JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS Vol. 51, No. 4, Aug. 2016, pp. 1193–1229
COPYRIGHT 2016, MICHAEL G. FOSTER SCHOOL OF BUSINESS, UNIVERSITY OF WASHINGTON, SEATTLE, WA 98195
doi:10.1017/S002210901600051X

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

# Does Common Analyst Coverage Explain Excess Comovement?

Ryan D. Israelsen*

## Abstract

This article shows that correlated errors in news about fundamentals are an important, rational determinant of excess comovement. Individual analysts' forecast errors tend to be correlated across stocks. Using a proxy for correlated forecast errors based on analyst coverage, I find that stocks with similar sets of analysts exhibit more excess comovement, controlling for industry and other variables. Exogenous changes in commonality in analyst coverage around i) brokerage firm mergers and ii) additions to an index lead to changes in excess comovement. This information channel explains 10% to 25% of the increase in comovement around additions to the S&P 500 index.

## I. Introduction

Return comovement is a fundamental part of modern asset pricing theory. Covariances between assets determine their expected returns.[1] However, many studies have documented the existence of "excess comovement," or comovement above and beyond that predicted by standard models. These papers typically fall into one of two categories: i) those identifying groups of assets whose returns are correlated after controlling for risk (see, e.g., Karolyi and Stulz (1996), Kallberg and Pasquariello (2008)) and ii) those documenting increases in comovement

*Israelsen (corresponding author), risraels@indiana.edu, Indiana University, Kelley School of Business, Bloomington, IN 47405. Thanks to Azi Ben-Rephael, Hendrik Bessembinder (the editor), Utpal Bhattacharya, Dennis Capozza, Jonathan Cohn, Andrew Ellul, Eitan Goldman, Raffi Indjejikian, Marcin Kacperczyk, Francisco Palomino, Paolo Pasquariello, Uday Rajan, Tyler Shumway, Irina Stefanescu, Michela Verardo, Scott Yonker, Lu Zhang, Xiaoyan Zhang, an anonymous referee, and participants in seminars at Brigham Young University, Indiana University, University of Michigan, Utah State University, the 2012 American Finance Association meetings in Chicago, the 2010 Financial Management Association meetings in New York City, the 2010 Conference on Financial Economics and Accounting at the University of Maryland, and the 2011 State of Indiana Finance Conference at Notre Dame. I thank Hong Kee Sul and Jordan Winkler for help gathering star analyst data. This paper is adapted from chapter 2 of my dissertation. This paper previously circulated under the title "Does Correlated Analyst Coverage Explain Excess Comovement?"

[1] In theory, the covariance of an asset's return with the stochastic discount factor determines its expected return. Depending on the specific theoretical or empirical assumptions, this amounts to covariances with growth in the marginal utility of consumption (as in consumption-based asset pricing models), with the return of the market portfolio (as in the capital asset pricing model (CAPM)), with the returns of a set of factor portfolios (as in the arbitrage pricing theory (APT)), or with the wealth portfolio and a hedging portfolio (as in the intertemporal capital asset pricing model (ICAPM)). However, each of these can be represented by factor-mimicking portfolios consisting of sets of traded assets.

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

around events not necessarily associated with changes in fundamentals, such as the addition of a stock to an index.[2] Although most explanations for this phenomenon rely on irrational behavior on the part of investors or liquidity-induced demand by institutional investors,[3] this article shows that a largely ignored source of commonality, correlated errors in news about fundamentals, is an important, rational determinant of excess comovement.

Because there are scant data on individual investors' information sets, information-driven comovement has received relatively less attention than other sources of comovement. Instead of directly focusing on those consuming information (investors), this article focuses on those producing information (sell-side analysts). Analyst earnings forecast data are readily available for a wide range of stocks.

For a number of reasons, an individual analyst's earnings forecast errors are likely to be correlated across the stocks he or she covers. For example, the analyst may use the same model, methodology, or inputs when making earnings predictions, which will cause any error to propagate through to all of the forecasts. When errors in signals about future earnings or cash flows are positively correlated across stocks, expected conditional correlations of stock returns are higher than can be explained by fundamentals. Rational Bayesian updating leads to an increase in comovement when correlations in forecast errors increase. Even though the errors are known to be correlated, the effects of this correlation cannot be fully reversed when forming beliefs.

Unfortunately, estimating the correlation coefficient between the forecast errors of a pair of stocks is not trivial. Precise estimates require a long time series of forecast errors. Furthermore, forecast revision dates, fiscal periods, and earnings announcement dates are typically staggered across stocks, which may lead to misleading estimates as analysts update forecasts for one stock based on earnings news from another.

Instead of directly estimating correlations in earnings forecast errors for all stocks, I make a few simple assumptions to develop an easy-to-calculate, intuitive proxy for the correlation in consensus forecast errors for a pair of stocks:

$$\frac{N_{i,j}}{\sqrt{N_i N_j}}. \tag{1}$$

Here, $N_i$ and $N_j$ are the number of analysts covering stocks $i$ and $j$, respectively, and $N_{i,j}$ is the number of analysts covering both stocks. Because this is simply the number of analysts in common between two stocks, scaled by the geometric average number of analysts covering the stocks, I call this measure "common analyst coverage" and denote it as $\rho_{ij}^{\text{an}}$.

[2] Three examples are Barberis, Shleifer, and Wurgler (2005) with additions to the Standard & Poor's (S&P) 500 index, Green and Hwang (2009) with stock splits, and Boyer (2011) with stocks switching between value and growth indices.

[3] Barberis and Shleifer (2003) develop a model with style investors who choose stocks based on common characteristics, which generates excess comovement within style classes. There is empirical evidence to support this among both retail investors (Kumar and Lee (2006)) and institutional investors (Pirinsky and Wang (2004)). There are also several papers providing evidence that comovement may be partially due to liquidity-induced demand by institutions (e.g., Coval and Stafford (2007), Boyer and Zheng (2009)).

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

Because it is based on simplifying assumptions about individual analyst behavior, using the proxy for correlated earnings forecast errors involves a trade-off. The accuracy of the measure depends on the extent to which actual behavior deviates from these assumptions. Conversely, the proxy is simple to calculate and requires no estimation, whereas calculating the correlation directly using two time series of earnings forecasts and actuals involves estimation error and is potentially biased if dates are staggered. Furthermore, the proxy allows the correlation to change yearly (or even quarterly), whereas a direct estimate requires a lengthy time series. For a subset of stocks with a sufficiently long time series and similar fiscal year-ends and announcement dates, I directly calculate the correlations in earnings forecast errors and validate that the analyst-coverage-based measure is a good proxy for correlated earnings forecast errors. Additionally, for this subset, I find a positive relation between the direct measure and excess comovement. Finally, I show that individual analysts' behavior is consistent with the assumptions used in deriving the proxy.

Having validated the new measure, I use data from a much broader sample to test whether pairs of stocks with higher forecast error correlations exhibit more excess comovement. I begin by examining comovement in the cross section of stock returns. Controlling for risk based on standard asset pricing models, I find that cross-sectional increases in common analyst coverage among pairs of S&P 500 stocks are associated with increases in excess comovement during the subsequent year. A 1% increase in the proxy for correlated errors means an increase in excess comovement of about 0.4%. Results are similar for stocks within the same industry and controlling for a variety of firm characteristics.

Of course, analyst coverage is clearly endogenous. Analysts tend to cover firms within a single industry and with similar exposure to systematic risk. To the extent that analysts are better at picking stocks with similar risk characteristics than are standardly used asset pricing models, the cross-sectional relation between analyst coverage and excess comovement could simply be capturing unmeasured systematic risk not accounted for by risk factors or other controls. To more thoroughly rule out this alternative hypothesis, I choose two types of events that result in time-series changes in analyst coverage for pairs of stocks, but are unlikely to be associated with changes in firms' fundamentals.

First, I examine the turnover of analysts following 15 mergers of brokerage firms. These events have been used as natural experiments to examine the effects of competition on the bias of analysts' forecasts (e.g., Hong and Kacperczyk (2010)) and on the effect of analyst coverage on firms' innovation (e.g., He and Tian (2013)). These mergers generate variation in analyst coverage that is likely exogenous. Analyst coverage naturally changes following mergers as analysts are made redundant, but the changes are unlikely to be related to any changes in fundamentals of the firms covered by a specific analyst. Focusing on stocks covered by both the acquiring firm's analyst and the target firm's analyst before the merger, and by the single remaining analyst postmerger, I find that excess comovement typically drops for pairs of affected stocks and that the drop is most pronounced where the loss in coverage has the most impact.

Second, I examine changes in analyst coverage around the addition of stocks to the S&P 500 index, where betas for newly added stocks generally increase.

Barberis et al. (2005) argue that firm fundamentals do not change in short windows around these events. I show that there is a fundamental change in the information structure associated with the newly added firms following the addition. In particular, the number of analysts covering the firms increases, on average, and these new analysts tend to be those already covering other S&P 500 stocks. Common analyst coverage between the newly added stocks and existing S&P 500 stocks increases significantly by about 18% within the first year following the event. Furthermore, I find a positive cross-sectional relation between the change in common analyst coverage and the change in excess comovement. Pairs of stocks whose sets of analysts become more similar tend to see larger increases in excess comovement. The magnitude of the effect is 10%–25% of the increase in betas and is consistent with the hypothesis that some level of comovement may be driven by commonality in information, specifically, correlations in forecast errors.

To further validate that common analyst coverage captures information, and not simply analysts' endogenous choice of firms, I modify the proxy in two ways. First, I place more weight on star analysts and on analysts whose signals are more precise. In these cases, the relation between excess comovement and the modified proxies is stronger, suggesting that information is an important part of the effect. Second, instead of focusing on pairs of stocks covered by the same analyst, I examine stocks each covered by a different analyst at the same brokerage firm or analysts that worked together at the same brokerage firm at the same time in the past. Because these are different analysts covering a broad range of stocks, endogeneity in the choice of coverage is not likely to be an issue. However, because they are socially connected, they may have had the same training, may use the same models or inputs, or may simply share information or opinions, their errors are likely to be correlated. Here, I find a statistically significant, albeit weaker, relation between the modified measures and excess comovement.

The enactment of Regulation Fair Disclosure (Reg FD) in 2000 was meant to eliminate analysts' access to private, firm-specific information. As a result, the information revealed by their forecasts is more likely to be systematic or industry-wide information. If this information is common to the stocks the analyst is covering, any error in this information will also be common. Consistent with this information story, I find that the relation between common analyst coverage and excess comovement increases after the implementation of Reg FD.

Finally, I investigate whether a trading strategy based on common analyst coverage and past returns can generate positive, abnormal returns. I find that stock returns that deviate from those of their "peers" (as defined by the common analyst coverage variable) tend to revert during the following month. A portfolio that is long stocks that have had abnormally low returns relative to peers and short stocks that have had abnormally high returns relative to peers generates abnormal returns of 1% per month in even the largest 10% of stocks. A similar strategy that uses Global Industry Classification Standard (GICS) industries to define peers does not generate statistically significant abnormal returns.

This article makes several contributions to the literature on information and comovement. Its main contribution lies in identifying an underexplored source of excess comovement, correlated errors in news about fundamentals, providing evidence both in the cross section and in the time series around exogenous changes in

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

analyst coverage. Additionally, it develops and validates a proxy for the correlation in consensus forecast errors between two stocks. Finally, it partially explains the increase in comovement around additions to the S&P 500 index (as in Barberis et al. (2005)) as a result of an increase in correlated information.

The article proceeds as follows: In Section II, I develop the hypotheses to be tested and explain the intuition. In Section III, I introduce the common analyst coverage measure, validate it as a proxy for correlated forecast errors, and describe the data used in the article. In Section IV, I use a variety of approaches to test the hypotheses that correlated forecast errors are associated with excess comovement. In Section V, I conclude.

## II. Hypotheses

To examine whether the structure of the information production environment has a significant effect on the comovement of asset prices, I develop and test three closely related hypotheses. The first is the central hypothesis of the article, and the final two involve modifications of the basic proxy for correlations in earnings forecast errors.

Under rational Bayesian updating, correlations in forecast errors lead to correlations in expected cash flows and returns. This is true even if actual earnings are uncorrelated. This is illustrated with following example. Consider the case of two firms, $i$ and $j$, that will release earnings $\pi_i$ and $\pi_j$, respectively, at the end of the period. Assume $\pi_i$ and $\pi_j$ are normally distributed with known variances $\sigma_i^2$ and $\sigma_j^2$, and that the two are uncorrelated.[4] Assume also that there are earnings forecasts, $f_i = \pi_i + e_i$ and $f_j = \pi_j + e_j$, where the forecast errors $e_i$ and $e_j$ are normally distributed with means of zero and variances $\sigma_{ei}^2$ and $\sigma_{ej}^2$. The two forecast errors have a potentially nonzero covariance of $\sigma_{e12}$. Given the priors and forecasts, the covariance and correlation between the Bayesian estimates of $\pi_i$ and $\pi_j$ are

$$\text{(2)} \qquad \operatorname{cov}(\pi_i, \pi_j | f_i, f_j) = \frac{\sigma_{eij}\sigma_i^2\sigma_j^2}{\left(\sigma_i^2 + \sigma_{ei}^2\right)\left(\sigma_j^2 + \sigma_{ej}^2\right) - \sigma_{eij}^2}$$

and

$$\text{(3)} \qquad \operatorname{corr}(\pi_i, \pi_j | f_i, f_j) = \frac{\sigma_{eij}\sigma_i^2\sigma_j^2}{\sqrt{\sigma_i^2\sigma_j^2\left[\sigma_{e2}^2\left(\sigma_i^2 + \sigma_{ei}^2\right) - \sigma_{eij}^2\right]\left[\sigma_{ei}^2\left(\sigma_j^2 + \sigma_{ej}^2\right) - \sigma_{eij}^2\right]}}.$$

When earnings forecast errors are uncorrelated (i.e., when $\sigma_{eij} = 0$), the covariance and correlation of the Bayesian estimates are both equal to 0. However, when there is a positive correlation in earnings forecast errors, both are positive. Furthermore, both increase as $\sigma_{eij}$ increases.[5] The intuition is straightforward. When the consensus earnings forecast errors for stocks $i$ and $j$ are

[4] Allowing the two to be directly correlated or to be based on a linear factor structure does not change the basic result, but the formulas become more cumbersome.

[5] These two results can be seen by taking the derivative of each with respect to $\sigma_{eij}$ and using the restriction that the variance–covariance matrix is positive definite. In other words, $\sigma_{eij}^2 < \sigma_{ei}^2\sigma_{ej}^2$.

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

positively correlated, both forecasts are likely to be simultaneously too high or too low, but a rational Bayesian agent will not know which is true. As long as the signals are informative about the means, both will be used to update expectations of future cash flows and returns, which will both likely be too high or too low. The agent cannot fully undo the correlation in the signals, and as a result, the updated expectations will be subject to some level of correlation.

To the extent that common analyst coverage is a good proxy for correlated earnings forecast errors, the following hypothesis should hold:

*Hypothesis 1.* Pairs of stocks with higher common analyst coverage, $\rho_{ij}^{\text{an}}$, will exhibit more excess comovement.

As described in the Introduction, I test this hypothesis both in the cross section and in the time series around exogenous changes in analyst coverage using a proxy. Additionally, I test it using direct estimates of correlations in earnings forecast errors for a subset of stocks.

An obvious alternative to the correlated information hypothesis is the idea that analysts may be better at picking similar stocks than are our asset pricing models. If this is the case, the common analyst coverage variable will be a measure of how similar stocks are in terms of exposure to systematic risk. If our models fail to capture these risk exposures or if we do not measure risk correctly, there will be an association between common analyst coverage and "excess" comovement. To rule out this alternative, I control for GICS industry and various firm characteristics and develop several more hypotheses.

The ideal way to rule out the alternative hypothesis of unmeasured risk would be to exogenously change analyst coverage without changing the fundamental risk in the underlying stocks and then examine the corresponding changes in excess comovement. If the structure of information production affects comovement, there should be predictable changes in excess comovement with exogenous changes in analyst coverage.

To more carefully test this hypothesis, I examine two types of events that are associated with changes in analyst coverage, but not necessarily changes in fundamentals: i) mergers of brokerage firms and the resulting turnover in analysts and analyst coverage and ii) additions of stocks to the S&P 500 index. Both of these events have been used in the literature as natural experiments in explaining the behavior of traders or analysts.

