Crowded Analyst Coverage

Charles Martineau and Marius Zoican*

University of Toronto

March 19, 2020

Abstract

Analyst stock coverage is "crowded:" the most-covered 5% U.S. equities amount to 25% of earnings forecasts. Coverage clustering persists after controlling for factors commonly associated with analyst following, and correlates with investors' demand for information. Is information supply optimally distributed in markets? We build a model where analysts compete for scarce investor attention, providing forecasts that enhance investor learning. If clearing prices are not perfectly revealing, strategic complementarity effects dominate: analysts prefer to share a crowded space rather than "going against the wind" to cover relatively neglected assets. As investors' learning capacity increases, analyst coverage clustering is always excessive.

Keywords: analyst coverage, rational inattention, career concerns, information processing, learning
JEL Codes: G11, G24, G40, D83, M41

*Charles Martineau is affiliated with University of Toronto Scarborough and Rotman School of Management. Marius Zoican is affiliated with University of Toronto Mississauga and Rotman School of Management. Charles can be contacted at charles.martineau@rotman.utoronto.ca. Marius can be contacted at marius.zoican@rotman.utoronto.ca. Address: Rotman School of Management, 105 St. George Street, Toronto, Ontario; Canada M5S 3E6. We have greatly benefited from discussion on this research with Alexandre Corhay, Peter Cziraki, Craig Doidge, Jerôme Dugast, Thierry Foucault, Nicolas Inostroza, Itay Goldstein, Jim Goldman, Vincent Grégoire, Vincent Van Kervel, Tom McCurdy, Jordi Mondria, Shannon Liu, Chayawat Ornthanalai, Christine Parlour, Jorge Sabat, Liyan Yang, Mao Ye, and Yan Xiong. We are grateful to seminar participants at Université Paris-Dauphine, University of Sydney, University of Toronto Rotman School of Management and Department of Economics, 2019 Santiago Finance Workshop, for insightful comments.

Crowded Analyst Coverage

Abstract

Analyst stock coverage is "crowded:" the most-covered 5% U.S. equities amount to 25% of earnings forecasts. Coverage clustering persists after controlling for factors commonly associated with analyst following, and correlates with investors' demand for information. Is information supply optimally distributed in markets? We build a model where analysts compete for scarce investor attention, providing forecasts that enhance investor learning. If clearing prices are not perfectly revealing, strategic complementarity effects dominate: analysts prefer to share a crowded space rather than "going against the wind" to cover relatively neglected assets. As investors' learning capacity increases, analyst coverage clustering is always excessive.

Keywords: analyst coverage, rational inattention, career concerns, information processing, learning

JEL Codes: G11, G24, G40, D83, M41

1 Introduction

Is the supply of information optimally distributed in the cross-section of securities? In this paper, we argue that rational inattention on investors' side, coupled with the existence of career concerns for analysts, leads to sub-optimal "crowding" of analyst coverage. Indeed, in 2017 the most-covered 5% of stocks in the United States amassed 25.38% of the total analyst coverage, as measured by the number of earnings forecasts. Stocks such as PayPal (42 forecasts) and Amazon (39 forecasts) receive three and ten times as much analyst attention as the median S&P 500 stock (13 forecasts) and the median S&P 1500 stock (4 forecasts), respectively.

Stock forecast clustering, or "crowded" coverage, could simply reflect investor preferences. Limited attention drives investors to learn more about some securities and neglect others (Van Nieuwerburgh and Veldkamp, 2010). As documented by Frankel, Kothari, and Weber (2006) and Groysberg, Healy, and Maber (2011), analysts aim to maximize investor attention and generate trading fees for their brokerage houses. In this case, by focusing on a small subset of stocks, analysts simply cater to their customers. Some level of forecast clustering is therefore to be expected.

Excessively "crowded" analyst coverage might indicate a misallocation of resources as information supply is costly. Analysts play an essential role to produce information by converting low- to high-precision signals (Dugast and Foucault, 2018). However, they have limited amount of computer and brain power, which they can direct towards crunching unstructured data (e.g., Tweets), financial statement data, as well as interpreting news releases for a large number of public firms. Harford, Jiang, Wang, and Xie (2018) empirically show that analysts' coverage choices are driven by career concerns, and their incentives can therefore depart from investors' objectives. To the extent that investors value diversification, information production resources should be allocated to achieve the highest marginal benefit. Is the 40th forecast on a "popular" firm indeed more valuable for investors than the fifth forecast on the median S&P 500 stock?

In this paper, we build a noisy rational expectation equilibrium (noisy REE) model of portfolio optimization, with endogenous information acquisition and endogenous information supply to better understand analysts' incentives and competition. This paper's contribution is twofold. First, we show that if investors have limited attention, analysts do not cover stocks equally, but cluster their forecasts on a subset of securities. That is, information supply in financial markets is concentrated in some assets. The "crowded" coverage result is driven by strategic complementarity effects in analyst reporting choices. Second, we find that information supply is sub-optimally distributed in the cross-section of securities. While investors welcome some degree of "crowded" coverage towards assets with highly uncertain payoffs, analysts' forecasts cluster either too little or too much in equilibrium relative to the first best, and not necessarily on the assets where the marginal benefit of coverage is largest.

Section 2 lays down the facts. Analyst coverage tends to cluster, particularly in stocks with large market capitalization. Importantly, forecast clustering persists even within the S&P 500 and S&P 1500 universes: The first 10% of stocks by number of forecasts account for 20% and 28% of total coverage, respectively. Further, coverage crowding is a persistent pattern over the past decades, and does not decline as more analysts join the industry. Coverage clustering in the cross-section of stocks persists when we control for factors such as firm size, industry, and time fixed effects: it is not simply the case that large firms are able to attract more coverage by virtue of their size.

Finally, the coverage share of a firm is correlated with the information demand, or attention, from institutional investors. We follow Ben-Rephael, Da, and Israelsen (2017) and proxy investor attention by active Bloomberg terminal searches for firm-specific news. We find that the coverage share is 4.42 basis points higher for high-attention stocks (top quintile) than for low-attention stocks (bottom quintile), which amounts to half of the unconditional average coverage share.

In our model, risk-averse investors can learn: that is, acquire signals about the assets' stochastic dividends. More precise signals allow investors to reduce fundamental uncertainty about assets' payoffs. However, learning is not free: investors have limited attention, or learning capacity, which they need to allocate across securities. If investors spend more of their limited attention for a given asset, they receive a more precise signal about its payoff.

It is more difficult to learn about some stocks than others. In particular, investors learn more easily about assets with higher analyst coverage. That is, if more analysts focus on a stock, investors can spend less attention to obtain an equally precise signal. There are, however, decreasing returns to scale from analyst coverage, with each additional analyst having a lower marginal impact on learning costs. Information supply is endogenous: analysts compete over investor attention. Finally, analysts also have limited resources and choose to cover a single stock.

Strategic complementarity effects between analysts' forecast choices lead to crowded coverage in equilibrium. On the one hand, if more analysts cover the same asset, it is easier for an investor to acquire signals about the respective security. Therefore, she allocates a larger share of her limited capacity to learn about better covered assets. It follows that analysts have an incentive to cluster their forecasts and maximize investor attention to the respective security. On the other hand, going "against the wind" can also be attractive: issuing forecasts about a less well-covered asset gives an analyst quasi-monopoly power over investors' attention directed towards that stock.

What drives the analysts' choice? It turns out that the key friction driving the cross-sectional distribution of information supply is limited investor attention capacity. If attention is relatively more scarce, strategic complementarity effects dominate and coverage tends to become more crowded. A constrained investor cannot learn about too many assets at once. She has a preference for either ex ante riskier assets, where signals are more valuable, or better covered assets, where signals are cheaper. To fix intuition, assume nine analysts cover *FutureTech, Inc.*, an exiting technology stock, and none cover *SameOld Ltd*. The tenth analyst to make a coverage choice is better off receiving

1/10th of the attention directed towards *FutureTech* than becoming the only source of information about *SameOld*. Investors are so limited in their attention that they will learn little, if anything, about the less "exciting" stock. In this scenario, coverage begets coverage.

If investors are less attention-constrained, they acquire signals in equilibrium about a larger number of stocks. Therefore, analysts have lower incentives to cluster and higher incentives to differentiate from the crowd. They are able to attract enough investor attention even if they issue forecasts about less popular stocks. Consequently, analyst coverage becomes less crowded.

In which securities do analyst forecasts cluster? For equal coverage, investors prefer to learn about ex ante riskier assets. Therefore, securities with larger fundamental uncertainty are more likely to be crowded. This is in line with Barth, Kasznik, and McNichols (2001), who show that analysts have a preference for firms with a higher share of intangible assets. However, if stocks have similar levels of risk, investors' preference to learn about one security or another is weaker: multiple equilibria emerge in which analysts may crowd in both risky or safe securities.

Coverage clustering is to some extent dampened by strategic substitutability effects driven by informative market clearing prices. In the model, asset prices aggregate private signals and represent a public, free source of information for investors. If more analysts cover a stock, then investors acquire higher-precision signals about the stocks' payoff. At the same time, the stock price becomes more informative as it aggregates less noisy signals. A more informative stock price allows investors to shift capacity to learn about other assets, generating incentives for analysts to spread out their coverage.

Finally, crowded coverage is not necessarily benign. We find that if investors' learning capacity is large enough, equilibrium coverage is excessively skewed relative to the investors' first best. If investors face a weak capacity constraint, they optimally learn about all stocks. Consequently, they would prefer a less skewed analyst coverage distribution than the equilibrium outcome. For intermediate capacity levels, and sufficiently informative prices, strategic substitutability effects imply that analysts might actually cluster *too little* relative to the optimal level.

Related literature

Closest to our paper, Van Nieuwerburgh and Veldkamp (2010) model investors' portfolio choice under a learning capacity constraint and find that limited attention leads to under-diversification. We depart from their model in two important ways. First, not all securities are equally easy to learn about, and therefore attention is more valuable in some stocks than in others. Second, and most important, we introduce endogenous information supply, that is, analyst forecast choices. Investors' learning costs depend on analyst coverage, which in turn depends on the investor attention they can attract.

Dugast and Foucault (2018) analyze information acquisition in the context of Big Data. Investors are nowadays faced with an abundance of low-precision signals in the form of unstructured data: social media posts, satellite images, press releases. As such raw signals become cheaper, they crowd out more precise signals such as thorough analyst forecasts, leading to worse price discovery. Our paper yields a mirror prediction: if analysts coverage is crowded in a particular stock, it lowers the cost of high-precision signals, leading to better price discovery. [1]

Veldkamp (2006) argues that there are strategic complementarities in investors' demand for information. Since acquiring information has high fixed costs and low marginal costs (i.e., cheap to replicate), investors tend to learn about the same assets at the same time, leading to "media frenzies." We complement this channel by introducing an agency layer between investors, who ultimately hold the portfolio, and analysts who compete on investor attention and endogenously determine the elasticity of information supply. We show that limited investor attention enhances strategic complementaries, amplifying the agency problem.

Farboodi, Matray, and Veldkamp (2018) document that that while price informativeness of large S&P 500 firms improved over time, the price informativeness of the average public U.S. firm declined. Our paper can explain such gap through an endogenous information supply channel (in line with, e.g., Harford, Jiang, Wang, and Xie, 2018, who find that stocks with higher coverage have more transparent information environments).

Our paper relates to the rational inattention literature, as it implies investors optimally choose to learn less about some securities than others. Sims (2003) argues that investors face a cognitive capacity limit, and are therefore unable to process all available information. Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016) build a model where investment managers are skilled, but have limited attention. They conjecture that the attention constraints leads to more cross-sectional dispersion in fund investment strategies and a higher co-variance between fund holdings and systematic risk. Begenau, Farboodi, and Veldkamp (2018) argue that when investors are able to process more information, investment costs in large firms drop by relatively more than for small firms, and large firms grow larger. Mondria (2010) shows that rational inattention can lead to excess asset return comovement.

Morris and Shin (2002) and Goldstein and Yang (2019) argue that better information disclosure (e.g., mandatory disclosure, stress tests) can worsen price discovery since traders rely too much on public information at the expense of private signals. Our model complements this channel, showing that better disclosure leads to information supply crowding and further worsens price discovery for a subset of stocks.

We also contribute to a literature studying analysts' coverage choices and the impact of analyst

[1] Further, Zhu (2019) finds that the availability of unstructured data reduces information acquisition costs for investors, leading to improved price discovery and managers' investment decisions.

coverage. What determines the amount and quality of coverage a stock receives from an analyst? Bhushan (1989) presents a simple model of total expenditure by investors on analyst services and documents that firm size, institutional ownership, and stock price volatility are some of the most common characteristics correlating with analyst coverage. The literature on analyst coverage since then examines the determinants of coverage from the investor and the analyst perspective. From the investor perspective, Barth, Kasznik, and McNichols (2001) find that investors demand more analyst coverage for firms with high intangibles (e.g., high R&D expenditure) and more difficult to examine.[2]Because investors have limited attention, Hameed, Morck, Shen, and Yeung (2015) show that investors prefer to follow bellwether firms, i.e., firms with fundamentals that highly covary with their respective industry fundamentals. To meet the demand, analyst coverage is higher for these firms. Our model takes into account investors' limited attention and is the central friction determining analyst crowdedness.

Built on the premise that analysts also have limited attention, e.g., limited time, energy, and resources, which stocks an analyst decide to follow will be influenced by its career concerns (e.g., compensation).[3] Groysberg, Healy, and Maber (2011) find that analysts compensation and future employment prospect depends on their reputation ("All star" status recognized by the Wall Street Journal) and ability to generate commission revenue. Analysts are more likely to choose to follow firms with significant trading volume, and institutional ownership has they represent a more lucrative source of income for their brokerage house. Analysts might also prefer to put more effort into following larger firms because they are more visible and attract more institutional following, and therefore, their performance in researching these firms also may have a more substantial impact on their compensation (Hong and Kubik, 2003). In line with the empirical findings, we model analysts as maximizing investors' attention as a proxy for their monetary compensation.

Do analysts have superior predictive power? Das, Guo, and Zhang (2006) argue that analysts have superior predictive power and they select to cover stocks with better future prospects. In the same spirit, Derrien and Kecskes (2013) document that an exogenous drop in analyst coverage leads to higher information asymmetry. Finally, Lee and So (2017) uncover further evidence for analysts' ability to predict returns: firms with abnormally high analyst coverage outperform low-coverage firms by 80 basis points per month.

The impact of analyst coverage goes beyond predicting returns. Andrade, Bian, and Burch (2013) argue that analyst coverage can mitigate asset bubbles by coordinating investors' beliefs, particularly so when there is little disagreement between analysts. Merkley, Michaely, and Pacelli

[2]A recent literature documents theoretically and empirically higher investor attention to macroeconomic news (Andrei and Hasler, 2015; Boguth, Grégoire, and Martineau, 2019; Fisher, Martineau, and Sheng, 2019; Benamar, Foucault, and Vega, 2019) and to firms (Gargano and Rossi, 2018; Ben-Rephael, Carlin, Israelsen, and Da, 2019) when uncertainty about fundamentals is high.

[3]Empirical evidence of limited analyst attention include Clement (1999), Lauren, Lou, and Malloy (2014), and DeHaan, Shevlin, and Thornock (2015)

(2017) find empirical support that sell-side analyst coverage play an important role in promoting market efficiency. In the same spirit, Chung and Jo (1996) and Chen, Harford, and Lin (2015) document that analysts improve firm governance by reducing agency costs related to the separation of ownership and control.

2 Stylized facts

In this Section we document two main empirical patterns relating to the cross-sectional distribution of analyst coverage. First, analyst coverage clusters in a relatively small number of stocks, even after controlling for firm size and industry. Second, analyst coverage is positively correlated with information demand from institutional investors. Such patterns are persistent over time: the data set we use spans more than three decades. The empirical regularities discussed in this section are a natural starting point for the model introduced in Section 3.

