Economic linkages inferred from news stories and the predictability of stock returns

Anna Scherbina
University of California, Davis

Bernd Schlusche
Federal Reserve Board

AEI Economics Working Paper 2016-05
February 2016

Economic Linkages Inferred from News Stories and the Predictability of Stock Returns

Anna Scherbina* Bernd Schlusche†

UC Davis *Federal Reserve Board*

First draft: December 3, 2013

This draft: December 28, 2015

ABSTRACT

We show that news stories contain information about economic linkages between firms and document that information diffuses slowly across linked stocks. Specifically, we identify linked stocks from co-mentions in news stories and find that linked stocks cross-predict one another's returns in the future. Our results indicate that information can flow from smaller to larger stocks and across industries. Content analysis of common news stories reveals many types of firm linkages that have not been previously studied. We find that the cross-predictability in returns remains even after firm pairs with customer-supplier ties are removed. Results show that both limited attention and slow processing of complex information contribute to slow information diffusion.

JEL classification: G10, G12, G14, G17

Keywords: News Media, Soft Information, Linked Stocks, Information Leadership, Lead-Lag Effect, Complex Information, Limited Attention, Market Efficiency

*Address: Graduate School of Management, University of California, Davis, One Shields Avenue, Davis, CA 95616. E-mail: ascherbina@ucdavis.edu. Phone: (530) 754-8076.

†Address: Board of Governors of the Federal Reserve System, 20th Street and Constitution Avenue NW, Washington, DC 20551. E-mail: bernd.schlusche@frb.gov. Phone: (202) 452-2591.

We are grateful to seminar participants at Pennsylvania State University, Tulane University, Michigan State University, Temple University, Johns Hopkins University, the Federal Reserve Board, Georgetown University, George Mason University, the University of Delaware, American University, Acadian Asset Management, BlackRock, Barclays, State Street Global Advisors, the Citi Quant Conference, the Financial Research Association early ideas session, the Western Finance Association Annual Meetings, the 7th Annual Florida State University SunTrust Beach Conference, the European Finance Association Meetings, and the 12th German Economists Abroad Conference, as well as our discussants, Oguzhan Ozbas (WFA), Umit Gurun (SunTrust Conference), and Jordi Mondria (EFA), for helpful comments. We are grateful to Thompson Reuters for making the TRNA dataset available for our research. The analysis and conclusions set forth are those of the authors alone and do not indicate concurrence by the Board of Governors or other members of the research staff.

Economic Linkages Inferred from News Stories and the Predictability of Stock Returns

ABSTRACT

We show that news stories contain information about economic linkages between firms and document that information diffuses slowly across linked stocks. Specifically, we identify linked stocks from co-mentions in news stories and find that linked stocks cross-predict one another's returns in the future. Our results indicate that information can flow from smaller to larger stocks and across industries. Content analysis of common news stories reveals many types of firm linkages that have not been previously studied. We find that the cross-predictability in returns remains even after firm pairs with customer-supplier ties are removed. Results show that both limited attention and slow processing of complex information contribute to slow information diffusion.

JEL classification: G10, G12, G14, G17

Keywords: News Media, Soft Information, Linked Stocks, Information Leadership, Lead-Lag Effect, Complex Information, Limited Attention, Market Efficiency

I. Introduction

A corporate news event often affects not only the firm at the center of that event but also a number of other companies in similar circumstances. On April 17, 2013, the Supreme Court handed a 9-0 victory to Royal Dutch Petroleum (RDP) in the Kiobel v. Royal Dutch Petroleum case. RDP was sued by 12 Nigerian citizens who claimed that the firm cooperated with the Nigerian government in the 1990s to brutally crush the resistance to oil development in the country. The case attracted considerable attention, in part, because it was deciding whether foreign citizens may seek compensation in U.S. courts for human-rights violations committed by firms outside the United States. About 150 such lawsuits have been filed in the United States in the past three decades, and these lawsuits were costly to settle as well as damaging to firms' reputations.[1] On September 30, 2012, Reuters reported on the firms that had filed amicus briefs in support of RDP and speculated that "a ruling against Kiobel could wipe out lawsuits pending against companies such as Exxon Mobil Corp, Rio Tinto Plc and Nestle, which are accused by private plaintiffs of helping governments violate human rights in Indonesia, Papua New Guinea and Ivory Coast, respectively."[2] Other business publications concurred that in the case of an RDP victory, virtually all similar lawsuits would be dismissed.[3]

Unlike the firm at the center of an unfolding event, such as RDP in the example above, other firms also affected by the news event are less visible. Media coverage, however, may reveal the companies that are tied to the firm at the center of the news development because journalists following a particular set of firms and/or events devote themselves fully to investigating the relevant news and to providing a broad analysis. For example, media

[1] *The Economist*, "The Shell Game Ends," April 20, 2013.

[2] Indeed, on August 8, 2012 (and on February 3, 2013, due to the Court's decision to hear additional arguments) a number of firms filed amicus briefs in support of RDP. These included Caterpillar, ConocoPhillips, General Electric, Honeywell International, IBM, Monsanto Company (these firms filed a joint brief), Coca-Cola, KBR, Engility Corporation, Chevron, and Rio Tinto Plc. A subset of these firms wrote in their amicus briefs that they had similar ongoing lawsuits.

[3] See, e.g., Forbes.com, "Supreme Court Observations: Kiobel v. Royal Dutch Petroleum & the Future of Alien Tort Litigation," April 18, 2013.

coverage of the RDP lawsuit clearly helped identify firms with similar legal issues that were also affected by the news, even though they were not at the center of the unfolding event. The introduction of Regulation Fair Disclosure (Reg FD) likely further facilitated the news media's ability to identify firm linkages. In accordance with Reg FD, firms are required to report significant developments that have the potential to affect their stock price (e.g., new business deals with other firms) to a broad audience and with a minimal delay, and press releases are a recommended way to comply. Some press releases, considered to be of interest to readers, are simply reprinted in the business press. All told, news coverage is likely to contain soft information about interlinkages and business similarities between firms that may not be easily available elsewhere.

In this paper, we show that information diffuses slowly across linked firms and that economic linkages between firms identified through co-mentions in news stories are useful in predicting stock returns.[4] If a firm experiences a news shock, its stock price will react quickly, but the linked stocks that are also affected by the news may react with a delay due to slow processing of complex information and limited investor attention. Indeed, the amount of corporate news that hits the market every day is immense. Based on a nearly complete set of corporate press releases for the period between April 2006 and August 2009, Neuhierl, Scherbina, and Schlusche (2013) document that, on an average day, a total of 281 valuation-relevant news items are released by all U.S.-based firms, which alone may explain, to a large extent, why information diffuses slowly. That study additionally shows that a large fraction of the press releases contains non-routine news whose impact on prices is hard to quantify. Of course, it is even more difficult to quantify the impact of a firm's news on its economically linked firms. We show that the attention of sophisticated investors helps ensure a quicker price reaction to the news of their linked firms. Moreover, frequent co-mentions in news stories, which make investors more aware of firm linkages, ensure quicker information

[4]Throughout the paper, we use the terms "linked stocks" and "economically linked stocks" interchangeably.

diffusion. These findings underscore that limited investor attention and slow processing of complex information both contribute to slow information diffusion across linked stocks.

Our approach to identifying stock linkages is novel in that it exploits the news media's soft information to uncover linkages between firms. We show that information available ex-ante from news coverage can be used to exploit the lead-lag relation in the returns of linked stocks. We begin by identifying, prior to each month t, all stocks $j = 1, \ldots, J^i$ that were co-mentioned with stock i in a news story in the Thomson-Reuters News Analytics (TRNA) dataset (having first removed news stories that would randomly group firms together) in the preceding three (or six or 12) months. Once stock linkages have been identified, we determine the direction of the information flow. For this purpose, we assume that information flows from stocks whose turnover is higher than their normal level since high turnover should indicate that a stock experienced a news event (we refer to these stocks as leaders) to their linked stocks with a normal level of turnover (followers). We then calculate the "linked-stock signal" as the equal-weighted average of the prior month's returns of the leaders and hypothesize that the price of followers will subsequently move in the direction of their linked-stock signal. Our results show that the linked-stock signal indeed possesses reliable predictive ability for stock returns and that this predictive ability is not subsumed by previously known return predictors at either the firm or industry levels.

Our setting allows us to provide new insights into the process of information diffusion in the stock market. Even though, due to the financial media's preference to cover large firms, our sample is heavily tilted towards larger stocks, we still find significant delays in which stock prices incorporate new information from their linked stocks. Nevertheless, for an average investor, prices may be "approximately" efficient (see, e.g., Lo (2004)). We show that trading on linked-stock signals is costly due to high portfolio turnover, and only sophisticated traders can profit from this strategy.[5] Moreover, as trading costs have decreased and access to

[5]Gatev, Goetzmann, and Rouwenhorst (2006) investigate the performance of the "pairs trading" investment strategy and conclude that arbitrageurs with low transaction costs can earn a post-transaction-cost

information has improved over time (for example, for a fee, traders can now obtain machine-readable news), one would expect prices to become more efficient. Consistently, we show that the predictive ability of the linked-stock signal has declined over time.

This paper contributes to the literature on lead-lag effects in stock returns. Studies in this literature typically use ex-ante stock characteristics to explain the lead-lag return patterns. Lo and MacKinlay (1990)—the paper that started this literature—show that returns of large firms tend to predict returns of small firms.[6] Subsequent studies identify various measures of investor attention that are associated with information leadership. Specifically, Brennan, Jegadeesh, and Swaminathan (1993), Badrinath, Kale, and Noe (1995), and Chordia and Swaminathan (2000) use analyst coverage, institutional ownership, and trading volume, respectively, to show that common information diffuses slowly from stocks with high levels of investor attention to those with low levels. More recently, Cohen and Lou (2012) show that information diffuses slowly from single-segment firms to multi-industry conglomerates, and Menzly and Ozbas (2010) and Cohen and Frazzini (2008) find that information travels slowly between firms linked by the supply chain.[7] In contrast, Scherbina and Schlusche (2013) do not rely on stock characteristics or firms' positions in the supply chain to identify leaders and followers. Instead, their identification is derived from the out-of-sample ability of leaders to Granger-cause the returns of their followers.

The approach here is most similar to Scherbina and Schlusche (2013) in that we do not rely on ex-ante available stock characteristics or required disclosure documents to identify lead-lag relations. Moreover, by analyzing the content of common news, we are able to observe additional types of economic linkages between firms, such as business partnerships,

profit from trading on this strategy. They suggest that these profits are a compensation to arbitrageurs for enforcing market efficiency. The same argument may be applied here as well.

[6]Hou (2007) shows that this effect works within industries—i.e., large stocks lead small stocks within the same industry. This effect is most recently confirmed by DeMiguel, Nogales, and Uppal (2014).

[7]More recent work documents excessive contemporaneous return correlations among stocks with common institutional ownership (Anton and Polk (2014)), common analyst coverage (Israelsen (2013) and Muslu, Rebello, and Xu (2014)), and textually similar financial reports and news stories (Hoberg and Phillips (2012) and Box (2014)), and attributes the excessive correlations to similarity in trading or the information environment.

similar legal issues, common regulatory exposures, labor and production commonalities, and so on. Importantly, even after we remove firm pairs tied by supply-chain links, the cross-predictability in returns remains. While most topic categories that we construct are generally too small to be analyzed individually, we document the cross-predictability in returns between firms linked exclusively by similar legal issues and by a collection of other identified linkages that we describe as "shared goals and operational similarities."

Identifying firm linkages from news offers additional advantages. First, this methodology allows researchers to fill the gaps in data availability induced by limited disclosure (for example, firms are required by the SEC to report the identity only of the customers that each comprise more than 10% of a firm's consolidated sales revenues; hence, slightly less important customers will be missing from the dataset). Second, it makes it possible to uncover transitory leaders (for example, RDP's leadership should disappear shortly after the Supreme Court has reached its decision and journalists stop co-mentioning RDP with firms facing similar legal issues), whereas lead-lag relationships are assumed to be long-lived in the prior literature. Third, our methodology permits within-industry bets, while in the lead-lag literature intra-industry bets are, for the most part, precluded. Finally, similarly to Cohen and Lou (2012) and Scherbina and Schlusche (2013), we are able to show that smaller stocks can lead returns of larger stocks.

We also contribute to the relatively new strand of literature that investigates the news media's role in financial markets. This literature examines the extent to which qualitative information in news articles—in particular, the tone of the language—is associated with subsequent stock returns.[8] The literature also shows that news coverage helps improve price efficiency (see, e.g., Peress (2008) and Peress (2014)) and reduces firms' cost of capital (e.g.,

[8]Tetlock (2007) focuses on news articles about the broad stock market and documents that the linguistic tone predicts aggregate stock prices. Similarly, Tetlock, Saar-Tsechansky, and Macskassy (2008) find that content in firm-specific news articles is correlated with the returns of individual stocks. Following a different identification strategy, Dougal, Engelberg, Garcia, and Parsons (2012) show that the identity of columnists writing for the *Wall Street Journal* significantly predicts returns of the Dow Jones Industrial Average. Engelberg and Parsons (2011) document that trading behavior of investors in a number of local markets in response to the same event differs depending on local media coverage.

