The Acquisition of Earnings Information: Along the Extensive Margin*

Roberto C. Gutierrez Jr.[a,†]
and
Charles Gaa[b]

August 13, 2019

Abstract

Information drives financial markets, yet the choices that market participants must make when acquiring their information remain largely unexamined empirically. We observe how many firms' earnings releases the *Wall Street Journal* chooses to cover each month. The extent of the coverage, controlling for firm and earnings characteristics, covaries negatively with the business cycle and positively with the equity risk premium, consistent with theories of costly acquisition of payoff-relevant information. The extensiveness measure is not simply capturing the rate of diffusion of revealed earnings news, but rather serves as a proxy for times when investors are more actively learning about the earnings and valuations of small stocks. Mutual funds trade small stocks better when the extent of coverage is high, consistent with skill being revealed when prices are less noisy.

*We are grateful for comments from Ioannis Branikas, Diane Del Guercio, Michal Dzielinski, Eric Kelley, Kyle Peterson, Rob Ready, Brady Twedt, and seminar participants at the EFA Meeting (Austria), Pacific Northwest Finance Conference (University of Washington), University of Tennessee, and Tulane University. Gutierrez acknowledges financial support from the Finance and Securities Analysis Center at the University of Oregon and the James E. Bramsen Faculty Fellowship. The views expressed in this paper are those of the authors. No responsibility for them should be attributed to the Bank of Canada.

[a]Lundquist College of Business, University of Oregon.(rcg@uoregon.edu)

[†]Corresponding author.

[b]Bank of Canada. (cgaa@bank-banque-canada.ca)

Information drives financial markets. Every decision that an agent makes is conditional on some information set. Since information gathering is costly, agents must make choices about which information they will acquire, or ignore. How these choices influence their subjective beliefs, and ultimately market prices, are important interactions to understand. While there are rich theoretical literatures on information acquisition and information asymmetry across agents in financial markets, empirical testing lags behind the theoretical work because of the latency of an agent's information set. In this study, we measure one dimension of the information acquisition activity of a prominent news agent and examine how information gathering temporally covaries with economic conditions and with the informativeness of stock prices.

Earnings releases are widely viewed as important information events for firms. Moreover, "earnings seasons" are opportunities to learn more about the economic state of particular industries and the state of the economy overall. Hundreds, sometimes thousands, of firms release their earnings information each month. How do agents filter these batches of information? Specifically, does the number of firm-earnings signals that an agent chooses to observe change over time?

Empirically examining the extensive margin of an agent's information gathering — the number of firms observed — is novel. Recent studies motivate potential drivers of the extensive margin as outcomes of optimal information gathering. Peng and Xiong (2006) and Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016) model information about the payoffs of risky assets with a factor structure, i.e., macroeconomic, industry, and firm-specific components. Agents select which information signals to acquire by trading off the marginal benefits of acquiring a signal with the marginal costs. These studies show that greater information-processing capacity should be devoted to learning about more systematic components of payoffs rather than more idiosyncratic ones. The more systematic a component is, the more beneficial it is to know. This view seems to match the common practice of market participants paying more attention to "bellwether" firms (typically large firms) whose earnings contain broader signals about their own industries and the economy at large.[1]

Turning to the temporal drivers of endogenous information acquisition, the models of Bansal and Shaliastovich (2011), Andrei and Hasler (2015), and Kacperczyk et al. (2016) show that the optimal gathering of information about risky assets will vary countercyclically to the economy. In other words, information is more valuable to investors when the uncertainty of asset payoffs and the price of risk increase.

Motivated by these studies, we investigate one agent's monthly decisions on how far to extend their acquisition of earnings information. For each earnings release in a given month, *The Wall Street Journal* (WSJ) must decide whether or not to report the news event to its subscribers. Admittedly, the WSJ is one agent amongst many market participants, but we believe that observing this particular agent offers unique advantages. First, the WSJ is a prominent and prestigious information agent. Second, the WSJ is tasked with filtering the flow of financial information for its subscribers. The WSJ must be attuned to the financial markets and provide its subscribers with useful information. So, like other agents in the marketplace, the WSJ must select how many firm-earnings signals to extract from a batch of earnings releases, since not all earnings releases can be covered. Third, by observing a single agent over twenty years, rather than aggregating the decisions of multiple agents over the sample period (e.g., other newspapers), we are seemingly relying on a more stationary process of information choices.

To isolate the temporal dynamics in the extensiveness of the WSJ's earnings coverage, we employ a logit model to remove the static components of the WSJ's decision to cover an earnings release. Since some firms and some earnings events have a greater likelihood than others of being covered, such as larger firms and firms with earnings surprises, we estimate the probability of the WSJ's covering each earnings report based on a variety of firm and earnings characteristics. Given a batch of earnings releases in a given month, we predict

[1] Beyond the anecdotes of bellwether firms, many studies find that information at earnings announcements spills over to the prices of related firms. For example, see the studies of Ramnath (2002), Hameed et al. (2015), Savor and Wilson (2016), and Patton and Verardo (2012).

the number of firm-earnings events that we expect will receive WSJ coverage that month.[2] We measure the deviation of the actual number of earnings reports receiving WSJ coverage from the predicted number, as a fraction of the predicted number. This residual percentage of earnings coverage is our measure of the temporal adjustments the WSJ makes to the extensiveness of its earnings coverage, holding firm and earnings characteristics constant. Additionally, we are careful to distinguish the extent of the WSJ's earnings coverage from the monthly linguistic tone of its coverage, as tone has been found to covary with stock returns. That is, we separate the WSJ's decision to gather earnings information from the outcome of the WSJ's processing of that information.[3]

For some perspective, note that the variation in the extent of coverage that we are identifying is small compared to the total number of firms that release earnings each month. In our sample, only about 11% of firms receive WSJ coverage of their earnings. The number of firms per month that release earnings is 915 on average, with 97 firms per month receiving coverage. Moreover, one standard deviation of the extent of coverage (controlling for firm and earnings characteristics) is 29 firms per month on average. As noted earlier, larger firms are more likely to receive newspaper coverage, as their signals tend to be both more beneficial to know and less costly to acquire. While the actual number of earnings releases covered by the WSJ is roughly split between firms that have the largest twenty percent of equity size and the remaining firms, the vast majority of the temporal variation in coverage that we measure occurs outside the largest stocks. That is, smaller stocks are the marginal firms for earnings coverage. The view we take is that the extent of the monthly coverage is the outcome of the WSJ's marginal evaluation of the number of firm-earnings signals

[2] Fang and Peress (2009), Solomon (2012), Ahern and Sosyura (2014), and others find newspaper coverage of firms to tilt toward larger firms, while Solomon and Soltes (2012) find coverage to tilt toward firms with earnings surprises. We also control for other characteristics, such as industry, analyst coverage, dispersion of analysts' forecasts, recent stock returns, as well as seasonalities.

[3] Tetlock (2007), Tetlock et al. (2008), Solomon (2012), Gurun and Butler (2012), and others find that the tone (positive/negative) of media coverage covaries with future stock returns. We control for the aggregate monthly tone of the WSJ articles by employing a multinomial logit model to estimate the probabilities of negative and of non-negative tone. The monthly residual tone of the coverage is unrelated to the temporal dynamics in stock returns that we examine.

needed to convey to its subscribers the information content in a given month's batch of earnings.[4]

We find that the temporal adjustments that the WSJ makes to the extent of its earnings coverage covary with macroeconomic conditions. Coverage extends to more firms when output gap (detrended industrial production) is lower, the consumption-wealth ratio of Lettau and Ludvigson (2001) is higher, and during NBER recessions. That is, the extensive margin moves countercyclically. Further support for a countercyclical margin comes from the fact that the measure covaries positively with the equity risk premium. A one standard deviation increase in the extent of coverage is associated with a roughly 2% increase in the excess stock market return over the next six months. In sum, the behavior of the WSJ's coverage of earnings is consistent with theories of optimal costly information gathering, and hence with a demand-driven information motive for the WSJ's reporting.

We then turn our attention to the information content of stock prices. Essentially, we are jointly testing (i) whether the WSJ's information-acquisition activity serves as a proxy for the information acquisition of the mass of stock market investors and (ii) whether prices are more informative about earnings when the acquisition of earnings information increases. Since smaller firms lie on the extensive margin, we focus on the pricing of the smallest quintile of stocks.

The well-documented post-earnings announcement drift (PEAD) in the returns of small stocks decreases with the extensiveness of earnings coverage, where earnings surprises are defined as deviations of reported earnings from the consensus forecast of analysts. A long-short portfolio formed in months when the extent of coverage is the least earns an abnormal return of nearly 6% over 60 days, while a long-short portfolio formed in the months when the extent of coverage is the greatest earns roughly 2%. The decline in PEAD is consistent with prices immediately after announcements more fully impounding earnings news when coverage expands to more firms.

[4]We abstract from other determinants of the WSJ's coverage decisions, such as other news events besides earnings that are competing for journalists' time and for newspaper space, as well as possible time variation in overall reporting capacity of the WSJ.

To gauge how actively investors are updating small-stock prices in months with greater WSJ coverage, we examine the price reactions upon the announcement of earnings surprises. We find that earnings response coefficients (ERC) for small stocks decline with the extensiveness of coverage. This is inconsistent with earnings news merely diffusing more quickly through the marketplace. In months when coverage extends to more firms, investors appear to be better anticipating earnings forecast errors (measured relative to the consensus forecasts of analysts).[5]

One potential channel for this active learning that has received support in the literature is a systematic learning channel. Prior studies find that intra-industry information from the earnings releases of early announcers flows to the stock prices of related firms (Ramnath, 2002; Hameed et al., 2015). Moreover, Loh and Stulz (2018) find that such spillovers are stronger in bad economic states, which echoes the predictions of Kacperczyk et al. (2016) that investors' information acquisition activity should be countercyclical with a relatively greater capacity devoted to common information signals.

Regardless of the particular learning channels that are engaged here, our findings show empirically that information acquisition is an important mechanism in understanding asset pricing. Additionally, we find that the cross-sectional return dispersion within small stocks increases with the extent of earnings coverage. This further supports the notion that the WSJ-based measure identifies times when investors are more actively gathering information on small stock payoffs and updating prices accordingly.

Lastly, having established that the informativeness of small stock prices about earnings increases with the extensiveness measure, we leverage this relation to examine the investing skills of mutual funds. Recent studies uncover that funds possess skill to fundamentally evaluate stocks (Baker et al., 2010; Jiang and Zheng, 2018). Moreover, Glode (2011) and Kacperczyk et al. (2016) highlight that the well-documented weak unconditional performance of mutual funds may be concealing elements of true underlying skill. Indeed, during

[5] Studies of firm-level attention to earnings releases support an information diffusion channel (e.g., Hirshleifer et al. (2009); Boulland et al. (2017)). The market-level measure of the extent of earnings coverage that is developed here captures a different channel. While only a handful of small firms receive WSJ earnings coverage each month, the extensiveness measure covaries with the returns of hundreds of stocks.

times when stock prices are better reflections of earnings information, we find that active mutual funds perform better. Specifically, we take long positions each quarter in the small stocks having the largest increases in the number of funds choosing to hold them, and short positions in the stocks having the largest decreases. The Carhart four-factor alpha of this long-short portfolio over the next quarter averages 17 basis points per month when the extent of earnings coverage is above its sample average and negative 33 basis points when the extent of coverage is below its sample average. In short, conditioning on the information state of stock prices confirms the findings of prior studies that at least some mutual funds are skilled at stock valuation.[6]

This study makes contributions to a growing literature on information acquisition and to a large literature still struggling to understand the economic rents of actively managed mutual funds. We directly measure the information acquisition activity of the WSJ along the extensive margin. The extensive margin is found to expand across more and smaller stocks as macroeconomic conditions weaken. The novel measure proposed here covaries with several dynamics in the returns of small stocks, implying that it can serve as a proxy for the information gathering of stock investors more broadly. As the information literature predicts, conditioning on the state of the information environment is important. Using this extensiveness measure as a state variable for the earnings content of small-stock prices, we find that mutual funds perform better when prices are better reflections of earnings (i.e., when the extent of coverage is high), suggesting that noise, and perhaps noise traders, obfuscate the true skill of some mutual funds.

