Crowded stock coverage

Charles Martineau and Marius Zoican*

University of Toronto

April 18, 2019

Abstract

Analyst stock coverage is "crowded:" In 2017, 5% of U.S. equities amounted to 25% of earnings forecasts. Is information supply optimally distributed in financial markets? We build a model where limited-attention investors endogenously learn about securities. Analysts compete for scarce investor attention, providing forecasts that reduce learning costs. Coverage crowding emerges through strategic complementarity effects. For limited investor attention, analysts prefer to share a crowded space rather than "going against the wind" to cover more opaque assets. However, coverage skewness is excessive from the investors' perspective. The implications echo documented patterns: analysts cluster in large stocks with significant intangible assets.

Keywords: analyst coverage, investor attention, learning, information supply, asset pricing
JEL Codes: G11, G14, G23, D83

*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.

Crowded stock coverage

Abstract

Analyst stock coverage is "crowded:" In 2017, 5% of U.S. equities amounted to 25% of earnings forecasts. Is information supply optimally distributed in financial markets? We build a model where limited-attention investors endogenously learn about securities. Analysts compete for scarce investor attention, providing forecasts that reduce learning costs. Coverage crowding emerges through strategic complementarity effects. For limited investor attention, analysts prefer to share a crowded space rather than "going against the wind" to cover more opaque assets. However, coverage skewness is excessive from the investors' perspective. The implications echo documented patterns: analysts cluster in large stocks with significant intangible assets.

Keywords: analyst coverage, investor attention, information processing, learning, asset pricing
JEL Codes: G11, G14, G23, D83

1 Introduction

Is the supply of information optimally distributed in the cross-section of securities? In 2017, the top 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). From Groysberg, Healy, and Maber (2011), analysts aim to maximize investor attention. Therefore, by focusing on a small subset of stocks, analysts simply cater to their customers. Some level of forecast clustering is therefore to be expected.

However, excessively "crowded" analyst coverage indicates a misallocation of resources. First, information supply is costly. A limited amount of computer and brain power is directed towards crunching unstructured data (e.g., Tweets), financial statement data, as well as interpreting news releases for a large number of public firms. Essentially, analysts produce information by converting low- to high-precision signals (Dugast and Foucault, 2018). 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?

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: large stocks are more likely to attract more analysts, widening the coverage gap.

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 for assets with highly uncertain payoffs, analysts' forecasts cluster too much in equilibrium relative to the first best, and not necessarily on the assets where the marginal benefit of coverage is largest.

We build a portfolio optimization model with endogenous information acquisition and endogenous information supply. Before trading, 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.

In our model, 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 a stock has better analyst coverage, investors can spend less attention to obtain a signal of a given precision. There are decreasing returns to scale from analyst coverage: the more analysts cover a stock, the less their marginal impact on learning costs. Information supply is endogenous: analysts compete over investor attention. Each analyst chooses one stock to cover with the objective to maximize the share of attention he receives.

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 investor to acquire signals about the respective security. Therefore, she allocates a larger share of her limited capacity to learn about assets with more analyst forecasts. Therefore, analysts have an incentive to cluster their forecasts and maximize investor attention to the respective security. On the other hand, going "against the wind" and issuing forecasts about a less well-covered asset gives an analyst monopoly power over investors' attention directed towards that asset.

What drives the analysts' choice? It turns out that the key friction driving the distribution information supply is the investors' attention capacity. If attention is 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 a "star" stock, and none cover a "ordinary" stock. The next analyst to make a coverage choice is better off receiving one tenth of the attention directed towards the star stock than becoming the only source of information about the other security. Investors are so limited in their attention that they will learn little, if anything, about the "ordinary" 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. 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. Further, common shocks to asset risk due to for example, systematic risk or market-wide implementation of disclosure standards, enhance analyst crowding. If risk across stocks is more correlated, investors have a weaker preference to learn about any particular stock. Therefore, they would allocate most attention to the stocks with the highest coverage, which are cheaper to learn about. Analysts who differentiate from the crowd receive less attention, which reinforces strategic complementarity effects and leads to more coverage clustering.

The model yields a number of empirical predictions. Analyst coverage tends to cluster in large capitalization, high trading volume securities with better price discovery. Further, due to intensive coverage, crowded securities are conditionally safer than securities with little coverage. The model implications match empirical patterns in Section 2, suggesting that analyst forecasts cluster in large, blue-chip index stocks.

Finally, crowded coverage is not necessarily benign. We find that 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.

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 stock than others. Second, and most important, we introduce an endogenous information supply through analyst forecasts. Investors' learning costs depend on analyst coverage, which in turn depends on the investor attention they can attract.

Our model implications echo the findings of Farboodi, Matray, and Veldkamp (2018). Their paper documents 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. The authors argue that the result is driven by a size effect: investors are better off learning about larger firms, as they represent a more relevant risk factor. Our paper generates a very similar prediction, i.e., a negative relationship between price informativeness and market capitalization. However, we obtain the result through an information supply channel, where both price discovery and size are co-determined in equilibrium.

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. Zhu (2019) finds that the availability of unstructured data reduces information acquisition costs for investors, leading to improved price discovery and managers' investment decisions. Veldkamp (2006) argues that there are strategic complementarities in investors' demand for information. Since acquiring information has high fixed costs and low marginal costs, 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 endogenously determine the elasticity of information supply.

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.

Investor attention matters for asset prices. Da, Engelberg, and Gao (2011) finds that investor attention, as measured by Google Search Volume, leads to a short-term increase in stock prices and a long-term reversal. Benamar, Foucault, and Vega (2019) measure attention by "short URL" clicks and document that high information demand leads to a higher price response to news. In line with our results on investor preferences, the authors conclude that investors are more likely to collect information when uncertainty about asset payoffs is higher. Studying macroeconomic announcements, Fisher, Martineau, and Sheng (2019) find that bad news draw more attention than good news; also, changes in investor attention are able predict announcement surprises and stock returns. Ben-Rephael, Da, and Israelsen (2017) document that the speed of price discovery increases in the attention by institutional investors, as measured by the magnitude of the post earnings drift. Andrei and Hasler (2015) argue that as investors attention increases, they can incorporate more fundamental information into prices: as a result, return volatility increases. In a dynamic model, Andrei and Hasler (2019) show that investors allocate more attention to an asset the further away variables that predict its return are from their long term mean. Investor attention seems to focus on analyst coverage: Lawrence, Ryans, and Sun (2017) use Yahoo! Finance data to show that investors' demand for analyst information significantly exceeds the demand for SEC filings or financial statement data.

We also contribute to a literature studying analysts' coverage choices and the impact of analyst coverage. In line with our model implications, Barth, Kasznik, and McNichols (2001) find that analyst coverage clusters in stocks with more intangible assets, which arguably have more fundamental payoff uncertainty. McNichols and O'Brien (1997) find that analysts are more likely to cover stocks with encouraging future prospects. Groysberg, Healy, and Maber (2011) find that analyst compensation is driven primarily by reputation ("All Star" status, recognition by *The Wall Street Journal*) and less so by forecast accuracy. Indeed, analyst compensation is designed to reward actions that increase brokerage and customer engagement. 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 may go 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. 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 relevant empirical patterns pertaining to the cross-sectional distribution of analyst coverage over more than three decades. The empirical regularities we discuss are a natural starting point for the model introduced in Section 3.

2.1 Data

We retrieve the analyst coverage 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.

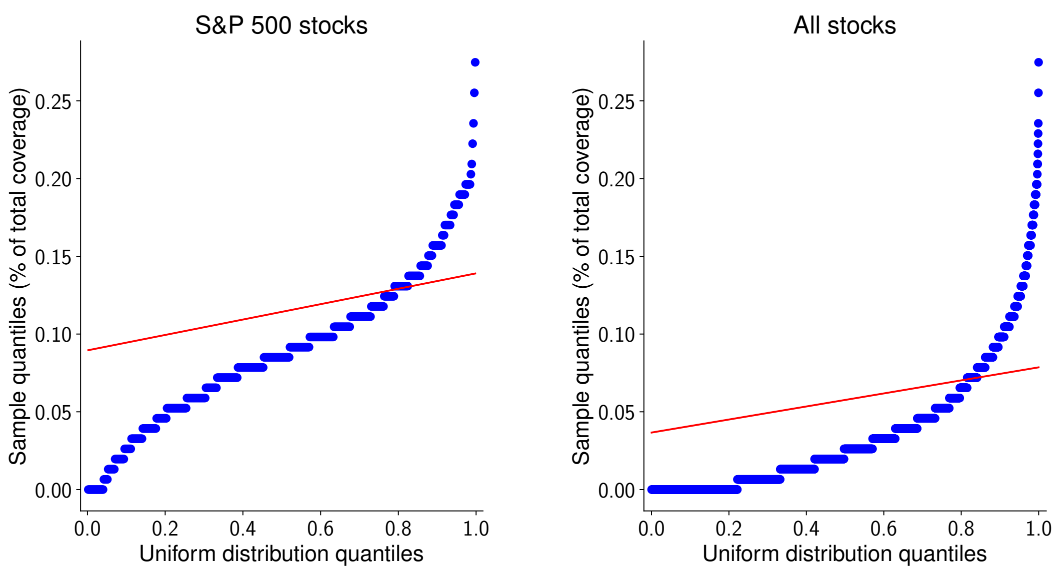
The sample period is from January 1, 1984 to December 31, 2017 for a total of 15,661 unique firms and 550,361 earnings announcements.

2.2 Analyst coverage patterns in the United States

Figure 1 illustrates the distribution of stock analyst coverage for 2017, the final year in our sample. Each observation corresponds to an unique stock. For each stock, coverage is measured as the number of analyst forecasts across the year. We are benchmarking the sample coverage distribution against an uniform distribution, corresponding to a scenario where each stock receives an equal number of forecasts. The salient feature of Figure 1 is a fat right tail, indicating significant coverage clustering. The top 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 condition on 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.

