How News and Its Context Drive Risk and Returns Around the World ☆

Charles W. Calomiris, *[,a,b] Harry Mamaysky [c]

[a] Columbia Business School, 3022 Broadway, Uris Hall 801, New York, NY 10027, United States

[b] National Bureau of Economic Research, 1050 Massachusetts Avenue, Cambridge, MA 02138, United States

[c] Columbia Business School, 3022 Broadway, Uris Hall 803, New York, NY 10027, United States

Abstract

We develop a classification methodology for the context and content of news articles to predict risk and return in stock markets in 51 developed and emerging economies. A parsimonious summary of news, including topic-specific sentiment, frequency, and unusualness (entropy) of word flow, predicts future country-level returns, volatilities, and drawdowns. Economic and statistical significance are high and larger for year ahead than monthly predictions. The effect of news measures on market outcomes differs by country type and over time. News stories about emerging markets contain more incremental information. Out-of-sample testing confirms the economic value of our approach for forecasting country-level market outcomes.

JEL classification: G12, G15, G17

Keywords: empirical asset pricing, international markets, financial news media, natural language processing

* Corresponding author.
Email addresses: cc374@columbia.edu (C. Calomiris), hm2646@columbia.edu (H. Mamaysky).

☆ We gratefully acknowledge support from the Program for Financial Studies and the Bank of England, and excellent research assistance from Yong Wang, Minchen Zheng, and Sirui Wang. We thank the Thomson Reuters Corp. for graciously providing the data that was used in this study. For helpful comments we thank the referee, Dmitry Livdan, Kent Daniel, Robert Hodrick, Leif-Anders Thorsrud, Diego Garcia, Sanjiv Das, and seminar participants at Catholic University of Chile, the 2016 RIDGE/ Banco Central del Uruguay Workshop on Financial Stability, the AlphaSimplex Group, the 2017 News & Finance Conference at Columbia, Chapman University, Villanova University, Ohio State University, the Federal Reserve Bank of Kansas City, Arizona State University, University of Arizona, the Global Risk Institute, the University of Colorado, the Columbia Machine Learning in Finance Workshop, the 2017 Society for Economic Measurement Conference, Cornell University, and the 2017 Cleveland Fed and University of Maryland Financial Stability & FinTech Conference.

1. Introduction

What is news and how is it associated with changes in stock market returns and risks? This is a fundamental question in asset pricing and has been the subject of decades of research (for example, Fama et al., 1969; Roll, 1984). Recently, financial economists have brought new tools to bear on this question, including the analysis of the relationships between market outcomes for individual stocks or US stock market indexes and various aspects of the words that appear in newspaper articles and other textual sources (for example, Tetlock 2007, 2010, 2011; Tetlock, Saar-Tsechansky, and Macskassy, 2008; Garcia 2013). The promising early work in the literature linking textual analysis and stock returns has raised more questions that it has answered. This paper addresses nine important sets of questions about the connections between news and market outcomes.

First, and perhaps most importantly, how should one best measure news using word flow? One approach, which we adopt here, is to apply atheoretical methods (i.e., those with no a priori position regarding which particular words should be the focus of the analysis) to organize the flow of words in a comprehensive and unconstrained manner to see which parts of word flow matter for market outcomes. An alternative approach (for example, Baker, Bloom, and Davis, 2016) is to identify, based on a priori criteria, key lists of words or combinations of key words to see how their presence matters for market outcomes. A major advantage of the former approach is that it does not require researchers to know in advance what aspects of word flow are most relevant. The atheoretical approach also avoids data mining risks by imposing discipline on the process by which text is analyzed.

Second, which aspects of word flow should be the focus of measurement? There is a large literature showing that "sentiment" has explanatory power for returns. Articles that contain words with preidentified positive sentiment value (as measured by a sentiment "dictionary") are associated with positive returns, while those with negative value are associated with negative returns. But sentiment is only one dimension of word flow. The frequency of the appearance of certain words or phrases (compared to their past frequency) may also be relevant, and it may also be that the context in which words appear (which we label "topics") is important to the meaning of word flow. In addition to measuring sentiment, the contextual frequency of word flow, and the way sentiment matters differently depending on context, other aspects of text flow may be relevant. Glasserman and Mamaysky (2018a) show that the unusualness (entropy) of word strings may have predictive power for market outcomes, especially when interacted with sentiment. As we show below, the effects of measured sentiment and frequency do vary across topic categories, so this decomposition of sentiment may be particularly useful in forecasting applications. Our empirical approach will include these various measures of text flow and their interactions and explore their incremental information content relative to nontextual variables often included in asset pricing studies.

Third, the patterns that link frequency, topics, sentiment, and entropy measures of word flow with market outcomes may vary over time. In this paper, we capture changes over time using a dividing point that is identified by principal components analysis. We show that the mapping from word flow measures to market outcomes changed somewhat after the global financial crisis. We present results for the entire sample period (1998-2015) and for two sub-periods (April 1998-February 2007, and March 2007-December 2015). We further explore dynamic changes in coefficients using a rolling elastic net regression–which allows for model selection and coefficient shrinkage–for out-of-sample forecasting. We find that allowing coefficients to change over time is important for out-of-sample forecasting.

Fourth, given the potential importance of identifying topical context, how should one identify topics? Within the set of atheoretical means of identifying topics there are two common methods, namely the Louvain (Blondel et al., 2008) and latent Dirichlet allocation (LDA, see Blei, Ng, and Jordan, 2003) approaches, as we discuss below. The Louvain method assigns salient words to mutually exclusive topic areas based on word co-occurrence (that is, each word belongs to only one topic area). The LDA method allows words to appear in more than one topic area (on a probabilistic basis). After verifying with some exploratory analysis that our regression findings are similar under either approach, we focus on the Louvain method. The Louvain approach, variants of which have been applied in diverse fields from sociology (Rule, Cointet, and Bearman, 2015) to marketing (Netzer et al., 2012), has the major advantages of much faster computational speed, which results from the simplicity of a mutually exclusive approach to assigning words to topic areas as well as greater ease of interpretability.[1]

Fifth, does the effect of our word flow measures operate through a risk channel? Our findings suggest that when a word flow measure predicts positive expected returns, it also predicts a reduction in risk. Word flow measures largely divide into "good" and "bad" news, where bad news implies lower expected returns and higher risk and good news implies the opposite. The fact that news tends to have opposite effects on expected returns and risk (e.g., expected returns higher, while risk lower) suggests that the factors captured by news flow are not priced risks (priced factors should affect expected returns and risk in the same direction).

Sixth, how should one measure risk? As is well known, if the returns process is characterized by Brownian motion and normality of the error term, then the standard deviation of returns (say, over a particular month) will be a sufficient statistic for risk. Those assumptions,

[1] Rule, Cointet, and Bearman (2015) discuss the pros and cons of various topic classification approaches and reach a conclusion similar to ours–that co-occurrence approaches are appealing due to their simplicity and the ease of interpretability of results.

however, generally are rejected, especially for emerging market (EM) countries, which exhibit pronounced momentum and nonnormality, both with respect to skew and kurtosis (see Bekaert et al., 1998; Karolyi, 2015, and Ghysels, Piazzi, and Valkanov, 2016). Given those facts, to capture risk, in addition to using the standard deviation of returns (*sigma*), we also employ the "maximum one-year drawdown." This measures, at any point in time, the maximum percentage decline that occurs from the current index value during the next year. This measure also is intended to capture the fact that "downside risk" may be treated differently from "upside risk" (the standard deviation of returns treats them as identical).

Seventh, the existing literature typically focuses on short-term analysis of individual US companies or the US stock market as a whole.[2] Do empirical patterns that apply to individual company stocks or the aggregate U.S. index also apply to other countries? When analyzing news and stock market behavior, should countries be lumped together or analyzed separately? We analyze the aggregate monthly stock market returns and risks for 52 countries.[3] There is a great deal of evidence suggesting that returns processes differ between emerging markets (EMs) and developed markets (DMs). Furthermore, the amount of risk and the nature of the news that drives risk differ between EMs and DMs (Beim and Calomiris, 2001; Karolyi, 2015), which reflects differences in political contexts, the ranges of government policy choices, differences in information production for securities markets, different market liquidity (Calomiris, Love, and Martinez-Peria, 2012), differences in legal environment and corporate governance (La Porta et al., 1998), different fiscal, monetary and exchange rate regimes (Calvo and Reinhart, 2002), differences in sovereign default risk (which is absent in most DMs but is relatively high in EMs, as described in Cruces and Trebesch, 2013), and differences in the frequency and severity of

[2] A notable exception is Froot et al. (2017), who analyze media reinforcement effects at the country index level.
[3] We only use 51 countries in our panel regressions because we exclude Iceland, which experienced a drawdown of 95%. Including this outlier affected coefficient magnitudes in our regression models, and therefore we excluded it.

banking crises (Laeven and Valencia, 2014). EMs suffer larger and more frequent major drawdowns of stock returns than DMs (Kaminsky and Schmukler, 2008). For all these reasons, we divide countries into EMs and DMs and perform separate panel analyses of each group of countries.

Eighth, what source of news should one use? Given our global interest (across EMs and DMs) we need an English language news source covering many countries. Thomson Reuters generously provided their entire database of news articles from 1996 through 2015.

Finally, over what time frame should word flow predict risk and return? Much of the existing finance literature on the effects of sentiment on individual stocks' returns have focused on high-frequency predictions. Glasserman and Mamaysky (2018a) are an exception; using the US stock market index, they find word flow predicts risk over the course of several months. Similarly, Sinha (2016) and Heston and Sinha (2017) find that it can be useful to aggregate over longer periods of time when analyzing news for individual stocks. They find that when aggregating news over a week, rather than a day, one substantially lengthens the time horizon over which news forecasts returns. Weekly news predicts returns for 13 weeks, while daily news predicts returns for only two days. Motivated by these findings, we aggregate news at a monthly horizon, examine both one-month-ahead and one-year-ahead predictions, and show that our country-level measures exhibit stronger predictive power for one-year-ahead returns and drawdowns than for one-month-ahead forecasts of return and volatility.

Section 2 describes how we derive measures of word flow used in the study and provides a list of variables and sources for them. Section 3 presents regression results. Section 4 presents out-of-sample tests of our model. Section 5 concludes by summarizing our findings.

2. Data construction, variable definitions, and summary statistics

The analysis in this paper combines three types of data–market, macro, and news–all of which are aggregated into a single data set at the month-country level. Our country-level stock market index data are obtained from Bloomberg. Table 1 shows the mapping from each of our DM and EM countries to the corresponding stock market index. All index returns are converted into US dollar terms using end-of-day exchange rates from Bloomberg. For a given country, we calculate its one-month ahead return (*return*), its one-year ahead return ($return^{12}$), its realized monthly volatility (*sigma*), and its one-year ahead drawdown (*drawdown*) using these US dollar returns. Our macro data, such as interest rates, GDP growth rates, and credit ratios, are obtained from a myriad of sources, like the World Bank and the International Monetary Fund, as detailed in the Appendix. Our textual data source is the Thomson Reuters Machine Readable News archive. This archive includes all Reuters News articles from 1996 to 2015, from which we use only the English language news. The measures of textual content we employ are constructed by us, as described below (also see the Appendix for more details).

Thomson Reuters News Analytics (TRNA) offers its own version of a sentiment measure as a commercial product, which has been used by Sinha (2016) and Heston and Sinha (2017), among others. The TRNA sentiment measure captures similar content to the sentiment measures we construct, but the TRNA sentiment measure is only available for a fraction of the articles in the Thomson Reuters database and only from 2003.[4] For this reason, we use our own sentiment

[4] In response to a comment received after we completed our analysis, we purchased the TRNA sentiment data and compared regression results for the post-2003 period using our sentiment measure with those based on the TRNA sentiment measure. A detailed comparison is provided in Online Appendix Tables A5-A9. We find that the two measures are correlated (correlation coefficients of the two approaches to measuring sentiment are generally greater than 0.3 and less than 0.4, depending on topical context), and using the TRNA measure in our regression framework yields somewhat similar findings to those reported here, but the use of our measure generally results in more precise estimation and higher *R*-squared.

measures constructed directly from the raw text of the Thomson Reuters Machine Readable News archive.

2.1. Construction of text measures

Our text processing can be broken up into four parts: (i) corpus selection and cleaning, (ii) construction of the document term matrix and topic classification, (iii) extraction of n-grams to allow for calculation of entropy, and (iv) calculation of article-level sentiment, topic, and entropy measures. Here we present a high-level overview of the process. The Appendix contains more technical and methodological detail.

In the first step, we select our text corpus and then clean and preprocess it. For the EM analysis, our corpus consists of all articles tagged by Thomson Reuters with the *N2:EMRG* code, which indicates an article about an emerging market country. Our EM corpus consists of 5mm unique articles. Our DM corpus consists of all articles about the countries identified as developed market economies in Table 1. The DM corpus consists of 12mm unique articles. All textual analysis in the paper is done separately for the EM and DM corpora.

In the second step, we calculate the document term matrix for the corpus under consideration. The document term matrix, with rows corresponding to articles and columns corresponding to words, counts the number of times a given word appears in a given article. For a given document term matrix, let us write $D_{j,w}$ for the number of times word w appears in article j. We restrict the words whose occurrences we count to our *econ word* list. This is a list of 1,242 stemmed words, bigrams, and trigrams[5] that are descriptive of either market or economic phenomena (prestemming examples include *barriers*, *currency*, *parliament*, *macroeconomist*, and *World Bank*). This list was derived as follows: we began with the index from Beim and

[5] Bigrams and trigrams (or 3-grams) are word phrases of length two and three, respectively.

Calomiris' Emerging Financial Markets textbook. We then searched for words that co-occurred frequently in our articles database with the words in that list. The list itself, as well as the classification of each word into a topic, is available from the authors.

To define topic groups, we use the document term matrix to measure the tendency of groups of words to occur in articles together–we refer to this tendency as *co-occurrence*. Information about the co-occurrence of words, as measures by cosine similarity (see Appendix), is stored in the co-occurrence matrix (a symmetric matrix with a row and column for each of our econ words). The co-occurrence matrix defines a network of our 1,242 econ words, to which we apply the Louvain community detection algorithm to find nonoverlapping clusters (i.e., a word can belong to only one cluster) of related words–we refer to these clusters as topics and label each one with what appears to us to be a natural topic title. Details of this procedure are given in the Appendix, but intuitively, we are looking for groups of words that tend to co-occur in articles more frequently than would be expected by chance. This procedure yields five topics for each of our DM and EM corpora.[6] Fig. 1 and 2 show the most frequently occurring words in each of our EM and DM topics.[7] For the EM corpus, we find five word-groupings, which we label as markets (*Mkt*), governments (*Govt*), commodities (*Comms*), corporate governance and structure (*Corp*), and macroeconomic topics (*Macro*). For the DM corpus, we find similar, but not identical topics: markets (*Mkt*), governments (*Govt*), commodities (*Comms*), corporate governance and structure (*Corp*), and the extension of credit (*Credit*).

Table 2 shows that the word overlap between the topics we identify in our EM and DM corpora is often, though not always, sizable. Our measure of word overlap is the Jaccard index,

[6] We found that recalculating topics over different subsamples of our data yielded very similar word groupings to those that were obtained over the entire sample. See the Appendix for more details.
[7] Fig. 3 shows that the original Louvain clustering produced over 40 word groupings for each corpus, yet only five of these contained more than just a few words. We place words from the smaller groupings into the five large ones for each corpus. This is discussed in greater detail in the Appendix.

which for two sets *A* and *B*, reports how many elements there are in their intersection divided by the number of elements in their union. The rows of the table correspond to DM topics, and the columns correspond to EM topics. For example, we see when we compare the *Govt* topic between our EM and DM corpora that 82% of all words common to the two topics are present in each topic separately. This indicates that the words that tend to co-occur in government-related articles in our EM and DM samples are quite similar. Similarly, the Jaccard overlap between the *Mkt* topic in our EM and DM samples is 59%. There is some overlap in the *Comms* topic as well. We also note that there is a large overlap (of 46%) between the *Corp* topic in EM and the *Credit* topic in DM. Our EM *Macro* topic has no close analog in any of the DM topics (the closest is the DM *Comms* topic)–suggesting that news about EM economies tend to focus on topics of macroeconomic interest in a way that articles about DM economies do not. Perhaps this is because macroeconomic institutions in DM economies are more settled than their EM counterparts and therefore require less news coverage.

Tables 3 (for EM) and 4 (for DM) show four sample headlines of articles classified as belonging to each of the topics we identify in our analysis, which provide some examples of how our identified topics relate to articles used in our analysis.[8] For example, in the emerging market corpus an article titled "Clinton says Putin can build strong, free Russia" is classified as being in the *Govt* topic. A Portuguese language article entitled "Sao Paulo volta a registrar inflacao no comeco de marco" is classified–seemingly correctly–in the *Macro* topic. Presumably this article was included, despite the fact that it is not in English, because the relevant stemmed Portuguese words are identical to their stemmed English counterparts. While we explicitly select only English language articles from the Thomson Reuters data set, some of their language metadata is apparently incorrect. In the developed markets corpus, most of our sample articles seem to be

[8] In these tables an article is classified into topic τ if between 80%-90% of its econ words belong to that topic.

classified correctly based on their headlines. For example “BRIEF-NQ Mobile announces termination of proposed divestment of Beijing Tianya” is in the *Corp* topic.

Thomson Reuters’s articles cover a wide range of topics. For example, sports articles are included, although they are often discussed from the perspectives of the economic or business implications of the sports-related event, which explains why sports articles have positive weights in the topic areas we identify. We consider restricting our sample of articles to those that were more narrowly focused on business, economics, and politics topics, but we find that doing so slightly reduces the explanatory power of news for stock returns and risk, and so we retain the full sample of news articles for our analysis.

The third step of our textual analysis is the extraction of n-grams. We use n-grams, or more specifically 4-grams, to construct a measure of the entropy of a given article, following closely the methodology proposed in Glasserman and Mamaysky (2018a). An n-gram is a collection of *n* contiguous words.[9] We do not allow n-grams to cross sentence boundaries–so these are n-word phrases that appear entirely in a single sentence. For a given 4-gram, we calculate the probability of observing the fourth word in the phrase conditional on seeing the first three words. This conditional probability is estimated from a training corpus as follows

$$m = \frac{\hat{c}(w_1,w_2,w_3,w_4)+1}{\hat{c}(w_1,w_2,w_3)+10} \;, \tag{1}$$

where $\hat{c}$ counts how frequently a given 4-gram or 3-gram occurred in a training corpus. Adding 1 to the numerator and 10 to the denominator is a simple way to handle cases in which the three-word phrase that begins the 4-gram was not seen in the training corpus. In the Appendix, we discuss why this 1:10 rule is an appropriate choice.

[9] The phrase “collection of *n* contiguous” is an example of a 4-gram.

For a given month *t*, the training corpus includes all articles from either the EM or DM corpus that appear from month *t-27* to *t-4* (we discuss this window choice in the Appendix). For example, consider the 4-gram "central bank cuts interest." Our conditional probability measure for this 4-gram would be high if the word "interest" very often followed the phrase "central bank cuts" in our training corpus. If many other words also followed the phrase "central bank cuts," then *m* would be small, and we would consider this 4-gram unusual. We extend the concept of entropy at the 4-gram level to the article level by calculating the negative average log probability of all 4-grams appearing in that article. For a given article *j*, this measure is given by

$$H_j = -\sum_i p_{j,i} \log m_i \quad , \tag{2}$$

where $p_{j,i}$ is the fraction of all 4-grams appearing in article *j* represented by the i^{th} 4-gram, and m_i is *i'*s conditional probability from the training corpus. This measure is also known as the cross-entropy of *m* with respect to *p*, and we will often refer to it as *entropy* in our analysis.

Intuitively, we characterize an article as unusual if it contains language that is unlikely to have been seen in the past. We conjecture that such new language may be needed to describe new market or economic phenomena, and that the presence of the latter may indicate heightened (or perhaps reduced) market risks. In the same way that the context of a news story might matter for its market relevance, the entropy of the news story may matter as well.

Finally, we combine our topic analysis with article-level sentiment. Our article-level sentiment measure for article *j* is defined as

$$s_j = \frac{POS_j - NEG_j}{a_j} \quad , \tag{3}$$

where POS_j, NEG_j, and a_j are the number of positive, negative, and total words in the article. We use the Loughran-McDonald (2011) sentiment word lists to classify words as being positive or negative. This is the standard measure of sentiment that has been used in the finance literature (see, for example, Garcia, 2013). Tables 3 (for EM) and 4 (for DM) show s_j for the sample articles discussed earlier. In each topic, we report two sample articles with a very negative sentiment as well as two sample articles with a very positive sentiment. For example, in the DM corpus the article "Euro rises above $1.07 against dollar on war" in the *Mkt* topic plausibly receives a very negative sentiment value of -0.20.[10]

For topic τ, let us define $e_{\tau,j}$ as the number of econ words in article *j* that fall into topic τ and e_j as the total number of econ words in article j. Then $f_{\tau,j} = e_{\tau,j}/e_j$ defines the fraction of article *j*'s econ words that fall into a specific topic (recall topics are defined as nonoverlapping sets of econ words). We refer to $f_{\tau,j}$ as the frequency of topic τ in article *j*. We can decompose an article's sentiment into a context-specific sentiment measure via

$$s_{\tau,j} = f_{\tau,j} \times s_j \quad . \tag{4}$$

For example, an article with a sentiment measure of -3% that was mostly about government issues with $e_{Govt,j}/e_j = 90\%$ would have a government-specific sentiment of -2.7%. And its sentiment allocation to the other topics would be close to zero. Note also that since $\sum_j f_{\tau,j} = 1$ we'll have $\sum_\tau s_{\tau,j} = s_j$, which justifies our use of the term "decomposition."

In this paper, we are interested in testing whether topical context matters for the impact of news. Does negative or positive sentiment matter differently for forecasting future market outcomes when it occurs in news stories about governments than when it occurs in news stories about macroeconomics?

---

[10] Sometimes the lack of semantic context causes our sentiment classification to assign an inappropriate value, given the actual meaning of the article. For example, the article "BRIEF-Moody's revises Pulte's outlook to stable from positive," which appears (appropriately) in the *Credit* topic, is assigned a very positive sentiment score of 0.23 because it contains words like "positive" and "stable"–both positive sentiment words in Loughran-McDonald–though clearly being moved to stable outlook from a positive outlook is a mildly negative credit event. We regard these errors as inevitable noise in identifying sentiment that arises from the inherent complexity of combinations of words and the consequent difficulty in coding sentiment of phrases using sentiment values of individual words.

We also explored whether topic-specific sentiment interacts with entropy in its effects on market outcomes. Following Glasserman and Mamaysky (2018a), we compute article-level context-specific sentiment interacted with entropy as follows

$$SentEnt_{\tau,j} = f_{\tau,j} \times H_j \times s_j. \tag{5}$$

This measure–which differentiates between topic sentiment on usual or unusual news days – turns out to not be useful in our empirical results. We discuss later why this might be the case.

2.1.1. Aggregation of article data at the daily and monthly level

Once we have article-level data–either entropy, context specific sentiment or entropy, or topic frequency–we aggregate these into a country-level daily measure by weighting by the number of words (total, not just econ words) in the article in question divided by the total number of words in all articles about that country on a given day. For example, daily topic sentiment is

$$s_\tau \equiv \sum_j \frac{a_j}{a} \times s_{\tau,j} \ , \tag{6}$$

where a is the total number of words in all articles mentioning a given country on a given day. The analogous definition is applied for article entropy and frequency. The monthly measure of either sentiment, entropy, or topic frequency for a given country is the simple average of that month's daily measures.