The paper proceeds as follows. The next section details the data sources and the measure of the extensiveness of the WSJ's coverage of earnings. We then examine how the extent of coverage covaries with the macroeconomic state in section 2. An analysis of the adjustments that are made along the extensive margin is undertaken in section 3. How the extent of earnings coverage covaries with the information content of stock prices is

[6]We follow Edelen, Ince, and Kadlec (2016) and measure changes in the number of mutual funds that hold a stock rather than changes in the aggregate share of fund ownership for several reasons that we discuss later.

investigated in section 4. Section 5 evaluates the conditional performance of the small-stock picks of mutual funds. Section 6 provides concluding thoughts.

1 Methodology and Data

1.1 Corporate earnings events

Our sample of quarterly earnings reports includes 233,348 firm-earnings events from I/B/E/S, covering the period from October 1984 to December 2005. The base data set of earnings and *The Wall Street Journal* (WSJ) articles to be described below were collected and used by Gaa (2008).[7]

To model the probability of each firm-earnings release being covered by the WSJ, we gather data from I/B/E/S, CRSP, and Compustat on firm-level characteristics, such as size, analyst coverage, recent stock performance, book-to-market ratio of equity (BE/ME), and industry, as well as earnings-specific characteristics, such as earnings surprise and pre-announcement forecast dispersion. Additionally, we omit earnings reports from firms having a negative BE/ME or having a stock price less than one dollar two days prior to the earnings announcement date. The complete set of variables and their sources are provided in Appendix A.1.

1.2 WSJ coverage of earnings

Our measure of the extent to which the WSJ covers earnings is based on how many firm-earnings releases are selected by the WSJ to receive a news article that covers the event. We begin with 68,102 "earnings" news articles from Factiva (code: c151) having at least 100 words. Requiring 100 words ensures that the earnings article represents material information acquisition. The computational linguistics program Rainbow by McCallum (1996) is then used to identify the articles which are about a specific firm's quarterly

[7] How the information gathering process measured here might have changed after this sample period due to News Corp's acquisition of the WSJ in 2007 and potential changes in the information environment for firms due to Regulation Fair Disclosure, Sarbanes Oxley, and the rise of social media is an interesting question for subsequent studies.

earnings release, as some of these earnings articles flagged by Factiva are about industry-level earnings trends, regulatory changes, restatements, accounting scandals, etc. The Naive-Bayesian text categorization is trained on a set of 500 articles and uses a unigram and bigram "bag of words" approach. Only articles with a posterior probability greater than 0.5 of being about a firm's earnings release remain in the final sample of 49,113 articles.

Although the tone of the WSJ articles is not of primary concern in this study, we want to distinguish the decision to cover a given firm-earnings release from the outcome of the WSJ's processing of the earnings information. These distinct aspects may be correlated, and evidence suggests that tone has explanatory power for stock returns (e.g., Tetlock (2007), Tetlock et al. (2008), Solomon (2012), and Gurun and Butler (2012)). To control for tone, we apply another round of text categorization to disambiguate "negative" and "nonnegative" articles, again trained on a set of 500 articles. We classify articles to be negative if the posterior probability of being negative is greater than 0.5, while all other articles are classified as nonnegative.[8]

Finally, for each of the earnings reports, we search for a related news article published in the WSJ within two days on either side of the I/B/E/S announcement date. If at least one article corresponding to a given earnings report is found within this 5-day window, that earnings report is considered to be "covered". We observe coverage for approximately 11% of the announcements in our sample, implying that the typical firm's quarterly earnings release is ignored by the WSJ.

1.3 Measuring the extensiveness of earnings coverage

Our goal is to measure monthly changes in the extent of the WSJ's coverage of earnings releases. Table 1 provides summary statistics to illuminate what typical earnings coverage looks like. On average, 11% of earnings reports released in a given month are covered by

[8]We test Rainbow's classification accuracy by randomly excluding 100 articles from the training set, re-estimating the model, and then checking accuracy for the excluded articles. Across 100 trials, the average accuracy is 88.5% for the first-stage classification ("earnings/not-earnings") and 83.4% for the second-stage ("negative/non-negative").

the WSJ (97 articles out of 915 events). The monthly number of earnings reports that receive coverage has a mean of 97 with a relatively large standard deviation of 80, which is largely driven by the monthly changes in the volume of earnings reports released each month. While normalizing the number of firms receiving earnings coverage by the number of earnings releases in that month can account for the volume changes, such a normalization implicitly assumes that each earnings report is equally likely to receive coverage in the WSJ. However, this is clearly not the case. Newspaper coverage is biased toward larger firms, certain industries, and firms with earnings surprises (e.g., Fang and Peress (2009), Solomon (2012), Solomon and Soltes (2012), and Ahern and Sosyura (2014)). To identify marginal changes in earnings coverage given these biases, we want to control for the characteristics of the firms that release earnings and the characteristics of the earnings information.[9]

To filter the monthly flow of earnings releases, we employ a multinomial logistic regression where the three responses are: negative coverage, nonnegative coverage, and no coverage. Classifications of each article as negative or nonnegative are done using the Rainbow program discussed in section 1.2. The set of firm and earnings characteristics we use as determinants are detailed in the Appendix section A.1. Output of the multinomial logit is shown in Appendix A.2. Using this model, we estimate the probability of coverage for each firm-earnings report. In short, firm and earnings characteristics provide a good deal of information about the probability of WSJ coverage. The (McFadden) pseudo R^2 is 0.27. The primary determinants of WSJ coverage are firm size, industry, and the number of analysts covering the firm. These three variables alone account for a pseudo R^2 of 0.22. We discuss the determinants of explained coverage and the composition of residual coverage more deeply in section 3.

We measure temporal changes in the extensiveness of the WSJ's earnings coverage as the percentage deviation of the actual number of firm-earnings events that receive coverage from the predicted number to receive coverage each month. We label this percentage

[9]Normalizing the number of firms that receive coverage in a given month by the number of firms that report earnings in that month produces a time series that does not covary with the return dynamics we examine in later sections. Hence, adjusting for static biases in the WSJ's coverage is necessary to increase the signal-to-noise ratio of the WSJ's coverage decisions.

residual as EEC, for the "extent of earnings coverage".

$$EEC_t = \frac{\sum_{k=1}^{K_t}(C_k - \hat{C}_k)}{\sum_{k=1}^{K_t}\hat{C}_k} \tag{1}$$

where K_t is the total number of earnings releases observed in month t, C_k is an indicator variable equal to one if earnings report k is associated with a WSJ article and zero if no coverage is observed, and $\hat{C}_k$ is the sum of the predicted probabilities of negative and nonnegative coverage respectively over the earnings releases in month t.

For some perspective, in 1985, the first full year of our sample, the mean monthly number of earnings releases is 410; the mean number of releases receiving WSJ coverage is 43 per month and the predicted number is 40. In 2005, the last year of our sample, the mean monthly number of earnings reports released is 1099, while the mean number of releases receiving WSJ coverage is 104 per month and the predicted number is 155. In Table 1, we see that the mean monthly value of EEC is close to zero at 0.03, but EEC varies a good deal as indicated by its monthly standard deviation of 0.26.

Figure 1 plots EEC (along with CAY). We can visually see the variability in EEC, and we can also see that EEC displays some persistence. The AR(1) coefficient of EEC is 0.62.[10] This is far below the near unit-root behaviors of the macroeconomic variables shown in Table 1, except one. Many studies raise concerns about spurious predictability of stock returns based on such highly persistent measures (e.g., Boudoukh et al. (2008)). We address these concerns by conducting Monte Carlo simulations to assess potential size distortions in our test statistics arising from examining a variable with an AR(1) coefficient of 0.62.

To control for the tone of the WSJ's coverage of earnings, i.e. the outcome of the WSJ's information processing beyond the earnings characteristics that we can observe, we form

[10]The plot of EEC displays small monthly reversals around an underlying lower-frequency cycle. Recall that we include calendar month controls in the logit model. We have verified that EEC is centered near zero across the first, second, and third months of quarters, respectively. Hence, the monthly reversals in EEC are not due to a monthly nor a within-quarter seasonality. The monthly reversals in EEC are consistent with a binding coverage constraint.

a residual measure of tone as:

$$NetEEC_t = \frac{\sum_{k=1}^{K_t}(C_k^{NNeg} - \hat{C}_k^{NNeg}) - \sum_{k=1}^{K_t}(C_k^{Neg} - \hat{C}_k^{Neg})}{\sum_{k=1}^{K_t}\hat{C}_k} \tag{2}$$

where C_k^{NNeg} is a dummy variable equal to one if earnings report k is associated with a nonnegative article, and zero otherwise, while C_k^{Neg} is a dummy variable equal to one if earnings report k is associated with a negative article and zero otherwise. Since each article is characterized to be either negative or nonnegative, $C_k = \left(C_k^{NNeg} + C_k^{Neg}\right)$. The predicted number of firms to receive either negative or non-negative WSJ coverage each month is labeled with a "hat". NetEEC declines as the percentage residual tone of coverage in month t becomes more negative.

EEC has a 0.30 correlation with NetEEC. Why the extensiveness of earnings coverage is positively correlated with the tone is not pursued in this study. For the current purposes, we simply wish to disentangle the decision to acquire information from the outcome of processing that information (beyond the earnings-level characteristics we control for in the logit). In all regression tests, we include both EEC and NetEEC. In the single-variable exercises, such as the figures, we examine EEC orthogonalized with respect to NetEEC. For ease of exposition, however, we will refer to the orthogonalized version of the extensiveness measure in the text simply as "EEC."

2 Extent of Coverage and Macroeconomic Conditions

2.1 Countercyclical information acquisition

The endogenous information models of Bansal and Shaliastovich (2011), Andrei and Hasler (2015), and Kacperczyk et al. (2016) predict that optimal information acquisition by investors moves countercyclically to the economy, as the variance of asset payoffs increases and the price of risk increases. That is, information is more beneficial to have during weaker economic conditions. In this section, we examine how the extensive margin of the WSJ's earnings coverage covaries with macroeconomic conditions. Our proxies for the state of the macroeconomy are NBER recession periods, output gap, and the consumption-wealth ratio. Only two brief declines in economic activity are identified by the NBER during our sample, July 1990 to March 1991 and March 2001 to November 2001. Output gap is measured as the deviation from a linear and quadratic trend in the log of industrial production (Cooper and Priestley (2009)). Turning to a forward-looking measure of economic conditions, we use the consumption-wealth ratio (CAY) of Lettau and Ludvigson (2001). Lettau and Ludvigson (2001, 2002) find that CAY forecasts future stock returns and future growth in corporate investment (private nonresidential fixed investment), suggesting that CAY tracks risk premia and varies countercyclically.[11]

Table 2 reports correlations with these measures of macroeconomic conditions. The correlations indicate that information acquisition by the WSJ is extending farther when economic conditions are poorer. Specifically, EEC is higher during NBER recessions, when CAY is higher, and when output gap is lower. The magnitudes of these correlations are 0.26, 0.48, and -0.27, respectively. The plots of EEC against CAY and against output gap are shown in Figures 1 and 2 (along with NBER recessions as shaded bars). The comovements of EEC with CAY and with output gap are each striking.[12]

Beyond the prediction that the extent of information acquisition moves counter to the business cycle, other time-series covariates are not as clearly identified from the existing information literature. To provide some additional stylized facts, Table 2 also shows correlations with three financial-market variables that have been linked to different aspects of macroeconomic conditions. The term spread of interest rates is measured as the yield of the 10-year Treasury bond minus the yield of the 3-month Treasury bill, and embeds investors' forecasts of future interest rates. The logarithm of the aggregate stock-market dividend yield is a well-studied valuation-ratio predictor of the equity premium.[13] Lastly, the sentiment index of Baker and Wurgler (2006) may capture periods when stock price

[11] CAY is the transitory deviation from a common stochastic trend in logs of aggregate consumption and wealth (that includes human capital). The measure arises from the log-linearization of the intertemporal budget constraint of investors.