For the final two hypotheses, I explore four deviations from the basic proxy to reduce the likelihood that the relation is driven by risk. If the effect is caused by correlated information as opposed to unmeasured risk, it ought to be the strongest where forecasts are given the most weight by investors. There are a few ways to measure the precision of information. One natural measure of information precision is whether the analyst making the forecasts is a "star analyst" as measured by *Institutional Investor* magazine. If the effect is driven by unmeasured risk, there should be no difference across the two types of analysts. If the effect is driven by information, it should be stronger among the subset of star analysts. Likewise, past earnings forecast errors provide another measure of the precision of an analyst's forecasts. A measure that places more weight on analysts whose earnings forecasts are more precise ought to provide a stronger relation. Thus, I test the following hypothesis:

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

*Hypothesis 2.* The relation between analyst coverage and excess comovement is stronger where information is more precise.

Social networks have been used to explain aspects of chief executive officer investment behavior (Fracassi (2016)), to examine the role of boards in choosing and monitoring executives (Adams and Ferreira (2007), Subrahmanyam (2008)), and to investigate the ability of connected analysts to gather superior information (Cohen, Frazzini, and Malloy (2010)). For a variety of reasons, social connections among analysts may affect analyst forecasts and, ultimately, return comovement. To the extent that analysts working for the same brokerage firm use similar methodology, information, or inputs, or simply process information in the same manner, we might expect their forecast errors to be correlated. Additionally, because of the wider coverage at the brokerage firm level, mismeasured industry-related risk factors are less likely to be an issue. Individual analysts tend to cover a small percentage of the entire universe of stocks. Any given brokerage firm, however, will employ analysts covering a much larger fraction of the S&P 500 firms.

Similarly, analysts who worked together in the past at the same firm may also have correlated forecast errors. First, these analysts may have remained in contact with each other and may be sharing information. Second, at the previous firm, they may have learned similar valuation techniques or information-gathering and -processing techniques, which they both may have retained to some degree. Third, to the extent that individual firms choose to hire specific types of analysts, employment history is a proxy for analyst "type." If similar types of analysts interpret information in the same way, this should filter down to the forecasts. Finally, along the same lines, there may be some self-selection on the part of analysts when choosing employers. Similar types may be attracted to the same firm. This intuition motivates the final hypothesis:

*Hypothesis 3.* Pairs of stocks with analysts that are socially connected will exhibit more excess comovement.

To test this hypothesis, I modify the basic analyst coverage measure to focus on pairs of stocks whose analysts work for the same brokerage firm or whose analysts worked together in the past.

## III. Data Sources, Variable Construction, and Validation

I use the unique analyst and broker identification numbers from the Institutional Brokers' Estimate System (IBES) database to determine which analysts are making earnings per share forecasts. I gather stock return and (6-digit) GICS industry code data from the Center for Research in Security Prices (CRSP) database. Data for the 3 Fama and French (1993) factors and the Carhart (1997) momentum factor along with data on additions to the S&P 500 index are from Wharton Research Data Services. Mutual funds holdings data are from the Thomson Reuters Mutual Fund Holdings database and institutional investor holdings are from the Thomson Reuters Institutional Holdings (13F) database. Star analyst data are from October issues of *Institutional Investor* magazine. Unless otherwise noted, the sample covers 1995 through 2008. The beginning of the sample is limited by the

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

availability of GICS industry codes. GICS codes from 1995 to 1998 are backfilled. These industry definitions have been shown to be the most relevant in explaining industry classifications by analysts. The use of Standard Industrial Classification codes and a longer sample does not qualitatively change results.

## A. Common Analyst Coverage

The primary focus of this article is the covariance structure of analysts' earnings forecast errors. It is possible to estimate this matrix directly using realized earnings forecast errors. However, given the staggered nature of both earnings announcement dates and analyst forecast dates, this is a difficult task in practice. Instead of estimating the covariance matrix directly for all stocks, I take an indirect approach using observable analyst coverage data and then validate the proxy using direct estimates for a subset of stocks. As mentioned, analysts often use models or standardized methodologies when making projections about earnings. Many of the inputs used in their models, such as projected gross domestic product or industry growth, may be used across multiple stocks and among analysts within the same firm. Additionally, analysts may interpret information in the same manner for multiple stocks. To the extent that each analyst uses the same methodology, model, or inputs across stocks, we might expect any systematic errors to filter through to the earnings forecasts across stocks for which they make projections.

With this in mind, assume for simplicity that the signal, or consensus forecast, of the $i$th firm is the simple average of the forecasts of $N_i$ analysts. Also, assume that forecast errors across analysts are uncorrelated but that each analyst's forecast errors are perfectly correlated across stocks and have variances of $\sigma_a^2$ for any stock.[6] Then, the variance of the consensus forecast error, $e_i$, for the $i$th stock is equal to $\sigma_a^2/N_i$. Let $N_{i,j}$ be the number of shared analysts between the $i$th and $j$th stocks. Then, the covariance of the forecast errors for stocks $i$ and $j$ is:

$$
\begin{aligned}
(4) \qquad \operatorname{cov}(e_i, e_j) &= \begin{cases} 0 & \text{if } N_{i,j} = 0, \\ \dfrac{\sigma_a^2}{N_i N_j} & \text{if } N_{i,j} = 1, \\ \dfrac{2\sigma_a^2}{N_i N_j} & \text{if } N_{i,j} = 2, \\ \vdots & \\ \dfrac{\min(N_i, N_j)\sigma_a^2}{N_i N_j} & \text{if } N_{i,j} = \min(N_i, N_j) \end{cases} \\
&= \frac{N_{i,j}}{N_i N_j}\sigma_a^2,
\end{aligned}
$$

[6]The assumption that individual analyst forecast errors are perfectly correlated is made for simplicity. Allowing an imperfect correlation across stocks will change the definition of $\rho_{ij}^{\text{an}}$, but the comparative statics used in developing the hypotheses and tests will not change. For example, if we assume that forecast errors at the individual analyst level across stocks have a correlation of $\alpha > 0$, it is easy to show that $\rho_{ij}^{\text{an}} = \alpha N_{i,j}/\sqrt{N_i N_j}$. Changing $\alpha$ will change the magnitude of the effect, but not the sign or the nature of the relation.

and thus the correlation of the consensus forecast errors is:

$$(5) \qquad \rho_{ij}^{\text{an}} \equiv \text{corr}(e_i, e_j) = \frac{N_{i,j}}{\sqrt{N_i N_j}}.$$

This coefficient, which ranges between 0 and 1, is similar to measures that have been used to quantify the "similarity" of sets.[7] Thus, this quantity is a practical measure of the similarity of the sets of analysts covering the two stocks and a theoretical proxy for the correlation between the forecast errors for the two stocks. Using the IBES database, this can be easily calculated for any pair of stocks and requires no estimation.

The cross-sectional results focus on S&P 500 stocks, which is where analyst coverage is the most concentrated. Each calendar year, for every pair of stocks listed on the S&P 500 index from 1995 through 2007, I calculate the common analyst coverage measure, $\rho_{ij}^{\text{an}}$, by counting the number of unique analysts making earnings forecasts in the IBES database, $N_i$ and $N_j$, and the number of analysts in common within these two groups, $N_{i,j}$. There are about 1 million stock-pair observations over the 13-year period. The first row in Panel A of Table 1 shows the summary statistics for $\rho_{ij}^{\text{an}}$. On average, this variable is about 2.7%. This number is relatively low. However, the distribution is skewed with a standard deviation of about 8.4%. Analysts tend to specialize in a particular industry, so naturally pairs of firms within the same industry generally have relatively high coefficients.

To test Hypotheses 2 and 3, I strategically modify common analyst coverage to create five new variables: i) common star analyst coverage, $\rho_{ij}^{\text{st}}$; ii) common nonstar analyst coverage, $\rho_{ij}^{\text{-st}}$; iii) precision-weighted common analyst coverage, $\rho_{ij}^{\text{wgt}}$; iv) common broker coverage, $\rho_{ij}^{\text{br}}$; and v) connected analyst coverage, $\rho_{ij}^{\text{ca}}$. The denominator in these five measures, $\sqrt{N_i N_j}$, is identical to that of the analyst coverage measure. However, the numerator is replaced with, respectively: i) the number of star analysts in common; ii) the number of nonstar analysts in common; iii) the number of analysts in common, weighted by the precision of their past forecasts; iv) the number of analysts from the same brokerage firm in common; and v) the number of analysts in common who used to work together at the same firm but no longer do.

The differences in the measures are illustrated by the following example. Assume there are two stocks, IBM and MSFT, and five analysts. Analysts A, B, and C cover IBM, and analysts A, D, and E cover MSFT. Analysts B and D work together at one brokerage firm, analysts C and E work together at another firm, and analysts A and E used to work together but no longer do. Also, analyst A is the only star analyst. Then, common analyst coverage, $\rho_{ij}^{\text{an}}$, between IBM and MSFT is equal to ⅓ (the number of analysts in common, 1, divided by the geometric average number of analysts across the two stocks, 3). Because the only analyst covering both stocks is a star analyst, common star analyst coverage, $\rho_{ij}^{\text{st}}$, is also

[7] Two popular measures of similarity are Sørensen's (1948) similarity coefficient, $2N_{i,j}/(N_i + N_j)$, and the Jaccard (1901) similarity coefficient, $N_{i,j}/(N_i + N_j - N_{i,j})$. The measure in this article differs from Sørensen's measure only in the sense that a geometric average is used in lieu of an arithmetic average.

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

TABLE 1
Summary Statistics

Table 1 reports summary statistics for the common analyst coverage variables and the stock return correlation coefficients for pairs of S&P 500 stocks from 1995–2007. Panel A presents common analyst coverage, common star analyst coverage, common nonstar analyst coverage, precision-weighted common analyst coverage, common broker coverage, and connected analyst coverage. The correlation coefficients used in Panel B are estimated using raw returns, and residuals from the capital asset pricing model (CAPM), the Fama–French (1993) 3-factor model, and the Fama–French 3-factor model plus the Carhart (1997) momentum factor. The three models are fit using 1 year's worth of daily data and are applied to the following year's data to estimate the correlation coefficients. Panel C presents statistics for common mutual fund holdings along with summary statistics for DLNSIZE, which is the absolute value difference in the natural logarithms of the market values of equity; DLNBTM, which is the absolute value difference in the natural logarithms of the book-to-market equity ratios; DPRICE, which is the absolute value difference in lagged stock prices; DLEVERAGE, which is the absolute value difference in leverage ratios; DROE is the absolute value difference in returns on equity; and STATE, which is a dummy variable indicating the headquarters of the two firms are in the same state. Analyst data are from the Institutional Brokers' Estimate System (IBES) database and *Institutional Investor* magazine, stock returns are from the Center for Research in Security Prices, the Fama–French and Carhart factor data are from Wharton Research Data Services, mutual fund holdings data are from Thomson Reuters, and earnings forecasts are from IBES. Means, standard deviations, minimums, and maximums are expressed in percentage terms.

| | Description | *N* | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|---|
| *Panel A. Common Analyst Coverage Measures* | | | | | | |
| $\rho_{ij}^{an}$ | Analyst coverage | 970,611 | 2.69 | 8.41 | 0.00 | 100.00 |
| $\rho_{ij}^{st}$ | Star analysts | 970,611 | 0.21 | 1.57 | 0.00 | 57.21 |
| $\rho_{ij}^{-st}$ | Nonstar analysts | 970,611 | 2.48 | 7.40 | 0.00 | 100.00 |
| $\rho_{ij}^{wgt}$ | Weighted analysts | 970,611 | 1.37 | 8.25 | 0.00 | 138.64 |
| $\rho_{ij}^{br}$ | Brokerage coverage | 970,611 | 46.72 | 19.41 | 0.00 | 189.74 |
| $\rho_{ij}^{ca}$ | Connected analysts | 970,611 | 61.68 | 35.09 | 0.00 | 313.53 |
| *Panel B. Stock Return Correlations* | | | | | | |
| $\rho_{ij}^{ret}$ | Raw returns | 970,611 | 26.97 | 15.65 | −34.62 | 96.19 |
| $\rho_{ij}^{capm}$ | CAPM alphas | 970,611 | 8.56 | 18.03 | −56.37 | 97.45 |
| $\rho_{ij}^{ff3}$ | FF3 alphas | 970,611 | 16.87 | 21.06 | −59.55 | 97.51 |
| $\rho_{ij}^{ff4}$ | FF3 + Carhart alphas | 970,611 | 19.10 | 22.25 | −55.42 | 98.68 |
| *Panel C. Other Variables* | | | | | | |
| $\rho_{ij}^{mf}$ | Mutual fund holdings | 970,611 | 39.96 | 9.24 | 0.00 | 87.83 |
| DLNSIZE | Abs ln size diff | 838,562 | 1.10 | 0.81 | 0.00 | 4.92 |
| DLNBTM | Abs ln B/M diff | 838,562 | 0.87 | 0.75 | 0.00 | 13.40 |
| DPRICE | Abs price diff | 838,562 | 0.64 | 0.52 | 0.00 | 4.91 |
| DLEVERAGE | Abs leverage diff | 838,562 | 0.19 | 0.15 | 0.00 | 0.91 |
| DROE | Abs ROE diff | 838,562 | 0.61 | 1.21 | 0.00 | 5.74 |
| STATE | Same state dummy | 838,562 | 0.05 | 0.22 | 0.00 | 1.00 |

equal to ⅓, and common nonstar analyst coverage, $\rho_{ij}^{-st}$, is equal to 0. Because two pairs of analysts work at the same firm, common broker coverage, $\rho_{ij}^{br}$, is equal to ⅔. Note that I do not include the link between analysts A and A across the two stocks. For this measure, the analyst must be different. Finally, connected analyst coverage, $\rho_{ij}^{ca}$, is equal to ⅓ because one pair of analysts used to work together but no longer do.

An analyst is considered to be a star analyst if he or she is ranked 1, 2, or 3 in one of the "All-American Research" teams in the October issue of *Institutional Investor* magazine from the previous year. A pair of analysts is considered to be socially connected if both analysts worked at the same firm at some point in the past but no longer work with each other at the same firm. Because of the latter condition, this measure is orthogonal to common analyst coverage and common broker coverage. Finally, the precision-weighted common analyst coverage variable is defined as the reciprocal of the analyst's mean, scaled, absolute earnings

forecast error during the previous year:[8]

$$\text{PRECISION}_i = \frac{1}{\frac{1}{N_i}\sum_{j\in F_i}\frac{|e_j|}{P_{j,m-1}}},$$

where $F_i$ is the set of earnings forecasts made by analyst $i$ during year $t-1$, $N_i$ is the number of forecasts, $e_j$ is the earnings forecast error for stock $j$, and $P_{j,m-1}$ is the price of the stock at the end of the month before the earnings announcement. For each analyst, I include the most recent forecast before the earnings announcement for every stock in the IBES database in $F_i$. Precision is measured using all stocks covered by an analyst, not stock by stock. I include two filters on this procedure. First, to reduce the noise in the measure, I exclude any analysts who have made forecasts for fewer than 4 firm-quarters in the previous year. Second, to avoid infinite precision, I winsorize the precision variables each year at 1% and 99%.

Once I have estimates of precision, I create a precision-weighted common analyst coverage measure, $\rho_{ij}^{\text{wgt}}$, by creating weights, $w_{il}$, for analyst $l$ covering stock $i$ from the $M$ analysts covering each stock ($w_{il} = \text{PRECISION}_l / \sum_{m=1}^{M}\text{PRECISION}_{im}$) and allocating the coverage based on these weights. In particular, I define the measure as

$$\rho_{ij}^{\text{wgt}} \equiv \frac{\sqrt{\left(\sum_{l\in S} w_{il}N_i\right)\left(\sum_{l\in S} w_{jl}N_j\right)}}{\sqrt{N_iN_j}} = \sqrt{\sum_{l\in S} w_{il}\sum_{l\in S} w_{jl}}, \tag{6}$$

where $S$ is the set of analysts covering both stocks $i$ and $j$. Note that when analysts all have the same precision, this reduces to the basic measure, $\rho_{ij}^{\text{an}}$. Summary statistics for all five alternative measures are presented in Panel A of Table 1.

Because star analysts make up only a small subset of all analysts, it is not surprising that the average value of this variable is much smaller than the general measure at 0.2%. The standard deviation is 1.6%. The nonstar analyst variable is 2.5% on average and has a standard deviation of 7.4%. The average of the sum of the two is equal to the average of the basic analyst coverage variable by construction, as $\rho_{ij}^{\text{st}} + \rho_{ij}^{-\text{st}} = \rho_{ij}^{\text{an}}$. The precision-weighted measure is slightly smaller on average than the basic measure at 1.4%, suggesting that analysts who cover a larger number of stocks are likely to have less precise forecasts overall than those who specialize in a small number of stocks. The standard deviation is similar to that of the basic measure.