Figure 1: **Analyst coverage clustering**

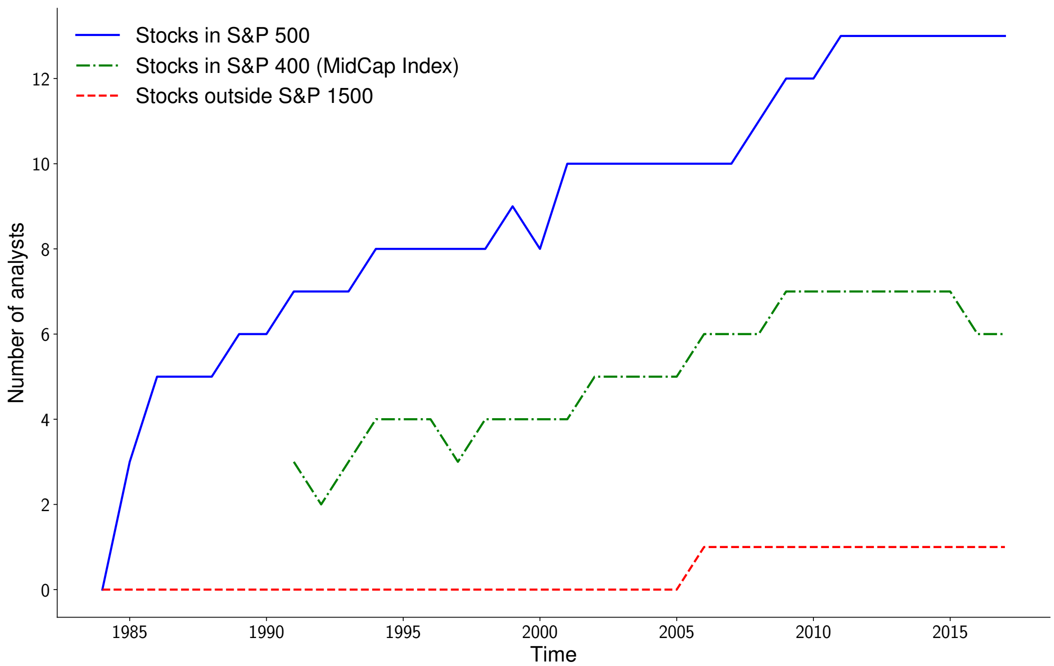
This figure is a quantile-quantile plot of the number of earnings analyst forecasts from I/B/E/S for the year 2017, aggregated for all four quarters. For each quarter, we take into account the last forecast update for a stock-analyst pair. Each observation corresponds to a single stock. We benchmark the earning forecast distribution against the uniform distribution.



Analyst coverage is crowded particularly in large stocks. From Figure 2, the median stock 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.

Figure 2: **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.



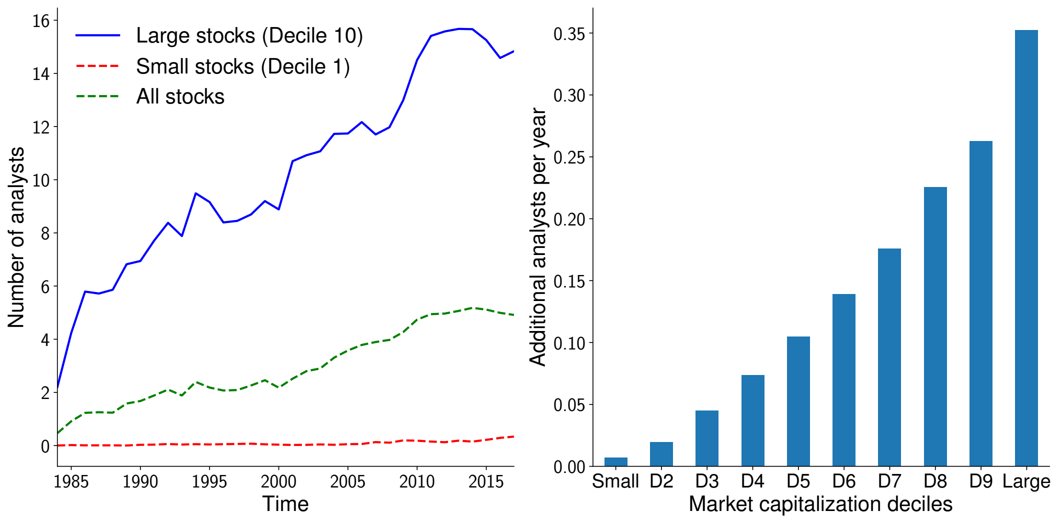
Coverage clustering persists over time. We sort equities in our sample into ten deciles by market capitalization for each year in the sample. Figure 3 illustrates the year-to-year analyst coverage growth rate for each size decile. Large-cap stocks are not only better covered on average, they are also more likely on the margin to receive additional coverage. From 1984 to 2017, a large firm attracts 0.35 new analysts each year, compared to 0.14 new analysts for the average stock, and essentially zero coverage growth for the smallest securities in the sample. The result suggests either lower costs for analysts to cover large stocks (for example, due to higher quality financial statements) or strategic complementarity effects that give analysts incentives to cover the same stocks as their peers. In Section 3, we focus on the second channel and build a model that explains coverage clustering, even if the information production cost is constant across securities.

Figure 3: **Analyst coverage and market capitalization**

This figure plots, in the left panel, the average number of analysts issuing an earnings forecast for the top and bottom 10% stocks in the sample sorted by market capitalization as well as for the average stock in the sample, for each year between 1984 and 2017. In the right panel, we plot the year-to-year coverage growth rate for each market cap decile, that is the trend coefficient θ in

$$\text{Coverage}_{d,t} = \text{Average coverage}_{d} + \theta \times t + \text{error},$$

where d indexes market cap deciles and t runs over years. Analyst coverage and market capitalisation data are provided by I/B/E/S and Compustat, respectively.

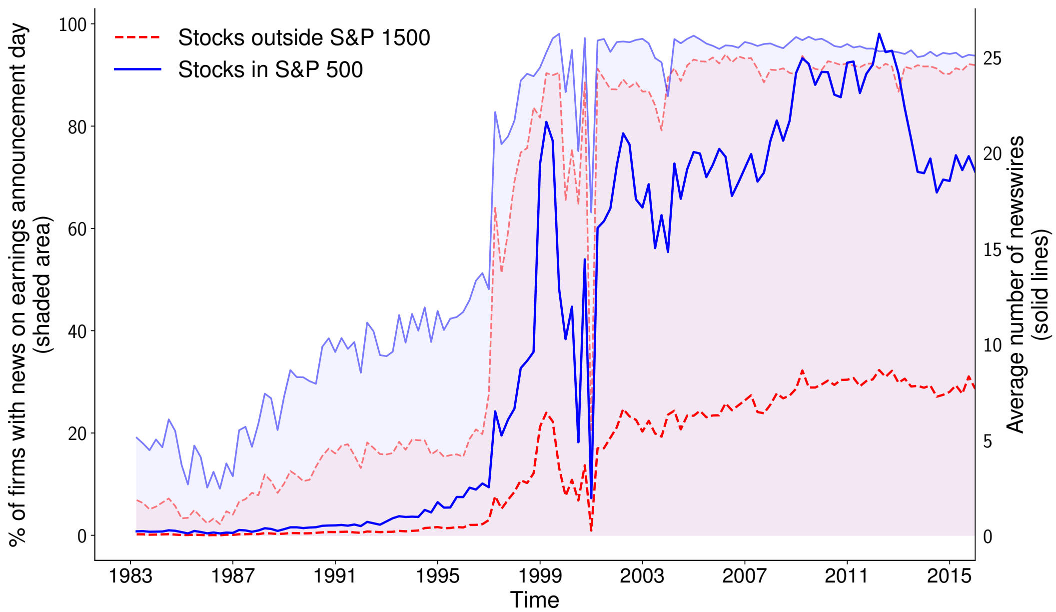


While Figure 3 suggests that coverage crowding increases over time, the probability that small-cap stocks receive any coverage at all increases over time. As in Martineau (2019), we use the Dow Jones newswire dataset to obtain the average number of newswire articles and the fraction of stocks with at least one newswire on the announcement day. Figure 4 displays the evolution of newswire coverage over time. On average, both S&P 500 stocks and stocks outside the S&P 1500 receive more news coverage. At the end of 1996, there are on average 2.73 (0.81) news articles on earnings announcement day, compared to 19.01 (7.67) at the end of 2015 for S&P 500 stocks (respectively for stocks outside the S&P 1500).

The breadth of stock newswire coverage increases over time across the entire stock universe. In 1983Q1, more than 90% of stocks outside S&P 1500 and 75% of S&P 500 stocks had no news coverage on earnings announcement day. At the end of 2015, more than 90% of stocks in both categories receive at least one newswire article on earnings announcement day.

Figure 4: **News on earnings announcement days**

This figure illustrates the average number of newswire articles (solid lines) and the fraction of earnings announcements (shaded areas) with at least one newswire article on earnings announcement day between 1983Q1 and 2015Q4, for S&P 500 stocks and stocks outside S&P 1500, respectively.



3 Model

Assets. Consider a four-period economy, where time is indexed by $t \in \{-1, 0, 1, 2\}$. There are two risky assets available in unit supply, labeled H and L. Each asset $i \in \{H, L\}$ pays off a stochastic dividend $v_i \sim \mathcal{N}\left(\mu_i, \tau_i^{-1}\right)$ at $t = 2$, where dividends are independently distributed with mean μ_i and precision τ_i. Without loss of generality, we assume that $\tau_H > \tau_L$: 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.

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

$$U_{\mathbf{I}} = -\mathbb{E}\left[\exp\left(-\rho \tilde{W}_2\right)\right]. \tag{1}$$

At $t = 1$ each investor chooses portfolio weights $\mathbf{w} = (w_H, w_L, 1 - w_H - w_L)$ for the high- and low-precision risky assets, respectively for the risk-free asset. Investors are endowed with one unit of wealth at $t = 0$, that is $W_0 = 1$. Finally, investors are atomistic and cannot individually impact prices: at $t = 1$, investors take as given the equilibrium asset prices and the risk-free rate.

Learning technology. Investors can acquire unbiased signals $s_i = v_i + \epsilon_i$ about each stock i, where $\epsilon_i \sim \mathcal{N}\left(0, (\hat{\tau}_i - \tau_i)^{-1}\right)$ at $t = 0$, before choosing portfolio shares. It follows that learning about asset i allows the investor to increase the stochastic payoff precision from τ_i to $\hat{\tau}_i$, where

$$\hat{\tau}_i = \text{var}^{-1}\left(v_i \mid \mathcal{F}_{t=1}\right) \geq \text{var}^{-1}\left(v_i \mid \mathcal{F}_{t=0}\right) = \tau_i, \tag{2}$$

and $\mathcal{F}_t$ is the investor information set at time t.