[12] Note that EEC and NetEEC are orthogonalized with respect to each other before their correlations with the other variables are estimated.

[13] Campbell (2018) provides excellent reviews of the literatures on the term spread and dividend yield.

dislocations from fundamentals are more systematically abundant. In particular, Stambaugh et al. (2012) find that the returns to many cross-sectional return anomalies are greater when sentiment is high. These and other variables used in the paper are defined in section A.1.2 of the Appendix.

Table 2 shows that EEC is positively correlated with the term spread and the dividend yield and negatively correlated with sentiment. Each of these correlations also suggests a countercyclical interpretation for EEC. However, the reported p-values are not adjusted for the strong serial correlations in these variables that was noted earlier. To examine the marginal explanatory power of these variables, using more reliable test statistics, we regress EEC on the set of macroeconomic variables from Table 2 as well as some additional variables (defined in the Appendix). The p-values are adjusted for spurious rejection rates of the null hypothesis due to the high serial autocorrelations in the covariates and residuals. To do this, we turn to Monte Carlo simulations. We form 10,000 simulated samples of an independently and normally distributed random variable with an AR(1) coefficient that matches that of the dependent variable. We regress our simulated dependent variable on the actual sample of the macro variables. We then calculate the frequency of observing Newey-West t-statistics with six lags that are greater than a given t-statistic found in the actual sample (or less than a given t-statistic that is negative). We multiply the frequency by two to arrive at a simulated p-value for a two-tailed test.

In Table 3, we see that the set of macro variables captures a large fraction of the temporal variation in the WSJ's extent of earnings coverage, with an R-squared of 0.41. CAY and output gap robustly display marginal covariances with EEC, with signs that continue to indicate a countercyclical pattern in the extensive margin. However, the marginal explanatory power of sentiment and recessions are indistinguishable from zero. Interestingly, the sign on the dividend yield switches from a positive correlation in the prior table to a negative relation in the multiple regression. There is obviously some multicollinearity amongst these right-hand side measures. Which covariates are expected to be stronger and whether nonlinearities in economic conditions might be expected are interesting avenues for future research. We say more on this in the next section.

We now turn to examining the relation between the extent of earnings coverage and aggregate stock pricing. Does EEC move with the equity risk premium, which is also expected to increase both with the uncertainty in payoffs as well as the price of risk? To investigate, we regress various windows of the log of future returns of the CRSP VW stock index in excess of the T-bill rate on EEC and NetEEC. Table 4 finds that EEC tracks the equity premium. EEC covaries positively with excess stock returns over months (+1,+6) at the ten-percent level and over months (+1,+12) at the five-percent level (using adjusted p-values). Examining the explanatory power of EEC jointly across multiple horizons provides a more powerful, and more stringent, test (Boudoukh et al. (2008)). Panel B of Table 4 shows the adjusted p-value testing the null hypothesis that the coefficients on EEC are jointly zero over months (+1,+6) and (+1,+12) to be 2.9%. Hence, EEC robustly comoves with the equity risk premium.

The economic magnitude of this relation is notable. The reported coefficients in Table 4 are with respect to standardized coefficients and can be easily interpreted. A one standard deviation increase in EEC is associated with an increase in excess stock returns of 1.89% over the following six months and 3.66% over the following twelve months, which are economically meaningful. Figure 3 plots means of excess stock returns across quintiles of EEC. The explanatory power of EEC for future stock returns over months (+1,+6) and (+1,+12) is visible across the full range of quintiles. Six-month returns vary from about 2% to 7%, and twelve-month returns vary from about 2% to 14%.

In sum, the WSJ extends its earnings coverage across more firms during weaker economic times, based on a variety of measures of the state of the economy. This behavior is consistent with a catering, demand-driven reasoning for the WSJ's coverage of firms' earnings. Models of optimal information acquisition predict that investors (subscribers) have a countercyclical desire to increase their acquisition of payoff-relevant information.

The WSJ seems to provide more information when the additional information is more beneficial to subscribers.

Lastly, the above findings suggest that the earnings coverage of the WSJ is not a proxy for the attention of irrational retail traders. The "ostrich effect" predicts that attention is procyclical, i.e., greater when stock returns have been high. By measuring the number of logins to investment accounts, Sicherman et al. (2016) provide evidence that retail traders seek less information about their portfolios as the returns of the stock market decline, possibly because investors want to delay realizing a bad outcome.[14] Regressions like those in Table 4 with lagged excess stock returns as the dependent variable, instead of future returns, indicate that the extent of the WSJ's earnings coverage does not covary with lagged returns over either a $[-6, -1]$ or a $[-12, -1]$ window.

2.2 Extensive margin and measuring of the economic state

In the prior section, we examined how EEC covaries with several measures of the state of the economy. The fact that several measures are used speaks to the difficulty in measuring economic conditions. One just has to look at the mosaic of indicators that are commonly used to diagnose the health of the economy to recognize the complexity of filtering estimates of the underlying latent state variable. The information channel is a promising and nascent mechanism for furthering our understanding of economic signals.

For example, instability over time in many of the interrelations among macro variables is well documented (Welch and Goyal, 2008; Stock and Watson, 2003). Recognizing that information must be acquired and that agents choose when and how to gather information can potentially explain some amount of parameter instability. Bansal and Shaliastovich (2011) make such a point explicitly. They argue that jumps in stock prices coincide with investors' endogenously choosing to gather information about the state of the economy at that particular point in time. The acquired information produces a discrete change in their expectations of fundamentals, which produces a jump in prices. The price jumps are

[14]Retail trading has been linked to Google search volumes and to articles in local newspapers by Da et al. (2011) and Engelberg and Parsons (2011), respectively.

disconnected from the presumed smooth movements in underlying fundamentals. Hence, learning can generate instability in covariances between prices and fundamentals.

To provide future research with additional stylized facts to consider, we examine whether EEC provides incremental linear information about the equity premium beyond that of the set of macro variables considered in Table 3. Recall that the set of macro variables correlates sizably with the temporal variation in the extent of earnings coverage, with an R-squared of 0.41. To assess the marginal covariance of EEC with the equity risk premium, we rerun the regressions of excess stock market returns on EEC and NetEEC adding the set of eleven continuous macroeconomic variables.

The right half of Table 4 displays the results of this larger regression for the two horizons. EEC provides linear incremental information about the equity premium over months $(+1, +6)$ at a ten-percent level of significance using the adjusted p-value, but statistically provides no incremental information over the $(+1, +12)$ window. The joint test across both windows in Panel B remains significant at a 5% level. Hence, the extensive margin of the WSJ's earnings coverage offers information on the equity risk premium that is not linearly spanned by the set of macro variables. Moreover, the extensive margin seems to be a shorter-run signal of the equity premium than the macro set.

3 Adjustments along the extensive margin

To better understand the temporal variation in the measure of the WSJ's extent of earnings coverage, it first seems instructive to note the main drivers of predicted coverage. The predominant determinants of the probability of WSJ coverage for a given earnings announcement in the multinomial logit model are firm size, industry, and the number of analysts covering the firm. A reduced multinominal logit model employing only these three variables — size, industry, and analysts' coverage — produces a pseudo R^2 of 0.22, falling from 0.27 when using the full specification that is shown in the appendix section A.2.

Each of these three variables intuitively reflects a tradeoff between the costs of acquiring information and the benefits of learning from that information. Larger firms are generally more important to investors. The stocks of larger firms have greater (value) weightings in typical investment portfolios, and larger firms typically have a greater number of shareholders. Also, the scale and scope of large-firms' operations makes them more likely to be considered "bellwethers" for other firms in their industries. Relatedly, particular industries, whether cyclical or defensive, can be better indicators of the economy's condition. Additionally, some industries may contain firms whose performances are more highly correlated with each other, and hence, require fewer firm-level signals to understand. Lastly, the number of sell-side analysts that cover a firm reveals the outcomes of the cost-benefit analysis made by these research divisions in assessing whether to gather information on a particular firm. Their assessments are positively correlated with the WSJ's own assessments to cover a firm.

Figure 2 shows that the WSJ adjusts the extent of its earnings coverage over time, as the marginal benefits of reporting on firm-earnings vary. Because smaller firms are typically costlier to learn about, smaller firms are likely candidates to lie on the extensive margin of the WSJ's earnings coverage. And given that many asset-pricing anomalies are well-documented to be stronger for small firms, this dimension of the extensive margin of coverage seems fruitful to investigate.

To begin, we separate sample months into three bins based on EEC measured across all stocks. Months with EEC above the 75th percentile are labeled "high," months below the 25th percentile are labeled "low," and remaining months are labeled "normal." Each month we recalculate the extent of earnings coverage at the size-quintile level, rather than across all stocks. That is, we determine the number of firms within a given quintile of size for which the full-specification full-sample multinomial logit predicts coverage, and then form EEC–Q which measures the percentage deviation of actual coverage from predicted coverage within size quintile Q.

Panel A of Table 5 reports the mean monthly values of EEC–Q within the low, normal, and high states of coverage extensiveness. The most striking result is the dramatic increase in the extent of coverage for the smaller stocks in high-EEC months. Coverage of earnings releases for the smallest stocks in these months balloons to 91%. The extent of coverage across the remaining four quintiles falls steadily down to 18% in the largest stocks. Panel B shows the number of firm-earnings reports per month that are actually covered by the WSJ. The smallest firms receive the least amount of actual coverage, with only 10 firms on average in the high-EEC state. The coverage of just one additional firm from the smallest quintile, however, is a much larger deviation from the baseline probability of coverage. On the other end, the largest firms receive much greater coverage on average, with the actual number of firms in the highest quintile averaging 45 per month in the high-EEC state.

Analogously, in the low-EEC state, the extent of coverage recedes the least for the largest firms. In the low state, the largest firms average negative 19%, and the smallest firms average negative 37%. While the variation across the size quintiles in the low state is not as dramatic as that in the high state, nor monotonic, the extent of coverage is generally contracting as firm size declines.

Table 5 also shows that the adjustments in earnings coverage across high-coverage months to low-coverage months is strongly decreasing across the size quintiles. The smallest stocks experience a 91% increase in residual coverage in the high-EEC months and a 37% decrease in the low-EEC months. This spread between high and low states falls monotonically with firm size. Hence, the marginal coverage decisions of the WSJ, expressed as a percentage deviation from predicted coverage, are essentially driven by the coverage choices outside the largest stocks. The correlation between the full-sample EEC and the extent of coverage within the lower four size quintiles is 0.95, while the correlation between the full-sample EEC and the extent of coverage within the largest quintile only is just 0.57. In sum, the adjustments made by the WSJ to the extensiveness of its earnings coverage can be characterized as expanding and contracting over the smaller deciles of firms.

4 Extensive Margin and the Information Content of Prices

Does the expansion and contraction of earnings coverage by the WSJ covary with the informativeness of stock prices? Prices will reflect the information that investors possess.

When the mass of investors in the stock market is acquiring less information about firms' earnings, stock prices will embed less of the earnings information, all else equal. In this section, we examine how prices react to earnings surprises as the extent of the earnings-information acquisition of the WSJ changes, and then broaden the analysis to consider more general information-acquisition activities.

4.1 Post-earnings announcement drift

The tests herein consider the speed with which prices update to earnings news. One common assessment of how quickly stock prices impound earnings surprises is the post-earnings announcement drift in stock returns (hereafter "PEAD"; Bernard and Thomas (1989)). To the extent that PEAD is due to the market's slow reaction to earnings news, rather than to mismeasurement of expected returns, we expect the drift in returns to be smaller for earnings announced in months when the extensiveness of earnings coverage is greater.