The mean of the broker-level variable is much larger than that of the analyst measure at about 47%. Its standard deviation is about 19%. Although individual

[8]The use of the root mean square (scaled) forecast error in the measure of precision yields very similar results.

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

analysts focus on industries, brokerage firms take a broader approach. The correlation coefficient between the two measures is 0.34. The final row of Panel A in Table 1 shows that the connected analyst measure has a mean of 62% and a standard deviation of 35%. Because there are potentially $N_i \times N_j$ pairs of analysts at the same broker or formerly at the same broker, these two variables are not constrained at 100%.

## B. Comovement Measures

In this article, I define comovement as the correlation of returns or alphas between two stocks. The pairwise methodology has been used in the literature to measure and explain excess comovement (e.g., Kallberg and Pasquariello (2008)). An alternative approach typically involves estimating betas relative to a market index or a subset of similar stocks. The analyst coverage measure in this study does not easily lend itself to the latter method. Because the analyst coverage variable is defined between a pair of stocks and not with respect to an index, the pairwise variable is a more natural measure when using panel data. Both measures involve scaled covariances of pairs of stocks; a beta is simply a weighted average of individual asset covariances divided by the variance of the entire index or portfolio of assets.

Excess comovement, $\rho_{ij}^{\text{alpha}}$, is defined as the correlation coefficient between two firms' realized alphas using an asset pricing model:

$$\rho_{ij}^{\text{alpha}} = \frac{\sum_{t=1}^{T} e_{it} e_{jt}}{\sqrt{\left(\sum_{t=1}^{T} e_{it}^2\right)\left(\sum_{t=1}^{T} e_{jt}^2\right)}}, \tag{7}$$

where $e_{it}$ and $e_{jt}$ are the daily residuals from the asset pricing equation:

$$r_{it} - r_{ft} = \alpha_i + \beta_{i1} f_{1t} + \beta_{i2} f_{2t} + \ldots + \beta_{iN} f_{Nt} + e_{it}. \tag{8}$$

Every calendar year from 1996 through 2008, I find all stocks that are part of the S&P 500 index during that entire year and the previous year. The factor loadings, $\beta_{i1}, \beta_{12}, \ldots, \beta_{Nt}$, are estimated using data from the previous calendar year. Three models are used: the CAPM; the Fama–French (1993) 3-factor model (FF3), which includes the market return, the size factor, and the growth factor; and a 4-factor model (FF4) that adds the Carhart (1997) momentum factor to the Fama–French factors. For comparison purposes, I also calculate correlations in raw returns, $\rho_{ij}^{\text{ret}}$.

Panel B of Table 1 provides summary statistics for these estimates (expressed in percentages) using daily returns data. There are an average of 390 stocks per year that are part of the index both the current and previous years and have been assigned GICS codes. This leads to 970,611 yearly correlations over the sample period, or about 75,000 per year. The average correlation using raw returns is about 27%. Excess comovement among these S&P 500 stocks based on the CAPM is about 8.6% on average, and about 16.9% and 19.1% using the FF3 and FF4, respectively. In a few cases, there is some negative excess comovement, and the maximum for each variable exceeds 96%.

Daily returns may be subject to microstructure issues such as nonsynchronous trading. Though not reported here, the results of the article are quali-

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

tatively unchanged when using excess comovement estimated at weekly frequencies. In most of the empirical tests, the dependent variables are measured 1 year before comovement to avoid any endogeneity problems.

## C. Control Variables

To control for other potential sources of excess comovement, I create several additional measures. One is related to trading behavior of mutual fund managers and the rest are characteristic-based measures of similarity.

Comovement in stock returns may be partially driven by price pressure from style investors or mutual fund managers tracking an index. If a pair of stocks is held by the same set of mutual funds, they are more likely to receive (near) simultaneous price pressure when these funds experience inflows or outflows of capital. Boyer and Zheng (2009) find a positive and significant correlation between mutual fund flows and returns. Coval and Stafford (2007) find that funds that experience large outflows or inflows tend to adjust their positions accordingly, which may cause price pressure on the stocks in their portfolios. Furthermore, Barberis et al. (2005) attribute the increase in comovement in stocks that are added to the S&P 500 index to the trading behavior of style investors, index funds, or other institutional investors.

To control for this source of comovement, I form a variable based on the trading behavior of mutual funds. Every calendar year, I determine the holdings of mutual funds from the Thomson Reuters Mutual Fund Holdings database. For each pair of stocks, I define common mutual fund holdings, $\rho_{ij}^{\text{mf}}$ as:

$$\rho_{ij}^{\text{mf}} \equiv N_{i,j}^{\text{mf}} \Big/ \sqrt{N_i^{\text{mf}} N_j^{\text{mf}}}, \tag{9}$$

where $N_i^{\text{mf}}$ is the number of mutual funds that own shares of stock $i$, and $N_{i,j}^{\text{mf}}$ is the number of mutual funds that hold both stocks $i$ and $j$ at time $t$. Panel C of Table 1 shows that the average value of the common mutual fund holdings variable is 40%. Therefore, on average, the S&P 500 stocks included in my sample share about 40% of their mutual fund investors. These numbers are as low as 0% and as high as 88%.

Finally, I control for variables that are potentially related to the comovement of firms' cash flows and discount rates that may not be captured by our standard asset pricing model. Specifically, I include the following variables for each pair of stocks: GICS, which is a dummy variable indicating that the two stocks are in the same GICS industry; STATE, which is a dummy variable indicating the headquarters of the two firms are in the same state; DLNSIZE, which is the absolute value difference in the natural logarithms of the market values of equity (from CRSP I use the number of shares outstanding, SHROUT, times the end-of-period price, PRC); DLNBTM, which is the absolute value difference in the natural logarithms of the book (the sum of Compustat items CEQ, TXDB, ITCH, and PRCA) to market equity ratios; DPRICE, which is the absolute value difference in lagged stock prices (CRSP item PRC); DLEVERAGE, which is the absolute value difference in leverage ratios, where leverage ratio is defined as total long-term debt (Compustat item LTT) plus total debt in current liabilities (Compustat item DLC)

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

scaled by total assets (Compustat item AT); and DROE, which is the absolute value difference in returns on equity (defined as Compustat item NI divided by SEQ). Panel C of Table 1 provides summary statistics for these variables.

### D. Validation of Analyst Coverage Measure

In this section, I use a subset of stocks to show that the analyst coverage measure is a good proxy for correlations in earnings forecasts errors. An ideal test of my central hypothesis would employ a direct measure of correlated forecast errors. Unfortunately, because information about economy- and industry-wide profitability is revealed with each firm's earnings release, only pairs of stocks with the same fiscal year-end and earnings announcement date are likely to provide precise estimates. For example, if IBM releases earnings that are much higher than expected, some of this may be idiosyncratic, but some may be due to higher than expected industry-level demand. Rational analysts will adjust their earnings forecasts for other firms in IBM's industry. As a result, estimated forecast errors for the other stocks are likely to be smaller than their true values. In contrast, if another firm, say Microsoft, releases earnings on the same day as IBM, analysts have no time to update their forecasts and any common error will show up in both forecasts. Therefore, the fiscal periods and announcement dates should coincide. This reduces the number of stock pairs in the sample. Furthermore, a sufficiently long time series of earnings forecast errors is necessary to estimate a correlation coefficient with reasonable precision; therefore, stocks with short histories are also excluded.

To validate that common analyst coverage is a good measure of correlated earnings forecast errors, I perform three tests. First, I test the validity of the assumption that individual analysts' forecast errors tend to be positively correlated across their assigned stocks. This is a necessary condition for the accuracy of the proxy. If there is no correlation in individual forecast errors for the same analyst across stocks, the consensus errors are unlikely to be correlated. Second, I form a "more fundamental" measure by directly estimating consensus-level earnings forecast correlations for pairs of stocks and examine whether these are positively correlated with the analyst coverage measures. If there is not a positive relation between these two, the proxy is probably not valid. Third, for this subset, I use the direct measure to examine the main hypothesis that correlated forecast errors lead to excess comovement. Table 2 presents these tests.

In Panel A of Table 2, I calculate the average earnings forecast error correlations for individual analysts. Correlations are estimated at the analyst-stock-pair level for pairs of stocks with the same fiscal year-ends, for analysts making revisions on the same date for both stocks. Correlations are estimated using each revision. Averages and standard errors of the mean are presented for correlations estimated using at least 5, 10, 15, 20, 25, and 30 revisions. As the minimum number of revisions increases, the number of analyst-stock-pair observations decreases from 21,151 to 253. However, the average correlation increases monotonically from 0.132 to 0.241. Thus, the correlations estimated with more precision are higher. These findings are consistent with the assumption used to derive the proxy for correlated earnings forecast errors.

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

TABLE 2

Validation-Earnings Forecast Errors, Analyst Coverage, and Excess Comovement

Panel A of Table 2 reports the average correlations of annual earnings forecast errors by individual analysts between pairs of covered stocks. Panel B reports estimated correlations between common analyst coverage, $\rho_{ij}^{an}$, and the correlation of consensus earnings forecast errors for pairs of stocks, $\rho_{ij}^{error}$. Panel C reports coefficient estimates from regressing returns or alphas on correlated forecast errors, $\rho_{ij}^{error}$. Correlations in Panel A are estimated at the analyst-stock-pair level for pairs of stocks with the same fiscal year-ends for analysts who make revisions on the same date for both stocks. Analyst-stock-pair correlations are included in the mean if there are at least as many observations as the number reported in the first column. In Panel B, only pairs of stocks that share the same fiscal period end date are used. Each month, the consensus quarterly earnings forecast error is estimated. Next, correlations of the consensus earnings forecast error, $\rho_{ij}^{error}$, are calculated for each stock with at least as many months of data as in the first column. Then, for each stock pair, common analyst coverage, $\rho_{ij}^{an}$, is measured during the same period as consensus earnings forecast error. Finally, the correlation coefficient is estimated and is presented along with *p*-values in columns 3 and 4. Earnings forecast error correlations used in Panel C correspond to those from Panel B where at least 5 years of monthly data are available. The correlation coefficients for the dependent variables are estimated yearly using daily raw returns (RAW) or the daily residuals from the capital asset pricing model (CAPM), the Fama–French (1993) 3-factor model (FF3), and the Fama–French 3-factor model augmented with the Carhart (1997) momentum factor (FF4) for each pair of S&P 500 stocks that are in the index during the calendar and following years. Factor loadings for the asset pricing model are estimated using data from the current year. Earnings forecasts data and analyst coverage are from the Institutional Brokers' Estimate System. Returns data are from the Center for Research in Security Prices and factor data from Wharton Research Data Services. Coefficients and standard errors (in parentheses) in Panel C are expressed in percentage terms. The sample covers 1984–2012.

*Panel A. Average Earnings Forecast Error Correlation by Analyst*

| Min. No. of Revisions Used to Estimate Correlations | No. of Corr. in Mean | Avg. Corr. | Std. Err. |
|---|---|---|---|
| 5 | 21,151 | 0.132 | 0.004 |
| 10 | 5,295 | 0.173 | 0.006 |
| 15 | 2,006 | 0.199 | 0.009 |
| 20 | 966 | 0.233 | 0.012 |
| 25 | 498 | 0.238 | 0.016 |
| 30 | 253 | 0.241 | 0.022 |

*Panel B. corr($\rho_{ij}^{an}$, $\rho_{ij}^{error}$) Using Consensus*

| Min. No. of Months Used to Estimate $\rho_{ij}^{error}$ | *N* | Corr. | *p*-Value |
|---|---|---|---|
| 24 | 93,093 | 0.043 | <0.0001 |
| 48 | 6,534 | 0.104 | <0.0001 |
| 60 | 2,869 | 0.109 | <0.0001 |
| 72 | 851 | 0.172 | <0.0001 |
| 96 | 151 | 0.306 | 0.0001 |
| 120 | 23 | 0.463 | 0.0261 |

*Panel C. Comovement and Correlated Forecast Errors*

| | RAW | CAPM | FF3 | FF4 |
|---|---|---|---|---|
| $\rho_{ij}^{error}$ | 7.06 | 3.75 | 6.78 | 7.23 |
| | (1.63) | (1.72) | (1.66) | (1.68) |
| No. of obs. | 2,869 | 2,869 | 2,869 | 2,869 |

Panel B of Table 2 provides direct evidence that common analyst coverage is a good proxy for correlated earnings forecast errors. This analysis is restricted to pairs of stocks with the same fiscal period-end and with the same earnings announcement date. Each month, consensus earnings forecast errors are calculated for both stocks as the arithmetic average earnings forecast minus the actual earnings. Next, the correlation of the consensus earnings forecast error, $\rho_{ij}^{error}$, is estimated for each stock pair with at least 2, 4, 5, 6, 8, and 10 years of monthly observations. Naturally, the number of observations drops as the constraint increases, from 93,093 pairs to 23 pairs. Finally, the correlation between $\rho_{ij}^{an}$ and $\rho_{ij}^{error}$ is calculated. As the length of the time series is increased, the correlation between the proxy and the direct measure increases monotonically, from 0.043 to 0.463. Again, where estimation error is likely to be the smallest, the correlation between the proxy and the direct measure is the highest.

In Panel C of Table 2, I use the direct measure to show that there is a positive relation between correlated earnings forecast errors and excess comovement. In this panel, I use the subsample of pairs of stocks from Panel B with at least 60 months of data. For these 2,869 observations, I regress four comovement measures, RAW, CAPM, FF3, and FF4, on the direct measure, $\rho_{ij}^{error}$. The coefficients on $\rho_{ij}^{error}$ in the univariate regressions vary between 3.75 and 7.23, and all are sta-

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

tistically significant. In all four cases, there is a strong and statistically significant relation between correlated forecast errors and excess comovement. Taken together, all three tests provide support for i) the use of common analyst coverage as a proxy for correlated forecast errors and ii) the positive relation between correlated earnings forecast errors and excess comovement. The Appendix provides additional supporting evidence using the direct measure.

# IV. Results

## A. Cross-Sectional Tests

### 1. Analyst Coverage and Excess Comovement

To test the hypothesis that an increase in common analyst coverage is associated with an increase in comovement, I regress excess comovement on the common analyst coverage variable along with various sets of control variables. To rule out endogeneity in analyst coverage, I focus on whether these variables can predict comovement the following year by lagging all of the independent variables 1 year.[9] Table 3 presents the results of these regressions using each of the following three methods to estimate excess comovement: using raw returns, the CAPM, and the FF4.[10]

The first three columns provide estimates of the basic model that includes only common analyst coverage and year fixed effects. In all three cases, the coefficients, expressed as percentages, are positive and highly significant. A 1% increase in common analyst coverage is associated with an increase in comovement of 0.34% and excess comovement of 0.46% (under the CAPM) and 0.41% (under the FF4). In economic terms, a 1-standard-deviation increase in common analyst coverage is associated with an increase in excess comovement of about 3% to 4% during the subsequent year, a significant increase. For example, under the CAPM, starting at the average level of excess comovement, this means an increase from 8.6% to 12.4%. The standard errors, which are reported in parentheses, show that the coefficients are all statistically significant at the 1% level. To avoid a potential bias in the estimated standard errors resulting from using a panel data set, I follow the advice in Thompson (2011) and cluster by year and firm.[11] The adjusted $R^2$ is 0.43 when using raw returns, 0.34 when using the CAPM, and 0.50 when using the FF4 to define excess comovement. Therefore, using these asset pricing models as a baseline, there is a positive relation between common analyst coverage and excess comovement.

Another way to evaluate the economic significance of an increase in comovement from 8.6% to 12.4% is to examine the impact of this change on an investor's choice of portfolio weights.[12] A loss in the benefits of diversification tends to reduce weights in the optimal portfolio. Under the CAPM, with reasonable assumptions, a change in stock return correlations of this magnitude between a pair

[9] Results are similar when using contemporaneous data.

[10] Excess comovement defined using the FF3 is not included in most regressions for the sake of brevity. Results using this model are very similar to those using the FF4.

[11] Clustering by year, firm pair, and firm does not significantly change standard errors.

[12] I am grateful to the anonymous referee for making this suggestion.