Investors endogenously choose the signal precision, and consequently the posterior precision $\hat{\tau}_i$, but they have limited attention and therefore cannot completely eliminate asset volatility. We follow Van Nieuwerburgh and Veldkamp (2010) in modelling the learning technology. In particular, the investor **I** is endowed with attention (learning) capacity $K > 0$ which she can allocate across assets, subject to the capacity constraint at $t = 0$:

$$\sum_{i \in \{H,L\}} \kappa_i \left(\hat{\tau}_i - \tau_i\right) \leq K. \tag{3}$$

Some stocks are relatively easier to learn about than others: To capture such heterogeneity, we introduce a stock-specific learning cost parameter $\kappa_i > 0$ in the capacity constraint (3). A lower κ_i, all else equal, allows the investor to allocate less attention to stock i for a given precision gain $\hat{\tau}_i - \tau_i$. In Van Nieuwerburgh and Veldkamp (2010), it is equally easy to learn about any asset, that is $\kappa_i = 1$, 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 $i \in \{H, L\}$. 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:

$$\kappa\left(n_i\right) = \frac{1}{n_i}. \tag{4}$$

If more analysts cover a stock, investors learn more easily about its future payoffs. However, since $\dfrac{\partial^2 \kappa}{\partial n^2} > 0$, 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 only learn $t = 1$ through analyst reports: if no analysts cover a stock ($n_i = 0$), the cost of learning is arbitrarily large.

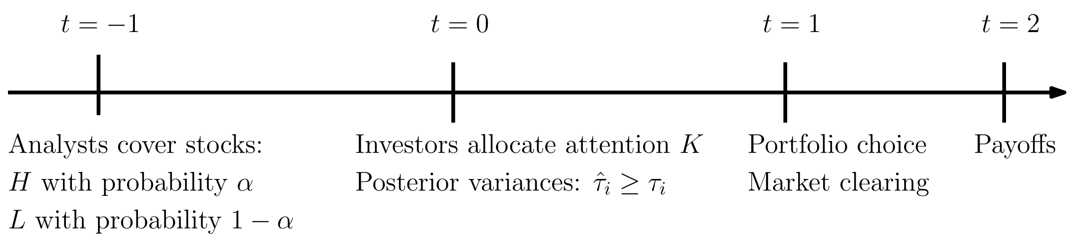
Each analyst j chooses to cover a single stock $i_j^\star$ to maximize the expected share of investor attention they capture. In equilibrium, investors allocate capacity $C_i \equiv \kappa_i^\star \left(\hat{\tau}_{i,t}^\star - \tau_i\right)$ to stock i, which is shared by the $n_i^\star$ analysts covering the stock. Formally, the analyst's problem can be written as

$$i_j^\star = \arg\max_i \mathbb{E}\left[\frac{1}{n_i^\star}\underbrace{\kappa_i^\star\left(\hat{\tau}_{i,t}^\star - \tau_i\right)}_{C_i}\right], \tag{5}$$

where $\sum_i C_i \leq K$.

Timing. Figure 5 summarizes the sequence of events at each time $t \in \{-1, 0, 1, 2\}$.

Figure 5: **Model timing**



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 $t = -1$: 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 ex-post precision $\hat{\tau}_i^\star$ for each stock i, (iii) optimal portfolio weights, $\mathbf{w}^\star = (w_H^\star, w_L^\star, 1 - w_H^\star - w_L^\star)$ at $t = 1$, and (iv) market clearing prices $\mathbf{p} = (p_H, p_L)$ at $t = 1$, 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 optimal investor portfolio and market clearing prices at $t = 1$, taking as given the posterior distribution (i.e., after learning) of the stochastic dividend. Next, Section 4.2 yields the optimal learning choice of the investor in each stock, for a given cross-sectional analyst coverage distribution. Finally, Section 4.3 solves for the analysts' equilibrium asset coverage strategy.

4.1 Optimal portfolio and market clearing

The investor is initially endowed with initial wealth W_0. She spends $p_i w_i$ to purchase asset i, and earns the risk-free return r on any remaining endowment. At $t = 2$, she receives the payoff v_i for each unit of asset i in her portfolio. The investor's wealth at the terminal date can be written as

$$W_2 = \left(W_0 - \sum_i p_i w_i \right) \times r + \sum_i w_i v_i. \tag{6}$$

Without loss of generality, we normalize the risk-free return to $r = 1$. From equations (1) and (6), the investor's portfolio choice problem at $t = 1$ becomes

$$\max_{\{w_i\}_i} -\mathbb{E}\left[\exp(-\rho r W_0 - \rho \sum_i w_i (v_i - p_i) \right]. \tag{7}$$

At the investment stage, since the investor learned about the asset at $t = 0$, the payoff for asset i is normally distributed with posterior mean $\hat{\mu}_i$ and posterior precision $\hat{\tau}_i^\star \geq \tau_i$. The posterior mean $\hat{\mu}_i$ adjusts in the direction of the signal s_i; the magnitude of the adjustment increases in the signal precision. Formally,

$$\hat{\mu}_i = \mu + \left(1 - \frac{\tau_i}{\hat{\tau}_i^\star} \right) (s_i - \mu). \tag{8}$$

Using the properties of the log-normal distribution, the equilibrium portfolio weights that maximize the investors' utility in (7) are given by

$$w_i^\star = \frac{\hat{\mu}_i - p_i}{\rho \left(\hat{\tau}_i^\star \right)^{-1}} \text{ for } i \in \{H, L\}. \tag{9}$$

The equilibrium price $p_i^\star$ sets equal the demand and supply for each stock. Since both stocks are available in unit supply, it follows that $p_i^\star$ solves the market clearing condition

$$\hat{\tau}_i^\star \frac{\hat{\mu}_i - p_i^\star}{\rho} = 1 \Longrightarrow p_i^\star = \hat{\mu}_i - \rho \left(\hat{\tau}_i^\star \right)^{-1}. \tag{10}$$

4.2 Optimal signal acquisition

The optimal portfolio in (9) and, consequently, equilibrium market prices in (10) critically depend on the posterior asset precision $\hat{\tau}_i^\star$. We impose a natural no-forgetting restriction as in Van Nieuwerburgh and Veldkamp (2010): the equilibrium posterior precision is weakly greater than the prior precision, $\hat{\tau}_i^\star > \tau_i$. That is, investors cannot "forget" prior information about one asset to relax the attention constraint and obtain in turn a more precise signal about the other asset. Equivalently, the variance of any signal s_i needs to be non-negative. Lemma 1 states the investors' learning objective function at $t = 0$, conditional on analyst coverage, capacity constraints, and no-forgetting restrictions.

Lemma 1. (Learning objective function.) *The representative investor's optimal learning problem corresponds to maximizing the product of posterior asset precisions, or equivalently, its natural logarithm* $\sum_i \ln(\hat{\tau}_i^\star)$.

From Lemma 1, the investors' Lagrangian problem at $t = 0$ is to choose posterior variances $\hat{\tau}_i$ that maximize

$$\mathcal{L} = \sum_i \ln(\hat{\tau}_i^\star) + \psi \left[K - \sum_i \frac{1}{n_i} (\hat{\tau}_i - \tau_i) \right] + \sum_i \lambda_i (\hat{\tau}_i - \tau_i), \tag{11}$$

where ψ is the Lagrange multiplier of the attention capacity constraint and λ_i are Lagrange multipliers of the no-forgetting condition for each stock i. The capacity constraint immediately follows from equations (3) and (4), that is

$$K - \sum_i \frac{1}{n_i} (\hat{\tau}_i - \tau_i) \geq 0. \tag{12}$$

In equilibrium, investors might learn about zero, one, or both stocks.[1] 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:

$$\mathcal{M} \equiv \{ i \in \{H, L\} \mid \hat{\tau}_i^\star - \tau_i > 0 \} \tag{13}$$

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 posterior asset value precision, conditional on the number of analysts covering each stock, n_i.

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

$$\hat{\tau}_i^\star = n_i \left(\frac{K}{m} + \frac{1}{m} \sum_{s \in \mathcal{M}} \frac{\tau_s}{n_s} \right). \tag{14}$$

For each stock $i \in \{H, L\} \setminus \mathcal{M}$, for which the investor does not acquire a signal, the equilibrium posterior precision is equal to the prior precision, $\hat{\tau}_i^\star = \tau_i$.

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

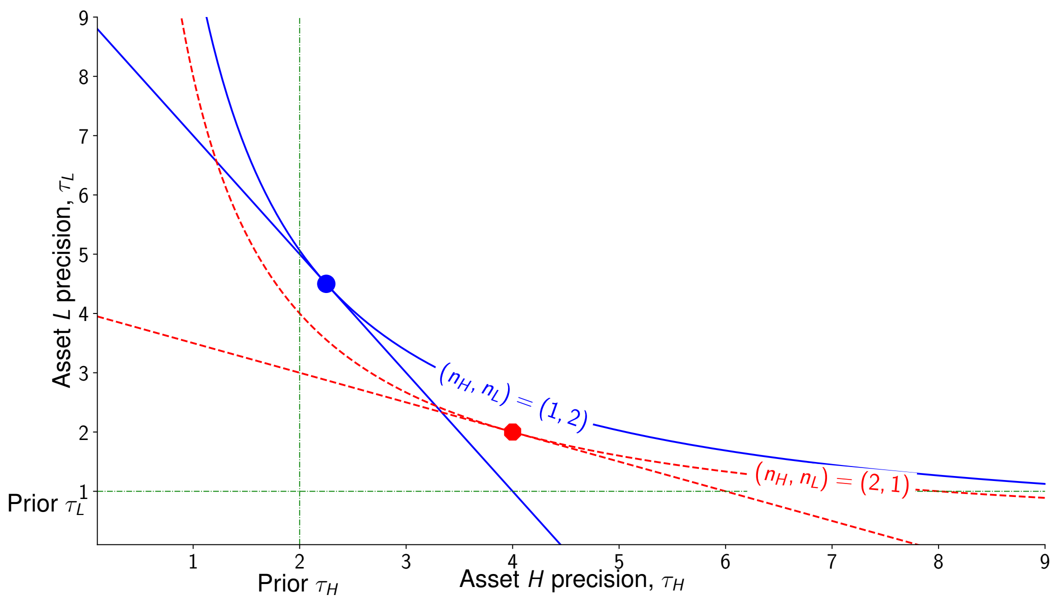
First, from Proposition 1, a higher attention capacity K yields a higher posterior precision $\hat{\tau}^{\star}$ for all assets in the investors' learning set $\mathcal{M}$. Second, investors learn about assets such that the posterior variances are proportional to the analyst coverage,

$$\frac{\hat{\tau}_i^{\star}}{\hat{\tau}_j^{\star}} = \frac{n_i}{n_j}, \forall i, j \in \mathcal{M}. \tag{15}$$

Finally, there are learning spillovers across stocks. If the prior precision τ_s 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 6 illustrates the investors' optimization problem for two different analyst coverage configurations.