Each month we sort the firms announcing earnings in that month according to their standardized unexpected earnings (SUE), defined as the announced earnings per share minus the median analyst forecast within thirty days prior to the announcement divided by the pre-announcement stock price. We then form a PEAD portfolio by taking a long position in the stocks in the highest decile of SUE and a short position in the stocks in the smallest decile of SUE. Daily abnormal returns to each stock are adjusted for size and book-to-market effects using 5×5 benchmark portfolios, with the quintile breakpoints determined from NYSE stocks only.[15] We then cumulate the abnormal daily returns (CAR) for each stock over a window beginning two days after the announcement and ending sixty days after the announcement. The CAR are then equally weighted within each month's long and short SUE portfolios respectively.

We sort calendar months over the sample period into quintiles based on EEC and then report the mean abnormal profits of the long-minus-short SUE portfolios. The left panel of Figure 4 reveals that the average PEAD profits within the smallest stocks decrease from

[15] We thank Kenneth French for providing the benchmark data on his website.

nearly 6% over the $[+2, +60]$ window for earnings surprises occurring in months with the least extensive coverage down to roughly 2% for earnings surprises occurring in months with the most extensive coverage. In short, the extensiveness of earnings coverage gets a good deal of traction in explaining temporal variation in PEAD.[16]

To assess the statistical significance of this negative relation, we employ monthly regressions. Each month we regress the cross section of cumulative abnormal stock returns over days $[+2, +60]$ on the SUE for each firm announcing earnings in that month. Then, we regress the time series of monthly cross-sectional coefficients from the first-stage regression on EEC and NetEEC. Table 6 shows the second-stage results. We see that PEAD profits for small stocks decrease with EEC at a 1% level of significance (using simulated p-values). In sum, when the extent of earnings coverage is greater, the price of a small stock immediately after the firm's earnings announcement tends to embed the earnings surprise more fully.

The strong comovement of the PEAD of small stocks with the extensiveness of earnings coverage is strikingly different from the nearly flat PEAD plot for large stocks on the right side of Figure 4. Consistent with the findings of section 3, that small stocks lie on the extensive margin, the expansion and contraction of earnings coverage by the WSJ apparently provides insights into the pricing of small stocks, but not large stocks.

Our ability to examine the PEAD of large stocks is however a bit hindered because there are months in which only a few large stocks release earnings. As a consequence, we examine quintiles of SUE to produce the right side of figure 4, rather than deciles as we do for small stocks on the left side. We also require a minimum of 10 firm-earnings releases each month for the long and short portfolios. This filter reduces the number of usable large-stock sample months from 255 to 170. As figure 4 shows, large-stock PEAD does not temporally covary with the extent of the WSJ's earnings coverage.

[16] The nonlinearity displayed by PEAD profits in quintile 3 may indicate a nonlinearity between asset pricing and information acquisition that is a byproduct of endogenizing information choices. Equilibria in these models are typically complex as they are outcomes of constrained optimizations that involve cost-benefit thresholds, corner solutions, dynamic feedback, complementarities, and other features.

We further investigate a relation between large-stock PEAD and EEC using the cross-sectional regression approach. A potential benefit with the regression approach is that we can extract more information when confronted with just a few stocks each month than is possible with the portfolio approach (since long and short stocks are pooled together each month in the regressions). Unfortunately, SUE is closer to zero for large stocks which can produce greater variability in first-stage coefficients. Although the results are not tabulated, we detect no relation between large-stock PEAD and EEC. The failure to detect a covariance between the extensiveness of earnings coverage and the return dynamics of large stocks is a pervasive finding throughout this study. In short, adjustments to the extent of earnings coverage are silent about the temporal dynamics of the returns of large stocks, but they do speak to the return dynamics of small stocks.

Finally, as done in the equity premium analysis, we provide additional stylized facts by including the set of eleven continuous macroeconomic variables. The results are reported in the second column of Table 6. With respect to PEAD, EEC also provides linearly incremental information over the span of the macro variables. This may be related to earnings information acquisition being more of a shorter-run signal for aggregate stock returns, as noted earlier.

4.2 How active is the information gathering?

The preceding findings for PEAD suggest that the state of information acquisition plays an important role in the pricing of small stocks. When EEC increases, PEAD decreases, suggesting that small-stock prices are more quickly embedding earnings news. However, is this merely a news diffusion channel or a more active state of information gathering by investors?

Prior studies have also examined the price response upon the release of earnings — the earnings response coefficient (hereafter, ERC) — in tandem with PEAD. As the information in a given earnings announcement diffuses more quickly through the marketplace, the density of the price response shifts closer to the event date. That is, studies of *firm-level* attention and the media's role in the stock market find that ERC increases with attention while PEAD decreases (e.g., Peress (2008); Hirshleifer et al. (2009); DellaVigna and Pollet (2009); Drake et al. (2015); Boulland et al. (2017)).[17]

If the market level measure of the acquisition of earnings information that we develop were solely a proxy for the speed of diffusion of the earnings surprise, we would expect EEC to covary positively with earnings response coefficients. However, recall that only eleven percent of firms that release earnings in a given month are covered on average by the WSJ, and then only a mere handful are from the smallest quintile of firms. Hence, by construction, EEC is not directly measuring the diffusion of a given small firm's SUE across investors. And since the media coverage for the smallest firms by outlets besides the WSJ is also rare, we do not expect a simple diffusion channel to be the driving mechanism in this study.

EEC may capture a more active state of information acquisition whereby some portion of the earnings surprise (measured with respect to analysts' consensus forecasts) is anticipated before the earnings announcement. If enough pre-announcement learning were occurring when EEC is high, then we may find that EEC covaries negatively with earnings response coefficients. A systematic learning channel is one plausible way that such learning can occur. In fact, there is a large literature supporting intra-industry information spillovers from one firm's earnings information to the stock prices of related firms.

In particular, Ramnath (2002) finds that the stock prices of late announcing firms respond to the earnings news of early announcing firms within the same industry. Moreover, analysts' revise their earnings forecasts for the later firms in accordance with the earnings surprise of the industry's first announcer. However, Ramnath finds that neither investors (i.e., prices) nor analysts fully respond to the earnings news of the first announcer. Additionally, Hameed et al. (2015) find that the prices of stocks without large analyst coverage

[17] A positive covariance between ERC and firm-level media coverage may also be due in part to reverse causality whereby greater price movements draw the attention of the media. Because so few stocks receive WSJ coverage, this is not a concern with the temporal covariance between the market-level EEC and stock returns.

(as would be the case for the smallest quintile of stocks) respond to the analysts' forecast revisions of bellwether firms in their industries.

In this cross-firm learning environment, an increase in earnings information spillovers from early announcers to late announcers would reduce the price impacts of the later earnings surprises, as a systematic portion of the later firms' earnings surprise would be revealed through the earlier industry news that is extracted from the early announcers' earnings. Moreover, Loh and Stulz (2018) find some evidence that spillovers of earnings information into other firms' stock prices are stronger in bad economic times.[18] Their findings suggest that information spillovers would be greater when EEC is high, since we find in section 2.1 that EEC moves countercyclically.

This systematic-learning scenario is also consistent with the theoretical predictions of Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016) in which investors allocate more attention in bad times to gathering common information relative to firm-specific information. While we focus in this study on the acquisition of earnings information, this does not exclude that investors may be acquiring information on the payoffs of groupings of firms from signals other than earnings releases. In sum, when investors are sufficiently more active in their learning about the systematic components of earnings surprises, the price response to earnings news upon the announcement will decline.[19]

4.3 Earnings Response Coefficients

To provide more of a gauge on how actively investors are learning about earnings and embedding that information into small-stock prices, we examine the event-window reactions of stock prices to earnings surprises. We established in section 4.1 that PEAD declines with EEC, indicating that prices immediately after earnings announcements are more fully reflecting earnings surprises in months when the extent of earnings coverage by the WSJ

[18] Loh and Stulz (2018) examine the price reactions of stocks to the revisions of analysts' buy/sell recommendations of related stocks, not to revisions of earnings forecasts per se.

[19] Admittedly, we do not empirically eliminate the alternative channel whereby investors are learning more about firm-specific components of earnings in anticipation of a given firm's earnings announcement, and that this one-off learning occurs with such extent across subsets of investors that the average informativeness of small-stock prices increases. However, this channel runs counter to the literature noted above, and seems tenuous to us based on the inefficiency of such a learning process.

increases. As discussed above, an indicator of an even more active state of learning by investors would be finding that earnings response coefficients decrease when the extent of coverage is greater, as prices better anticipate upcoming earnings releases.

Again we measure earnings surprises relative to the consensus of analysts' forecasts. Each month, we regress the cross section of cumulative abnormal returns for a stock over days $[-1, +1]$ on the stock's SUE. The time series of the mean regression coefficients from each month (i.e. the mean monthly ERC) is then regressed on EEC and NetEEC.

The right-side of Table 6 reports the second-stage results. Price reactions to earnings surprises are lower in the months when EEC is higher, at a five-percent level of significance (using simulated p-values). This finding is inconsistent with increases in EEC being solely associated with an increase in the diffusion of SUE across the marketplace. Instead, the ERC finding indicates that investors are actively learning about earnings, and presumably through more systematic channels.

Interestingly, the last column of Table 6 also shows that EEC has no incremental information for ERC above that contained in the linear span of the macro variables. This qualitatively differs from the equity premium and PEAD findings (and the remaining findings still to be presented). As noted before, the structure of the interrelations among the macro variables and information acquisition activity is an avenue for future research. Perhaps the ERC relation is less subject to nonlinearities, or perhaps this is a spurious failure to reject linear spanning in one of five tests.

Regardless, the extent of the WSJ's earnings coverage covaries negatively with the variation in earnings response coefficients, implying that EEC captures times when investors are more actively learning, and hence, when prices better anticipate earnings news. In the next sections, we investigate another aspect of small-stock returns to further examine whether the informativeness of stock prices increases with EEC.

4.4 Return Dispersion

An underpinning of information models is that an agent's beliefs are updated when information is acquired, and in this study, we are considering updates to stock valuations. Our previous findings for ERC and PEAD are consistent with investors acquiring more information about the earnings of small stocks when the WSJ is increasing the extent of its earnings coverage. With more information being impounded into prices, the cross-sectional dispersion of returns should be greater as the valuations for a greater number of firms are being updated.[20] Hence, we examine the relation between stock-return dispersion and EEC.

We begin by sorting months into quintiles based on EEC. For each quintile, the left panel of Figure 5 reports the mean monthly standard deviation of the cross section of returns for firms in the smallest quintile of market value of equity. We see that return dispersion among small stocks increases notably with EEC, from about 19% per month in the months with the lowest EEC to 24% in the months with the greatest EEC. To assess statistical significance, we regress the monthly time-series of return dispersion on EEC and NetEEC. We see in Table 7 that the relation between EEC and return dispersion in small stocks is statistically strong, with adjusted p-values below one percent.

For completeness, both Figure 5 and Table 7 examine return dispersion among the largest stocks. The range of return dispersion for large stocks in Figure 5 is from a little below 8% per month to a little above 9%. And Table 7 confirms a flat relation between EEC and large-stock return dispersion. This is consistent once again with the marginal adjustments of earnings coverage mapping into the pricing of smaller stocks only.

In sum, the increased return dispersion of small stocks when EEC is greater is consistent with an information channel. Investors seem to be acquiring more information about small stocks in months when the WSJ is expanding its earnings coverage to more firms, resulting in price adjustments that are larger than in other months.

[20] Any heterogeneity across investors' signals (and beliefs) would seemingly amply this effect.