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

TABLE 3

Excess Comovement and Analyst Coverage

Table 3 reports ordinary least squares regression coefficients and standard errors from regressing 1-year-ahead stock return correlations between pairs of stocks on combinations of common analyst coverage, $\rho_{ij}^{an}$; common mutual fund holdings, $\rho_{ij}^{mf}$; and various control variables. Common analyst coverage is defined as $N_{i,j}/\sqrt{N_i N_j}$, where $N_{i,j}$ is the number of common analysts for stocks $i$ and $j$, and $N_i$ and $N_j$ are the total number of analysts covering stocks $i$ and $j$, respectively. Common mutual fund holdings is defined similarly, substituting mutual fund investors for analysts. GICS is an indicator variable that is equal to 1 if the pair of stocks share the same 6-digit Global Industry Classification Standard code, and 0 otherwise; DLNSIZE is the absolute value difference in the natural logarithms of the market values of equity; DLNBTM is the absolute value difference in the natural logarithms of the book-to-market equity ratios; DPRICE is the absolute value difference in lagged stock prices; DLEVERAGE is the absolute value difference in leverage ratios; DROE is the absolute value difference in returns on equity; and STATE is a dummy variable indicating the headquarters of the two firms are in the same state. Year fixed effects are included but not reported. The correlation coefficients are estimated yearly using daily raw returns (RAW) or the daily residuals from the capital asset pricing model (CAPM) and the Fama–French (1993) 3-factor model augmented with the Carhart (1997) momentum factor (FF4) for each pair of S&P 500 stocks that are in the index during the calendar and following years. Factor loadings for the asset pricing model are estimated using data from the previous year. Returns data are from the Center for Research in Security Prices and factor data from Wharton Research Data Services. Analyst forecast data are from the Institutional Brokers' Estimate System. Standard errors are clustered at the firm and year levels and are listed in parentheses. All numbers are expressed in percentage terms except for number of observations and adjusted $R^2$. There are 970,611 observations used from 1995–2007.

| Variable | RAW | CAPM | FF4 | RAW | CAPM | FF4 | RAW | CAPM | FF4 |
|---|---|---|---|---|---|---|---|---|---|
| $\rho_{ij}^{an}$ | 33.50 | 45.60 | 40.80 | 19.80 | 28.20 | 25.40 | 18.40 | 25.80 | 22.90 |
| | (3.63) | (4.34) | (3.86) | (4.09) | (4.79) | (4.35) | (3.78) | (4.17) | (3.69) |
| GICS | | | | 10.30 | 12.90 | 11.40 | 10.00 | 12.40 | 10.80 |
| | | | | (1.92) | (2.18) | (2.16) | (1.77) | (1.96) | (1.97) |
| DLNSIZE | | | | −0.07 | −0.64 | −0.53 | 0.82 | 0.90 | 1.10 |
| | | | | (0.18) | (0.14) | (0.20) | (0.28) | (0.49) | (0.57) |
| DLNBTM | | | | −1.27 | −1.40 | −1.13 | −1.07 | −1.06 | −0.77 |
| | | | | (0.25) | (0.40) | (0.48) | (0.24) | (0.33) | (0.42) |
| DPRICE | | | | −2.65 | −2.37 | −3.41 | −2.66 | −2.40 | −3.44 |
| | | | | (0.32) | (0.58) | (0.65) | (0.35) | (0.62) | (0.64) |
| DLEVERAGE | | | | −1.31 | −8.67 | −6.22 | −0.87 | −7.91 | −5.42 |
| | | | | (1.34) | (2.44) | (2.43) | (1.27) | (2.24) | (2.26) |
| DROE | | | | −1.20 | 1.05 | 0.11 | −1.32 | 0.84 | −0.12 |
| | | | | (0.56) | (0.62) | (0.55) | (0.56) | (0.59) | (0.50) |
| STATE | | | | 1.04 | 0.59 | 0.29 | 1.03 | 0.57 | 0.27 |
| | | | | (0.28) | (0.48) | (0.55) | (0.28) | (0.47) | (0.54) |
| $\rho_{ij}^{mf}$ | | | | | | | 17.20 | 29.60 | 31.50 |
| | | | | | | | (2.74) | (7.47) | (7.51) |
| $N$ | 970,611 | 970,611 | 970,611 | 838,562 | 838,562 | 838,562 | 838,562 | 838,562 | 838,562 |
| Adj. $R^2$ | 0.433 | 0.336 | 0.502 | 0.456 | 0.355 | 0.515 | 0.463 | 0.371 | 0.527 |

of stocks reduces the optimal portfolio weights of the two stocks by about 4%. If there is a similar increase in all stocks in an industry of 10 stocks, as might happen as the result of an additional analyst, the optimal portfolio weights for all 10 stocks drop by about 17%.[13]

## 2. Industries and Characteristics

To rule out the alternative explanation that analysts tend to cover firms with systematic risk characteristics that are not captured by our asset pricing models, I perform several tests. In the first, most direct test, I include an industry-related

[13] In the CAPM, with an investor with negative exponential utility and risk aversion parameter, $\gamma$, and a risk-free return of 0, the optimal portfolio weights are proportional to $(1/\gamma)\Sigma^{-1}(\overline{z})$, where $\Sigma$ is the variance–covariance matrix of returns and $\overline{z}$ is a vector of expected returns (see Ingersoll (1987), p. 98). Setting $\gamma$ to 3 and the expected returns, variances, and covariances equal to their sample means for 500 stocks and then varying the covariance such that correlations increase from 8.5% to 12.4% yield these results.

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

dummy variable in the equation and control for other firm characteristics that may proxy for risk. Specifically, I create the variable GICS, which is an indicator variable set to 1 if the two firms share the same 6-digit GICS code, and 0 otherwise. Boni and Womack (2006) show that the partitions generated using GICS classification industries are a good proxy for the way analysts define themselves by their coverage choices. Additionally, I include the variables DLNSIZE, DLNBTM, DPRICE, DLEVERAGE, DROE, and STATE. If characteristics explain risk above and beyond that explained by factor loadings and GICS industry, the coefficients on these additional variables should be negative, with the exception of the state dummy, which should be positive.

The middle three columns in Table 3 provide the results when controlling for these additional variables. A pair of stocks in the same GICS industry has excess comovement of between 10% and 13% more than those from different industries. Therefore, there does seem to be some level of industry-related comovement not captured by the asset pricing models. Additionally, all of the control variables either have the expected sign or are not statistically significant. However, using all three measures of comovement, the coefficient on the common analyst coverage variable remains positive and significant, though magnitudes are smaller, ranging between 19.8% and 28.2%. To the extent that these industry definitions and characteristics control for unmeasured risk, excess comovement increases with common analyst coverage.

### 3. Institutional Price Pressure

It is possible that simultaneous price pressure from mutual funds or other institutional investors is driving the results. If mutual funds tend to specialize by industry or if analysts cater to institutions, common analyst coverage could be indirectly capturing this. To rule this out, I add common mutual fund holdings, $\rho_{ij}^{\text{mf}}$, to the analysis and examine whether common analyst coverage still has a significant effect on comovement. The final three columns of Table 3 provide estimates of these models. The coefficients on common analyst coverage drop slightly to 18.4% using raw returns, 25.8% using the CAPM, and 22.9% using the FF4, but all three are still very statistically significant. In fact, the precision of the estimates is larger.

The regression coefficients of the mutual fund measure are positive and statistically significant in all three cases, between 17% and 32%. A 1-standard-deviation increase in common mutual fund holdings is associated with an increase in excess comovement of about 2% to 3% during the following year. This may be due to simultaneous price pressure or mutual fund managers choosing similar stocks, above and beyond that explained by similar characteristics and factor loadings. Analysts may indeed cater to particular investors, but the positive coefficient on $\rho_{ij}^{\text{an}}$ does not seem to be simply a proxy for the trading behavior of mutual funds.[14] Though not reported here, the same is true of the trading behavior of institutions as measured by their 13F filings.

[14] In untabulated results, I examine potential nonlinearity in the impact of mutual fund holdings and find that the coefficient on common analyst coverage is virtually unchanged.

## B. Exogenous Changes in Analyst Coverage

### 1. Endogeneity

The cross-sectional tests cannot completely rule out endogeneity in analyst coverage as an alternative hypothesis for the relation between analyst coverage and excess comovement. As mentioned, if the asset pricing models we use to control for risk are imperfect, and analysts choose to cover similar firms, the positive relation between excess comovement and common analyst coverage may simply be mismeasured systematic risk. An ideal test would involve randomly reassigning firms to analysts and examining the resulting changes in comovement. Though this is not entirely possible, such a natural experiment is approximated by the change in analyst coverage resulting from i) mergers of brokerage houses and ii) additions of stocks to stock market indexes. In this section, I use both types of events to provide additional evidence that correlated forecast significantly contribute to excess comovement.

### 2. Turnover around Brokerage House Mergers

Mergers of brokerage houses have been used in the literature as natural experiments that are not related to the fundamentals of stocks covered by the associated analysts. For example, Hong and Kacperczyk (2010) use exogenous analyst turnover around 15 mergers to examine the effect of competition on analyst earnings forecast bias. Following a merger, analyst coverage may change due to analyst turnover at the resulting firm. Coverage may increase or decrease depending on how assignments are allocated to the remaining analysts. Because nothing is likely to have changed in terms of the fundamental risk characteristics of the underlying firms covered by these brokerage houses, these events provide an ideal way to test whether analyst coverage leads to increased excess comovement and is not simply capturing unmeasured risk.

Wu and Zang (2009) show that turnover following brokerage house mergers is somewhat mechanical. When there are redundant analysts, the target firm's analyst is much more likely to be removed than is the acquirer's analyst. Conversely, the analyst reshuffling that occurs around mergers may offer remaining analysts the opportunity to initiate coverage of stocks that have recently become more similar to their current set of stocks. Furthermore, managers' choices of which analysts to dismiss and which stocks to reassign are not entirely mechanical. Though unlikely, these decisions may be associated with changes in the fundamentals of the underlying stocks. As a result, I strategically choose pairs of stocks in my sample to rule out the possibility that this behavior is driving the empirical results.

First, to rule out any strategic behavior in the choice of which analysts are removed, I include only pairs of stocks with coverage by an analyst at the acquirer and an analyst at the target before the merger, and by only one of the two original analysts after the merger. The analyst dropping coverage must have had his or her employment terminated by the merged firm. I exclude any cases in which another analyst from the merged firm initiates coverage of the stock and in which both of the analysts remain at the firm or both are fired. As a result, the decision of which analyst to terminate is irrelevant. The change in the proxy for correlated forecast errors will be the same regardless of which analyst is removed. Limiting

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

the sample to these stocks should also rule out the role of the brokerage firm in choosing which stocks are affected. Included are stocks that had coverage by both firms before the merger.

Second, to rule out the outside chance that pairs of stocks in my sample become fundamentally less similar around the mergers, I further restrict the sample to stocks that had no change in analyst coverage by analysts other than the two employed by the merging firms. Because analyst coverage remains relatively constant, these sets of stocks are those least likely to have had changes in fundamentals around the merger. In other words, under the alternative hypothesis of endogeneity in the choice of analyst coverage, this criterion chooses the set of stocks least likely to be affected.

To summarize, using the 15 mergers in Hong and Kacperczyk (2010), I choose the subset of stocks that meet the following three criteria: i) covered by one analyst from each of the two brokerage firms before the merger; ii) covered by one of these two analysts from the merged firm after the event, with the remaining analyst leaving the firm; and iii) no change in coverage from analysts unaffiliated with the two merging firms. In other words, the stock loses coverage from exactly one analyst: the redundant analyst from one of the two merging firms. There are a total of 1,726 stock-merger observations across the 15 mergers that meet the first two criteria.[15] Most do not meet the third during the pre- and postevent window. Those that meet all three criteria result in 3,567 pairwise observations. For each pair of stocks, I calculate changes (denoted with a delta, $\Delta$) in common analyst coverage, common mutual fund holdings, and excess comovement before and after the events, and I examine the effect of the changes in analyst coverage on comovement. Additionally, because institutional holdings may change around these mergers, and this may potentially be related to the actual merger itself, I use the Thomson Reuters Institutional Holdings (13F) database to create a common institutional holdings variable, $\Delta\rho_{ij}^{\text{in}}$, which is analogous the mutual fund variable, and include it in the regression, as well.

Factor loadings are estimated using data beginning during the year before the preevent window. Panel A of Table 4 presents summary statistics for these changes around the 15 mergers for all stocks that meet the criteria. The change in analyst coverage by merging firms results in an average change in common analyst coverage of $-14\%$. Because I have included in my sample only stocks that lose a single analyst, the changes are all negative, ranging from $-82\%$ to $0\%$. The standard deviation for the change in analyst coverage is about 10.8%. Mutual fund and institutional holdings variables do not change much on average, only about 2.2% and $-1.6\%$, respectively, but both have large dispersions. Raw return correlations decrease around the event by 1%, on average. After controlling for risk, excess comovement decreases on average when using the CAPM ($-0.3\%$) and the FF3 ($-0.4\%$) and increases on average using the FF4 (3.9%). However, the standard deviations of the change in excess comovement are relatively large, 15% to 20%. Twelve-month windows are used before and after the mergers, but results are similar using 6-month windows.

[15] I thank Marcin Kacperczyk and Harrison Hong for making these data available.

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

TABLE 4

Excess Comovement: Brokerage House Mergers

Table 4 reports summary statistics (Panel A) and ordinary least squares regression coefficients and standard errors (Panel B) from regressing the change in stock return correlations between stocks on changes in common analyst coverage, $\Delta\rho_{ij}^{an}$; common mutual fund holdings, $\Delta\rho_{ij}^{mf}$; and common institutional holdings, $\Delta\rho_{ij}^{in}$, before and after 15 brokerage house mergers between 1994 and 2005 (see Hong and Kacperczyk (2010), Table III for a list). Firm fixed effects are included, but not reported. The common analyst coverage variable is calculated using the Institutional Brokers' Estimate System database and equation (5) in the text. The correlation coefficients are estimated using daily raw returns (RAW) or residuals from the capital asset pricing model (CAMP alphas), the Fama–French (1993) 3-factor model (FF3 alphas), and a model using the Fama–French factors plus the Carhart (1997) momentum factor (FF4 alphas), for 6 months before and after a merger. Returns data are from the Center for Research in Security Prices and factor data from Wharton Research Data Services. Mutual fund and institutional holdings data are from the Thomson Reuters databases as described in the data section. There are 3,567 observations used. Stocks are included only if i) they are covered by an analyst from both merging firms before the merger, ii) they are covered by a single analyst from the merged firm after the merger, and iii) analyst coverage from nonmerging firms does not change. All numbers are expressed in percentage terms except for number of observations and adjusted $R^2$. Standard errors are in parentheses.

*Panel A. Summary Statistics*

| Variable | Change in: | *N* | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|---|
| $\Delta\rho_{ij}^{an}$ | Common analyst coverage | 3,567 | −14.0 | 10.8 | −81.6 | 0.0 |
| $\Delta\rho_{ij}^{mf}$ | Common mutual fund holdings | 3,567 | 2.2 | 12.1 | −67.3 | 66.9 |
| $\Delta\rho_{ij}^{in}$ | Common institutional holdings | 3,567 | −1.6 | 17.5 | −80.1 | 62.9 |
| $\Delta\rho_{ij}^{ret}$ | Raw return correlations | 3,567 | −1.0 | 15.9 | −150.3 | 112.9 |
| $\Delta\rho_{ij}^{capm}$ | CAPM alpha correlations | 3,567 | −0.3 | 15.6 | −128.4 | 116.2 |
| $\Delta\rho_{ij}^{ff3}$ | FF3 alpha correlations | 3,567 | −0.4 | 16.2 | −106.9 | 130.0 |
| $\Delta\rho_{ij}^{ff4}$ | FF4 alpha correlations | 3,567 | 3.9 | 19.8 | −148.0 | 108.9 |

*Panel B. Regression Coefficients*

| Variable | RAW | | | CAPM Alphas | | |
|---|---|---|---|---|---|---|
| $\Delta\rho_{ij}^{an}$ | 10.4 | 9.9 | 8.9 | 7.2 | 6.9 | 6.3 |
| | (2.7) | (2.7) | (2.7) | (2.7) | (2.7) | (2.7) |
| $\Delta\rho_{ij}^{mf}$ | | 12.4 | | | 7.5 | |
| | | (2.4) | | | (2.4) | |
| $\Delta\rho_{ij}^{in}$ | | | 12.0 | | | 6.5 |
| | | | (1.8) | | | (1.8) |
| Adj. $R^2$ | 0.242 | 0.248 | 0.252 | 0.227 | 0.229 | 0.230 |
| Variable | FF3 Alphas | | | FF4 Alphas | | |
| $\Delta\rho_{ij}^{an}$ | 11.8 | 11.5 | 10.6 | 12.5 | 12.1 | 11.5 |
| | (2.8) | (2.8) | (2.8) | (3.1) | (3.1) | (3.1) |
| $\Delta\rho_{ij}^{mf}$ | | 8.1 | | | 8.5 | |
| | | (2.5) | | | (2.7) | |
| $\Delta\rho_{ij}^{in}$ | | | 8.9 | | | 7.7 |
| | | | (1.9) | | | (2.0) |
| Adj. $R^2$ | 0.236 | 0.239 | 0.151 | 0.382 | 0.384 | 0.385 |

To examine whether exogenous changes in analyst coverage lead to increases in excess comovement, I run a cross-sectional regression of the change in (excess) comovement on changes in common analyst coverage, $\Delta\rho_{ij}^{an}$. The first columns of each quadrant in Panel B of Table 4 present the coefficients, standard errors, and overall fit of these regressions using each of the four measures of comovement. The first quadrant shows that increases in common analyst coverage that result around brokerage firm mergers are associated with a marginal increase in raw return correlation of 10.4%. The final three quadrants show that after controlling for risk, this effect is still strong: 7.2% using CAPM alphas, 11.8% using FF3 alphas, and 12.5% using FF4 alphas. All of these coefficients are statistically significant at the 1% level and adjusted $R^2$ statistics range between 23% and 38%.