Figure 6: **Optimal investor learning**

This figure illustrates the investors' signal acquisition problem at $t = 0$. We plot the attention capacity constraint in equation (3) 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: $\tau_H = 2$, $\tau_L = 1$, and $K = 2$.



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

$$C_i\left(n_i, n_{-i}\right) = \frac{1}{n_i}\left(\hat{\tau}_i^{\star} - \tau_i\right) = \max\left\{\frac{K}{m} + \left[\frac{1}{m}\sum_{s\in\mathcal{M}}\frac{\tau_s}{n_s} - \frac{\tau_i}{n_i}\right], 0\right\}. \tag{16}$$

A salient state variable in equation (16) is the analyst coverage-weighted prior asset precision, that is $\frac{\tau_i}{n_i}$. 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 $\frac{\tau_i}{n_i}$ is below average. That is, investors learn more about stocks that are ex-ante more uncertain (lower τ_i) or stocks that are better covered by analysts (higher n_i).

If the state variable $\frac{\tau_i}{n_i}$ is very large, either due to stock i having low ex-ante payoff uncertainty, low analyst coverage, or both, investors choose not to learn about it and allocate zero capacity. From equations (14) and (16), it follows that if $C_i = 0$, then the posterior and prior precisions for stock i are the same.

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, $\overline{K} > \underline{K} > 0$:*

$$\begin{aligned} \underline{K} &\equiv \max\left\{\frac{\tau_H}{2} - \tau_L, \tau_L - \frac{\tau_H}{2}\right\} \text{ and} \\ \overline{K} &\equiv \tau_H - \frac{\tau_L}{2}. \end{aligned} \tag{17}$$

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

Table 1: **Investors' learning set** $\mathcal{M}$

Analyst coverage	$K < \underline{K}$	$K \in (\underline{K}, \overline{K}]$	$K > \overline{K}$
$(n_H, n_L) = (3, 0)$		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 1) and two configurations where analysts cover both securities (second and third row in Table 1).

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 a 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 with relatively low $\frac{\tau}{n}$: well-covered (and consequently easy to learn about) stocks with 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, $K > \overline{K}$.

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 $\underline{K} < \overline{K}$. Even if attention capacity is limited such that learning about one stock is optimal, investors acquire signals about high precision stock H if

$$\frac{\tau_H}{n_H} \leq \frac{\tau_L}{n_L} \text{ or, equivalently, } \tau_H \leq 2\tau_L, \tag{18}$$

and acquire signals about L if $\tau_H > 2\tau_L$.

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

$$\begin{aligned} U_j(H) &= \sum_{n=0}^{2} \binom{2}{n} \alpha^n (1-\alpha)^{2-n} C_H(n+1, 2-n) \\ &= \alpha^2 \frac{1}{3} C_H(3,0) + 2\alpha(1-\alpha)\frac{1}{2} C_H(2,1) + (1-\alpha)^2 C_H(1,2). \end{aligned} \tag{19}$$

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 (19), that is,

$$
\begin{aligned}
U_j(L) &= \sum_{n=0}^{2} \binom{2}{n} \alpha^n (1-\alpha)^{2-n} C_H(n, 3-n) \\
&= \alpha^2 C_H(2,1) + 2\alpha(1-\alpha)\frac{1}{2} C_L(1,2) + (1-\alpha)^2 \frac{1}{3} C_L(0,3). \qquad (20)
\end{aligned}
$$

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

$$
U_j(H, \alpha^\star) = U_j(L, \alpha^\star) \text{ for all } j, \qquad (21)
$$

where $U_j(H, \alpha^\star)$ and $U_j(L, \alpha^\star)$ are defined in equations (19) and (20), respectively. If there is no $\alpha \in (0,1)$ that solves equation (21), 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 $K \leq 3\left(\tau_H - \frac{\tau_L}{2}\right)$, *all analysts cover the low-precision security* L.

(ii) For $K > 3\left(\tau_H - \frac{\tau_L}{2}\right)$, *analysts cover the high-precision security* H *with probability*

$$
\alpha^\star = \frac{1}{2} - \frac{9(\tau_H - \tau_L)}{8K - 6(\tau_H + \tau_L)} < \frac{1}{2}, \qquad (22)
$$

and the low-precision security L with probability $1-\alpha^{\star}$.

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 (16), 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 $K\leq 3\left(\tau_L-\frac{\tau_H}{2}\right)$, 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 $K\in\left(3\left(\tau_L-\frac{\tau_H}{2}\right),3\left(\tau_H-\frac{\tau_L}{2}\right)\right]$, all analysts cover the low-precision security L.

(iii) For $K>3\left(\tau_H-\frac{\tau_L}{2}\right)$, analysts cover the high-precision security H with probability

$$\alpha^{\star}=\frac{1}{2}-\frac{9\left(\tau_H-\tau_L\right)}{8K-6\left(\tau_H+\tau_L\right)}<\frac{1}{2}, \tag{23}$$

and the low-precision security L with probability $1-\alpha^{\star}$.

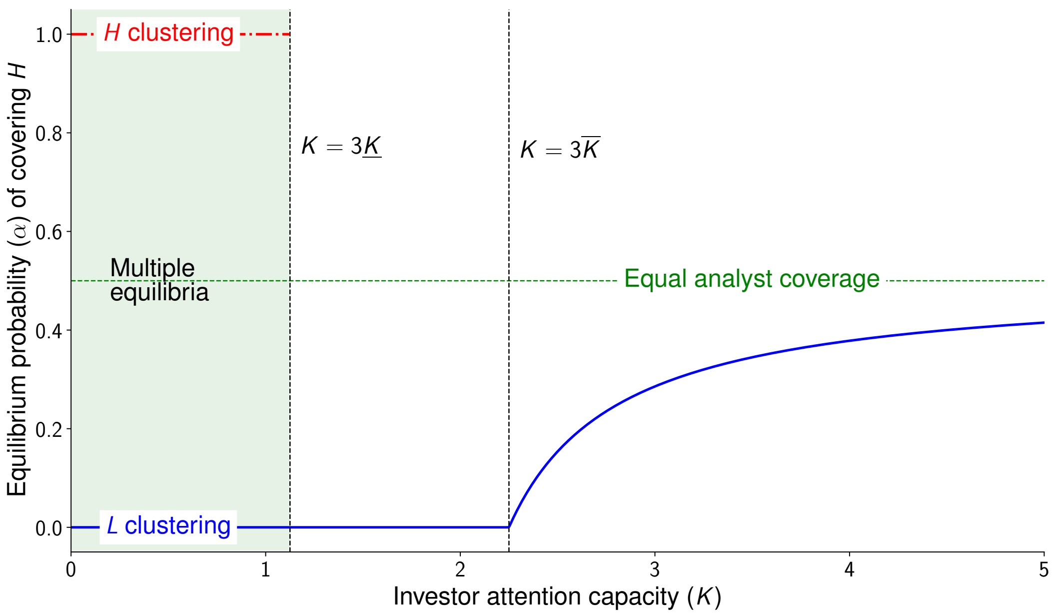
If assets have similar prior uncertainty (i.e., if $\frac{\tau_H}{\tau_L}\leq 2$), investors' preference to learn about L is weaker. In particular, if $\eta=(2,1)$ such that asset H has more coverage, investors allocate more capacity to H than to L. Therefore, especially for limited investor attention, 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, $K\leq 3\left(\tau_L-\frac{\tau_H}{2}\right)$, 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 the equilibrium strategies coincide with those in Proposition 2.

Figure 7 illustrates that the equilibrium probability of asset H coverage, α, weakly increases in investor capacity K, while never exceeding one half.

Figure 7: **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: $\tau_H = 2$ and $\tau_L = 1$.



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

$$\text{Clustering}\left(\alpha\right) = 1 - 4\alpha\left(1-\alpha\right), \tag{24}$$

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

$$\text{Skewness} = \frac{1-2\alpha}{\sqrt{\alpha\left(1-\alpha\right)}}, \tag{25}$$

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) *For an arbitrarily large investor capacity K, analysts are equally likely to cover either stock.*

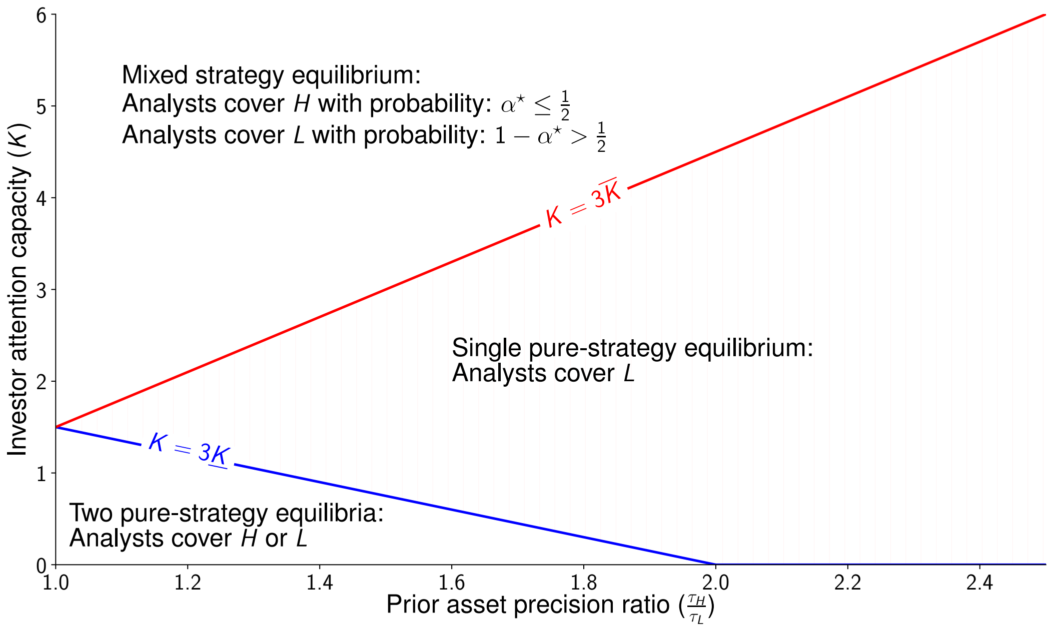
Corollary 2 emphasizes that analyst coverage clustering is a consequence of limited investor capacity. Indeed, from Propositions 2 and 3, if $K > 3\left(\tau_H - \frac{\tau_L}{2}\right)$ analysts choose either stock with positive probability $\alpha^\star$ and, moreover,

$$\lim_{K\to\infty} \alpha^\star = \lim_{K\to\infty} \frac{1}{2} - \frac{9\left(\tau_H - \tau_L\right)}{8K - 6\left(\tau_H + \tau_L\right)} = \frac{1}{2}. \tag{26}$$