Fang and Peress (2009)). In this study, we uncover a different aspect of the news media's role in financial markets: We provide evidence that news stories contain soft information that helps identify economic linkages between firms.

The paper is organized as follows: Section II describes the data. Section III explains our methodology for identifying linked stocks. Section IV documents the ability of linked stocks to cross-predict one another's returns. Section V concludes.

II. Data

Stock-specific data, as well as asset pricing factors, are obtained from the Center for Research in Security Prices (CRSP).[9] Accounting data are obtained from the merged CRSP/Computstat database. Institutional ownership information comes from Thomson Reuters' 13f files. Analyst coverage is obtained from I/B/E/S.

Industry classifications are obtained from Kenneth French's web site.[10] Throughout the paper, we use 38 industry classifications, but our results are nearly unchanged when using 12 industry classifications instead. Table A1 in the Appendix presents, for our sample, the average fraction of firms in each industry. The industry classified as "Irrigation Systems" drops out of our sample after data restrictions explained below are imposed, reducing the number of industries to 37. Additionally, whenever portfolio sorts are performed within industries or when leaders are required to belong to a different industry than their followers, we drop stocks in the industry identified as "Other" because of the implied heterogeneity (however, as can be seen from the table, this industry has very few stocks).

[9] We adjust stock returns for delisting in order to avoid survivorship bias (Shumway (1997)) as follows: When a stock is delisted, we use the delisting return from CRSP, if available. Otherwise, we assume the delisting return to be -100%, unless the reason for delisting is coded as 500 (reason unavailable), 520 (went to OTC), 551-573, 580 (various reasons), 574 (bankruptcy), or 584 (does not meet exchange financial guidelines). In these cases, we assume that the delisting return is -30%.

[10] `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/det_48_ind_port.html`.

Finally, the news data are available from the Thompson-Reuters News Analytics (TRNA) dataset for the period April 1996 through December 2012. The dataset links news stories to TRNA's firm IDs (which we link to ticker symbols and then to permno's) and assigns news topic codes to each news item. Each distinct news story about a firm is labeled with a unique primary news access code (PNAC). The TRNA dataset provides news headlines and news sources, as well as a variety of quantitative scores for each news item, such as the sentiment score. The sentiment score takes on values +1, -1, or 0, indicating whether a story is positive, negative, or neutral for each firm mentioned in the news story. (This score is based on the predominance of positive/negative words that pertain to each firm mentioned in the text). We will rely on the sentiment score when deciding whether or not the co-mentioned stocks are competitors. A more detailed description of the TRNA dataset is provided in Scherbina and Schlusche (2013).

The tables and figures presented throughout most of the paper cover the period from July 1996 to January 2013, unless stated otherwise, as the initial three months are used to identify stocks linked through common news stories, and the linked-stock returns are forecasted one month ahead (the period starts in October 1996 when a six-month identification window is used and in April 1997 when a 12-month identification window is used).

III. Identifying Linked Stocks from New Stories

A group of stocks linked to stock i in month t comprises all publicly traded stocks $j = 1, \ldots, J^i$ that were co-mentioned with stock i in at least one news story in the Thomson-Reuters News Analytics (TRNA) dataset during a trailing three- (six- or 12-) month period, excluding the last five trading days in month t. We exclude these days to ensure that all news have become available to the broad market and that investors had sufficient time to identify linked stocks before forming predictive signals.[11]

[11]Throughout the paper, we refer to this period as the "identification period."

Figure 1 provides an example of two groups of linked stocks; all stocks connected with lines were co-mentioned in at least one news story over the identification window. In the example, stock A is linked to stocks B, C, and D, and stock B is linked to stocks A and E. Stocks A and E are not directly linked, but are linked via a "second-tier" linkage. We will discuss second-tier linkages in Section IV, Subsection H.

For the purposes of this analysis, we discard news stories that could randomly co-mention firms. Specifically, we remove stories about technical analysis, large price movements, index changes, credit rating changes, listings and delistings of equity, shareholder meetings, analyst recommendations, investment fund news, and trade order imbalances. When these conditions are not imposed and all news are kept, more unrelated stock pairs are erroneously identified through common news coverage as economically linked stocks. However, erroneously identified linkages simply add noise to the estimates of the linked-stock signal; results are not significantly different when all news stories are kept. Some large news stories are transmitted in parts, and we consider only the complete story in order not to double count the same news.[12]

Table I presents descriptive statistics for the TRNA dataset as well as descriptions of the stocks linked by common news stories. As shown in Panel A, the TRNA dataset contains a total of almost 5.5 million unique news stories. This sample is reduced by roughly a third after removing news that would randomly group stocks together. Of the remaining news sample, just over 14% mention more than one firm. As reported in Panel B of the table, this subset of news stories co-mentions 2.78 firms, on average. The median number of firms mentioned is two.

Despite removing what we consider to be irrelevant stories for our purposes, we are still concerned that firms may be grouped together in a story by coincidence, for example,

[12]Specifically, we remove stories with topic codes INSI, STX, HOT, INDX, AAA, LIST1, USC, MEVN, RCH, FUND, and DBT. We also discard stories with item genre 'Imbalance' (with different capitalization possibilities). Finally, at Reuters' suggestion, we delete news observations for which the variable "more_news" takes on values 'M' or 'm' and for which the variable "update_sz" is greater than 8500 in order to only consider complete news stories rather than the parts in which they were transmitted.

when reporting on market conditions or unrelated events (e.g., contemporaneous earnings announcements). Therefore, we further narrow our news sample to news stories that co-mention exactly two firms, which reduces our sample to 331,232 unique news stories.

As we will discuss in more detail in Section IV, our return predictability signal hinges on the presumption that after a news shock for a given firm, the prices of its linked stocks would move in the same direction. Hence, in order to minimize the likelihood that linked stocks would move in opposite directions, we eliminate news about competitors. The elimination is based on two conditions. First, we argue that a news story mentions competitors if the sentiment scores assigned to each of the two firms differ by an absolute value of two (i.e., firms are considered competitors if a news story about them is interpreted as "positive" for one firm and "negative" for the other). This requirement eliminates about 10% of the stories that co-mention exactly two firms. Second, we discard news stories that contain variations of the words "rival" or "competitor" in the headline, as these stories are likely to cover competing firms. As a validity check, in untabulated results, we confirm that the presence of the words "rival" or "competitor" in the headline is associated with more than twice as high a probability that the two firms mentioned in the text receive opposite sentiment scores. These two conditions reduce our dataset to 299,060 unique news stories, which is our final dataset.

Panels C through G of the table provide additional descriptive statistics for the final dataset. Panel C shows that the majority of firm pairs are co-mentioned just once during the identification window. Panel D shows that the probability that both stocks in the linked pair are in the same industry is greater than implied by chance. Since the news media tends to cover large firms more extensively, the number of linkages increases with a stock's market capitalization (Panel E), and linked stocks tend to be larger than unlinked stocks (Panel F). Panel F furthermore shows that only 20.51% (27.58%) of stock-months are linked when three-month (six-month) identification window is used. Panel G shows that economic linkages between stocks show significant persistence over time (the length of the identification window used for this table is three months). Specifically, the probability that a particular pair of linked stocks in month τ continues to be identified as such in month $\tau + 12$ is about 22.86%. Even five years later, this probability is 18.43%. For stock pairs that had at least two common news stories over the identification period and thus, presumably, are more closely linked, the persistence is even stronger, with the probability of linkage in month $\tau + 12$ equal to 42.76%.

IV. Return Predictability

A. Constructing predictive signals

Having identified linked stock pairs, we need to determine the direction of the information flow. Stocks with an abnormal turnover likely have experienced a news event.[13] Stocks with turnover within the normal range likely have not. Therefore, we assume return leaders to be stocks with an above-normal turnover in month t and, conversely, followers to be stocks with turnover within a normal range in month t, and information to flow from thus-identified leaders to the followers linked to these leaders. In determining turnover thresholds for leaders and followers, we balance the benefits of precisely identifying leaders and followers against the costs of reducing the sample size. Specifically, we assume that leaders are the stocks whose turnover in month t is above their own median turnover over the trailing 12-month window, $t - 11$ to t, and followers are the stocks whose turnover is below their own 75th percentile turnover over the same window.[14]

[13] Perhaps, a more straightforward method for identifying stock-specific news events would be to check whether a stock had a news story in the TRNA dataset. However, we are wary that due to labor and space constraints not all firm-specific news will be covered in the TRNA dataset.

[14] If both stocks in a linked pair had a normal level of turnover in month t, we assume that there is no return leader among the stocks in the pair and ignore this linkage when constructing a predictive signal for the following month's followers' returns. In addition, if both stocks' month-t turnover falls between the 50th and 75th percentiles, we assume that the information flow could be bi-directional. We have experimented

To construct a predictive signal in month t for the month-$t+1$ return of each follower stock i, we calculate the weighted average month-t return of its linked leader stocks $j = 1, \ldots, J^i$:

$$Signal_t^i = \sum_{j=1}^{J_t^i} \omega_j Ret_t^j, \tag{1}$$

where Ret_t^j is stock j's return in month t, excluding the last five trading days of the month to mitigate concerns regarding non-synchronous trading as well as to allow investors sufficient time to process news stories in order to compute linked-stock signals, and ω_j is the weight assigned to stock j's return. In the baseline specification, signaling stocks' returns are equal-weighted, in which case $\omega_j = 1/J^i$. In robustness checks, we consider alternative weighting schemes for the linked-stock returns, such as weighting returns by the number of co-mentions during the identification window or by the leaders' market capitalization in month $t-2$.

Finally, as is standard in the asset pricing literature, we impose two restrictions on the stocks whose returns we are predicting with linked-stock signals (the follower stocks). These stocks must be common shares of U.S.-incorporated firms (i.e., stocks with share codes 10 or 11) and they must be priced at or above \$5 per share at the end of month t.

B. Baseline specification

In the baseline specification, we set the identification window to three months, starting at the beginning of month $t-2$ and ending five trading days before the end of month t. We then identify linked stocks from co-mentions in the news over the trailing identification window, determine leaders and followers, and calculate linked-stock signals as previously described. Next, we sort stocks into quintiles *within* each industry based on the linked-stock signal; stocks in the same quintile are combined across industries into portfolios, thereby ensuring

with slightly varying turnover cutoffs, and these variations do not significantly affect the results. These results are available upon request.

an equal industry representation in each portfolio 1 through 5.[15] (Forming portfolios within industries helps us dismiss the concern that we are simply picking up the large-small stock lead-lag effect within industries or the industry momentum effect.) Stocks are held in the portfolios for one month, and the process is repeated in the subsequent month. The timeline for the portfolio formation is illustrated in Figure 2.

Table II reports various abnormal return measures—excess return, market alpha, and three-, four-, and five-factor alphas—for equal-weighted portfolios (Panel A) and value-weighted portfolios (Panel B) as well as return differentials between the extreme quintile portfolios.[16] The results indicate that signals from linked stocks have significant predictive ability for stock returns for both equal- and value-weighted portfolios. Portfolios that contain stocks with low linked-stock signals earn low returns, while portfolios that contain stocks with high signals earn high returns. Portfolio returns increase gradually with the linked-stock signal. Return differentials between quintiles 5 and 1, reported in the bottom row of each table, are positive and statistically significant, with their annualized five-factor alphas equal to more than 9% for both weighting schemes. Return differentials are somewhat lower for value-weighted than for equal-weighted portfolios. While large stocks receive more extensive news coverage that reveals useful information about stock linkages, which should make the linked-stock signal more informative, larger stocks are typically more efficiently priced. Our results show that the second effect slightly dominates. Panels A and B of the Appendix Table A2 present the factor loadings for equal- and value-weighted portfolios, respectively. The factor loadings are about the same across various quintile portfolios, which makes the factor loadings of the return differential between portfolios 1 and 5 statistically insignificant for both equal- and value-weighted portfolios.[17]

[15] If in a given month there are fewer than five eligible stocks in an industry, that industry/month observation is discarded.

[16] The reported t-statistics are adjusted for autocorrelation in returns using the Newey and West (1987) methodology, and, for each specification, the number of lags is determined as the cubic root of the number of observations in the time series.

[17] It is possible that it is some omitted systematic risk factor pertinent to the linked stocks, rather than an idiosyncratic news shock to one of the linked stocks, which affects the linked stocks' returns in both months t

As shown in the second column, stocks in the extreme quintile portfolios tend to be linked to fewer stocks than those in the middle portfolios. This is to be expected, as the more leaders a stock has, the more likely the weighted average return of these leaders ends up in the middle portfolio.

Importantly, if the lead-lag effect in the returns of linked stocks was driven not by information flows but rather, for example, by correlated trading, then the portfolio returns in the months following portfolio formation would reverse and the performance of the long-short portfolios in subsequent months would be negative. However, we do not observe any return reversal. In fact, the average monthly raw return differential between months $t+2$ and $t+6$ is slightly positive but insignificant for both equal- and value-weighted portfolios.[18]

The subsample analysis, presented in Panel C of Table II, shows that return differentials between high- and low-linked-stock-signal portfolios have declined over time. While the five-factor alphas of the return differentials for the period from July 1996 to December 2003 are 1.45% and 1.19% per month for equal- and value-weighted portfolios, respectively, they drop to 0.35% and 0.37% per month, respectively, for the period from January 2004 to January 2013. This finding is consistent with market participants becoming more adept over time at processing complex information about firms' interconnections.

