponding table notes.

Appendix B Proofs and Derivations

B.1 Proof of Proposition 1

Under the symmetric benchmark,

$$\begin{aligned}\text{Var}\left(\frac{1}{N}\sum_{j=1}^{N}\eta_{ij}\middle|\mathcal{F}_i\right) &= \frac{1}{N^2}\left[N\sigma_{\eta,i}^2 + N(N-1)\rho_i\sigma_{\eta,i}^2\right] \\ &= \sigma_{\eta,i}^2\left[\rho_i + \frac{1-\rho_i}{N}\right]. \end{aligned} \tag{B.1}$$

Let

$$C_i \equiv \frac{b_i^2}{\sigma_{\eta,i}^2} > 0.$$

Equation (4) becomes

$$\kappa_i(N) = C_i\frac{N}{\rho_i N + 1 - \rho_i}.$$

Therefore,

$$\begin{aligned}\kappa_i(N+1) - \kappa_i(N) &= C_i\left[\frac{N+1}{\rho_i(N+1)+1-\rho_i} - \frac{N}{\rho_i N + 1 - \rho_i}\right] \\ &= \frac{C_i(1-\rho_i)}{(\rho_i N + 1 - \rho_i)[\rho_i(N+1)+1-\rho_i]} > 0 \end{aligned} \tag{B.2}$$

for $N \geq 1$ and $\rho_i \in [0, 1)$.

If $\rho_i > 0$,

$$\lim_{N\to\infty}\kappa_i(N) = \frac{b_i^2}{\sigma_{\eta,i}^2\rho_i}.$$

If $\rho_i = 0$,

$$\kappa_i(N) = \frac{b_i^2}{\sigma_{\eta,i}^2}N.$$

B.2 Proof of Proposition 2

Define

$$h_i(p) \equiv D_i(p) - L_i(p).$$

Under the conditions of Proposition 2, $h_i(p)$ is continuous and strictly decreasing on $\mathcal{P}_i = [\underline{p}_i, \overline{p}_i]$, with

$$h_i(\underline{p}_i) > 0, \qquad h_i(\overline{p}_i) < 0.$$

The intermediate value theorem implies that there exists a unique $p_i^* \in (\underline{p}_i, \overline{p}_i)$ satisfying

$$h_i(p_i^*) = 0.$$

Strict monotonicity implies

$$h_i(p) > 0 \quad \text{for } p < p_i^*, \qquad h_i(p) < 0 \quad \text{for } p > p_i^*.$$

From equation (7),

$$\frac{\kappa_i'(p)}{\kappa_i(p)} = h_i(p).$$

Because $\kappa_i(p) > 0$,

$$\kappa_i'(p) > 0 \quad \text{for } p < p_i^*, \qquad \kappa_i'(p) < 0 \quad \text{for } p > p_i^*.$$

Thus, p_i^* is the unique maximizer of $\kappa_i(p)$ on $\mathcal{P}_i$.

B.3 Derivation and Proof for the Relevance-Decay Model

For the equal-weight aggregate of the first k sources,

$$s(k) = \frac{1}{k} \sum_{g=1}^{k} (b_g u_i + \eta_g) .$$

Using $b_g = b_0 x^{g-1}$,

$$\begin{aligned} B(k) &= \frac{1}{k} \sum_{g=1}^{k} b_g \\ &= \frac{b_0}{k} \sum_{g=1}^{k} x^{g-1} \\ &= \frac{b_0}{k} \frac{1 - x^k}{1 - x}, \end{aligned} \tag{B.3}$$

which gives equation (10).

From equation (9),

$$\eta_g = c_g f + \sqrt{1 - c_g^2}\, e_g.$$

Thus,

$$\frac{1}{k} \sum_{g=1}^{k} \eta_g = \frac{1}{k} \left(\sum_{g=1}^{k} c_g \right) f + \frac{1}{k} \sum_{g=1}^{k} \sqrt{1 - c_g^2}\, e_g.$$

Mutual independence gives

$$\begin{aligned} \Omega(k) &= \frac{\sigma^2}{k^2} \left[\left(\sum_{g=1}^{k} c_g \right)^2 + \sum_{g=1}^{k} (1 - c_g^2) \right] \\ &= \frac{\sigma^2}{k^2} \left[k + \left(\sum_{g=1}^{k} c_g \right)^2 - \sum_{g=1}^{k} c_g^2 \right]. \end{aligned} \tag{B.4}$$