2.1 Data

We retrieve analyst coverage data based on quarterly analyst earnings forecasts from I/B/E/S. We select only analyst forecasts issued over the 90 days before the earnings announcement date. If analysts revise their forecasts during this interval, we use only their most recent forecasts. We follow Livnat and Mendenhall (2006) and impose the following selection criteria for each firm earnings announcement i for firm quarter q:

1. The earnings announcement date is reported in Compustat.
2. The price per share is available from Compustat as of the end of quarter q and is greater than \$1 and the stock market capitalization is greater than \$5 million.
3. The firm's shares are traded on the New York Stock Exchange (NYSE), American Stock Exchange, or NASDAQ.
4. Accounting data, specifically total assets and market capitalization, is available in Compustat at the end of December of the previous calendar year.

The sample period is from January 1, 1985 to December 31, 2017 for a total of 15,208 unique firms and 525,834 earnings announcements. We then take the average number of quarterly forecasts for each firm and year.

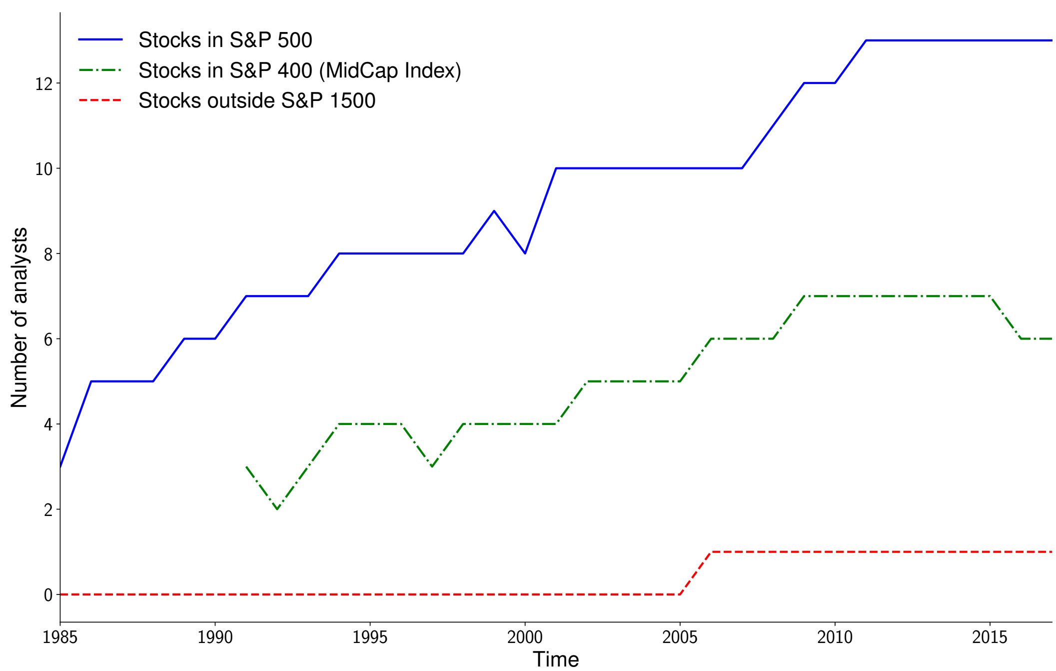
Further, we follow Ben-Rephael, Da, and Israelsen (2017) and construct a measure of investor attention based on institutional demand for information from the Bloomberg terminal. The Bloomberg measure of attention combines the number of times terminal users actively search, and subsequently read, news articles about a particular stock.[4] Bloomberg assigns a score of 10 when users search for news and a score of 1 when users read a news article. The scores are then aggregated into hourly counts for each stock and Bloomberg transforms these counts into daily measures varying from 0 (low abnormal attention) to 4 (high abnormal attention).

2.2 Analyst coverage patterns in the United States

Figure 1 shows the median analyst coverage for S&P 500 (large cap) stocks, S&P 400 (mid cap) stocks, and stocks outside the S&P 1500 index.[5] Stocks outside the S&P 1500 index had no earnings coverage before 2006, and only a single forecast per quarter between 2006 and 2017. In contrast, S&P 500 stocks coverage increased four-fold over the past three decades: from three analysts in 1985 to 13 in 2017. Mid-cap stocks, that is, securities in the S&P 400 index, lie in between large- and small-caps with a median of five forecasts in 2017. Therefore, large stocks receive relatively more analyst coverage than small stocks, and the effect is persistent over time.

Figure 1: **Median analyst coverage in the cross-section of stocks**

This figure plots the median number of analysts issuing an earnings forecast for stocks in the S&P 500 (large cap) and S&P 400 (mid cap) indices, as well as for stocks outside the S&P 1500 index, for each year between 1984 and 2017. Analyst coverage data are provided by I/B/E/S.



[4]For more detail on the construction of the measure, we refer the reader to Ben-Rephael, Da, and Israelsen (2017) and to `https://www3.nd.edu/~zda/AIA_App.pdf` on how to retrieve the data.

[5]Merkley, Michaely, and Pacelli (2017) document that since 2000, the number of analysts in the United States remained stable, at about 3000 to 3500 analysts.

The most-covered 5% of stocks account for 25.38% of the total number of forecasts. Moreover, crowded coverage has a certain "fractal" property: Forecast clustering is preserved, albeit to a lower extent, if we restrict our sample to index stocks, which are typically better covered. For S&P 1500 stocks and S&P 500 stocks, the top 5% of stocks account for 18% and 12% of analyst coverage, respectively.

Importantly, coverage clustering is not fully driven by firm size. Indeed, Hong, Lim, and Stein (2000) document that analyst coverage is strongly and positively correlated with market capitalization, and further the cross-sectional distribution of firm size is significantly right-skewed (Cabral and Mata, 2003). To provide evidence that analyst coverage clustering is a phenomenon beyond firm size and other variables commonly associated with analyst coverage, we regress the coverage share for each firm-year on market capitalization, institutional ownership, firm age, turnover, stock price volatility, and firm industry, that is

$$\begin{aligned}\text{Coverage share}_{i,t} &= \beta_0 + \beta_1 \log\left(\text{MarketCap}_{i,t}\right) + \beta_2 \text{InstitutionalOwnership}_{i,t} + \beta_3 \text{Age}_{i,t} \\ &\quad + \beta_4 \text{Turnover}_{i,t} + \beta_5 \text{Volatility}_{i,t} + \text{Industry}_{i,t} + \text{Year}_t + \text{error}_{i,t}, \qquad (1)\end{aligned}$$

where i runs over firms and t runs over years from 1985 to 2017.[6]

Coverage share is measured as the fraction of analysts covering stock i, that is

$$\text{Coverage share}_{i,t} = \frac{\text{Number of forecasts from unique analysts}_{i,t}}{\sum_i \text{Number of forecasts from unique analysts}_{i,t}}. \qquad (2)$$

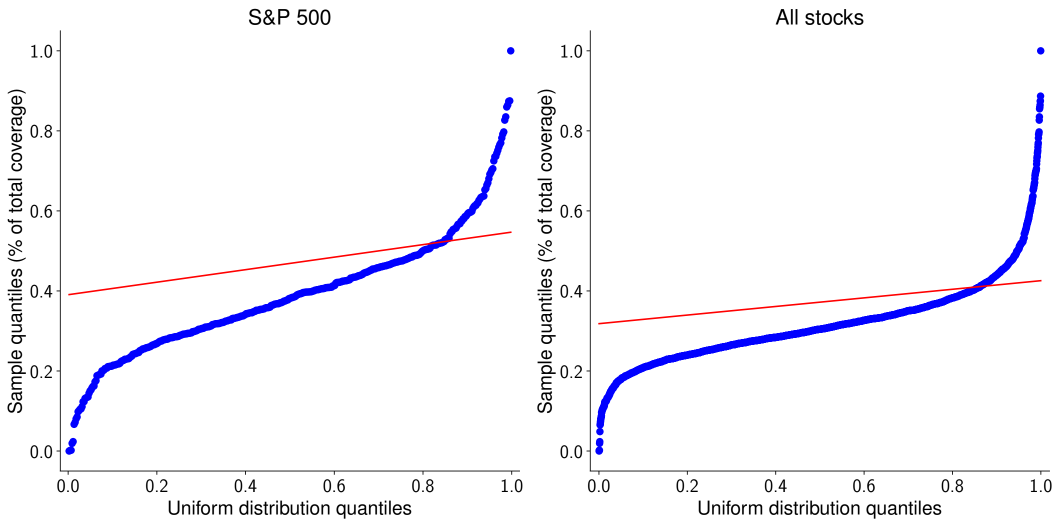
Figure 2 illustrates the distribution of regression residuals in a quantile-quantile plot for the year of 2017. We benchmark the sample "excess" coverage distribution against an uniform distribution, corresponding to a scenario where each stock receives an equal number of forecasts.

The salient feature of Figure 2 is a fat right tail, indicating significant coverage clustering even after controlling for size, industry, and time fixed effects, both for S&P 500 as well as the entire universe of stocks. Therefore, analysts have incentives to "crowd" their coverage that cannot be fully explained by a preference for large stocks. We note that since we only consider the *last* quarterly forecast for each analyst-stock pair, our clustering results are not affected by the frequency of news (i.e., a scenario in which some stocks might simply have more noteworthy events during the quarter, and therefore have more forecast updates).

[6] Market capitalization is the average market capitalization in year t, institutional ownership is the average quarterly institutional ownership in year t, the firm age proxy is the number of years the stock-PERMNO has been covered in CRSP, turnover is the average daily turnover in year t, and share volatility is the standard deviation of daily stock returns over year t. Industry dummies correspond to the 4-digit Global Industry Classification Standard.

Figure 2: **Analyst coverage clustering**

This figure is a quantile-quantile plot of residual earnings analyst coverage from equation (1) for the year of 2017. For each quarter, we take into account the last forecast update for a stock-analyst pair. Each observation corresponds therefore to a single stock. We benchmark the residual earning forecasts distribution against the uniform distribution. Data is from I/B/E/S. The straight red line is the expected order statistic scaled by the standard deviation of the sample plus the mean.



Information demand might drive analyst coverage choices. Recent empirical research find that analysts prefer to cover stocks that attract the most institutional investor attention (see, for example, Groysberg, Healy, and Maber, 2011; Harford, Jiang, Wang, and Xie, 2018).

We use the Bloomberg terminal investor attention data (Ben-Rephael, Da, and Israelsen, 2017) to relate information demand and analyst coverage. The Bloomberg attention measure has two important advantages. First, since it heavily weights active *searches* for firm information, it is a good proxy for information *demand* rather than for information supply – for example, the availability of headlines about a particular company.[7] Second, the measure is focused on institutional investors since they are more likely to use Bloomberg terminals: Ljungqvist, Marston, Starks, Wei, and Yan (2007), for example, show that analysts are sensitive to investment banking and brokerage pressure.

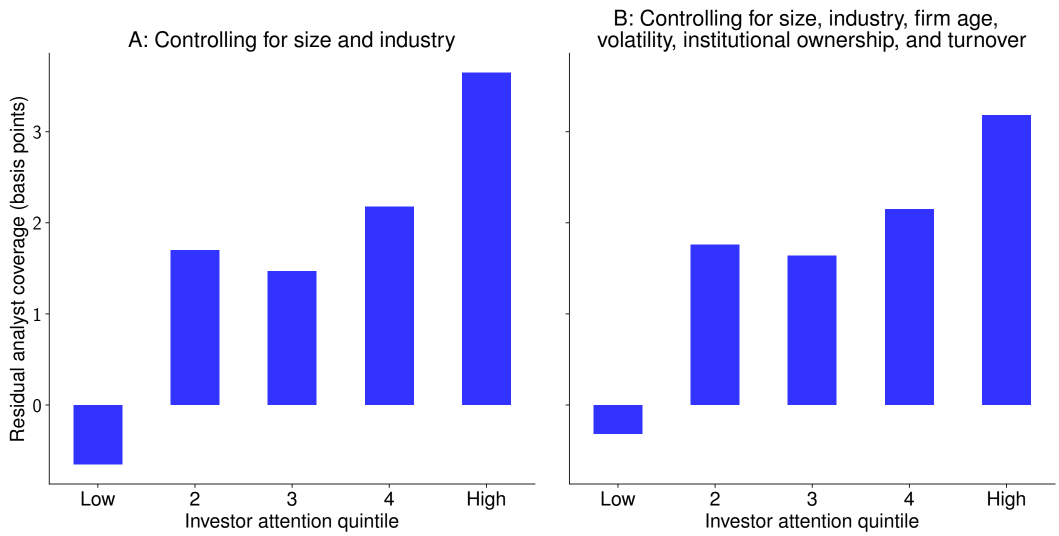
In Figure 3, we sort the average residual analyst coverage from equation (1) by investor attention, across S&P 500 stocks in 2017. In the left panel we sort "excess" coverage (i.e., after controlling for size and industry) by lagged abnormal investor attention, measured in 2016. In the right panel, we further control for additional variables specified in (1). Both panels show a positive and almost monotonic relationship between analyst coverage and investor attention. The economic magnitude of the effect is large: the "excess" coverage share is 4.42 basis points higher for high-attention stocks

[7]For example, the number of "clicks" on news items about a company also depends of the supply of news, not just on information demand.

than for low-attention stocks, amounting to almost half of the unconditional average coverage share (8.89 basis points).

Figure 3: **Investor attention and analyst coverage**

This figure plots the average residual analyst coverage in 2017 for S&P 500 firms by the investor attention quintile as measured in 2016. Residual analyst coverage are obtained after controlling for contemporaneous size and industry in Panel A and also controlling for institutional ownership, turnover, and volatility in Panel B (see equation 1). The institutional investor attention data is from the Bloomberg terminal as in Ben-Rephael, Da, and Israelsen (2017).



3 Model

Assets and markets. Consider a four-period economy, where time is indexed by $t \in \{-1, 0, 1, 2\}$. There are two risky assets available in zero net supply, labeled H and L. Each asset $i \in \{H, L\}$ pays off a stochastic dividend $v_i \sim \mathcal{N}\left(0, \tau_{vi}^{-1}\right)$ at $t = 2$, where dividends are independently distributed with mean zero and precision τ_{vi}. Without loss of generality, we assume that $\tau_{vH} > \tau_{vL}$: Therefore, H is a high-precision (low volatility) asset, and L is a low-precision (high volatility) asset. Additionally, agents can lend and borrow freely at a risk-free rate of r, which we normalize to one.

Each asset is traded on an auction market which clears at $t = 2$, just before the stochastic dividend is revealed. All trades are executed at the market clearing price.

Investors. There is a unit continuum of investors **I** with CARA utility over wealth at the terminal date $t = 2$, and a risk-aversion coefficient of γ. That is, the investors' expected utility can be written as

$$U_{\mathbf{I}} = -\mathbb{E}\left[\exp\left(-\gamma \tilde{W}_2\right)\right]. \tag{3}$$

At $t = 1$, conditional on all available information, each investor j chooses her demand $\mathcal{D}_{ij}(\cdot)$ for each asset $i \in \{H, L\}$ as a function of the market clearing price.

Noise traders participate in the asset market alongside investors. The noise trader demand x_i is independently distributed across assets i as a random normal variable with variance τ_x^{-1}: $\tilde{x}_i \sim \mathcal{N}\left(0, \tau_x^{-1}\right)$ The presence of noise traders guarantees that the market clearing price does not perfectly reveal the asset payoff, and generates incentives for traders to independently acquire information about the asset payoff.