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

If individual analysts cater to institutional investors, it could be that the change in the analyst coverage variable is capturing the effect of a change in investors. With this motivation, I include the change in common mutual fund holdings, $\Delta\rho_{ij}^{\text{mf}}$, and common institutional holdings, $\Delta\rho_{ij}^{\text{in}}$, in the regression in the second column of each panel. The coefficients on both the mutual fund variable and institutional investor variables are positive and statistically significant in explaining the cross section of changes in comovement around these mergers. However, the coefficients on the change in the analyst coverage variable are essentially unaffected by the addition of these variables. Thus, there is a positive and significant relation between common analyst coverage and excess comovement around brokerage house mergers, suggesting that changes in the information structure itself are responsible for some of the changes in excess comovement. In the Appendix, I repeat this exercise using correlations in forecast errors and find consistent results.

## 3. Additions to the S&P 500 Index

Barberis et al. (2005) find that stocks' betas with respect to the S&P 500 index tend to increase after being added to that index. They argue that nothing fundamental changes with the stock in the short window over which they estimate these betas and ascribe the increased correlation to the trading behavior of "style investors," investors who trade stocks based on characteristics or classifications. Style investors and institutional investors buying and selling stocks to track the S&P 500 index may very well result in an increase in a firm's measured beta. However, there may be another source of comovement. A change in the information production environment may also lead to changes in comovement.

Hedge and McDermott (2003) show that additions to the S&P 500 index are associated with increases in analyst coverage. If these new analysts are also those covering other S&P 500 stocks, forecast errors may become more correlated. With this in mind, I calculate $\rho_{ij}^{\text{an}}$ each quarter from 12 quarters before to 12 quarters after the event. To avoid cases in which firms were simply not initially in the sample, I use data from the 647 firms from 1984 through 2011 that had at least one analyst in the IBES database before the event. Table 5 shows the average values

TABLE 5

Common Analyst Coverage before and after Additions to the S&P 500 Index

Table 5 reports the average common analyst coverage measure for stocks, $\overline{\rho_{ij}^{\text{an}}}$, before and after being added to the S&P 500 index, plus the percentage change and difference in these means, and the *t*-statistic and *p*-value for these differences. The means are calculated for the 1st quarter before and after the event through the 12th quarter before and after. There are 687 stocks used from 1984 through 2011. Means and differences are expressed in percentage terms.

| Qtr. | Before | After | %Change | Diff. | *t*-Stat. | *p*-Value |
|---|---|---|---|---|---|---|
| 1 | 1.01 | 1.08 | 6.4 | 0.06 | 2.0 | 0.0215 |
| 2 | 0.98 | 1.07 | 8.8 | 0.09 | 2.6 | 0.0044 |
| 3 | 0.94 | 1.09 | 16.0 | 0.15 | 4.6 | <0.0001 |
| 4 | 0.92 | 1.08 | 17.6 | 0.16 | 5.0 | <0.0001 |
| 5 | 0.91 | 1.07 | 17.2 | 0.16 | 4.8 | <0.0001 |
| 6 | 0.90 | 1.07 | 18.6 | 0.17 | 5.0 | <0.0001 |
| 7 | 0.87 | 1.05 | 21.1 | 0.18 | 5.5 | <0.0001 |
| 8 | 0.84 | 1.04 | 23.6 | 0.20 | 6.0 | <0.0001 |
| 9 | 0.79 | 1.03 | 29.5 | 0.23 | 7.1 | <0.0001 |
| 10 | 0.79 | 1.04 | 31.2 | 0.25 | 7.0 | <0.0001 |
| 11 | 0.77 | 1.02 | 33.0 | 0.25 | 7.2 | <0.0001 |
| 12 | 0.70 | 1.03 | 48.3 | 0.34 | 9.6 | <0.0001 |

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

before and after the event, plus the percentage increases, differences, $t$-statistics, and $p$-values. On average, these measures of correlated information increase when a stock is added to the index. There is an 18% increase in $\rho_{ij}^{\text{an}}$ for these new S&P 500 stocks from 1 year before to 1 year after the addition. From 12 quarters before to 12 quarters after, there is a 48% increase. Figure 1 plots this variable in event time. There is a clear increase in common analyst coverage leading up to the event. On average, the common analyst coverage increases with respect to the S&P 500 stocks when stocks are added to the index.

If the increase in correlated forecast errors is responsible for some of the increase in return comovement, we would expect this increase to be largest where the increase in common analyst coverage is greatest. Accordingly, I regress the changes in excess comovement defined by the four measures on changes in the analyst coverage variable from the calendar years before and after the event. Panel A of Table 6 presents summary statistics for the 273,547 pairwise observations. The first four rows show summary statistics for changes in the comovement, $\Delta\rho_{ij}^{\text{ret}}$, $\Delta\rho_{ij}^{\text{capm}}$, and $\Delta\rho_{ij}^{\text{ff4}}$. Average changes in comovement are 4.39%, 0.59%, and 0.34%, respectively. The next row shows that common analyst coverage with respect to S&P 500 stocks increases by about 0.06% on average. However, these changes are as large as 49.66%. The average change in common mutual fund holdings is 10.97%. Panel B presents results from regressing changes in comovement on changes in analyst coverage and mutual fund coverage around the additions. In the simple regression of $\Delta\rho_{ij}^{\text{ret}}$ on $\Delta\rho_{ij}^{\text{an}}$, the coefficients are 2.61% using raw returns and more than 6% using either the CAPM or the FF4. All three coefficients are statistically significant. As shown in the final three columns, changes in mu-

FIGURE 1
Common Analyst Coverage: Additions to S&P 500 Index

Figure 1 presents the average common analyst coverage, $\rho_{ij}^{\text{an}}$, measured quarterly, between firms added to the S&P 500 index from 1982 through 2011 and all other S&P 500 firms by quarter from the 12th quarter before the addition date to the 12th quarter after the date. Analyst data are from the Institutional Brokers' Estimate System database. S&P 500 constituent data are from the Center for Research in Security Prices. The values plotted in this graph are found in Table 5.

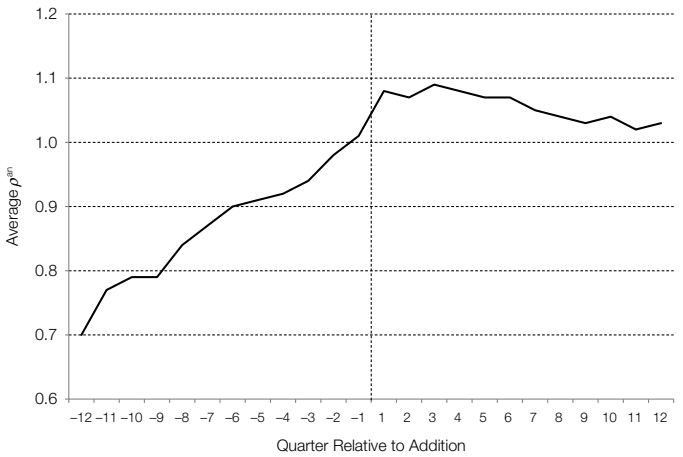


TABLE 6

Excess Comovement: Additions to the S&P 500 Index

Table 6 reports summary statistics (Panel A) as well as ordinary least squares regression coefficients and standard errors from regressing the change in stock return correlations between stocks that are added to the S&P 500 index and all other stocks in the index on changes in common analyst coverage, $\Delta\rho_{ij}^{an}$, before and after the addition date (Panel B) and from regressing changes in S&P 500 betas of added stocks on changes in average common analyst coverage between the added stock and all other S&P 500 stocks (Panel C). Firm fixed effects are included in the regressions in Panel A and year fixed effects in Panels A and B. The common analyst coverage variable is calculated using the Institutional Brokers' Estimate System database and equation (5). The correlation coefficients are estimated using daily raw returns (RAW), residuals from the capital asset pricing model (CAPM), or residuals from a model using the Fama–French (1993) factors plus the Carhart (1997) momentum factor (FF4). S&P 500 betas are estimated using a year's worth of daily returns on the equal-weighted S&P 500 index as in Barberis, Shleifer, and Wurgler (2005). Returns data are from the Center for Research in Security Prices and factor data from Wharton Research Data Services. Mutual fund holdings data are from the Thomson Reuters databases as described in the data section. There are 273,547 pairwise observations used in Panel B, which correspond to the 647 stock-level observations used in Panel C. All numbers are expressed in percentage terms except for number of observations and adjusted $R^2$. Standard errors are in parentheses.

*Panel A. Summary Statistics*

| Variable | $N$ | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| $\Delta\rho_{ij}^{ret}$ | 273,547 | 4.39 | 15.37 | −76.25 | 79.39 |
| $\Delta\rho_{ij}^{capm}$ | 273,547 | 0.59 | 10.95 | −68.39 | 67.61 |
| $\Delta\rho_{ij}^{ff4}$ | 273,547 | 0.34 | 10.21 | −64.94 | 61.41 |
| $\Delta\rho_{ij}^{an}$ | 273,547 | 0.06 | 1.73 | −41.94 | 49.66 |
| $\Delta\rho_{ij}^{mf}$ | 273,547 | 10.97 | 10.99 | −35.73 | 100.00 |
| $\Delta\beta_i^{SP500}$ | 647 | 0.16 | 0.48 | −1.92 | 2.82 |
| $\Delta\overline{\rho_{ij}^{an}}$ | 647 | 0.05 | 0.24 | −0.89 | 1.43 |

*Panel B. Changes in Stock Return Correlations*

| Variable | RAW | CAPM | FF4 | RAW | CAPM | FF4 |
|---|---|---|---|---|---|---|
| $\Delta\rho_{ij}^{an}$ | 2.61 | 6.71 | 6.30 | 2.62 | 6.72 | 6.30 |
| | (1.09) | (1.19) | (1.13) | (1.09) | (1.19) | (1.13) |
| $\Delta\rho_{ij}^{mf}$ | | | | 2.99 | 2.28 | 1.73 |
| | | | | (0.25) | (0.28) | (0.26) |
| No. of obs. | 273,547 | 273,547 | 273,547 | 273,547 | 273,547 | 273,547 |
| Adj. $R^2$ | 0.625 | 0.053 | 0.015 | 0.625 | 0.053 | 0.016 |

*Panel C. Changes in Betas:* $\Delta\beta_i^{SP500} = \alpha + \gamma \times \Delta\overline{\rho_{ij}^{an}} + e$

| Variable | 1984–2011 | 1984–1997 | 1998–2011 |
|---|---|---|---|
| $\Delta\overline{\rho_{ij}^{an}}$ | 20.64 | 11.82 | 29.68 |
| | (5.11) | (5.97) | (8.04) |
| Std. coeff. | 0.156 | 0.117 | 0.191 |
| No. of obs. | 647 | 278 | 369 |
| Adj. $R^2$ | 0.244 | 0.207 | 0.264 |

tual fund holdings are associated with changes in comovement, but their inclusion does not diminish the effect of changes in analyst coverage.

Though this article defines comovement as a pairwise measure, it is natural to examine how much of the increase in S&P 500 betas can be explained by changes in analyst coverage. Although the results from Panel B of Table 6 are suggestive, it is not obvious that the relation will hold using covariances with the index. Therefore, I estimate value-weighted S&P 500 betas for stocks before and after their addition as well as the average common analyst coverage between added stocks and all S&P 500 stocks before and after the addition. This latter measure is calculated by taking the equal-weighted average of $\rho_{ij}^{an}$ between each stock, $i$, added to the index and the other stocks in the index around the addition. The final two rows of Panel A provide summary statistics for changes in these two variables around the additions.

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

On average, S&P 500 betas increase by about 0.16 for newly added stocks, with a standard deviation of 0.48. This is roughly the same average change that Barberis et al. (2005) document for 1997–2000 using daily returns. The change in average common analyst coverage between the added stock and all other stocks in the S&P 500 index is 0.05, with a standard deviation of 0.24. I regress changes in beta, $\Delta\beta_i^{SP500}$, on changes in average common analyst coverage, $\Delta\overline{\rho_{ij}^{an}}$, and year fixed effects for the entire period, as well as for the subperiods 1984–1997 and 1998–2011. Panel C of Table 6 provides coefficient estimates and standard errors as well as standardized coefficients.

There is a positive, statistically significant impact of changes in common analyst coverage on changes in S&P 500 betas for the entire period. Furthermore, this effect is statistically significant in both subperiods, and much stronger from 1998 to 2011. The standardized coefficient measures show that a 1-standard-deviation cross-sectional increase in common analyst coverage is associated with an increase in S&P 500 beta of about 15% of its standard deviation, or about 0.07. This is about half the magnitude of the average change in betas. The standardized coefficient is 0.117 in the first subperiod and 0.191 in the second. Moreover, according to the adjusted $R^2$, the models explain between 20% and 26% of the variation in changes in S&P 500 betas. Therefore, depending on the measure and period, between about 11% and 25% of the change in betas around additions to the S&P 500 index can be explained by changes in analyst coverage. Overall, the effects of changes in common analyst coverage on excess comovement are consistently positive and significant around additions to the index, using either pairwise correlations or betas, suggesting that increases in analyst coverage do lead to increases in comovement that are distinct from changes in fundamental risk characteristics. The Appendix repeats this exercise using correlations in forecast errors and finds similar results.

## C. Alternative Measures of Analyst Coverage

In this section, I modify the basic common analyst coverage variable to further rule out that endogeneity in analysts' choice of coverage is driving the result. First, I place more emphasis on important analysts or on analysts with more precise forecasts. Second, I minimize the endogeneity problem by focusing on different analysts at the same firm or analysts who once worked together at the same firm.