To illustrate the profitability of a trading strategy based on the linked-stock signal, Figure 3 plots the cumulative return of a \$1 investment over the period from July 1996 to January 2013. A portfolio with returns equal to those of the equal-weighted long-short portfolio would have ended up worth \$5.38 on January 31, 2013; for the value-weighted long-short portfolio, this number would have been \$4.65. As explained earlier, these returns have no significant

and $t+1$. To make sure that this is not the case, we verify that there is no return continuation for the leaders from month t to month $t+1$. Specifically, we sort *leaders* into quintiles based on their month-t return and show that the return differential between the extreme portfolios is statistically indistinguishable from zero in month $t+1$. Moreover, in the cross-sectional regressions in Table V, we control for followers' month-t return and show that the regression coefficient on the own lagged return is negative rather than positive.

[18]For equal-weighted portfolios, this average monthly raw return differential is 0.10% (t-statistic=1.02), with the five-factor alpha of 0.02% (t-statistic=0.20); for value-weighted portfolios, these numbers are 0.14% (t-statistic=0.91) and 0.10% (t-statistic=0.75), respectively.

loading on any of the five priced factors. For reference, a \$1 investment in the S&P 500 index would have ended up worth only \$2.23 over the same period (and since the long-short returns represent a zero-investment strategy, *excess* returns on the S&P 500 index over Treasury bill returns would represent a more "fair" comparison; a \$1 investment at that rate of return would have ended up worth only \$1.43.)

C. Alternative specifications and robustness checks

In the following, we test whether the return predictability is sensitive to variations in the methodology used to form portfolios and to identify linked stocks. The results for various alternative specifications are reported in Table III.

As a first robustness check, we modify the weighting scheme used for the calculation of linked-stock signals. In Panel A, leaders' returns are weighted by the number of co-mentions with their follower over the three-month identification window, and, in Panel B, leaders' returns are value-weighted using their market capitalization in month $t-2$. Comparing the return differentials to those of the baseline specification in Table II shows that both alternative weighting methods reduce the return differentials, and more so for value-weighted portfolios. As for weighting by the number of co-mentions, as done in Panel A, more frequent co-mentions in the news could indicate a more tightly linked stock pair. At the same time, frequent co-mentions increase market participants' awareness of the linkage between the stocks, and, hence, speed up information transmission. Our results indicate that the latter effect dominates.[19] Hence, the news media facilitates the information flow between stocks. When leaders' returns are value-weighted, as done in Panel B, more weight is given to large leaders. Large firms are more visible to market participants, and it is less likely that their relevant news will be overlooked. In sum, both sets of results point to limited investor attention as a contributing factor for slow information transmission between linked stocks.

[19] In unreported results, we check the return predictability over shorter horizons and find that when stocks are co-mentioned more frequently, there is indeed less of a delay in the price reaction of linked stocks.

In Panels C and D, we introduce slight changes to the identification of linked stocks and of leaders and followers. In Panel C, we no longer discard news stories we assume to cover competing firms. As expected, the return differentials for both weighting schemes fall relative to the baseline specification. The reason is that a news shock to a firm could have an opposite-sign effect on its competitor's return, which would mute the predictive ability of the linked-stock signal. In Panel D, we no longer impose turnover conditions to identify leaders and followers among linked stocks. As a result, the return differentials are reduced, especially for value-weighted portfolios. Hence, our turnover restrictions indeed help to identify the direction of the information flow between linked firms.

In Panels E and F, we broaden our sample of common news stories. Instead of identifying firm linkages exclusively from news stories that co-mention exactly two firms, we allow all news stories that co-mention up to five firms in Panel E, and up to 10 firms in Panel F. With more firms co-mentioned, it is more likely that these firms are co-mentioned by coincidence rather than because they are truly linked; thus the return differentials are lower than in the baseline specification. This is especially apparent for equal-weighted portfolios, whose return differentials drop more in Panel F than in Panel E. Consistent with the fact that the news media devotes less space and effort to the coverage of small firms, it appears that small firms tend to be co-mentioned in stories that mention several other firms purely by chance (perhaps because all firms in the story happen to have a news event on the same day) rather than because these firms are economically linked. These results substantiate the previously expressed concern that news stories that co-mention many stocks convey less valuable information about firm linkages than those that co-mention only two stocks.

Next, we increase the length of our identification window. Instead of the three-month window of the baseline specification, we use a six-month window in Panel G and a 12-month window in Panel H (as before, we end the window five trading days before the month-end). The six-month window works slightly better than the three-month window. While the return differentials are about the same in magnitude as in Table II, their t-statistics are larger because portfolios now contain more stocks. In contrast, the 12-month window produces somewhat lower return differentials. An identification window that is too short will miss some still-in-effect stock linkages. In contrast, an identification window that is too long, will either pick up stock linkages that have already dissolved or linkages that have become well known to market participants, thereby ensuring quicker information transmission.

In the last four panels, Panels L - O, we introduce a number of restrictions on groups of linked stocks in order to identify various channels of information flows. With these restrictions in place, the sample size declines considerably, and it helps to have longer identification windows to increase the number of linked stock pairs. For this reason, we report results for the six-month identification window in addition to the three-month identification window of the baseline specification. In Panels L and M, for each follower, we limit the set of leaders to those that belong to a different industry and report results for the three-month and six-month identification windows, respectively. Portfolio return differentials decrease noticeably but remain significant for equal-weighted portfolios, despite a significant decrease in sample size. As in the example provided in the introduction, our results show that information can flow across industries and that the news media are able to identify inter-industry linkages. Results indicate that both inter- and intra-industry linkages are important for predicting linked-stock returns.

In Panels N and O, the set of leaders is limited to those that are smaller than their follower for the three- and six-month identification windows, respectively. Again, portfolio return differentials decrease relative to the baseline specification, but they remain significant for both equal- and value-weighted portfolios, which is remarkable given there are now significantly fewer stocks in each portfolio. Hence, in contrast to the findings in the lead-lag literature, which showed that large stocks lead the returns of small stocks, we show that smaller stocks experiencing news events can lead returns of larger economically linked stocks.

In sum, the results in this section highlight that limited attention contributes to slow information diffusion; that one has to exercise some care in how to extract information on stock linkages from news stories; and that information can flow in unexpected directions.

D. The content of common news stories

Using news stories to identify linked stocks allows us to analyze the content of common news stories and thereby observe various ties that can exist between firms. We show that firm linkages uncovered from news stories tend to capture similar business fundamentals rather than common investor sentiment. Moreover, we are able to check whether firm linkages other than customer-supplier ties, which have been studied in prior literature, can also give rise to the cross-predictability in stock returns and should, therefore, be studied in more detail in future work.

We classify news into topics by using key words for news headlines, as well as topic codes assigned by Thompson-Reuters. The algorithm to classify news stories into topic codes, which is described in detail in Appendix A1, consists of three steps. In the first step, we parse the headlines in our news dataset (that is, the dataset of common news stories that co-mention exactly two firms that we classify as non-competitors) into words; remove articles, prepositions, and conjunctions; and then rank the remaining words by the number of headlines in which they appear. In the second step, we assign frequently used words to one of 15 news topics. In the third step, we classify all news stories based on our key words and a subset of relevant Thompson-Reuters topics codes into one of the 15 news topics.

All news topics are defined to be mutually exclusive, and their hierarchy is presented in Appendix A1. If a news story can be classified into more than one topic, it is assigned to the topic closest to the top of the list. Moreover, if a firm pair has more than one common news story over the identification window, and the news stories are assigned to different topics, the nature of the linkage between the firms is classified into the topic closest to the top of the list. The hierarchy of news topics was determined based on whether the topic was studied in the prior literature (perhaps even in a different context) and on how narrowly and precisely the topic can be defined.

Following this approach, we are able to classify over 60% of the news stories in our sample. All classified stories describe some aspect of similar business fundamentals or direct economic ties that exist between firms. "Legal" is the largest topic category, containing 11.14% of news in our sample. Also prevalent are stories about natural resources, geopolitical developments, and regulation, encompassing 10.09%, 9.57%, and 7.66% of the sample, respectively. Of note, news stories about supply-chain relationships, which have been investigated extensively in prior literature, are not prevalent in our dataset and represent only 4.41% of the sample. (See Appendix A1 for more details.)

In order to show that firm linkages other than the supply-chain relations can give rise to the cross-predictability in returns, for the results in this section, we remove firm pairs that were assigned to the topic "Supply chain" at any time over a trailing three-*year* window. Additionally, in order to ensure that we do not inadvertently overlook supply-chain linkages between remaining firm pairs, we discard all unclassified news stories. Since these restrictions significantly reduce our sample of news stories, we increase the identification window to six months in order to obtain more linked pairs.

Panels A - C of Table IV report portfolio returns for linked stock pairs whose linkages belong to the categories described in each panel heading. In Panel A, we keep all firm pairs that remain after the restrictions described above are implemented. Each signal-sorted portfolio contains roughly 63 stocks, which represents a 32% decrease relative to the portfolio size of the unrestricted sample with a six-month identification window in Panel G of Table III. Though the return differentials in Panel A are lower than those in the unrestricted sample, they are economically meaningful and statistically significant. These results indicate that economic linkages other than customer-supplier linkages can induce cross-predictability in returns.

In Panel B, we consider only firm pairs tied by the most populous news category, "Legal." The number of stocks per portfolio declines to about 20, which makes portfolio returns rather volatile and, thereby, reduces the statistical significance levels of the extreme-portfolio return differentials. Nonetheless, equal-weighted portfolio return differentials are very large and statistically significant. Though value-weighted return differentials are positive, they are insignificant. These results suggest that legal similarities induce a considerable cross-predictability in returns.

Finally, Panel C contains all stock pairs from Panel A that do not appear in Panel B. The linkages between these firm pairs can be broadly described as "similar operations and shared goals" since the common news stories describe firms operating in similar environments (e.g., conducting business in same countries; being similarly affected by regulations; facing similar labor issues; etc.) and having common business objectives (e.g., through business partnerships, parent-subsidiary relations, and cross-financing). The number of stocks in each portfolio is roughly 20% lower than in Panel A. The return differentials are close in magnitude to those in Panel A.[20]

All told, these results show that a wide variety of firm linkages, and not only the linkages that exist along the supply chain, can induce cross-predictability in stock returns. In particular, we are able to show that linkages based on similar legal exposures and on operational similarities and shared business goals induce cross-predictability in stock returns. As the news sample continues to grow over time, it will become possible to investigate more narrow types of firm interlinkages. Importantly, our ability to produce significant return differentials focusing deliberately on news about fundamentals rather than similar investor

[20] In a contemporaneous paper, Cao, Chordia, and Lin (forthcoming) show that firms linked by alliances cross-predict each other's returns. We have tried removing firm pairs in the category "Partnership" in Panel C, and the return differentials remain significant.

sentiment suggests that common fundamentals are a significant driver of the lead-lag effect in stock returns.[21]

E. Placebo tests

In order to rule out the possibility that, rather than the linked-stock signal, a different stock characteristic, omitted from the portfolio sorts, drives the cross-predictability of stock returns, we conduct a placebo test of the cross-predictability of returns. Since Panel G of Table I shows that stocks that are mentioned in a common news story continue to be co-mentioned in the subsequent five years, we check whether the cross-predictability in returns existed *before* rather than after a stock pair was first co-mentioned in the news, speculating that the cross-predictability should not yet exist at that time.

Specifically, we pretend that each observed co-mention in the news had occurred 12 months earlier and make sure that this hypothetical co-mention had occurred at least six months prior to the *first* time that the two firms were co-mentioned in the news. If both stocks in the linked pair are present in the CRSP dataset at that time, we proceed to compute linked-stock signals and form portfolios exactly as we did in the baseline specification of Table II.

As we expected, now that the linkage dates are moved one year back, we obtain insignificant portfolio return differentials. Specifically, the five-factor alpha of the equal-weighted return differential is 0.12% (with a t-statistic of 0.45), and the five-factor alpha of the value-weighted return differential is -0.33% (with a t-statistic of -0.66). The placebo results, therefore, confirm our conjecture that it is the information about economic linkages embedded in the news co-mentions rather than some stock-specific characteristics that explains

[21]In the future, as the news dataset continues to grow and news coverage becomes more comprehensive, the return predictability could be improved by selecting only stocks with relevant linkages for each firm-specific news event by matching the nature of the linkages to the content of the news (e.g., using the example from the introduction, when RDP receives news about its courtroom development, we will select only stocks linked to RDP through the "legal" channel).

why linked stocks cross-predict each other's returns. Moreover, these results suggest that the TRNA dataset contains rather timely reports on new firm linkages and that the majority of these linkages likely has not existed long prior to the first co-mention in a TRNA news story.