Learning technology. At $t = 0$, before submitting her demand function, each investor j can acquire unbiased signals $s_{i,j} = \tilde{v}_i + \epsilon_{i,j}$, where $\epsilon_{i,j} \sim \mathcal{N}\left(0, \tau_{\epsilon i,j}^{-1}\right)$. Investors endogenously choose the signal precision, but have finite learning capacity and therefore cannot completely eliminate payoff risk.

We follow Van Nieuwerburgh and Veldkamp (2010) in modelling the investors' learning constraints. In particular, each investor $j \in \mathbf{I}$ is endowed with learning capacity $K > 0$ (a measure for cognitive processing power) which she can allocate across the two assets to acquire signals of precision $\tau_{\epsilon i,j}$. A "learning budget" constraint immediately obtains: that is, the weighted sum of precisions must be lower than the capacity K:

$$\sum_{i \in \{H,L\}} \phi_i \tau_{\epsilon i,j} \leq K. \tag{4}$$

Some stocks are relatively easier to learn about than others: To capture such heterogeneity, we introduce a stock-specific learning cost parameter $\phi_i > 0$ in the capacity constraint (4). A lower ϕ_i, all else equal, allows the investor to allocate less attention to stock i for a given signal precision $\tau_{\epsilon i}$. In Van Nieuwerburgh and Veldkamp (2010), it is equally easy to learn about any asset, that is $\phi_i = 1$, for any asset i.

Analyst coverage. There are $S = 3$ stock analysts in the economy. At $t = -1$, each analyst j chooses to cover a single asset $i \in \{H, L\}$. [8] Analyst coverage reduces investors' cost to learn about a given security: If n_i is the number of analysts covering stock i, then the marginal cost of learning about stock i is:

$$\phi(n_i) = \frac{1}{n_i}. \tag{5}$$

If more analysts cover a stock, investors learn more easily about its future payoffs. However, since $\dfrac{\partial^2 \phi}{\partial n^2} > 0$, analyst coverage has decreasing returns to scale. That is, each additional analyst covering a particular stock generates less incremental value for investors. Further, we assume that investors

[8] This assumption is consistent with e.g., Harford, Jiang, Wang, and Xie (2018) who argue that analysts have limited processing capacity.

only learn at $t = 1$ through analyst reports: if no analysts cover a stock ($n_i = 0$), the cost of learning is arbitrarily large. In Appendix C, we provide further motivation for our choice of ϕ. We argue that analyst coverage "compresses" information and reduces the necessary cognitive space to process a signal with a given variance, in a similar way as file compression reduces required storage space on a computer.

Each analyst j chooses to cover a single stock $i_j^\star$ to maximize the expected share of investor attention they capture. In equilibrium, investors allocate capacity $C_i \equiv \phi_i^\star \tau_{\epsilon i,t}^\star$ to stock i, which is shared by the $n_i^\star$ analysts covering the stock. Formally, the analyst's problem can be written as

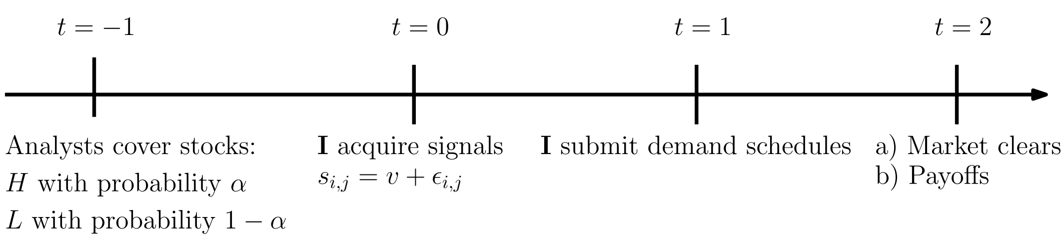
$$i_j^\star = \arg\max_i \mathbb{E}\left[\frac{1}{n_i^\star}\underbrace{\phi_i^\star \tau_{\epsilon i,t}^\star}_{C_i}\right], \tag{6}$$

where $\sum_i C_i \leq K$.

Analyst career concerns. In the model, the analysts' objective function (6) can be interpreted to emerge from career concerns (Holmström, 1999). Hong and Kubik (2003) show that analysts' careers depend less on forecast accuracy and more on their ability to promote stocks and generate investment banking business. In the same spirit, Frankel, Kothari, and Weber (2006) document that analysts direct their efforts towards researching firms with large trading volumes and institutional ownership to maximize commission fee revenue for brokerage houses. Finally, Groysberg, Healy, and Maber (2011) find that analyst compensation is related to investment-banking contributions, and reputation effects (for example, "All-Star" recognition) and not necessarily to the quality of the forecasts. Such reputation and business-generating incentives are consistent with analysts competing on investor attention. In particular, in light of the empirical evidence, the analysts' objective may not coincide with the investors' problem (i.e., reduce variance across the portfolio).

Timing. Figure 4 summarizes the sequence of events at each time $t \in \{-1, 0, 1, 2\}$.

Figure 4: **Model timing**



We closely follow Verrecchia (1982) and Goldstein and Yang (2017) in that we model a noisy rational expectation economy (noisy REE) with endogenous information acquisition. As in Van Nieuwerburgh and Veldkamp (2010), the cost of acquiring information is modeled as a capacity constraint on the sum of signal precisions.

Equilibrium. We are looking for stable subgame-perfect Nash equilibria in pure and mixed strategies.

> **Definition 1.** An *equilibrium* of the game consists of (i) each analyst j's choice of which stock to cover at $t=-1$, represented by a potentially mixed strategy $(\alpha_j, 1-\alpha_j)$, where α_j is the probability of covering the high-precision stock H; (ii) the investors' learning choices at $t=0$, summarized by the signal precision choice $\tau_{\epsilon i}^{\star}$ for each stock i, (iii) optimal demand schedules from investors at $t=1$, and (iv) market clearing prices $\mathbf{p}=(p_H, p_L)$ at $t=2$, such that no agent is strictly better off by deviating.

4 Equilibrium

We solve the game by backward induction. In Section 4.1 we solve for the market clearing prices at $t=2$ and the optimal investor demand schedule at $t=1$, taking the investor signal precision as given. Next, Section 4.2 yields the optimal learning choice of the investor in each stock, for a given cross-sectional distribution of analyst coverage. Finally, Section 4.3 solves for the analysts' equilibrium asset coverage strategy.

4.1 Market clearing and optimal demand schedules

Following Goldstein and Yang (2017), we conjecture that the market clearing price of asset i is a linear function of the stochastic dividend $\tilde{v}_i$ and the noise trading demand $\tilde{x}_i$. For convenience of exposition, and since the asset payoffs are independent, we drop the subscript i in the discussion below and state our linear conjecture as

$$\tilde{p} = p_v \tilde{v} + p_x \tilde{x}. \tag{7}$$

Next, we verify the linear price conjecture and obtain closed-form expressions for p_v and p_x. First off, note that the information revealed by the clearing price $\tilde{p}$ is equivalent to a signal $s_p = \tilde{v} + \rho^{-1}\tilde{x}$, where $\rho \equiv \dfrac{p_v}{p_x}$. Since investors have CARA utility and stochastic quantities are normally distributed, it immediately follows from Verrecchia (1982) that the optimal demand schedule from investor j is

$$\mathcal{D}_j\left(p \mid s_j\right) = \frac{\mathbb{E}\left[\tilde{v} \mid s_p, s_j\right] - p}{\gamma \mathrm{var}\left[v \mid s_p, s_j\right]}. \tag{8}$$

From Bayes' rule, and from the properties of the normal distribution, we obtain the expectation and variance in (8) as

$$\begin{aligned} \mathbb{E}\left[\tilde{v} \mid s_p, s_j\right] &= \frac{1}{\tau_v + \tau_{\epsilon,j} + \rho^2\tau_x}\left(\tau_{\epsilon,j}s_j + \rho^2\tau_x s_p\right) \text{ and} \\ \text{var}\left[v \mid s_p, s_j\right] &= \frac{1}{\tau_v + \tau_{\epsilon,j} + \rho^2\tau_x}. \end{aligned} \tag{9}$$

Consequently, the optimal demand schedule obtains by replacing the quantities in (9) in equation (8):

$$\mathcal{D}_j\left(p \mid s_j\right) = \frac{1}{\gamma}\left(\tau_{\epsilon,j}s_j + \rho^2\tau_x \underbrace{\left(\tilde{v} + \rho^{-1}\tilde{x}\right)}_{s_p} - p\left(\tau_v + \tau_{\epsilon,j} + \rho^2\tau_x\right)\right). \tag{10}$$

The equilibrium price is pinned down by the market clearing condition

$$\int_j \mathcal{D}_j\left(p \mid s_j\right) + \tilde{x} = 0. \tag{11}$$

Since there is a continuum of investors, their private signals average to the true payoff value, that is $\int_j s_j = v$. Rearranging equation (11) and solving for p, we obtain

$$\tilde{p} = \underbrace{\frac{\tau_\epsilon + \rho^2\tau_x}{\tau_v + \tau_\epsilon + \rho^2\tau_x}}_{p_v}\tilde{v} + \underbrace{\frac{\rho\tau_x - \gamma}{\tau_v + \tau_\epsilon + \rho^2\tau_x}}_{p_x}\tilde{x}, \tag{12}$$

which verifies the original linear price conjecture. To close equation (12), we note that solving $\rho = \frac{p_v}{p_x}$ yields the unique solution $\rho = \gamma^{-1}\tau_\epsilon$.

4.2 Optimal signal acquisition

At $t = 0$, each investor chooses signal precision $\tau_{\epsilon,i}$ for each asset i such that she maximizes her expected utility over wealth at $t = 2$. The expected utility at $t = 0$ is conditional on investors taking the optimal actions at $t = 1$: that is, to submit the demand schedule in (8) given the market clearing price in (12). It follows that the investors' problem is

$$\max_{\tau_{\epsilon,i}} \mathbb{E}^{t=0}_{\tilde{s}_{i,j},\tilde{p}_i}\mathbb{E}^{t=1}_{\tilde{v}_i}\left\{-\exp\left(-\gamma\sum_i\left[\left(\tilde{v}_i - \tilde{p}_i\right)\frac{\mathbb{E}\left[\tilde{v}_i \mid p_i, s_{i,j}\right] - p_i}{\gamma\text{var}\left[\tilde{v}_i \mid p_i, s_{i,j}\right]}\right)\right]\right\} \tag{13}$$

We assess the inner expectation at $t = 1$ using the properties of the log-normal distribution and rewrite the investors' problem in terms of the expected squared Sharpe ratios of each asset, conditional on the signal received:

$$\max_{\tau_{\epsilon,i}} \mathbb{E}^{t=0}_{\tilde{s}_{i,j},\tilde{p}_i} \left\{ -\exp\left(-\frac{1}{2} \sum_i \left[\frac{\left(\mathbb{E}\left[\tilde{v}_i \mid p_i, s_{i,j}\right] - \tilde{p}_i\right)^2}{\operatorname{var}\left[\tilde{v}_i \mid p_i, s_{i,j}\right]} \right) \right] \right\}. \tag{14}$$

From Lemma 2 in Verrecchia (1982, p. 1421) and equation (9), the $t = 0$ expectation in equation (14) reduces to

$$\max_{\tau_{\epsilon,i}} \prod_i \left(\underbrace{\left(\tau_{\epsilon,i} + \tau_{vi} + \rho_i^2 \tau_x\right)}_{\equiv \hat{\tau}_i} \frac{\tau_{vi} + p_{x,i}^2 \tau_x^{-1}}{\tau_{vi}} \right). \tag{15}$$

The first term in equation (15) is the posterior precision for asset i, given the precision of the signal ($\tau_{\epsilon,i}$) and the information extracted from the market clearing price ($\rho_i^2 \tau_x$). Since each investor has zero mass, she takes the price as given: That is, an individual investors' precision choice does not impact p_x or ρ, which are determined at market clearing through aggregating all investors' signals.

From equation (15) and following Van Nieuwerburgh and Veldkamp (2010), the investors' Lagrangian problem at $t = 0$ is to choose signal precisions $\tau_{\epsilon i}$ that maximize the product of posterior asset precisions, subject to capacity constraints:

$$\mathcal{L} = \sum_i \ln\left(\tau_{\epsilon i} + \tau_v + \rho^2 \tau_x\right) + \psi \left[K - \sum_i \frac{1}{n_i} \tau_{\epsilon i} \right] + \sum_i \lambda_i \tau_{\epsilon i}, \tag{16}$$

where ψ is the Lagrange multiplier of the attention capacity constraint and λ_i are Lagrange multipliers of constaints ensuring signal precisions are positive. The capacity constraint immediately follows from equations (4) and (5), that is

$$K - \sum_i \frac{1}{n_i} \tau_{\epsilon i} \geq 0. \tag{17}$$

In equilibrium, investors might learn about zero, one, or both stocks.[9] We define $\mathcal{M}$ to be the investors' "learning set," that is, the subset of stocks for which the equilibrium posterior variance is strictly higher than the prior variance:

$$\mathcal{M} \equiv \{ i \in \{H, L\} \mid \tau_{\epsilon i} > 0 \} \tag{18}$$

and we let $m \leq 2$ to be cardinality of $\mathcal{M}$. Therefore, investors acquire signals about m stocks. Proposition 1 states the partial equilibrium signal precision, conditional on the number of analysts covering each stock, n_i.

[9] Learning about zero stocks is equivalent to acquiring a zero-precision, or infinite variance, signal s_i at $t = 0$.

Proposition 1. (Posterior variances.) *For each stock $i \in \mathcal{M}$, for which the investor acquires a signal, the equilibrium posterior precision is*

$$\hat{\tau}_i \equiv \tau^\star_{\epsilon i} + \tau_{vi} + \rho_i^2 \tau_x = n_i \left(\frac{K}{m} + \frac{1}{m} \sum_{s \in \mathcal{M}} \frac{\tau_{vs} + \rho_s^2 \tau_x}{n_s} \right). \tag{19}$$

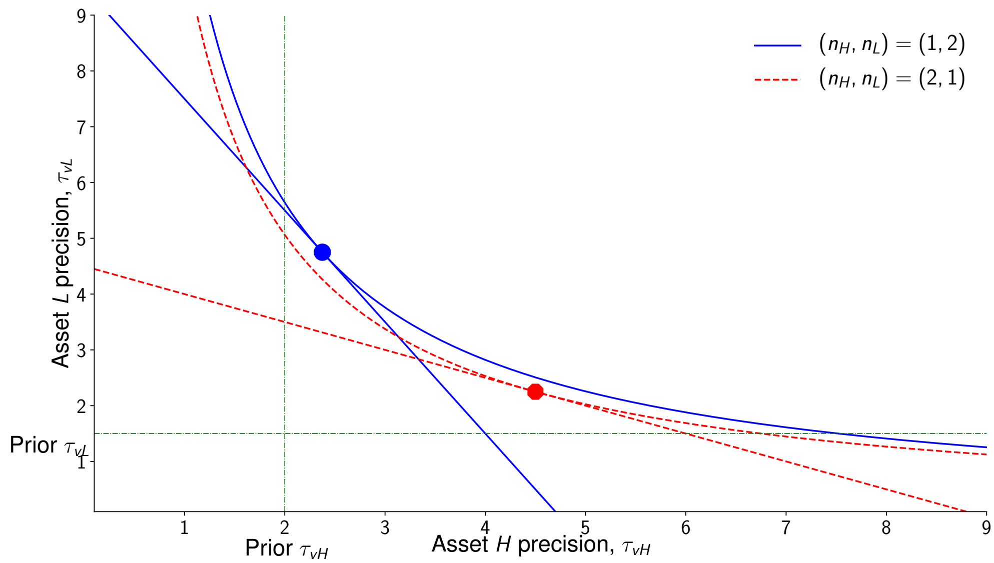
First, from Proposition 1, a higher attention capacity K yields a higher posterior precision $\hat{\tau}^\star$ for all assets in the investors' learning set $\mathcal{M}$. Second, investors learn about assets such that the posterior variances are proportional to the analyst coverage,

$$\frac{\tau^\star_{\epsilon i} + \tau_{vi} + \rho_i^2 \tau_x}{\tau^\star_{\epsilon j} + \tau_{vj} + \rho_j^2 \tau_x} = \frac{n_i}{n_j}, \forall i, j \in \mathcal{M}. \tag{20}$$

Finally, there are learning spillovers across stocks. If the prior precision τ_{vs} of a stock $s \in \mathcal{M}$ increases, the benefit from acquiring a signal about s is lower. The investor learns less about stock s, freeing up capacity to improve the posterior precision of other stocks. Otherwise, if more analysts cover stock s (i.e., n_s is larger), the marginal cost of learning about s decreases. Consequently, a larger share of the investors' learning capacity is directed towards stock s, and the posterior precision of the other asset is lower. Figure 5 illustrates the investors' optimization problem for two different analyst coverage configurations and for fully uninformative prices (i.e., for an arbitrarily small τ_x).