As investor capacity K increases, $\alpha^\star$ is larger and therefore analyst coverage becomes less crowded. Figure 8 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 8: **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 $\frac{\tau_H}{\tau_L}$. Parameter values: $\tau_L = 1$.



5 Model implications

5.1 Analyst coverage and posterior uncertainty

Van Nieuwerburgh and Veldkamp (2010) show that if all assets are equally easy to learn about (equivalent to, in the context of our model, equal analyst coverage) then the equilibrium posterior variance should be constant for all stocks. Therefore, investors learn more in equilibrium about assets with ex ante higher variance. From Propositions 2 and 3, analysts tend to concentrate their coverage in such assets, which as a result become cheaper to learn about. Consequently, stocks with higher prior variance (i.e., in the absence analyst coverage) tend to have lower *posterior* variance, that is once they are covered by analysts.

Such a "learning overshooting" mechanism can explain why blue chip stocks, for example members of the S&P 500 index, benefit empirically from wider analyst coverage. If we simultaneously observe equilibrium returns and analyst coverage at $t = 1$, there is a negative correlation between implied risk and coverage. That is, analyst reports are biased towards *ex-post* safer stocks.

In a pure strategy equilibrium with $\alpha = 1$ (all analysts cover the low-precision asset L), it follows from equation (14) that

$$\hat{\tau}_L \geq \hat{\tau}_H \Longrightarrow 3K + \tau_L \geq \tau_H \Longrightarrow K \geq \frac{\tau_H - \tau_L}{3}, \tag{27}$$

and therefore the posterior precision of L is higher than the posterior precision of H for sufficiently large investor capacity K. Clearly, in a pure strategy equilibrium with $\alpha = 0$, both the prior and posterior precision of H are highest: investors learn about the ex ante safest asset, making it even safer.

We focus next on the mixed strategy equilibrium: If $K > 3\left(\tau_H - \frac{\tau_L}{2}\right)$, analysts cover stock H with probability $\alpha^\star \in \left(0, \dfrac{1}{2}\right)$. The posterior precision of asset H is:

$$\begin{aligned}\hat{\tau}_H &= (1-\alpha)^3 \times \tau_H + \alpha^3 \times (3K + \tau_H) + 3\alpha^2(1-\alpha) \times 2\left(\frac{K}{2} + \frac{1}{2}\left(\frac{\tau_H}{2} + \tau_L\right)\right) \\ &\quad + 3\alpha(1-\alpha)^2 \times \left(\frac{K}{2} + \frac{1}{2}\left(\tau_H + \frac{\tau_L}{2}\right)\right)\end{aligned} \tag{28}$$

With probability $(1-\alpha)^3$, all three analysts cover stock L. There is no learning about stock H and the posterior precision equals the prior precision, τ_H. With probability α^3, all three analysts cover stock H. The entire capacity K is used to learn about stock H, that is $\dfrac{1}{3}(\hat{\tau}_H - \tau_H)$ and the posterior precision is $\hat{\tau}_H = 3K + \tau_H$. With probabilities $3\alpha^2(1-\alpha)$ and $3\alpha(1-\alpha)^2$ exactly two, respectively one analyst covers stock H and the posterior variance is given by equation (14). Further, the posterior precision of asset L is, using the same reasoning,

$$\begin{aligned}\hat{\tau}_L &= (1-\alpha)^3 \times (3K + \tau_L) + \alpha^3 \times \tau_L + 3\alpha^2(1-\alpha) \times \left(\frac{K}{2} + \frac{1}{2}\left(\frac{\tau_H}{2} + \tau_L\right)\right) \\ &\quad + 6\alpha(1-\alpha)^2 \left(\frac{K}{2} + \frac{1}{2}\left(\tau_H + \frac{\tau_L}{2}\right)\right)\end{aligned} \tag{29}$$

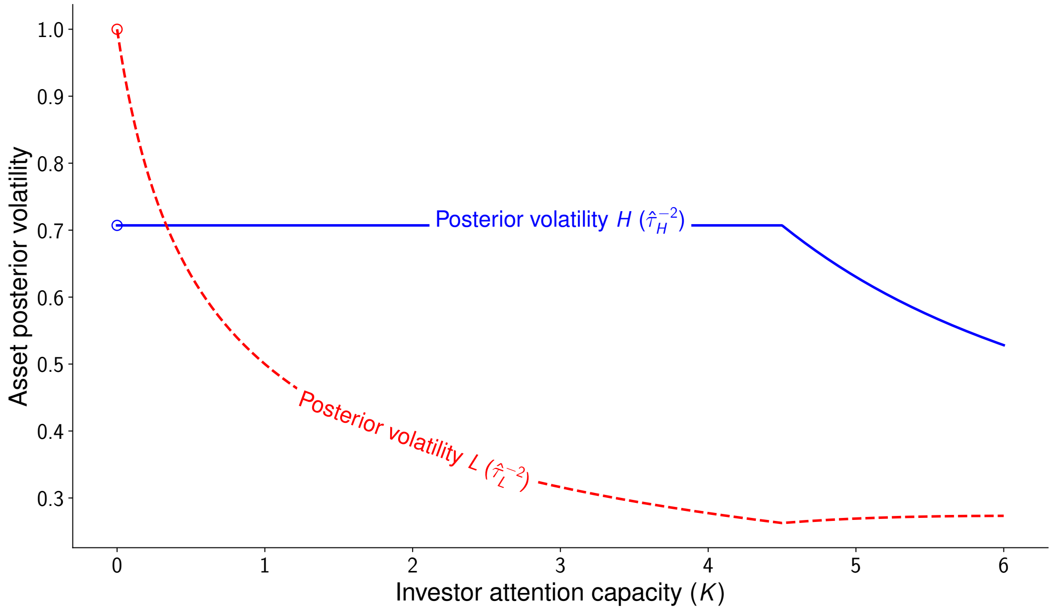
It immediately follows that the posterior precision of stock L is larger than the posterior precision of asset H, since $K > 3\left(\tau_H - \frac{\tau_L}{2}\right)$ and

$$\hat{\tau}_L - \hat{\tau}_H = \frac{(\tau_H - \tau_L)(23K + 3(\tau_H + \tau_L))}{4K - 3(\tau_H + \tau_L)} > 0. \tag{30}$$

Therefore, unless investor learning capacity is very limited ($K < \dfrac{\tau_H - \tau_L}{3}$), there is a positive correlation between analyst coverage clustering and posterior asset payoff precision. Figure 9 illustrates that, if investors attention capacity is large enough, the ex ante riskiest asset becomes the safest following the analyst coverage choices and investor signal acquisition.

Figure 9: **Posterior asset volatility**

This figure illustrates the posterior asset volatility, $\sqrt{\frac{1}{\tau_i}}$, for stock i as a function of investor attention capacity K for both H and L assets. Parameter values: $\tau_L = 1$ and $\tau_H = 2$.



5.2 Analyst coverage and systematic risk

Payoff uncertainty is likely to be correlated in the cross-section of stocks. First, shocks to the macro-economy and market risk affect all assets. Second, innovative industries, such as technology firms, face more uncertainty than established industries, such as utilities. Third, regulatory reforms might improve disclosure standards across all companies (for example, the eXtensible Business Reporting Language, XBRL, introduced in 2009: see Dong, Li, Lin, and Ni, 2016) In this section, we study the impact of a common shock to payoff uncertainty on analyst crowding in equilibrium.

We impose the following structure on asset uncertainty: the prior precision parameters for both stocks have a common component ξ, standing in for systematic risk, and an idiosyncratic component θ_i, that is

$$\tau_H = \xi + \theta_H \text{ and}$$
$$\tau_L = \xi + \theta_L. \tag{31}$$

Corollary 3. (Common precision shock.) *An increase in the common component of stock precision, ξ, leads to more coverage clustering in equilibrium.*

The intuition of Corollary 3 is that strategic complementarity effects in analyst coverage are relatively more important when systematic risk is the dominant risk component, since investors have weaker preferences to learn about one particular (relatively risky) asset. From equation (16), investors allocate more attention to stocks that are (i) ex ante relatively more uncertain (higher τ_i) or (ii) are better covered by analysts. If the systematic component of uncertainty is larger, investors are more sensitive to analyst coverage. We can decompose equation (16) to show the impact of common and idiosyncratic uncertainty on capacity shares,

$$C_i\left(n_i, n_{-i}\right) = \frac{1}{n_i}\left(\hat{\tau}_i^\star - \tau_i\right) = \frac{K}{m} + \underbrace{\xi\left(\frac{1}{m}\sum_{s\in\mathcal{M}}\frac{1}{n_s} - \frac{1}{n_i}\right)}_{\text{common component}} + \underbrace{\left(\frac{1}{m}\sum_{s\in\mathcal{M}}\frac{\theta_s}{n_s} - \frac{\theta_i}{n_i}\right)}_{\text{idiosyncratic component}}, \tag{32}$$

if $\hat{\tau}_i^\star > \tau_i$. Therefore, an increase in ξ leads to investors allocating more capacity to better covered stocks. The strategic complementarity is reinforced, and analyst coverage becomes more crowded.