5 Conditional Performance of Mutual Funds

Results of the prior sections combine to suggest that EEC captures periods when investors are more actively updating stock prices to better reflect information about earnings. In this section, we examine the trading performance of a particular group of investors — mutual funds — conditional on the extent of earnings information acquisition. We are motivated by two elements of the literature. First, mutual funds have been found to possess skill at forecasting earnings. Baker et al. (2010) find that the buying and selling of stocks by mutual funds leads future earnings surprises. Jiang and Zheng (2018) find that the ability to forecast earnings surprises varies across funds and that some funds are persistently better at this than others. Our second motivation comes from Kacperczyk et al. (2014, 2016) who argue that funds allocate more information-gathering capacity during recessions to learning about common payoffs than about firm-specific payoffs, and that this optimal behavior provides a potential explanation for why the performance of mutual funds on average is greater during recessions (Moskowitz, 2000; Glode, 2011). Since our earlier finding for earnings response coefficients is consistent with more systematic small-stock information being used by investors in times when EEC is high, and that EEC moves countercyclically, it seems natural to test whether mutual funds small-stock picks perform better when EEC is greater.[21]

Following Edelen, Ince, and Kadlec (2016), we estimate the change in aggregate mutual fund demand for a given stock as the change in the number of funds that are holding that stock from last quarter to this quarter. This measure aggregates across the individual fund-level buy/sell signals, ignoring the sizing of the changes which would otherwise weight the aggregation toward reflecting the signals of the largest funds. Also, this measure captures only entry and exit decisions of funds, which reasonably reveal more misvaluation-based trading than adjustments to ongoing holdings do (Baker et al., 2010).

[21] As Kacperczyk et al. (2016) note, a long literature seeks to understand what rents are extracted in the money-management industry, particularly when the ability to generate positive alpha for clients unconditionally seems weak in the historical sample. For example, see the recent studies by Berk and van Binsbergen (2015) and Pastor, Stambaugh, and Taylor (2017).

We collect data from Thomson Reuters on the quarterly stock holdings of actively managed domestic equity mutual funds.[22] Following the methodology of Edelen, Ince, and Kadlec (2016), we scale the quarterly change in the number of funds that hold a given stock by the mean number of mutual funds that are holding the stocks of small firms at the end of the prior quarter. This scaling captures the rate of change in the number of funds that hold a stock based on the typical number of funds that are holding a similar stock, which facilitates cross-sectional and temporal comparisons. Finally, the scaled measure is winsorized quarterly at the 1% and 99% values.

Because changes in fund holdings are quarterly, we recalculate EEC and NetEEC at quarterly intervals so that the state of earnings information acquisition will match the horizon over which we measure the net signals and trades of funds. To maintain power for our tests, we still use monthly observations for stock returns. To avoid spurious rejections, we again employ simulations to adjust the p-values for the serial autocorrelations in EEC and NetEEC (and the macro variables when included). Stocks with prices below five dollars at the end of the quarter are removed.

Our assessment of how well mutual funds are trading small stocks begins by estimating a cross-sectional regression of a given month's stock returns (adjusted for size and book-to-market effects using 5×5 matched portfolios) on the most recent calendar quarter's change in the number of funds holding each stock. The time series of monthly first-stage coefficients is then regressed on EEC measured over the prior quarter, contemporaneous with the observed changes in quarterly fund holdings. That is, this exercise examines the quarter $t + 1$ cross-sectional performance of the small-stock buy/sell signals processed by mutual funds in quarter t, conditional on the information state in quarter t.

[22] We identify domestic equity funds using Lipper Prospectus objective codes equal to EI, EIEI, ELCC, G, GI, LCCE, LCGE, LCVE, LSE, MC, MCCE, MCGE, MCVE, MLCE, MLGE, MLVE, MR, S, SCCE, SCGE, SCVE, SESE, SG, or Strategic Insight objective codes equal to AGG, GMC, GRI, GRO, ING, SCG, or Weisenberger objective codes equal to GCI, IEQ, IFL, LTG, MCG, SCG, G, G-I, G-I-S, G-S, G-S-I, GS, I, I-G, I-G-S, I-S, I-S-G, S, S-G-I, S-I, S-I-G. We remove funds ever having an index fund flag of D in the CRSP dataset. Only holdings reported for the ends of calendar quarters are used. Funds with TNA less than ten million dollars are removed. Shares are adjusted for stock splits occurring between the report date and the filing date.

Panel A of Table 8 reports the regression results. The trading performance of mutual funds within the small-stock universe increases with EEC, at the five-percent level of significance (using simulated p-values). To provide an economic interpretation of this relation, we sort small stocks each quarter into quintiles based on changes in the number of funds holding a given stock. Stocks in the highest quintile are held long, and stocks in the lowest quintile are held short. Monthly abnormal profits of long-short portfolios over each of the next three months (one quarter) are calculated. We then compare the profits across high and low states of the information environment, sorting calendar quarters based on the median value of EEC over the sample period. Finally, the time series of long-short profits is filtered through a four-factor model of returns using Mkt-RF, SMB, HML, and UMD, allowing for loadings to vary across information states.[23]

Panel B of Table 8 reports the four-factor alphas of the long-short portfolios across information states. The mean monthly abnormal profit of this mutual-fund-trade-mimicking portfolio is 17 basis points per month when the portfolio is formed in quarters with high EEC, and −0.33 basis points per month when formed in quarters with low EEC. These means differ at a five-percent level of significance (t=1.95, not tabulated). The sizable spread in alpha across states is impressive, revealing that the information environment matters for our understanding of mutual fund performance.

This finding extends those of Kacperczyk et al. (2016) by linking mutual fund performance to a direct measure of information acquisition activity. When the extensive margin of earnings information acquisition expands to include more firms, and hence small-stock prices better reflect the earnings of firms, mutual funds perform better. As Kacperczyk et al. (2016) note, investing skill in mutual funds has been difficult to detect unconditionally. Our findings are consistent with some mutual funds having skill. They are also consistent with theories whereby the return performance of investors having true skill at fundamentally evaluating stocks is impeded by the trading of uninformed, noise investors (e.g., Shleifer and Vishny (1997)).

[23]We thank Kenneth French for providing these data on his website.

Finally, we observe a couple more stylized facts surrounding EEC. First, Panel A of Table 8 shows that the covariation between EEC and mutual-fund performance is incremental to the linear span of the eleven macro variables. With respect to the mutual fund performance literature, therefore, conditioning on the state of the macroeconomy is insufficient at capturing the state of information acquisition activity, at least for the extensive margin of earnings information. Including the NBER recession indicator as a twelfth explanatory variable does not alter this conclusion. Also, the long-short portfolio only loads on the UMD factor, and this loading varies across states, declining from 0.20 in the low EEC state to 0.10 in the high EEC state (at a ten-percent level of significance, not tabulated). Momentum trading may be related to an information channel, with momentum trading declining as prices better reflect fundamental values.

6 Conclusion

Measuring the state of the information environment in the marketplace is an important avenue for future research. Obtaining empirical measures of information acquisition, however, has long been an obstacle. Using the observable signals of financial agents to serve as a proxy for dimensions of the information state seems promising. In this study, we empirically measure the extensive margin of the WSJ's coverage of earnings, i.e., how many firms in a given month the WSJ chooses to cover out of the many hundreds that release their earnings. This is the first empirical investigation of the extensive margin that we are aware of. This extensive margin moves countercyclical to the economy and covaries with a number of return dynamics of small stocks. The traction this measure gets in both the level and cross-section of stock returns is notable.

When information acquisition extends to more firms, we find that the stock prices of small firms better reflect earnings information. Conditioning on this dimension of the information state of the marketplace reveals that some mutual funds possess skill at evaluating stock fundamentals. The unconditional performance of mutual funds cannot detect this skill.

A Appendix

A.1 Variable definitions

A.1.1 Firm and Earnings Characteristics

Earnings announcement dates, actual earnings, and analysts' forecasts of earnings are from I/B/E/S. Stock prices, returns, and trading volume are from CRSP. Book value of equity is from the CRSP/Compustat Merged Database. Institutional ownership is obtained from Thomson Reuters.

UE quantiles are indicator variables. Each quarterly earnings release is assigned to one of 11 quantiles based on its earnings surprise, where the surprise is the announced EPS minus the median analyst earnings forecast within 30 days prior to the announcement normalized by the closing stock price two days prior. Quantiles 1 to 5 rank the negative surprise announcements into equal-sized quintiles. Quantile 6 consists of zero-surprise announcements where announced earnings equal the median of analysts' forecasts. Quantiles 7 to 11 rank the positive surprise announcements into equal-sized quintiles. Indicator variables are formed for each quantile other than the zero-surprise quantile 6, which serves as the base case.

Loss is a dummy variable equal to 1 if the announced earnings is negative.

$Stdev(analysts'\ forecasts)$ is the standard deviation of analysts' EPS forecasts during the 30 calendar days prior to the announcement, with each forecast normalized by the closing stock price two days prior to the announcement. At least two covering analysts are required.

$log(analysts'\ coverage)$ is the natural logarithm of one plus the number of distinct analysts' forecasts observed over the 30 calendar days prior to the announcement.

$log(ME)$ is the natural logarithm of the number of shares outstanding multiplied by the firm's closing stock price two days prior to the announcement.

$log(value\ of\ trading)$ is the natural logarithm of a stock's mean dollar value of daily trading volume over the 60 trading days prior to the announcement.

$Beta$ is the estimated coefficient from a regression of a firm's daily stock return on the S&P 500 return over the 60 days prior to the announcement.

$Recent\ returns$ is a stock's mean daily return over the 60 trading days prior to the announcement.

$Stdev(recent\ returns)$ is the standard deviation of a stock's daily returns over the 60 trading days prior to the announcement.

BE/ME is the firm's book value of equity from the fiscal year ending in the previous calendar year divided by its market value of equity from December 31 minus the value-weighted average book-to-market ratio of all announcing firms over the rolling three month period ending in the current month.

$Distraction$ is the announcement day's decile rank (in a given calendar quarter) based on the number of earnings announcements released from other firms on the same day. See Hirshleifer et al. (2009).

$Institutional\ ownership$ is the percentage of shares held by institutions at the end of the previous calendar year obtained from 13F filings.

$Seasonality\ and\ industry\ dummies$ are three indicator variables for month-of-the-year, day-of-the-week, and the 49 Fama-French industries, respectively. Firms are assigned to industries using SIC codes from Compustat.

A.1.2 Macroeconomic Variables

$PDND$ is the value-weighted dividend premium from Baker and Wurgler (2004). (Downloaded from Jeffrey Wurgler's website.)

$log(Div.\ Yield)$ is the natural logarithm of the market dividend yield (aggregate dividends for months t to $t-11$, divided by total market capitalization in month t for NYSE, AMEX, and Nasdaq firms). (Downloaded from Michael Roberts' website.)

$log(Net\ Payout\ Yield)$ is the natural logarithm of the total net payout yield (where the equity issuance yield is aggregate net equity issues for months t to $t-11$, divided by total market capitalization in month t). (See Boudoukh et al. (2007). Downloaded from Michael Roberts' website.)

$Risk\text{-}free\ rate$ is the US 90-day T-Bill rate. (Downloaded from Michael Roberts' website.)

CAY is the estimated quarterly deviation from the long-run log aggregate consumption wealth ratio. (See Lettau and Ludvigson (2001). Downloaded from Sidney Ludvigson's website.)

B/M is the book value of equity divided by its market value for the Dow Jones Industrial Average. (Downloaded from Amit Goyal's website.)

$Default\ Spread$ is the difference between the BAA and AAA corporate bond yields from FRED. (Downloaded from Amit Goyal's website.)

$Term\ Spread$ is the yield on the 10-year Treasury bond minus the yield on the 3-month Treasury bill. (Downloaded from Amit Goyal's website.)

$Equity\ Share\ of\ New\ Issues$ is the dollar amount of equity new issues divided by the dollar amount of total new issues (debt plus equity) described in Baker and Wurgler (2000). (Downloaded from Jeffrey Wurgler's website.)

$Sentiment$ is the sentiment index from Baker and Wurgler (2006), which is based on the principal component of 6 sentiment proxies. (Downloaded from Jeffrey Wurgler's website.)

Output Gap is the estimated residual from a linear and quadratic monthly time trend in the natural logarithm of US Industrial Production over the sample period. See Cooper and Priestley (2009).