Figure 5: **Optimal investor learning**

This figure illustrates the investors' signal acquisition problem at $t = 0$. We plot the attention capacity constraint in equation (4) and tangent isoutility curve if (a) two analysts cover H and one analyst covers L (red dashed line) and (b) two analysts cover L and one analyst covers H (blue solid line). Parameter values: $\tau_{vH} = 2$, $\tau_{vL} = 1$, and $K = 2$. For simplicity, we illustrate the investors' learning problem in the limiting case where $\tau_x \to 0$, i.e., prices are not informative due to arbitrarily large noise trading.



Conditional on the number of analysts covering each stock $\eta = (n_i, n_{-i})$, the share of learning capacity that investors allocate to stock i is

$$C_i\left(n_i, n_{-i}\right) = \frac{1}{n_i}\tau^{\star}_{\epsilon i} = \max\left\{\frac{K}{m} + \left[\frac{1}{m}\sum_{s\in\mathcal{M}}\frac{\tau_{vs} + \rho_s^2\tau_x}{n_s} - \frac{\tau_{vi} + \rho_i^2\tau_x}{n_i}\right], 0\right\}. \tag{21}$$

A salient state variable in equation (21) is

$$\Psi_i \equiv \frac{\tau_{vi} + \rho_i^2\tau_x}{n_i}, \tag{22}$$

that is the precision of public information about asset i, scaled by the number of analysts covering i. Public information include the prior belief τ_{vi} and the information contained in the market clearing price. If investors share their attention equally across each asset in m, each stock is allotted $C_i = \frac{K}{m}$. However, in equilibrium, investors would tilt attention towards stocks for which Ψ_i is below average. That is, investors learn more about stocks that are ex-ante more uncertain (lower τ_{vi}), for which the price has a less informative signal (large $\rho_i^2\tau_x$), or about stocks that are better covered by analysts (higher n_i).

If the state variable Ψ_i is very large, either due to stock i having either low payoff uncertainty conditional on public information, or low analyst coverage, or both, investors choose not to learn about it and allocate zero capacity. From equations (19) and (21), it follows that if $C_i = 0$, then the posterior and prior precisions for stock i are the same – the clearing price itself is uninformative since traders acquire no signals.

To obtain a closed form solution for the equilbrium signal precision $\tau^{\star}_{\epsilon i}$, $i \in \{H, L\}$, we recognize that $\rho_i = \tau_{\epsilon i}\gamma^{-1}$ (from equation 12) and solve for a fixed point of equation (19). That is, the equilibrium signal precision chosen each investor j is a symmetric best response to other investors' signal choices, reflected in the market price.

The precision of the price signal, after substituting for ρ, is equal to $\frac{\tau_x}{\gamma^2}\tau^{\star,2}_{\epsilon,i}$. For exposition simplicity we define

$$\xi \equiv \frac{\tau_x}{\gamma^2}, \tag{23}$$

a measure proportional to price informativeness. A higher ξ stands in for a more informative market clearing price.

Table 1 enumerates the equilibrium closed-form solutions for $\tau^{\star}_{\epsilon i}$, for each asset i and every possible analyst coverage combination.

Table 1: **Investor optimal signal precisions**

Analyst coverage (n_H, n_L)	Equilibrium signal precision ($\tau^{\star}_{\epsilon H}$ and $\tau^{\star}_{\epsilon L}$)
$(3, 0)$	$\tau^{\star}_{\epsilon H} = 3K$ and $\tau^{\star}_{\epsilon L} = 0$
$(2, 1)$	$\tau^{\star}_{\epsilon H} = \frac{1}{\xi}\left(\sqrt{2\xi\left(4K^2\xi + 6K - \tau_{vH} + 2\tau_{vL}\right) + 4} - 2\right) - 2K$ $\tau^{\star}_{\epsilon L} = \frac{1}{2\xi}\left(2 - \sqrt{2\xi\left(4K^2\xi + 6K - \tau_{vH} + 2\tau_{vL}\right) + 4}\right) + 2K$
$(1, 2)$	$\tau^{\star}_{\epsilon H} = \frac{1}{2\xi}\left(2 - \sqrt{2\xi\left(4K^2\xi + 6K + 2\tau_{vH} - \tau_{vL}\right) + 4}\right) + 2K$ $\tau^{\star}_{\epsilon L} = \frac{1}{\xi}\left(\sqrt{2\xi\left(4K^2\xi + 6K + 2\tau_{vH} - \tau_{vL}\right) + 4} - 2\right) - 2K$
$(0, 3)$	$\tau^{\star}_{\epsilon H} = 0$ and $\tau^{\star}_{\epsilon L} = 3K$

Corollary 1. (Learning set.) *Let $\eta = (n_H, n_L)$ be the distribution of analysts covering stocks H and L,*

respectively, with $n_H + n_L = 3$*. We define two learning capacity thresholds,* $\kappa_1 > \kappa_0 > 0$*:*

$$\kappa_0 \equiv \max\left\{\frac{\sqrt{1+4\xi\left(2\tau_{vL}-\tau_{vH}\right)}-1}{4\xi}, \frac{\sqrt{1+2\xi\left(\tau_{vH}-2\tau_{vL}\right)}-1}{2\xi}\right\} \text{ and}$$

$$\kappa_1 \equiv \frac{\sqrt{1+4\xi\left(2\tau_{vH}-\tau_{vL}\right)}-1}{4\xi}. \tag{24}$$

Conditional on the analyst coverage distribution, investors weakly expand their learning set $\mathcal{M}$ *if capacity* K *increases, as tabulated below:*

Table 2: **Investors' learning set** $\mathcal{M}$

Analyst coverage	$K < \kappa_0$	$K \in (\kappa_0, \kappa_1]$	$K > \kappa_1$
$(n_H, n_L) = (3, 0)$		Learn about H	
$(n_H, n_L) = (2, 1)$	L or H	$\{H, L\}$	$\{H, L\}$
$(n_H, n_L) = (1, 2)$	L	L	$\{H, L\}$
$(n_H, n_L) = (0, 3)$		Learn about L	

Corollary 1 describes the investors' learning set $\mathcal{M}$ as a function of the analyst coverage distribution. There are four possible configurations for analyst coverage: two configurations where all analysts cover the same security (first and last row in Table 2) and two configurations where analysts cover both securities (second and third row in Table 2).

Naturally, if $\eta \in \{(3,0), (0,3)\}$, that is if analyst coverage focuses on a single stock, investors only learn about the respective security. The marginal cost of learning about the other, zero coverage, asset is arbitrarily large. Otherwise, if coverage is balanced, that is if $\eta \in \{(2,1), (1,2)\}$, investors learn about both securities only if their capacity K is large enough. For low learning capacity, investors only acquire signals about stocks that are either well-covered (and consequently easy to learn about) or have high prior uncertainty.

If $\eta = (1,2)$, investors have strong incentives to learn about stock L, as it is both better covered by analysts and more uncertain ex ante than stock H. Therefore, investors will start acquiring signals about stock H only if they have significant capacity, $K > \kappa_1$.

However, if $\eta = (2,1)$ the trade-off is less straightforward. Learning about stock H is easier due to better coverage; at the same time, signals about L are more valuable due to the larger prior uncertainty. Investors have stronger incentives to simultaneously learn about both stocks, and require a lower minimum capacity to do so, that is $\kappa_0 < \kappa_1$. Even if attention capacity is limited such that learning about one stock is optimal, investors acquire signals about high precision stock

H if

$$\frac{\tau_{vH}}{n_H} \leq \frac{\tau_{vL}}{n_L} \text{ or, equivalently, } \tau_{vH} \leq 2\tau_{vL}, \tag{25}$$

and acquire signals about L if $\tau_{vH} > 2\tau_{vL}$.

4.3 Endogenous analyst coverage

In this section we solve for the analysts' optimal coverage choice at $t = -1$. The analysts face the following trade-off as they compete to capture investors' attention. On the one hand, there are strategic complementarities in stock coverage. If more analysts cover an asset, it is cheaper for an investor to learn about the respective security, and she will allocate a larger share of her limited attention capacity to acquire signals about it. The strategic complementarity encourages analyst coverage clustering to maximize "the size of the pie," that is the share of investor attention to any given security. On the other hand, if analysts focus on the same stock, they need to share the investors' attention. If investors have sufficient learning capacity to learn about both stocks, it may be optimal for analysts to deviate from such "crowding" behavior and specialize in less-covered stocks. If analysts coverage is crowded, the direction of coverage clustering depends on the ex ante uncertainty about asset payoffs: all else equal, investors prefer to learn about low precision assets (i.e., stock L).

We are searching for both pure and mixed strategies equilibria. Let α and $1-\alpha$ be the probabilities that analyst j covers stock H and L, respectively. Analyst j's expected utility from covering stock H is

$$\begin{aligned} U_j(H) &= \sum_{n=0}^{2} \binom{2}{n} \alpha^n (1-\alpha)^{2-n} \frac{1}{n+1} C_H(n+1, 2-n) \\ &= \alpha^2 \frac{1}{3} C_H(3,0) + 2\alpha(1-\alpha) \frac{1}{2} C_H(2,1) + (1-\alpha)^2 C_H(1,2). \end{aligned} \tag{26}$$

With probability α^2, the two other analysts also cover stock H. In this case, analyst j captures one third of the attention allocated by investors to stock H. With probability $2\alpha(1-\alpha)$, the two other analysts cover different stocks. Therefore, if analyst j covers security H, the coverage distribution becomes $\eta = (2,1)$ and analyst j captures half of the attention allocated to H. With probability $(1-\alpha)^2$, the two other analysts cover L. Consequently, analyst j has a "monopoly" position in the high-precision security as he captures the entire attention share that investors allocate to asset H.

The expected utility from covering stock L is the mirror image of equation (26), that is,

$$
\begin{aligned}
U_j(L) &= \sum_{n=0}^{2} \binom{2}{n} \alpha^n (1-\alpha)^{2-n} \frac{1}{3-n} C_H(n, 3-n) \\
&= \alpha^2 C_H(2,1) + 2\alpha(1-\alpha) \frac{1}{2} C_L(1,2) + (1-\alpha)^2 \frac{1}{3} C_L(0,3) . \qquad (27)
\end{aligned}
$$

From Corollary 1, if all analysts cover stock i then investors allocate their entire attention to that stock. Therefore, it immediately follows that $C_H(3,0) = C_L(0,3) = K$. Otherwise, investor attention shares C_i depend on the level of learning capacity and in particular on whether investor choose to acquire signals about both stocks or not.

In a totally mixed strategy equilibrium with $\alpha^\star \in (0,1)$, analysts are indifferent between covering either of the two stocks. Therefore, $\alpha^\star$ is pinned down by the equilibrium condition

$$
U_j(H, \alpha^\star) = U_j(L, \alpha^\star) \text{ for all } j, \qquad (28)
$$

where $U_j(H, \alpha^\star)$ and $U_j(L, \alpha^\star)$ are defined in equations (26) and (27), respectively. If there is no $\alpha \in (0,1)$ that solves equation (28), a pure strategy equilibrium emerges.

Two types of analyst coverage equilibria can arise, as a function of prior asset variances. If the cross-sectional distribution of prior asset variances is very dispersed, that is if $\frac{\tau_H}{\tau_L} > 2$, a *high-dispersion* equilibrium emerges, in which analysts are more likely to cover the low-precision stock L. Proposition 2 describes the analysts' optimal coverage choice in this case. Conversely, if prior asset variances are not too different, if $\frac{\tau_H}{\tau_L} \leq 2$, a *low-dispersion* equilibrium emerges, described in Proposition 3, in which analysts can cluster their coverage in high-precision asset H if investor capacity is low, but have a bias for the low-precision asset L if investor capacity is high.

Proposition 2. (High-dispersion equilibrium) *If $\frac{\tau_H}{\tau_L} > 2$, that is for a large prior variance ratio, a* single-clustering equilibrium *emerges with the following analyst strategies:*

(i) For $K \leq \overline{K}$, all analysts cover the low-precision security L.

(ii) For $K > \overline{K}$, analysts cover the high-precision security H with probability $\alpha^\star < \frac{1}{2}$. Conversely, analysts cover the low-precision security L with probability $1 - \alpha^\star > \frac{1}{2}$.

where $\overline{K} \equiv \frac{3}{14\xi} \left(\sqrt{1 + 14\xi (2\tau_{vH} - \tau_{vL})} - 1 \right)$.

If the prior asset variances are very different, investors have a strong preference for learning about the low-precision asset L. In particular, they allocate a larger attention share to L even if the high-precision asset is better covered by analysts. From equation (21), the investors' bias for asset

L has a larger relative impact on attention allocation if investors' capacity K is low. Therefore, for a low K, investors spend no or little attention to learn about a relatively very safe asset. Consequently, analysts have little incentives to differentiate and cover the high-precision asset H: Even if they do, they cannot capture the investors' attention. Instead, analysts "crowd" their coverage in the relatively risky security L. As the investors' capacity increases, their preference for learning about L becomes relatively weaker and they become willing to allocate more attention to H. As a result, analysts have stronger incentives to differentiate and they cover the high-precision stock with positive probability in equilibrium. However, since investors have a strong preference for asset L, analyst crowding persists: the probability to cover H is $\alpha^\star < \frac{1}{2}$, and therefore analysts are biased towards covering stock L.

Proposition 3. (Low-dispersion equilibrium) *If $\frac{\tau_H}{\tau_L} \leq 2$, that is for a small prior variance ratio, a* low-dispersion equilibrium *emerges with the following analyst strategies:*

(i) For $K \leq \underline{K}$, there are two pure strategy equilbria: analysts cover either stock L or stock H with probability one. A third equilibrium, in mixed strategies, is not stable.

(ii) For $K \in \left(\underline{K}, \overline{K}\right]$, all analysts cover the low-precision security L.

(iii) For $K > \overline{K}$,analysts cover the high-precision security H with probability $\alpha^\star < \frac{1}{2}$. Conversely, analysts cover the low-precision security L with probability $1 - \alpha^\star > \frac{1}{2}$.

where $\underline{K} \equiv \frac{3}{14\xi}\left(\sqrt{1 + 14\xi\left(2\tau_{vL} - \tau_{vH}\right)} - 1\right)$ *and* $\overline{K} \equiv \frac{3}{14\xi}\left(\sqrt{1 + 14\xi\left(2\tau_{vH} - \tau_{vL}\right)} - 1\right)$.