F. Cross-sectional regressions

In this section, we run a number of Fama and MacBeth (1973) cross-sectional regressions to provide further evidence of the previously documented predictive ability of the linked-stock signal. The regression setting allows us to include other variables known to predict returns. We show that these control variables do not subsume the linked-stock signal's predictive ability, which further confirms that we have identified an independent source of return predicability.[22] The linked-stock signal is computed every month as described in the baseline specification.

The results of the cross-sectional regressions are reported in Table V. All but the first specification include the month-t stock return and the stock's value-weighted industry return as additional cross-sectional return predictors. In all specifications but the third, we also include size, book-to-market, and the average stock return over six months from $t-6$ to $t-1$. Specifications (7)-(9), in addition, contain interaction terms between the linked-stock signal and additional control variables meant to capture the level of attention from sophisticated market players; these regression specifications also include the stock characteristic in the interaction term as an additional control variable.

The regression coefficient on the linked-stock signal is highly significant in all specifications. The estimated coefficient ranges from 0.018 to 0.036 in magnitude; these magnitudes are in line with our portfolio results. The coefficients on the linked-stock signal are more statistically significant than the coefficients on the lagged industry return or the firm's own

[22]The control variables are described in detail in Appendix A2.

lagged return in all regression specifications. The p-value of the linked-stock signal never rises above 1%.

In regression specifications (7)-(9), we check whether a high level of sophisticated investor attention helps speed up information processing since sophisticated investors may be more skilled at identifying linked stocks and interpreting their relevant news. In specification (7), our proxy for sophisticated investor attention is a dummy variable that equals one if institutional ownership of a stock is above the median value in the cross-section of stocks at the most recent quarter-end, and zero otherwise. The coefficient on the interaction term between this proxy and the linked-stock signal is negative and significant at the 5% level. This suggests that the one-month-ahead predictive ability of the linked-stock signal is lower for stocks with high levels of institutional ownership, and, hence, the attention from institutional investors appears to speed up information processing.

In specifications (8) and (9), we use less direct proxies for sophisticated investor attention—analyst coverage and size, respectively—and also interact these attention proxies with the linked-stock signal. As before, our proxies are dummy variables that are equal to one if these variables are above the median value in the cross-section of stocks at the end of month t and zero otherwise. The coefficient on the interaction between the linked-stock signal and the high-analyst-coverage dummy is negative, but insignificant. The coefficient on the interaction term between the linked-stock signal and the large-size dummy is negative and statistically significant at the 10% level.

Overall, the results confirm the robustness of the linked-stock signal as a return predictor and show that the attention from sophisticated investors helps speed up the price reaction to linked-stock signals. Therefore, the slow reaction to the news of linked stocks may be, at least in part, caused by slow processing of complex information.

G. Break-even trading costs

Thus far, we have abstracted from trading costs. Even though we have documented significantly positive monthly excess returns of the long-short portfolios formed on the linked-stock signal, high trading costs associated with trading on this strategy may render these simple trading strategies unprofitable. In the following, we estimate break-even trading costs that would set the post-transaction-cost five-factor alphas of the portfolio return differentials equal to zero. Our estimates of break-even trading costs are expressed in units of return, i.e., as a percentage cost per dollar traded. In order to conserve space, we estimate break-even trading costs only for the baseline trading strategy that corresponds to Table II.

We assume that trading costs of the same magnitude are incurred when a stock is bought or sold and that these costs are identical across stocks (though in reality trading costs decrease with stocks' liquidity and increase with the number of shares traded). Table A4 in the Appendix shows that the extreme portfolio assignments exhibit a slight persistence over a one-month period (which disappears over a two-month period); hence, portfolio turnover is slightly lower than 100% in any given month. All else equal, this introduces a slight advantage for value-weighted trading strategies since they do not incur rebalancing costs for buy-and-hold portfolios. However, value-weighted returns are somewhat lower than equal-weighted returns. On net, our estimate of break-even trading costs is slightly lower for value-weighted than for equal-weighted portfolios. Specifically, our estimates of break-even trading costs are 0.22% for equal-weighted portfolios and 0.20% for value-weighted portfolios.[23]

To put these estimates into perspective, using the TAQ dataset for the period from January 1983 to August 2001, Sadka and Scherbina (2007) estimate an average effective spread of 0.25% for a typical stock and a typical trade. This number is somewhat higher than both our estimates of break-even trading costs. Hence, for an average investor, prices appear to lie within the no-arbitrage bounds around the fair value imposed by trading costs.

[23]In reality, value-weighted strategies, of course, involve lower trading costs because larger amounts are traded in large stocks, which are typically more liquid.

In contrast, sophisticated institutional investors are more skilled at minimizing trading costs than an average trader in the TAQ dataset, and their typical trading costs may easily fall below our break-even estimates.[24] Hence, sophisticated investors trading on this strategy may earn a profit, which can be viewed as a compensation for ensuring price efficiency.

The pre-arbitrage-cost profitability of the linked-stock-signal trading strategy should be closely related to the costs of arbitrage. Arbitrage costs fell over time: Trading costs have declined, and machine-readable news platforms have reduced the labor intensity involved in processing news. Our result that the pre-arbitrage-cost profitability of the linked-stock signal has declined over time is consistent with falling arbitrage costs.

H. Firms connected by second-tier links

If a firm's news coverage is limited, as may be the case for smaller firms, it is possible that second-tier linkages in the news may also contain useful information about economic linkages between firms. Figure 1 illustrates a second-tier linkage between firms A and E: these firms are not directly linked via a co-mention in a news story but are linked through their respective direct linkages to firm B. It is quite possible that the information from second-tier links is too muted to contribute sufficiently to the return predictability. Whether or not it is useful to collect the data on second-tier links is an empirical question.

In order to check whether stocks connected through second-tier linkages still cross-predict each other's returns, we identify all stock pairs that are not directly linked over the identification window but share a first-tier linkage. We then form the predictive signal for each stock by calculating equal-weighted returns from all second-tier linkages, having imposed the

[24] Frazzini, Israel, and Moskowitz (2012) describe sophisticated trading algorithms designed by institutional investors to minimize trading costs and tabulate a monthly implementation shortfall (that is, the paper return lost due to trading costs) incurred by a large institutional money manager. When trades are weighted by their dollar amount, the average implementation shortfall is only 0.20% a month, which is substantially below our long-short monthly portfolio returns.

same turnover conditions as in the baseline specification. We then proceed to form monthly portfolios based on the signals from second-tier links.

Table VI reports portfolio results for the three-, six-, and 12-month identification windows. The number of stocks in the second-tier linked portfolios is about one-third larger than that in the respective portfolios constructed based on first-tier links. This is because there are more second-tier than first-tier links, and more linked stocks meet the criteria for being classified as leaders or followers. The equal-weighted return differentials are significantly positive in all three specifications. The five-factor alpha for the value-weighted return differentials is significant for the six-month identification window, marginally significant for the three-month identification window, and insignificant for the 12-month identification window. In unreported Fama-MacBeth regressions, we verify that the signal from second-tier linked stocks has an incremental forecasting ability over the signal from first-tier links.[25]

These results suggest that the information about links-of-links clearly adds value for the identification of economic linkages for smaller stocks. This is not surprising since small stocks receive less coverage than large stocks in business press. For large stocks, the information about economic linkages is more readily available through the direct co-mentions in the news.

V. Conclusion

This paper documents that economically linked stocks cross-predict each other's returns. And, while it may be impossible to uncover economic linkages between firms only from required disclosures, we show that economic linkages can be identified through media coverage. Specifically, we argue that a pair of stocks is economically linked if the two stocks are co-mentioned in a news story over a trailing identification window. Having identified groups

[25] In particular, when the signal from second-tier links is included along with the linked-stock signal in the Fama-MacBeth regression, corresponding to specification (2) of Table V, its regression coefficient is 1.97, with a t-statistic of 1.87 (the linked-stock signal remains highly significant, with a regression coefficient of 2.14 and a corresponding t-statistic of 3.89). Full regression results are available upon request.

of linked stocks, we conjecture the direction of the information flow based on turnover conditions: firms with above-average turnover likely have experienced a news event, making them return leaders, and firms with turnover in the normal range likely have not, making them followers. We compute the predictive signal for the followers by equal-weighting the prior-month's returns of their leaders. Results show that the linked-stock signal is a reliable predictor of a follower's next-month return and that its predictive ability is robust to various firm- and industry-level controls.

Information processing frictions contribute to the slow diffusion of information between linked stocks. The predictive ability of the linked-stock signal is weaker for firm pairs that are frequently co-mentioned in the news. It is also lower when the linked-stock signal is formed by value-weighting leaders' returns, thereby assigning lower weights to the returns of smaller leaders, whose news are more likely to go unnoticed. This suggests that limited investor attention contributes to slow information diffusion. Moreover, the presence of sophisticated investors reduces the predictive power of the linked-stock signal, which suggests that the slow reaction to the news of linked stocks may be, in part, explained by slow processing of complex information.

Our methodology allows us to identify lead-lag return relations without relying on ex-ante stock characteristics. As a result, we are able to identify return leaders that are smaller than their followers and to detect short-lived lead-lag relations. Moreover, our approach allows for within-industry bets. Such bets are largely precluded in the lead-lag literature, as the predictive signals in that literature tend to be correlated within industries.

Our estimations show that high portfolio turnover makes it difficult to earn significant profits from trading on linked-stock signals. However, sophisticated investors with low trading and information processing costs may be earning a profit, and this profit can be viewed as compensation for ensuring price efficiency.

An important advantage of identifying firm linkages from co-mentions in the news is the ability to study the content of common news stories in order to learn about the nature of firm linkages. We are able to classify common news stories along a variety of dimensions: similar operations; similar regulatory environments; similar technologies; similar labor, production and infrastructure issues; business partnerships; and so on. Importantly, we document that not only customer-supplier ties, but also other types of linkages can give rise to the cross-predictability of stock returns. Legal similarities appear to be a particularly important type of linkage in our sample. As the news dataset continues to grow, it will allow for a detailed study of each type of stock linkage.

References

Anton, Miguel, and Christopher Polk, 2014, Connected stocks, *Journal of Finance* 69, 1099–1127.

Badrinath, S. G., Jayant R. Kale, and Thomas H. Noe, 1995, Of shepherds, sheep, and the cross-autocorrelations in equity returns, *Review of Financial Studies* 8, 401–30.

Box, Travis, 2014, Stock price comovement and the market for information, Working paper.

Brennan, Michael J, Narasimhan Jegadeesh, and Bhaskaran Swaminathan, 1993, Investment analysis and the adjustment of stock prices to common information, *Review of Financial Studies* 6, 799–824.

Cao, Jie, Tarun Chordia, and Chen Lin, forthcoming, Alliances and return predictability, *Journal of Financial and Quantitative Analysis* forthcoming.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.

Chordia, Tarun, and Bhaskaran Swaminathan, 2000, Trading volume and cross-autocorrelations in stock returns, *Journal of Finance* 55, 913–935.

Cohen, Lauren, and Andrea Frazzini, 2008, Economic links and predictable returns, *Journal of Finance* 63, 1977–2011.

Cohen, Lauren, and Dong Lou, 2012, Complicated firms, *Journal of Financial Economics* 104, 383–400.

DeMiguel, Victor, Francisco J. Nogales, and Raman Uppal, 2014, Stock return serial dependence and out-of-sample portfolio performance, *Review of Financial Studies* 27, 1031–1073.

Dougal, Casey, Joseph Engelberg, Diego Garcia, and Christopher Parsons, 2012, Journalists and the stock market, *Review of Financial Studies* 25, 639–679.

Engelberg, Joseph, and Christopher Parsons, 2011, The causal impact of media in financial markets, *Journal of Finance* 66, 67–97.

Fama, Eugene F., and Kenneth R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 46, 427–466.

——— , 1993, Common risk factors in the returns of stocks and bonds, *Journal of Financial Economics* 33, 3–56.

——— , 2000, Characteristics, covariances, and average returns: 1929-1997, *Journal of Finance* 55, 389–406.

Fama, Eugene F., and James MacBeth, 1973, Risk, return and equilibrium: Empirical tests, *Journal of Political Economy* 51, 55–84.

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

Frazzini, Andrea, Ronen Israel, and Tobias J. Moskowitz, 2012, Trading costs of asset pricing anomalies, Working paper.

Gatev, Evan G., William N. Goetzmann, and K. Geert Rouwenhorst, 2006, Pairs trading: Performance of a relative value arbitrage rule, *Review of Financial Studies* 19, 797–827.

Hoberg, Gerard, and Gordon M. Phillips, 2012, Industry choice and product language, Working paper.

Hou, Kewei, 2007, Industry information diffusion and the lead-lag effect in stock returns, *Review of Financial Studies* 20, 1113–1138.

Israelsen, Ryan D., 2013, Does common analyst coverage explain excess comovement?, Working paper.

Lo, Andrew W., 2004, The adaptive markets hypothesis, *Journal of Portfolio Management* 30, 15–29.

——— , and A. Craig MacKinlay, 1990, When are contrarian profits due to stock market overreaction?, *Review of Financial Studies* 3, 175–205.

Loughran, Tim, and Bill McDonald, 2011, When is a liability not a liability? Textual analysis, dictionaries, and 10-ks, *Journal of Finance* 66, 35–65.