A.2 Logit Model Results

Table A.1 Predicting WSJ coverage

Below are the estimated coefficients from the multinomial logit regression of $c_{i,t} \in (-1, 0, 1)$ on various explanatory variables, where $c_{i,t}$ equals -1 when the earnings report for firm i in month t receives negative coverage, 1 when it receives nonnegative coverage, and 0 when it receives no coverage (the base case). Variable definitions are in section A.1. The t-statistics (in parentheses) are computed using standard errors clustered by firm; * indicates significance at 10%; ** indicates significance at 5%;*** indicates significance at 1%.

	WSJ Coverage	
	-1	1
UE quantile 11	0.805***	0.530***
	(11.12)	(7.52)
UE quantile 10	0.411***	0.336***
	(6.40)	(6.05)
UE quantile 9	0.113*	0.248***
	(1.74)	(4.96)
UE quantile 8	-0.113*	0.137***
	(-1.77)	(3.00)
UE quantile 7	-0.462***	0.00240
	(-6.77)	(0.06)
UE quantile 5	0.112*	-0.0185
	(1.76)	(-0.39)
UE quantile 4	0.603***	-0.000930
	(9.47)	(-0.02)
UE quantile 3	0.916***	0.186***
	(14.22)	(3.04)
UE quantile 2	1.048***	0.260***
	(14.91)	(3.56)
UE quantile 1	1.272***	0.367***
	(16.05)	(3.32)
Loss dummy	1.041***	-1.449***
	(19.01)	(-17.63)
Stdev(analysts' forecasts)	0.0624***	-0.0905**
	(4.75)	(-2.53)
log(analysts' coverage)	0.410***	0.356***
	(10.57)	(8.18)
log(ME)	0.442***	0.598***
	(12.38)	(14.95)
log(value of trading)	0.226***	0.210***
	(7.82)	(6.49)
Beta	-0.0986***	-0.0146
	(-4.16)	(-0.55)

Recent returns	-0.00226***	-0.00177***
	(-4.55)	(-3.47)
Stdev(recent returns)	0.640	-10.01***
	(0.53)	(-6.10)
BE/ME	0.570***	0.339***
	(20.79)	(8.47)
Distraction	-0.0346***	-0.0637***
	(-2.99)	(-5.13)
Seasonality and industry (FF-49) dummies	Yes	Yes
Observations	233348	
Pseudo R^2	0.269	

References

Ahern, Kenneth R., and Denis Sosyura, 2014, Who writes the news? Corporate press releases during merger negotiations, *The Journal of Finance* 69, 241–291. [page 3], [page 9]

Andrei, Daniel, and Michael Hasler, 2015, Investor attention and stock market volatility, *Review of Financial Studies* 28 (1), 33–72. [page 2], [page 11]

Baker, Malcolm, Lubomir Litov, Jessica A. Wachter, and Jeffrey Wurgler, 2010, Can mutual fund managers pick stocks? evidence from their trades prior to earnings announcements, *Journal of Financial and Quantitative Analysis* 45, 1111–1131. [page 5], [page 26]

Baker, Malcolm, and Jeffrey Wurgler, 2000, The equity share in new issus and aggregate stock rteurns, *Journal of Finance* 55, 2219–2257. [page 32]

Baker, Malcolm, and Jeffrey Wurgler, 2004, A catering theory of dividends, *Journal of Finance* 59, 1125–1165. [page 31]

Baker, Malcolm, and Jeffrey Wurgler, 2006, Investor sentiment and the cross-section of stock returns, *Journal of Finance* 61 (4), 1645–1680. [page 12], [page 32], [page 44]

Bansal, Ravi, and Ivan Shaliastovich, 2011, Learning and asset-price jumps, *Review of Financial Studies* 24 (8), 2738–2780. [page 2], [page 11], [page 15]

Berk, Jonathan B., and Jules H. van Binsbergen, 2015, Measuring skill in the mutual fund industry, *Journal of Financial Economics* 118, 1 – 20. [page 26]

Bernard, Victor L, and Jacob K Thomas, 1989, Post-earnings-announcement drift: delayed price response or risk premium?, *Journal of Accounting Research* 27, 1–36. [page 19]

Boudoukh, Jacob, Roni Michaely, Matthew Richardson, and Michael R. Roberts, 2007, On the importance of measuring payout yield: Implications for empirical asset pricing, *Journal of Finance* 62 (2), 877–915. [page 32]

Boudoukh, Jacob, Matthew Richardson, and Robert F. Whitelaw, 2008, The myth of long-horizon predictability, *Review of Financial Studies* 21 (4), 1577–1605. [page 10], [page 14]

Boulland, Romain, François Degeorge, and Edith Ginglinger, 2017, News dissemination and investor attention, *Review of Finance* 21, 761–791. [page 5], [page 22]

Campbell, John Y., 2018, *Financial decisions and markets : A course in asset pricing* (Princeton University Press, Princeton). [page 12]

Cooper, Ilan, and Richard Priestley, 2009, Time-varying risk premiums and the output gap, *Review of Financial Studies* 22 (7), 2801–2833. [page 12], [page 32]

Da, Zhi, Joseph Engelberg, and Pengjie Gao, 2011, In search of attention, *Journal of Finance* 66 (5), 1461–1499. [page 15]

DellaVigna, Stefano, and Joshua M. Pollet, 2009, Investor inattention and Friday earnings announcements, *Journal of Finance* 64 (2), 709–749. [page 22]

Drake, Michael S., Darren T. Roulstone, and Jacob R. Thornock, 2015, The determinants and consequences of information acquisition via edgar, *Contemporary Accounting Research* 32, 1128–1161. [page 22]

Edelen, Roger M., Ozgur S. Ince, and Gregory B. Kadlec, 2016, Institutional investors and stock return anomalies, *Journal of Financial Economics* 119, 472 – 488. [page 6], [page 26], [page 27]

Engelberg, Joseph E.., and Christopher A. Parsons, 2011, The causal impact of media in financial markets, *Journal of Finance* 66 (1), 67–97. [page 15]

Fang, Lily, and Joel Peress, 2009, Media coverage and the cross-section of stock returns, *The Journal of Finance* 64, 2023–2052. [page 3], [page 9]

Gaa, Charles Clyde, 2008, *Media coverage and investor attention*, Ph.D. thesis, University of British Columbia. [page 7]

Glode, Vincent, 2011, Why mutual funds "underperform", *Journal of Financial Economics* 99, 546 – 559. [page 5], [page 26]

Gurun, Umit G., and Alexander W. Butler, 2012, Don't believe the hype: Local media slant, local advertising, and firm value, *The Journal of Finance* 67, 561–598. [page 3], [page 8]

Hameed, Allaudeen, Randall Morck, Jianfeng Shen, and Bernard Yeung, 2015, Information, analysts, and stock return comovement, *The Review of Financial Studies* 28, 3153–3187. [page 2], [page 5], [page 22]

Hirshleifer, David, Sonya Seongyeon Lim, and Siew Hong Teoh, 2009, Driven to distraction: Extraneous events and underreaction to earnings news, *Journal of Finance* 64 (5), 2289–2325. [page 5], [page 22], [page 31]

Jiang, Hao, and Lu Zheng, 2018, Active Fundamental Performance, *The Review of Financial Studies* 31, 4688–4719. [page 5], [page 26]

Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp, 2014, Time-varying fund manager skill, *The Journal of Finance* 69, 1455–1484. [page 26]

Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp, 2016, A rational theory of mutual funds' attention allocation, *Econometrica* 571–626. [page 1], [page 2], [page 5], [page 11], [page 23], [page 26], [page 28]

Lettau, Martin, and Sydney Ludvigson, 2001, Consumption, aggregate wealth, and expected stock returns, *Journal of Finance* 56, 815–849. [page 4], [page 12], [page 32], [page 39], [page 44]

Lettau, Martin, and Sydney Ludvigson, 2002, Time-varying risk premia and the cost of capital: An alternative implication of the q theory of investment, *Journal of Monetary Economics* 49, 31 – 66. [page 12]

Loh, Roger K., and Rene M. Stulz, 2018, Is sell-side research more valuable in bad times?, *The Journal of Finance* 73, 959–1013. [page 5], [page 23]

McCallum, Andrew K., 1996, Bow: A toolkit for statistical language modeling, text retrieval, classification and clustering, http://www.cs.cmu.edu/mccallum/bow. [page 7]

Moskowitz, Tobias J., 2000, Discussion, *The Journal of Finance* 55, 1695–1703. [page 26]

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 703–708.

Pastor, Lubos, Robert F. Stambaugh, and Lucian A. Taylor, 2017, Do funds make more when they trade more?, *The Journal of Finance* 72, 1483–1528. [page 26]

Patton, Andrew J., and Michela Verardo, 2012, Does beta move with news? firm-specific information flows and learning about profitability, *Review of Financial Studies* 25 (9), 2789–2839. [page 2]

Peng, Lin, and Wei Xiong, 2006, Investor attention, overconfidence and category learning, *Journal of Financial Economics* 80, 563 – 602. [page 1]

Peress, Joel, 2008, Media coverage and investors' attention to earnings announcements, working paper. [page 22]

Ramnath, Sundaresh, 2002, Investor and analyst reactions to earnings announcements of related firms: An empirical analysis, *Journal of Accounting Research* 40, 1351–1376. [page 2], [page 5], [page 22]

Savor, Pavel, and Mungo Wilson, 2016, Earnings announcements and systematic risk, *The Journal of Finance* 71, 83–138. [page 2]

Shleifer, Andrei, and Robert W. Vishny, 1997, The limits of arbitrage, *The Journal of Finance* 52, 35–55. [page 28]

Sicherman, Nachum, George Loewenstein, Duane J. Seppi, and Stephen P. Utkus, 2016, Financial attention, *Review of Financial Studies* 29, 863–897. [page 15]

Solomon, David, and Eugene Soltes, 2012, Managerial control of business press coverage, working paper. [page 3], [page 9]

Solomon, David H., 2012, Selective publicity and stock prices, *The Journal of Finance* 67, 599–638. [page 3], [page 8], [page 9]

Stambaugh, Robert F., Jianfeng Yu, and Yu Yuan, 2012, The short of it: Investor sentiment and anomalies, *Journal of Financial Economics* 104, 288 – 302, Special Issue on Investor Sentiment. [page 13]

Stock, James H., and Mark W. Watson, 2003, Forecasting output and inflation: The role of asset prices, *Journal of Economic Literature* 41, 788–829. [page 15]

Tetlock, Paul C., 2007, Giving content to investor sentiment: The role of media in the stock market, *Journal of Finance* 62 (3), 1139–1168. [page 3], [page 8]

Tetlock, Paul C., Maytal Saar-Tsechansky, and Sofus Macskassy, 2008, More than words: Quantifying language to measure firms' fundamentals, *Journal of Finance* 63 (3), 1437–1467. [page 3], [page 8]

Welch, Ivo, and Amit Goyal, 2008, A comprehensive look at the empirical performance of equity premium prediction, *Review of Financial Studies* 21 (4), 1455–1508. [page 15]