### 1. Analyst Heterogeneity

One of the assumptions used to derive common analyst coverage is that analysts' forecasts are equally precise. In reality, this is not true. If the excess comovement associated with analyst coverage is driven by information, and not by analysts' coverage choices, it ought to be the most pronounced among the analysts whose forecasts are given the most attention by investors. Conversely, if the result is driven by analysts choosing similar stocks, there should be no difference. To test this information hypothesis, I regress excess comovement on the common star analyst coverage measures, $\rho_{ij}^{st}$ and $\rho_{ij}^{-st}$, along with the same control variables from the base case. Results are in Table 7. In the first three columns, I examine the effect of common star analyst coverage without controlling for nonstar analysts. On the margin, increasing common star analyst coverage has a strong effect on

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

TABLE 7

Comovement and Analyst Coverage: Alternate Measures

Table 7 reports ordinary least squares regression coefficients and standard errors from regressing 1-year-ahead stock return correlations between pairs of stocks on combinations of common star analyst coverage, $\rho_{ij}^{st}$; common nonstar analyst coverage, $\rho_{ij}^{-st}$; the precision-weighted analyst coverage measure, $\rho_{ij}^{wgt}$; common mutual fund holdings, $\rho_{ij}^{mf}$; and various control variables. Common star analyst coverage is defined as $N_{i,j}^{st}/\sqrt{N_i N_j}$, where $N_{i,j}^{st}$ is the number of common star analysts for stocks $i$ and $j$, and $N_i$ and $N_j$ are the total number of analysts covering stocks $i$ and $j$, respectively. Common nonstar analyst coverage is defined similarly, substituting nonstar analysts ($N_{i,j}^{-st}$) in the numerator. The precision-weighted measure is defined in equation (6). The variable $\rho_{ij}^{mf}$ is defined in equation (9); GICS is a dummy variable indicating that the two stocks are in the same Global Industry Classification Standard industry; DLNSIZE is the absolute value difference in the natural logarithms of the market values of equity; DLNBTM is the absolute value difference in the natural logarithms of the book-to-market equity ratios; DPRICE is the absolute value difference in lagged stock prices; DLEVERAGE is the absolute value difference in leverage ratios; DROE is the absolute value difference in returns on equity; and STATE is a dummy variable indicating the headquarters of the two firms are in the same state. Year fixed effects are included but not reported. The correlation coefficients are estimated yearly using daily raw returns (RAW), or the daily residuals from the capital asset pricing model (CAPM) and the Fama–French (1993) 3-factor model augmented with the Carhart (1997) momentum factor (FF4) for each pair of S&P 500 stocks that is in the index during the calendar and following years. Factor loadings for the asset pricing model are estimated using data from the current year. Returns data are from the Center for Research in Security Prices and factor data from Wharton Research Data Services. Analyst data are from the Institutional Brokers' Estimate System, and star analyst data are from *Institutional Investor* magazine. All numbers are expressed in percentage terms except for number of observations and adjusted $R^2$. Standard errors are clustered at the firm and year levels and are listed in parentheses. There are 838,562 observations used from 1995–2007.

| Variable | RAW | CAPM | FF4 | RAW | CAPM | FF4 | RAW | CAPM | FF4 |
|---|---|---|---|---|---|---|---|---|---|
| $\rho_{ij}^{st}$ | 70.40 | 84.90 | 76.60 | 42.30 | 40.90 | 38.10 | | | |
| | (7.55) | (11.19) | (8.79) | (7.16) | (10.16) | (8.90) | | | |
| $\rho_{ij}^{-st}$ | | | | 15.20 | 23.80 | 20.80 | | | |
| | | | | (4.48) | (5.15) | (4.65) | | | |
| $\rho_{ij}^{wgt}$ | | | | | | | 21.20 | 27.60 | 24.40 |
| | | | | | | | (1.48) | (2.42) | (2.38) |
| GICS | 12.10 | 16.00 | 14.00 | 9.59 | 12.10 | 10.60 | 8.67 | 11.20 | 9.84 |
| | (1.19) | (1.30) | (1.44) | (1.57) | (1.73) | (1.76) | (1.15) | (1.14) | (1.21) |
| DLNSIZE | 0.83 | 0.92 | 1.12 | 0.81 | 0.90 | 1.10 | 0.81 | 0.89 | 1.10 |
| | (0.27) | (0.49) | (0.57) | (0.28) | (0.49) | (0.57) | (0.27) | (0.49) | (0.57) |
| DLNBTM | −1.09 | −1.09 | −0.79 | −1.07 | −1.06 | −0.77 | −1.07 | −1.07 | −0.77 |
| | (0.24) | (0.33) | (0.42) | (0.24) | (0.33) | (0.42) | (0.24) | (0.33) | (0.42) |
| DPRICE | −2.67 | −2.41 | −3.46 | −2.66 | −2.39 | −3.44 | −2.65 | −2.39 | −3.44 |
| | (0.35) | (0.61) | (0.64) | (0.35) | (0.62) | (0.64) | (0.35) | (0.61) | (0.64) |
| DLEVERAGE | −0.96 | −8.06 | −5.55 | −0.87 | −7.91 | −5.42 | −0.82 | −7.87 | −5.38 |
| | (1.27) | (2.24) | (2.26) | (1.27) | (2.24) | (2.26) | (1.28) | (2.25) | (2.26) |
| DROE | −1.33 | 0.83 | −0.13 | −1.33 | 0.83 | −0.12 | −1.35 | 0.80 | −0.15 |
| | (0.59) | (0.61) | (0.52) | (0.57) | (0.59) | (0.51) | (0.59) | (0.61) | (0.52) |
| STATE | 1.14 | 0.73 | 0.41 | 1.05 | 0.59 | 0.29 | 1.05 | 0.60 | 0.30 |
| | (0.28) | (0.48) | (0.55) | (0.29) | (0.48) | (0.55) | (0.28) | (0.47) | (0.54) |
| $\rho_{ij}^{mf}$ | 17.60 | 30.30 | 32.10 | 17.10 | 29.60 | 31.50 | 17.00 | 29.40 | 31.30 |
| | (2.60) | (7.46) | (7.54) | (2.71) | (7.49) | (7.53) | (2.60) | (7.52) | (7.61) |
| $N$ | 838,562 | 838,562 | 838,562 | 838,562 | 838,562 | 838,562 | 838,562 | 838,562 | 838,562 |
| Adj. $R^2$ | 0.461 | 0.366 | 0.525 | 0.464 | 0.371 | 0.527 | 0.465 | 0.371 | 0.527 |

excess comovement. The marginal effect of an increase in common star analyst coverage on the raw return correlation between two stocks is 70.4%. Using the CAPM alphas and FF4 alphas to define excess comovement, the marginal effect is 84.9% and 76.6%, respectively. These coefficients are more than 3 times as large as those estimated using the basic measure.

In the three middle columns of Table 7, I include the nonstar analyst variable in the regression. With the addition of this variable, the coefficient on the star analyst measure drops to 42.3% using raw returns, 40.9% using CAPM alphas, and 38.1% using FF4 alphas. The coefficient on the nonstar analyst variable is less than the coefficient on the general measure, but still statistically significant, between 15.2% and 23.8%. Separating the general variable into the two subsets yields results consistent with the hypothesis that the excess comovement associ-

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

ated with common analyst coverage is driven by correlated information, and not simply by unmeasured risk.

As an alternative to the binary, star versus nonstar classification of analysts, I use the precision-weighted analyst coverage measure, $\rho_{ij}^{\text{wgt}}$, to test whether the comovement is driven by commonality in the information production process or by something other than information. Because this more sophisticated measure accounts for the past performance of each analyst, the marginal effect on comovement should be larger than the effect using the equal-weighted measure. The basic measure essentially ignores the heterogeneity in the level of accuracy across analysts. If commonality in information plays a role in comovement, a pair of stocks with one analyst in common who makes uninformative forecasts should have lower comovement than the same pair of stocks would have if the common analyst's forecasts were more precise, but still correlated.

The final three columns of Table 7 provide coefficients from the model using the precision-weighted measure along with the control variables. The marginal effect of an increase in the precision-weighted measure is 21.2%, 27.6%, and 24.4% for the raw returns, CAPM alphas, and FF4 alphas, respectively. These are all slightly larger than was found with the basic measure, though the differences are not statistically significant. In untabulated results, I find that this is also true using the two other specifications from the base case. Thus, accounting for the precision of forecasts using both star analyst rankings and track records provides evidence in favor of the information hypothesis.

### 2. Social Networks and Excess Comovement

Individual analysts tend to cover only a small fraction of the entire universe of stocks, typically a single industry. Any given brokerage firm, however, may employ analysts covering a much larger percentage of the S&P 500 firms. To the extent that analysts working for the same firm use similar methodology, information, or predictions, or process information in the same manner, we might expect their forecast errors to be correlated.[16] The common broker coverage variable, $\rho_{ij}^{\text{br}}$, may provide additional power in explaining excess comovement. Additionally, because of the wider coverage, mismeasured industry-related risk factors are less likely to be an issue.

To test this third hypothesis I repeat the regressions in the third specification of Table 3, replacing the analyst-level variable with the brokerage-level measure. Results are in the middle three columns of Table 8. For reference, the results using the analyst measure are repeated in the first three columns. The coefficients on the brokerage coverage variable are all positive and significant, ranging from about 3% to 8%. Not surprisingly, the effects of this indirect measure of correlated information are smaller than the direct effect associated with the analyst variable. However, because brokerage houses cover a much broader set of stocks than do individual analysts, this is consistent with an information-driven channel driving some level of excess comovement, and not an unmeasured risk story.

[16] See the Appendix for empirical evidence that forecast errors are, indeed, correlated across analysts within the same brokerage firm.

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

TABLE 8

Comovement and Analyst Coverage: Social Networks

Table 8 reports ordinary least squares regression coefficients and standard errors from regressing 1-year-ahead stock return correlations between pairs of stocks on combinations of common analyst coverage, $\rho_{ij}^{an}$; common brokerage coverage, $\rho_{ij}^{br}$; common connected analyst coverage, $\rho_{ij}^{ca}$; common mutual fund holdings, $\rho_{ij}^{mf}$; and various control variables. Common analyst coverage is defined as $N_{i,j}/\sqrt{N_i N_j}$, where $N_{i,j}$ is the number of common analysts for stocks $i$ and $j$, and $N_i$ and $N_j$ are the total number of analysts covering stocks $i$ and $j$, respectively. Common brokerage coverage is defined similarly using brokerage firms in lieu of analysts. Common connected analyst coverage is equal to $N_{i,j}^{ca}/\sqrt{N_i N_j}$, where $N_{i,j}^{ca}$ is the number of social connections between analysts covering stocks $i$ and $j$. Analysts are connected if both worked at the same firm at the same time in the past but no longer do. The variable $\rho_{ij}^{mf}$ is defined in equation (9); GICS is a dummy variable indicating that the two stocks are in the same Global Industry Classification Standard industry; DLNSIZE is the absolute value difference in the natural logarithms of the market values of equity; DLNBTM is the absolute value difference in the natural logarithms of the book-to-market equity ratios; DPRICE is the absolute value difference in lagged stock prices; DLEVERAGE is the absolute value difference in leverage ratios; DROE is the absolute value difference in returns on equity; and STATE is a dummy variable indicating the headquarters of the two firms are in the same state. Year fixed effects are included but not reported. The correlation coefficients are estimated yearly using daily raw returns (RAW), or the daily residuals from the capital asset pricing model (CAPM) and the Fama–French (1993) 3-factor model augmented with the Carhart (1997) momentum factor (FF4) for each pair of S&P 500 stocks that is in the index during the calendar and following years. Factor loadings for the asset pricing model are estimated using data from the current year. Returns data are from the Center for Research in Security Prices and factor data from Wharton Research Data Services. Analyst forecast data are from the Institutional Broker's Estimate System. All numbers are expressed in percentage terms except for number of observations and adjusted $R^2$. Standard errors are clustered at the firm and year levels and are listed in parentheses. There are 838,562 observations used from 1995–2007.

| Variable | RAW | CAPM | FF4 | RAW | CAPM | FF4 | RAW | CAPM | FF4 |
|---|---|---|---|---|---|---|---|---|---|
| $\rho_{ij}^{an}$ | 18.40 | 25.80 | 22.90 | | | | | | |
| | (3.78) | (4.17) | (3.69) | | | | | | |
| $\rho_{ij}^{br}$ | | | | 3.31 | 6.03 | 7.85 | | | |
| | | | | (1.16) | (1.38) | (1.50) | | | |
| $\rho_{ij}^{ca}$ | | | | | | | 1.58 | −0.02 | 1.66 |
| | | | | | | | (0.79) | (0.46) | (0.78) |
| GICS | 10.00 | 12.40 | 10.80 | 24.30 | 36.80 | 37.70 | 25.10 | 36.60 | 38.40 |
| | (1.77) | (1.96) | (1.97) | (4.42) | (5.96) | (6.62) | (4.43) | (5.73) | (6.42) |
| DLNSIZE | 0.82 | 0.90 | 1.10 | 13.50 | 17.70 | 15.50 | 13.90 | 18.70 | 16.70 |
| | (0.28) | (0.49) | (0.57) | (1.10) | (1.55) | (1.70) | (1.08) | (1.56) | (1.62) |
| DLNBTM | −1.07 | −1.06 | −0.77 | 1.08 | 1.39 | 1.57 | 1.08 | 1.21 | 1.44 |
| | (0.24) | (0.33) | (0.42) | (0.36) | (0.44) | (0.52) | (0.37) | (0.39) | (0.49) |
| DPRICE | −2.66 | −2.40 | −3.44 | −1.06 | −0.81 | −0.64 | −1.04 | −0.74 | −0.58 |
| | (0.35) | (0.62) | (0.64) | (0.25) | (0.26) | (0.35) | (0.25) | (0.26) | (0.36) |
| DLEVERAGE | −0.87 | −7.91 | −5.42 | −3.02 | −2.97 | −3.66 | −3.09 | −3.03 | −3.77 |
| | (1.27) | (2.24) | (2.26) | (0.30) | (0.65) | (0.65) | (0.30) | (0.66) | (0.67) |
| DROE | −1.32 | 0.84 | −0.12 | 0.29 | −7.23 | −4.02 | 0.45 | −7.24 | −3.87 |
| | (0.56) | (0.59) | (0.50) | (1.53) | (2.09) | (2.05) | (1.56) | (2.09) | (2.04) |
| STATE | 1.03 | 0.57 | 0.27 | −1.95 | 0.83 | −0.22 | −1.91 | 0.84 | −0.17 |
| | (0.28) | (0.47) | (0.54) | (0.64) | (0.61) | (0.61) | (0.63) | (0.63) | (0.63) |
| $\rho_{ij}^{mf}$ | 17.20 | 29.60 | 31.50 | 0.34 | 0.13 | −0.02 | 0.36 | 0.33 | 0.15 |
| | (2.74) | (7.47) | (7.51) | (0.28) | (0.31) | (0.37) | (0.28) | (0.31) | (0.36) |
| $N$ | 838,562 | 838,562 | 838,562 | 838,562 | 838,562 | 838,562 | 838,562 | 838,562 | 838,562 |
| Adj. $R^2$ | 0.463 | 0.371 | 0.527 | 0.457 | 0.362 | 0.523 | 0.457 | 0.362 | 0.522 |

A more stringent test is provided by the connected analyst coverage measure, $\rho_{ij}^{ca}$. The final three columns of Table 8 provide the results of the same set of regressions using this "social networking" variable as an independent variable. The results weakly support the information story. When comovement is defined as the simple pairwise correlation of stock returns, the coefficient on $\rho_{ij}^{ca}$ is 1.48% and is statistically significant at the 5% level. After controlling for market risk, using the CAPM, the relation is no longer statistically significant. However, when using the FF4 to define excess comovement, the coefficient is 1.66% and is statistically significant. These coefficients are much smaller than those for both the analyst-level measure and the broker-level measure. However, given the degree of separation between the analysts used to define this measure, the positive,

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

significant results provide additional evidence that some degree of comovement is driven by the structure of information production, and not simply a result of analysts covering similar stocks. This may be a result of direct information sharing, self-selection, selection of similar "types" by the previous firm, or similar training. The Appendix provides additional evidence along these lines.

## D. Regulation Fair Disclosure

In Aug. 2000, the U.S. Securities and Exchange Commission implemented Reg FD (see Heflin, Subramanyam, and Zhang (2003) for a review), which changed the way analysts collect and process data. Public firms were now required to disclose information to all investors at the same time. Because this regulation limited the extent of analysts' informational advantages, they were forced to rely less on private, firm-specific information and more on their own abilities to process public information. Because firm-specific private information becomes more difficult to obtain, the relative amount of firm-specific private information in forecasts drops. Thus, correlations across individual analysts' forecasts should increase. As a result, their forecasts across stocks are more likely to be correlated with macro-level information. This should lead to an increase in the effect of $\rho_{ij}^{\text{an}}$ on return correlations.