Because

$$c_g = \sqrt{\rho_0}\, y^{g-1},$$

we have

$$\sum_{g=1}^{k} c_g = \sqrt{\rho_0}\frac{1-y^k}{1-y}$$

and

$$\sum_{g=1}^{k} c_g^2 = \rho_0\frac{1-y^{2k}}{1-y^2}.$$

Substitution into equation (B.4) yields

$$\Omega(k) = \frac{\sigma^2}{k} + \frac{\sigma^2\rho_0}{k^2}\left[\frac{(1-y^k)^2}{(1-y)^2} - \frac{1-y^{2k}}{1-y^2}\right],$$

which gives equation (11).

B.3.1 Proof of Proposition 3(i)

Since $x \in (0,1)$,

$$0 < 1 - x^k \leq 1.$$

Therefore,

$$B(k)^2 = \frac{b_0^2}{k^2}\left(\frac{1-x^k}{1-x}\right)^2 \leq \frac{b_0^2}{k^2(1-x)^2}.$$

Moreover,

$$\left(\sum_{g=1}^{k} c_g\right)^2 - \sum_{g=1}^{k} c_g^2 = 2\sum_{1\leq g<h\leq k} c_g c_h \geq 0.$$

Equation (B.4) therefore implies

$$\Omega(k) \geq \frac{\sigma^2}{k}.$$

Consequently,

$$0 \leq \kappa(k) = \frac{B(k)^2}{\Omega(k)} \leq \frac{b_0^2}{\sigma^2(1-x)^2}\frac{1}{k}.$$

It follows that

$$\lim_{k\to\infty} \kappa(k) = 0.$$

B.3.2 Proof of Proposition 3(ii)

For $k = 1$,

$$B(1) = b_0, \qquad \Omega(1) = \sigma^2,$$

and therefore

$$\kappa(1) = \frac{b_0^2}{\sigma^2}.$$

For $k = 2$,

$$B(2) = \frac{b_0(1+x)}{2}.$$

Also,

$$\begin{aligned}\Omega(2) &= \frac{\sigma^2}{2} + \frac{\sigma^2\rho_0}{4}\left[(1+y)^2 - (1+y^2)\right] \\ &= \frac{\sigma^2}{2}(1+\rho_0 y). \end{aligned} \tag{B.5}$$

Hence,

$$\kappa(2) = \frac{b_0^2(1+x)^2}{2\sigma^2(1+\rho_0 y)}.$$

It follows that

$$\begin{aligned}\kappa(2) > \kappa(1) &\iff \frac{(1+x)^2}{2(1+\rho_0 y)} > 1 \\ &\iff (1+x)^2 > 2(1+\rho_0 y) \\ &\iff x^2 + 2x - 1 > 2\rho_0 y. \end{aligned} \tag{B.6}$$

B.3.3 Proof of Proposition 3(iii)

Suppose

$$x^2 + 2x - 1 > 2\rho_0 y.$$

By part (ii),

$$\kappa(2) > \kappa(1) > 0.$$

By part (i),

$$\lim_{k\to\infty} \kappa(k) = 0.$$

Therefore, there exists a finite integer $\bar{k} \geq 3$ such that

$$\kappa(k) < \kappa(2) \qquad \text{for all } k \geq \bar{k}.$$

Let $G_i \geq \bar{k}$, and let

$$k^* \in \arg\max_{k\in\{1,\ldots,G_i\}} \kappa(k).$$

Because

$$\kappa(2) > \kappa(1),$$

we have $k^* \neq 1$. Because

$$\kappa(k) < \kappa(2) \qquad \text{for all } k \geq \bar{k},$$

we also have $k^* < \bar{k}$. Hence,

$$k^* \in \{2, \ldots, \bar{k} - 1\}.$$

Since $G_i \geq \bar{k}$,

$$1 < k^* < G_i.$$

Thus, every maximizer is interior whenever the available source set satisfies $G_i \geq \bar{k}$.