2.2. Data summary and preliminary analysis

Table 5 contains a brief description of the variables used in our analysis, and Table 6 contains summary statistics for those variables from 1998 to 2015. Compared to DM, EM returns were higher (1.04% versus 0.65% per month), more volatile (21.48% versus 19.20% annualized volatility), and subject to higher drawdowns (17.4% versus 15.3%). As reported by

Ghysels, Piazzi, and Valkanov (2016), EM returns are also more right skewed as *retpl* (the positive portion of returns) averages 3.9% for EM and only 2.8% for DM, and *retpl* is also more persistent for emerging markets with an AR(1) coefficient of 0.12 versus 0.05 for developed markets. Emerging markets grew faster (*gdp*), had higher inflation (*gdpdefaltor*), higher nominal interest rates (*rate*), and lower private sector debt to GDP ratios (*cp*). The average number of articles per day (*artcount)* for EM countries is 26.0 and for DM countries is 106.7. The fraction of these articles dealing with *Corp*, *Govt*, and *Mkt* topics are similar, and EM countries have many more *Comms* articles (15.9% for EM versus only 2.7% for DM). Finally, the average article-level entropy for both corpora is roughly 2.45.

2.2.1. Structural break around the financial crisis

Fig. 4 and 5 provide factor loadings and plots for each topic category related to the first two principal components for the 140 EM (five series for 28 countries) and 120 DM (five series for 24 countries) time series of country-month-topic sentiment. The first principal component (both for EMs and DMs) tracks the aggregate time series of market sentiment. For both EMs and DMs, the second principal component appears as a step function with a break at the timing of the global financial crisis, and it has different factor loadings (both in sign and in absolute value) across different topic areas. *Govt* sentiment enters negatively and *Mkt* sentiment enters positively for the second principal component. That means that, prior to 2007, the sentiment score of market topic-related articles was more positive than government topic-related articles. That higher relative magnitude reversed after 2007, and the sentiment score of market topic-related articles became relatively negative in comparison with government topic-related articles. In our regression findings below, we find important breaks in regression coefficients (and some reversals in sign) that are related to this structural break in 2007-2008.

3. Empirical findings

Here we present our empirical findings about the connections between various measures of word flow and our measures of expected *return*, the standard deviation of returns (*sigma*), and cumulative downside risk (*drawdown*). As a starting point for our analysis, following Tetlock, Saar-Tsechansky, and Macskassy's (2008) and Hendershott, Livdan, and Schurhoff's (2015) analysis of company returns, we perform an event study of country stock returns around days in which sentiment scores for news for a given country are extremely positive or extremely negative. Specifically, we identify days for which positive or negative sentiment lies in the top decile of the historical distribution, and we do this for each of the five topical categories, separately for EMs and DMs. Fig. 6 and 7 plot cumulative abnormal returns (for EMs and DMs, respectively) for the ten days prior to and subsequent to the identified event dates (which appear as day 0 in the figures). Abnormal returns for each country are the residuals from regressing that country's US dollar returns on a constant and the appropriate MSCI index (either DM or EM) over the entire sample period.[11] We plot these abnormal returns separately for positive and negative news dates, along with standard error bands.[12] We also plot (in between the positive and negative top deciles) the results for the middle decile (45^{th}-55^{th} percentile) as a placebo.[13]

Interestingly, the plots for EMs and DMs are quite similar for the four common topical categories (*Mkt, Govt, Comms, and Corp*) and, surprisingly, are also quite similar for the fifth

[11] When running lagged regressions prior to the event date as the control, we noticed that the pre-event estimated alpha was correlated with the news event itself. Positive (negative) news days tended to be preceded by positive (negative) alphas. Because of this, the pre-event window was not an appropriate baseline return model, and therefore we used a regression over the entire sample as the control.

[12] Our standard errors are calculated under the assumption of serial and cross-sectional independence of events. Both assumptions are problematic in our data. Furthermore, it is possible that the pre-event country index performance has a causal relationship to the news event itself. Proper inference in this setting is beyond the scope of the present paper, and our standard errors should be interpreted with this caution in mind.

[13] The decile cutoffs are calculated over the entire sample. Note that the numbers of events in our three decile buckets are not the same. We bucket by the daily sentiment in each of the topic categories. Some of these event dates are either on nontrading days (e.g., weekends) or within ten days of the start or end of the sample. We do not include such event days for the calculation of abnormal returns.

(dissimilar) topical category (*Macro* for EMs and *Credit* for DMs). For both sets of countries, the patterns of cumulative abnormal returns around event dates are often similar for negative and positive news, although there are also some interesting asymmetries. Positive and negative cumulative returns tend to occur in advance of, respectively, positive and negative big news days, with the exception of negative news days for *Govt* and *Comms* in DMs and also positive news days for *Comms* in DMs.

One noteworthy aspect of the event studies is that news events appear to cause more of a market reaction in our DM sample than in our EM sample (note the bigger event-day price jump in the former compared to the latter). This reflects either more timely reporting by Reuters in their developed market news bureaus or information leakage (perhaps due to weaker regulatory enforcement) in EM economies.

It is interesting to compare our event studies to those in Tetlock, Saar-Tsechansky, and Macskassy (2008)–their Fig. 3. Our country level abnormal returns, relative to their US firm level abnormal returns, have more pronounced pre- and post-event drifts around negative news events–a finding that seems to hold for both EM (for *Mkt* and *Comms* topics) and DM (for *Mkt* and *Credit* topics) markets. In Tetlock, Saar-Tsechansky, and Macskassy (2008) abnormal returns on stocks seem to be very weakly mean reverting following negative news. Both their results and ours–in some cases–show a weak positive drift after positive news events. This more pronounced country-level drift after negative news is potential, though tentative, evidence of the relative micro efficiency and macro inefficiency of markets (see Glasserman and Mamaysky, 2018b for a theoretical exploration of this question).

In results not reported here, we investigated whether these extreme positive and negative news days are predictable based on prior days' sentiment scores. We found no evidence of a pre-event drift in sentiment–sentiment did not decrease (increase) in the ten days leading up to a bottom (top) decile negative sentiment day. Our evidence suggests that news reports respond more slowly to underlying market or economic developments than do returns. This does not imply, however, that word flow measures lack predictive content for returns. Indeed, as our monthly analysis below shows, lagged word flow measures (including sentiment) do have predictive content for *return*, *sigma*, and *drawdown*.

3.1 Panel regression analysis of risk and return in EMs and DMs

Tables 7-12 report regressions employing country-month observations, divided into EM and DM samples, for our three dependent variables (*return*[12], *sigma,* and *drawdown*).[14] We regress month *t* values of the dependent variables on lagged (either *t-1* or *t-2*) values of our explanatory variables. Our regressions are panels with country-month data and country fixed effects. Section A.6 in the Appendix discusses some associated econometric issues. In each table, we report nine different regressions, which consist of three regressions for each of three time periods: April 1998-December 2015 (the entire sample period), April 1998-February 2007 (the pre-global financial crisis period), and March 2007-December 2015 (the post-global financial crisis period).

Within each time period we first report a baseline regression, which includes control variables (nontextual predictors of the three dependent variables). Controls include two lags of monthly returns (for the *sigma* regressions, we use $retmi = \max(-return, 0)$ instead based on the findings in Bekaert and Hoerova, 2014), two lags of monthly volatility, and single lags of other financial, macroeconomic, and electoral cycle control variables, all of which are described

[14] Regression results for one-month ahead returns (*return*) are in Online Appendix Tables A3 and A4.

in Table 5. We include indicator variables that capture electoral timing by dividing time periods into preelection and postelection periods, as described in Section 2 and the Appendix.[15]

In addition to the baseline regression, for each time period, we report two additional regressions that examine the incremental predictive power of various word flow measures. Each of these specifications includes country-level monthly entropy ($entropy_{t-1}$), the monthly average of daily article counts ($artcount_{t-1}$), and the monthly frequency measure f_τ for each topic. The first specification (in column labeled *Sent*) includes each topic sentiment measure in its simple form, i.e., $s_{\tau,j}$. The second specification (labeled *SentEnt*) includes the entropy interacted versions of the sentiment variables, $SentEnt_{\tau,j}$ from Eq. (5). In the tables, we label rows showing the loadings on $s_{\tau,j}$ and $SentEnt_{\tau,j}$ as *sMkt*, *sGovt*, and so on; the column heading specifies whether these refer to the simple or the entropy interacted topic sentiment. All sentiment measures, except *entropy*, are normalized to have unit variance at the country level. The Appendix provides more details about our regression specifications.

Our findings with respect to baseline variables are consistent with prior studies and will not be commented on further here.[16] Coefficient values differ across EMs and DMs, and overall, *return*, *sigma*, and *drawdown* tend to be more predictable for DMs (as measured by higher R-squared). This confirms the view that the nature of news, and the range of potential news

[15] There is a large literature on forecasting country-level returns. The general conclusion has been that stock-level effects are also present at the country level. For example, lagged valuation ratios and lagged interest rates have all been shown to forecast country-level returns (see Asness, Liew, and Stevens, 1997; Ang and Bekaert, 2007; Angelidis and Tessaromatis, 2017; Hjalmarsson 2010, among many others). Also, momentum and reversal effects have been documented (Asness, Moskowitz, and Pedersen, 2013 and Richards, 1997). We control for these effects and also introduce other market (exchange rate changes) and macroeconomic (inflation, GDP growth, etc.) variables as additional controls.

[16] We observe, as have others (e.g., Fama and French 1988) very little forecasting power for one-month ahead returns. One interesting finding is that GDP growth is negatively associated with returns and positively associated with drawdown, especially in the later part of the sample. We can think of several explanations for this finding. First, positive GDP growth may raise the probability of contractionary monetary policy, which may be bad news for stocks. Second, GDP growth may serve as a proxy for states of the world in which coefficients on other variables in the model (such as momentum or value) would change if the model permitted time-varying parameters.

outcomes, differ in EMs and DMs (reflecting important differences in the political and economic environments, which are reflected in returns outcomes). Further observations follow.

Similarity of effect for returns and risk. When a word flow measure has a positive (negative) effect on *return*, it often tends to have a negative (positive) effect on *sigma* and a negative effect on *drawdown*. In other words, news contained in word flow is often either “good” or “bad” for all three dependent variables, where good news increases *return* and reduces risk measured either by *sigma* or *drawdown*. In fact, we never observe a coefficient on a text variable in a *return* regression that is of the same sign (and statistically significant) as the same variable’s coefficient in a *sigma* or *drawdown* regression.

*Incremental R-squared***.** The economic importance of word flow measures (incremental contribution to *R*-squared) tends to be relatively small for *return* and *sigma*, both in EMs and DMs, compared to their contribution to $return^{12}$ and *drawdown*. Volatility (*sigma*) is the most predictable of the three dependent variables, with values ranging from 0.45 to 0.48 in DMs and from 0.32 to 0.40 for EMs. The usefulness of baseline control variables is especially high for predicting *sigma* in DMs, while the incremental contribution of word flow to *sigma* is small in DMs and EMs.

In EMs, the economic importance of word flow is higher for all three return measures, but it is especially high for $return^{12}$ and *drawdown*. In DMs, *R*-squared increases for $return^{12}$ and *drawdown*, respectively, too (rising from 0.16 to 0.21, and from 0.26 to 0.32 for the sample period as a whole). In EMs, the absolute value of the increase is slightly larger, but the increase in *R*-squared as a proportion of baseline *R*-squared is much larger: for the sample period as a whole, including text measures roughly doubles the *R*-squared for both $return^{12}$ and *drawdown* (from 0.07 to 0.13 and from 0.06 to 0.12). For the precrisis period, that difference between EMs and DMs is even greater: for EMs, *R*-squareds for $return^{12}$ and *drawdown* rise from 0.02 to 0.13 and from 0.08 to 0.22, while for DMs these increase from 0.27 to 0.30 and from 0.40 to 0.45. We interpret this as confirming that the nature of news tends to be different in EMs and DMs: in EMs, where events reported in the news often contain information about fundamental shifts in political and economic regimes (which is relatively absent in DMs), the incremental value of tracking word flow is greater.

Effects of specific text measures. The impacts of individual text flow measures on annual returns and drawdowns often are economically large. In DMs, individual text measures are mostly significant for one-year ahead returns (as shown in Table 7) during the period after 2007. During that period, a one standard deviation increase in *entropy* is associated with a 3.9% higher return over the next year (the product of its standard deviation, 0.17, and its coefficient, 23.17). A standard deviation increase in *sMkt* is associated with a 5.1% increase in $return^{12}$, while a standard deviation increase in *sGovt* is associated with a 3.9% reduction in $return^{12}$.

Magnitudes for drawdowns (shown in Table 11) are comparable for the aforementioned variables (and signs are opposite), with the exception of the drawdown consequences of an increase in *entropy*, which are about half as large in absolute value. Additionally, in the drawdown regressions for the earlier subperiod, *entropy* and *sCorp* are statistically significant. A one standard deviation increase in *entropy* now forecasts an increase in drawdown (of roughly the same absolute value, and the opposite sign as observed for the later period). A one standard deviation increase in *sCorp* forecasts a 1.5% decrease in drawdown.

In EMs, as shown in Table 8, more text flow measures are statistically significant for one-year ahead returns. A standard deviation increase in *artcount* forecasts a 10.5% decline in returns in the early subperiod; there is no significant effect in the later subperiod. *Entropy* does not enter significantly in either subperiod. *fMkt* switches signs from a large negative returns effect (-11.0% per standard deviation) in the earlier period to a large positive effect (8.8%) in the later period. *fGovt* enters negatively in the earlier subperiod with a large magnitude (-10.4%), but it does not enter in the later period. *sCorp* enters negatively in the later period (-8.2%) but not in the earlier period. *fCorp* enters negatively in the earlier period but not in the later period. *fMacro* does not enter significantly in either subperiod, but its sign is consistently positive, and for the combined period, it shows a large and statistically significant effect of 5.9%.

More variables are statistically significant in the EM drawdown regressions (Table 12), often in both subperiods. Coefficient magnitudes are similarly large and, when statistically significant, are of opposite sign to those observed in the returns regressions. A one standard deviation increase in *entropy* flips from forecasting an increase in drawdowns of 5.9% (0.17 x 34.498) in the earlier subperiod to forecasting a decrease of 3.8% in the later subperiod. *sComms* enters negatively in drawdowns, which mainly reflects its forecasting power pre-2007.

Entropy interactions. We do not find that interacting sentiment measures with entropy, the *SentEnt* specification, adds much explanatory power. Coefficient magnitudes and *R*-squareds sometimes rise and sometimes fall across *Sent* and *SentEnt* specifications, but the changes tend to be small; interacting sentiment with entropy adds little. By itself, however, *entropy* enters as a significant in-sample predictor of *drawdown* for DMs and EMs in both subperiods, of *sigma* in EMs for the pre-2007 subperiod, of *return* in DMs in the post-2007 subperiod and of EMs in both subperiods, and of $return^{12}$ for DMs in the post-2007 period.[17]

[17] As we discuss below, out-of-sample results shown in Fig. 10-12 and Appendix Fig. A2 confirm the importance of including *entropy* in the model. We find that *entropy* is chosen for inclusion in the parsimonious elastic net model,

Time variation in coefficients. Consistent with our principal component discussion in Section 2, we find important differences in coefficient values for word flow measures over time–that is, differences between the pre-2007 and post-2007 periods. Tables A1 and A2 in the Online Appendix summarize our panel results by subperiod. A "+" ("-") in the table indicates that an explanatory variable enters with a positive (negative) coefficient and is significant at the 10% level or better. The symbol "Ø" indicates that the explanatory variable is not present in that specification (for example, $return_{t-1}$ is not present in the *sigma* panels). For DMs, negative coefficients on *return* for *fGovt* and *sGovt* are a feature of the post-2007 subperiod, as is the positive coefficient for *return* for *sMkt* and *sCorp*. For EMs, positive *fGovt* is associated with larger *drawdown* in the earlier subperiod, but not in the later. For EMs, the coefficient on *entropy* in the $return^{12}$ regression is zero across the two subperiods, while the coefficients on *entropy* in the *drawdown* regressions flip from positive to negative. For DM $return^{12}$, *entropy* matters (positively) only in the post-2007 period. *Entropy* has no effect on DM *sigma*. For DM *drawdown*, *entropy* flips from positive significant to negative significant as we move from the earlier to the later subperiod. This sign flipping for *entropy* is examined in more detail in Section 3.2 below.

Sign of sentiment and market outcomes. Coefficients for sentiment or frequency can be positive or negative, depending on the topic area and period. There is no general finding that positive sentiment is always associated with good news. In DMs and EMs, positive *sGovt* or *fGovt* can be bad news and positive *sCorp* or *fCorp* can also be a negative news event, whereas *positive sMkt* is typically good news for DMs and positive *sComms* and *fMacro* are typically good news for

for both EMs and DMs, for *return* and $return^{12}$, for *drawdown*, and for *sigma*, although its importance and its sign vary over time. See also the discussion in Section 3.2.

EMs. Clearly, there is something to be gained by considering the context in which positive or negative sentiment is expressed. Note that sentiment is statistically significant as bad news only in the later subperiod (although frequency of market, government, and corporate news is negative in EMs in the earlier subperiod). *sCorp* has a significant negative sign for EM *return* and $return^{12}$ and marginally negative for DM $return^{12}$ and a significant positive sign for *drawdown* in EMs only for the later subperiod; *sGovt* has a significant negative sign for DM *return* and marginally negative for $return^{12}$ and a significant positive sign for *drawdown* for the full sample and the later subperiod.

One interpretation of our findings on sentiment is that negative sentiment can indicate good news if the negative sentiment is describing problems that government actions are trying to address. The notion that negative sentiment in the context of government responses is reflecting positive policy news events could also explain the postcrisis timing of the surprising coefficients for sentiment. In Section 3.4, we show that *Govt* and *Corp* sentiment and frequency both predict increases in future economic policy uncertainty (Baker, Bloom, and Davis 2016) for DMs, which suggests that a policy channel is potentially at work. A similar pre- and postcrisis difference in influence could explain the observed sign flipping with respect to *entropy*. In the precrisis period, unusual word flow generally indicates risky times, but in the context of the postcrisis period, unusual word flow may be associated with unprecedented policy actions.

3.2. Pre- and postcrisis differences in the meaning of news flow

To address this pre- and postcrisis interpretation of the two anomalies observed above– the negative news content of *sCorp* and *sGovt* in the post-2007 subperiod, and the flipping of the sign on *entropy* to imply positive news content in the post-2007 subperiod–we take a closer look at the changing patterns of co-occurrence among entropy, sentiment, and topical frequency over time. To examine the nature of the role of crisis influences, we divide the post-February 2007 time period into two subperiods: the global crisis period from March 2007 to August 2011 (the midpoint of the post-February 2007 period) and the subperiod after August 2011. By splitting the post-February 2007 period in half, we are able to investigate whether post-crisis differences reflect changes that persist throughout the period or changes that are only related to the onset of the global financial crisis. As before, we consider EM and DM countries separately.

Fig. 8 and 9 display our results for DM and EM countries, respectively. We find that there are, indeed, changes in the patterns of co-occurrence among entropy, sentiment, and topical frequency across time, and that these differ in interesting ways for EMs and DMs. In each figure, we plot sentiment and frequency by topic first for all country-days within each subperiod and, additionally, for country-day observations in the top fifth percentile of entropy. Each chart shows the difference between the average country-day sentiment or frequency in that subperiod/entropy grouping (e.g., the early subsample high-entropy group, or the late subsample average-entropy group) and the full-sample average, normalized by the full-sample standard deviation. For example, the top-left chart in Fig. 8 shows that in the 1996—2007 time period for DMs, average government sentiment was 0.15 standard deviations lower than the full-sample average, whereas credit sentiment was 0.05 standard deviations higher. Our focus is on how sentiment and frequency by topic vary across time and across high versus typical entropy days. Our interpretation is that high-entropy days contain particularly informative news flow and are therefore worth singling out for analysis.

With respect to the top two panels of Fig. 8 and 9, using all the articles in each sub-period, we observe substantial changes over time in topical frequency and topic-specific sentiment, which differ between EMs and DMs. This variation could account for the fact that our regression specifications in Tables 7-12 gained little from including interactions between entropy and sentiment (contrary to Glasserman and Mamaysky, 2018a). It may be that modeling sentiment as topic-specific and including topical frequency as a regressor, in an environment with such dramatic change over subperiods in topical frequencies, captures much of the interaction between entropy and sentiment that would not otherwise be captured.

Conditional on observing high-entropy country-days in EMs, the relative frequencies of the five topics are nearly constant over time (bottom row, Fig. 9). High-entropy days in EMs are associated with fewer market- and more government-related articles. Interestingly, high-entropy days in EM are associated with generally lower sentiment levels across all topics except *Mkts* relative to average-entropy days. Furthermore, high-entropy EM days exhibit important changes over time in topic-specific sentiment (third row, Table 9). In high-entropy days, government topic-related sentiment becomes less negative during the March 2007-August 2011 subperiod than it was before, commodities-related sentiment scores become much more negative, and other topics show little change. In other words, unusual news related to commodities during the height of the global crisis tended to be negative in EMs. News related to government had slightly less negative sentiment during high-entropy days than it had in the first period. EM country discussions related to government (which always tend to be sentiment negative in high entropy days) are less sentiment negative during the height of the global crisis. After August 2011, topic-specific sentiment for high-entropy days in EMs reverts to its pre-March 2007 pattern.

As in EMs, high-entropy days in DMs (third row, Table 8) are typically associated with lower sentiment in all topic areas (except *Mkts*). The subperiods patterns for DMs during high-entropy days are somewhat different however. First, for the pre-March 2007 subperiod, the high-entropy-day topical sentiment scores are quite similar to those of EMs. Second, as in EMs, during the second subperiod, government-related articles on high-entropy days are less negative than before, although they are still very negative relative to average entropy days in that sub-period. But for DMs, all the other topical areas on high-entropy days become more negative in their sentiment scores during the post-February 2007 period (with commodities-related sentiment scores showing the least change). It is not surprising that unusual DM news days related to commodities during the crisis were less negative than for EMs, given that DMs tend to be users rather than producers of commodities relative to EMs. Neither is it surprising that DMs, where the global crisis originated (with housing and banking crises originating in the US, Ireland, Spain, and the UK), are the countries where unusual news during the post-February 2007 period became particularly negative for market, corporate, and credit topics.

Even more striking is the fact that DM sentiment patterns for high-entropy days did not revert to the pre-March 2007 patterns as they did in EMs. Instead, DMs saw a continuation of the post-February 2007 topic-specific patterns for sentiment scores. It appears that the changes in the structure and content of news related to the onset of the crisis were more persistent in DMs, where the crisis and policy reactions to it were more long lasting. In additional tests not reported here, we investigated whether that persistence of DM sentiment negativity for the four non-government topic areas (relative to the pre-March 2007 subperiod) is driven by a subsample of Eurozone or European countries. We found that it was not isolated to Europe or the Eurozone but reflected persistent changes associated with the crisis that applied to DMs more generally.

The patterns observed in Fig. 8 and 9 reinforce the interpretation that the two anomalies reported in Tables 7-12–the negative news content of *sCorp* and *sGovt* in the post-2007 subperiod and the flipping of the sign on *entropy* to imply positive news content in the post-2007 subperiod–are related to how news coverage and its meaning change during a crisis.

3.3. Persistence of effects and endogeneity of news: a panel VAR approach

It is noteworthy that the measured effects of news are greatest at long (one-year) time horizons. This implies that our news measures likely capture fundamental economic influences rather than transitory "animal spirits" (see also Sinha, 2016).[18] In the Online Appendix, we provide another perspective on the duration of news relevance by constructing panel vector autoregressive (panel VAR) models, separately for DMs and EMs, which measure the linkages among sentiment, entropy, monthly return, and monthly volatility.[19] These results are reported in Fig. A3-A6. This approach is also useful for gauging the extent to which news may itself reflect past market outcomes. Because of the need to constrain the dimension of the model, we collapse the various topic-sentiment measures into a single sentiment index, which–by ignoring the topical context–understandably reduces the measured importance of sentiment, compared to the results reported above. We report two versions of the model: one that puts the news variables first in the ordering (sentiment and entropy, followed by returns and volatility) and the other that puts the return and volatility measures first, followed by sentiment and then by entropy.