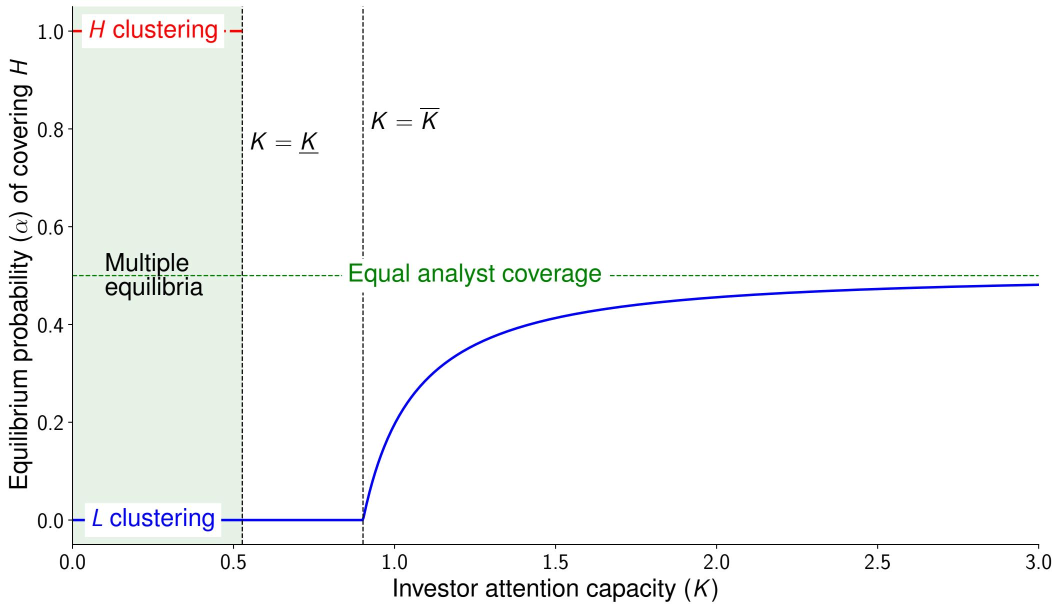
If assets have similar prior uncertainty (i.e., if $\frac{\tau_H}{\tau_L} \leq 2$), investors' preference to learn about L is weaker. In particular, if $\eta = (2, 1)$ such that asset H has more coverage, investors allocate more capacity to H than to L. Therefore, especially for limited investor attention, strategic complementarity effects are stronger. There is scope for analyst crowding in stock H, if each analyst believes its competitors will also cover the high-precision asset.

For low attention capacity, $K \leq \underline{K}$, investors have a strong preference to learn about the stock with more coverage: Two "crowded" equilibria arise where analysts cluster either in the low- or the high-precision stock. If K increases, investors would optimally allocate more attention to less well-covered stocks. Strategic complementarity effects are weakened in the following sense: Even if stock H is covered by two out of three analysts, and therefore cheap to learn about, it is optimal for the third analyst to deviate and cover L as he is able to attract enough investor attention. If investors' attention capacity K increases even further, the equilibrium strategies coincide with those in Proposition 2.

Figure 6 illustrates that the equilibrium probability of asset H coverage, α, weakly increases in investor capacity K, while never exceeding one half.

Figure 6: **Equilibrium analyst coverage probabilities**

This figure illustrates the equilibrium coverage probability for stock H, that is $\alpha^{\star}$, as a function of investor attention capacity K. Parameter values: $\tau_{vH} = 2$ and $\tau_{vL} = 1$, $\tau_x = 1.5$, and $\gamma = 1$.



From both Propositions 2 and 3, strategic complementarity effects are stronger when investor attention is limited. If learning capacity is low, investors primarily value stocks that are well-covered by analysts. In this case, analysts are better off sharing a crowded space rather than differentiating to cover new securities, as in the latter case they cannot the investors' attention. In contrast, if investor capacity K increases, analysts have stronger incentives to reduce coverage clustering and are more likely to cover the ex-ante safer security H.

To measure analyst crowding, we define a clustering measure based on the variance of analyst coverage. If both stocks are equally likely to be covered, that is if $\alpha = \frac{1}{2}$, the variance of analyst coverage variance reaches its maximum at $\frac{1}{4}$. We define

$$\text{Clustering}\,(\alpha) = 1 - 4\alpha\,(1-\alpha)\,, \tag{29}$$

which takes the value zero if both stocks are equally covered and one if only one stock is covered. Alternatively, we note that the skewness of a Bernoulli distribution is

$$\text{Skewness} = \frac{1-2\alpha}{\sqrt{\alpha\,(1-\alpha)}}, \tag{30}$$

which decreases in α. A lower α corresponds therefore to a more skewed analyst coverage distribution. However, skewness is not well-defined in the case pure strategy equilibria.

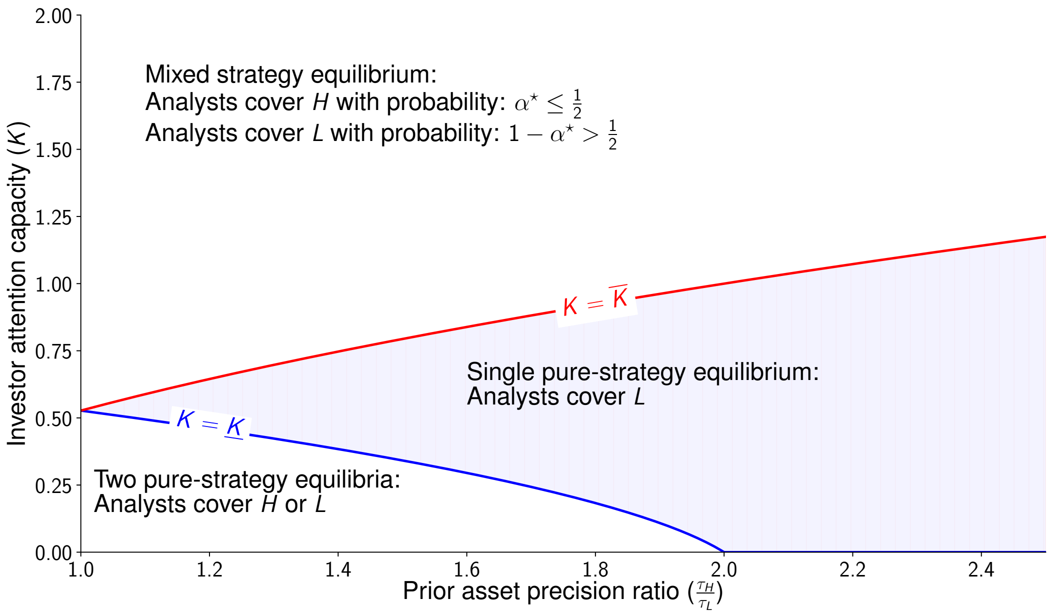
Corollary 2. (Capacity constraint) *The equilibrium probability of an analyst covering stock H, $\alpha^\star$, increases in the attention capacity K. For an arbitrarily large investor capacity K, analysts are equally likely to cover either stock, that is,*

$$\lim_{K\to\infty} \alpha^\star = \frac{1}{2}. \tag{31}$$

Corollary 2 emphasizes that analyst coverage clustering is a consequence of limited investor capacity. As investor capacity K increases, $\alpha^\star$ is larger and therefore analyst coverage becomes less crowded. Figure 7 illustrates the different equilibria of the game in the parameter space defined by investor capacity K and prior asset precision ratio. For low capacity and assets with similar risk, multiple pure strategy equilibria with full analyst coverage clustering arise. For large enough investor attention capacity, there is an unique mixed strategy equilibrium where analysts favor the low-precision security L.

Figure 7: **Equilibrium regions**

This figure illustrates the number and structure of analyst coverage equilibria, as a function of investor attention capacity K and asset prior variance ratio $\frac{\tau_{vH}}{\tau_{vL}}$. Parameter values: $\tau_{vL} = 1$, $\tau_x = 1.5$, and $\gamma = 1$.



Discussion. We emphasize two features of the theoretical setting. First, we do not introduce firm size: the assets are identical apart from payoff uncertainty. Therefore, our model is able to explain how coverage "crowding" can persist even for firms with similar market capitalization – in line with the empirical patterns we document in Section 2.

Second, introducing "free rider" effects to the model can reinforce crowding.[10] If the cost of covering a stock decreases in the number of analysts issuing forecasts for that firm, for example if analysts can recycle information from competitors' forecasts, they have additional incentives to co-ordinate. We conjecture that high-cost (low-skill) analysts are more likely to be "followers" and free-ride on low-cost (high-skill) competitors. Since low-skill analysts are unlikely to produce substantial information on their own, the marginal social cost of free-riding is likely to be low. We show that free-riding is not a necessary condition to obtain coverage clustering in equilibrium. Further, equally skilled analysts (with zero cost to produce a forecast) still have incentives to coordinate.

5 Model implications

5.1 Noise trading and analyst clustering

In equilibrium, investors learn about the asset payoff both directly, by acquiring private signals, as well as indirectly through the market clearing price. From the perspective of an individual investor, the two sources of information are substitutes. How does the amount of noise in prices influence analysts' coverage choices? Corollary 3 establishes that a more precise price signal (i.e., equivalent lower noise trading activity) leads to less coverage clustering in equilibrium.

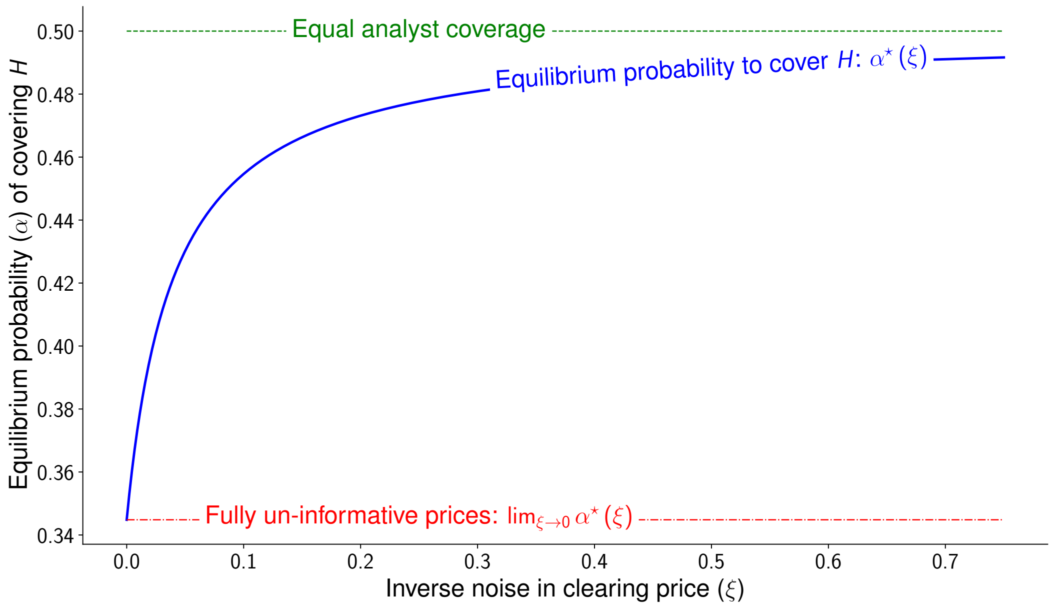
> **Corollary 3.** (Noise trading and clustering) *The equilibrium probability of an analyst covering stock H, $\alpha^\star$, decreases in the amount of noise trading (that is, increases in ξ) . For an arbitrarily small variance of noise trading, analysts are equally likely to cover either stock, that is,*
>
> $$\lim_{\xi \to 0} \alpha^\star = \frac{1}{2}. \tag{32}$$

An informative market price generates a strategic substitutability channel for analyst coverage choices. If analysts focus their coverage on asset i, investors are able to acquire relatively high-precision signals about asset i at a low cost, since focused analyst coverage relaxes their capacity constraint. Consequently, the price of asset i becomes relatively more informative (since it incorporates higher precision signals). Since investors can observe the clearing price of asset i for free, they have an incentive to re-allocate attention capacity to asset $-i$, for which the price remains relatively less informative. In equilibrium, analysts coverage turns towards asset $-i$ as well, since analysts choices follow the investor demand.

[10] Anecdotal evidence from equity analysts is that it is difficult to have access to reports from other analysts. Analysts usually obtain reports from other analysts through their clients. With high-profile brokers, investors must have an on-going relationship with the broker to have access to the reports.

Figure 8: **Analyst coverage choices and noise trading**

This figure illustrates the equilibrium coverage probability for stock H, that is $\alpha^{\star}$, as a function of the scaled inverse noise trading $\xi = \tau_x \gamma^{-2}$. Parameter values: $\tau_{vH} = 2$ and $\tau_{vL} = 1$, $K = 7.5$.



The strategic complementarity channel discussed in Section 4 enhances coverage clustering, as analysts cover the same stocks as their peers. In contrast, the strategic substitutability channel *reduces* clustering, as analysts focus on stocks for which they expect the price to be less informative. Figure 8 illustrates that, in equilibrium, the strategic complementarity channel dominates: any positive amount of noise in clearing prices is sufficient to generate analyst coverage clustering. However, coverage becomes less "crowded" if prices become more informative.

5.2 Optimal analyst coverage

Are investors' and analysts' objectives perfectly aligned? In Section 4, we show that analyst coverage tends to "crowd" in certain stocks in equilibrium. In this section, we analyze if such analyst clustering corresponds to the investors' preferred coverage distribution. From equation (16), the investors' objective function is the product of posterior precisions. From Proposition 1, the equilibrium posterior precision for asset i is

$$\hat{\tau}_i = \tau_{\epsilon i}^{\star} + \tau_{vi} + \xi \left(\tau_{\epsilon i}^{\star}\right)^2 , \tag{33}$$

where $\tau_{\epsilon i}^{\star}$ is the equilibrium signal precision in Table 1.

Clearly, for any capacity K, if analysts only cover one stock, investors prefer coverage to be clustered on asset L than on asset H. For $K < \underline{K}$, from Corollary 1, the investor only learns about one stock. Learning about stock L leads to a larger product of posterior precisions since

$$\underbrace{\left(3K + \tau_{vL} + 9\xi K^2\right)}_{\hat{\tau}_{vL}} \tau_{vH} > \underbrace{\left(3K + \tau_{vH} + 9\xi K^2\right)}_{\hat{\tau}_{vH}} \tau_{vL}. \tag{34}$$

Therefore, $\eta = (0, 3)$ dominates $\eta = (3, 0)$ and investors prefer that, conditional on full coverage clustering, that analysts favor asset L. From Proposition 3, a pure strategy equilibrium exists where, for low investor capacity, analysts choose to always cover asset H. Such equilibrium, if it arises, is always sub-optimal from the investors' perspective.

We define the investor-optimal analyst coverage distribution as the coverage distribution $(\alpha_{\text{optimal}}, 1 - \alpha_{\text{optimal}})$ that maximizes the product of posterior variances, given optimal learning behavior at $t = 1$. That is,

$$\alpha_{\text{optimal}} = \arg\max_{\alpha} \sum_{n=0}^{3} \binom{3}{n} \alpha^n (1-\alpha)^{3-n} \hat{\tau}_H (n, 3-n)\, \hat{\tau}_L (n, 3-n), \tag{35}$$

where $\hat{\tau}_H (n, 3-n)$ and $\hat{\tau}_L (n, 3-n)$ are computed using the optimal investor signal acquisition strategy in Corollary 1 for each n, and the corresponding posterior variances in Proposition 1.

There is generally no closed form solution to the problem in (35), and therefore we approximate α_{optimal} numerically. We find that relaxing the attention constraint (i.e., increasing K) leads to excessive coverage clustering.

Figure 9 plots the equilibrium analyst clustering defined in equation (29) and the optimal coverage clustering found numerically by maximizing the product in (35) against investor attention capacity.