From Propositions 2 and 3, analysts choose to cover the same stock for $K \leq 3\left(\tau_H - \frac{\tau_L}{2}\right)$ and randomize between stocks with probabilities $\alpha^\star$ and $1-\alpha^\star$ if $K > 3\left(\tau_H - \frac{\tau_L}{2}\right)$, where $\alpha^\star$ is defined in equation (22). The capacity threshold increases in ξ since

$$3\left(\tau_H - \frac{\tau_L}{2}\right) = 3\left(\frac{\xi}{2} + \theta_H - \frac{\theta_L}{2}\right) \nearrow \xi, \tag{33}$$

leading to a wider region for K where crowded pure strategy equilibria emerge. At the same time, we can re-write $\alpha^\star$ as

$$\alpha^\star = \frac{1}{2} - \frac{9\left(\theta_H - \theta_L\right)}{8K - 6\left(2\xi + \theta_H + \theta_L\right)} < \frac{1}{2}, \tag{34}$$

which decreases in ξ. Therefore, from equation (25), it follows that larger common uncertainty leads to a more skewed analyst coverage. Figure 10 illustrates the relationship between systematic risk and equilibrium coverage clustering.

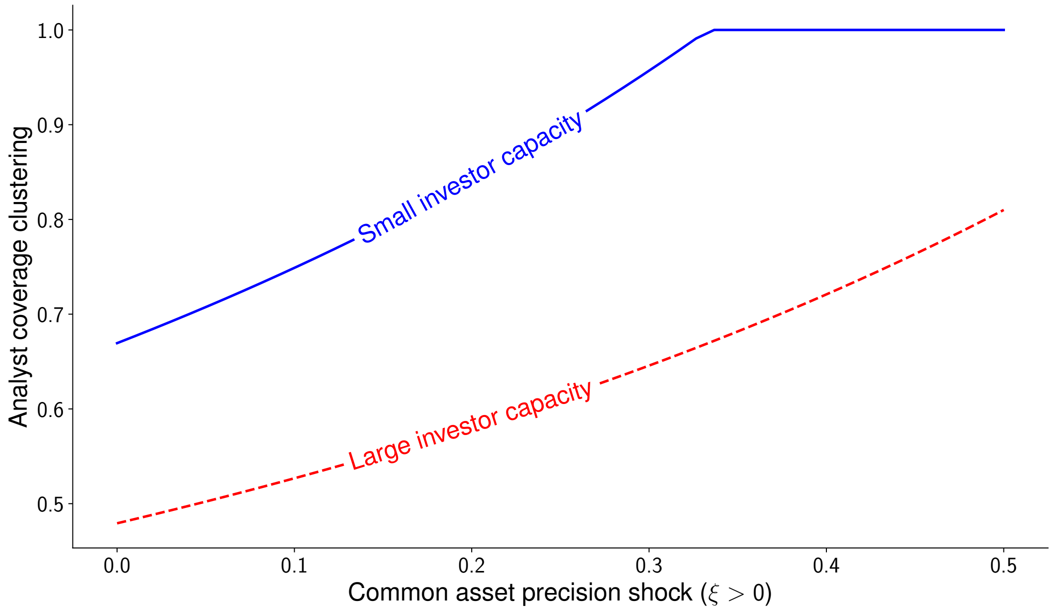
5.3 Analyst coverage, price discovery, and market capitalization

The model yields testable predictions for trading quantities such as price discovery or market capitalization. We measure price discovery as the inverse variance of price errors for stock i, that is

$$\text{Price discovery}_i = \text{var}^{-1}\left(p_i - v_i\right). \tag{35}$$

Figure 10: **Analyst clustering and systematic risk**

This figure illustrates the equilibrium analyst coverage clustering, defined in equation (24) as a function of a positive common precision shock $\xi > 0$ for both assets. We illustrate the relationship for two levels of investor attention capacity, $K = 5$ (blue solid line) and $K = 5.5$ (red dashed line). Parameter values: $\tau_L = 1$ and $\tau_H = 2$.



From equations (8) and (10) it follows that the price error depends on the true value of the asset and the precision of the signal,

$$p_i - v_i = \mu + \left(1 - \frac{\tau_i}{\hat{\tau}_i}\right)\left(\underbrace{v_i + \epsilon_i}_{\text{signal } s_i} - \mu\right) - \frac{\rho}{\hat{\tau}_i} - v_i$$

$$= \text{Non-stochastic term} + \left(1 - \frac{\tau_i}{\hat{\tau}_i}\right)\epsilon_i - \frac{\tau_i}{\hat{\tau}_i} v_i. \tag{36}$$

Since v_i and ϵ_i are by construction independent, it follows that the variance of price errors can be written as

$$\text{var}\,(p_i - v) = \left(1 - \frac{\tau_i}{\hat{\tau}_i}\right)^2 \frac{1}{\hat{\tau}_i - \tau_i} + \frac{\tau_i^2}{\hat{\tau}_i^2}\frac{1}{\tau_i} = \frac{1}{\hat{\tau}^i}. \tag{37}$$

Therefore, price discovery is proportional to the posterior asset precision. In Section 5.1 we showed there is a positive correlation between analyst clustering and posterior precision. Therefore, we predict a positive correlation between asset clustering and price discovery.

We turn next to market capitalization. Since the asset supply is equal to one, the market capitalization for each stock simply coincides with its price. The expected price p_i at the start of the game can be decomposed, from equation (36), into a "zero-coverage" price and a positive learning component,

$$\mathbb{E}\left[p_i \mid \mathcal{F}_0\right] = \mu - \frac{\rho}{\hat{\tau}_i} = \underbrace{\mu - \frac{\rho}{\tau_i}}_{\substack{\text{zero coverage}\\ \text{price}}} + \underbrace{\frac{\rho}{\tau_i} - \frac{\rho}{\hat{\tau}_i}}_{\substack{\text{learning component}\\ >0}}. \tag{38}$$

From equation (38), higher analyst coverage increases the ex-post asset precision $\hat{\tau}_i$. As uncertainty is resolved, investors require lower expected returns in the future (from $t = 1$ to $t = 2$). Consequently, the price $t = 1$ increases with analyst coverage, leading to higher market capitalization. Consistent with the empirical evidence, the model predicts that analyst clustering occurs in securities with large market capitalization.

Finally, the portfolio share w_i can be interpreted as the trading volume at $t = 1$. From equation 9, stocks with lower posterior volatility have a larger share in investors' portfolio, and therefore higher volume at $t = 1$. Therefore, the model predicts that analyst clustering occurs in securities with large trading volume, and the relationship should be stronger if a larger share of investors are passive index trackers who are likely to invest more in large market cap securities.

5.4 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 Lemma 1, the investors' objective function is the product of posterior precisions.

Clearly, for any capacity K, if analysts only cover one stock, the investors prefers 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

$$\underbrace{(3K + \tau_L)}_{\hat{\tau}_L} \tau_H > \underbrace{(3K + \tau_H)}_{\hat{\tau}_H} \tau_L. \tag{39}$$

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,

$$\alpha_{\text{optimal}} = \arg\max_{\alpha} \sum_{n=0}^{3} \binom{3}{n} \alpha^n (1-\alpha)^{3-n} \hat{\tau}_H (n, 3-n) \hat{\tau}_L (n, 3-n), \tag{40}$$

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 maximizing a polynomial of degree six, but the function is well-behaved on $\alpha \in [0, 1]$ and a maximum can be found numerically. We note that, even for very large attention capacity K, investors always prefer that analysts cover asset L more than asset H. To see this, we evaluate the K partial derivative of the posterior precision product in equation (40) at $\alpha = \frac{1}{2}$ and obtain

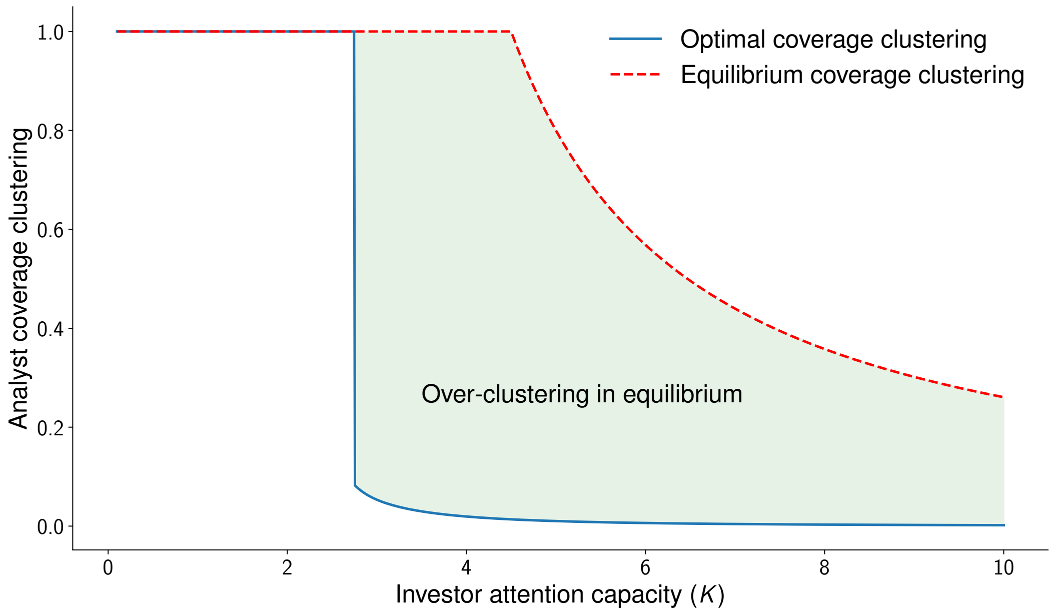
$$\left.\frac{\partial \hat{\tau}_L \hat{\tau}_H(\alpha)}{\partial K}\right|_{\alpha=\frac{1}{2}} = -\frac{15}{128}(\tau_H - \tau_L)(8K + 3(\tau_H + \tau_L)) < 0. \tag{41}$$

Therefore, if $\alpha = \frac{1}{2}$, one can increase the posterior precision product by setting α slightly less than one half. Therefore, some clustering is always preferred unless assets have equal prior variance.