We find that, with minor exceptions, the effects of sentiment and entropy on returns are similar for EMs and DMs. When sentiment and entropy are first in the ordering, they both produce positive return responses in the first two months after the shock with no evidence of subsequent mean reversion–suggesting both are capturing long-term news and not transitory

---

[18] Shiller (2017) argues that animal spirits can, in fact, have large fundamental economic effects. In the present work, we are not able to distinguish effects of long-lived animal spirits from news that forecasts economic fundamentals, but we are able to reject the view that the news that drives market changes reflects short-lived animal spirits. Shapiro, Sudhof, and Wilson (2018) show that text-based sentiment measures forecast future macroeconomic outcomes in the US; Thorsrud (2016) presents similar evidence for Norway.

[19] We estimate the VAR using monthly data with two lags and country fixed effects. The variables in the VAR have units given in Table 6, except the sentiment measure, which scales to have unit variance. We are constrained to include only variables that capture monthly variation. In particular, we do not include 12-month returns or drawdowns in the VAR models. However, our impulse response functions allow us to gage the persistence of shocks to returns and volatility over many months.

animal spirits. When sentiment and entropy are second in the ordering, there is less evidence of persistent effects on returns, but this is because we do not differentiate sentiment according to its topical context; it is likely that this approach aggregates positive and negative responses across different topics.[20] Similarly, for both EMs and DMs, positive sentiment forecasts drops in realized volatility that persist for over a year after the initial shock. This is true, regardless of the ordering of the VAR. In the case of DMs, entropy shocks depress realized volatility for several months when entropy is the second variable in the system and have no effect when entropy is last in the system. Interestingly, in the case of EMs, entropy shocks increase realized volatility, regardless of the ordering of the variables. Sentiment and entropy are also dynamically related: shocks to either of these variables produces protracted negative results in the other.

We also find that intertemporal influence flows in both directions. Shocks to returns and volatility have significant, and sometimes protracted, influences on sentiment and entropy. Return shocks increase sentiment and decrease entropy, while realized volatility shocks decrease sentiment and increase entropy, again regardless of the ordering. These results highlight the importance of examining long-term cumulative effects of news on returns and drawdowns and of including lagged measures of returns and volatility, as we do in our above models that evaluate the predictive importance of news flow for future returns, volatility, and drawdowns.[21]

[20] Our panel regressions in Table 7-12 and our out-of-sample results in Section 4 both show that topic sentiment is important for future market outcomes, even after controlling for lagged returns and volatility. With only 18 years of monthly data, we do not believe we can reliably estimate a VAR with topic-specific sentiment (with two lags, this requires estimating two 8x8 coefficient matrixes).

[21] Our modeling of the effects of text on returns, volatility, and drawdown in Tables 7-12 employs one lag of each of the 12 text measures, but two lags of returns and of volatility. In results not reported here, we also experimented with adding additional lags of text measures. Doing so slightly improves the statistical significance of text measures in some cases, raises *R*-squared slightly, and sometimes diminishes the importance of nontext measures. Overall, the effect of adding additional lags of text measures is small and usually divides the explanatory power captured in the one-lag specifications across the greater number of lags of the text measures in the expanded version. We report

3.4. Comparison with Baker, Bloom, and Davis' a priori approach

The Baker, Bloom, Davis (BBD) (2016) index of economic policy uncertainty (*EPU*) measures the frequency with which newspapers in a given country mention the words "economy" and "uncertainty," along with references to political acts or actors in the same article. For a sub-sample of EM and DM countries, it is possible to compare our approach to measuring news with that of BBD (2016). Those countries include 11 DMs (the US, Canada, Germany, the UK, Italy, France, Spain, Netherlands, Japan, Australia, and Ireland), and seven EMs (Chile, Mexico, China, South Korea, Brazil, Russia, and India). Although these DM and EM subsamples are a small proportion of the total number of countries in our sample, they represent a very large proportion of the total economic activity in the larger sample, and therefore this is a highly relevant subsample. Our sample time frame is from 1998 to 2015.

In Table 13 we show that our word measures can explain substantial future variation in the BBD uncertainty measure (Table 13 shows a panel regression of the time *t* value of *EPU* on time *t*-1values of macro control variables and our text measures). It is interesting to note that we explain a much larger portion of future *EPU* variation in the DM sample than the EM sample. In DMs, higher *Mkt* and *Credit* sentiment forecast lower *EPU*, while higher values of *Govt* and *Corp* sentiment and frequency forecast higher *EPU*. In EMs, higher *entropy*, higher *Macro* sentiment and frequency, and higher *Corp* sentiment all forecast lower *EPU*, while higher article count forecasts higher *EPU*. For EMs, in the baseline model, dollar appreciation (*dexch*) forecasts higher *EPU*, though this effect is subsumed by our text measures.

Of greater interest is the explanatory power of the BBD economic policy uncertainty measure (EPU_{t-1}) for *return*[12], *sigma*, and *drawdown*, both by itself and in regressions that

only the one-lag specifications of text measures because doing so is more conservative, as it avoids falsely attributing effects to text measures that can be explained by lagged volatility or returns.

include our word flow measures.[22] Tables 14 and 15 evaluate the incremental explanatory power of the BBD measure for those three variables. For each variable we report four regressions: a baseline regression that includes neither our word flow measures nor the BBD measure, a second regression that includes only the BBD measure (*EPU*), a third that includes only our word flow measures, and a fourth that includes both the BBD measure and our word flow measures.

For DMs, in the second regressions for each variable in Table 14, the BBD measure does exhibit incremental explanatory power, but the effects are small. *R*-squared for $return^{12}$ is increased by 0.009, for *sigma* by 0.002, and for *drawdown* by 0.011. In contrast, including our word flow measures raise R-squared by much larger amounts. Furthermore, as shown in the fourth regressions for each dependent variable, in the presence of our measures, the BBD measure loses its statistical significance. In other words, the part of the BBD measure that contains incremental explanatory power for $return^{12}$, *sigma*, and *drawdown* is subsumed by our word flow measures, and our word flow measures also contain additional explanatory power.

For EMs, the second regression results shown in Table 15, for $return^{12}$, *sigma*, and *drawdown*, show that the BBD measure adds almost no incremental explanatory power for all three variables relative to the baseline. Furthermore, the *EPU* measure never enters significantly in any of the specifications. In contrast, for EMs, adding our word flow measures meaningfully increases *R*-squared for all three specifications. We conclude that our atheoretical approach provides a more effective means of distilling the information contained in news stories that is relevant for market return and risk.

4. Out-of-sample tests

[22] It should be noted that Baker, Bloom, and Davis (2106) argue that economic policy uncertainty is useful in forecasting macroeconomic– not market–outcomes.

There are two important reasons to explore out-of-sample forecasting properties of our model. First, given the substantial variation over time in coefficient estimates reported in Tables 7-12, it is unclear whether a forward-looking application of our model would produce useful forecasts of market return and risk. Second, the baseline and augmented models from Section 3 contain many explanatory–text and non-text–variables that make them susceptible to overfitting in any given sample.

When overfitting is a concern, the typical solution is to penalize coefficient estimates by shrinking their absolute value based on an objective function that weighs each (normalized) coefficient's contribution to explanatory power (which receives a positive weight) against the magnitude of that coefficient (which receives a negative weight). We use the elastic net estimator (implemented in the *glmnet* package of Hastie and Qian, 2016), which combines the least absolute shrinkage and selection operator (lasso) regression, introduced by Tibshirani (1996), with a ridge regression, to ameliorate this overfitting problem. In our panel setting, we estimate rolling five-year regressions using the elastic net objective function, which is given by

$$\min_{\beta} \frac{1}{2N} \sum_{i,t} \left(y_{i,t} - x'_{i,t-1}\beta\right)^2 + \lambda(\alpha\|\beta\|_1 + (1-\alpha)\|\beta\|_2^2/2) \ , \qquad (7)$$

where N is the total number of observations in the regression, $y_{i,t}$ is the response variable, $x_{i,t-1}$ is a vector of the predictors, $\|\beta\|_1$ is the L1-norm of the coefficients (the sum of the absolute values of the β vector), and $\|\beta\|_2^2$ is the L2-norm squared (the sum of the squares of the β coefficients).[23] We include country fixed effects by constructing demeaned y's and x's within each country grouping–so a constant in the above regression is not necessary. The choice of λ determines the penalty applied to the blended L1- and L2-norms of the coefficients. This

[23] One important subtlety in the out-of-sample estimation for 12-month ahead returns and drawdowns is to truncate the measured 12-month ahead outcomes in the pre time-(*t+1*) estimation window to ensure that they do not overlap with the *t+1* through *t+12* outcome that we are trying to forecast out-of-sample.

parameter is selected in each 60-month window to minimize the cross-validation error. We set $\alpha = 0.75$, though this choice has little effect on the predictions obtained from the model (setting $\alpha < 1$ improves the numerical behavior of the algorithm, as discussed in Hastie and Qian, 2016).

Even a powerful model selection procedure has a hard time when confronted with too many explanatory variables and a relatively small data set. Therefore, we impose some structure on our estimation by using only a subset of our nontext variables for the out-of-sample tests: one-month returns (for our volatility model we use only the negative portion of returns), one-month realized volatility, our value measure, the private sector credit-to-GDP ratio, and the local interest rate. Except the credit-to-GDP measure, all of these proxy for well-known asset pricing effects. The credit-to-GDP ratio was very important in the in-sample regressions (perhaps because of its predictive power for returns around the financial crisis), and so we keep it for the out-of-sample tests. These five variables, with a country fixed effect, constitute our out-of-sample baseline model. By selecting variables with known forecasting power for the baseline model, we effectively raise the bar for our text measures to add any value.[24] We keep all our text measures for the out-of-sample tests but drop commodity frequency and sentiment, which were unimportant in most full-sample regressions. By dropping only two text measures, as opposed to many nontext measures, we believe we are being conservative in our out-of-sample tests.

An elastic net regression performs both model selection–many of the β's can be set to zero–and shrinkage–the nonzero coefficient estimates are smaller than their Ordinary Least Squares (OLS) counterparts. A measure of the degree to which elastic net coefficients are smaller than their OLS counterparts is the ratio $\|\beta\|_1/\|\beta^{OLS}\|_1$. In our empirical results this ratio (reported in the upper left-hand corners of Fig. 10-12 and A2) ranges from close to zero, to

[24] Had we selected more non-text variables, the out-of-sample performance of the baseline model would be degraded because the elastic net would have too many degrees of freedom. Choosing variables that we know will work a priori makes the baseline model a tougher comparison.

nearly 100%, meaning that the elastic net sometimes chooses an optimal in-sample model with no explanatory variables (often this happens for our 1 month return forecasting regressions) and sometimes chooses a model with coefficient estimates almost as large as their OLS counterparts (for example, in many windows for forecasting 12 month ahead returns).

Fig. 10-12 show significant changes over time in the elastic net coefficient estimates for the variables in our model, including the text measures. Coefficient magnitudes, when non-zero, are large and similar to the statistically significant coefficients identified in Tables 7-12 and have similar temporal patterns. For example, the flip in the sign of *entropy* for EM and DM drawdowns and 12-month ahead returns is visible in the elastic net estimates. It is important to bear in mind that multicollinearity (which, by construction, is not apparent in elastic net estimates) leads to some noncomparability of coefficient magnitudes reported in Tables 7-12 and in Fig. 10-12. Nevertheless, the elastic net results reinforce the message of Tables 7-12 about coefficient magnitudes and their variation over time. One-year ahead *return* and *drawdown* display similar pictures (with opposite signs) for individual variables for EMs and DMs. For example, both sets of results show similar changes in model fit over time. *Value* plays an important but varying role in the regressions for both EMs and DMs, as do *sGovt* and *sMkt*. *Rate* is important, but varying, in DMs. *Artcount* and *fMacro* are important, but varying, in EMs. Note that in out-of-sample tests, as well as in our panel analysis, positive *sMkt* tends to be good news for future market outcomes, whereas positive *sGovt*, *fGovt*, *sCor,p* and *fCorp* tend to be bad news.

It is noteworthy that some variables have very similar coefficient estimates for EMs and DMs (such as *entropy*, *sGovt*, and *sCorp*), whereas others (like *rate* and *artcount*) only seems to matter in one group (*rate* for DM and *artcount* for EM) but not in the other. The *entropy* measure, in particular, is associated with very consistent coefficient estimates between EMs and DMs in all four of our forecasting specifications.

4.1. Trading strategy based on out-of-sample model predictions

To evaluate the economic importance of our text measures, we analyze how useful textual information would be to a mean-variance optimizing investor who already had access to our baseline model's out-of-sample forecasts. In other words, we assume an investor forms at each time *t* an estimate of future returns and volatility using the five variables that constitute our baseline out-of-sample model. We then additionally allow this investor to condition, using only out-of-sample data, on our text measures. We refer to this as the CM model. Finally, we also allow an investor to estimate next period's mean return and volatility for a given country using only historical return data. We refer to this as the naïve model (this model is just a rolling country fixed effect).

Following Campbell and Thompson (CT, 2008), we assume a myopic mean-variance investor whose allocation to country *i* at time *t* is

$$w_t^i = \frac{E_t[r_{t+1}^i - r_{f,t}]}{\gamma \times var_t(r_{t+1}^i - r_{f,t})} \ , \tag{8}$$

where r_{t+1}^i is country *i*'s next period return to a dollar investor (we discuss the horizon of this momentarily), and $r_{f,t}$ is the US six-month T-bill rate. This weight is applied to the month *t+1* excess return of country *i*. Like Campbell and Thompson, we cap w_t^i at 1.5, but unlike CT, we allow short selling by imposing a floor of -1.5. A floor of zero makes sense in the CT setting because they analyze the allocation between cash and the stock market, but in our context, negative information in our signals about a given country's stock returns is useful and ought to be used in the trading strategy, which is feasible given our focus on country-level stock index trading. Furthermore, we set $\gamma = 5$ (it is 3 in their paper) because with lower risk aversion our weights often hit the 1.5/-1.5 boundary rendering inter-model variation less important. Finally, to aggregate country weights into a portfolio at time *t*, we divide all w_t^i's by the number of countries for which we have a time *t* signal. These weights are applied to time *t*+1 returns. The net portfolio position is invested in the US six-month T-bill.

The conditional moments in Eq. (8) are calculated using either the out-of-sample naïve, baseline, or CM model (the latter two are estimated using the elastic net model). We use the model's 12-month ahead return estimate to proxy for the forward-looking monthly return expectation (estimating the model using one-month ahead returns does not identify the proper dependencies in the data because the time horizon is too short–as we discuss further below), and we use the square of the model's one-month ahead volatility estimate for the conditional variance (both quantities are reported in annualized terms). The portfolio is held for one month and then reconstituted at time *t+1* (to then realize time *t*+2 returns) based on the month *t+1* ending information. Since we use a five-year window, our first portfolio is formed in April of 2003 and our last is formed in December of 2015. The aggregate amount invested varies between -0.5 and 1.4 times the portfolio capital, so the strategy we have parameterized is not overly levered. We run the strategy separately for our EM and DM countries.

We intentionally ignore portfolio level optimization (e.g., correlations across countries) and also employ (as do Campbell and Thompson) a myopic investment rule, to isolate the informational content of our text measures. Our approach follows DeMiguel, Garlappi, and Uppal (2007) in using (i) rolling five-year estimates of conditional moments to form myopic mean-variance portfolios and (ii) an equal-weighted portfolio (in our case, equally weighted across countries though each country allocation varies according to Eq. 8). We use overlapping

12-month observations to estimate how expected returns depend on our predictors to address the well-known problems of estimation error with using short-horizon returns.[25] In fact, Britten-Jones, Neuberger, and Nolte (2011) show that a predictive model with overlapping returns can be transformed into a predictive model for one-period ahead returns but with a properly transformed set of regressors. Our use of the untransformed 12-month ahead forecast in a monthly rebalanced myopic portfolio is certainly suboptimal[26] but is transparent and captures enough of the underlying structure in the data to lead to meaningful results.

To evaluate the economic significance of our results, we estimate the three-factor international asset pricing model suggested by Brusa, Ramadorai, and Verdelhan (2017), henceforth BRV. Our factors are the net total return of the MSCI global index, the return on a currency carry trade, and the return on an investment in the US dollar funded by borrowing against a basket of global currencies.[27] Table 16 shows the results of these regressions. The CM strategy generates lower market exposures than the baseline and naïve models. Both strategies have minimal exposures to the currency carry factor and are short the dollar (which is a mechanical outcome of a net long in foreign stock markets without hedging the currency exposure). We note that for both EM and DM strategies, the naïve strategy (i.e., rolling country-level means and volatilities) leads to very poor investment outcomes. The baseline model is better and leads to an economically significant 6.8% annual alpha for DM countries (in general,

---

[25] For example, Cochrane and Piazzesi (2005) claim on p. 139 that to see the underlying economic structure in their model for bond returns they can't use monthly observations and must use overlapping annual ones. Also see Britten-Jones, Neuberger, and Nolte (2011).

[26] See Barberis (2000) for a comparison of myopic versus dynamic portfolio rules in the presence of return predictability.

[27] The currency carry trade and US dollar index return data are available from Lustig and Verdelhan's websites but do not cover our entire sample. Instead, for the carry trade we use the Deutsche Bank Currency Carry USD Total Return Index, and for the US dollar index we use the US Dollar Index. The US Dollar Index is adjusted to have a negative 1.8% per year carry to match the average return of the US dollar index obtained from Lustig and Verdelhan. This adjusted dollar index and the Deutsche Bank carry trade index match the BRV factors very closely in the part of the sample where they overlap. The MSCI returns and both currency series are obtained from Bloomberg.

with a single series with 153 monthly observations we will not have much power against the null), although the baseline model delivers a much weaker 3.27% alpha for EM countries. The CM model, which augments the baseline model with our text measures, is the best performer in both samples, with an annual alpha of 8.8% in each–this is a very large economic effect, though the alpha is statistically significant only in the EM sample.

Given our interest in the incremental information content of textual measures, perhaps the more interesting aspect of our analysis is not the absolute values of the alphas but whether the difference between the alphas of the CM and baseline models is large. As Panel B of Table 16 shows, the difference is 2% per year for DMs and 5.5% per year for EMs. Both differences are clearly important economically, especially so for EM countries. Differences are also statistically significant at standard levels.[28] This confirms our finding from the in-sample panels that our text-based measures yielded incrementally more predictive power for the EM countries. In a carefully constructed out-of-sample test we have shown that by using modern model selection techniques, we are able to use the information content of our text-based country-level measures to meaningfully improve on investment performance.

5. Conclusion

We develop an atheoretical approach for capturing news through various word flow measures, including sentiment, frequency, unusualness (entropy), and the topical context in which these word flow outcomes occur. We apply that approach to 51 countries over the time period 1998 to 2015. We find that it is possible to develop a parsimonious and flexible approach to extract from news flow information that is useful for forecasting equity market risk and

[28] The reason we have much more power to reject the null that the differences are zero is because the residuals from the CM and baseline models are highly positively correlated (over 90%), which leads to their difference having very little volatility.

returns. We find that news contained in our text flow measures forecasts one-year ahead returns and drawdowns. One interpretation of this finding is that word flow captures "collective unconscious" aspects of news that are not understood at the time articles appear but that capture influences on the market that have increasing relevance over time. It may be that these unconscious aspects of news even influence fundamental economic behavior in ways that produce changes in returns and risks, as conjectured by Shiller (2017).

We consider the importance of topical context by giving all news articles weights according to the topics they cover. Topical context is defined using the Louvain method for grouping words into clusters, or word groups. In our sample, there are five such topic clusters for EMs and five for DMs, four of which are common to both sets of countries.

It is useful to divide news analysis of countries by considering EMs and DMs separately, because the basic statistical properties of news and returns are different for the two sets of countries, as are the relevant topics for news stories.

Principal components analysis of topic areas suggests a possible change in coefficient values occurs during the onset of the global financial crisis. We divide our sample period into two at February 2007 to take this change into account, and we find that coefficient values on various word flow measures do change over time.

Our word flow measures (sentiment, frequency, and entropy) capture important aspects of news that are relevant for returns, volatility, and drawdown risk, and have incremental predictive power over and above a baseline specification of standard control variables. Coefficient magnitudes of text flow measures are often large. News tends to divide into good or bad news that is relevant both for returns or for risk (measured either by volatility or drawdown).

The predictive content of sentiment, frequency, and entropy not only vary over time but are also context specific. Depending on the topic area of the article in which word flow appears and the timing, some positive sentiment news days appear as negative news events.

Word flow measures tend to have greater incremental predictive power (measured in terms of percentage improvement in *R*-squared) for understanding returns and risks in EMs, although they also have important incremental predictive power for returns and drawdowns in DMs.

We compare the predictive power of our atheoretical approach to analyzing the information content of news with the Baker, Bloom, and Davis approach to measuring economic policy uncertainty through an a priori identification of key words. We find that our approach is correlated with the BBD measure. The BBD measure, however, has much less incremental explanatory power for returns, volatility, and drawdown risk than our word flow measures, and in regressions that include both the BBD measure and our measures, the BBD measure loses statistical significance.

We perform out-of-sample testing using an elastic net regression to investigate whether our model is economically useful despite the large number of explanatory variables and the time variation in estimated coefficient parameters. From the standpoint of out-of-sample trading strategies, the additional alpha generated by using text flow measures is greater in EMs. For both DMs and EMs, text measures contribute significantly to improvements in out-of-sample forecasts relative to a baseline model that excludes text measures.

We conclude that the meaning of news flow can be captured usefully through a small number of atheoretical measures (sentiment, frequency, and entropy). The meaning of those measures for stock market risk and return vary over time, vary across EMs and DMs, and vary according to the topical context in which sentiment and frequency are measured. Thus, it is important to distinguish across country types and topical contexts, and permit coefficient estimates to vary over time, when using text to forecast risk and return. Nevertheless, we find that it is possible to construct a parsimonious and flexible forecasting model that maps usefully from these atheoretical, context-specific measures of news flow into equity market risk and return.

<u>Appendix</u>

A.1. Text preprocessing for sentiment and document term matrix extraction

Cleaning the data involves (i) converting all text into lowercase, (ii) removing stop words (e.g., *it*, *out*, *so*, and *the*) though not negating stop words (like *no*, *nor*, and *not*), (iii) tokenizing the text (e.g., converting *Boston-based* to *Boston* and *based* as separate words), (iv) entity replacement (for example, *International Monetary Fund* → *IMF*, numbers → tokens *_n_*, *_mn_*, and *_bn_*), (v) sentiment negation using the Das and Chen (2007) algorithm, (vi) punctuation removal, and (vii) word stemming (which converts inflected words, like *cats* or *speaking* into their root form).

Once data have been cleaned, we select all relevant English language articles for either the EM or DM corpus. Articles in the Thomson Reuters archive are often revised several times after their initial publication. Such article chains, i.e., the initial article and subsequent revisions, are labeled with a unique Primary News Access Code (PNAC) code. For each PNAC code, we select only the final article in the sequence. If we were more focused on a high-frequency analysis, it would be more natural to select the first, rather than last, article, but for the time horizon of our analysis (monthly), we believe that the final article is likely to have the richest information content and the several hour lag from first to last article in a chain will not have a meaningful effect on our results.

A.2. Construction of econ word list

The initial list of economics words (which we refer to as econ words) was compiled by the authors by looking at every word in the index of Beim and Calomiris (2001) and then subjectively selecting words with important economic or market-related meaning. This yielded

237 words. Then using the articles from the Thomson Reuters corpus from 1996-2015 that were tagged by the publisher as being about emerging markets (having a *qcode* of *N2:EMRG*), we analyzed all words occurring more than 3,000 times in any given year. This yielded 3,831 unique words. We ranked these words by their cosine similarity (see definition in the next section) to our original set of 237 words, averaged over all years in which these words appeared more than 3,000 times. Out of those words with an average cosine distance greater than 0.015 (which can be thought of as roughly a correlation of 1.5%), the authors and their research assistants selected an additional set of words that co-occurred very frequently with the original set of words. We then culled our list to eliminate redundancy (words that have a common word stem).

We then added 59 more commodity-related words by looking up the commodity groups from the IMF's Indices of Primary Commodity Prices,[29] 18 subjectively determined housing-related words, and 8 law-related words.