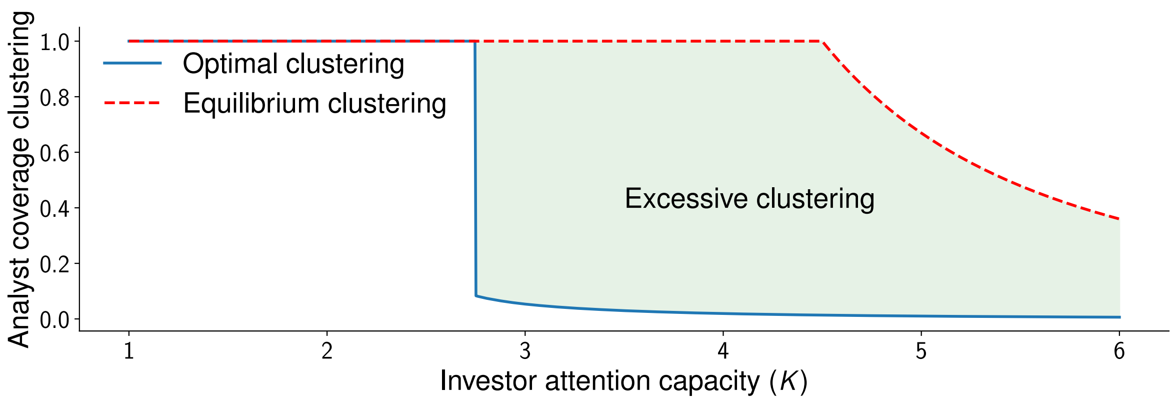
We distinguish between three regions of investor capacity K. First, for low values of K, analyst equilibrium behavior coincides with investors' preference and clustering is optimal. Investors are sufficiently attention-constrained such that it is optimal to only focus on one stock (i.e., the low-precision stock).

For intermediate levels of investor capacity, and only if the clearing price is informative enough, analysts may actually not focus their coverage enough relative to the investors' first best (i.e., "insufficient" clustering). An informative clearing price might dissuade analysts to cover the same assets as their peers. Crowded coverage in asset i generates a cheap public signal about the payoff of asset i, and allows investors to learn more about other stocks. Analysts have an incentive to cover "neglected" stocks, even though it could be sub-optimal from the perspective of attention-constrained investors.

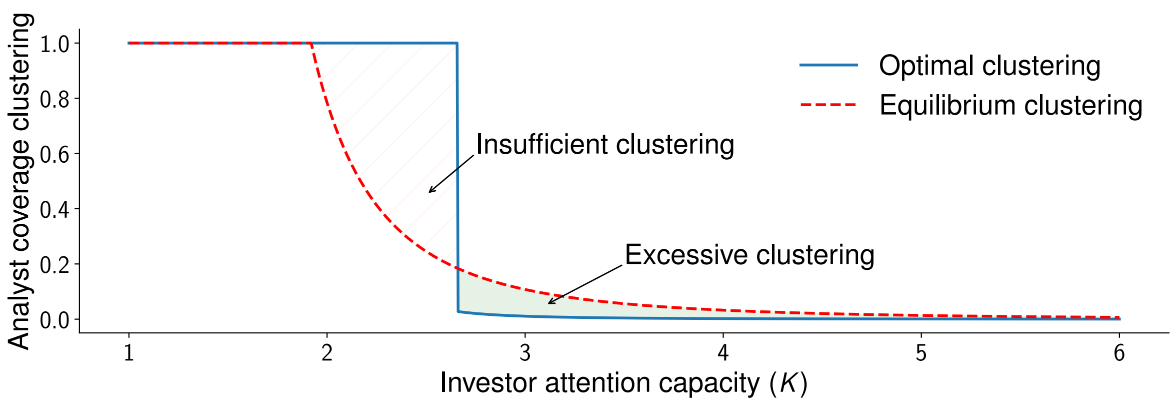
Finally, as investor capacity increases further, strategic complementarities dominate the welfare channel. For larger K, coverage clustering in equilibrium decreases too slowly compared to investors' optimum. Analyst crowding is "sticky" and persists even if the investors' attention constraint is relaxed. This is due to strategic complementarity effects in coverage choice that give incentives to analysts to share limited attention in crowded but popular stocks rather than opening up new information markets which might not attract enough attention at first.

Figure 9: **Sub-optimal analyst clustering**

This figure illustrates the equilibrium analyst coverage clustering, defined in equation (29) as well as the optimal analyst clustering from the investors' perspective, as a function of the investor capacity K and for two levels of noise trading demand: $\tau_x = 10^{-5}$ (top panel) and $\tau_x = 0.3$ (bottom panel). Parameter values: $\tau_{vL} = 1$ and $\tau_{vH} = 2$, $\gamma = 1$.



(a) Low price informativeness: $\tau_x = 10^{-5}$



(b) High price informativeness: $\tau_x = 0.3$

5.3 Price informativeness and coverage clustering

Analyst coverage clustering generates cross-sectional implications for price informativeness. We follow Verrecchia (1982) and defines price informativeness as the contribution of the clearing price to the posterior precision $\hat{\tau}_i$, that is

$$\text{Price informativeness}_i = \frac{\left(\tau_{\epsilon i}^{\star}\right)^2 \tau_x}{\gamma^2} \equiv \xi \left(\tau_{\epsilon i}^{\star}\right)^2, \tag{36}$$

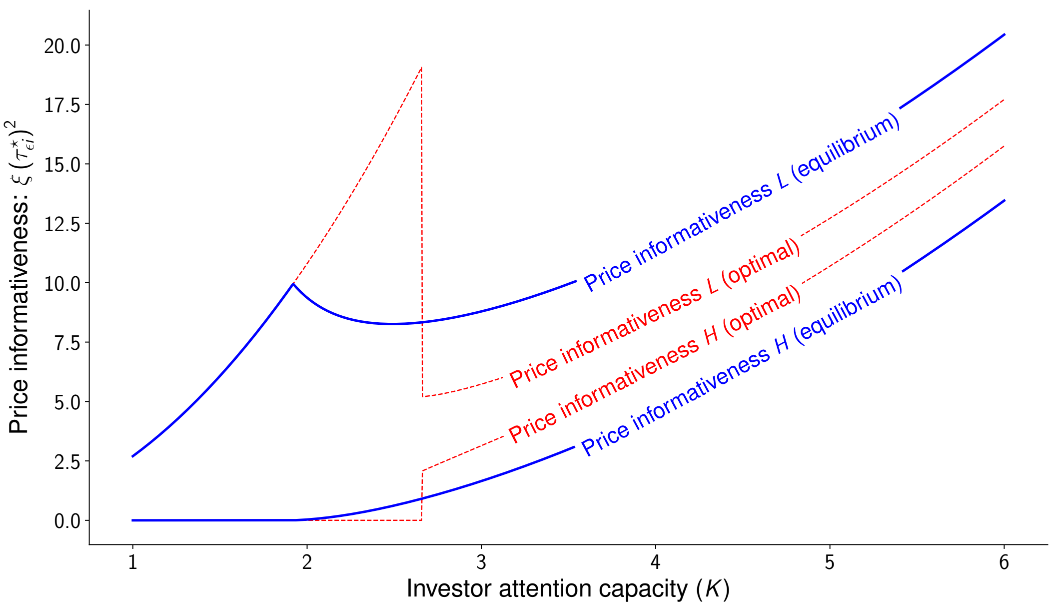
since $\xi = \tau_x \gamma^{-2}$. Figure 10 plots the expected price informativeness for each stock $i \in \{H, L\}$,

$$\mathbb{E}\text{Price informativeness}_i = \sum_{n=0}^{3} \binom{3}{n} \alpha^n (1-\alpha)^{3-n} \xi \left(\tau_{\epsilon i}^{\star}(n, 3-n)\right)^2, \tag{37}$$

for both the equilibrium $\alpha^\star$ and the optimal α that solves (35).

Figure 10: **Price informativeness**

This figure illustrates the expected price informativeness level for stocks H and L, as defined in equation (36), as a function of investor attention capacity K. We evaluate the expectation under both the equilibrium coverage probabilities $(\alpha^\star, 1-\alpha^\star)$ and the probabilities corresponding to the optimal coverage: $\left(\alpha_{\text{optimal}}, 1-\alpha_{\text{optimal}}\right)$ Parameter values: $\tau_{vH} = 2$ and $\tau_{vL} = 1$, $\tau_x = 0.3$, $\gamma = 1$.



We note that both insufficient and excessive clustering map to cross-sectional inefficiencies in price informativeness. For intermediate levels of investor capacity K, analysts tend to spread out their coverage too much, leading to a less informative price about asset L relative to the optimal level. For large investor capacity, analysts tend to "crowd" their coverage, leading to a less informative price about asset H relative to the optimal level.

6 Conclusion

This paper documents significant analyst coverage clustering in the U.S. equities universe. In 2017, popular stocks received up to ten times more analyst attention than the median S&P 1500 stock. Analyst crowding persist even after controlling for factor commonly associated with coverage (e.g., firm size and industry), and is correlated with institutional investors' attention. We build a model of endogenous learning and stock coverage in financial markets to study how such "crowded" coverage emerges in equilibrium. Can forecast clustering be fully explained by investor preferences, or does it signal inefficiencies in information supply?

The first-order insight emerging from our results is that crowded coverage can emerge in a model where stock analysts compete on scarce investor attention. If the availability of forecasts relaxes their learning constraints, investors are naturally drawn towards stocks with better coverage. Therefore, analysts are better off sharing a crowded space rather than trying to differentiate and cover relatively more obscure assets. Especially for limited investor attention, capturing a small fraction of high information demand for popular stocks is more profitable for analysts than capturing the entire information demand for a more obscure stock, which is low in absolute value. Coverage clustering emerges as a result.

The implications of our model match empirical patterns. We predict that analyst clustering occurs in stocks with lower return volatility and more intangible assets (as shown, e.g., by Barth, Kasznik, and McNichols, 2001).

Finally, from a normative perspective, our results suggest that crowded coverage can indicate sub-optimal information production. While investors always prefer some coverage clustering, particularly towards stocks with more uncertain fundamentals (e.g., intangible assets), the equilibrium level of analyst crowding is excessive, particularly so if investors have more attention to spare.

References

Andrade, Sandro C., Jiangze Bian, and Timothy R. Burch, 2013, Analyst Coverage, Information, and Bubbles, *Journal of Financial and Quantitative Analysis* 48, 1573–1605.

Andrei, Daniel, and Michael Hasler, 2015, Investor Attention and Stock Market Volatility, *The Review of Financial Studies* 28, 33–72.

Barth, Mary E., Ron Kasznik, and Maureen F. McNichols, 2001, Analyst Coverage and Intangible Assets, *Journal of Accounting Research* 39, 1–34.

Begenau, Juliane, Maryam Farboodi, and Laura Veldkamp, 2018, Big data in finance and the growth of large firms, *Journal of Monetary Economics* 97, 71–87.

Ben-Rephael, Azi, Bruce Carlin, Ryan D. Israelsen, and Zhi Da, 2019, Information consumption and asset pricing, *Journal of Finance, Forthcoming*.

Ben-Rephael, Azi, Zhi Da, and Ryan D. Israelsen, 2017, It Depends on Where You Search: Institutional Investor Attention and Underreaction to News, *The Review of Financial Studies* 30, 3009–3047.

Benamar, Hedi, Thierry Foucault, and Clara Vega, 2019, Demand for Information, Macroeconomic Uncertainty, and the Response of U.S. Treasury Securities to News, *Manuscript*.

Bhushan, Ravi, 1989, Firm characteristics and analyst following, *Journal of accounting and economics* 11, 255–274.

Boguth, Oliver, Vincent Grégoire, and Charles Martineau, 2019, Shaping expectations and coordinating attention: The unintended consequences of fomc press conferences, *Journal of Financial and Quantitative Analysis* 54, 2327–2353.

Cabral, Luis, and Jose Mata, 2003, On the evolution of the firm size distribution: Facts and theory, *American Economic Review* 93, 1075–1090.

Chen, Tao, Jarrad Harford, and Chen Lin, 2015, Do analysts matter for governance? Evidence from natural experiments, *Journal of Financial Economics* 115, 383–410.

Chung, Kee H, and Hoje Jo, 1996, The Impact of Security Analysts' Monitoring and Marketing Functions on the Market Value of Firms, *The Journal of Financial and Quantitative Analysis* 31, 493–512.

Clement, Michael B., 1999, Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter?, *Journal of Accounting and Economics* 27, 285–303.

Das, Somnath, Re-Jin Guo, and Huai Zhang, 2006, Analysts' Selective Coverage and Subsequent Performance of Newly Public Firms, *The Journal of Finance* 61, 1159–1185.

DeHaan, Ed, Terry Shevlin, and Jacob Thornock, 2015, Market (in) attention and the strategic scheduling and timing of earnings announcements, *Journal of Accounting and Economics* 60, 36–55.

Derrien, Francois, and Ambrus Kecskes, 2013, The Real Effects of Financial Shocks: Evidence from Exogenous Changes in Analyst Coverage, *The Journal of Finance* 68, 1407–1440.

Dugast, Jérôme, and Thierry Foucault, 2018, Data abundance and asset price informativeness, *Journal of Financial Economics* 130, 367–391.

Farboodi, Maryam, Adrien Matray, and Laura Veldkamp, 2018, Where Has All the Big Data Gone?, *Manuscript*.

Fisher, Adlai J., Charles Martineau, and Jinfei Sheng, 2019, Macroeconomic Attention and Announcement Risk Premia, *Manuscript*.

Frankel, Richard, S. P. Kothari, and Joseph Weber, 2006, Determinants of the informativeness of analyst research, *Journal of Accounting and Economics*.

Gargano, Antonio, and Alberto G. Rossi, 2018, Does it pay to pay attention?, *Review of Financial Studies* 31, 4595–4641.

Goldstein, Itay, and Liyan Yang, 2017, Information Disclosure in Financial Markets, *Annual Review of Financial Economics* 9, 101–125.

——— , 2019, Good disclosure, bad disclosure, *Journal of Financial Economics* 131, 118–138.

Groysberg, Boris, Paul M. Healy, and David A. Maber, 2011, What Drives Sell-Side Analyst Compensation at High-Status Investment Banks?, *Journal of Accounting Research* 49, 969–999.

Hameed, Allaudeen, Randall Morck, Jianfeng Shen, and Bernard Yeung, 2015, Information, analysts, and stock return comovement, *The Review of Financial Studies* 28, 3153–3187.

Harford, Jarrad, Feng Jiang, Rong Wang, and Fei Xie, 2018, Analyst Career Concerns, Effort Allocation, and Firms' Information Environment, *The Review of Financial Studies* 32, 2179–2224.

Holmström, Bengt, 1999, Managerial Incentive Problems: A Dynamic Perspective, *The Review of Economic Studies* 66, 169–182.

Hong, Harrison, and Jeffrey D. Kubik, 2003, Analyzing the Analysts: Career Concerns and Biased Earnings Forecasts, *The Journal of Finance* 58, 313–351.

Hong, Harrison, Terence Lim, and Jeremy C. Stein, 2000, Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies, *The Journal of Finance* 55, 265–295.

Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp, 2016, A Rational Theory of Mutual Funds' Attention Allocation, *Econometrica* 84, 571–626.

Lauren, Cohen, Dong Lou, and Christopher Malloy, 2014, Playing favorites: How firms prevent the revelation of bad news, *Working paper*.

Lee, Charles M.C., and Eric C. So, 2017, Uncovering expected returns: Information in analyst coverage proxies, *Journal of Financial Economics* 124, 331–348.

Livnat, Joshua, and Richard R. Mendenhall, 2006, Comparing the post-earnings announcement drift for surprises calculated from analyst and time series forecasts, *Journal of Accounting Research* 44, 177–205.