To test this proposition, I create the variable REG_FD, which indicates whether the observation takes place after the implementation of Reg FD. This variable equals 1 for 2001–2007, and 0 otherwise. Because observations from 2000 are not included in the regression, the number of observations drops slightly to 922,294. The regressions with GICS industry dummies are repeated and results are reported in Table 9. Using all four asset pricing models, there is an average

TABLE 9

Regulation Fair Disclosure: Comovement and Analyst Coverage

Table 9 reports ordinary least squares regression coefficients and *t*-statistics from regressing stock return correlations between pairs of stocks on common analyst coverage, $\rho_{ij}^{\text{an}}$, along with an industry dummy variable, GICS, and the variable, REG_FD, which is equal to 1 if the observation is after Regulation Fair Disclosure was instituted in 2000, and 0 otherwise. Additionally, REG_FD is interacted with the common analyst coverage variable. Common analyst coverage is defined as $N_{i,j}/\sqrt{N_i N_j}$, where $N_{i,j}$ is the number of common analysts for stocks $i$ and $j$, and $N_i$ and $N_j$ are the total number of analysts covering stocks $i$ and $j$, respectively. The variable GICS is an indicator variable that is equal to 1 if the pair of stocks share the same 6-digit Global Industry Classification Standard code, and 0 otherwise. The correlation coefficients are estimated yearly using daily raw returns (RAW) (column 1), or the daily residuals from the capital asset pricing model (CAPM) (column 2), the Fama–French (1993) 3-factor model (FF3) (column 3), and the Fama–French 3-factor model augmented with the Carhart (1997) momentum factor (FF4) (column 4) for each pair of S&P 500 stocks that are in the index during the calendar and previous years. Factor loadings for the asset pricing model are estimated using data from the previous year. Returns data are from the Center for Research in Security Prices, factor data from Wharton Research Data Services, and analyst data are from the Institutional Brokers' Estimate System. There are 922,294 observations used from 1995–1999 and 2001–2007. Observations from 2000 are omitted. All numbers are expressed in percentage terms except for number of observations and adjusted $R^2$. Standard errors are clustered by firm and year and are reported in parentheses.

| Variable | RAW | CAPM | FF3 | FF4 |
|---|---|---|---|---|
| $\rho_{ij}^{\text{an}}$ | 5.2 | 30.1 | 30.2 | 34.2 |
| | (8.8) | (7.6) | (8.7) | (12.2) |
| REG_FD | 10.0 | 3.7 | 2.7 | 3.2 |
| | (3.6) | (3.7) | (4.2) | (6.0) |
| $\rho_{ij}^{\text{an}} \times$ REG_FD | 10.7 | 32.9 | 37.6 | 53.2 |
| | (10.2) | (12.4) | (13.8) | (22.3) |
| GICS | 13.4 | 5.34 | 3.86 | −1.58 |
| | (3.7) | (4.0) | (4.5) | (8.4) |
| *N* | 922,294 | 922,294 | 922,294 | 922,294 |
| Adj. $R^2$ | 0.158 | 0.113 | 0.101 | 0.100 |

increase in excess comovement ranging from 3% to 10% after the adoption of Reg FD. However, except for the model with raw returns, none of these is statistically significant. Conversely, the coefficient on the interaction between this dummy variable and the common analyst coverage variable ranges from 10% to 53%, and is statistically significant when using the three asset pricing equations to determine excess comovement. Therefore, after Reg FD is implemented, common analyst coverage becomes much more important in determining excess comovement in the cross section of returns. This is consistent with individual analysts who are forced to rely relatively more on public or aggregate information, making common mistakes across stocks. As shown in the Appendix, results are similar when using a direct measure of earnings forecast errors.

## E. Trading Strategy

Because common analyst coverage is associated with comovement, I examine whether there is a profitable trading strategy based on this relation. As noted, analyst coverage should be correlated with comovement for two reasons. First, analysts tend to focus on stocks with similar exposures to risk. Second, any common error in an analyst's forecasts should be reflected in the correlations across all of the stocks covered by the analyst. Chen, Chen, and Li (2008) examine deviations in returns of sets of stocks with high past correlations. They find that these deviations are reversed in the following month. Anton and Polk (2014) form peer stocks using institutional ownership as a measure of "connectedness" of stocks and find similar results. In the same spirit as these two studies, I use analyst coverage as a proxy for comoving, or "peer," firms and test the hypothesis that stocks whose returns deviate from their peers in the current month tend to revert in the following month. Analyst coverage can be precisely measured over small windows of time. Accordingly, trading strategies based on this measure may be more profitable than those based other measures that are imprecisely estimated over longer periods. As a benchmark, I also evaluate returns using a similar strategy based on GICS industry codes, which, as previously noted, are a good measure of industries from the standpoint of analysts.

It is well documented that the returns of small firms tend to lag returns of larger firms, but not vice versa (see Lo and MacKinlay (1990), Brennan, Jegadeesh, and Swaminathan (1993), and Chordia and Swaminathan (2000)). To avoid capturing this intraindustry lead–lag effect based on size, I restrict my trading portfolio to stocks in the largest size decile. At the end of each month from Jan. 1994 through Nov. 2010, I calculate $\rho_{ij}^{\text{an}}$ for every pair of stocks listed on the New York Stock Exchange/American Stock Exchange/National Association of Securities Dealers Automated Quotations using all analysts who made earnings forecasts during the previous 12 months. Next, for each stock $i$, I sort on $\rho_{ij}^{\text{an}}$, choose the top 30 similar stocks based on this measure, and denote them as "peer" stocks.[17,18] I form an equal-weighted portfolio of peers and calculate stock returns in the current month for both the stock and its peer portfolio. Then, I

[17] The sample begins in 1994 because that is the earliest date for which GICS codes are available. Using analyst coverage to define peers, we can go back as far as 1984. Results using that measure are similar over both periods.

[18] Stocks with only a few analysts may have fewer than 30 peers.

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

independently sort each stock into quintiles based on i) its own return that month and ii) the return of the peer portfolio that month. This results in 25 portfolios. Finally, I calculate the equal-weighted and value-weighted stock returns during the subsequent month within each of the 25 portfolios. Panel A of Table 10 provides these averages as well as the standard errors using value-weighted portfolios of only the largest 10% of stocks based on the market capitalization in the previous month.

Controlling for the stocks' own past returns, subsequent returns generally increase as peers' past returns increase. The final two columns of each panel in

TABLE 10
Trading Strategy

Table 10 reports value-weighted 1-month portfolio percentage returns for 25 portfolios of the largest 10% of stocks sorted into quintiles based on their own 1-month lagged return and on the equal-weighted 1-month lagged return of their peers. A stock's peers are either 30 stocks with the most similar analyst coverage based on $\rho_{ij}^{an}$ (Panel A) or the set of stocks in the same Global Industry Classification Standard (GICS) industry (Panel B). Additionally, for each quintile, returns and alphas are reported for the strategy that is long one extreme, the highest peer return (HP) or lowest lagged return (L), and short the other corresponding extreme portfolio. Finally, returns and alphas are presented for the strategy that is long the "low lagged return/high peer lagged return" portfolio (L–HP) and short the "high lagged return/low peer lagged return" portfolio (H–LP). Alphas are calculated using the Fama–French (1993) 3-factor model augmented with the Carhart (1997) momentum factor, the Pastor and Stambaugh (2003) liquidity factor and the short-term reversal factor. Standard errors are in parentheses. Data are from 1994–2010.

*Panel A. Peers Based on Common Analyst Coverage*

| | Peer Lagged Return Quintile | | | | | HP – LP | |
|---|---|---|---|---|---|---|---|
| Lagged Return Quintile | LP | 2 | 3 | 4 | HP | Return | Alpha |
| L | 0.62 | 0.89 | 1.08 | 1.21 | 1.42 | 0.79 | 0.69 |
| | (0.60) | (0.53) | (0.45) | (0.38) | (0.44) | (0.52) | (0.54) |
| 2 | 0.86 | 1.18 | 1.00 | 1.52 | 1.22 | 0.35 | 0.33 |
| | (0.47) | (0.37) | (0.34) | (0.36) | (0.40) | (0.40) | (0.42) |
| 3 | 0.65 | 0.95 | 0.95 | 0.77 | 1.24 | 0.59 | 0.50 |
| | (0.39) | (0.35) | (0.33) | (0.34) | (0.40) | (0.42) | (0.44) |
| 4 | 0.72 | 0.39 | 0.66 | 0.70 | 0.46 | −0.26 | −0.34 |
| | (0.37) | (0.34) | (0.34) | (0.33) | (0.44) | (0.40) | (0.42) |
| H | 0.13 | 0.87 | 0.33 | 0.25 | 0.53 | 0.40 | 0.15 |
| | (0.44) | (0.36) | (0.41) | (0.47) | (0.51) | (0.46) | (0.48) |
| *L–H* | | | | | | | |
| Return | 0.50 | 0.02 | 0.75 | 0.95 | 0.89 | (L–HP) – (H–LP) | |
| | (0.51) | (0.46) | (0.45) | (0.46) | (0.47) | Return | Alpha |
| Alpha | 0.37 | 0.07 | 0.61 | 0.89 | 0.91 | 1.29 | 1.06 |
| | (0.53) | (0.49) | (0.47) | (0.48) | (0.48) | (0.38) | (0.40) |

*Panel B. Peers Based on GICS Code*

| | Peer Lagged Return Quintile | | | | | HP – LP | |
|---|---|---|---|---|---|---|---|
| Lagged Return Quintile | LP | 2 | 3 | 4 | HP | Return | Alpha |
| L | 0.65 | 0.96 | 1.24 | 1.75 | 0.72 | 0.07 | 0.14 |
| | (0.57) | (0.54) | (0.47) | (0.41) | (0.48) | (0.48) | (0.50) |
| 2 | 0.85 | 1.18 | 1.03 | 1.24 | 1.23 | 0.38 | 0.28 |
| | (0.44) | (0.41) | (0.39) | (0.39) | (0.37) | (0.40) | (0.42) |
| 3 | 0.60 | 0.88 | 0.91 | 1.11 | 1.20 | 0.60 | 0.58 |
| | (0.38) | (0.33) | (0.34) | (0.35) | (0.41) | (0.39) | (0.41) |
| 4 | 0.50 | 0.54 | 0.72 | 0.83 | 0.55 | 0.05 | 0.03 |
| | (0.37) | (0.37) | (0.36) | (0.38) | (0.41) | (0.42) | (0.44) |
| H | 0.36 | 0.81 | 0.52 | 0.28 | 0.37 | 0.01 | −0.20 |
| | (0.43) | (0.38) | (0.47) | (0.45) | (0.55) | (0.52) | (0.54) |
| *L–H* | | | | | | | |
| Return | 0.29 | 0.15 | 0.72 | 1.47 | 0.35 | (L–HP) – (H–LP) | |
| | (0.47) | (0.48) | (0.57) | (0.46) | (0.54) | Return | Alpha |
| Alpha | 0.01 | 0.29 | 0.72 | 1.40 | 0.35 | 0.36 | 0.15 |
| | (0.48) | (0.50) | (0.60) | (0.48) | (0.55) | (0.40) | (0.42) |

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

Table 10 present the average raw returns and alphas for the strategy that is long the high peer (HP) lagged return portfolio and short the low peer (LP) lagged return portfolio for each of the four lagged return quintiles. Alphas are calculated using the FF4, the Pastor–Stambaugh (2003) liquidity factor, and the short-term reversal factor.[19] Regardless of stocks' own past returns, future returns are generally higher when peer lagged returns are high than when peer lagged returns are low; aside from the lagged peer quintile 4, differences range between 15 and 79 basis points (bps). However, these differences are not statistically significant. Of course, because the strategy focuses only on the largest size decile, the number of stocks in each portfolio is relatively small, about 20, on average.

Controlling for peer past returns, future stock returns tend to decrease as past returns increase. The final two rows of each panel provide average returns and alphas for the strategy that is long the low (L) lagged return portfolio and short the high (H) lagged return portfolio for each of the four lagged peer return quintiles. Estimates range from 2 bps to 95 bps per month, though these differences are statistically significant only in the highest peer lagged return quintiles.

If analyst coverage is associated with comovement, a stock whose returns are extremely low (L) compared to its peers (HP) or extremely high (H) compared to its peers (LP) might represent a temporary divergence. If these deviations revert within a month, a trading strategy that is long the L–HP portfolio and short the H–LP portfolio should be profitable. Large stocks in the L–HP portfolio have returns of about 1.42% per month compared to only 0.13% for stocks in the H–LP portfolio. The long/short portfolio returns and alphas associated with this strategy are provided in the lower right quadrant of Panel A of Table 10. The strategy generates an average value-weighted raw return of 1.29% per month and an alpha of 1.06%. Both are statistically significant at the 1% level. These numbers are smaller than those generated using past correlations to define peers (Chen et al. (2008)), and using institutional ownership to define peers (Anton and Polk (2014)). However, the samples in these studies include firms of all sizes and are likely to capture the size-based lead–lag effects first documented by Lo and MacKinlay (1990); the studies do not include the short-term reversal factor in their models.

To test whether this trading strategy is simply capturing lead–lag relations within an industry or industry momentum, I repeat the exercise using GICS codes to define peers. Panel B of Table 10 presents the resulting portfolio returns. The same general patterns hold as when peers are defined using analyst coverage. However, the magnitudes tend to be smaller, especially across the peer lagged return quintiles. Furthermore, the spread between the L/HP and H/LP is only 36 bps in raw terms and 15 bps after controlling for risk, suggesting that information transmission based on analyst coverage that is orthogonal to industry classifications is the culprit, particularly in the largest stocks. Taken together, these results are consistent with the notion that common analyst coverage leads to excess comovement. Stocks that deviate from these patterns quickly revert.

[19] All factors are courtesy of Kenneth French's Web site (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html).

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

## V. Conclusion

This article is the first to show that correlated earnings forecast errors lead to excess comovement. When errors in news about fundamentals are correlated across stocks, stock returns tend to comove more than predicted by standard asset pricing models. This article develops a new, easy-to-calculate measure of correlated information based on analyst coverage. A 1% increase in this measure leads to a 0.41% to 0.46% increase in excess comovement before controlling for industry, and a 0.25% to 0.28% increase after controlling for industry. This measure is strongest when accounting for analyst precision, as measured using a subset of star analysts or weighting by the accuracy of past earnings forecasts, and does not seem to be simply a proxy for risk exposure, style investors, or correlations in unexpected earnings or earnings revisions. This effect is significant around exogenous changes in analyst coverage as a result of brokerage firm mergers or additions to the S&P 500 index, where there is a change in the information structure but not necessarily in the risk exposure of the stock. The effect increases after the implementation of Reg FD. It is weaker but still present when measured using brokerage firms instead of analysts, and when using a social networking measure based on past working relationships between analysts. This suggests that in addition to adjusting for risk, the structure of information dissemination should be accounted for when defining excess comovement.

## Appendix. Correlated Forecast Errors and Comovement

In this Appendix, I repeat much of the analysis from the body of the article using a more direct estimate of commonality, the correlation in earnings forecast errors across stocks, and show that the results are generally consistent with those found using the proxy. Using a subset of data for which correlations in earnings forecast errors are likely to be more precisely estimated, I find that i) correlations decrease when analyst coverage drops following mergers of brokerage firms and increase when analyst coverage increases following additions to the S&P 500 index; ii) these changes can explain changes in excess comovement in the same pairs of firms; iii) analysts who work together, and analysts who used to work together, at the same brokerage firm tend to have correlated earnings forecast errors; and iv) following the implementation of Reg FD, correlated forecast errors have become a more important determinant of excess comovement. This is an extension of the validation section.

As mentioned, there are several drawbacks to estimating correlations in earnings forecast errors relative to using the common analyst coverage proxy. First, calculating the measure requires two time series of earnings forecasts and actuals, and involves estimation error, and correlations can be estimated only infrequently. Moreover, because fiscal periods, earnings forecast dates, and earnings announcement dates are typically staggered, there is a potential bias in the estimate. To alleviate these concerns, I examine a subset of stocks for which earnings announcements are made on the same day and fiscal periods end on the same day. Additionally, I examine pairs of analysts making forecasts on the same days for the same stocks. Although this greatly reduces the number of observations used in the analysis, it provides a validation that the proxy based only on analyst coverage does indeed capture some measure of common information and is likely to be a valid measure for a broad set of stocks.

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

Table A1 examines how changes in correlated forecast errors following analyst turnover as a result of mergers of brokerage firms affect excess comovement. This is analogous to the tests presented in Table 4. There are 1,187 pairs of stocks that satisfy the same conditions as those in Table 4 (i.e., i) they are covered by an analyst from both merging firms before the merger, ii) they are covered by a single analyst from the merged firm after the merger, and iii) analyst coverage from nonmerging firms does not change), as well as the restrictions that the fiscal periods for both firms end on the same dates. Then, using the monthly consensus forecast errors, I define changes in correlations, $\Delta\rho_{ij}^{\text{error}}$, as the difference in the correlation of consensus forecast error from before and after the merger for pairs of firms with at least 8 observations. Panel A shows that, on average, these correlations decrease by 2.5% following the reduction in analyst coverage around mergers. On average, excess comovement does not change significantly, though there is a large amount of variation within the sample.

In Panel B of Table A1, changes in comovement around the merger are regressed on changes in forecast error correlations. For all four measures of comovement, the coefficient on $\Delta\rho_{ij}^{\text{error}}$ is positive and, with the exception of the CAPM, statistically significant. Thus, the pairs of stocks whose forecast errors become less correlated as a result of the merger see a reduction in comovement. These results are consistent with Table 4, which suggests that the common analyst coverage measure is a good proxy for correlated forecast errors.

Next, I repeat the above analysis but instead examine changes in forecast error correlations around additions to the S&P 500 index. This is analogous to Table 6. Again, I require firms to have the same fiscal periods and estimate changes in correlations in forecast errors between pairs of stocks following the addition. Because many of these firms are young and have little or no analyst coverage before the addition to the index, only 359 pairs are in the sample. This is one disadvantage to using a correlation as opposed to the coverage-based measure.