We find that analyst coverage is excessively clustered, particularly so if the attention constraint is relaxed (i.e., for a large K). Figure 11 plots the equilibrium analyst clustering defined in equation (24) and the optimal coverage clustering found numerically by maximizing the product in (40) against investor attention capacity. For low capacity, analyst equilibrium behavior coincides with investors' preference and clustering is optimal. However, as investor capacity increases, coverage clustering in equilibrium decreases too slowly compared to investors' optimum. Once a certain capacity threshold is exceeded, investors would optimally switch very fast from full coverage of a single stock to a balanced coverage distribution. However, 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 11: **Excessive analyst clustering**

This figure illustrates the equilibrium analyst coverage clustering, defined in equation (24) as well as the optimal analyst clustering from the investors' perspective, as a function of the investor capacity K. Parameter values: $\tau_L = 1$ and $\tau_H = 2$.



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. 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 my 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 larger stocks, with lower return volatility, better price discovery, 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.

——— , 2019, Dynamic Attention Behavior under Return Predictability, *Management Science* Forthcomin.

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, 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*.

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.

Da, Zhi, Joseph Engelberg, and Pengjie Gao, 2011, In Search of Attention, *The Journal of Finance* 66, 1461–1499.

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.

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.

Dong, Yi, Oliver Zhen Li, Yupeng Lin, and Chenkai Ni, 2016, Does Information-Processing Cost Affect Firm-Specific Information Acquisition? Evidence from XBRL Adoption, *Journal of Financial and Quantitative Analysis* 51, 435–462.

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 the Stock Market, *Manuscript*.

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.

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

Lawrence, Alastair, James P Ryans, and Estelle Y Sun, 2017, Investor Demand for Sell-Side Research, *The Accounting Review* 92, 123–149.

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.

Martineau, Charles, 2019, The Evolution of Financial Market Efficiency: Evidence from Earnings Announcements, *Manuscript*.

McNichols, Maureen, and Patricia C O'Brien, 1997, Self-Selection and Analyst Coverage, *Journal of Accounting Research* 35, 167–199.

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.

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_i, $i \in \{H, L\}$	Prior precision of asset i (at time $t = -1$), without coverage.
μ_i	Expected return of the risky asset i.
ρ	Investors' CARA risk-aversion.
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 $t = 1$).
α	Probability of each analyst to cover the high-precision asset H.
κ_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 $\sum_i C_i = K$.
w_i, p_i	Portfolio share and market-clearing price of asset i at $t = 1$.

B Proofs

Lemma 1

Proof. The proof is identical to the first step in Appendix A.2 in Van Nieuwerburgh and Veldkamp (2010). Particularly, at $t = 0$ investors choose learning shares to maximize expected utility over posterior expected returns $\hat{\mu}_i$, using the optimal portfolio weights in equation (9):

$$U_{\mathbf{I},t=0} = -\mathbb{E}_{t=0}\left[\exp\left(-\frac{1}{2}\sum_i \frac{(\hat{\mu}_i - p_i r)^2}{(\hat{\tau}_i)^{-1}}\right)\right]. \tag{B.1}$$

From equation (8), at $t = 0$ the posterior mean is normally distributed around the unconditional expectation, $\mathbb{E}_{t=0}\hat{\mu}_i = \mu_i$, and with variance $\tau_i^{-1} - \hat{\tau}_i^{-1}$ However, since $\hat{\mu}_i$ enters the investors' objective as a squared term, the expectation to maximize is of a Wishart variable and not of a log-normal variable. We apply the formula for expectation of a Wishart and obtain the result in equation (27) in Van Nieuwerburgh and Veldkamp (2010), p. 801:

$$U_{\mathbf{I},t=0} = -\left(\frac{\prod_i \hat{\tau}_i}{\prod_i \tau_i}\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\sum_i \frac{(\mu_i - p_i r)^2}{\tau_i^{-1}}\right). \tag{B.2}$$

Maximizing the expected utility in (B.2) is equivalent to maximizing the product of posterior variances, $\prod_i \hat{\tau}_i$. $\square$

Proposition 1

Proof. We compute the first order condition of Lagrangian $\mathcal{L}$ in equation (11), with respect to each posterior precision $\hat{\tau}_i$:

$$\mathcal{L} = \sum_i \ln\left(\hat{\tau}_i^\star\right) + \psi\left[K - \sum_i \frac{1}{n_i}\left(\hat{\tau}_i - \tau_i\right)\right] + \sum_i \lambda_i \left(\hat{\tau}_i - \tau_i\right). \tag{B.3}$$

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

$$\frac{\partial \mathcal{L}}{\partial \hat{\tau}_i} = \frac{1}{\hat{\tau}_i} - \frac{1}{n_i}\psi + \lambda_i \Longrightarrow \hat{\tau}_i^\star = \frac{1}{\frac{1}{n_i}\psi - \lambda_i}. \tag{B.4}$$

From the complementary slackness condition, if $i \in \mathcal{M}$ such that the investor acquires a signal about i, the Lagrange multiplier of the no-forgetting constraint, λ_i, is zero and therefore

$$\hat{\tau}_i^\star = \frac{n_i}{\psi}, \forall i \in \mathcal{M}. \tag{B.5}$$

Finally, we solve for ψ from the capacity constraint:

$$\sum_{i\in\mathcal{M}} \frac{1}{n_i}\left(\frac{n_i}{\psi} - \tau_i\right) = K \Longrightarrow \frac{1}{\psi} = \frac{1}{m}\left(K + \sum_{i\in\mathcal{M}} \frac{\tau_i}{n_i}\right). \tag{B.6}$$

From equations (B.5) and (B.6), it follows that

$$\hat{\tau}_i^\star = n_i\left(\frac{K}{m} + \frac{1}{m}\sum_{i\in\mathcal{M}} \frac{\tau_i}{n_i}\right), \tag{B.7}$$

which concludes the proof. $\square$

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 no-forgetting constraints to establish when investors choose to learn about both stocks if $\eta = (2, 1)$. Using the posterior variances from Proposition 1, the two no-forgetting constraints can be written as:

$$\hat{\tau}_H = 2\left(\frac{K}{2} + \frac{1}{2}\left(\frac{\tau_H}{2} + \tau_L\right)\right) > \tau_H \Longrightarrow K > \frac{\tau_H}{2} - \tau_L \text{ and} \tag{B.8}$$

$$\hat{\tau}_L = \left(\frac{K}{2} + \frac{1}{2}\left(\frac{\tau_H}{2} + \tau_L\right)\right) > \tau_L \Longrightarrow K > \tau_L - \frac{\tau_H}{2}. \tag{B.9}$$

We note that, depending on whether $\tau_L \lesseqgtr \frac{\tau_H}{2}$, one of the thresholds is negative and the other is positive. Therefore, both no-forgetting constraints are satisfied if $K \geq \underline{K} = \max\left\{\frac{\tau_H}{2} - \tau_L, \tau_L - \frac{\tau_H}{2}\right\}$. Otherwise, investors learn about one asset only. In particular, investors optimally learn about asset H if

$$\hat{\tau}_H \tau_L \geq \tau_H \hat{\tau}_L \Longrightarrow (2K + \tau_H)\,\tau_L \geq \tau_H\,(K + \tau_L) \Longrightarrow \tau_H \leq 2\tau_L, \tag{B.10}$$

and optimally learn about asset L if $\tau_H > 2\tau_L$.

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

$$\hat{\tau}_H = \left(\frac{K}{2} + \frac{1}{2}\left(\tau_H + \frac{\tau_L}{2}\right)\right) > \tau_H \Longrightarrow K > \tau_H - \frac{\tau_L}{2} \text{ and} \tag{B.11}$$

$$\hat{\tau}_L = 2\left(\frac{K}{2} + \frac{1}{2}\left(\tau_H + \frac{\tau_L}{2}\right)\right) > \tau_L \Longrightarrow K > \frac{\tau_L}{2} - \tau_H. \tag{B.12}$$

Since $\tau_H > \tau_L$, it follows that investors learn about both assets if $K > \overline{K} \equiv \tau_H - \frac{\tau_L}{2}$. It immediately follows that $\overline{K} > \underline{K}$. $\square$

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 $\frac{\tau_H}{\tau_L} \geq 2$.

Region 1: $K \leq \frac{\tau_H}{2} - \tau_L$. From Corollary 1, investors acquire signals about H if and only all three analysts cover stock H. From equations (16), (19), (20), and Corollary 1, the expected analyst utilities from covering H and L are, respectively:

$$U(H) = \alpha^2 \frac{K}{3} \text{ and}$$
$$U(L) = \alpha^2 K + 2\alpha(1-\alpha)\frac{K}{2} + (1-\alpha)^2\frac{K}{3}. \tag{B.13}$$

It immediately follows that $U(L) > U(H)$ 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: $K \in \left(\frac{\tau_H}{2} - \tau_L, \tau_H - \frac{\tau_L}{2}\right]$. From Corollary 1, if $\eta = (2,1)$ investors learn about both securities and if $\eta = (1,2)$ investors only learn about L. From equations (16), (19), (20), and Corollary 1, the expected analyst utilities from covering H and L are, respectively:

$$\begin{aligned} U(H) &= \alpha^2 \frac{K}{3} + 2\alpha(1-\alpha)\frac{1}{2}\left[\frac{K}{2} + \frac{1}{2}\left(\tau_L - \frac{\tau_H}{2}\right)\right] + (1-\alpha)^2 \times 0 \text{ and} \\ U(L) &= \alpha^2\left[\frac{K}{2} + \frac{1}{2}\left(\frac{\tau_H}{2} - \tau_L\right)\right] + 2\alpha(1-\alpha)\frac{K}{2} + (1-\alpha)^2 \frac{K}{3}. \end{aligned} \tag{B.14}$$

The value of $\alpha^\dagger$ that solves $U(H) = U(L)$ is

$$\alpha^\dagger = \frac{4K}{2K + 3(2\tau_L - \tau_H)} > 2, \tag{B.15}$$

since $\tau_L - \frac{\tau_H}{2} < 0$. Therefore, the two utility functions do not cross for any $\alpha \in [0,1]$. Indeed, $U(L) > U(H)$ for any α since

$$\begin{aligned} U(L, \alpha = 0) - U(H, \alpha = 0) &= \frac{K}{3} > 0 \text{ and} \\ U(L, \alpha = 1) - U(H, \alpha = 1) &= \frac{1}{2}\left(\frac{\tau_H}{2} - \tau_L\right) + \frac{K}{6} > 0. \end{aligned} \tag{B.16}$$

Therefore, it is optimal for each analyst to cover the low-precision security L with probability one for any $K \le \tau_H - \dfrac{\tau_L}{2}$.