Menzly, Lior, and Oguzhan Ozbas, 2010, Market segmentation and cross-predictability of returns, *Journal of Finance* 65, 1555–1580.

Muslu, Volkan, Michael Rebello, and Yexiao Xu, 2014, Sell-side analyst research and stock comovement, *Journal of Accounting Research* 52, 911–954.

Neuhierl, Andreas, Anna Scherbina, and Bernd Schlusche, 2013, Market reaction to corporate press releases, *Journal of Financial and Quantitative Analysis* 48, 1207–1240.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive-definite, heteroscedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Pástor, Ľuboš, and Robert F. Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 111, 642–685.

Peress, Joel, 2008, Media coverage and investors' attention to earnings announcements, Working paper.

———, 2014, The media and the diffusion of information in financial markets: Evidence from newspaper strikes, *Journal of Finance* 69, 2007–2043.

Sadka, Ronnie, and Anna Scherbina, 2007, Analyst disagreement, mispricing and liquidity, *Journal of Finance* 62, 2367–2403.

Scherbina, Anna, and Bernd Schlusche, 2013, Cross-firm information flows and the predictability of stock returns, Working paper.

Shumway, Tyler, 1997, The delisting bias in CRSP data, *Journal of Finance* 52, 327–340.

Tetlock, Paul C., 2007, Giving content to investor sentiment: The role of media in the stock market, *Journal of Finance* 62, 1139–1168.

———, Maytal Saar-Tsechansky, and Sofus Macskassy, 2008, More than words: Quantifying language to measure firms' fundamentals, *Journal of Finance* 63, 1437–1467.

Appendix

A1. Classifying news stories into topics

We classify news into topics using key words in news headlines as well as relevant topic codes assigned by Thompson Reuters. The news dataset is the one used in the baseline analysis (i.e., the set of news stories that co-mention exactly two firms that we consider to be non-competitors).

We classify news stories into topics in three steps. In the first step, we parse all headlines into words and discard articles, preposition, and conjunctions. We then rank the remaining words by the frequency of appearance in *unique* headlines and keep all words that appeared in at least 100 headlines. For reference, "update" is the most frequently used word, and it appears in 97,276 headlines. The word "layoffs" appears in 100 headlines and is therefore among the least frequently used words that we consider.

In the second step, we assign each word to one of 15 news topics, the selection of which is heavily influenced by the topic list used in Neuhierl, Scherbina, and Schlusche (2013). For example, we deem the word "update" to be uninformative about the content of the news but assign the word "layoffs" to the topic "Labor, production and infrastructure." Our list of keywords includes variations in spelling and hyphenation, as well as words that have a common root with the original keyword. We further augment our list of keywords with keywords used in Neuhierl, Scherbina, and Schlusche (2013) for the relevant topics; for the topic "Legal," we augment our keyword list with keywords from the online legal wordlist, with finance terms removed (see, e.g., Loughran and McDonald (2011)).[26] We also identify relevant narrowly-defined topic codes in Thompson Reuters to be used in conjunction with our keywords to classify news stories into topics. Table A3 contains a comprehensive list of key words and TRNA topic codes that we use for each topic.

[26] `http://www.enchantedlearning.com/wordlist/legal.shtml`

In the third step, we classify news stories into 15 topics based on both keywords and the Thompson-Reuters topic codes. If a story can be classified into more than one topic, we use the topic closest to the top of the list. For example, if a story describes a merger objection lawsuit, the story is assigned to the topic "M&A" since it is higher on the hierarchy than the topic "Legal." (Our topic hierarchy is based on how frequently a particular topic was studied in the prior literature and how narrowly it can be defined.) If a headline cannot be classified, it is assigned to the category "Other."

Listed below are the topic descriptions, ordered by the topic hierarchy, from highest to lowest, with the percentage of news stories in our sample assigned to that topic provided in parentheses. We also include headlines of representative news stories and in parentheses provide the names of the firms mentioned, as well as the story's PNAC and date.

1. **Supply chain (4.41%).** News involving customers and suppliers.
   EXAMPLE: Headline: "ExxonMobil says it is the exclusive supplier for 33 Caterpillar(reg) lubricant" (Firms: Exxon Mobil and Caterpillar; PNAC: n0322NATN1; Date 22MAR1997).

2. **Partnerships (3.66%).** News about corporate partnerships, strategic alliances, and licensing deals.
   EXAMPLE: Headline: "Unilever says 50-50 venture with Pepsico <PEP.N> to expand Lipton Tea distribution" (Firms: Unilever and Pepsico; PNAC: nAAT003618; Date: 30JUL2004).

3. **M&A (5.01%).** News, updates, and speculations about corporate mergers and acquisitions.
   EXAMPLE: Headline: "Government argues Staples merger cannot be undone" (Firms: Staples and Office Depot; PNAC: n6114DAL; Date 01JUN1997).

4. **Parent/subsidiary relations (1.06%).** Stories about parent companies and their subsidiaries.
   EXAMPLE: Headline: "Time Warner Inc. declares spin-off dividend of Time Warner Cable Inc. shares and announces March 27 effective date for one-for-three reverse stock split <TWC.N><TWX.N>" (Firms: Time Warner and Time Warner Cable; PNAC: nBw266205a; Date: 26FEB2009).

5. **Legal (11.14%).** News about various legal issues regarding accounting fraud, labor, production and environmental problems, class action lawsuits, SEC's concerns and investigations, criminal investigations, etc.

EXAMPLE: Headline: “FOCUS-Agribusiness put on defensive by legal threat” (Firms: E.I. DuPont de Nemours & Co. and Astrazeneca Plc; PNAC: n5439663; Date: 13SEP1999).

6. **Regulation (7.66%).** News about government and agency regulations, as well as political actions affecting corporations.
EXAMPLE: Headline: “U.S. Senate panel targets offshore profits and taxes” (Firms: Microsoft and Hewlett Packard; PNAC: nL1E8KIGMB; Date: 18SEP2012).

7. **Labor, production, and infrastructure (2.90%).** News about the firm’s labor force, products, operations, and infrastructure.
EXAMPLE: Headline: “UK’s Britannia gas field would have to shut if Forties North Sea oil pipeline is closed due to planned strike” (Firms: BP and ConocoPhillips; PNAC: nL21545896; Date: 21APR2008).

8. **Executive compensation and corporate governance (0.01%).** News about executive compensation and firms’ corporate governance.
EXAMPLE: Headline: “WPP <WPP.L> shareholders back executive bonus plan” (Firms: Omnicom Group and Interpublic Group Cos; PNAC: nLAT001047; Date: 02SEP1999).

9. **Management news (0.05%).** News about changes in top management: promotions, retirements, firings, managers changing firms.
EXAMPLE: Headline: “H&R Block announces management change” (Firms: H&R Block and Fiserv; PNAC: nWEN3804; Date: 07NOV2005).

10. **Common customer (0.32%).** News about common customers, typically, U.S. federal and state governments and the military and other countries’ government projects and militaries.
EXAMPLE: Headline: “GE and Juniper Networks to develop family of rugged, secure network appliances for military vehicles <GE.N><JNPR.O>” (Firms: General Electric and Juniper Networks; PNAC: nMKW18241a; Date: 07NOV2011).

11. **Cross-investments (0.67%).** News about cross-investments, firms leasing each other’s assets and extending loans to each other.
EXAMPLE: Headline: “TowerJazz and GE Capital sign definitive asset based loan agreement to provide up to 4 billion yen credit line (approximately $50 million)” (Firms: General Electric and Tower Semiconductor Ltd; PNAC: nWNAB9038; Date: 09MAY2012).

12. **Natural resources (10.09%).** News that mentions raw materials used as production inputs, as well as natural and environmental disasters affecting firms’ operations.
EXAMPLE: Headline: “RPT-India Panna-Mukta fields, shut by explosion, likely to resume oil output by next week” (Firms: Bunge Ltd and InterOil Corp; PNAC: nDEL001551; Date: 12JUN2008).

13. **Energy (2.58%).** News that mentions energy inputs.
EXAMPLE: Headline: “Gas strike threatens to shut Sri Lanka industry” (Firms: Royal Dutch Petroleum and Sheldahl Co; PNAC: nCOL001356; Date: 07MAY1997).

14. **Technology (3.02%).** News that mentions various production/operation technologies. EXAMPLE: Headline: "ATM security flaws could be a jackpot for hackers" (Firms: Diebold Inc and NCR Corp; PNAC: nN25138100; Date: 25JUN2010).

15. **Geopolitical (9.57%).** News about firms' foreign operations and geopolitical events, regional conflicts, sovereign policies, etc., affecting these operations.
EXAMPLE: Headline: "Kiev local favouritism seen spurring Motorola exit" (Firms: Motorola Solutions and Royal KPN; PNAC: nFLLB41CON; Date: 01APR1997).

16. **Other (37.84%).** News that could not be classified into any of the above categories.

A2. Variable definitions and estimations

This appendix provides detailed descriptions of the variables used in our cross-sectional regressions. Unless specified otherwise, all variables are calculated at the month-end.

Previous month's industry return *(Ind. Ret)* is defined as the value-weighted industry return over the previous month.

Size *(Size).* A stock's size is defined as the product of the price per share and the number of shares outstanding, expressed in thousands of dollars and measured at the end of the previous month.

Book-to-market ratio *(Book/Market).* Following Fama and French (1992, 1993, 2000), the book-to-market equity ratio is computed at the end of June of each year as the book value of stockholders' equity, plus deferred taxes and investment tax credit (if available), minus the book value of preferred stock, scaled by the market value of equity. Depending on availability, we use the redemption, liquidation, or par value (in that order) to estimate the book value of preferred stock for the previous fiscal-year end. The market value of equity is the product of share price and the number of shares outstanding at the end of December of the previous fiscal year.

Previous month's stock return *(Rev).* This short-term reversal predictor is defined as the stock return over the previous month.

Stock's momentum return *(Mom. Ret).* This variable is meant to capture the return continuation and is calculated as the average own stock return from month $t-7$ to $t-2$.

Institutional Ownership *(Inst. Ownership)* is defined as the percentage of total shares outstanding owned by institutions, computed using the data in the Institutional Holdings (13F) dataset and measured at the end of the previous month.

Analyst Coverage *(Analyst Cov.)* is defined as the number of analysts issuing annual earnings forecasts for the current fiscal year, computed using the I/B/E/S dataset and measured at the end of the previous month.

Leader signal *(Leader Signal)* is defined as the monthly frequency signal computed from stocks that have Granger-caused a given stock's monthly returns over a trailing 12 months, as described in Scherbina and Schlusche (2013).

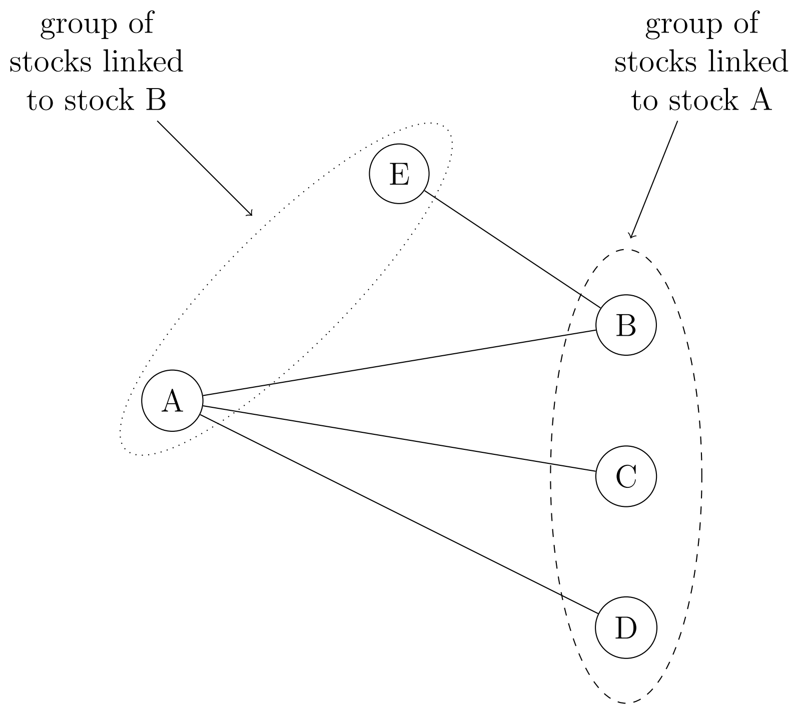

Figure 1. Examples of Linked Stocks. This figure provides an example of two groups of linked stocks. The dashed ellipse contains the stocks directly linked to stock A, and the dotted ellipse contains stocks directly linked to stock B. Stocks A and E are not directly linked but are linked through a second-tier linkage.