TABLE A1

Correlated Forecast Errors: Brokerage House Mergers and Comovement

Table A1 reports summary statistics (Panel A) and ordinary least squares regression coefficients and standard errors (Panel B) from regressing the change in stock return correlations between stocks on changes in consensus forecast error correlations, $\Delta\rho_{ij}^{\text{error}}$, before and after 15 brokerage house mergers between 1994 and 2005 (see Hong and Kacperczyk (2010), Table 3 for a list). Firm fixed effects are included but not reported. Correlations in consensus forecast errors are calculated using quarterly forecasts in the Institutional Brokers' Estimate System for pairs of stocks with at least 8 months of data. The return correlation coefficients are estimated using daily raw returns (RAW) or residuals from the capital asset pricing model (CAPM), the Fama–French (1993) 3-factor model (FF3), and a model using the Fama–French factors plus the Carhart (1997) momentum factor (FF4) for 1 year before and after a merger. Returns data are from the Center for Research in Security Prices and factor data from Wharton Research Data Services. Stocks are included only if i) they are covered by an analyst from both merging firms before the merger, ii) they are covered by a single analyst from the merged firm after the merger, and iii) analyst coverage from nonmerging firms does not change. All numbers are expressed in percentage terms except for number of observations and adjusted $R^2$. Standard errors are in parentheses.

*Panel A. Summary Statistics*

| Variable | Change in: | *N* | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|---|
| $\Delta\rho_{ij}^{\text{error}}$ | Forecast error correlations | 1,187 | −2.4 | 39.8 | −134.9 | 147.5 |
| $\Delta\rho_{ij}^{\text{ret}}$ | Raw return correlations | 1,187 | 2.1 | 20.3 | −121.9 | 116.7 |
| $\Delta\rho_{ij}^{\text{capm}}$ | CAPM alpha correlations | 1,187 | 0.4 | 23.3 | −155.7 | 104.3 |
| $\Delta\rho_{ij}^{\text{ff3}}$ | FF3 alpha correlations | 1,187 | 0.4 | 26.8 | −192.7 | 172.3 |
| $\Delta\rho_{ij}^{\text{ff4}}$ | FF4 alpha correlations | 1,187 | 0.6 | 27.8 | −177.0 | 172.3 |

*Panel B. Regressions*

| Variable | RAW | CAPM | FF3 | FF4 |
|---|---|---|---|---|
| $\Delta\rho_{ij}^{\text{error}}$ | 4.93 | 1.51 | 6.21 | 6.99 |
| | (1.47) | (1.70) | (1.95) | (2.03) |
| No. of obs. | 1,187 | 1,187 | 1,187 | 1,187 |
| Adj. $R^2$ | 0.214 | 0.164 | 0.169 | 0.192 |

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

Panel A of Table A2 shows results consistent with Table 6. Not only does coverage increase, but the correlation in forecast errors also increases. Panel B presents regressions of changes in comovement on changes in correlated forecast errors. The coefficients in all four specifications are positive and are statistically significant in the three measures of excess comovement. Thus, changes in correlations in consensus forecast errors around additions of stocks to the S&P 500 index lead to changes in excess comovement.

TABLE A2

Correlated Forecast Errors: Additions to the S&P 500 Index and Comovement

Table A2 reports summary statistics (Panel A) as well as ordinary least squares regression coefficients and standard errors (Panel B) from regressing the change in stock return correlations between stocks that are added to the S&P 500 index and all other stocks in the index on changes in consensus forecast error correlations, $\Delta\rho_{ij}^{error}$, before and after the addition date (Panel B). Firm fixed effects are included but not reported. Correlations in consensus forecast errors are calculated using quarterly forecasts in the Institutional Brokers' Estimate System for pairs of stocks with at least 8 months of data. The return correlation coefficients are estimated using daily raw returns (RAW) or residuals from the capital asset pricing model (CAPM), the Fama–French (1993) 3-factor model (FF3), and a model using the Fama–French factors plus the Carhart (1997) momentum factor (FF4) for 1 year before and after the addition. Returns data are from the Center for Research in Security Prices and factor data from Wharton Research Data Services. There are 359 pairwise observations used. All numbers are expressed in percentage terms except for number of observations and adjusted $R^2$. Standard errors are in parentheses.

*Panel A. Summary Statistics*

| Variable | Change in: | *N* | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|---|
| $\Delta\rho_{ij}^{error}$ | Forecast error correlations | 359 | 1.26 | 60.37 | −144.06 | 145.42 |
| $\Delta\rho_{ij}^{ret}$ | Raw return correlations | 359 | 7.17 | 38.98 | −90.03 | 165.06 |
| $\Delta\rho_{ij}^{capm}$ | CAPM alpha correlations | 359 | 1.56 | 40.92 | −106.33 | 130.13 |
| $\Delta\rho_{ij}^{ff3}$ | FF3 alpha correlations | 359 | 1.71 | 42.06 | −104.87 | 140.38 |
| $\Delta\rho_{ij}^{ff4}$ | FF4 alpha correlations | 359 | −0.28 | 42.09 | −129.16 | 128.06 |

*Panel B. Regressions*

| Variable | RAW | CAPM | FF3 | FF4 |
|---|---|---|---|---|
| $\Delta\rho_{ij}^{error}$ | 4.78 | 12.03 | 9.26 | 7.90 |
| | (3.39) | (3.52) | (3.64) | (3.65) |
| No. of obs. | 359 | 359 | 359 | 359 |
| Adj. $R^2$ | 0.565 | 0.464 | 0.483 | 0.408 |

Next, I validate the alternate proxies by examining whether analysts who work at the same broker and analysts who used to work together at the same broker have correlated forecast errors. For each pair of analysts at the same brokerage firm, I keep forecasts/revisions on the same date for firms with the same fiscal periods and subsequent earnings announcements on the same dates. Then, I calculate forecast errors and estimate correlation coefficients. Panel A of Table A3 shows mean correlations and standard errors for pairs of analysts at the same broker with at least 5, 10, 15, 20, 25, and 30 revisions. As was the case when using errors for the same analyst (see Table 2), as the sample is restricted to pairs of analysts with longer time series, the mean correlation increases: from about 6% using the 337,170 pairs of analysts with at least 5 revisions to 27% using the 63 pairs of analysts with at least 30 revisions. Panel B repeats this analysis but restricts the pairs of analysts to those who used to work together at the same brokerage firm, but no longer do. Here, there are far fewer pairs of analysts with forecasts/revisions and announcements that lie on the same date. Only 202 pairs have multiple observations. Though the mean correlation is positive, it is not statistically significant. However, when all observations are pooled together (including pairs of analysts with only one set of forecast errors) the estimated correlation is 12% and is highly significant.

Finally, I investigate whether the impact of correlated forecast errors on comovement changes after the implementation of Reg FD in 2000. The regression is identical to that

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

TABLE A3

Correlated Forecast Errors and Social Networks

Table A3 reports the average correlation of annual earnings forecast errors across stocks by pairs of analysts from the same brokerage firm (Panel A) and across pairs of analysts who worked together in the past at the same brokerage firm but no longer do (Panel B). Correlations are estimated for each pair of stocks and analysts for which i) fiscal year-ends coincide, ii) revisions are made on the same date for both stocks, and iii) earnings announcements are on the same date. Analyst-stock-pair correlations are included in the mean if there are at least as many observations as the number reported in the first column of each panel. The total number of correlations included in each sample as well as the average correlation and standard error of the mean are reported. Additionally, in Panel B, all 2,936 observations are used to estimate a pooled correlation and a *p*-value. Earnings forecasts data and analyst coverage are from the Institutional Brokers' Estimate System and cover 1984–2012.

*Panel A. Pairs of Analysts at Same Broker*

| Min. No. of Revisions Used to Estimate Correlations | No. of Corr. in Mean | Avg. Corr. | Std. Err. |
|---|---|---|---|
| 5 | 337,170 | 0.059 | 0.001 |
| 10 | 39,322 | 0.089 | 0.002 |
| 15 | 4,822 | 0.159 | 0.005 |
| 20 | 1,033 | 0.198 | 0.010 |
| 25 | 269 | 0.231 | 0.020 |
| 30 | 63 | 0.271 | 0.040 |

*Panel B. Pairs of Connected Analysts*

| Min. No. of Revisions Used to Estimate Correlations | No. of Corr. in Mean | Avg. Corr. | Std. Err. |
|---|---|---|---|
| 2 | 202 | 0.034 | 0.067 |
| 3 | 39 | 0.100 | 0.106 |
| 4 | 5 | −0.045 | 0.305 |
| | No. in Corr. | Corr. | *p*-Value |
| (Pooled) | 2,936 | 0.118 | <0.0001 |

from Table 9, replacing the variable $\rho_{ij}^{\text{an}}$ with $\rho_{ij}^{\text{error}}$, the correlation in consensus forecast errors. Table A4 provides results from these regressions. The dummy variable, REG_FD, is equal to 1 if the observation is after Reg FD was instituted, and 0 otherwise. As can be seen by examining the interaction between REG_FD and $\rho_{ij}^{\text{error}}$, the impact of correlated forecast errors on comovement increases following the introduction of Reg FD. Although the sample is much smaller than was used in Table 9, the results are consistent.

TABLE A4

Correlated Forecast Errors: Regulation Fair Disclosure and Comovement

Table A4 reports ordinary least squares regression coefficients and *t*-statistics from regressing stock return correlations between pairs of stocks on consensus forecast error correlations, $\rho_{ij}^{\text{error}}$, along with the variable, REG_FD, which is equal to 1 if the observation is after Regulation Fair Disclosure was instituted in 2000, and 0 otherwise. Additionally, REG_FD is interacted with the consensus forecast error correlation. Correlations in consensus forecast errors are calculated using quarterly forecasts in the Institutional Brokers' Estimate System for pairs of stocks with at least 8 months of data. GICS is an indicator variable that is equal to 1 if the pair of stocks share the same 6-digit Global Industry Classification Standard code, and 0 otherwise. The correlation coefficients are estimated yearly using daily raw returns (RAW) (column 1) or the daily residuals from the capital asset pricing model (CAPM) (column 2), the Fama–French (1993) 3-factor model (FF3) (column 3), and the Fama–French 3-factor model augmented with the Carhart (1997) momentum factor (FF4) (column 4) for each pair of S&P 500 stocks that are in the index during the calendar and previous years. Factor loadings for the asset pricing model are estimated using data from the previous year. Returns data are from the Center for Research Security Prices, factor data are from Wharton Research Data Services, and analyst data are from the Institutional Brokers' Estimate System. There are 22,927 observations used from 1995 through 1999 and 2001 through 2007. Observations from 2000 are omitted. All numbers are expressed in percentage terms except for number of observations and adjusted $R^2$. Standard errors are clustered by firm and year and are reported in parentheses.

| Variable | RAW | CAPM | FF3 | FF4 |
|---|---|---|---|---|
| $\rho_{ij}^{\text{error}}$ | 1.48 | 1.13 | 1.58 | 1.83 |
| | (0.76) | (0.80) | (0.84) | (0.88) |
| REG_FD | 4.03 | 0.79 | −0.12 | −0.05 |
| | (0.61) | (0.35) | (0.32) | (0.31) |
| $\rho_{ij}^{\text{error}} \times$ REG_FD | 5.80 | 4.15 | 3.62 | 2.56 |
| | (1.07) | (1.10) | (1.04) | (1.10) |
| GICS | 16.99 | 24.79 | 21.56 | 21.21 |
| | (0.79) | (0.89) | (0.88) | (0.86) |
| Adj. $R^2$ | 0.058 | 0.089 | 0.070 | 0.067 |
| *N* | 22,927 | 22,927 | 22,927 | 22,927 |

https://doi.org/10.1017/S002210901600051X Published online by Cambridge University Press

## References


Adams, R., and D. Ferreira. "A Theory of Friendly Boards." *Journal of Finance*, 62 (2007), 217–250.

Anton, M., and C. Polk. "Connected Stocks." *Journal of Finance*, 69 (2014), 1099–1127.

Barberis, N., and A. Shleifer. "Style Investing." *Journal of Financial Economics*, 68 (2003), 161–199.

Barberis, N.; A. Shleifer; and J. Wurgler. "Comovement." *Journal of Financial Economics*, 75 (2005), 283–317.

Boni, L., and K. Womack. "Analysts, Industries, and Price Momentum." *Journal of Financial and Quantitative Analysis*, 41 (2006), 85–109.

Boyer, B. "Style Related Comovement: Fundamentals or Labels?" *Journal of Finance*, 66 (2011), 307–332.

Boyer, B., and L. Zheng. "Investor Flows and Stock Market Returns." *Journal of Empirical Finance*, 16 (2009), 87–100.

Brennan, M.; N. Jegadeesh; and B. Swaminathan. "Investment Analysis and the Adjustment of Stock Prices to Common Information." *Review of Financial Studies*, 6 (1993), 799–824.

Carhart, M. "On Persistence in Mutual Fund Performance." *Journal of Finance*, 52 (1997), 57–82.

Chen, H.; S. Chen; and F. Li. "Firm-Level Return Comovement." Working Paper, University of British Columbia (2008).

Chordia, T., and B. Swaminathan. "Trading Volume and Cross-Autocorrelations in Stock Returns." *Journal of Finance*, 55 (2000), 913–935.

Cohen, L.; A. Frazzini; and C. Malloy. "Sell-Side School Ties." *Journal of Finance*, 65 (2010), 1409–1437.

Coval, J., and E. Stafford. "Asset Fire Sales (and Purchases) in Equity Markets." *Journal of Financial Economics*, 86 (2007), 479–512.

Fama, E., and K. French. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics*, 33 (1993), 3–56.

Fracassi, C. "Corporate Finance Policies and Social Networks." *Management Science*, forthcoming (2016).

Green, C., and B. Hwang. "Price-Based Return Comovement." *Journal of Financial Economics*, 93 (2009), 37–50.

He, J., and X. Tian. "The Dark Side of Analyst Coverage: The Case of Innovation." *Journal of Financial Economics*, 109 (2013), 856–878.

Hedge, S., and J. McDermott. "The Liquidity Effects of Revisions to the S&P Index: An Empirical Analysis." *Journal of Financial Markets*, 6 (2003), 413–459.

Heflin, F.; K. Subramanyam; and Y. Zhang. "Regulation FD and the Financial Information Environment: Early Evidence." *Accounting Review*, 78 (2003), 1–37.

Hong, H., and M. Kacperczyk. "Competition and Bias." *Quarterly Journal of Economics*, 25 (2010), 1683–1725.

Ingersoll, J. *Theory of Financial Decision Making,* Vol. 3. Lanham, MD: Rowman & Littlefield (1987).

Jaccard, P. "Étude Comparative de la Distribution Florale dans une Portion des Alpes et des Jura." *Bulletin del la Société Vaudoise des Sciences Naturelles*, 37 (1901), 547–579.

Kallberg, J., and P. Pasquariello. "Time-Series and Cross-Sectional Excess Comovement in Stock Indexes." *Journal of Empirical Finance*, 15 (2008), 481–502.

Karolyi, A., and R. Stulz. "Why Do Markets Move Together? An Investigation of U.S. Japan Stock Return Co-Movements." *Journal of Finance*, 51 (1996), 951–986.

Kumar, A., and C. Lee. "Retail Investor Sentiment and Return Comovements." *Journal of Finance*, 61 (2006), 2451–2486.

Lo, A., and C. MacKinlay. "When Are Contrarian Profits Due to Stock Market Overreaction?" *Review of Financial Studies*, 3 (1990), 175–205.

Pastor, L., and R. Stambaugh. "Liquidity Risk and Expected Stock Returns." *Journal of Political Economy*, 111 (2003), 642–685.

Pirinsky, A., and Q. Wang. "Institutional Investors and the Comovement of Equity Prices." Working Paper, Texas A&M University (2004).

Sørensen, T. "A Method of Establishing Groups of Equal Amplitude in Plant Sociology Based on Similarity of Species and Its Application to Analyses of the Vegetation on Danish Commons." *Biologiske Skrifter/Kongelige Danske Videnskabernes Selskab*, 5 (1948), 1–34.

Subrahmanyam, A. "Social Networks and Corporate Governance." *European Financial Management*, 14 (2008), 633–662.

Thompson, S. "Simple Formulas for Standard Errors that Cluster by Both Firm and Time." *Journal of Financial Economics*, 99 (2011), 1–10.

Wu, J., and A. Zang. "What Determine Financial Analysts' Career Outcomes during Mergers?" *Journal of Accounting and Economics*, 47 (2009), 59–86.