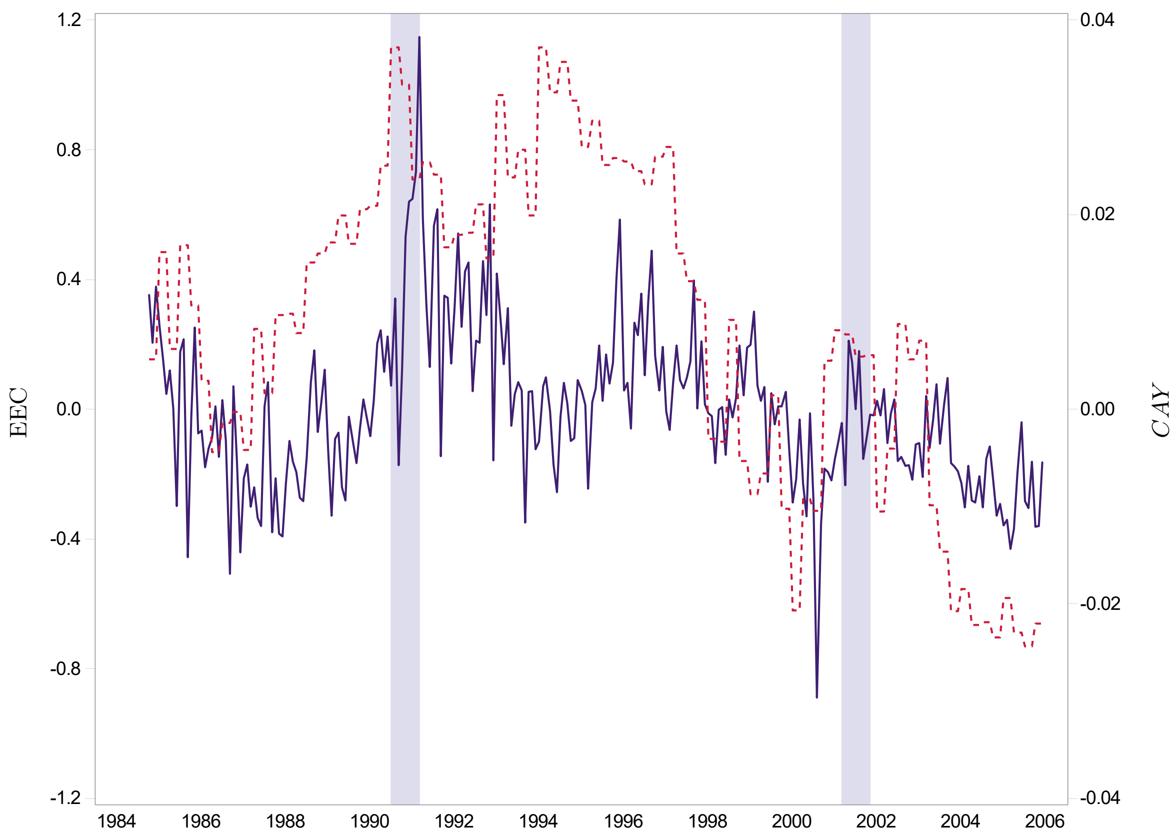

Figure 1: Extent of Coverage Covaries With CAY. The monthly EEC series (solid line) is plotted against the quarterly CAY series (dotted line). EEC is orthogonalized with respect to tone (NetEEC). CAY is the consumption-wealth ratio of Lettau and Ludvigson (2001). The shaded regions are NBER recession periods.

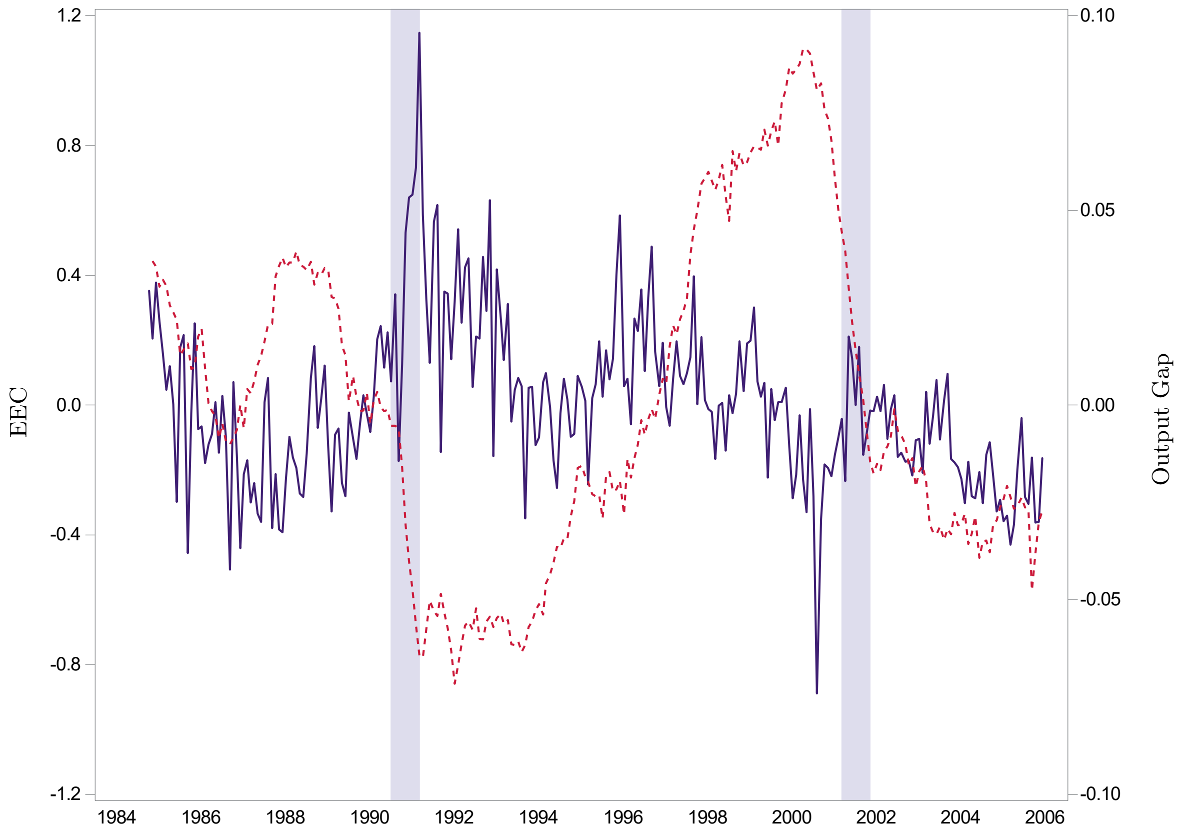

Figure 2: Extent of Coverage Covaries Negatively With Output Gap. The monthly EEC series (solid line) is plotted against the monthly output gap series (dotted line). EEC is orthogonalized with respect to tone (NetEEC). Output gap is the residual from a linear and quadratic time trend in the natural logarithm of industrial production. The shaded regions are NBER recession periods.

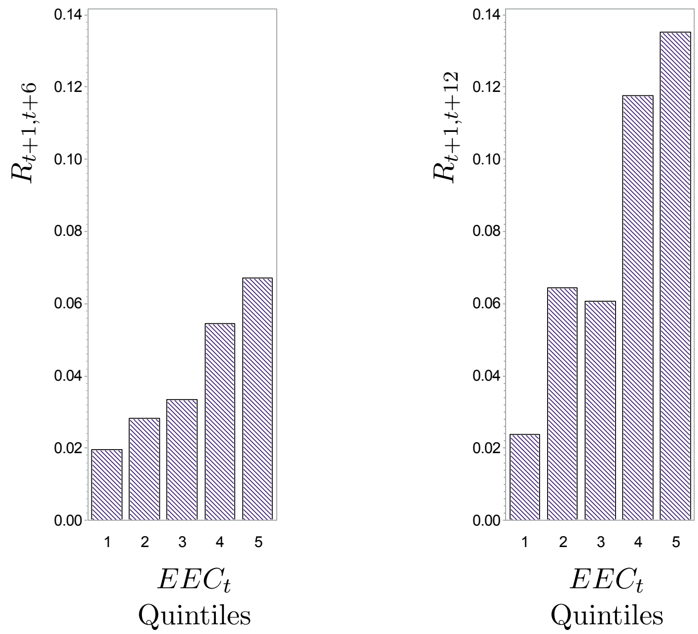

Figure 3: Extent of Coverage and Equity Risk Premium. Months are sorted into quintiles based on EEC, orthogonalized with respect to tone (NetEEC). Means of excess stock returns within each quintile are plotted for months (+1,+6) in the left panel and for months (+1,+12) in the right panel.

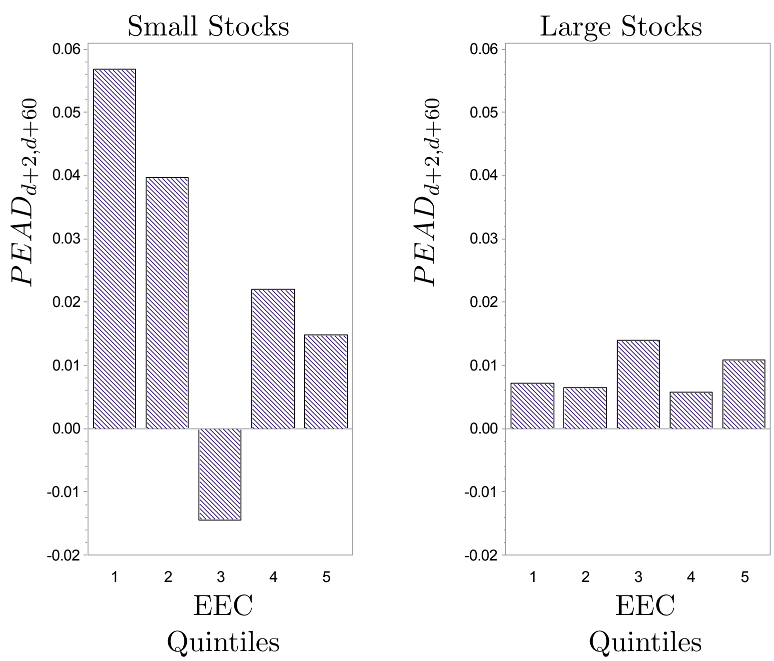

Figure 4: Extent of Coverage and Post-Earnings Announcement Drift in Returns. Months are sorted into quintiles based on EEC, orthogonalized with respect to tone (NetEEC). Means of PEAD profits within small stocks for days (+2,+60) are plotted in the left panel ($\leq 20^{th}$ percentile using NYSE breakpoints) and within large stocks in the right panel ($> 80^{th}$ percentile using NYSE breakpoints).

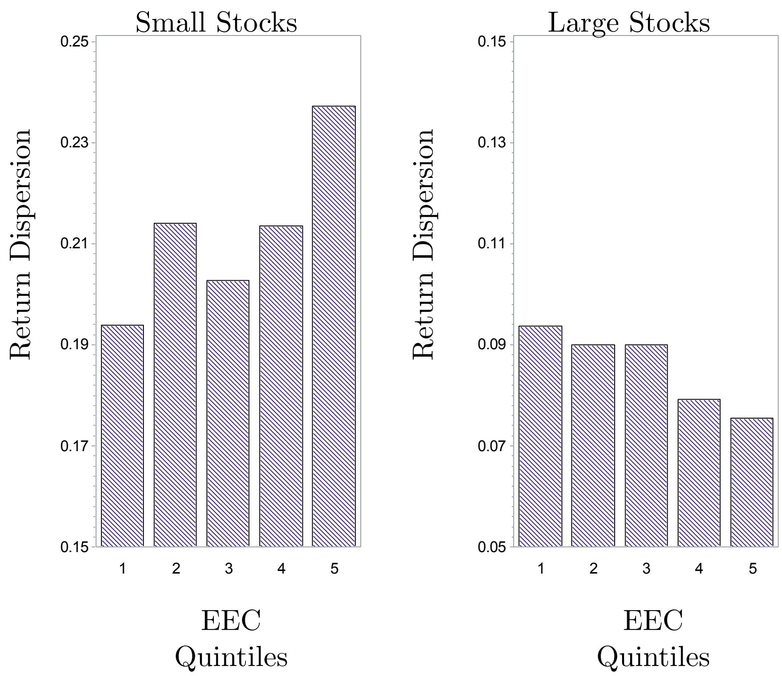

Figure 5: Extent of Coverage and Return Dispersion. Months are sorted into quintiles based on EEC, orthogonalized with respect to tone (NetEEC). Means of monthly cross-sectional standard deviation of returns within small stocks are plotted in the left panel ($\leq 20^{th}$ percentile of market cap using NYSE breakpoints) and within large stocks in the right panel ($> 80^{th}$ percentile of market cap using NYSE breakpoints). Different axis scales are used for small and large stocks.

Table 1
Summary Statistics

Below are summary statistics for various monthly measures of the WSJ's coverage of earnings and of macroeconomic measures from October 1984 to December 2005. The expected number of nonnegative articles and of negative articles are estimated with a multinomial logit (see section 1.3). EEC is the deviation of the actual number of firm-earnings covered in a given month from the expected number to be covered, while NetEEC is the deviation of the actual net tone of the coverage from the expected net tone (see equations 1 and 2). Equity premium is the monthly return on the CRSP value-weighted index minus the one-month T-Bill return in percentage terms. The macroeconomic measures in the lower portion are defined in Appendix A.1.