As a final step for identifying econ words, we collected the most frequently occurring 500 bigrams in every year of our Thomson Reuters emerging markets article set. This yielded 2,052 unique bigrams (for example, the two most frequently occurring bigrams were "Reuters message," which we deemed not useful and "central bank" which we deemed useful) and for the 100 bigrams that we subjectively determined to be economically relevant, we replaced the bigram with a single token which would then appear in our document term matrixes (and therefore in our topic analysis). For example, the bigram "central bank" was replaced with the token *central_bank*. We repeated the same analysis for the top 500 trigrams in every year. These yielded 3,740 unique trigrams, of which we determined 13 to be relevant (for example, the

[29] See https://www.imf.org/external/np/res/commod/Table1a.pdf.

most frequently occurring retained trigram was "International Monetary Fund"). There were many fewer retained trigrams than bigrams because having a three-word phrase introduces a much greater degree of context than does a two-word phrase, which renders many of the examined trigrams not broadly useful.

This process yields a total of 1,242 unique tokens (words and tokenized bigrams and trigrams) for constructing our document term matrixes.

A.3. Topic extraction using the document term matrix

We consider two words to be closely connected–or to co-occur–if there are many articles in which the two words appear together. Our measure of co-occurrence is the cosine similarity between two words. This similarity measure is computed as $\frac{D_j' D_i}{\|D_i\| \|D_j\|}$, where D_i is the i^{th} column of the document term matrix, and $\|D_i\|$ is the Euclidean norm of the i^{th} column. The cosine similarity has several nice properties: it is zero for words that never occur together in the same document, it is 1 for a word relative to itself, and it is between zero and one for words that, conditional on how often they occur, tend to occur in articles together. Let us refer to the symmetric matrix whose element $A_{i,j}$ corresponds to the cosine similarity between words i and j as the co-occurrence matrix. The matrix A defines a network of our econ words, where the strength of the link between two words corresponds to their cosine similarity.

We are now interested in extracting the structure of this network by finding non-overlapping clusters of words (i.e., a word appears in only one cluster) that tend to occur together frequently. We will refer to such word clusters as *topics*. Here we follow the approach of Newman and Girvan (2004) and Newman (2006) and cluster our word network so as to maximize network modularity–which we do via the Louvain algorithm (see Blondel et al., 2008). For a given partition of a network into k communities, let us define the $k \times k$ symmetric matrix e as having its (i,j)th element equal to the fraction of all edge weights in the network that connect members of communities i and j. The modularity of the network, $Q = \text{Trace}\ e - |e^2|$, where $|\cdot|$ indicates the sum of matrix elements, is a measure of the extent to which intracluster links tend to occur more frequently than at random. The Louvain algorithm is a particularly effective maximization heuristic for finding network partitions to maximize modularity.

Fig. 3 shows the initial clustering produced by the Louvain method for our EM and DM co-occurrence network. Clusters are ordered from the largest (by number of words) to the smallest, with the number of words in a cluster on the y-axis. As is evident, the algorithm naturally produces five large clusters for the EM and DM corpora, as well as a collection of several dozen much smaller clusters. Following the initial Louvain clustering, we then place each word from a small cluster (i.e., one outside of the top five) into one of the top five clusters so as to maximize network modularity. This process then yields five EM and DM clusters–each of which is a subset of our 1,242 econ words.

To investigate the time stability of our clustering algorithm, we repeated our topic extraction over successive four-year windows of our DM and EM corpora (recall the data set runs from 1996 to 2015). In each four-year window we recalculated the modularity maximizing word clusters using the Louvain method. To compare the subsample word categories to the full sample ones, we use the best match method described in Section 2.2 of Meila (2007). Consider two sets of word topics, *C* and *C'*, defined over the same set of words. For each topic *k* in *C,* we find the topic *k'* in *C'* that has the maximum word overlap with *k*, while making sure that the mapping is injective–i.e., that a topic *k'* in *C'* only gets mapped into once (if at all, because *C* and *C'* may not have the same number of topics). We then count the total number of words in the best-matched topics in *C* and *C'* and divide this by the total number of econ words appearing in both clusterings. This measure tells us what fraction of all our econ words fall into the same topic category in *C* and *C'*, where "same" means the best-matched categories.

In the five 4-year subsamples of our DM corpus, we find that the fractions of words matched from each subsample set of categories to the same full sample ones are 70% (1996-1999), 79% (2000-2003), 80% (2004-2007), 84% (2008-2011), and 78% (2012-2015), respectively. So approximately 80% of all our econ words get placed into the same topic in the full sample and in each of our subsamples. For our EM corpus, these fractions are 72%, 77%, 74%, 78%, and 67%. In the last four years (2012-2015) of our EM sample, the full sample *Mkt* topic is split between the subsample *Macro* and *Mkt* topics, and the full sample *Comms* topic is split between the subsample *Comms* and *Mkt* topics; these account for the somewhat lower word overlap in this subperiod.

It should be noted that under the null that the full-sample clustering is identical to each subsample, we still wouldn't expect to empirically find 100% cluster overlap due to sampling variation. Furthermore, we do not weigh overlapping words by frequency of occurrence, and if we were to do so, we would find higher percent overlap than the reported numbers. We interpret these results as indicating that the topics we identify over the full sample are quite robust, and a subsample-level analysis identifies very similar topics to the full sample. In the paper, we present all our results using our full-sample topics.

We also investigated the potential usefulness of an alternative method–(LDA)–for defining topic areas. In LDA, words are not assigned to mutually exclusive groups, but rather a group is defined as a probability distribution over all the words. We performed a pilot study to investigate the sensitivity of our results to the use of the Louvain method as opposed to LDA. We found no major qualitative differences in the resultant topics from the two methods. However, the Louvain method is much faster. For example, the document term matrix for our EM sample has 4,994,729 rows and 1,240 columns (our EM sample has no occurrences of two of our econ words). The computation time for the Louvain method, which involves computing cosine similarity for all word pairs and finding clusters to maximize network modularity, is 40 seconds. Computing the LDA clustering with five topics for only a single month (February 2007) took 113.8 seconds, and computing LDA for all of 2007 took 1,708 seconds. Assuming LDA scales roughly linearly, it would take approximately 10 hours for our entire sample. We, therefore, decided to focus on the Louvain method, given its relative ease of computation.

A.4. n-grams and conditional probabilities

The count operators $\hat{c}$ in Eq. (1) return the number of occurrences of a given 4-gram w_1, w_2, w_3, w_4 and its starting 3-gram w_1, w_2, w_3 in the training corpus. For month *t*, the training corpus for our EM (DM) entropy measure contains all EM (DM) articles in our sample in the two-year period from month *t*-27 to month *t*-4. We use a rolling two-year window to keep the size of the information set for our entropy calculations constant at all months in our sample. The reason we skip months *t*-3, *t*-2, and *t*-1 is to treat a 4-gram and its starting 3-gram that both appeared for the first time in month *t*-3, and neither of which had appeared in our corpus before, as being unusual for the next three months (such a 4-gram would be assigned *m*=0.1 in months *t*-3, *t*-2, *t*-1, and *t*). In month *t*+1, this 4-gram would be assigned a much higher value of *m* because month *t*-3 would now have entered the training corpus.

We refer to the 1 and 10 present in Eq. (1) for m as the 1:10 rule. Continuing with the example from the prior paragraph, when we observe a 4-gram and its starting 3-gram for the first time, we need a rule for assigning an appropriate conditional probability. We would like to treat a new never before seen n-gram as being a representative member of the set of never before seen n-grams. For every month t, we find in the EM corpus all 4-grams that do not appear in month t's training corpus. We then compute the m for each never before seen n-gram for the remainder of our sample, i.e., from month t until December 2015. Tabulating these m's for never before seen n-grams across all months (we have close to 100 million observations) produces the distribution shown in Fig. A1 in the Online Appendix. The median value of this distribution is 0.083, and the mean value is 0.28. Therefore, our choice of 1:10 rule would assign an m roughly equal to this median to a 4-gram that is encountered in month t but not in month t's training corpus.

We experimented using n-grams that drop stop words and using n-grams that retain stop words. While the results were similar using the two methods, we chose to use n-grams that retain stop words because these often preserve more of the article's semantics. For example, the phrase "business sentiment has improved of late" yields the 4-grams "business sentiment has improved," "sentiment has improved of," and "has improved of late" when stop words are retained. With stop words removed, we have "business sentiment improved late," which may convey a different meaning that the original statement.

A.5. Market price and macro data

Our price data come from Bloomberg. Table 1 shows the mapping from each country, as well as for the MSCI EM and DM index, to its corresponding Bloomberg ticker. All stock price data are converted into US dollar terms using end-of-day exchange rates. Price data are converted into total returns, *return*, by adding in the dividend yield from the prior 12 months accrued over the horizon of the return calculation (either weekly or daily). Our realized volatility variable, *sigma*, is computed by Bloomberg over the last 20 business days of every month. Our drawdown measure, *drawdown*, is computed as the maximum negative return realized by an investment in a given market index over the ensuing 252 trading days. For a given return *return*, we define *retmi* (*retpl*) as $\max(-return, 0)$ ($\max(return, 0)$), i.e., the absolute value of the negative (positive) portion of returns.

To maximize the number of observations for which we have data, we construct our *value* variable to be an accounting-free measure of country-level stock valuation. We borrow ideas from Asness, Moskowitz, and Pedersen. (2013) and define *value* for a country in a month *t* as the average level of the US dollar price of the country's stock market index from 5.5 years to 4.5 years prior to *t*, divided by month *t*'s closing US dollar price of the country index. We obtain similar results in panel regressions where we use the market-to-book ratio as the value measure, but in this case, we lose many observations from our sample because of the lack of book equity data.

Below we document the methodology and data sources underlying our macro data.

- Rate of growth of real GDP **(***gdp*): Quarterly real GDP growth rate data are obtained from International Financial Statistics of the International Monetary Fund. Annual data are used only when quarterly data is not available. The series is calculated as year-over-year percent changes.
- Rate of growth of GDP deflator (*gdp_deflator*): Quarterly GDP deflator data are obtained from International Financial Statistics of the International Monetary Fund. Annual data

are used only when quarterly data are not available. The series is calculated as year-over-year percent changes.

- Credit-to-GDP ratio (*cp*): We look at domestic credit to private sector as a percentage of GDP. Annual-credit-to-GDP ratio data are obtained from World Development Indicators, World Bank. We use linear interpolation and Bank of International Settlements credit data to replace missing values.
- First-difference-of-credit-to-GDP ratio (*dcp*): First difference of *cp* at the monthly frequency.
- Interest rate (*rate*): For developing markets, we use monthly deposit rates from International Financial Statistics of the International Monetary Fund. Deposit rates refer to the weighted average rate offered by commercial and universal banks on three- to six-month time deposits in the national currency. For developed markets, we use government bond yield data from Datastream. The maturity of these yields ranges from 5 to 10 years, with 7 years being the average. We use quarterly data only when monthly data do not exist.
- Monthly percent change in local currency exchange rate versus the US Dollar (*dexch*): We obtain monthly exchange rate data from Datastream for EMs and from Bloomberg for DMs. All exchange rates are determined as the US dollar in terms of the local currency (for example, for Turkey, our exchange rate measure is 3.4537 on 12/8/2016). So a positive value of *dexch* represents a local currency (USD) depreciation (appreciation). The US series is set to zero. This variable is truncated at $\pm 50\%$.

- Preelection dummy (*pre*): The preelection dummy takes a value of one for all months in a six-month window prior to an election and a value of zero on the election month and all other months. We use the Database of Political Institutions for elections dates.
- Postelection dummy (*post*): The postelection dummy takes a value of one for all months in a six-month window after an election and a value of zero on the election month and on all over months. Any month that would receive a classification as both a pre- and postelection month is labeled as a preelection (but not a postelection) month. We use the Database of Political Institutions for elections dates.

A.6. Panel regressions

All panel regressions report robust standard errors using the White method. For the *return* and *sigma* panels, we cluster residuals by time to control for cross-sectional correlations between countries. In the *return*[12] and *drawdown* regressions, we use Thompson (2011) to cluster by both time and country to control for the serial correlation in our overlapping left-hand-side variables as well as for country correlations. We use the *plm* package from *R* for our panel data analysis.

Panel fixed effect regressions, with N individuals and T time observations, that include lagged, persistent independent variables as regressors (i.e., our one-month ahead volatility regressions in Tables 9 and 10) suffer from a bias in the AR coefficient estimates when N is large and T is small. Nickell (1981) shows that this bias in the AR(1) case is approximately equal to $-(1+\rho)/(T-1)$, where ρ is the AR(1) coefficient. Since our T is quite large (as large as 200), this bias, which only affects the lagged loadings on one-month realized volatility in one set of panels, is quite small.

Another problem exists with forecasting regressions that use lagged explanatory variables (like price ratios or interest rates) whose changes are correlated with return innovations (see Stambaugh, 1999 or Hjalmarsson, 2010 for an analysis in a multicountry setting). In this case, the coefficient estimate for the explanatory variable, while consistent, will be biased in small samples (in a panel setting Hjalmarsson, 2010 points out the bias will be of "second-order"). In our setting this issue may affect the interpretation of the coefficient loadings on our *rate* and *value* measures in our forecasting panels. Ang and Bekaert (2007) use a Monte Carlo study to show that the biases in these coefficient estimates cause them to underestimate the effects of rates and dividend yields on future country-level returns, and furthermore that the bias is quite small when T is 200 (as is the case here), especially when the forecasting period is a year or longer (as is the case for our 12-month ahead return and drawdown regressions). For these reasons, we do not believe that the Stambaugh bias is an important consideration in our setting.

Most importantly, the focus of our analysis in Tables 7-12 and A3-A4 is on the loadings on our sentiment measures. Because these measures are not mechanically related to past returns or drawdowns in the same way that price ratios and interest rates are, there is no reason to be concerned about bias in coefficient estimates for our text measures.

Finally, we investigate whether the nonnormality of our *drawdown* variable leads to incorrect inferences in the *drawdown* panels. The residuals from the *drawdown* panels, while not normal, appear quite symmetrical and are as close to normality as the residuals from our $return^{12}$ regressions. We conclude that nonnormality is unlikely to be problematic in our setting.

References

Ang, A., Bekaert, G., 2007. Stock return predictability: is it there? Review of Financial Studies 20, 651-707.

Angelidis, T., Tessaromatis, N., 2017. Global equity country allocation: an application of factor investing. Financial Analysts Journal 73, 55-73.

Asness, C., Liew, J., Stevens, R., 1997. Parallels between the cross-sectional predictability of stock and country returns. Journal of Portfolio Management 23, 79-87.

Asness, C., Moskowitz, T., Pedersen, L., 2013. Value and momentum everywhere. Journal of Finance 68, 929-985.

Baker, S. R., Bloom, N., Davis, S. J., 2016. Measuring economic policy uncertainty. Quarterly Journal of Economics 131, 1593-1636.

Barberis, N., 2000. Investing for the long run when returns are predictable. Journal of Finance 55, 225-264.

Beim, D. O., Calomiris, C. W., 2001. Emerging Financial Markets. Irwin-McGraw Hill, New York.

Bekaert, G., Erb, C., Harvey, C., Viskanta, T., 1998. Distributional characteristics of emerging market returns and asset allocation. Journal of Portfolio Management, 102-116.

Bekaert, G. Hoerova, M., 2014. The VIX, the variance premium and stock market volatility. Journal of Econometrics 183, 181-192.

Blei, D., Ng, A., Jordan, M., 2003. Latent Dirichlet allocation. Journal of Machine Learning Research 3, 993–1022.

Blondel, V., Guillaume, J.-L., Lambiotte, R., Lefebvre, E., 2008. Fast unfolding of communities in large networks. Journal of Statistical Mechanics 10, 10008.

Britten-Jones, M., Neuberger, A., Nolte, I., 2011. Improved inference in regression with overlapping observations. Journal of Business Finance & Accounting 38, 657-683.

Brusa, F., Ramadorai, T., Verdelhan, A., 2017. The international CAPM redux. Unpublished working paper. Oxford University.

Calomiris, C. W., Love, I., Martinez Peria, M.S., 2012. Global returns' sensitivities to crisis shocks. Journal of International Money and Finance 30, 1-23.

Calvo, G. A., Reinhart, C. M., 2002. Fear of floating. Quarterly Journal of Economics 117, 379-408.

Campbell, J., Thompson, S., 2008. Predicting excess stock returns out of sample: can anything beat the historical average? Review of Financial Studies 21, 1509-1531.

Cochrane, J., Piazzesi, M., 2005. Bond risk premia. American Economic Review 95, 138-160.

Cruces, J. J., Trebesch, C., 2013. Sovereign defaults: the price of haircuts. American Economic Journal: Macroeconomics 5, 85-117.

Das, S.R., Chen, M.Y., 2007. Yahoo! For Amazon: sentiment extraction from small talk on the Web. Journal of Management Science 53, 1375-1388.

DeMiguel, V., Garlappi, L., Uppal, R., 2007. Optimal versus naive diversification: how inefficient is the 1/N portfolio strategy? Review of Financial Studies 22, 1915-1953.

Fama, E., Fisher, L., Jensen, M. Roll, R., 1969. The adjustment of stock prices to new information, International Economic Review 10, 1-21.

Fama, E., French, K., 1988. Dividend yields and expected stock returns. Journal of Financial Economics 22, 3-25.

Froot, K., Lou, X., Ozik, G., Sadka, R., Shen, S., 2017. Media reinforcement in international financial markets. Unpublished working paper. Harvard Business School.

Garcia, D., 2013. Sentiment during recessions. Journal of Finance 68, 1267-1300.

Glasserman, P., Mamaysky, H., 2018a. Does unusual news forecast market stress? Unpublished working paper. Columbia Business School.

Glasserman, P., Mamaysky, H., 2018b. Investor information choice with macro and micro information. Unpublished working paper. Columbia Business School.

Ghysels, E., Plazzi, A., Valkanov, R., 2016. Why invest in emerging markets? the role of conditional return asymmetry. Journal of Finance 71, 2145-2191.

Hastie, T., Qian, J., 2016. Glmnet vignette. Unpublished working paper. Stanford University.

Hendershott, T., Livdan, D., Schurhoff, N., 2015. Are institutions informed about news? Journal of Financial Economics 117, 249-287.

Heston, S., Sinha, N., 2017. News vs. sentiment: predicting stock returns from news stories. Financial Analysts Journal 73, 67-83.

Hjalmarsson, E., 2010. Predicting global stock returns.' Journal of Financial and Quantitative Analysis 45, 49-80.

Kaminsky, G., Schmukler, S., 2008. Short-run pain, long-run gain: the effects of financial liberalization. Review of Finance 12, 253-292.

Karolyi, G. A., 2015. Cracking the Emerging Markets Enigma. Oxford University Press, New York.

La Porta, R., Lopez de Silanes, F., Shleifer, A., Vishny, R., 1998. Law and finance. Journal of Political Economy 106, 1113-1155.

Loughran, T., McDonald, B., 2011. When is a liability not a liability? textual analysis, dictionaries, and 10-Ks. Journal of Finance 66, 35-65.

Meila, M., 2007. Comparing clusterings–an information based distance. Journal of Multivariate Analysis 98, 873-895.

Netzer, O., Feldman, R., Goldenberg, J., Fresko, M., 2012. Mine your own business: market structure surveillance through text mining, Marketing Science 31, 1-23.

Newman, M.E.J., 2006. Modularity and community structure in networks. Proceedings of the National Academy of Sciences 103, 8577-8582.

Newman, M.E.J., Girvan, M., 2004. Finding and evaluating community structure in networks. Physical Review E, 69, 026113.

Nickell, S., 1981. Biases in dynamic models with fixed effects. Econometrica 49, 1417-1426.

Richards, A., 1997. Winner-loser reversals in national stock market indices: can they be explained? Journal of Finance 52, 2129-2144.

Roll, R., 1984. Orange juice and weather. American Economic Review 74, 861-880.

Rule, A., Cointet, J.-P., Bearman, P.S., 2015. Lexical shifts, substantive changes, and continuity in State of the Union discourse, 1790-2014. Proceedings of the National Academy of Sciences 112, 1790-2014.

Shapiro, A., Sudhof, M., Wilson, D., 2018. Measuring news sentiment. Unpublished working paper. Federal Reserve Bank of San Francisco.

Shiller, R. J., 2017. Narrative economics. American Economic Review 107, 987-1004.

Sinha, N. R., 2016. Underreaction to news in the US stock market. Quarterly Journal of Finance 6, 1-46.

Stambaugh, R., 1999. Predictive regressions. Journal of Financial Economics 54, 375-421.

Tetlock, P. C., 2007. Giving content to investor sentiment: the role of media in the stock market. Journal of Finance 62, 1139-1168.

Tetlock, P. C., Saar-Tsechansky, M., Macskassy, S., 2008. More than words: quantifying language to measure firms' fundamentals. Journal of Finance 63, 1437-1467.

Tetlock, P. C., 2010. Does public financial news resolve asymmetric information? Review of Financial Studies 23, 3520-3557.

Tetlock, P. C., 2011. All the news that's fit to reprint: do investors react to stale information? Review of Financial Studies 24, 1481-1512.

Thompson, S., 2011. Simple formulas for standard errors that cluster by both firm and time. Journal of Financial Economics 99, 1-10.

Thorsrud, L.A., 2016. Nowcasting using news topics: big data versus big bank. Upublished working paper. Norges Bank.

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society 58, 267-288.

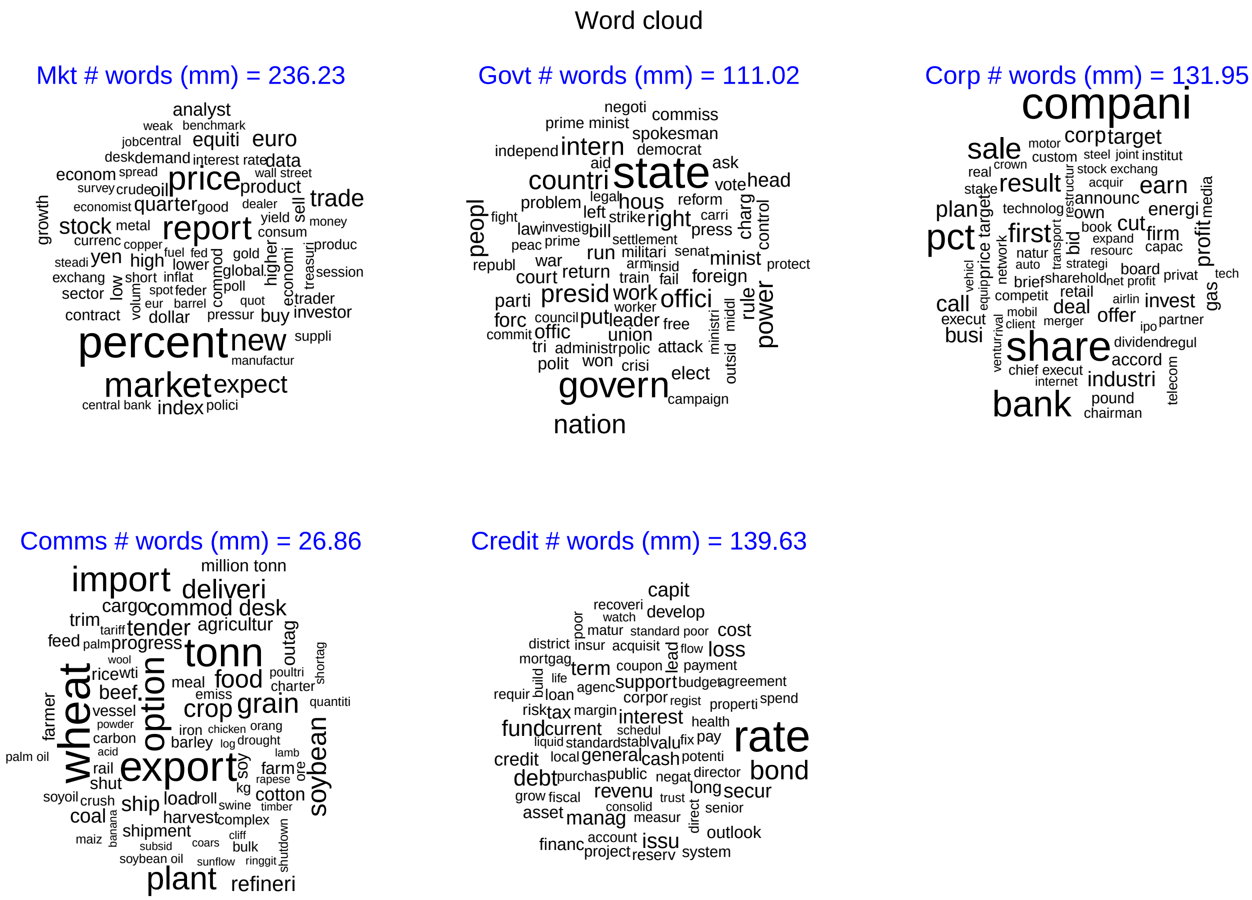


Fig. 1. Word cloud plots for topics extracted from the developed markets corpus using the Louvain clustering algorithm. Each cluster shows the number of occurrences (in millions) of its constituent words in the corpus. Larger font indicates words that occur more frequently within a given cluster.