B.4 Bayesian Updating and Information Gain

Conditional on $\mathcal{F}_i$,

$$u_i \sim N(0, \tau_u^{-1}), \qquad m_i = u_i + \omega_i, \qquad \omega_i \sim N(0, \kappa_i^{-1}),$$

with

$$u_i \perp \omega_i.$$

Gaussian conjugacy gives

$$E[u_i \mid m_i, \mathcal{F}_i] = \frac{\kappa_i}{\tau_u + \kappa_i} m_i$$

and

$$\mathrm{Var}(u_i \mid m_i, \mathcal{F}_i) = \frac{1}{\tau_u + \kappa_i}.$$

Therefore,

$$\lambda_i = \frac{\kappa_i}{\tau_u + \kappa_i}.$$

The proportional reduction in posterior variance is

$$\begin{aligned} 1 - \frac{\mathrm{Var}(u_i \mid m_i, \mathcal{F}_i)}{\mathrm{Var}(u_i \mid \mathcal{F}_i)} &= 1 - \frac{(\tau_u + \kappa_i)^{-1}}{\tau_u^{-1}} \\ &= 1 - \frac{\tau_u}{\tau_u + \kappa_i} \\ &= \frac{\kappa_i}{\tau_u + \kappa_i} \\ &= \lambda_i. \end{aligned} \tag{B.7}$$

Because $s_i = B_i m_i$, multiplication by the nonzero constant B_i does not change the squared conditional correlation. Moreover,

$$\mathrm{Cov}(u_i, m_i \mid \mathcal{F}_i) = \tau_u^{-1},$$

and

$$\mathrm{Var}(m_i \mid \mathcal{F}_i) = \tau_u^{-1} + \kappa_i^{-1}.$$

Thus,

$$\begin{aligned}
\mathrm{Corr}^2(u_i, s_i \mid \mathcal{F}_i) &= \mathrm{Corr}^2(u_i, m_i \mid \mathcal{F}_i) \\
&= \frac{\tau_u^{-2}}{\tau_u^{-1}(\tau_u^{-1} + \kappa_i^{-1})} \\
&= \frac{\kappa_i}{\tau_u + \kappa_i} \\
&= \lambda_i. \qquad \text{(B.8)}
\end{aligned}$$

Finally,

$$\frac{d\lambda_i}{d\kappa_i} = \frac{\tau_u}{(\tau_u + \kappa_i)^2} > 0.$$

Internet Appendix

When Is Aggregate Peer Information More Informative?

This Internet Appendix reports three sets of additional tests. First, it evaluates the robustness of the main findings to alternative variable constructions, empirical specifications, and sample definitions. Second, it examines potential alternative explanations for the results. Third, it provides additional evidence on the aggregation of residual peer information and the persistence of its predictive content across future-profitability horizons.

IA.1 Robustness and Additional Tests

IA.1.1 Alternative Peer Earnings Signal Measures

Table IA.1.

Alternative PES Measures

	Baseline	Log-transformed	Equal-weighted
PES $\times$ PID	0.276**	0.211*	0.214*
	(2.20)	(1.80)	(1.77)
PES $\times$ PID^2	−0.269**	−0.212*	−0.203*
	(−2.17)	(−1.90)	(−1.69)
Turning point	0.515	0.497	0.529
Slope at $PID = 0.10$	0.005	0.009	0.002
	(0.39)	(0.58)	(0.16)
Slope at $PID = 0.50$	0.051***	0.042**	0.039**
	(2.87)	(2.58)	(2.17)
Slope at $PID = 0.90$	0.012	0.008	0.011
	(0.75)	(0.61)	(0.73)

Notes: The log-transformed measure applies $\text{sign}(PES)\ln(1 + |PES|)$, and the equal-weighted measure assigns equal weight to each valid peer. All columns use the continuous quadratic specification and controls from Table 5. t-statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table IA.2.

PES Excluding Shared-Analyst Forecasts

	All forecasts	Excluding shared-analyst forecasts
PES $\times$ PID	0.215**	0.195*
	(2.04)	(1.94)
PES $\times$ PID^2	-0.195*	-0.174*
	(-1.92)	(-1.80)
Turning point	0.549	0.561
Slope at $PID = 0.10$	0.005	0.003
	(0.34)	(0.24)
Slope at $PID = 0.50$	0.044***	0.040***
	(3.02)	(2.96)
Slope at $PID = 0.90$	0.020	0.020
	(1.46)	(1.60)

Notes: Peer earnings surprises are reconstructed from I/B/E/S Detail History. The second column excludes forecasts issued by analysts who also cover the focal firm. Both columns use the continuous quadratic specification and controls from Table 5. t-statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

IA.1.2 Alternative PID Construction

Table IA.3.