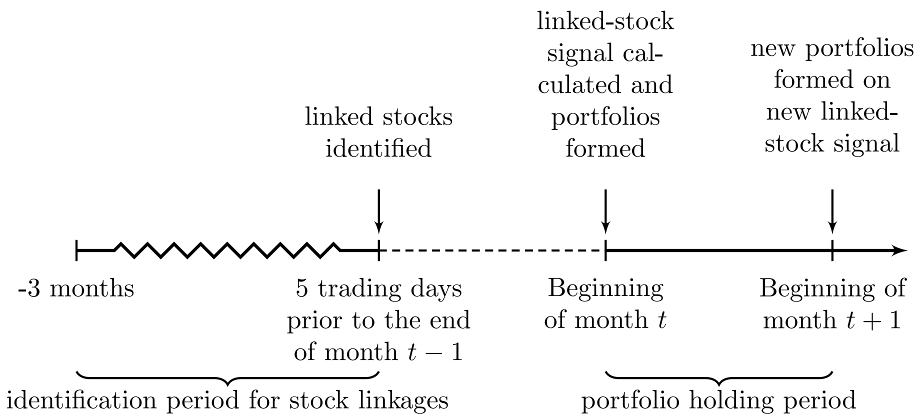

Figure 2. Timeline for Analysis. This figure presents the timeline for the identification of stock linkages, the calculation of predictive return signals, and the formation of portfolios. The zigzag line indicates a long period of time and the dashed line a short period of time relative to the period between months t and $t+1$.

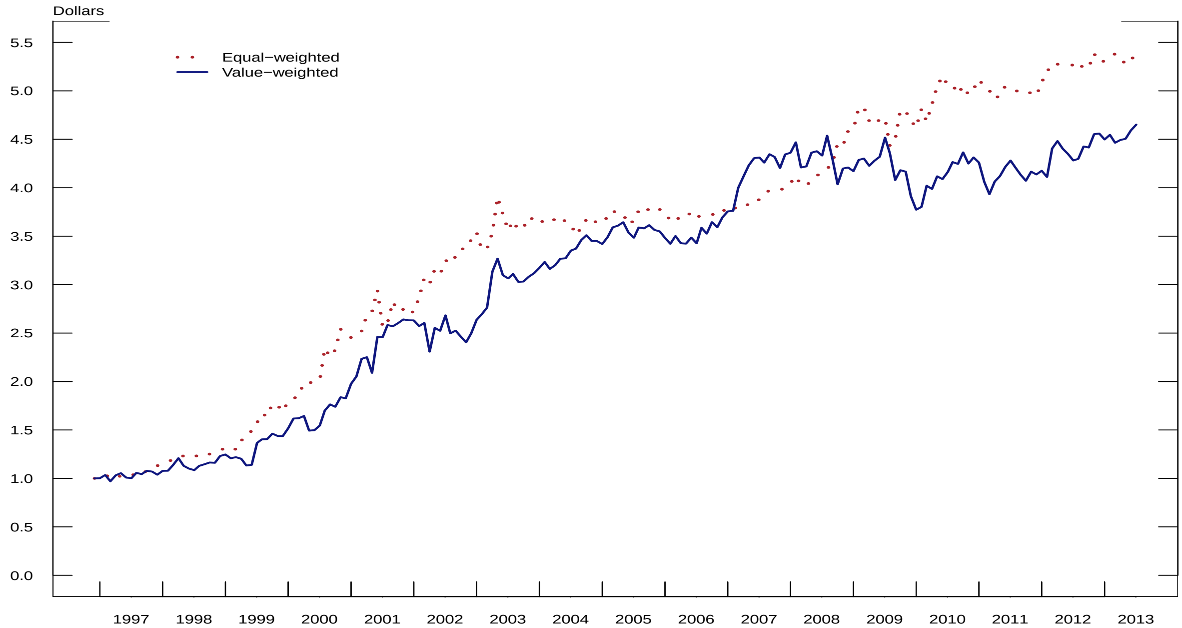

Figure 3. Cumulative returns. The charts plot, for equal- and value-weighted portfolios, the value of $1 invested in the beginning of the period at the return earned on a zero-investment strategy of buying stocks in the top and selling short stocks in the bottom linked-stock-signal quintile. Portfolios are formed as in the baseline specification of Table II. The time period is July 1996–January 2013.

Table I
Descriptive statistics

This table provides descriptive statistics for the TRNA dataset and for the news sample used in the main tests of the paper. It also presents a description of the stocks linked by common news stories. The investible CRSP universe comprises stocks with share codes 10 or 11. In Panels B through G, the statistics are calculated based on Sample 3 of Panel A. Panel G presents probabilities that a stock pair classified as linked by common news in month τ will remain linked in week $\tau + t$, given that both stocks are still present in the CRSP dataset at that time. A stock pair is classified as linked if (1) there was a common news story written exclusively about these two stocks in the previous three months and (2) the stocks are identified as non-competitors. The sample period is July 1996–December 2012.

Panel A: The news dataset

Total number of unique news stories	5,455,605
# of stories remaining after removing stories about trade order imbalances, index changes, bond ratings news, analyst recommendation revisions, etc.	3,689,918
Of these:	
# of stories that mention more than one firm (Sample 1)	521,845
# of stories that mention exactly 2 firms (Sample 2)	331,232
Of the stories that mention exactly 2 firms:	
# of stories that are about non-competitors (Sample 3)	299,060

Panel B: Statistics on the number of firms mentioned per story (based on Sample 1)

mean	median	75th percentile	95th percentile
2.78	2	3	5

Panel C: The distribution of the number of common news stories for linked pairs (computed in non-overlapping windows)

no. of common news	% of total	
	3-month window	6-month window
1	64.48%	62.20%
2	16.27%	16.29%
3	6.47%	6.51%
4	3.38%	3.67%
5	2.09%	2.31%
>5	7.30%	9.03%

Panel D: Fraction of linked firm pairs in the same industry

38 industry classifications used	40.86%
12 industry classifications used	49.06%

Panel E: The average number of stock linkages by CRSP size decile

CRSP size decile	Identification window used: 3 months	6 months
1	0.07	0.13
2	0.08	0.15
3	0.10	0.19
4	0.13	0.24
5	0.18	0.32
6	0.22	0.38
7	0.27	0.48
8	0.34	0.63
9	0.56	1.02
10	2.47	4.32

Panel F: Linked vs. unlinked firm-months in the investible CRSP universe

	Identification window used			
	3 months		6 months	
	linked	unlinked	linked	unlinked
Fraction of investible CRSP universe	20.51%	79.49%	27.58%	72.42%
Average market capitalization ($, mil.)	11,624.64	1,157.73	9,413.94	977.84
Average CRSP size decile	7.57	4.97	7.34	4.80
Median CRSP size decile	8	5	8	5

Panel G: Persistence of linkages between stock pairs

Number of months in the future (t)	All stock pairs: common news count	All stock pairs: prob.	Stock pairs with at least two common stories: common news count	Stock pairs with at least two common stories: prob.
3	2.44	27.10%	5.31	50.87%
6	2.43	24.57%	4.59	45.91%
9	2.44	23.34%	4.41	44.09%
12	2.45	22.86%	4.28	42.76%
...				
60	2.37	18.43%	5.02	33.83%

Table II
Portfolios formed on linked-stock signal (baseline specification)

This table presents monthly abnormal returns in month $t+1$ of portfolios sorted based on linked-stock signals calculated in month t. Linked stocks are identified over the trailing window that starts three months and ends five trading days before the start of month $t+1$. Return leaders are assumed to be the stocks whose turnover in month t is above their median turnover over months $t-11$ to t. Leaders' returns over month t, excluding the last five trading days, are equal-weighted to form the linked-stock signal for their followers (linked stocks whose turnover in month t is below their 75th percentile turnover over the months $t-11$ to t). All followers with linked-stock signals are sorted into quintile portfolios within each of the 36 industries. Portfolio returns are equal-weighted in Panel A and value-weighted in Panel B. The second column contains the average linked-stock signal. The third column shows the average number of linked stock leaders for the followers in that portfolio. The fourth column reports the average portfolio return in excess of the risk-free rate; the fifth column reports the market alpha; the sixth column reports the alpha of the Fama and French (1993) three-factor model; the seventh column reports the alpha of the four-factor model that also includes the Carhart (1997) momentum factor; and the eighth column reports the alpha of the five-factor model that in addition includes the liquidity factor of Pástor and Stambaugh (2003). The last row shows the return differential between the highest- and lowest-signal portfolios. Newey-West-adjusted t-statistics are reported in parentheses.

Panel A: Equal-weighted portfolios

Quintile	Linked-stock signal	Number of linked stocks	Excess return	Market alpha	3-factor alpha	4-factor alpha	5-factor alpha
1	-12.93%	1.45	0.13%	-0.42%	-0.53%	-0.45%	-0.46%
			(0.27)	(-2.31)	(-3.21)	(-2.76)	(-2.74)
2	-4.16%	1.95	0.46%	-0.06%	-0.20%	-0.11%	-0.11%
			(1.01)	(-0.44)	(-1.71)	(-0.98)	(-0.95)
3	0.36%	2.21	0.48%	-0.03%	-0.15%	-0.08%	-0.09%
			(1.07)	(-0.18)	(-1.18)	(-0.65)	(-0.67)
4	4.97%	2.04	0.59%	0.03%	-0.07%	-0.03%	-0.02%
			(1.31)	(0.22)	(-0.47)	(-0.22)	(-0.15)
5	16.29%	1.50	1.01%	0.45%	0.31%	0.35%	0.36%
			(2.14)	(2.42)	(2.26)	(2.50)	(2.51)
5-1			0.88%	0.87%	0.84%	0.80%	0.82%
			(4.65)	(4.57)	(4.06)	(3.87)	(3.82)

Panel B: Value-weighted portfolios

Quintile	Linked-stock signal	Number of linked stocks	Excess return	Market alpha	3-factor alpha	4-factor alpha	5-factor alpha
1	-11.11%	2.62	0.03%	-0.43%	-0.42%	-0.42%	-0.42%
			(0.06)	(-2.07)	(-2.02)	(-1.99)	(-2.00)
2	-3.38%	4.59	0.49%	0.02%	-0.01%	0.01%	-0.01%
			(1.16)	(0.12)	(-0.04)	(0.06)	(-0.07)
3	0.46%	5.54	0.54%	0.06%	0.14%	0.14%	0.11%
			(1.28)	(0.38)	(0.88)	(0.86)	(0.71)
4	4.49%	4.96	0.43%	-0.04%	0.03%	0.01%	0.02%
			(1.05)	(-0.21)	(0.21)	(0.07)	(0.11)
5	13.86%	2.97	0.87%	0.39%	0.40%	0.37%	0.36%
			(1.95)	(2.25)	(2.30)	(2.22)	(2.19)
5-1			0.84%	0.82%	0.81%	0.79%	0.78%
			(3.57)	(3.53)	(3.41)	(3.31)	(3.27)

Panel C: Subperiod analysis

	July 1996–December 2003					January 2004–January 2013			
	EW portfolios		VW portfolios			EW portfolios		VW portfolios	
Quintile	Excess return	5-factor alpha	Excess return	5-factor alpha	Quintile	Excess return	5-factor alpha	Excess return	5-factor alpha
1	-0.13%	-0.74%	-0.12%	-0.54%	1	0.29%	-0.30%	0.13%	-0.32%
	(-0.19)	(-2.37)	(-0.16)	(-1.30)		(0.42)	(-2.32)	(0.20)	(-1.69)
...					...				
5	1.40%	0.71%	1.33%	0.65%	5	0.66%	0.05%	0.49%	0.05%
	(2.15)	(2.51)	(2.18)	(2.27)		(0.98)	(0.41)	(0.76)	(0.27)
5-1	1.52%	1.45%	1.44%	1.19%	5-1	0.37%	0.35%	0.36%	0.37%
	(4.94)	(3.55)	(3.57)	(2.25)		(2.66)	(2.37)	(1.66)	(1.72)

Table III
Alternative specifications and robustness checks

This table presents monthly abnormal returns of portfolios sorted based on linked-stock signals. In the baseline specification, the identification window is three months, signals from linked leaders are equal-weighted, and portfolios are formed within 36 industries. Variations to this baseline specification are described in each panel heading. Each panel reports excess returns and five-factor alphas for equal- and value-weighted portfolios, as well as the return differentials between the highest- and lowest-signal portfolios in the last row. Newey-West-adjusted t-statistics are reported in parentheses.