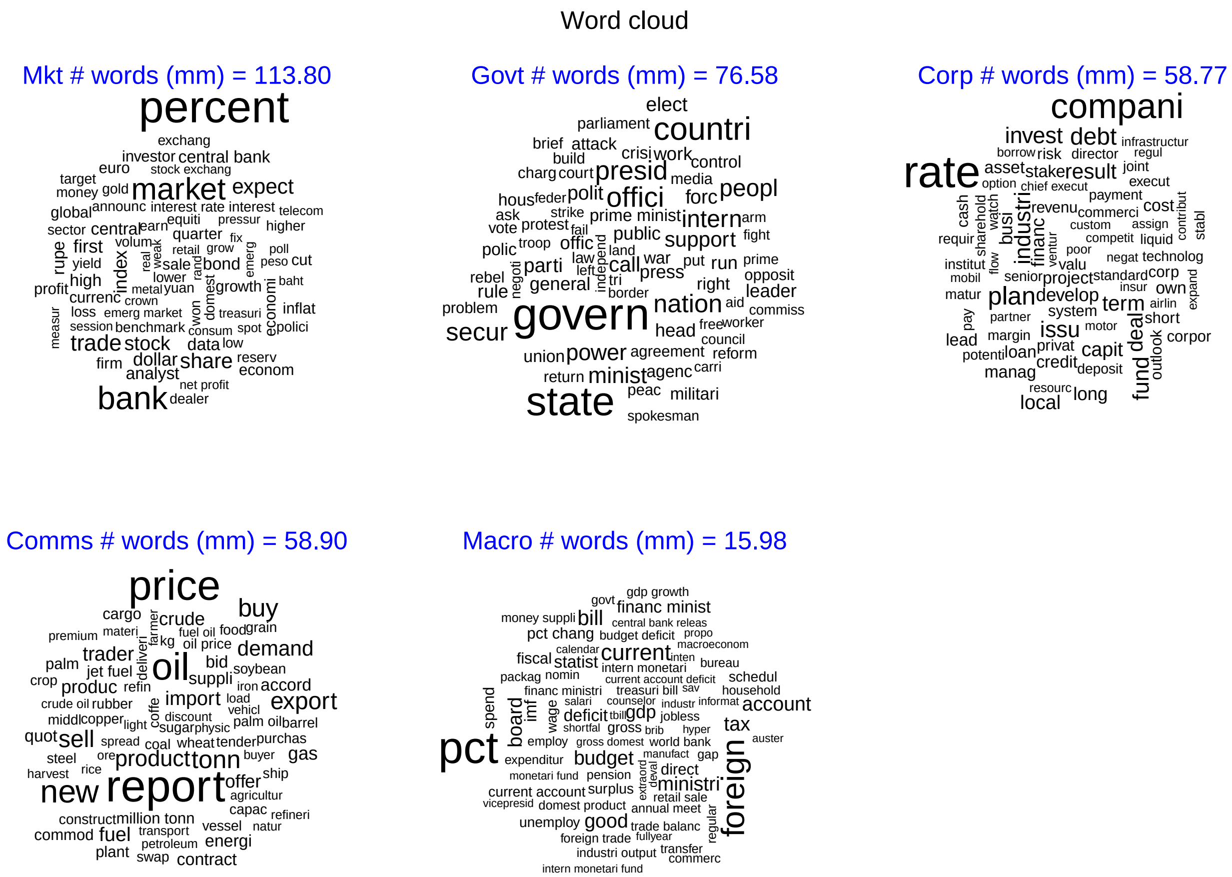


Fig. 2. Word cloud plots for topics extracted from the emerging markets corpus using the Louvain clustering algorithm. Each cluster shows the number of occurrences (in millions) of its constituent words in the corpus. Larger font indicates words that occur more frequently within a given cluster.

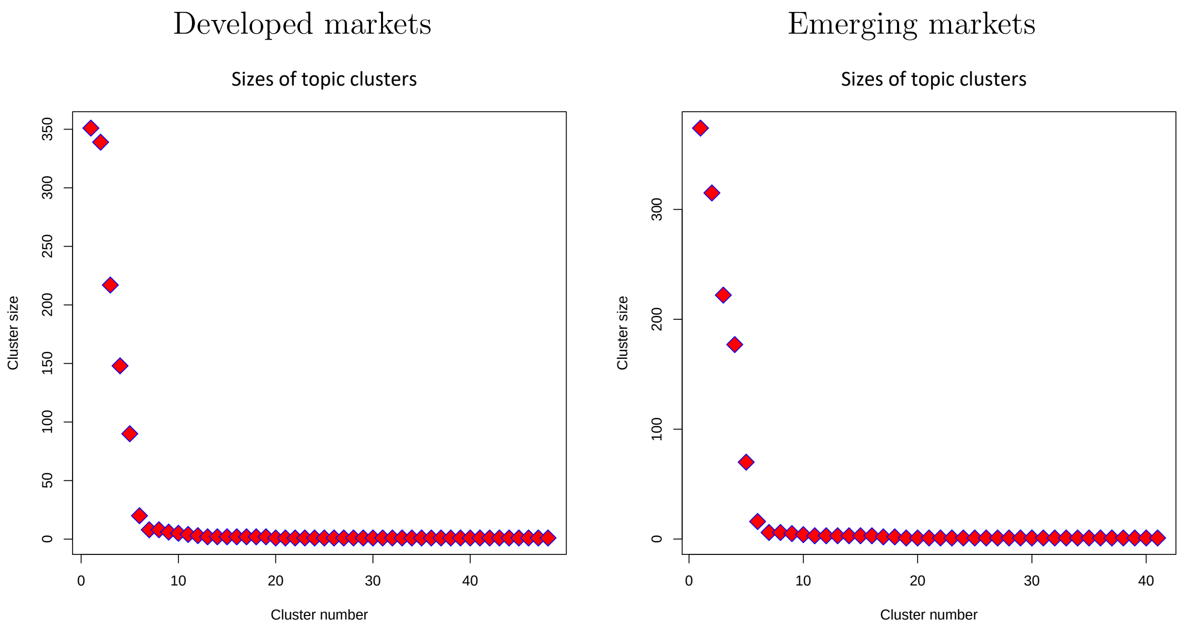
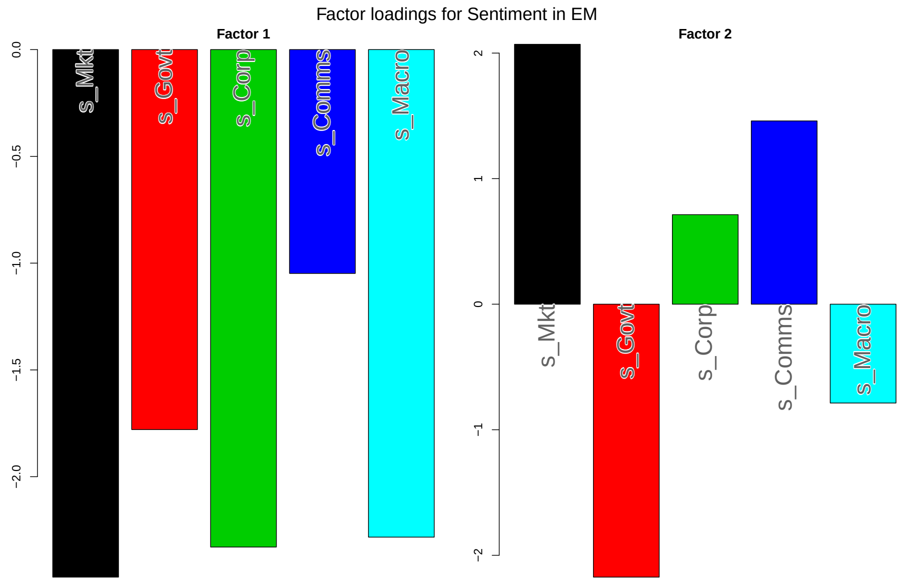


Fig. 3. Number of words in original set of clusters identified by the Louvain algorithm for the emerging and developed markets corpora, respectively. We allocate all words belonging to clusters outside of the top five into one of the top five clusters so as to maximize the resulting network modularity, subject to the constraint that the network consists of five clusters.

Factor decomposition of news topic sentiment in emerging markets



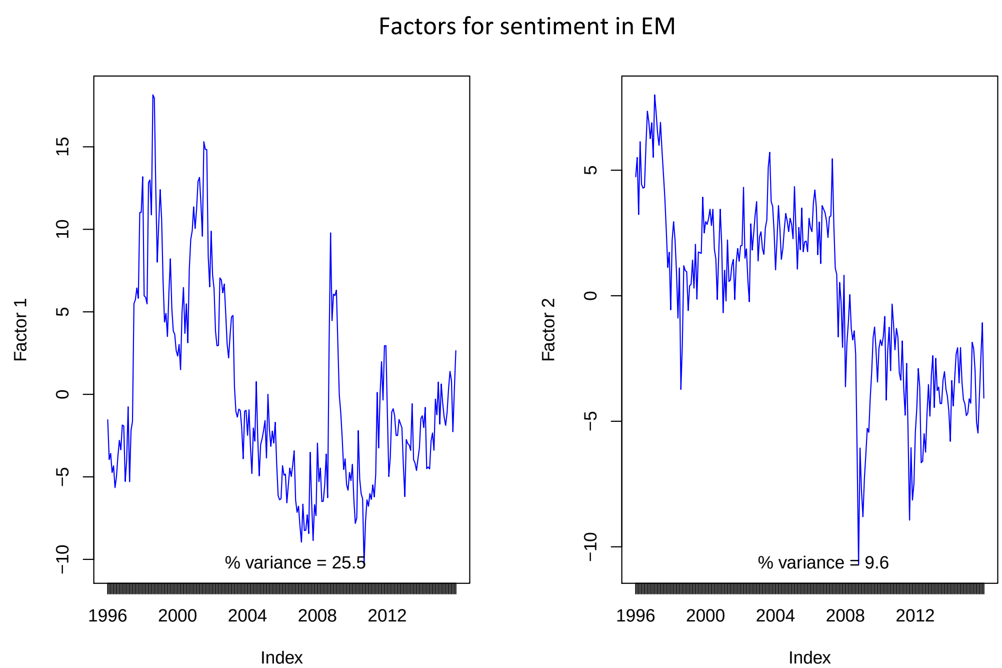
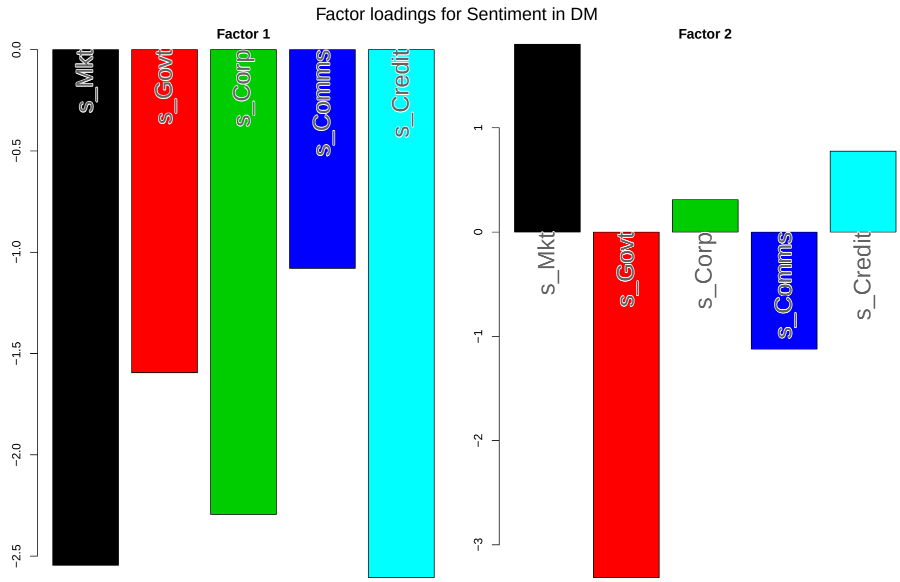


Fig. 4. The figure shows the two top factors from a principal components analysis of all country-topic sentiment series (i.e., $\#(countries) \times \#(topics)$) from emerging market countries. The top row shows the topic loadings of each factor aggregated by topic. Each topic bar is the sum of that topic's factor loadings across all countries in the sample. All country-topic sentiment series used in the principal components analysis were normalized to unit variance.

Factor decomposition of news topic sentiment in developed markets



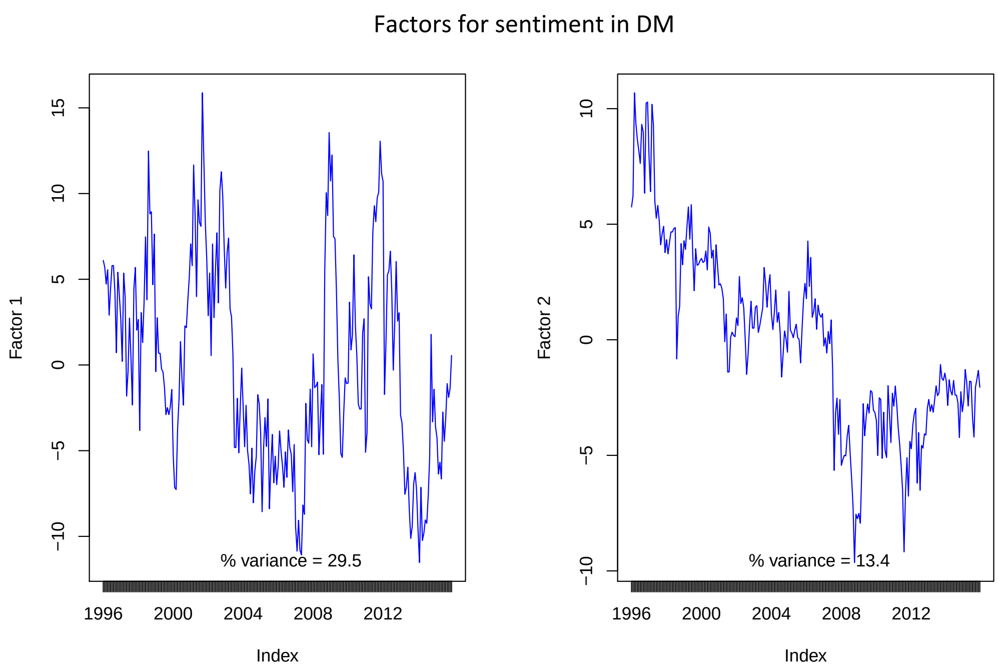


Fig. 5. The figure shows the two top factors from a principal components analysis of all country-topic sentiment series (i.e., $\#(countries) \times \#(topics)$) from developed market countries. The top row shows the topic loadings of each factor aggregated by topic. Each topic bar is the sum of that topic's factor loadings across all countries in the sample. All country-topic sentiment series used in the principal components analysis were normalized to unit variance.

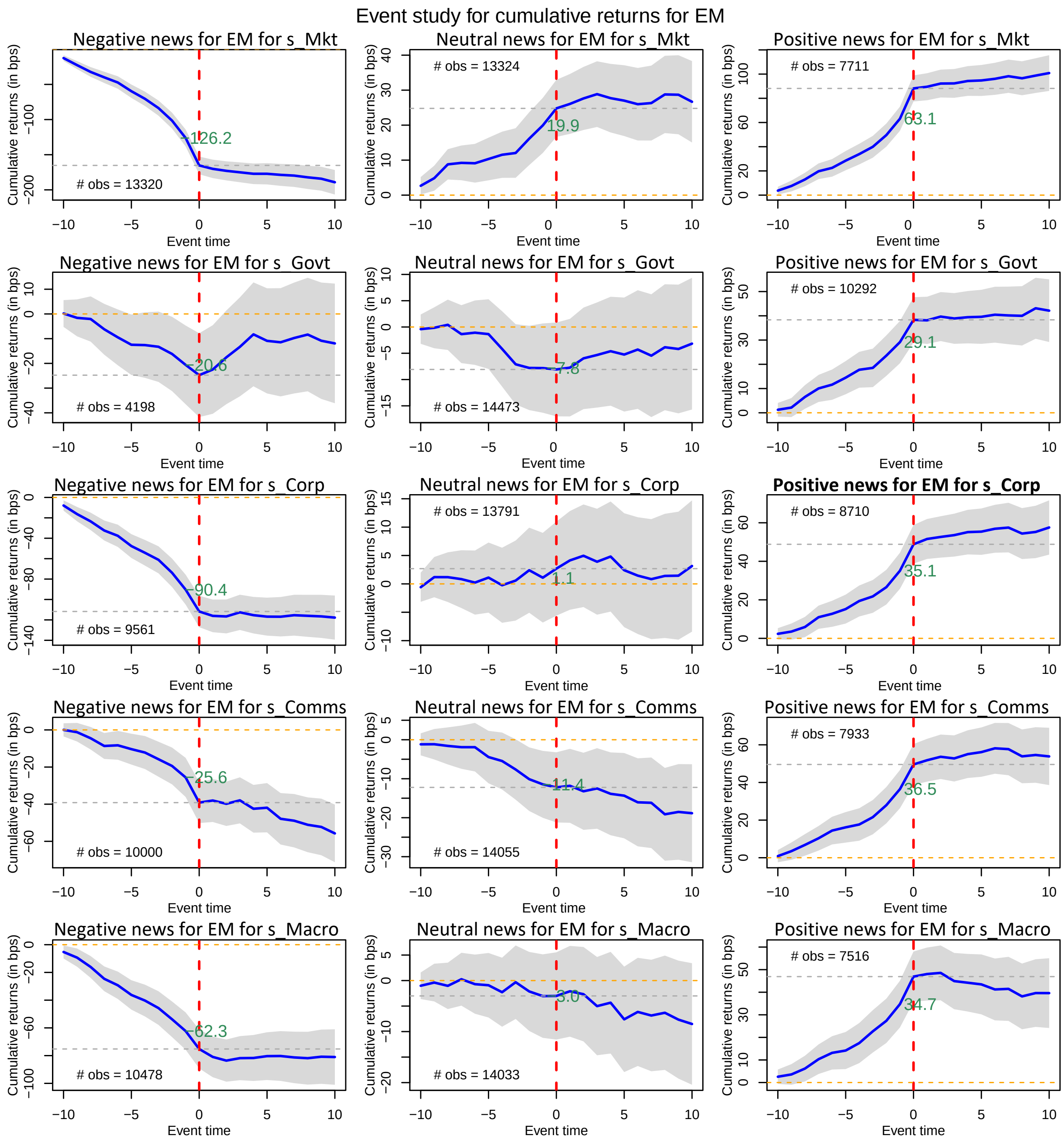


Fig. 6. Event studies of cumulative abnormal returns on days that are in the bottom, middle (45%-55%), and top deciles by sentiment for each topic. Topics are indicated by *s_[Topic]*, where *Topic* is one of *Mkt* (markets), *Govt* (government), *Corp* (corporate), *Comms* (commodities), and *Macro* (macro). Abnormal returns are the residuals from a regression of US dollar country index returns on a EM market index and a constant. Cumulative returns are shown in basis points, with two standard error bands, calculated under the assumption of independent observations. The cumulative return on the day prior to the event is labeled. The number of events in each study is shown on the plot. Events runs from January 1, 1996 to December 31, 2015.

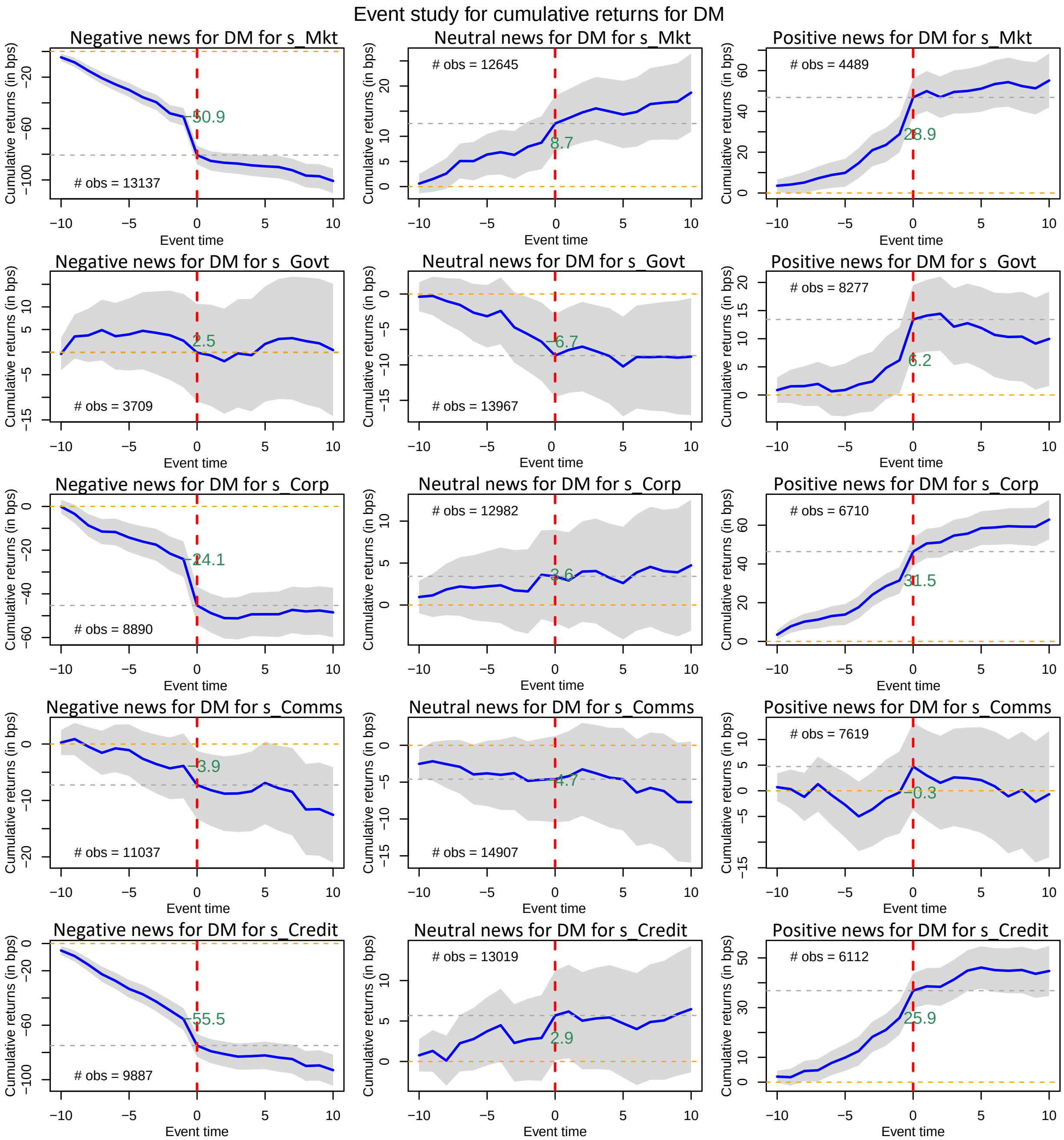


Fig. 7. Event studies of cumulative abnormal returns on days that are in the bottom, middle (45%-55%), and top deciles by sentiment for each topic. Topics are indicated by *s_[Topic]*, where *Topic* is one of *Mkt* (markets), *Govt* (government), *Corp* (corporate), *Comms* (commodities), and *Credit* (credit markets). Abnormal returns are the residuals from a regression of US dollar country index returns on a DM market index and a constant. Cumulative returns are shown in basis points, with two standard error bands, calculated under the assumption of independent observations. The cumulative return on the day prior to the event is labeled. The number of events in each study is shown on the plot. Events runs from January 1, 1996 to December 31, 2015.