PES Estimates Using GICS-Based PID

	Coefficient	t-statistic	p-value
PES $\times$ PID	0.316***	2.70	0.007
PES $\times$ PID^2	-0.310***	-2.78	0.005
Turning point	0.510		
PES slope at $PID = 0.10$	0.004	0.28	0.778
PES slope at $PID = 0.50$	0.056***	3.53	< 0.001
PES slope at $PID = 0.90$	0.009	0.59	0.554

Notes: PID is reconstructed using six-digit GICS industries instead of two-digit SIC industries. The specification is otherwise identical to the continuous quadratic specification in Table 5. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

IA.1.3 Robustness to PCA Windows and Link Weights

Table IA.4.

Robustness of the Residual Peer Earnings Signal

Specification	$PES^{Residual}$ × Low PID	$PES^{Residual}$ × Mid PID	$PES^{Residual}$ × High PID
Panel A. PCA estimation window			
48 months	−0.001 (−0.02)	0.032** (2.23)	−0.003 (−0.19)
72 months	−0.002 (−0.05)	0.036** (2.40)	0.007 (0.48)
96 months	0.006 (0.20)	0.036** (2.40)	0.010 (0.74)
Panel B. Weights used to aggregate residual peer signals			
Analyst-count weights	−0.010 (−0.36)	0.030** (2.31)	0.009 (0.87)
Capped link weights	−0.011 (−0.36)	0.029** (2.28)	0.012 (1.07)
Excluding the strongest peer link	−0.013 (−0.44)	0.026** (2.07)	0.016 (1.41)
Observations		96,374	

Notes: $PES^{Residual}$ is obtained after removing the variation captured by the first five historical principal components. Panel A varies the PCA estimation window, and Panel B varies the weights used to aggregate the residual peer signals. All specifications use the same 96,374 firm-month observations and the controls and fixed effects from Table 5. t-statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

IA.1.4 Robustness to Firm and Peer-Network Characteristics

Table IA.5.

Robustness to Firm and Peer-Network Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
PES, Mid PID	0.063***	0.064***	0.065***	0.071***	0.110***	0.111***
	(2.60)	(2.66)	(2.66)	(2.59)	(2.72)	(2.74)
Observed-network levels		Y	Y	Y	Y	Y
Size squared			Y	Y	Y	Y
$PES\times$ network characteristics				Y	Y	Y
$PES\times$ analyst coverage and size					Y	Y
$PES\times$ business concentration						Y
Public controls	Y	Y	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y	Y	Y
Month FE	Y	Y	Y	Y	Y	Y
Observations	114,078	114,078	114,078	114,078	114,078	114,078

Notes: This table examines whether the positive Mid-PID coefficient is explained by firm and peer-network characteristics. Observed-network levels include log observed peer count, log total shared-analyst weight, dominant-peer weight share, and peer-weight HHI. Network interactions additionally include the number of source industries. Business concentration is included in the public controls in every column; column (6) additionally interacts it with PES. Continuous variables are standardized by month before forming interactions. All specifications include the public controls and fixed effects from Table 5. Standard errors are two-way clustered by firm and month; t-statistics are in parentheses. Significance markers are based on asymptotic two-sided p-values. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

IA.1.5 Stability across Time Periods

Table IA.6.

Stability across Selected Time Periods

	PES		$PES^{Residual}$		
Sample	Mid PID	t	Mid PID	t	Observations
Full eligible sample	0.063***	2.60	0.038***	2.59	114,078
Excluding 2008	0.033*	1.82	0.028**	2.23	108,372
Excluding 2020 and 2021	0.069***	2.59	0.040**	2.57	101,764
Pre-2020	0.071**	2.43	0.034**	2.00	89,047

Notes: This table repeats the information-content regression in Table 5 across selected time periods. The dependent variable is the average change in the focal firm's quarterly profitability from quarter q to quarters $q+1$ through $q+4$. The reported coefficients are those on PES and $PES^{Residual}$ for Mid-PID firms. All specifications use the same controls and fixed effects as Table 5. Standard errors are two-way clustered by firm and month; t-statistics are reported in the adjacent columns. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

IA.1.6 Common-Shock Exclusion and Network Placebo

Table IA.7.