Ljungqvist, Alexander, Felicia Marston, Laura T. Starks, Kelsey D. Wei, and Hong Yan, 2007, Conflicts of interest in sell-side research and the moderating role of institutional investors, *Journal of Financial Economics* 85, 420–456.

Merkley, Kenneth, Roni Michaely, and Joseph Pacelli, 2017, Does the scope of the sell-side analyst industry matter? An examination of bias, accuracy, and information content of analyst reports, *The Journal of Finance* 72, 1285–1334.

Mondria, Jordi, 2010, Portfolio choice, attention allocation, and price comovement, *Journal of Economic Theory* 145, 1837–1864.

Morris, Stephen, and Hyun Song Shin, 2002, Social Value of Public Information, *American Economic Review* 92, 1521–1534.

Sims, Christopher A., 2003, Implications of rational inattention, *Journal of Monetary Economics* 50, 665–690.

Van Nieuwerburgh, Stijn, and Laura Veldkamp, 2010, Information acquisition and under-diversification, *Review of Economic Studies* 77, 779–805.

Veldkamp, Laura L., 2006, Media frenzies in markets for financial information, *American Economic Review* 96, 577–601.

Verrecchia, Robert E., 1982, Information Acquisition in a Noisy Rational Expectations Economy, *Econometrica* 50, 1415–1430.

Zhu, Christina, 2019, Big Data as a Governance Mechanism, *The Review of Financial Studies* 32, 2021–2061.

A Notation summary

Exogenous parameters and their interpretation.	
Parameter	Definition
τ_{vi}, $i \in \{H, L\}$	Prior precision of asset i (at time $t = -1$), without coverage.
τ_x	Inverse variance of noise trading demand.
γ	Investors' CARA risk-aversion.
ξ	Defined as $\tau_x \gamma^{-2}$, a scaled inverse of noise in prices.
$s_{i,j}$	Signal acquired by investor j in asset i.
K	Investors' capacity constraint.
Endogenous quantities and their interpretation.	
Variable	Definition
$\hat{\tau}_i$, $i \in \{H, L\}$	Posterior precision of asset i (at trading time $t = 1$).
$\tau_{\epsilon i}$	Precision of investors' private signal for asset i.
α	Probability of each analyst to cover the high-precision asset H.
ϕ_i	Marginal cost (in units of investor attention) to learn about asset i.
$\eta = (n_i, n_j)$	Number of analysts covering asset i and j, respectively.
C_i	Investor attention capacity share allocated to asset i, with $\sum_i C_i = K$.
Ψ_i	Precision of public information about asset i, scaled by n_i: $\frac{\tau_{vi}+\xi\tau_{\epsilon,i}^2}{n_i}$.
p_i	Market-clearing price of asset i at $t = 2$.
$\mathcal{D}_{ij}$	Demand for asset i from investor j, at $t = 1$.

B Proofs

Proposition 1

Proof. We compute the first order condition of Lagrangian $\mathcal{L}$ in equation (16), with respect to each posterior precision $\hat{\tau}_i$:

$$\mathcal{L} = \sum_i \ln\left(\tau_{\epsilon i} + \tau_v + \rho^2 \tau_x\right) + \psi\left[K - \sum_i \frac{1}{n_i}\tau_{\epsilon i}\right] + \sum_i \lambda_i \tau_{\epsilon i}. \tag{B.1}$$

That is, we can solve for the optimal posterior precision in terms of the Lagrange multipliers:

$$\frac{\partial \mathcal{L}}{\partial \tau_{\epsilon i}} = \frac{1}{\tau_{\epsilon i} + \tau_{vi} + \rho_i^2 \tau_x} - \frac{1}{n_i}\psi + \lambda_i. \tag{B.2}$$

From the complementary slackness condition, if $i \in \mathcal{M}$ such that the investor acquires a signal about i, the Lagrange multiplier of the positive precision constraint, λ_i, is zero and therefore

$$\tau_{\epsilon i}^{\star} = \frac{n_i}{\psi} - \left(\tau_{vi} + \rho_i^2 \tau_x\right), \forall i \in \mathcal{M}. \tag{B.3}$$

Finally, we solve for ψ from the capacity constraint:

$$\sum_{i \in \mathcal{M}} \frac{1}{n_i} \left(\frac{n_i}{\psi} - \left(\tau_{vi} + \rho_i^2 \tau_x\right) \right) = K \Longrightarrow \frac{1}{\psi} = \frac{1}{m} \left(K + \sum_{i \in \mathcal{M}} \frac{\tau_{vi} + \rho_i^2 \tau_x}{n_i} \right). \tag{B.4}$$

From equations (B.3) and (B.4), it follows that

$$\tau_{\epsilon i}^{\star} = n_i \left(\frac{K}{m} + \frac{1}{m} \sum_{i \in \mathcal{M}} \frac{\tau_{vi} + \rho_i^2 \tau_x}{n_i} \right) - \left(\tau_{vi} + \rho_i^2 \tau_x\right), \tag{B.5}$$

which concludes the proof. □

Corollary 1

Proof. First, whenever $n_i = 0$ and $n_{-i} = 3$, the marginal cost of acquiring signals about stock i is arbitrarily large and the investors can only learn about stock $-i$.

We use the constraint that $\tau_{\epsilon i} > 0$ to establish when investors choose to learn about both stocks if $\eta = (2, 1)$. Using the signal precisions from Table 1, the two constraints can be written as:

$$2 + 4K\xi > \sqrt{2\xi \left(4K^2\xi + 6K - \tau_{vH} + 2\tau_{vL}\right) + 4} > 2 + 2K\xi, \tag{B.6}$$

leading to the solution:

$$K > \kappa_0 \equiv \max \left\{ \frac{\sqrt{1 + 4\xi \left(2\tau_{vL} - \tau_{vH}\right)} - 1}{4\xi}, \frac{\sqrt{1 + 2\xi \left(\tau_{vH} - 2\tau_{vL}\right)} - 1}{2\xi} \right\} \tag{B.7}$$

We note that, depending on whether $\tau_L \lesseqgtr \frac{\tau_H}{2}$, one of the thresholds is negative and the other is positive. In particular, investors optimally learn about asset H if $\tau_{vH} \leq 2\tau_{vL}$ and optimally learn about asset L if $\tau_{vH} > 2\tau_{vL}$.

Next, we write the no-forgetting constraints if $\eta = (1, 2)$:

$$2 + 4K\xi > \sqrt{2\xi \left(4K^2\xi + 6K - \tau_{vL} + 2\tau_{vH}\right) + 4} > 2 + 2K\xi, \tag{B.8}$$

leading to the solution:

$$K > \kappa_1 \equiv \frac{\sqrt{1 + 4\xi\left(2\tau_{vH} - \tau_{vL}\right)} - 1}{4\xi}, \tag{B.9}$$

which is always positive since $\tau_{vH} > \tau_{vL}$. It immediately follows that $\kappa_1 > \kappa_0$. $\square$

Proposition 2

Proof. From Corollary 1, the investors' learning set depends on the capacity K. We solve for the optimal coverage choices for different regions of K, assuming $\frac{\tau_{vH}}{\tau_{vL}} \geq 2$.

Region 1: $K \leq \frac{\sqrt{1 + 2\xi\left(\tau_{vH} - 2\tau_{vL}\right)} - 1}{2\xi}$. From Corollary 1, investors acquire signals about H if and only all three analysts cover stock H. From equations (21), (26), (27), and Corollary 1, the expected analyst utilities from covering H and L are, respectively:

$$\begin{aligned} U\left(H\right) &= \alpha^2 \frac{K}{3} \text{ and} \\ U\left(L\right) &= \alpha^2 K + 2\alpha\left(1-\alpha\right)\frac{K}{2} + \left(1-\alpha\right)^2 \frac{K}{3}. \end{aligned} \tag{B.10}$$

It immediately follows that $U\left(L\right) > U\left(H\right)$ for any $\alpha \in [0,1]$ and therefore it is optimal for each analyst to cover the low-precision security L with probability one.

Region 2: $K \in \left(\frac{\sqrt{1 + 2\xi\left(\tau_{vH} - 2\tau_{vL}\right)} - 1}{2\xi}, \frac{\sqrt{1 + 4\xi\left(2\tau_{vH} - \tau_{vL}\right)} - 1}{4\xi}\right]$. From Corollary 1, if $\eta = (2,1)$ investors learn about both securities and if $\eta = (1,2)$ investors only learn about L. From equations (21), (26), (27), and Corollary 1, the expected analyst utilities from covering H and L are, respectively:

$$\begin{aligned} U\left(H\right) &= \alpha^2 \frac{K}{3} + 2\alpha\left(1-\alpha\right)\frac{1}{2} C_H\left(2,1\right) + \left(1-\alpha\right)^2 \times 0 \text{ and} \\ U\left(L\right) &= \alpha^2 C_L\left(2,1\right) + 2\alpha\left(1-\alpha\right)\frac{K}{2} + \left(1-\alpha\right)^2 \frac{K}{3}. \end{aligned} \tag{B.11}$$

The difference between the two utility functions is linear in α, that is

$$U\left(L\right) - U\left(H\right) = \frac{K}{3} + \alpha\left(\frac{K}{3} - C_H\left(2,1\right)\right). \tag{B.12}$$

We will prove that the two utility functions do not cross for any $\alpha \in [0,1]$. Indeed,

$$
\begin{aligned}
U(L,\alpha=0)-U(H,\alpha=0) &= \frac{K}{3} > 0 \text{ and} \\
U(L,\alpha=1)-U(H,\alpha=1) &= \frac{2K}{3} - C_H(2,1) > 0,
\end{aligned} \tag{B.13}
$$

since $\frac{2K}{3} - C_H(2,1)$ has two negative roots for $\tau_{vH} > 2\tau_{vL}$. Therefore, it is optimal for each analyst to cover the low-precision security L with probability one for any $K \leq \frac{\sqrt{1+4\xi(2\tau_{vH}-\tau_{vL})}-1}{4\xi}$.

Region 3: $K > \frac{\sqrt{1+4\xi(2\tau_{vH}-\tau_{vL})}-1}{4\xi}$. For large K, from Corollary 1, investors always acquire signals about a stock as long as there is at least one analyst covering it. From equations (21), (26), (27), and Corollary 1, the expected analyst utilities from covering H and L are, respectively:

$$
\begin{aligned}
U(H) &= \alpha^2\frac{K}{3} + 2\alpha(1-\alpha)\frac{1}{2}C_H(2,1) + (1-\alpha)^2 \times C_H(1,2) \text{ and} \\
U(L) &= \alpha^2 C_L(2,1) + 2\alpha(1-\alpha)\frac{1}{2}C_L(1,2) + (1-\alpha)^2\frac{K}{3}.
\end{aligned} \tag{B.14}
$$

The difference between the two utility functions is

$$
\Delta(\alpha) \equiv U(H) - U(L) = \alpha\left(\frac{2}{3}K - 2C_H(1,2) + C_H(2,1) - C_L(1,2)\right) + C_H(1,2) - \frac{K}{3}. \tag{B.15}
$$

After some algebraic manipulation, it can be shown that there are two capacity thresholds,

$$
\overline{K} = \frac{3}{14\xi}\left(\sqrt{1+14\xi(2\tau_{vH}-\tau_{vL})}-1\right) \tag{B.16}
$$

$$
\underline{K} = \frac{3}{14\xi}\left(\sqrt{1+14\xi(2\tau_{vL}-\tau_{vH})}-1\right) < \kappa_0. \tag{B.17}
$$

such that $\Delta(0) > 0$ if and only if $K > \kappa_0$ and $\Delta(1) > 0$ if and only if $K < \kappa_1$.

Note that there is an unique value of α for which the utility difference changes sign, that is,

$$
\alpha^\star = \frac{3\sqrt{2\xi(4K^2\xi+6K+2\tau_{vH}-\tau_{vL})+4} - 10K\xi - 6}{3\left(\sqrt{2\xi(4K^2\xi+6K+2\tau_{vH}-\tau_{vL})+4} + \sqrt{2\xi(4K^2\xi+6K-\tau_{vH}+2\tau_{vL})+4} - 4\right) - 20K\xi}. \tag{B.18}
$$

We tabulate the sign of the utility difference in equation (B.15) for $\alpha = 0$ and $\alpha = 1$ to establish the existence of pure and mixed strategy equilibria.

Learning capacity	$\Delta \equiv U(H) - U(L)$		Equilibrium
	$\alpha = 0$	$\alpha = 1$	
$K \leq \underline{K}$	$-$	$+$	Two equilibria: Cover L **or** H
$K \in (\underline{K}, \overline{K}]$	$-$	$-$	Cover L is dominant strategy.
$K > \overline{K}$	$+$	$-$	Mixed strategy

For $K \leq \underline{K}$, there are two pure strategy equilibria. From equation (B.15), if two analysts choose $\alpha = 0$, then $U(H) - U(L) < 0$ and therefore the third analyst also chooses $\alpha = 0$. Conversely, if two analysts choose $\alpha = 1$, then $U(H) - U(L) > 0$ and therefore the third analyst also chooses $\alpha = 1$. The mixed strategy equilibrium $(\alpha^\star, 1 - \alpha^\star)$ is not stable. To see that, we perturb the beliefs of one analyst to $\alpha + \epsilon$, with $\epsilon > 0$ arbitrarily small (i.e., a small bias towards covering H). In this case,

$$U(H) - U(L) > 0, \tag{B.19}$$

and the analyst is better off choosing H with probability one.

Since $\tau_H > 2\tau_L$, it follows that $\underline{K} < 0$ and therefore the first row in the table above does not correspond to a valid parameter region. Consequently, analysts cover L with probability one for $K \leq \overline{K}$. For $K > \overline{K}$, a mixed strategy emerges in equilibrium where analysts choose H with probability $\alpha^\star$ in equation (B.18).

Finally, we can write $\alpha^\star$ as

$$\alpha^\star = \frac{\beta_0(K)}{2\beta_0(K) + \beta_1(K)}, \tag{B.20}$$

where

$$\beta_0(K) = 10K\xi + 6 - 3\sqrt{2\xi\left(4K^2\xi + 6K + 2\tau_{vH} - \tau_{vL}\right) + 4} \text{ and} \tag{B.21}$$

$$\beta_1(K) = 3\left(\sqrt{2\xi\left(4K^2\xi + 6K + 2\tau_{vH} - \tau_{vL}\right) + 4} - \sqrt{2\xi\left(4K^2\xi + 6K - \tau_{vH} + 2\tau_{vL}\right) + 4}\right) > 0. \tag{B.22}$$

Since $\beta_0 > 0$ for any $K > \overline{K}$, it follows that $\alpha^\star < \frac{1}{2}$. □

Proposition 3

Proof. We proceed as in the proof of Propostion 2 and solve for the optimal coverage choices for different regions of K, assuming $\frac{\tau_H}{\tau_L} < 2$.

Region 1: $K \leq \dfrac{\sqrt{1 + 4\xi\left(2\tau_{vL} - \tau_{vH}\right)} - 1}{4\xi}$. From Corollary 1, investors acquire a signal about H alone if $\eta = (2, 1)$ and about L alone if $\eta = (1, 2)$. From equations (21), (26), (27), and Corollary 1, the expected analyst utilities from covering H and L are, respectively:

$$U\left(H\right)=\alpha^{2}\frac{K}{3}+2\left(1-\alpha\right)\alpha\frac{K}{2}+\left(1-\alpha\right)^{2}\times 0 \text{ and}$$
$$U\left(L\right)=\alpha^{2}\times 0+2\alpha\left(1-\alpha\right)\frac{K}{2}+\left(1-\alpha\right)^{2}\frac{K}{3}. \tag{B.23}$$

The difference between the utility functions is

$$U\left(L\right)-U\left(H\right)=\frac{K}{3}\left(1-2\alpha\right). \tag{B.24}$$

Two pure strategy equilibria emerge. If $\alpha=0$ for two analysts, $U\left(L\right)-U\left(H\right)>0$ and it is also optimal for the third analyst to cover L and select $\alpha=0$. Conversely, if $\alpha=1$ for two analysts, the third analyst also optimally covers H and selects $\alpha=1$. The mixed strategy equilibrium with $\alpha=\frac{1}{2}$ is not stable. To see that, we perturb the beliefs of one analyst to $\alpha+\epsilon$, with $\epsilon>0$ arbitrarily small (i.e., a small bias towards covering H). In this case,

$$U\left(L\right)-U\left(H\right)=U\left(L\right)-U\left(H\right)=\frac{K}{3}\left(1-2\alpha-2\epsilon\right)<0, \tag{B.25}$$

and therefore the third analyst chooses H with probability one.