How high-entropy days are different for DM

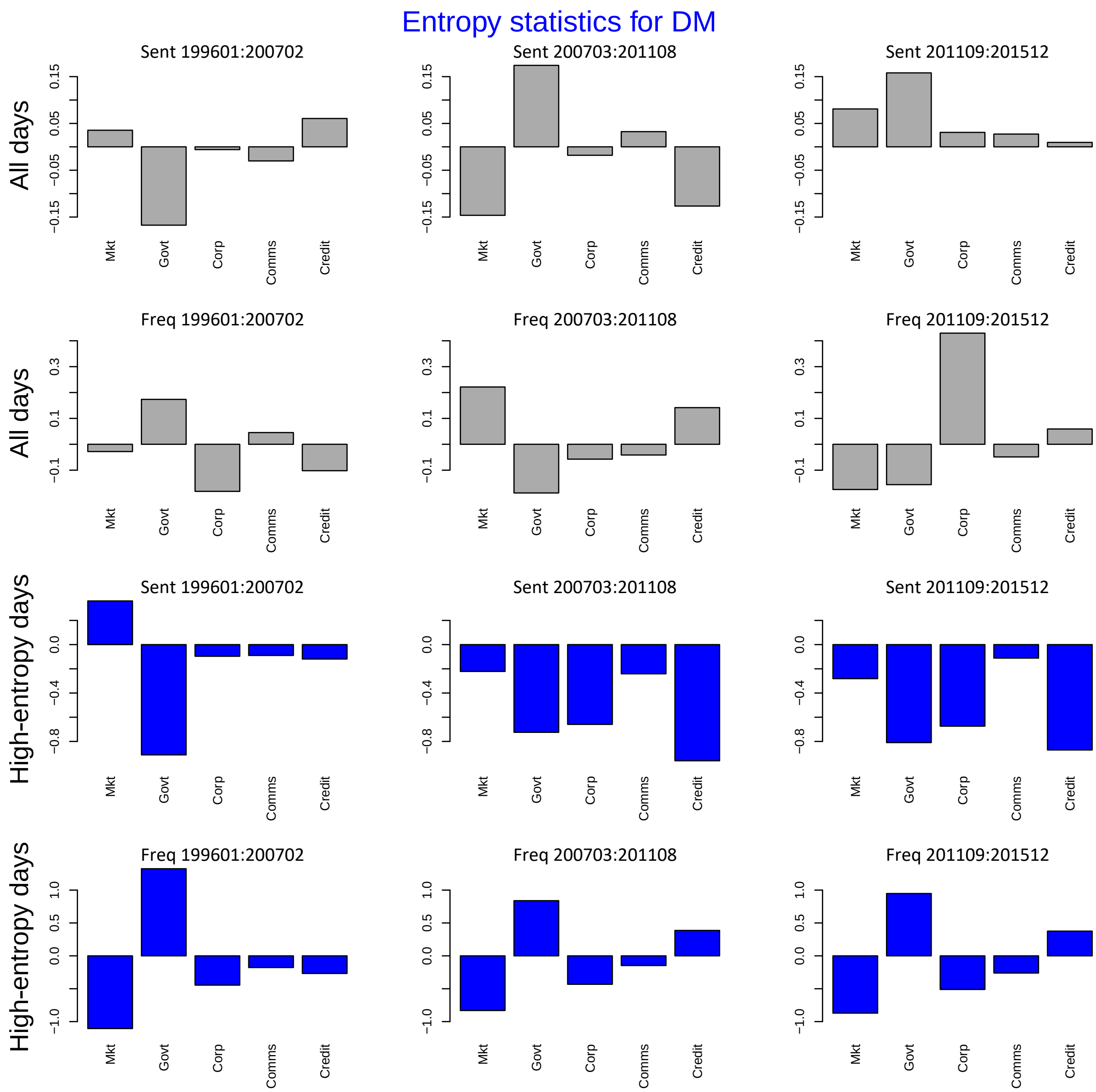


Fig. 8. The sample is split into three time periods: precrisis, crisis, and postcrisis. Within each subperiod, the first row shows the average sentiment by topic in that subsample minus the full-sample mean of topic sentiment, divided by the full-sample standard deviation of topic sentiment. The second row shows the same calculation but for topic frequency. The third row shows the same measure (sentiment) as in row 1 but restricted to country/day observations in the top 5% of full-sample entropy. The fourth row shows the same measure (frequency) as in row 2 but again for only country/day observations that are in the top 5% of full-sample entropy. In the figures, *Sent* refers to topic sentiment, and *Freq* refers to frequency of articles within that topic.

How high-entropy days are different for EM

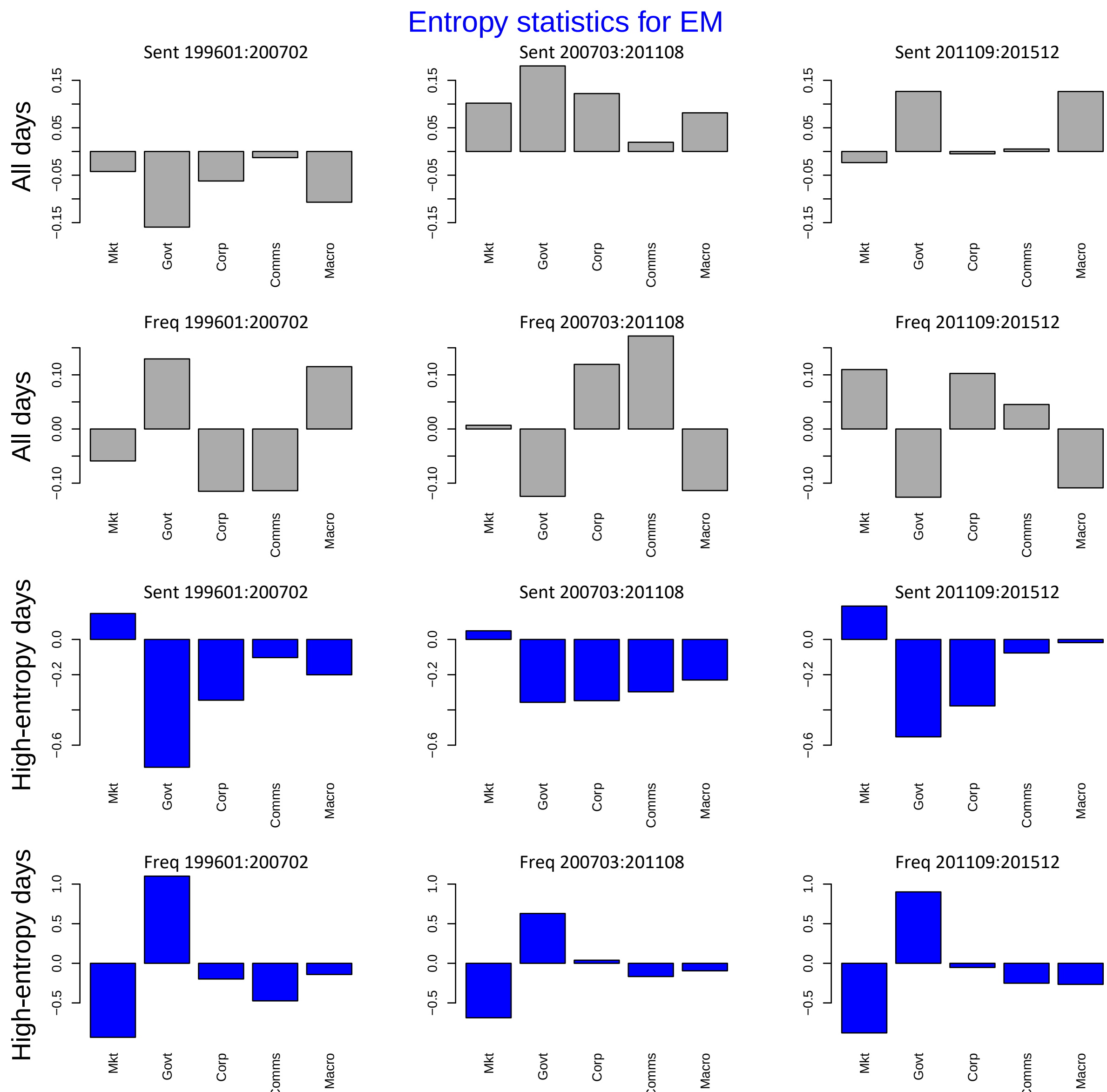


Fig. 9. The sample is split into three time periods: precrisis, crisis, and postcrisis. Within each subperiod, the first row shows the average sentiment by topic in that subsample minus the full-sample mean of topic sentiment, divided by the full-sample standard deviation of topic sentiment. The second row shows the same calculation but for topic frequency. The third row shows the same measure (sentiment) as in row 1 but restricted to country/day observations in the top 5% of full-sample entropy. The fourth row shows the same measure (frequency) as in row 2 but again for only country/day observations that are in the top 5% of full-sample entropy. In the figures, *Sent* refers to topic sentiment, and *Freq* refers to frequency of articles within that topic.

Coefficient time series from elastic net for 12-month returns

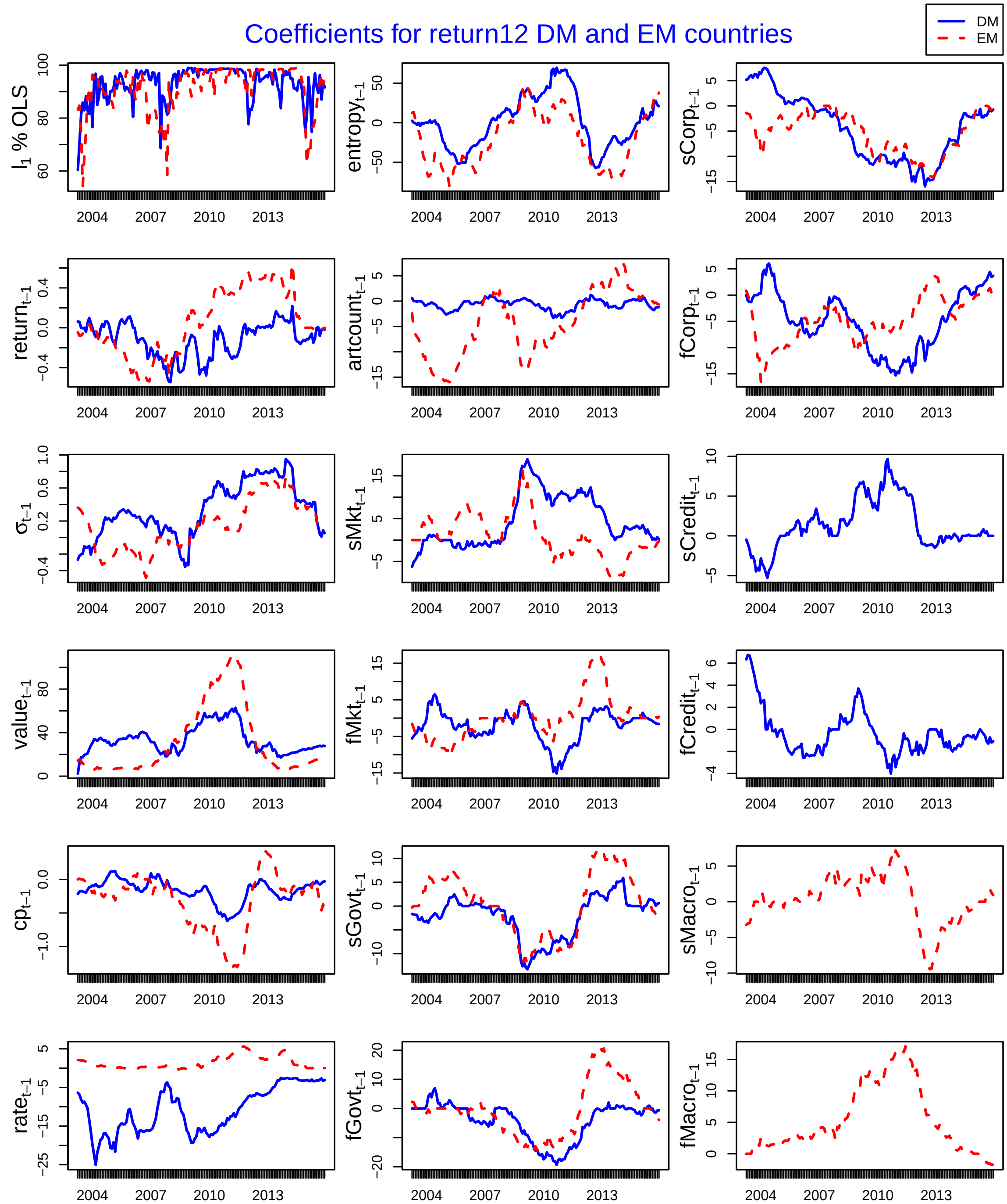


Fig. 10. The charts show the time series of coefficient estimates from a rolling elastic net regression to forecast 12-month returns. The chart labeled "l_1 % of OLS" gives the ratio of the elastic net coefficient l_1-norm to the OLS coefficient l_1-norm in every time period. Coefficient estimates refer to loadings on variables defined in Table 5. The elastic net regressions are run over rolling 60-month windows, with weighting parameter, λ, chosen to minimize cross-validation error. We set $\alpha = 0.75$, which represents a 0.75 weight on the lasso penalty and a 0.25 weight on the ridge regression penalty function. The out-of-sample forecasts start in March 2003 and go to December 2015.

Coefficient time series from elastic net for next 12-month drawdown

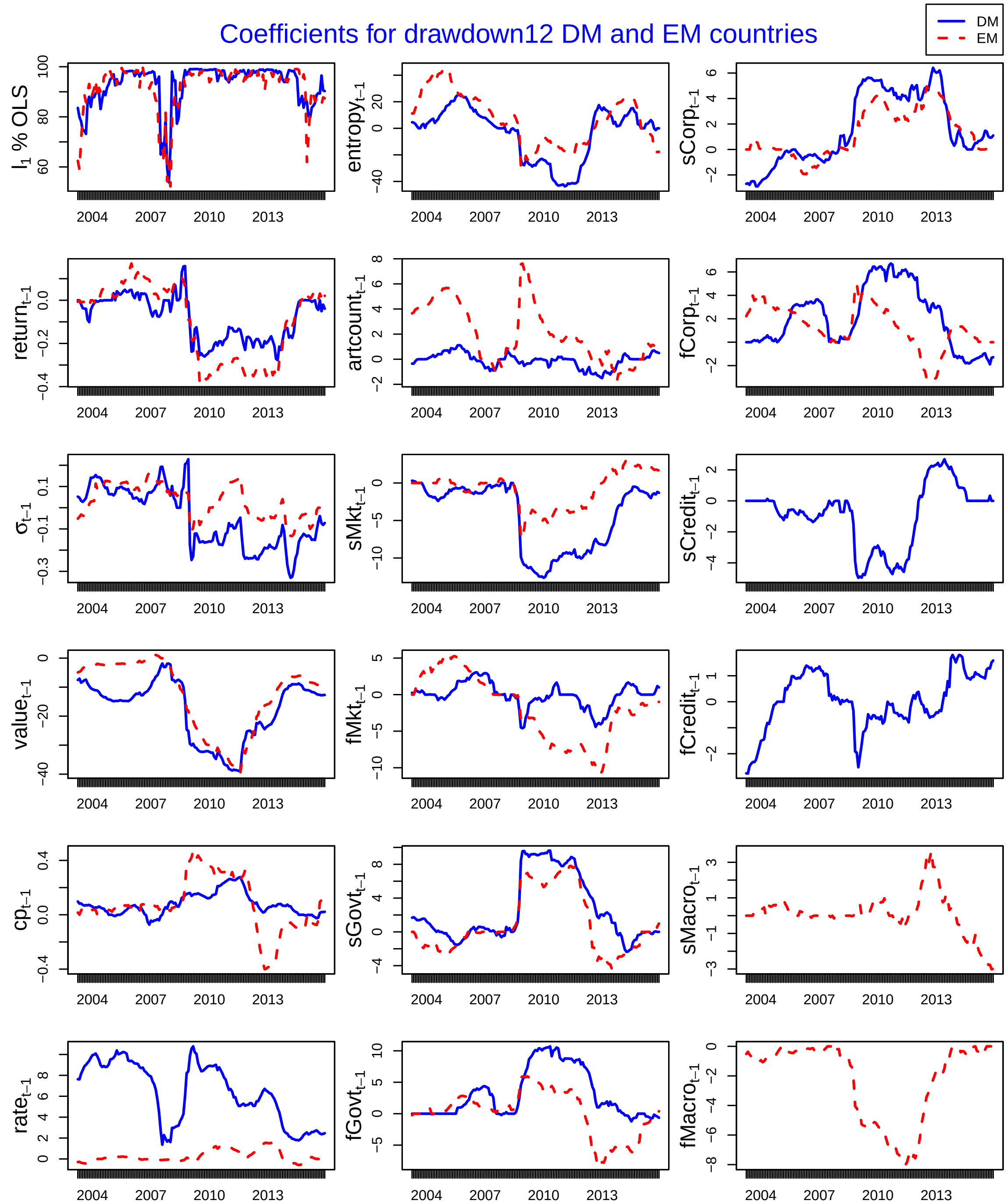


Fig. 11. The charts show the time series of coefficient estimates from a rolling elastic net regression to forecast next 12-month drawdown. The chart labeled "l_1 % of OLS" gives the ratio of the elastic net coefficient l_1-norm to the OLS coefficient l_1-norm in every time period. Coefficient estimates refer to loadings on variables defined in Table 5. The elastic net regressions are run over rolling 60-month windows, with weighting parameter, λ, chosen to minimize cross-validation error. We set $\alpha = 0.75$, which represents a 0.75 weight on the lasso penalty and a 0.25 weight on the ridge regression penalty function. The out-of-sample forecasts start in March 2003 and go to December 2015.

Coefficient time series from elastic net for realized volatility

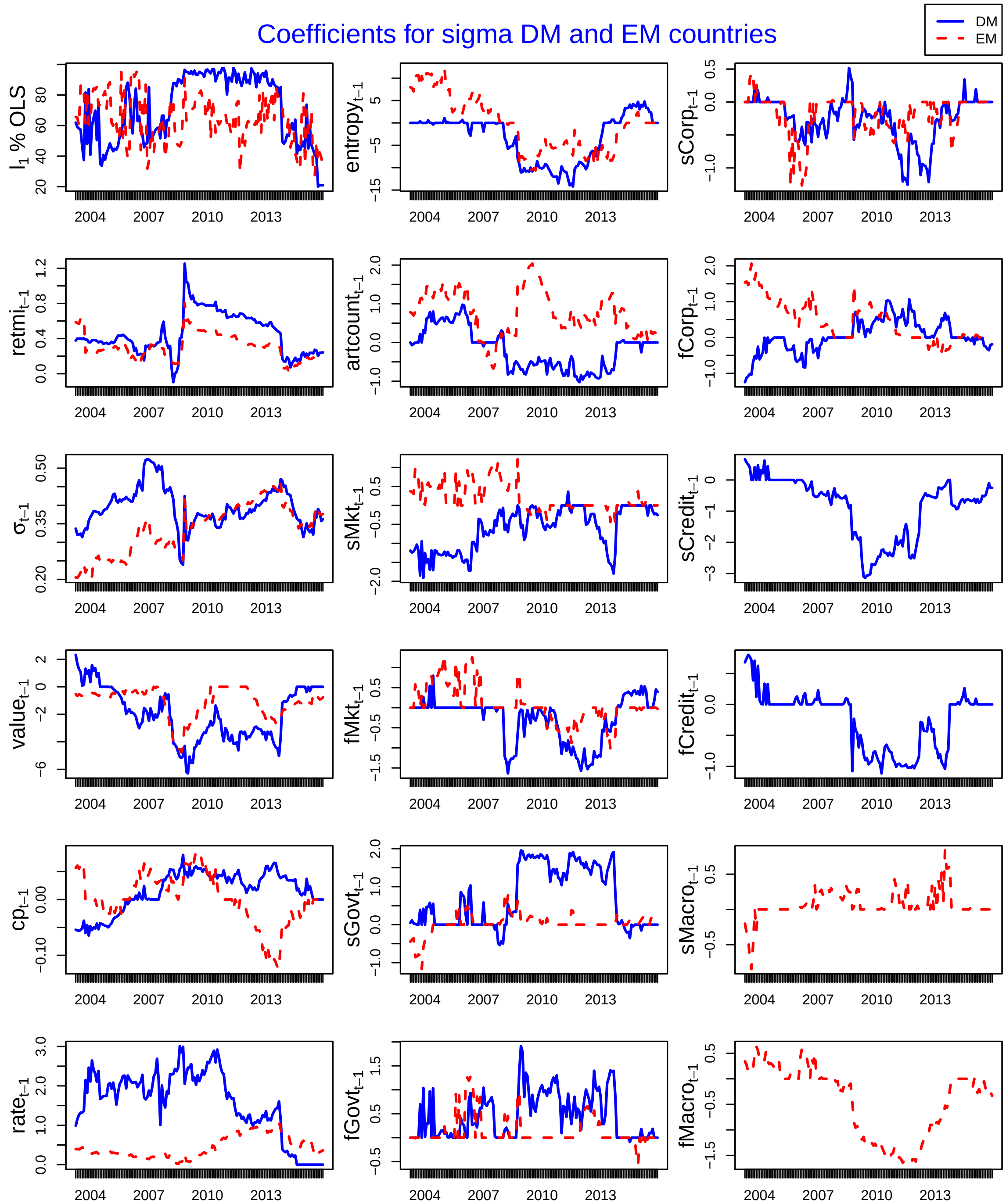


Fig. 12. The charts show the time series of coefficient estimates from a rolling elastic net regression to forecast realized volatility. The chart labeled "l_1 % of OLS" gives the ratio of the elastic net coefficient l_1-norm to the OLS coefficient l_1-norm in every time period. Coefficient estimates refer to loadings on variables defined in Table 5. The elastic net regressions are run over rolling 60-month windows, with weighting parameter, λ, chosen to minimize cross-validation error. We set $\alpha = 0.75$, which represents a 0.75 weight on the lasso penalty and a 0.25 weight on the ridge regression penalty function. The out-of-sample forecasts start in March 2003 and go to December 2015.

Table 1
List of EM and DM countries and their associated stock market index from Bloomberg (BBG index) as well as their country code (TR code) in the Thomson-Reuters Machine Readable News archive. The *EM* and *DM* rows refer to the MSCI EM and DM indexes, respectively, which are used to compute abnormal returns in the event study. Iceland text data is used in our event studies and in the topic clustering analysis but is excluded from all our panel and out-of-sample forecasting analysis.