Common-Shock Exclusion and Network-Match Placebo

Panel A. Source-industry common-shock exclusion

Signal construction	Low PID	Mid PID	High PID	Mid–Low p	Mid–High p
Original residual PES	0.017 (0.53)	0.037** (2.53)	0.000 (0.03)	0.567	0.018
SIC2-month-demeaned residual PES	0.017 (0.82)	0.022** (2.17)	0.006 (0.67)	0.826	0.140

Panel B. Network-match placebo

Signal assignment	Low PID	Mid PID	High PID
Actual analyst-linked network	0.016 (0.50)	0.038*** (2.59)	0.000 (0.02)
Placebo mean	0.004	0.001	−0.002
Placebo 5th–95th percentiles	[−0.025, 0.034]	[−0.012, 0.016]	[−0.020, 0.014]
Share positive	56.0%	47.0%	39.0%

Notes: Both panels use residual peer earnings signals after removing the first five historical principal components. Panel A subtracts the SIC2-by-month mean residual earnings surprise from each source firm before aggregation and uses 113,911 firm-month observations. Panel B randomly reassigns residual signals across firm-months matched by month and observed peer count while retaining their PID groups and controls. It reports 100 random assignments; the randomization p-value for the actual Mid-PID estimate is 0.010. All specifications include controls in Table 5, firm fixed effects, and month fixed effects. t-statistics are in parentheses.

Table IA.8.

Bridge from the Baseline Specification to the Peer-Relevance Specification

	(1) Same-sample baseline	(2) Add relevance level controls	(3) Add PES × relevance	(4) Final specification
Panel A. Implied PES coefficient at median peer relevance				
Low PID	-0.037**	-0.037**	-0.037**	-0.042**
	(-2.04)	(-2.04)	(-2.06)	(-2.29)
Mid PID	0.044***	0.044***	0.042***	0.042***
	(3.64)	(3.63)	(3.40)	(3.37)
High PID	-0.002	-0.002	-0.002	0.005
	(-0.11)	(-0.11)	(-0.12)	(0.34)
Panel B. Low-PID implied PES coefficient across peer relevance				
Relevance at the 25th percentile	-0.037**	-0.037**	-0.034*	-0.038**
Relevance at the 50th percentile	-0.037**	-0.037**	-0.037**	-0.042**
Relevance at the 75th percentile	-0.037**	-0.037**	-0.039**	-0.045**
Common estimation sample	Y	Y	Y	Y
Controls	Y	Y	Y	Y
Firm FE	Y	Y	Y	Y
Time FE	Y	Y	Y	Y
Relevance level controls	N	Y	Y	Y
PES× relevance	N	N	Y	Y
Redundancy terms	N	N	N	Y
Observations	167,133	167,133	167,133	167,133
Firms	6,097	6,097	6,097	6,097
Months	271	271	271	271

Notes:

All columns use the estimation sample from the final specification. Column (1) reports the baseline specification reestimated on this common sample. Column (2) adds the level controls for peer relevance. Column (3) additionally includes the PES×peer-relevance interactions within the Low-, Mid-, and High-PID groups. Column (4) further adds the peer-signal-redundancy controls used in the final specification. Peer relevance and peer-signal redundancy are defined in Appendix A. Panel A reports the implied PES coefficients evaluated at the median level of peer relevance. Panel B reports the implied Low-PID coefficients evaluated at the 25th, 50th, and 75th percentiles of peer relevance. Standard errors are two-way clustered by firm and month. t-statistics are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

IA.1.7 Future Profitability by Quarter

Table IA.9.

Future Profitability by Quarter

Forecast horizon	Low PID	Mid PID	High PID
Quarter $q+1$	0.011 (0.75)	0.044*** (2.81)	0.025 (1.58)
Quarter $q+2$	0.000 (0.02)	0.077*** (2.81)	0.036** (2.06)
Quarter $q+3$	0.002 (0.08)	0.036* (1.75)	0.011 (0.48)
Quarter $q+4$	0.011 (0.60)	0.059** (1.96)	−0.008 (−0.38)

Notes: This table repeats the specification in Table 5 separately for profitability changes from quarter q to quarters $q+1$ through $q+4$. All horizons use the same estimation sample, controls, and fixed effects. Each horizon-specific profitability change is winsorized separately by availability month. Because the baseline outcome is first averaged across the four horizons and then winsorized, the average of the horizon-specific coefficients need not equal the baseline coefficient. Standard errors are two-way clustered by firm and availability month; t-statistics are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.