Alternative signal weighting schemes

Panel A: Linked-stock returns in equation (1) are weighted by the number of co-mentions

	EW portfolios		VW portfolios	
Quintile	Excess return	5-factor alpha	Excess return	5-factor alpha
1	0.15%	-0.41%	0.27%	-0.16%
	(0.32)	(-2.64)	(0.54)	(-0.83)
...				
5	0.97%	0.33%	0.81%	0.27%
	(2.08)	(2.37)	(1.71)	(1.27)
5-1	0.82%	0.74%	0.53%	0.43%
	(4.55)	(3.76)	(1.96)	(1.57)

Panel B: Linked-stock returns in equation (1) are value-weighted by market capitalization in month $t-2$

	EW portfolios		VW portfolios	
Quintile	Excess return	5-factor alpha	Excess return	5-factor alpha
1	0.17%	-0.42%	0.58%	0.10%
	(0.35)	(-2.79)	(1.22)	(0.52)
...				
5	1.04%	0.39%	1.05%	0.48%
	(2.24)	(2.91)	(2.35)	(2.18)
5-1	0.87%	0.81%	0.48%	0.38%
	(4.99)	(4.54)	(2.13)	(1.80)

Relax restrictions on linked stocks

Panel C: Do not remove competitors

	EW portfolios		VW portfolios	
Quintile	Excess return	5-factor alpha	Excess return	5-factor alpha
1	0.15%	-0.44%	0.34%	-0.10%
	(0.32)	(-2.78)	(0.72)	(-0.51)
...				
5	0.97%	0.31%	0.88%	0.41%
	(2.03)	(2.11)	(1.97)	(2.22)
5-1	0.82%	0.76%	0.54%	0.52%
	(3.91)	(3.35)	(2.05)	(1.76)

Panel D: Do not impose turnover conditions on leaders and followers

	EW portfolios		VW portfolios	
Quintile	Excess return	5-factor alpha	Excess return	5-factor alpha
1	0.27%	-0.61%	0.68%	-0.00%
	(0.50)	(-4.14)	(1.41)	(-0.02)
...				
5	0.97%	0.21%	1.00%	0.45%
	(1.84)	(1.66)	(1.84)	(2.41)
5-1	0.70%	0.82%	0.33%	0.46%
	(3.36)	(4.27)	(1.09)	(1.73)

Include news stories that co-mention more than 2 firms

Panel E: Linkages are computed from news stories that co-mention up to 5 firms

Quintile	EW portfolios Excess return	EW portfolios 5-factor alpha	VW portfolios Excess return	VW portfolios 5-factor alpha
1	0.38%	-0.22%	0.41%	-0.16%
	(0.81)	(-1.47)	(0.92)	(-0.93)
...				
5	0.80%	0.12%	0.90%	0.39%
	(1.70)	(0.96)	(2.30)	(3.00)
5-1	0.43%	0.34%	0.49%	0.55%
	(1.98)	(1.69)	(2.80)	(2.81)

Panel F: Linkages are computed from news stories that co-mention up to 10 firms

Quintile	EW portfolios Excess return	EW portfolios 5-factor alpha	VW portfolios Excess return	VW portfolios 5-factor alpha
1	0.39%	-0.22%	0.34%	-0.21%
	(0.86)	(-1.38)	(0.80)	(-1.43)
...				
5	0.77%	0.10%	0.91%	0.40%
	(1.62)	(0.75)	(2.40)	(2.65)
5-1	0.38%	0.32%	0.57%	0.60%
	(1.79)	(1.50)	(3.09)	(2.96)

Increase the length of the trailing window to identify linked stocks

Panel G: Use a six-month identification window

Quintile	EW portfolios Excess return	EW portfolios 5-factor alpha	VW portfolios Excess return	VW portfolios 5-factor alpha
1	0.25%	-0.34%	0.24%	-0.23%
	(0.53)	(-2.42)	(0.55)	(-1.37)
...				
5	1.07%	0.46%	1.06%	0.56%
	(2.36)	(3.05)	(2.42)	(3.25)
5-1	0.82%	0.80%	0.82%	0.79%
	(4.53)	(4.09)	(4.04)	(4.04)

Panel H: Use a 12-month identification window

Quintile	EW portfolios Excess return	EW portfolios 5-factor alpha	VW portfolios Excess return	VW portfolios 5-factor alpha
1	0.38%	-0.25%	0.46%	-0.08%
	(0.78)	(-2.10)	(1.09)	(-0.57)
...				
5	0.91%	0.29%	1.03%	0.51%
	(2.02)	(2.77)	(2.39)	(2.89)
5-1	0.53%	0.53%	0.57%	0.60%
	(3.67)	(3.70)	(2.62)	(2.88)

Signals are computed exclusively from linked stocks in a different industry

Panel L: three-month identification window used

	EW portfolios		VW portfolios	
Quintile	Excess return	5-factor alpha	Excess return	5-factor alpha
1	0.21%	-0.33%	0.55%	0.09%
	(0.39)	(-1.79)	(1.04)	(0.31)
...				
5	0.71%	0.08%	0.69%	0.26%
	(1.49)	(0.52)	(1.29)	(1.18)
5-1	0.50%	0.41%	0.14%	0.17%
	(2.37)	(1.84)	(0.46)	(0.53)

Panel M: six-month identification window used

	EW portfolios		VW portfolios	
Quintile	Excess return	5-factor alpha	Excess return	5-factor alpha
1	0.36%	-0.24%	0.65%	0.14%
	(0.71)	(-1.26)	(1.33)	(0.50)
...				
5	0.95%	0.39%	0.98%	0.52%
	(2.07)	(2.12)	(2.11)	(2.31)
5-1	0.59%	0.63%	0.33%	0.38%
	(2.69)	(2.56)	(1.26)	(1.35)

Signals are computed exclusively from smaller linked stocks

Panel L: three-month identification window used

	EW portfolios		VW portfolios	
Quintile	Excess return	5-factor alpha	Excess return	5-factor alpha
1	0.38%	-0.26%	0.52%	0.06%
	(0.76)	(-1.13)	(1.03)	(0.20)
...				
5	1.04%	0.40%	1.08%	0.59%
	(2.19)	(2.13)	(2.24)	(2.22)
5-1	0.66%	0.66%	0.56%	0.53%
	(2.47)	(2.30)	(1.88)	(1.56)

Panel L: six-month identification window used

	EW portfolios		VW portfolios	
Quintile	Excess return	5-factor alpha	Excess return	5-factor alpha
1	0.41%	-0.25%	0.40%	-0.17%
	(0.89)	(-1.42)	(0.91)	(-0.72)
...				
5	0.94%	0.38%	0.91%	0.38%
	(2.05)	(2.45)	(2.00)	(2.08)
5-1	0.52%	0.63%	0.51%	0.55%
	(2.31)	(2.61)	(2.16)	(1.94)

Table IV

Conditioning on the content of common news stories, excluding firm pairs with customer-supplier ties

Common news stories are assigned to news topics (the methodology for assigning news topics is described in the text and in Appendix A1). Next, all firm pairs that had at least one common news story in the "Supply chain" topic in the trailing three-year window are excluded from the analysis. All common news stories that could not be assigned to a topic are also dropped from the analysis. A six-month identification window is used. Panel A includes all firm pairs linked by common news that could be assigned one of 14 news topics that remain after removing the "Supply chain" topic. Panel B includes only firm pairs with at least one common news story assigned to the topic "Legal" over the identification window; of these, firm pairs with at least one common news story assigned to topics "Partnerships," "M&A," or "Parent/subsidiary relations" over the identification window are dropped. Panel C contains firms pairs linked by news stories classified in the remaining 13 categories, (i.e., firm pairs that are in Panel A but not in Panel B). Using only linked firm pairs within each subsample, we proceed to form portfolios as described in the baseline specification: turnover conditions are imposed to identify leaders and followers, leader returns are equal-weighted, and signal-based quintile portfolios are formed within 36 industries. Each panel reports excess returns and five-factor alphas for equal- and value-weighted portfolios, as well as the return differentials between the highest- and lowest-signal portfolios in the last row. Newey-West-adjusted t-statistics are reported in parentheses.

Panel A: All defined news topics except for customer-supplier links

	EW portfolios		VW portfolios	
Quintile	Excess return	5-factor alpha	Excess return	5-factor alpha
1	0.22%	-0.41%	0.06%	-0.48%
	(0.44)	(-2.16)	(0.13)	(-2.31)
...				
5	0.70%	0.08%	0.53%	0.03%
	(1.52)	(0.62)	(1.23)	(0.15)
5-1	0.48%	0.49%	0.48%	0.51%
	(2.54)	(2.39)	(2.01)	(1.87)

Panel B: Only similar legal issues

EW portfolios		VW portfolios	
Excess return	5-factor alpha	Excess return	5-factor alpha
0.14%	-0.33%	0.73%	0.48%
(0.25)	(-0.97)	(1.31)	(0.97)
...			
1.57%	1.15%	1.08%	0.69%
(2.57)	(2.53)	(2.32)	(2.21)
1.43%	1.47%	0.35%	0.21%
(2.77)	(2.82)	(0.66)	(0.36)

Panel C: Only similar operations and shared goals

EW portfolios		VW portfolios	
Excess return	5-factor alpha	Excess return	5-factor alpha
0.28%	-0.35%	0.14%	-0.42%
(0.54)	(-1.92)	(0.31)	(-1.78)
...			
0.68%	0.07%	0.62%	0.20%
(1.48)	(0.53)	(1.43)	(0.91)
0.40%	0.42%	0.48%	0.62%
(2.04)	(1.95)	(1.72)	(2.02)

Table V
Cross-sectional regressions

This table presents the results of Fama and MacBeth (1973) regressions of stock returns in month $t+1$ on a set of explanatory variables available as of the end of month t. Unless stated otherwise, we follow the baseline identification method of Table II to compute the linked-stock signal. Control variables include log(mktcap), book/market, and the average monthly stock return over months $t-7$ to $t-2$, as well as, in the specifications with interaction terms, the dummy variable included in the interaction term. All explanatory variables are described in Appendix A2. The median cross-sectional value of each variable is denoted as *med*. Newey-West-adjusted t-statistics are reported in parentheses. All coefficients are multiplied by 100. The sample period is July 1996–January 2013.

Specification	(1)	(2)	(3)	(4)	(5)[¶]	(6)[§]	(7)	(8)	(9)
Linked-Stock Signal	2.15^a	2.31^a	1.79^a	2.49^a	1.63^a	1.77^a	3.56^a	2.66^a	2.98^a
	(3.89)	(4.52)	(3.30)	(4.56)	(3.19)	(3.75)	(3.79)	(3.08)	(3.81)
Rev		-2.61^b	-2.43^b	-2.67^b	-1.91^a	-3.17^a	-2.47^b	-2.57^b	-2.63^b
		(-2.53)	(-2.14)	(-2.47)	(-2.65)	(-3.60)	(-2.24)	(-2.47)	(-2.52)
Industry Ret		4.83^c	6.39^b	3.91	5.93^a	4.31^c	5.08^c	4.54^c	4.91^c
		(1.85)	(2.54)	(1.46)	(2.62)	(1.85)	(1.93)	(1.71)	(1.87)
$\mathbb{1}\{Inst.Ownership > med\}\times$							-2.47^b		
Linked-Stock Signal							(-2.24)		
$\mathbb{1}\{AnalystCoverage > med\}\times$								-1.03	
Linked-Stock Signal								(-0.85)	
$\mathbb{1}\{Size > med\}\times$									-1.67^c
Linked-Stock Signal									(-1.75)
Controls	Yes	Yes	No	Yes[†]	Yes	Yes	Yes	Yes	Yes

^{a}, ^{b}, and ^{c} indicate significance at the 1%, 5%, and 10% levels, respectively.

[†]In addition to other controls we include include the monthly-frequency leader signal from Scherbina and Schlusche (2013).

[¶]No turnover conditions are imposed on the followers.

[§]A six-month identification window is used, resulting in the October 1996–January 2013 sample period.

Table VI

Portfolios constructed based on signals from second-tier links

This table presents monthly abnormal returns of portfolios sorted based on the signals of stocks linked exclusively by second-tier links (or links of links). Stock pairs linked through second-tier links are not linked directly over the identification period, but share at least one common direct link (as stocks A and E in Figure 1). Once second-tier linked stock pairs are identified, predictive signals are calculated and portfolios are formed as in the baseline specification: turnover conditions are imposed to identify leaders and followers, leader returns are equal-weighted, and signal-based quintile portfolios are formed within 36 industries. Each panel reports excess returns and five-factor alphas for equal- and value-weighted portfolios, as well as the return differentials between the highest- and lowest-signal portfolios in the last row. Newey-West-adjusted t-statistics are reported in parentheses.

Panel A: Three-month identification window

	EW portfolios		VW portfolios	
Quintile	Excess return	5-factor alpha	Excess return	5-factor alpha
1	0.28%	-0.34%	0.41%	-0.15%
	(0.61)	(-2.75)	(0.94)	(-0.84)
...				
5	0.66%	0.03%	0.69%	0.23%
	(1.37)	(0.25)	(1.57)	(1.30)
5-1	0.38%	0.36%	0.29%	0.38%
	(2.48)	(2.29)	(1.10)	(1.69)

Panel B: Six-month identification window

	EW portfolios		VW portfolios	
Quintile	Excess return	5-factor alpha	Excess return	5-factor alpha
1	0.46%	-0.18%	0.42%	-0.19%
	(1.01)	(-1.71)	(0.98)	(-1.28)
...				
5	0.80%	0.15%	0.84%	0.32%
	(1.70)	(1.20)	(2.10)	(1.76)
5-1	0.33%	0.33%	0.43%	0.51%
	(2.15)	(2.10)	(1.79)	(2.24)

Panel C: 12-month identification window

	EW portfolios		VW portfolios	
Quintile	Excess return	5-factor alpha	Excess return	5-factor alpha
1	0.50%	-0.14%	0.59%	-0.02%
	(1.07)	(-1.43)	(1.37)	(-0.15)
...				
5	0.89%	0.24%	0.58%	0.06%
	(1.96)	(1.92)	(1.51)	(0.45)
5-1	0.39%	0.39%	-0.01%	0.08%
	(2.46)	(2.28)	(-0.05)	(0.42)

Table A1
Industries

This table presents the monthly average distribution of stocks across industries in our sample. The sample consists of common shares of U.S.-incorporated firms (stocks with share codes 10 or 11) that were priced above $5 per share and had at least one linked stock that we consider to be a non-competitor over a trailing three-month identification window. The sample period is April 1996–December 2012.