List of EM and DM countries

	EM countries				DM countries		
	Country	BBG index	TR code		Country	BBG index	TR code
1	Argentina	BURCAP	AR	1	Australia	AS52	AU
2	Brazil	IBOV	BR	2	Austria	ATX	AT
3	Chile	IGPA	CL	3	Belgium	BELPRC	BE
4	China (PRC)	SHCOMP	CN	4	Canada	SPTSX	CA
5	Colombia	COLCAP	CO	5	Denmark	KAX	DK
6	Czech Republic	PX	CZ	6	DM	MXWO	–
7	EM	MXEF	EMRG	7	Finland	HEX	FI
8	Estonia	TALSE	EE	8	France	CAC	FR
9	Ghana	GGSECI	GH	9	Germany	DAX	DE
10	Hong Kong	HSI	HK	10	Greece	ASE	GR
11	Hungary	BUX	HU	11	Iceland	ICEXI	IS
12	India	SENSEX	IN	12	Ireland	ISEQ	IE
13	Indonesia	JCI	ID	13	Italy	ITLMS	IT
14	Israel	TA-25	IL	14	Japan	NKY	JP
15	Kenya	NSEASI	KE	15	Luxembourg	LUXXX	LU
16	Malaysia	FBMKLCI	MY	16	Netherlands	AEX	NL
17	Mexico	INMEX	MX	17	New Zealand	NZSE	NZ
18	Nigeria	NGSEINDX	NG	18	Norway	OSEBX	NO
19	Peru	SPBL25PT	PE	19	Portugal	BVLX	PT
20	Philippines	PCOMP	PH	20	Singapore	STI	SG
21	Poland	WIG20	PL	21	Spain	IBEX	ES
22	Russia	INDEXCF	RU	22	Sweden	OMX	SE
23	Slovakia	SKSM	SK	23	Switzerland	SPI	CH
24	Slovenia	SBITOP	SQ	24	United Kingdom	UKX	GB
25	South Africa	JALSH	ZA	25	United States	SPX	US
26	South Korea	KOSPI	KR				
27	Thailand	SET50	TH				
28	Turkey	XU100	TR				
29	Ukraine	PFTS	UA				

Table 2
Comparison of overlap between developed and emerging market clusters obtained via the Louvain network algorithm. For two clusters, A and B, the corresponding entry in the table reports $\#(A \cap B)/\#(A \cup B)$, where $\#(X)$ is the number of elements in set X.

Similarity of developed and emerging market clusters

	Mkt (EM)	Govt (EM)	Corp (EM)	Comms (EM)	Macro (EM)
Mkt (DM)	0.59	0.01	0.01	0.13	0.05
Govt (DM)	0.01	0.82	0.01	0.01	0.04
Corp (DM)	0.10	0.02	0.23	0.07	0.04
Comms (DM)	0.01	0.01	0.03	0.33	0.21
Credit (DM)	0.04	0.04	0.46	0.03	0.06

Table 3
For each topic, we show sample articles j whose topic allocation, i.e., $e_{\tau,j}/e_j$, is between 80% and 90%. For all articles that satisfy this criterion, we show the top and bottom two articles by sentiment within each topic. The *Sent* column shows our sentiment measure s_j for each article.

Sample articles in each topic for emerging markets

Topic	Date	Sent	Headline
Mkt	1997-11-06	-0.22	Elbit Ltd<ELBT3.TA><ELBTF.O>Q3 loss $0.11 per share
Mkt	1996-02-16	-0.22	Uganda shilling weakens against dollar
Mkt	1999-09-06	0.12	Hungarian shares open higher on Dow gains
Mkt	2015-03-05	0.12	BUZZ-USD/THB eked out small gains
Govt	2011-03-16	-0.23	US objects to 'excessive force' in Bahrain
Govt	1997-09-18	-0.22	Tehran mayor rejects resignations of 12 mayors
Govt	2000-06-04	0.10	Clinton says Putin can build strong, free Russia
Govt	2008-04-03	0.11	Mugabe's party expects runoff, says he will win
Corp	2011-01-19	-0.25	BRIEF-Moody's downgrades Tunisia's to Baa3, outlook negative
Corp	2011-01-31	-0.25	BRIEF-Moody's downgrades Egypt to Ba2, negative outlook
Corp	2013-05-02	0.14	CORRECTED-TABLE-Philippines' sovereign credit rating history
Corp	2013-03-27	0.16	TABLE-Philippines' sovereign credit rating history
Comms	2008-09-12	-0.13	BP says Baku-Supsa oil pipeline remains shut
Comms	1996-05-09	-0.12	Russia's Novorossiisk oil port still shut by fog
Comms	2006-12-27	0.08	Great Offshore buys anchor-handling tug vessel
Comms	1997-06-26	0.08	Tunisia tender for 150,000 T U.S. wheat detailed
Macro	1996-03-07	-0.12	Hungary 1995 C/A deficit falls to $2.48 billion
Macro	2003-04-30	-0.11	Turkish Jan-Feb c/a deficit jumps to $1.178 bln
Macro	2006-03-10	0.00	Sao Paulo volta a registrar inflacao no comeco de marco
Macro	2012-09-11	0.01	CORRECTED-Lithuania current account surplus rises in June

Table 4
For each topic, we show sample articles whose topic allocation, i.e. $e_{\tau,j}/e_j$, is between 80% and 90%. For all articles that satisfy this criteria, we show the top and bottom two articles by sentiment within each topic. The *Sent* column shows our sentiment measure s_j for each article.

Sample articles in each topic for developed markets

Topic	Date	Sent	Headline
Mkt	2012-05-21	-0.20	BRIEF-FINRA Panel awards John Galinsky $3.5 mln in compensatory damages for breach of contract against Advanced Equities
Mkt	2003-03-25	-0.20	Euro rises above $1.07 against dollar on war
Mkt	1996-01-18	0.12	UK's Clarke confident about inflation, growth
Mkt	2010-11-02	0.12	BRIEF-Metro CEO cautiously optimistic for good christmas
Govt	2009-01-08	-0.30	BRIEF-UK Serious Fraud Office to probe Madoff's UK operations
Govt	2005-09-09	-0.25	Soccer-Former secretary's claim against English FA dismissed
Govt	2014-04-29	0.13	BUZZ-GBP-4/5 on UKIP to win a seat in 2015 UK elections
Govt	2013-09-20	0.13	BUZZ-GBP-5/4 UKIP win most votes in European election
Corp	2014-07-21	-0.15	BRIEF-Valeant Pharmaceuticals contacts Quebec and U.S. regulators about Allergan's false and misleading statements
Corp	2015-12-16	-0.15	BRIEF-NQ Mobile announces termination of proposed divestment of Beijing Tianya
Corp	1996-05-26	0.13	Rangatira has 9.77 pct stake in Advantage <ADV.NZ>
Corp	2015-08-11	0.14	BRIEF-Tom Tailor to improve earnings in 2016 - CEO
Comms	2002-04-17	-0.07	Australasia port conditions - Lloyds
Comms	2012-06-13	-0.07	Cooperatives cut German 2012 wheat crop forecast
Comms	2006-10-10	0.13	TAKE A LOOK- Weekly US state crop progress reports
Comms	2006-10-16	0.13	TAKE A LOOK- Weekly US state crop progress reports
Credit	1998-11-16	-0.29	TABLE - NeoPharm Inc <NEO.A> Q3 net loss
Credit	1998-07-10	-0.27	TABLE - NDC Automation Inc <AGVS.OB> Q2 loss
Credit	2012-02-21	0.22	BRIEF-Moody's revises euramax's outlook to stable from positive
Credit	2011-04-21	0.23	BRIEF-Moody's revises Pulte's outlook to stable from positive

Table 5

Data definitions summary. More detailed information on variable construction and data sources is available in the Appendix. *Topic* is one of *government*, *markets*, *macroeconomics* (EM only), *credit* (DM only), *commodities*, or *corporate events*.

Data definitions summary

Variable	Definition
return	Total monthly stock returns (in %) including capital gains and dividend yield
$return^N$	Cumulative stock returns from the start of month t to the end of month $t+N-1$
sigma	Rolling 20-day realized volatility reported in annualized terms
$drawdown^N$	For a $100 initial investment, the maximum loss, potentially 0, experienced over the subsequent N-month period (for 12-month drawdowns, we often omit N)
retmi	Negative portion of returns (i.e., $\max(-return, 0)$)
retpl	Positive portion of returns (i.e., $\max(return, 0)$)
value	Average stock index level from 4.5 to 5.5 years ago divided by current index level
gdp	Rate of growth of real GDP
$gdpdeflator$	Rate of change of the GDP deflator
cp	Private sector credit-to-GDP ratio
dcp	First difference of credit-to-GDP ratio
rate	Local currency rate: deposit rate for EM and five- to ten-year government bond yields for DM
dexch	Percent change in value of US Dollar in terms of local currency (positive values are local currency depreciations), truncated at ±50%
pre	Dummy variable set to one if month t is six or fewer months prior to an election
post	Dummy variable set to one if month t is six of fewer months after an election
entropy	Daily word count weighted average of article level H_j averaged over a month
artcount	Number of articles written about a country per day, averaged over a month
$s[Topic]$	Sentiment s_τ in a given month due to *Topic*
$f[Topic]$	Frequency f_τ of articles in a given month in *Topic*

Table 6

Data summary using country-month observations from April 1998 to December 2015 (the time frame for our panel regressions). For each variable, the table shows the mean, standard deviation, and the 5^{th} and 95^{th} percentiles of the pooled observations. The the one-lag autocorrelation coefficient $AR(1)$ is calculated at the country level and then averaged across all countries. N is the number of country-month observations in the pooled sample. Variable definitions are given in Table 5. The text measures, which except entropy are normalized to mean zero/unit variance in the panels, are not normalized here.

Emerging markets data summary

	mean	sd	5%	95%	AR(1)	N
$return$	1.044	9.316	-13.497	15.299	0.137	5549
$return^{12}$	16.796	42.392	-41.367	89.574		5549
$return^{24}$	35.742	73.016	-44.931	170.380		5297
$sigma$	21.482	14.282	7.085	47.108	0.538	5476
$drawdown^{12}$	17.398	17.373	-0.000	55.721		5583
$drawdown^{24}$	23.012	20.725	-0.000	64.639		5328
$retmi$	2.848	5.359	0.000	13.497	0.177	5549
$retpl$	3.891	5.993	0.000	15.299	0.122	5549
$value$	0.953	0.925	0.187	2.431	0.949	4917
gdp	4.157	4.836	-3.000	9.950	0.933	5964
$gdpdeflator$	7.133	8.984	-1.100	23.560	0.932	5964
cp	55.657	44.713	10.688	145.941	0.988	5964
dcp	1.153	7.918	-9.113	12.065	0.933	5964
$rate$	7.451	7.384	0.580	19.985	0.977	5964
$dexch$	0.365	3.695	-4.178	5.546	0.097	5955
pre	0.162	0.369	0.000	1.000	0.810	5964
$post$	0.157	0.363	0.000	1.000	0.802	5964
$entropy$	2.429	0.171	2.111	2.649	0.790	5964
$artcount$	26.022	30.146	3.200	73.349	0.763	5964
$sMkt$	-0.005	0.002	-0.009	-0.002	0.693	5964
$fMkt$	0.334	0.080	0.195	0.460	0.637	5964
$sGovt$	-0.007	0.004	-0.016	-0.002	0.565	5964
$fGovt$	0.290	0.096	0.163	0.480	0.596	5964
$sCorp$	-0.002	0.001	-0.004	-0.001	0.635	5964
$fCorp$	0.169	0.034	0.124	0.229	0.587	5964
$sComms$	-0.002	0.001	-0.004	-0.001	0.482	5964
$fComms$	0.159	0.064	0.082	0.289	0.696	5964
$sMacro$	-0.001	0.000	-0.002	-0.000	0.636	5964
$fMacro$	0.048	0.028	0.025	0.105	0.620	5964

Developed markets data summary

	mean	sd	5%	95%	AR(1)	N
$return$	0.646	6.671	-10.939	10.783	0.113	4803
$return^{12}$	8.983	27.341	-38.063	51.641		4792
$return^{24}$	18.835	42.251	-42.492	92.441		4757
$sigma$	19.197	11.427	7.470	41.669	0.627	4815
$drawdown^{12}$	15.299	15.461	-0.000	50.593		4817
$drawdown^{24}$	20.818	19.326	-0.000	61.306		4802
$retmi$	2.188	4.096	0.000	10.939	0.211	4803
$retpl$	2.835	3.914	0.000	10.783	0.049	4803
$value$	0.936	0.624	0.317	1.836	0.982	4432
gdp	2.083	3.450	-3.490	6.680	0.933	4898
$gdpdeflator$	1.917	2.762	-1.473	5.650	0.925	4898
cp	113.275	41.235	55.687	188.754	0.990	4898
dcp	2.503	11.916	-9.540	13.744	0.926	4898
$rate$	3.947	2.025	1.030	6.221	0.986	4899
$dexch$	0.025	3.037	-4.773	4.782	0.009	4686
pre	0.126	0.333	0.000	1.000	0.813	4899
$post$	0.122	0.327	0.000	1.000	0.806	4899
$entropy$	2.455	0.170	2.125	2.666	0.841	4899
$artcount$	106.700	214.462	12.787	316.793	0.547	4899
$sMkt$	-0.004	0.002	-0.007	-0.002	0.745	4899
$fMkt$	0.350	0.055	0.271	0.450	0.715	4899
$sGovt$	-0.004	0.002	-0.008	-0.001	0.619	4899
$fGovt$	0.236	0.055	0.150	0.328	0.613	4899
$sCorp$	-0.002	0.001	-0.003	-0.001	0.636	4899
$fCorp$	0.220	0.050	0.159	0.314	0.750	4899
$sComms$	-0.000	0.000	-0.001	-0.000	0.464	4899
$fComms$	0.027	0.013	0.012	0.052	0.554	4899
$sCredit$	-0.002	0.001	-0.004	-0.001	0.682	4899
$fCredit$	0.167	0.023	0.137	0.208	0.576	4899

Table 7

Panel regressions for developed market next 12-month returns. Results are shown for the base specification of the model, which excludes the text-based measures, and two text specifications, one that includes context-specific sentiment (in column Sent) and another that includes context-specific sentiment interacted with entropy (in column SentEnt). Results are reported for the entire sample as well as for the early and late subsamples. Observations are monthly. All text measures except *entropy* are normalized to unit variance. Variables (all of which are defined in Table 5) include realized volatility (*sigma*), monthly returns (*return*), negative portfolio of returns (*retmi*), ratio of five-year ago to current index level (*value*), year-over-year GDP growth (*gdp*), year-over-year inflation (*gdpdeflator*), private sector credit to GDP (*cp*), year-over-year change in *cp* (*dcp*), local currency interest rate (*rate*), percent US$ appreciation against local currency (*dexch*), pre- and postelection dummies (*pre* and *post*), country-month entropy (*entropy*), number of articles per country per month (*artcount*), and country-level article sentiment ($s[Topic]$) and frequency ($f[Topic]$) for *Topic* in markets, government, corporate sector, commodities, credit (DM), and macro (EM). All panels include country fixed effects, and standard errors are clustered either by time or by time and country (labeled "both"); the *stderr* row indicates the type of calculation. Coefficients labeled with "***", "**", and "*" are significant at the 1%, 5%, and 10% levels, respectively.

Developed markets: Forecasting panel for next 12-month returns

	Base	Sent	SentEnt	Base	Sent	SentEnt	Base	Sent	SentEnt
$sigma_{t-1}$	0.134	0.191	0.198	0.131	0.087	0.083	0.329*	0.372**	0.360**
$sigma_{t-2}$	0.037	0.076	0.079	-0.078	-0.011	-0.005	0.298	0.347*	0.332*
$return_{t-1}$	0.273	0.152	0.147	0.109	-0.026	-0.028	0.209	0.089	0.092
$return_{t-2}$	0.024	-0.070	-0.072	0.148	0.093	0.091	-0.110	-0.174	-0.173
$value_{t-1}$	21.240***	22.455***	22.563***	14.876***	13.254**	13.281**	25.244***	25.775***	25.869***
gdp_{t-1}	-0.473	-0.859	-0.904	0.232	0.269	0.224	-2.066***	-1.947***	-1.913***
$gdpdeflator_{t-1}$	0.878	0.486	0.465	0.580	0.499	0.460	0.056	0.119	0.126
cp_{t-1}	-0.276***	-0.220**	-0.224**	0.042	-0.016	-0.009	-0.301**	-0.292**	-0.298**
dcp_{t-1}	0.075	0.026	0.022	-0.177	-0.158	-0.158	0.270**	0.252*	0.257**
$rate_{t-1}$	-3.705***	-5.306***	-5.398***	-14.708***	-14.360***	-14.257***	-4.926***	-5.307***	-5.382***
$dexch_{t-1}$	0.392	0.418	0.414	-0.581	-0.705*	-0.711*	0.779	0.793	0.790
pre	-0.923	-2.502	-2.553	-3.088	-3.008	-3.003	2.835	2.010	2.102
post	-0.404	-1.365	-1.413	-2.708	-3.033	-2.968	2.283	1.991	2.067
$entropy_{t-1}$		20.794	23.123		-23.426	-21.945		23.165**	22.670*
$artcount_{t-1}$		-0.667	-0.709		0.379	0.299		-0.022	-0.008
$sMkt_{t-1}$		2.636	2.914		-1.356	-2.132		5.144*	5.637**
$fMkt_{t-1}$		0.755	0.864		-4.034	-4.135		1.245	1.231
$sGovt_{t-1}$		-3.268	-3.460*		-0.640	-0.965		-3.915*	-2.821
$fGovt_{t-1}$		-3.690	-3.840		-3.589	-3.809		-6.273	-6.006
$sCorp_{t-1}$		-2.767	-1.862		3.208	3.887*		-2.709	-2.885*
$fCorp_{t-1}$		-5.894*	-5.571*		-5.901	-5.760		-0.041	0.136
$sComms_{t-1}$		1.112	0.914		0.213	0.451		0.531	0.416
$fComms_{t-1}$		0.864	0.850		-1.537	-1.348		1.543	1.528
$sCredit_{t-1}$		3.764*	3.234		-1.007	-0.893		-0.743	-1.754
$fCredit_{t-1}$		0.092	0.084		0.853	0.832		-1.091	-1.295
$R2$	0.164	0.215	0.215	0.267	0.3	0.301	0.456	0.476	0.476
start	Apr 1998	May 1998	May 1998	Apr 1998	May 1998	May 1998	Mar 2007	Mar 2007	Mar 2007
end	Dec 2015	Dec 2015	Dec 2015	Feb 2007	Feb 2007	Feb 2007	Dec 2015	Dec 2015	Dec 2015
Nobs	4411	4395	4395	2003	1987	1987	2408	2408	2408
stderr	both	both	both	both	both	both	both	both	both

Table 8

Panel regressions for emerging market next 12-month returns. Results are shown for the base specification of the model, which excludes the text-based measures, and two text specifications, one that includes context-specific sentiment (in column Sent) and another that includes context-specific sentiment interacted with entropy (in column SentEnt). Results are reported for the entire sample as well as for the early and late subsamples. Observations are monthly. All text measures except *entropy* are normalized to unit variance. Variables (all of which are defined in Table 5) include realized volatility (*sigma*), monthly returns (*return*), negative portfolio of returns (*retmi*), ratio of five-year ago to current index level (*value*), year-over-year GDP growth (*gdp*), year-over-year inflation (*gdpdeflator*), private sector credit to GDP (*cp*), year-over-year change in *cp* (*dcp*), local currency interest rate (*rate*), percent US$ appreciation against local currency (*dexch*), pre- and postelection dummies (*pre* and *post*), country-month entropy (*entropy*), number of articles per country per month (*artcount*), and country-level article sentiment ($s[Topic]$) and frequency ($f[Topic]$) for $Topic$ in markets, government, corporate sector, commodities, credit (DM), and macro (EM). All panels include country fixed effects, and standard errors are clustered either by time or by time and country (labeled "both"); the *stderr* row indicates the type of calculation. Coefficients labeled with "***", "**", and "*" are significant at the 1%, 5%, and 10% levels, respectively.

Emerging markets: Forecasting panel for next 12-month returns

	Base	Sent	SentEnt	Base	Sent	SentEnt	Base	Sent	SentEnt
$sigma_{t-1}$	0.189	0.257**	0.256**	-0.045	0.255	0.254	0.486***	0.422***	0.417***
$sigma_{t-2}$	0.320**	0.373***	0.369***	-0.042	0.154	0.150	0.707***	0.662***	0.655***
$return_{t-1}$	0.298	0.199	0.206	0.091	-0.197	-0.184	0.440*	0.450**	0.464**
$return_{t-2}$	0.037	0.021	0.028	-0.099	-0.210	-0.200	0.111	0.131	0.142
$value_{t-1}$	4.061	4.075	4.181	3.657	8.939**	9.021**	9.371	7.417	7.521
gdp_{t-1}	-0.972	-0.884	-0.880	-0.849	-1.918*	-1.889*	-1.614*	-1.194*	-1.175
$gdpdeflator_{t-1}$	0.503*	0.387	0.389	1.078*	0.160	0.182	-0.267	-0.058	-0.061
cp_{t-1}	-0.487***	-0.311***	-0.307***	-0.320	-0.224	-0.212	-0.030	-0.125	-0.128
dcp_{t-1}	0.234	0.230	0.215	0.481	0.348	0.360	0.082	0.226	0.183
$rate_{t-1}$	-0.732*	-0.763*	-0.764*	-1.464*	-0.300	-0.317	-0.954	-1.146	-0.991
$dexch_{t-1}$	0.550	0.576	0.575	0.039	0.221	0.197	1.001**	0.811*	0.820*
pre	-0.277	-0.214	-0.203	-2.691	-5.186	-5.340	-1.580	-1.264	-1.258
$post$	-4.406	-4.149	-4.216	1.401	-0.147	-0.346	-7.895***	-7.018***	-7.167***
$entropy_{t-1}$		1.907	0.999		-48.869	-45.438		6.525	3.590
$artcount_{t-1}$		-5.785***	-5.788***		-10.478***	-10.520***		-1.198	-1.009
$sMkt_{t-1}$		3.076	3.117		3.776	2.705		1.090	1.290
$fMkt_{t-1}$		-3.584	-3.493		-11.018**	-11.430**		8.828***	8.925***
$sGovt_{t-1}$		-1.585	-0.717		-0.671	-0.884		-0.046	0.064
$fGovt_{t-1}$		-6.820	-6.200		-10.377**	-10.517**		3.901	3.733
$sCorp_{t-1}$		-7.044**	-6.705**		-1.812	-0.842		-8.206**	-7.954**
$fCorp_{t-1}$		-7.703***	-7.661***		-7.280*	-6.887*		2.577	2.483
$sComms_{t-1}$		1.260	1.139		3.433	3.826		0.037	-0.661
$fComms_{t-1}$		1.802	1.732		5.054	5.240		3.193	2.766
$sMacro_{t-1}$		2.778	1.765		2.515	1.515		-1.895	-1.625
$fMacro_{t-1}$		5.865**	5.503**		3.544	3.078		2.863	3.042
$R2$	0.0697	0.127	0.125	0.0213	0.13	0.128	0.212	0.264	0.263
start	Apr 1998	May 1998	May 1998	Apr 1998	May 1998	May 1998	Mar 2007	Mar 2007	Mar 2007
end	Dec 2015	Dec 2015	Dec 2015	Feb 2007	Feb 2007	Feb 2007	Dec 2015	Dec 2015	Dec 2015
Nobs	4853	4839	4839	2100	2086	2086	2753	2753	2753
stderr	both	both	both	both	both	both	both	both	both

Table 9

Panel regressions for developed market volatility. Results are shown for the base specification of the model, which excludes the text-based measures, and two text specifications, one that includes context-specific sentiment (in column Sent) and another that includes context-specific sentiment interacted with entropy (in column SentEnt). Results are reported for the entire sample as well as for the early and late subsamples. Observations are monthly. All text measures except *entropy* are normalized to unit variance. Variables (all of which are defined in Table 5) include realized volatility (*sigma*), monthly returns (*return*), negative portfolio of returns (*retmi*), ratio of five-year ago to current index level (*value*), year-over-year GDP growth (*gdp*), year-over-year inflation (*gdpdeflator*), private sector credit to GDP (*cp*), year-over-year change in *cp* (*dcp*), local currency interest rate (*rate*), percent US$ appreciation against local currency (*dexch*), pre- and postelection dummies (*pre* and *post*), country-month entropy (*entropy*), number of articles per country per month (*artcount*), and country-level article sentiment ($s[Topic]$) and frequency ($f[Topic]$) for *Topic* in markets, government, corporate sector, commodities, credit (DM), and macro (EM). All panels include country fixed effects, and standard errors are clustered either by time or by time and country (labeled "both"); the *stderr* row indicates the type of calculation. Coefficients labeled with "***", "**", and "*" are significant at the 1%, 5%, and 10% levels, respectively.