	Mean	Std. Dev.	AR(1)
Number of Earnings Released	915.09	706.12	−0.14
Number of Earnings Covered	96.60	79.56	−0.17
Number of Nonnegative Articles	66.39	59.99	−0.18
Number of Negative Articles	30.22	24.09	0.00
E[Number of Nonnegative Articles]	66.39	60.34	−0.16
E[Number of Negative Articles]	30.22	22.51	−0.01
Extent of Earnings Coverage (EEC)	0.03	0.26	0.62
NetEEC	0.00	0.19	0.43
Equity Premium	0.45	4.38	0.04
PDND	−0.12	0.11	0.93
Dividend Yield	0.02	0.01	0.99
Net Payout Yield	0.10	0.02	0.99
Risk-free rate	0.05	0.02	0.98
CAY	0.01	0.02	0.98
B/M	0.33	0.15	0.98
Default Spread	0.01	0.00	0.95
Term Spread	1.76	1.15	0.97
Equity Share of New Issues	0.12	0.06	0.61
Sentiment	0.10	0.59	0.94
Output Gap	0.00	0.04	0.99

Table 2
Correlations

EEC and NetEEC are each orthogonalized with respect to each other (except for their pairwise correlation). *Recession* is an NBER recession dummy. *CAY* is the deviation of the consumption-wealth ratio from its long-term trend sampled quarterly (Lettau and Ludvigson (2001)). Output gap is the residual from a linear and quadratic time trend in log U.S. industrial production. Term spread is the month-end 10-year Treasury yield minus the 3-month Treasury yield. Sentiment is from Baker and Wurgler (2006). Standard p-values are listed below each correlation coefficient.

	EEC	NetEEC	Recession	CAY	Output Gap	Term Spread	log(Div. Yield)	Sentiment
EEC	1.000							
NetEEC	0.302	1.000						
	(0.00)							
Recession	0.261	−0.223	1.000					
	(0.00)	(0.00)						
CAY	0.482	−0.054	0.154	1.000				
	(0.00)	(0.39)	(0.01)					
Output Gap	−0.272	0.336	−0.069	−0.271	1.000			
	(0.00)	(0.00)	(0.27)	(0.00)				
Term Spread	0.117	−0.212	0.002	0.058	−0.557	1.000		
	(0.06)	(0.00)	(0.97)	(0.35)	(0.00)			
log(Dividend Yield)	0.197	−0.083	-0.006	0.496	−0.345	0.238	1.000	
	(0.00)	(0.19)	(0.92)	(0.00)	(0.00)	(0.00)		
Sentiment	−0.174	0.049	0.216	−0.156	0.561	−0.084	−0.297	1.000
	(0.01)	(0.44)	(0.00)	(0.01)	(0.00)	(0.15)	(0.00)	

Table 3

Extent of Earnings Coverage and Macro Conditions

EEC is regressed on a set of macroeconomic variables, which are defined in section A.1.2 of the Appendix. The t-statistics (in parentheses) are calculated using Newey-West standard errors with six lags. The p-values shown in brackets are determined using 10,000 randomly generated samples of independent normally distributed variables with first-order serial correlation coefficients equal to that of EEC. Asterisks correspond to the simulated p-values, with * indicating significance at 10%,** indicating significance at 5%, and *** indicating significance at 1%.

PDND	0.004	Default Spread	3.709
	(1.80)		(0.37)
	[0.21]		[0.80]
log(Div. Yield)	-0.977^{***}	Term Spread	0.064
	(-4.96)		(1.98)
	[0.00]		[0.15]
log(Net Payout Yield)	-0.052	Equity Share of New Issues	0.435
	(-0.22)		(1.57)
	[0.87]		[0.24]
RF	11.037^{**}	Sentiment	-0.112
	(3.12)		(-1.92)
	[0.02]		[0.21]
CAY	7.523^{***}	Output Gap	-2.959^{**}
	(4.56)		(-2.92)
	[0.00]		[0.04]
B/M	0.927	Recession	0.164
	(2.14)		(1.92)
	[0.16]		[0.26]
Constant	-4.657		
	(-4.85)		
Observations	255		
R^2	0.41		

Table 4

Extent of Earnings Coverage and the Equity Premium

The dependent variable is the log return of the CRSP value-weighted index minus the one-month T-Bill rate across the future six and twelve months respectively. EEC_t and $NetEEC_t$ are the month t measures of the extent of coverage and of the tone of coverage, respectively, defined in equations (1) and (2). Both of these explanatory variables are standardized. The row labeled "Macro Controls" indicates whether or not the set of macroeconomic variables (shown in section A.1.2) are included as regressors. The t-statistics (in parentheses) are calculated using Newey-West standard errors with six lags. The p-values shown in brackets are determined using 10,000 randomly generated samples of independent normally distributed variables with first-order serial correlation coefficients respectively equal to that of EEC and NetEEC. Asterisks correspond to simulated p-values, with * indicating significance at 10%,** significance at 5%, and *** significance at 1%. Panel B reports the simulated p-values of the joint hypothesis that the coefficients on EEC in Panel A are both zero across the six-month and twelve-month horizons.

A. Regressions of Excess Stock Returns

	(+1,+6)	(+1,+12)	(+1,+6)	(+1,+12)
EEC	0.0189*	0.0366**	0.0214*	0.0208
	(1.97)	(2.51)	(2.36)	(1.63)
	[0.097]	[0.047]	[0.078]	[0.234]
NetEEC	0.0169	0.0109	0.0092	0.0066
	(1.62)	(0.59)	(1.03)	(0.78)
	[0.168]	[0.608]	[0.412]	[0.550]
Macro Controls	No	No	Yes	Yes
Constant	0.035	0.069	0.731	−0.489
	(2.84)	(3.43)	(1.21)	(−0.80)
Observations	255	255	255	255
R^2	0.08	0.08	0.33	0.50

B. Joint-Horizon Tests of Coefficients on EEC

Horizons	Simulated P-Value of Joint Significance
$(+1,+6) \cap (+1,+12)$	0.029
$(+1,+6) \cap (+1,+12)$ with Macro Controls	0.048

Table 5

Extent of Earnings Coverage within Size Quintiles

Months from October 1984 to December 2005 are ranked based on EEC into the lowest 25%, highest 25%, and remaining "normal" months. EEC–Q is calculated analogously to EEC but within only size quintile Q in a given month, from smallest (quintile 1) to largest (quintile 10). The full-sample logit specification is used to determine the likelihood of coverage. The means of EEC–Q and the means of the actual number of earnings reports covered within each decile are shown for the months with high EEC, normal EEC, and low EEC.

EEC	(Smallest) 1	2	3	4	(Largest) 5
Panel A: EEC–Q					
Low	−0.37	−0.48	−0.41	−0.25	−0.19
Normal	0.01	−0.07	0.04	0.06	0.02
High	0.91	0.62	0.41	0.32	0.18
Panel B: Actual Number of Firms Covered per Month					
Low	4.4	5.2	8.5	18.3	44.4
Normal	7.6	9.0	13.2	22.6	46.9
High	10.0	12.4	16.0	23.1	45.3

Table 6

Price Reactions to Earnings Surprises of Small Stocks

The left-panel dependent variable is the time-series of monthly coefficients from a cross-sectional regression of small-stock cumulative abnormal returns from day +2 to day +60 (PEAD) on standardized unexpected earnings (SUE). Regressions of the monthly first-stage coefficients on EEC and NetEEC measured in the month of the announcement are reported below. Small stocks are defined as those with a percentile of market capitalization less than or equal to the 20^{th} percentile of NYSE stocks. The right panel examines cumulative abnormal returns from day -1 to $+1$ (ERC) using an analogous method to the left panel. The row labeled "Macro Controls" indicates whether or not the set of macroeconomic variables listed in section A.1.2 are included as regressors. t-statistics (in parentheses) are computed using Newey-West standard errors with three lags. P-values shown in brackets are determined using 10,000 randomly generated samples of independent normally distributed variables with first-order serial correlation coefficients matching those of EEC and NetEEC respectively. Asterisks correspond to the simulated p-values; * indicates significance at 10%,** significance at 5%, and *** significance at 1%.

	PEAD		ERC	
EEC	-0.140^{***}	-0.184^{***}	-0.066^{**}	-0.023
	(-2.96)	(-3.68)	(-2.37)	(-1.12)
	[0.007]	[0.001]	[0.051]	[0.320]
NetEEC	0.050	0.075	0.015	0.017
	(1.33)	(1.52)	(0.70)	(0.71)
	[0.204]	[0.163]	[0.547]	[0.523]
Macro controls	No	Yes	No	Yes
Constant	0.173	-1.038	0.295	0.589
	(3.95)	(-0.36)	(10.52)	(0.46)
Observations	255	255	255	255
R^2	0.04	0.13	0.04	0.36

Table 7
Extent of Earnings Coverage and Return Dispersion

Monthly standard deviation of the cross section of returns is regressed on EEC, NetEEC, and macroeconomic controls. Dispersion is measured each month across large and small stocks respectively. Large and small are defined as above the 80^{th} percentile of market capitalization using NYSE breakpoints in the prior month and less than or equal to the 20^{th} percentile, respectively. The row labeled "Macro Controls" indicates whether or not the set of macroeconomic variables listed in section A.1.2 are included as regressors. t-statistics (in parentheses) are computed using Newey-West standard errors with six lags. P-values shown in brackets are determined using 10,000 randomly generated samples of two independent normally distributed variables with first-order serial correlation coefficients matching EEC and NetEEC respectively. Asterisks correspond to simulated p-values. * indicates significance at 10%,** significance at 5%, and *** significance at 1%.

	Return Dispersion Small Stocks		Return Dispersion Large Stocks	
EEC	0.016***	0.019***	−0.007	0.000
	(2.76)	(4.35)	(−1.54)	(0.09)
	[0.007]	[0.000]	[0.203]	[0.949]
NetEEC	−0.008	−0.006	0.008	0.004*
	(−1.72)	(−1.71)	(1.50)	(2.07)
	[0.113]	[0.137]	[0.180]	[0.053]
Macro Controls	No	Yes	No	Yes
Constant	0.212	−0.021	0.086	−0.190
	(37.66)	(−0.11)	(15.67)	(−1.57)
Observations	255	255	255	255
R^2	0.07	0.24	0.04	0.63

Table 8
Extent of Earnings Coverage and Mutual Fund Alpha
within Small Stocks

Each month, the cross section of quarterly small-stock abnormal returns is regressed on lagged quarterly changes in the number of mutual funds that hold a given stock. The first-stage coefficients are then regressed on EEC and NetEEC measured over the same horizon as the changes in fund holdings. Newey-West t-statistics use twelve lags and are shown in parentheses. P-values shown in brackets are determined using 10,000 randomly generated samples of two independent normally distributed variables with first-order serial correlation coefficients matching EEC and NetEEC respectively. Asterisks correspond to simulated p-values; * indicates significance at 10%,** significance at 5%, and *** significance at 1%. Panel B reports the mean monthly four-factor alphas of long-short portfolios, separated into quarters with above-average EEC and below-average EEC in the formation period. The portfolios are long (short) the stocks in the highest (lowest) quintile of changes in the number of funds holding the stock over the quarter, and are held for one quarter. Factor loadings are allowed to vary across the two states of EEC.

A. Regressions		
EEC	0.002**	0.002*
	(3.35)	(2.64)
	[0.05]	[0.07]
NetEEC	0.000	−0.001
	(0.93)	(−1.21)
	[0.64]	[0.41]
Macro controls	No	Yes
Constant	−0.001	0.057
	(−0.84)	(1.66)
Observations	252	252
R^2	0.05	0.12

B. Portfolio Alphas	
High EEC	0.17
	(0.93)
Low EEC	−0.33*
	(−1.83)