Stone, Clay and Glass Products	0.17%
Agriculture, Forestry, and Fishing	0.20%
Textile Mill Products	0.22%
Sanitary Services	0.27%
Lumber and Wood Products	0.28%
Tobacco Products	0.33%
Leather and Leader Products	0.34%
Furniture and Fixtures	0.45%
Glass	0.48%
Apparel and other Textile Products	0.60%
Paper and Allied Products	0.71%
Miscellaneous Manufacturing Industries	0.88%
Rubber and Miscellaneous Plastics Products	0.99%
Mining	1.00%
Construction	1.04%
Petroleum and Coal Products	1.11%
Fabricated Metal Products	1.25%
Public Administration	1.28%
Primary Metal Industries	1.42%
Printing and Publishing	1.92%
Telephone and Telegraph Communication	1.93%
Radio and Television Broadcasting	2.16%
Transportation Equipment	2.56%
Food and Kindred Products	2.96%
Wholesale	3.22%
Transportation	3.22%
Oil and Gas Extraction	3.73%
Instruments and Related Products	4.27%
Electric, Gas, and Water Supply	5.43%
Machinery, Except Electrical	5.55%
Retail Stores	7.07%
Electrical and Electronic Equipment	7.17%
Chemicals and Allied Products	8.34%
Finance, Insurance, and Real Estate	13.08%
Services	15.53%

Table A2
Factor loadings for baseline linked-stock signal portfolios

This table presents the five-factor model factor loadings for the baseline linked-stock portfolios of Table II. Panels A and B report the factor loadings for equal- and value-weighted portfolios, respectively, along with their alphas and the corresponding R^2s. Newey-West-adjusted t-statistics are reported in parentheses.

Panel A: Factor loadings for equal-weighted portfolios

Decile	alpha	MKT	SMB	HML	UMD	LIQ	R^2
1	-0.46%	1.07	0.46	0.08	-0.11	-0.01	86.54%
	(-2.74)	(23.09)	(4.36)	(0.97)	(-1.67)	(-0.27)	
2	-0.11%	1.05	0.34	0.20	-0.13	0.01	91.88%
	(-0.95)	(27.26)	(10.39)	(4.86)	(-6.04)	(0.57)	
3	-0.09%	1.01	0.34	0.17	-0.10	-0.01	89.93%
	(-0.67)	(23.44)	(9.12)	(3.50)	(-3.29)	(-0.55)	
4	-0.02%	1.12	0.31	0.13	-0.05	0.03	89.09%
	(-0.15)	(38.44)	(6.93)	(2.87)	(-1.45)	(1.01)	
5	0.36%	1.09	0.49	0.17	-0.06	0.03	89.92%
	(2.51)	(32.52)	(10.06)	(3.25)	(-1.35)	(1.38)	
5-1	0.82%	0.02	0.03	0.09	0.05	0.04	-0.00%
	(3.82)	(0.34)	(0.20)	(0.79)	(0.56)	(1.18)	

Panel B: Factor loadings for value-weighted portfolios

Decile	alpha	MKT	SMB	HML	UMD	LIQ	R^2
1	-0.42%	0.98	0.02	-0.05	0.00	-0.02	68.61%
	(-2.00)	(19.42)	(0.26)	(-0.59)	(0.05)	(-0.35)	
2	-0.01%	1.05	-0.05	0.10	-0.03	-0.08	74.46%
	(-0.07)	(14.46)	(-0.54)	(1.26)	(-0.79)	(-2.02)	
3	0.11%	1.06	-0.20	-0.12	0.01	-0.07	82.43%
	(0.71)	(24.24)	(-3.94)	(-2.03)	(0.26)	(-2.06)	
4	0.02%	1.04	-0.26	-0.07	0.03	0.02	78.76%
	(0.11)	(23.02)	(-5.08)	(-1.41)	(0.63)	(0.56)	
5	0.36%	1.06	-0.03	0.00	0.04	-0.04	77.07%
	(2.19)	(24.58)	(-0.39)	(0.02)	(0.64)	(-1.47)	
5-1	0.78%	0.08	-0.04	0.05	0.03	-0.03	-1.57%
	(3.27)	(1.24)	(-0.52)	(0.49)	(0.64)	(-0.57)	

Table A3
Key words and topic codes for the classification of headlines

This table lists key words and the TRNA topic codes used to classify news into topics. Variations of word combinations are presented in parentheses. In the text, the words making up key word combinations may be separated by other words. The set of keywords used includes variations in spelling, ending, and hyphenation of the keywords listed in the table.

Supply chain. Key words: customer, supplier, (agreement) to supply (sell to, buy from), supply to, buy from, supplier (customer) agreement (pact), multi-year agreement. TRNA topic codes: DEAL, DEAL1
Partnerships. Key words: partner, joint, consortium, collaboration, contract, alliance, distribution, outsource, licensing, common project, team up, join forces, distribution (licensing, contractual, exclusive, joint, mutual, temporary) agreement, sign (make, finalize, extend, strike, reach, craft) deal (agreement, contract, partnership, collaboration), agree on (work out) a deal, joint project (operations, venture), form (create, start) venture (collaboration, joint project), strategic deal (agreement), share profit (revenue), profit (revenue) sharing. TRNA topic codes: ALLCE.
M&A. Key words: M&A, merge, merger, acquisition, acquire, takeover, consolidation, combine, to purchase, to buy, bid on. TRNA topic codes: MGR.
Parent/subsidiary relations. Key words: subsidiary, corporate parent, parent company, carve-out, spin-off, split-off, divestiture. TRNA topic codes: DVST.
Legal. Key words: scandal, patent, ruling, lawsuit, settlement, antitrust, probe, judge, infringe, fraud, bankruptcy, privatization, law, recall, arbitration, verdict, cartel, audit, misstatement, restatement, defect, copyright, compliance, fiduciary, probe, investigate, subpoena, court, trial, smuggle, justice, attorney, acquit, affidavit, justice, allegation, arrest, assault, bail, bailiff, bankruptcy, circumstantial evidence, crime, complainant, confess, constitution, contract, continuance, counsel, court, crime, cross-examination, custody, damages, decree, defendant, defense, deposition, disbarment, docket, due process, entrapment, escrow, ethics, evidence, examination, exonerate, expunge, felony, jury, grievance, guilty, habeas, corpus, hearing, hearsay, immunity, incarceration, incompetent, indictment, infraction, injunction, innocent, jail, judge, judiciary, jurisdiction, jurisprudence, justice, larceny, lawsuit, lawyer, legal, legislation, leniency, liable, lien, litigation, manslaughter, marshal, mediation, misdemeanor, mistrial, murder, negligence, oath, objection, ordinance, overrule, paralegal, pardon, parole, perjury, petition, plaintiff, plea, precedent, probable cause, hearing, prison, probate, probation, prosecute, redress, rejoinder resolution, search warrant, sentence, sequester, settlement, sheriff, statute, subpoena, judgment, suit, suppress, testimony, theft, tort, transcript, trial, trustee, usury, verdict, voir dire, waiver, witness, zoning, off-shore tax, SEC, DOJ. TRNA topic codes: CASE1, CLASS, MNGISS, MONOP, JUDIC, LAW, ACB, HRGT, SCAM1, FAKE1, JUDIC, FRAUD1, REGS, CRIM, BRIB, DAT, CIV, CLJ.
Regulation. Key words: regulation, senate, Obama, Clinton, Bush, Yellen, Bernanke, justice, CDC, GAAP, environmental, parliament, democrats, republicans, cabinet, treaty, commissioner, Rotterdam, referendum, FDIC, Pentagon, Homeland, House of representatives, government regulation, U.S. House (Senate), House (Senate) bill, House (Senate) majority (minority). TRNA topic codes: POL, JOB, WASH, USDA, DEFOR, CEN, GFIN, HEA, WOM, CO2, AWLQ, PLCY, ENV, MCE, SDS, GFIN, FED, HREP, SEN, G20, G8, G7, MEVN.
Labor, production, and infrastructure. Key words: worker salary (pay, benefits), pay cut, layoff, wage increase, labor union, union negotiations, job (positions) cut (reduced, eliminated), strike, plant (factory, store, facilities, operations) closed (shut down, eliminated, opened), production, product (child, consumer) safety, product defect (recall), USDA, FDA, subsidy, stimulus, manufacturing, development, infrastructure, exit market, downsize, revamp, expand, expansion plan, innovate, restructure, new product (service, device), launch product (service, delivery); exclude "wage war." TRNA topic codes: RTM, WPAY, STAT, BKRT.
Executive compensation and corporate governance. Key words: executive (CEO) salary (bonus, pay, benefits, compensation, contract), corporate governance, governance issue (problem, failure), weak (poor) governance, poorly governed; exclude "governance services," "corporate governance firms."
Management developments. Key words: appoint, resign, demote, retire, election, management change (turnover), board turnover, change in management. TRNA topic codes: BOSS1.
Common customer. Key words: missile, military, prison, defense, olympics, bombardier, US Marines, US Navy, Air Force, Homeland, DHS, contract with, sell (supply) services to (customer agreement with) government (prison, jail). TRNA topic codes: DEF, DEFBUY.
Cross-investments. Key words: lease, loan, lend, financing, cross-financing, cross-holdings, credit facility, subsidiary, to fund (provide) funding (credit, capital), credit line, to purchase stake, line of credit, secure credit, infuse (invest) capital (equity, cash, money); exclude "finance chief (head)." TRNA topic codes: LOA, SFIN, STK.
Natural resources. Key words: names or metals and minerals used in production and manufacturing. TRNA topic codes: MIN, AGRI, ALU, AMCRU, ASCRU, ATMY, AUSCRU, BASMTL, BRGE, BRLY, BSMH, BUN, CANCRU, CBLT, CHR, CHS, CO2, COA, COC, COCOIL, COF, MIN, COFARA, COFROB, CONT, COR, COROIL, COT, CPPR, CRU, DAIR, DBULK, DIAM, DISTLL, EMACRU, FERR, FERT, GOL, GRA, H2O, HOIL, INDI, IRDM, IRN, JET, LATCRU, LEAD1, LITH, LIV, LNG, LPG1, MEAL MECRU, METL, MGS, MGSM, MINMTL, MLDM, MLK, MOG, NAP, NASCRU, NATU, NGL, NGS, NIOB, NKL, NRG, NSCRU, NSEA, NUC, OILS, OLVOIL, ORJ, PALL, PETC, PGM, PHOS, PLAS, PLAT, PNTOIL, POIL, POTH, PRCP, PREMTL, PROD PWR, RAPOIL, RAREE, RFO, RHDM, RHEN, RICE1, RLFT, RNW, RTNM, RUB, RUSCRU, SCRP, SEACRU, SFTS, SHFV, SLCN, SLK, SLVR, SNFOIL, SOIL, SORG, SOY1, SSTE, STE, SUG, TEA, TGSN, TIN1, TMBR, TMP, TNKR, TNTE, TTNM, URAN, USCRU, VNDM, WEA, WHT, WINE1, WND, WOO, ZNC, TWAVE, DFTS, QUAK.
Energy. Key words: pipeline, deepwater, sunpower, electricity, energy, gasoline, exploration, refinery, oil, gas, hydroelectric, solar, biofuel, ethanol. TRNA topic codes: DRIL, TRNSPT, OILG, OILI, EXPL, ENER, AFRCRU, BIOCEL, BIODSL, BIOETH, BIOF, BIOMS.
Technology. TRNA topic codes: LSCI, WWW, NSS, SPAC, GMO, ITEC, SCI.
Geopolitical. Key words: quake, tsunami, pandemic, forex, militant, province, cossack, Putin, Kremlin, unrest; names of world currencies; names of continents and large non-US geographical regions (e.g., Mediterranean), names of large islands and island chains, names of all world countries and their capitals, names of large non-US cities. TRNA topic codes: VIO, WAR, PIA, as well as topic codes for all countries and geographical regions.

Table A4
Portfolio transition probabilities

This table reports portfolio transition probabilities between months, t and $t+1$ and month t and $t+2$ for the baseline portfolios of Table II, calculated only if the stock exists in the CRSP universe at both points in time.

Portfolio in	Portfolio in month $t+1$					
month t	1	2	3	4	5	unassigned
1 (low signal)	0.14	0.13	0.11	0.12	0.13	0.38
2	0.10	0.15	0.15	0.15	0.11	0.33
3	0.09	0.15	0.17	0.16	0.10	0.33
4	0.09	0.15	0.16	0.15	0.10	0.34
5 (high signal)	0.11	0.12	0.13	0.13	0.12	0.38

Portfolio in	Portfolio in month $t+2$					
month t	1	2	3	4	5	unassigned
1 (low signal)	0.09	0.11	0.10	0.10	0.10	0.49
2	0.09	0.12	0.13	0.12	0.09	0.44
3	0.08	0.13	0.15	0.13	0.09	0.41
4	0.08	0.12	0.14	0.13	0.09	0.43
5 (high signal)	0.09	0.11	0.10	0.11	0.10	0.49