Region 3: $K > \tau_H - \frac{\tau_L}{2}$. 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 (16), (19), (20), and Corollary 1, the expected analyst utilities from covering H and L are, respectively:

$$\begin{aligned} U(H) &= \alpha^2 \frac{K}{3} + 2\alpha(1-\alpha)\frac{1}{2}\left[\frac{K}{2} + \frac{1}{2}\left(\tau_L - \frac{\tau_H}{2}\right)\right] + (1-\alpha)^2 \times \left[\frac{K}{2} + \frac{1}{2}\left(\frac{\tau_L}{2} - \tau_H\right)\right] \text{ and} \\ U(L) &= \alpha^2\left[\frac{K}{2} + \frac{1}{2}\left(\frac{\tau_H}{2} - \tau_L\right)\right] + 2\alpha(1-\alpha)\left[\frac{K}{2} + \frac{1}{2}\left(\tau_H - \frac{\tau_L}{2}\right)\right] + (1-\alpha)^2 \frac{K}{3}. \end{aligned} \tag{B.17}$$

The difference between the two utility functions is

$$U(L) - U(H) = \frac{1}{3}\alpha\left(K - \frac{3}{4}(\tau_H + \tau_L)\right) + \frac{1}{6}\left(3\left(\tau_H - \frac{\tau_L}{2}\right) - K\right). \tag{B.18}$$

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

$$\alpha^\star = \frac{2K - 6\tau_H + 3\tau_L}{4K - 3(\tau_H + \tau_L)} = \frac{1}{2} - \frac{9(\tau_H - \tau_L)}{8K - 6(\tau_H + \tau_L)}. \tag{B.19}$$

We tabulate the sign of the utility difference in equation (B.18) for $\alpha = 0$ and $\alpha = 1$ to establish the existence of pure and mixed strategy equilibria.

Learning capacity	$U(L) - U(H)$		$\alpha^\star$	Equilibrium
	$\alpha = 0$	$\alpha = 1$		
$K \leq 3\left(\tau_L - \frac{\tau_H}{2}\right)$	$+$	$-$	$\alpha^\star \in (0,1)$	Two equilibria: Cover L **or** H
$K \in \left(3\left(\tau_L - \frac{\tau_H}{2}\right), \frac{3}{4}(\tau_H + \tau_L)\right]$	$+$	$+$	$\alpha^\star > 1$	Cover L is dominant strategy.
$K \in \left(\frac{3}{4}(\tau_H + \tau_L), 3\left(\tau_H - \frac{\tau_L}{2}\right)\right]$	$+$	$+$	$\alpha^\star < 1$	Cover L is dominant strategy.
$K > 3\left(\tau_H - \frac{\tau_L}{2}\right)$	$-$	$+$	$\alpha^\star \in (0,1)$	Mixed strategy

For $K \leq 3\left(\tau_L - \frac{\tau_H}{2}\right)$, there are two pure strategy equilibria. From equation (B.18), if two analysts choose $\alpha = 0$, then $U(L) - U(H) > 0$ and therefore the third analyst also chooses $\alpha = 0$. Conversely, 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 $(\alpha^\star, 1 - \alpha^\star)$ 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,

$$U(L) - U(H) = \frac{1}{3}(\alpha^\star + \epsilon)\left(K - \frac{3}{4}(\tau_H + \tau_L)\right) + \frac{1}{6}\left(3\left(\tau_H - \frac{\tau_L}{2}\right) - K\right) < 0, \tag{B.20}$$

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

Since $\tau_H > 2\tau_L$, it follows that $\tau_H - \frac{\tau_L}{2} > 3\left(\tau_L - \frac{\tau_H}{2}\right)$ and therefore the first row in the table above does not correspond to a valid parameter region. Consequently, over the three capacity regions, analysts cover L with probability one for $K \leq 3\left(\tau_H - \frac{\tau_L}{2}\right)$. For $K > 3\left(\tau_H - \frac{\tau_L}{2}\right)$, a mixed strategy emerges in equilibrium where analysts choose H with probability $\alpha^\star$ in equation (B.19). $\square$

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 $\dfrac{\tau_H}{\tau_L} < 2$.

Region 1: $K \leq \tau_L - \dfrac{\tau_H}{2}$**.** From Corollary 1, investors acquire a signal about H alone if $\eta = (2,1)$ and about L alone if $\eta = (1,2)$. From equations (16), (19), (20), and Corollary 1, the expected analyst utilities from covering H and L are, respectively:

$$\begin{aligned} U(H) &= \alpha^2 \frac{K}{3} + 2(1-\alpha)\alpha\frac{K}{2} + (1-\alpha)^2 \times 0 \text{ and} \\ U(L) &= \alpha^2 \times 0 + 2\alpha(1-\alpha)\frac{K}{2} + (1-\alpha)^2 \frac{K}{3}. \end{aligned} \tag{B.21}$$

The difference between the utility functions is

$$U(L) - U(H) = \frac{K}{3}(1 - 2\alpha). \tag{B.22}$$

Two pure strategy equilibria emerge. If $\alpha = 0$ for two analysts, $U(L) - U(H) > 0$ 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,

$$U(L) - U(H) = U(L) - U(H) = \frac{K}{3}(1 - 2\alpha - 2\epsilon) < 0, \tag{B.23}$$

and therefore the third analyst chooses H with probability one.

It immediately follows that $U(L) = U(H)$ if and only if $\alpha = \dfrac{1}{2}$ and therefore analysts mix between covering each stock with equal probability.

Region 2: $K \in \left(\tau_L - \dfrac{\tau_H}{2}, \tau_H - \dfrac{\tau_L}{2}\right]$. From Corollary 1, if $\eta = (2,1)$ investors learn about both securities and if $\eta = (1,2)$ investors only learn about L. The expected analyst utilities from covering H and L are therefore the same as in (B.14). The value of $\alpha^\star$ that solves $U(H) = U(L)$ is

$$\alpha^\dagger = \frac{4K}{2K + 3(2\tau_L - \tau_H)}. \tag{B.24}$$

We evaluate $U(L) - U(H)$,

$$U(L) - U(H) = \frac{1}{12}\left[4K - \alpha\left(2K + 3(2\tau_L - \tau_H)\right)\right] \tag{B.25}$$

for $\alpha = 0$ and $\alpha = 1$, that is,

$$\begin{aligned} U(L, \alpha = 0) - U(H, \alpha = 0) &= \frac{K}{3} > 0 \text{ and} \\ U(L, \alpha = 1) - U(H, \alpha = 1) &= \frac{1}{2}\left(\frac{\tau_H}{2} - \tau_L\right) + \frac{K}{6}. \end{aligned} \tag{B.26}$$

If $K > 3\left(\tau_L - \frac{\tau_H}{2}\right)$, then the utility difference never switches sign. Therefore, $U(L) - U(H) > 0$ for any $\alpha \in [0,1]$ and analysts cover L with probability one.

If $K \leq 3\left(\tau_L - \frac{\tau_H}{2}\right)$, 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,

$$U(L) - U(H) = \frac{1}{12}\left[4K - \alpha^{\dagger}\underbrace{(3(2\tau_L - \tau_H) - 2K)}_{>0}\right] < 0, \tag{B.27}$$

and therefore the third analyst chooses H with probability one.

Two cases emerge:

1. If $\frac{\tau_H}{\tau_L} > \frac{7}{5}$, two pure strategy equilibria emerge for $K \in \left(\tau_L - \frac{\tau_H}{2}, 3\left(\frac{\tau_H}{2} - \tau_L\right)\right]$. Conversely, for $K \in \left(3\left(\frac{\tau_H}{2} - \tau_L\right), \tau_H - \frac{\tau_L}{2}\right]$, analysts choose to cover L with probability one.
2. If $\frac{\tau_H}{\tau_L} \leq \frac{7}{5}$, two pure strategy equilibria emerge for the entire region, $K \in \left(\tau_L - \frac{\tau_H}{2}, \tau_H - \frac{\tau_L}{2}\right]$.

Region 3: $K > \tau_H - \frac{\tau_L}{2}$**.** 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,

$$\alpha^{\star} = \frac{2K - 6\tau_H + 3\tau_L}{4K - 3(\tau_H + \tau_L)} = \frac{1}{2} - \frac{9(\tau_H - \tau_L)}{8K - 6(\tau_H + \tau_L)}. \tag{B.28}$$

For convenience, we tabulate again the sign of the utility difference in equation (B.18) for $\alpha = 0$ and $\alpha = 1$ to establish the existence of mixed strategy equilibria.

Learning capacity	$U(L) - U(H)$ $\alpha = 0$	$\alpha = 1$	$\alpha^{\star}$	Equilibrium
$K \leq 3\left(\tau_L - \frac{\tau_H}{2}\right)$	$+$	$-$	$\alpha^{\star} \in (0,1)$	Two equilibria: Cover L **or** H
$K \in \left(3\left(\tau_L - \frac{\tau_H}{2}\right), \frac{3}{4}(\tau_H + \tau_L)\right]$	$+$	$+$	$\alpha^{\star} > 1$	Cover L is dominant strategy.
$K \in \left(\frac{3}{4}(\tau_H + \tau_L), 3\left(\tau_H - \frac{\tau_L}{2}\right)\right]$	$+$	$+$	$\alpha^{\star} < 1$	Cover L is dominant strategy.
$K > 3\left(\tau_H - \frac{\tau_L}{2}\right)$	$-$	$+$	$\alpha^{\star} \in (0,1)$	Mixed strategy

As in Proposition 2, the mixed strategy equilibrium for $K \leq 3\left(\tau_L - \frac{\tau_H}{2}\right)$ is not stable. For $K > 3\left(\tau_H - \frac{\tau_L}{2}\right)$, a mixed strategy always arises in equilibrium. We combine the table above with the equilibrium strategy for the second region to complete the proof. □

Corollaries 2 and 3

Proof. Proofs are provided in the main text. □