Region 2: $K\in\left(\frac{\sqrt{1+4\xi\left(2\tau_{vL}-\tau_{vH}\right)}-1}{4\xi},\frac{\sqrt{1+4\xi\left(2\tau_{vH}-\tau_{vL}\right)}-1}{4\xi}\right]$**.** rom Corollary 1, if $\eta=(2,1)$ investors learn about both securities and if $\eta=(1,2)$ investors only learn about L. From equations (21), (26), (27), and Corollary 1, the expected analyst utilities from covering H and L are, respectively:

$$U\left(H\right)=\alpha^{2}\frac{K}{3}+2\alpha\left(1-\alpha\right)\frac{1}{2}C_{H}\left(2,1\right)+\left(1-\alpha\right)^{2}\times 0 \text{ and}$$
$$U\left(L\right)=\alpha^{2}C_{L}\left(2,1\right)+2\alpha\left(1-\alpha\right)\frac{K}{2}+\left(1-\alpha\right)^{2}\frac{K}{3}. \tag{B.26}$$

The difference between the two utility functions is linear in α, that is

$$U\left(L\right)-U\left(H\right)=\frac{K}{3}+\alpha\left(\frac{K}{3}-C_{H}\left(2,1\right)\right). \tag{B.27}$$

We will prove that the two utility functions do not cross for any $\alpha\in[0,1]$. Indeed,

$$U\left(L,\alpha=0\right)-U\left(H,\alpha=0\right)=\frac{K}{3}>0 \text{ and}$$
$$U\left(L,\alpha=1\right)-U\left(H,\alpha=1\right)=\frac{2K}{3}-C_{H}\left(2,1\right). \tag{B.28}$$

We note that $U(L, \alpha = 1) - U(H, \alpha = 1) > 0$ for any $K > \underline{K}$, where κ_1 is defined in (B.16). Therefore, if $K > \underline{K}$, then the utility difference never switches sign. It follows that $U(L) - U(H) > 0$ for any $\alpha \in [0, 1]$ and analysts cover L with probability one.

If $K \leq \underline{K}$, there are two pure strategy equilibria. If two analysts choose $\alpha = 0$, then $U(L) - U(H) > 0$ and therefore the third analyst also chooses $\alpha = 0$. Otherwise, if two analysts choose $\alpha = 1$, then $U(L) - U(H) < 0$ and therefore the third analyst also chooses $\alpha = 1$. The mixed strategy equilibrium $\left(\alpha^{\dagger}, 1 - \alpha^{\dagger}\right)$ is not stable. To see that, we perturb the beliefs of one analyst to $\alpha + \epsilon$, with $\epsilon > 0$ arbitrarily small (i.e., a small bias towards covering H). In this case,

$$U(L) - U(H) = \frac{K}{3} + \alpha^{\dagger}\left(\frac{K}{3} - C_H(2, 1)\right) < 0, \tag{B.29}$$

and therefore the third analyst chooses H with probability one.

Region 3: $K > \dfrac{\sqrt{1 + 4\xi(2\tau_{vH} - \tau_{vL})} - 1}{4\xi}$. For large K, from Corollary 1, investors always acquire signals about a stock as long as there is at least one analyst covering it. The results are the same as in the proof of Proposition 2. In particular, there is a unique possible mixed strategy probability,

$$\alpha^{\star} = \frac{3\sqrt{2\xi(4K^2\xi + 6K + 2\tau_{vH} - \tau_{vL}) + 4} - 10K\xi - 6}{3\left(\sqrt{2\xi(4K^2\xi + 6K + 2\tau_{vH} - \tau_{vL}) + 4} + \sqrt{2\xi(4K^2\xi + 6K - \tau_{vH} + 2\tau_{vL}) + 4} - 4\right) - 20K\xi}. \tag{B.30}$$

For convenience, we tabulate again the sign of the utility difference in equation (B.15) for $\alpha = 0$ and $\alpha = 1$ to establish the existence of mixed strategy equilibria.

Learning capacity	$U(L) - U(H)$ $\alpha = 0$	$\alpha = 1$	Equilibrium
$K \leq \underline{K}$	$+$	$-$	Two equilibria: Cover L **or** H
$K \in \left(\underline{K}, \overline{K}\right]$	$+$	$+$	Cover L is dominant strategy.
$K > \overline{K}$	$-$	$+$	Mixed strategy

As in Proposition 2, the mixed strategy equilibrium for $K \leq \underline{K}$ is not stable. For $K > \overline{K}$, a mixed strategy always arises in equilibrium. We combine the table above with the equilibrium strategy for the second region to complete the proof. □

Corollary 2

Proof. We can write $\alpha^{\star}$ as

$$\alpha^{\star} = \frac{\beta_0(K)}{2\beta_0(K) + \beta_1(K)}, \tag{B.31}$$

where

$$\beta_0(K) = 10K\xi + 6 - 3\sqrt{2\xi\left(4K^2\xi + 6K + 2\tau_{vH} - \tau_{vL}\right) + 4} \text{ and} \tag{B.32}$$

$$\beta_1(K) = 3\left(\sqrt{2\xi\left(4K^2\xi + 6K + 2\tau_{vH} - \tau_{vL}\right) + 4} - \sqrt{2\xi\left(4K^2\xi + 6K - \tau_{vH} + 2\tau_{vL}\right) + 4}\right) > 0. \tag{B.33}$$

We take the first derivative of $\alpha^\star$ with respect to K:

$$\frac{\partial \alpha^\star}{\partial K} = \frac{\beta_1(K)\,\beta_0'(K) - \beta_1'(K)\,\beta_0(K)}{\left(2\beta_0(K) + \beta_1(K)\right)^2}, \tag{B.34}$$

and we show that $\beta_1(K)\,\beta_0'(K) - \beta_1'(K)\,\beta_0(K) > 0$.

First, note that

$$\beta_1'(K) = \frac{\xi(8K\xi + 6)}{\sqrt{2\xi\left(4K^2\xi + 6K + 2\tau_{vH} - \tau_{vL}\right) + 4}} - \frac{\xi(8K\xi + 6)}{\sqrt{2\xi\left(4K^2\xi + 6K - \tau_{vH} + 2\tau_{vL}\right) + 4}} < 0. \tag{B.35}$$

To show that $\beta_0(K)$ is increasing we note that $\beta_0'(K)$ has no real roots and that

$$\beta_0'(0) = \xi\left(10 - \frac{9}{\sqrt{\xi\tau_{vH} - \frac{\xi\tau_{vL}}{2} + 1}}\right) > 0. \tag{B.36}$$

Further, since $\beta_0(\kappa_0) > 0$, it follows that $\beta_1(K)\,\beta_0'(K) - \beta_1'(K)\,\beta_0(K) > 0$ and therefore $\alpha^\star$ increases in K for $K > \overline{K}$.

Finally, since $\lim_{K\to\infty} \beta_1(K) = 0$, it immediately follows that $\lim_{K\to\infty} \alpha^\star = \frac{1}{2}$. □

Corollary 3

Proof. To show that $\alpha = \frac{\beta_0}{2\beta_0 + \beta_1}$ increases in ξ, it is enough to show that $\frac{\beta_1}{\beta_0}$ decreases in ξ. It can be shown that

$$\frac{\partial\left(\beta_1\beta_0^{-1}\right)}{\partial \xi} = c_0^2 \times \left(c_1 + c_2\right), \tag{B.37}$$

where

$$c_1 \equiv 3\left(-3\sqrt{2\xi\left(4K^2\xi + 6K + 2\tau_{vH} - \tau_{vL}\right) + 4} + 10K\xi + 6\right) \times \\ \times \left(\frac{8K^2\xi + 6K + 2\tau_{vH} - \tau_{vL}}{\sqrt{2\xi\left(4K^2\xi + 6K + 2\tau_{vH} - \tau_{vL}\right) + 4}} + \frac{-2K(4K\xi + 3) + \tau_{vH} - 2\tau_{vL}}{\sqrt{2\xi\left(4K^2\xi + 6K - \tau_{vH} + 2\tau_{vL}\right) + 4}}\right) \tag{B.38}$$

and

$$c_2 \equiv -3\left(10K - \frac{3\left(8K^2\xi + 6K + 2\tau_{vH} - \tau_{vL}\right)}{\sqrt{2\xi\left(4K^2\xi + 6K + 2\tau_{vH} - \tau_{vL}\right) + 4}}\right) \times$$
$$\times\left(\sqrt{2\xi\left(4K^2\xi + 6K + 2\tau_{vH} - \tau_{vL}\right) + 4} - \sqrt{2\xi\left(4K^2\xi + 6K - \tau_{vH} + 2\tau_{vL}\right) + 4}\right). \tag{B.39}$$

It remains to be shown that $c_1 + c_2 < 0$. This is indeed the case, since

$$\lim_{\xi \to 0} c_1 + c_2 = 0 \tag{B.40}$$

and

$$\frac{\partial\left(c_1 + c_2\right)}{\partial \xi} = -\frac{K\xi(6K + 10\tau_{vH} - 5\tau_{vL}) + 2K - 6\tau_{vH} + 3\tau_{vL}}{\sqrt{2K^2\xi^2 + 3K\xi + \xi\tau_{vH} - \frac{\xi\tau_{vL}}{2} + 1}} < 0 \tag{B.41}$$

for $K > \bar{K}$.

□

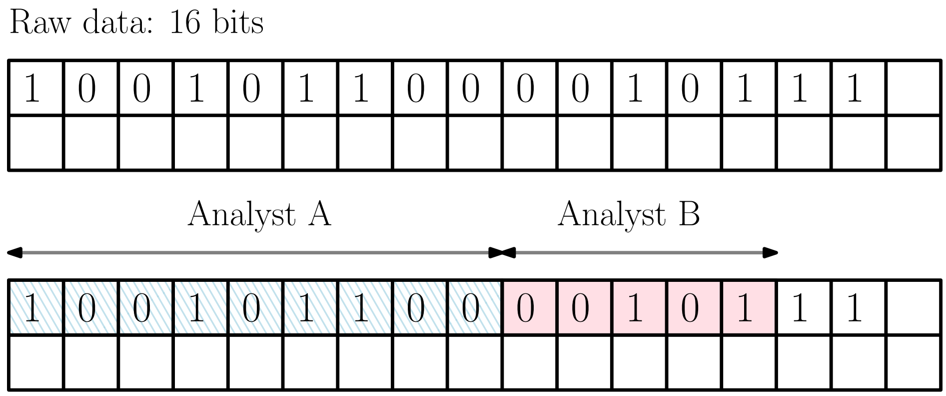
C Analyst coverage as information compression

Equations (4) and (5) imply that if more analysts cover a stock, investors spend less attention capacity to acquire a signal with given precision. Further, analyst coverage has decreasing returns to scale. In this Section, we motivate the assumption by arguing analyst coverage "compresses" information. For an investor, following analyst coverage instead of searching for and interpreting raw financial data frees up cognitive capacity – just like archiving computer files frees up hard-drive space.

In the spirit of Farboodi, Matray, and Veldkamp (2018), assume we can partition cognitive capacity into *bits*. Each bit can hold a single number (not necessarily zero or one), that is exactly one piece of information. Figure C.1 illustrates a situation where investors want to acquire an informative signal about the asset from a 16-bit "raw" data set. The size of the partitioned space (i.e., of $2 \times 17 = 34$ bits in the Figures below) corresponds to the attention capacity K in the model.

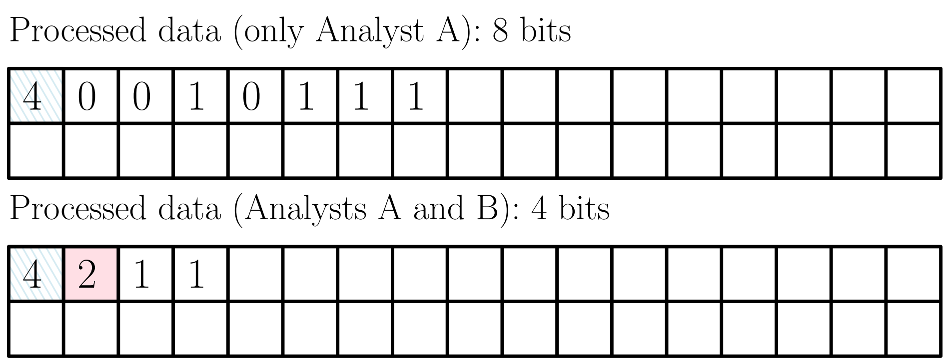
Each zero or one bit corresponds to one piece of unprocessed information: for example, one line in the balance sheet or profit-and-loss statement, one news headline, the earnings announcement, the last quarter earnings' consensus, or the market share of the company. For simplicity, assume that the investor only requires the sum of all bit values as a sufficient statistic to make a trading decision. To compute the sum, however, the investor needs to reserve 16 bits worth of cognitive capacity.

Figure C.1: **Raw signal analysis**



Analyst coverage helps the investor, in the context of this simple illustration, by providing "partial sums" of bits. For example, Analyst A focuses on interpreting the income statement (say, the first 9 out of the 16 bits). From Figure C.2, the investor only needs to reserve 8 bits of cognitive capacity: one for Analyst A's report and seven for the remaining information units not covered. If, additionally, Analyst B summarizes the following 5 bits (for example, relevant news items in the past quarter), then the investor can compress the 16-bit raw information into 4-bits of cognitive space: one each for the two analysts' reports, and 2 bits for information that is not covered by analyst A or B.

Figure C.2: **Processed signal analysis**



The example is chosen to be consistent with decreasing returns to scale from analyst coverage. Analyst A's helps the investor to reduce cognitive space consumption, for the same raw signal, by 8 bits, whereas Analyst B's marginal contribution to investors is of only 4 bits. Note that it doesn't need to be the case that Analyst B is less productive than Analyst A: it can be the case that A provides information about bits 1 through 9, and B provides information about bits 6 through 14. Since part of the information content overlaps, the *marginal* contribution of Analyst B is to provide information about bits 10 through 14.

In particular, if the marginal contribution of the n^{th} analyst is of summarizing $\dfrac{\Gamma}{n\,(n+1)}+1$ bits, where Γ is the bit length of the raw signal, then the processed signal length becomes $\dfrac{\Gamma}{n+1}$, in line with equation (5) which assumes analyst coverage improves information processing at rate n^{-1}.[11]

Finally, we note that even the marginal contribution of the two analysts in the illustration above is different, they both receive an equal share (i.e., one bit) of the investors' attention capacity, which is consistent with our modelling choice.

[11] In total, the first n analysts summarize $n+\dfrac{n}{n+1}\Gamma$ bits (by expanding a telescopic sum) and their forecasts take up n bits. Investors still need to parse $\Gamma-\left(n+\dfrac{n}{n+1}\Gamma\right)=\dfrac{1}{n+1}\Gamma-n$ "uncovered" bits of information, plus the n bits containing the analysts forecasts: in total, $\dfrac{1}{n+1}\Gamma$ bits.