Developed markets: Forecasting panel for volatility

	Base	Sent	SentEnt	Base	Sent	SentEnt	Base	Sent	SentEnt
$sigma_{t-1}$	0.411***	0.386***	0.384***	0.354***	0.320***	0.319***	0.400***	0.372***	0.370***
$sigma_{t-2}$	0.140**	0.126**	0.126**	0.182***	0.170***	0.170***	0.075	0.044	0.044
$retmi_{t-1}$	0.664***	0.631***	0.631***	0.652***	0.605***	0.610***	0.614**	0.593**	0.592**
$retmi_{t-2}$	0.000	-0.016	-0.018	-0.074	-0.077	-0.075	0.095	0.118	0.115
$value_{t-1}$	-2.127***	-2.492***	-2.531***	-1.880*	-2.602***	-2.601***	-2.391**	-2.450**	-2.484**
gdp_{t-1}	-0.184*	-0.136	-0.129	-0.115	-0.046	-0.044	-0.176	-0.143	-0.143
$gdpdeflator_{t-1}$	-0.018	0.020	0.023	0.059	0.096	0.097	0.174	0.152	0.149
cp_{t-1}	0.010	0.006	0.006	-0.032*	-0.023	-0.022	0.030	0.033*	0.031
dcp_{t-1}	-0.018	-0.012	-0.011	0.010	0.011	0.011	-0.048**	-0.047*	-0.045*
$rate_{t-1}$	0.759***	0.835***	0.852***	1.684***	1.605***	1.615***	0.797***	0.836**	0.852**
$dexch_{t-1}$	-0.233	-0.235	-0.237	-0.524***	-0.511***	-0.516***	-0.062	-0.071	-0.072
pre	0.025	-0.006	0.003	-0.087	-0.021	-0.020	0.006	-0.179	-0.147
$post$	0.063	0.059	0.081	0.687*	0.843**	0.846**	-0.479	-0.531	-0.513
$entropy_{t-1}$		-2.188	-2.689		1.050	0.421		-1.974	-2.661
$artcount_{t-1}$		0.054	0.070		0.153	0.143		-0.600	-0.572
$sMkt_{t-1}$		-0.739	-0.793		-1.295*	-1.218*		-0.475	-0.616
$fMkt_{t-1}$		-0.341	-0.339		-0.348	-0.316		-0.022	-0.067
$sGovt_{t-1}$		0.616	0.565		-0.211	-0.206		1.204	1.067
$fGovt_{t-1}$		0.313	0.285		-0.165	-0.233		0.749	0.702
$sCorp_{t-1}$		0.153	0.154		0.194	0.197		-0.668	-0.614
$fCorp_{t-1}$		0.074	0.044		-0.344	-0.396		-0.416	-0.449
$sComms_{t-1}$		0.093	0.130		-0.135	-0.105		0.323	0.339
$fComms_{t-1}$		-0.240	-0.221		-0.224	-0.230		-0.295	-0.275
$sCredit_{t-1}$		-0.604	-0.647		0.087	-0.029		-0.402	-0.373
$fCredit_{t-1}$		-0.087	-0.107		0.334	0.309		-0.417	-0.424
$R2$	0.473	0.479	0.479	0.466	0.475	0.475	0.453	0.459	0.459
start	Apr 1998	May 1998	May 1998	Apr 1998	May 1998	May 1998	Mar 2007	Mar 2007	Mar 2007
end	Dec 2015	Dec 2015	Dec 2015	Feb 2007	Feb 2007	Feb 2007	Dec 2015	Dec 2015	Dec 2015
Nobs	4422	4406	4406	2003	1987	1987	2419	2419	2419
stderr	by time	by time	by time	by time	by time	by time	by time	by time	by time

Table 10

Panel regressions for emerging market volatility. Results are shown for the base specification of the model, which excludes the text-based measures, and two text specifications, one that includes context-specific sentiment (in column Sent) and another that includes context-specific sentiment interacted with entropy (in column SentEnt). Results are reported for the entire sample as well as for the early and late subsamples. Observations are monthly. All text measures except *entropy* are normalized to unit variance. Variables (all of which are defined in Table 5) include realized volatility (*sigma*), monthly returns (*return*), negative portfolio of returns (*retmi*), ratio of five-year ago to current index level (*value*), year-over-year GDP growth (*gdp*), year-over-year inflation (*gdpdeflator*), private sector credit to GDP (*cp*), year-over-year change in *cp* (*dcp*), local currency interest rate (*rate*), percent US$ appreciation against local currency (*dexch*), pre- and postelection dummies (*pre* and *post*), country-month entropy (*entropy*), number of articles per country per month (*artcount*), and country-level article sentiment ($s[Topic]$) and frequency ($f[Topic]$) for *Topic* in markets, government, corporate sector, commodities, credit (DM), and macro (EM). All panels include country fixed effects, and standard errors are clustered either by time or by time and country (labeled "both"); the *stderr* row indicates the type of calculation. Coefficients labeled with "***", "**", and "*" are significant at the 1%, 5%, and 10% levels, respectively.

Emerging markets: Forecasting panel for volatility

	Base	Sent	SentEnt	Base	Sent	SentEnt	Base	Sent	SentEnt
$sigma_{t-1}$	0.356***	0.352***	0.352***	0.280***	0.264***	0.264***	0.413***	0.402***	0.402***
$sigma_{t-2}$	0.116***	0.114***	0.113***	0.126***	0.117***	0.116***	0.044	0.030	0.029
$retmi_{t-1}$	0.529***	0.526***	0.524***	0.641***	0.650***	0.648***	0.340***	0.330***	0.329***
$retmi_{t-2}$	-0.024	-0.026	-0.026	-0.132*	-0.125	-0.126	0.098	0.113	0.113
$value_{t-1}$	-0.731**	-0.722***	-0.739***	-0.274	-0.541*	-0.551*	-1.650***	-1.450***	-1.467***
gdp_{t-1}	-0.019	-0.017	-0.015	0.010	0.078	0.079	-0.032	-0.044	-0.040
$gdpdeflator_{t-1}$	0.044	0.057	0.057	-0.005	0.064	0.066	0.088	0.102	0.102
cp_{t-1}	-0.035**	-0.047***	-0.047***	0.004	0.010	0.011	-0.076**	-0.083***	-0.084***
dcp_{t-1}	0.063***	0.062***	0.061***	0.050	0.054	0.054	0.076**	0.085**	0.084**
$rate_{t-1}$	0.367***	0.385***	0.384***	0.447***	0.399***	0.395***	0.625***	0.714***	0.706***
$dexch_{t-1}$	-0.186**	-0.178**	-0.178**	-0.242**	-0.213*	-0.215*	-0.032	-0.031	-0.032
pre	0.211	0.281	0.281	-0.091	0.342	0.345	0.718	0.762	0.753
$post$	-0.012	-0.037	-0.029	0.540	0.709	0.720	-0.135	-0.253	-0.238
$entropy_{t-1}$		0.288	0.260		10.974***	10.982***		-6.050	-6.123
$artcount_{t-1}$		0.589**	0.578*		0.261	0.265		1.108**	1.081**
$sMkt_{t-1}$		0.443	0.432		0.989*	0.943		0.215	0.185
$fMkt_{t-1}$		0.786	0.771*		1.131*	1.099*		0.059	0.021
$sGovt_{t-1}$		0.323	0.395		-0.230	-0.219		0.486	0.666
$fGovt_{t-1}$		1.065*	1.107**		0.696	0.684		0.515	0.617
$sCorp_{t-1}$		-0.068	-0.144		-0.115	-0.163		-0.166	-0.335
$fCorp_{t-1}$		0.683*	0.664*		1.445***	1.429***		-0.010	-0.057
$sComms_{t-1}$		-0.344	-0.305		-0.416	-0.304		-0.290	-0.228
$fComms_{t-1}$		0.358	0.381		0.266	0.324		0.514	0.543*
$sMacro_{t-1}$		-0.088	-0.101		-0.114	-0.168		0.124	0.120
$fMacro_{t-1}$		0.031	0.021		0.571	0.534		-0.654**	-0.677**
$R2$	0.374	0.376	0.376	0.318	0.328	0.328	0.384	0.391	0.392
start	Apr 1998	May 1998	May 1998	Apr 1998	May 1998	May 1998	Mar 2007	Mar 2007	Mar 2007
end	Dec 2015	Dec 2015	Dec 2015	Feb 2007	Feb 2007	Feb 2007	Dec 2015	Dec 2015	Dec 2015
Nobs	4844	4830	4830	2092	2078	2078	2752	2752	2752
stderr	by time	by time	by time	by time	by time	by time	by time	by time	by time

Table 11

Panel regressions for developed market drawdowns. Results are shown for the base specification of the model, which excludes the text-based measures, and two text specifications, one that includes context-specific sentiment (in column Sent) and another that includes context-specific sentiment interacted with entropy (in column SentEnt). Results are reported for the entire sample as well as for the early and late subsamples. Observations are monthly. All text measures except *entropy* are normalized to unit variance. Variables (all of which are defined in Table 5) include realized volatility (*sigma*), monthly returns (*return*), negative portfolio of returns (*retmi*), ratio of five-year ago to current index level (*value*), year-over-year GDP growth (*gdp*), year-over-year inflation (*gdpdeflator*), private sector credit to GDP (*cp*), year-over-year change in *cp* (*dcp*), local currency interest rate (*rate*), percent US$ appreciation against local currency (*dexch*), pre- and postelection dummies (*pre* and *post*), country-month entropy (*entropy*), number of articles per country per month (*artcount*), and country-level article sentiment (*s*[*Topic*]) and frequency (*f*[*Topic*]) for *Topic* in markets, government, corporate sector, commodities, credit (DM), and macro (EM). All panels include country fixed effects, and standard errors are clustered either by time or by time and country (labeled "both"); the *stderr* row indicates the type of calculation. Coefficients labeled with "***", "**", and "*" are significant at the 1%, 5%, and 10% levels, respectively.

Developed markets: Forecasting panel for drawdowns

	Base	Sent	SentEnt	Base	Sent	SentEnt	Base	Sent	SentEnt
$sigma_{t-1}$	0.098	0.016	0.011	0.068	0.033	0.035	-0.005	-0.099	-0.094
$sigma_{t-2}$	0.171**	0.110	0.112	0.156**	0.097	0.095	0.081	-0.014	-0.003
$return_{t-1}$	-0.392***	-0.261**	-0.264**	-0.186*	-0.091	-0.095	-0.414***	-0.268*	-0.272*
$return_{t-2}$	-0.073	-0.008	-0.007	-0.031	-0.011	-0.010	-0.065	-0.016	-0.012
$value_{t-1}$	-12.271***	-13.805***	-13.878***	-6.073***	-6.611***	-6.589***	-15.672***	-16.588***	-16.749***
gdp_{t-1}	0.165	0.412	0.445	-0.092	0.040	0.071	0.792**	0.794**	0.780**
$gdpdeflator_{t-1}$	-0.286	-0.081	-0.069	-0.349	-0.241	-0.218	0.361	0.156	0.149
cp_{t-1}	0.185***	0.153***	0.156***	-0.008	0.033	0.027	0.149**	0.144**	0.144**
dcp_{t-1}	-0.070	-0.038	-0.034	0.074	0.068	0.068	-0.144	-0.115	-0.113
$rate_{t-1}$	3.057***	3.885***	3.964***	7.657***	7.476***	7.403***	4.185***	4.427***	4.535***
$dexch_{t-1}$	-0.398	-0.334	-0.343	-0.094	-0.001	-0.008	-0.471	-0.424	-0.430
pre	1.376	1.947	1.981	2.030	2.089	2.096	0.065	0.250	0.240
$post$	0.128	0.351	0.391	0.484	0.848	0.822	-0.572	-1.020	-1.060
$entropy_{t-1}$		-11.165*	-13.357**		11.666*	9.826		-12.148*	-13.010**
$artcount_{t-1}$		0.149	0.217		-0.057	-0.001		-0.840	-0.819
$sMkt_{t-1}$		-4.132***	-4.348***		-0.499	0.080		-6.765***	-7.292***
$fMkt_{t-1}$		-0.866	-0.893		0.589	0.715		-1.330	-1.388
$sGovt_{t-1}$		4.576***	4.303***		0.551	0.708		6.013***	4.945***
$fGovt_{t-1}$		3.860***	3.818***		0.857	0.938		5.574***	5.261***
$sCorp_{t-1}$		1.195	0.632		-1.537*	-2.011**		0.343	0.484
$fCorp_{t-1}$		2.614**	2.323**		1.381	1.267		-0.062	-0.310
$sComms_{t-1}$		-0.783	-0.509		-0.654	-0.681		0.281	0.410
$fComms_{t-1}$		-0.098	-0.022		0.386	0.291		0.340	0.416
$sCredit_{t-1}$		-1.654	-1.207		-0.252	-0.404		1.610*	2.341***
$fCredit_{t-1}$		0.382	0.379		-0.429	-0.424		0.868	0.944
$R2$	0.263	0.323	0.322	0.404	0.45	0.453	0.389	0.448	0.448
start	Apr 1998	May 1998	May 1998	Apr 1998	May 1998	May 1998	Mar 2007	Mar 2007	Mar 2007
end	Dec 2015	Dec 2015	Dec 2015	Feb 2007	Feb 2007	Feb 2007	Dec 2015	Dec 2015	Dec 2015
Nobs	4422	4406	4406	2003	1987	1987	2419	2419	2419
stderr	both	both	both	both	both	both	both	both	both

Table 12

Panel regressions for emerging market drawdowns. Results are shown for the base specification of the model, which excludes the text-based measures, and two text specifications, one that includes context-specific sentiment (in column Sent) and another that includes context-specific sentiment interacted with entropy (in column SentEnt). Results are reported for the entire sample as well as for the early and late subsamples. Observations are monthly. All text measures except *entropy* are normalized to unit variance. Variables (all of which are defined in Table 5) include realized volatility (*sigma*), monthly returns (*return*), negative portfolio of returns (*retmi*), ratio of five-year ago to current index level (*value*), year-over-year GDP growth (*gdp*), year-over-year inflation (*gdpdeflator*), private sector credit to GDP (*cp*), year-over-year change in *cp* (*dcp*), local currency interest rate (*rate*), percent US$ appreciation against local currency (*dexch*), pre- and postelection dummies (*pre* and *post*), country-month entropy (*entropy*), number of articles per country per month (*artcount*), and country-level article sentiment ($s[Topic]$) and frequency ($f[Topic]$) for *Topic* in markets, government, corporate sector, commodities, credit (DM), and macro (EM). All panels include country fixed effects, and standard errors are clustered either by time or by time and country (labeled "both"); the *stderr* row indicates the type of calculation. Coefficients labeled with "***", "**", and "*" are significant at the 1%, 5%, and 10% levels, respectively.

Emerging markets: Forecasting panel for drawdowns

	Base	Sent	SentEnt	Base	Sent	SentEnt	Base	Sent	SentEnt
$sigma_{t-1}$	0.092	0.066	0.069	0.107*	0.007	0.009	-0.022	-0.027	-0.024
$sigma_{t-2}$	0.023	-0.011	-0.008	0.099	0.014	0.014	-0.134*	-0.155**	-0.153**
$return_{t-1}$	-0.221***	-0.163**	-0.172**	-0.084	-0.012	-0.017	-0.298***	-0.269***	-0.282***
$return_{t-2}$	-0.076	-0.066	-0.071	0.054	0.062	0.059	-0.166	-0.164*	-0.171*
$value_{t-1}$	-2.105*	-1.885**	-1.964**	0.192	-1.698**	-1.712**	-6.564**	-5.499**	-5.586**
gdp_{t-1}	0.364**	0.336	0.332	0.386	0.739***	0.731***	0.588**	0.480**	0.484**
$gdpdeflator_{t-1}$	0.060	0.057	0.056	-0.113	0.214	0.211	0.418*	0.342*	0.350*
cp_{t-1}	0.124**	0.044	0.042	0.102	0.057	0.055	-0.084	-0.060	-0.062
dcp_{t-1}	0.070	0.038	0.041	-0.021	0.018	0.014	0.143	0.072	0.082
$rate_{t-1}$	0.531***	0.575***	0.573***	0.651***	0.196	0.195	1.933***	2.082***	2.060***
$dexch_{t-1}$	-0.305**	-0.318**	-0.313**	-0.059	-0.095	-0.090	-0.442**	-0.371*	-0.370*
pre	-1.259	-1.335	-1.333	-0.004	1.313	1.371	-0.624	-0.828	-0.867
$post$	0.166	-0.225	-0.162	-0.620	0.175	0.258	0.991	0.076	0.163
$entropy_{t-1}$		-5.269	-5.015		34.498***	33.656***		-22.328**	-21.475**
$artcount_{t-1}$		2.791***	2.804***		3.908***	3.930***		2.487**	2.420**
$sMkt_{t-1}$		-1.787	-1.391		0.331	0.822		-0.933	-0.589
$fMkt_{t-1}$		0.066	0.143		4.248**	4.404**		-5.188***	-5.206***
$sGovt_{t-1}$		2.277**	1.995**		0.175	0.233		2.439	2.455*
$fGovt_{t-1}$		3.787***	3.561***		3.534**	3.525**		0.057	0.024
$sCorp_{t-1}$		2.872**	2.564**		-0.541	-0.735		3.309***	2.854**
$fCorp_{t-1}$		3.011***	2.906***		3.278***	3.143***		-1.427	-1.567
$sComms_{t-1}$		-1.422*	-1.309*		-1.628***	-1.707***		-1.407	-0.994
$fComms_{t-1}$		-0.421	-0.379		-0.699	-0.767		-1.182	-1.015
$sMacro_{t-1}$		-1.283	-1.132		-0.235	-0.284		-0.571	-0.848
$fMacro_{t-1}$		-1.941**	-1.922**		0.155	0.146		-2.195**	-2.356**
$R2$	0.0616	0.121	0.118	0.0798	0.221	0.223	0.173	0.252	0.249
start	Apr 1998	May 1998	May 1998	Apr 1998	May 1998	May 1998	Mar 2007	Mar 2007	Mar 2007
end	Dec 2015	Dec 2015	Dec 2015	Feb 2007	Feb 2007	Feb 2007	Dec 2015	Dec 2015	Dec 2015
Nobs	4853	4839	4839	2100	2086	2086	2753	2753	2753
$stderr$	both	both	both	both	both	both	both	both	both

Table 13
Regression of the BBD country-level policy uncertainty measures in month t on one-month lags of macro control and one-month lags of our text measures. Results are shown for the EM sample (Brazil, Chile, China, India, Mexico, Russia, and South Korea) and the DM sample (USA, Canada, Germany, UK, Italy, France, Spain, Netherlands, Japan, Australia, and Ireland). All text measures except *entropy* are normalized to unit variance. Variables (all of which are defined in Table 5) include realized volatility (*sigma*), monthly returns (*return*), negative portfolio of returns (*retmi*), ratio of five-year ago to current index level (*value*), year-over-year GDP growth (*gdp*), year-over-year inflation ($gdpdeflator$), private sector credit to GDP (*cp*), year-over-year change in *cp* (*dcp*), local currency interest rate (*rate*), percent US$ appreciation against local currency (*dexch*), pre- and postelection dummies (*pre* and *post*), country-month entropy (*entropy*), number of articles per country per month (*artcount*), and country-level article sentiment ($s[Topic]$) and frequency ($f[Topic]$) for $Topic$ in markets, government, corporate sector, commodities, credit (DM), and macro (EM). All panels include country fixed effects, and standard errors are clustered either by time or by time and country (labeled "both"); the *stderr* row indicates the type of calculation. Coefficients labeled with "***", "**", and "*" are significant at the 1%, 5%, and 10% levels, respectively.

	Base EM	Full EM	Base DM	Full DM
gdp_{t-1}	-0.021**	0.002	-0.083***	-0.033***
$gdpdeflator_{t-1}$	0.002	0.000	-0.024*	-0.011
cp_{t-1}	0.015***	0.014***	0.001	0.000
dcp_{t-1}	-0.003	-0.005	-0.008***	-0.004***
$rate_{t-1}$	0.039***	0.028***	-0.175***	-0.227***
$dexch_{t-1}$	0.028***	0.012	0.017	-0.005
pre	0.172**	0.166**	0.033	0.057
$post$	0.137*	0.107	0.112**	0.046
$entropy_{t-1}$		-1.159***		-0.151
$artcount_{t-1}$		0.064*		-0.003
$sMkt_{t-1}$		-0.074		-0.374***
$fMkt_{t-1}$		-0.062		0.021
$sGovt_{t-1}$		-0.060		0.095*
$fGovt_{t-1}$		0.011		0.199**
$sCorp_{t-1}$		-0.094*		0.184***
$fCorp_{t-1}$		0.044		0.204***
$sComms_{t-1}$		-0.026		-0.024
$fComms_{t-1}$		0.107		-0.001
$sMacro_{t-1}$		-0.184***		
$fMacro_{t-1}$		-0.134***		
$sCredit_{t-1}$				-0.290***
$fCredit_{t-1}$				0.032
$R2$	0.102	0.192	0.134	0.347
start	Apr 1998	May 1998	Apr 1998	May 1998
end	Dec 2015	Dec 2015	Dec 2015	Dec 2015
Nobs	1433	1427	2249	2240
stderr	by time	by time	by time	by time

Table 14

Results of the panel regressions from Tables A3, 7, 9, and 11 modified to include the BBD policy uncertainty index as an explanatory variable. Results are shown for the full period regressions. Countries include: USA, Canada, Germany, UK, Italy, France, Spain, Netherlands, Japan, Australia, and Ireland. All text measures except *entropy* are normalized to unit variance. Variables (all of which are defined in Table 5) include realized volatility (*sigma*), monthly returns (*return*), negative portfolio of returns (*retmi*), ratio of five-year ago to current index level (*value*), year-over-year GDP growth (*gdp*), year-over-year inflation (*gdpdeflator*), private sector credit to GDP (*cp*), year-over-year change in *cp* (*dcp*), local currency interest rate (*rate*), percent US$ appreciation against local currency (*dexch*), pre- and postelection dummies (*pre* and *post*), country-month entropy (*entropy*), number of articles per country per month (*artcount*), and country-level article sentiment ($s[Topic]$) and frequency ($f[Topic]$) for *Topic* in markets, government, corporate sector, commodities, credit (DM), and macro (EM). All panels include country fixed effects, and standard errors are clustered either by time or by time and country (labeled "both"); the *stderr* row indicates the type of calculation. Coefficients labeled with "***", "**", and "*" are significant at the 1%, 5%, and 10% levels, respectively.

Comparison to Baker, Bloom and Davis for developed markets

	$return^{12m}$-1	$return^{12m}$-2	$return^{12m}$-3	$return^{12m}$-4	*sigma*-1	*sigma*-2	*sigma*-3	*sigma*-4	*drawdown*-1	*drawdown*-2	*drawdown*-3	*drawdown*-4
$sigma_{t-1}$	0.123	0.196	0.308**	0.318**	0.412***	0.399***	0.367***	0.364***	0.055	0.003	-0.103	-0.108
$sigma_{t-2}$	0.019	0.033	0.148	0.147	0.106*	0.102*	0.065	0.065	0.195**	0.185**	0.058	0.058
$return_{t-1}$	0.158	0.144	-0.139	-0.138					-0.430***	-0.420***	-0.192*	-0.193*
$return_{t-2}$	-0.020	-0.039	-0.154	-0.156					-0.113	-0.100	-0.024	-0.023
$retmi_{t-1}$					0.799***	0.780***	0.721***	0.717***				
$retmi_{t-2}$					0.023	0.014	0.009	0.007				
$value_{t-1}$	27.871***	30.436***	32.057***	32.439***	-2.778**	-3.356***	-3.743***	-3.933***	-15.693***	-17.532***	-19.774***	-19.983***
gdp_{t-1}	-1.181**	-1.206**	-1.347***	-1.348***	-0.271*	-0.265*	-0.199	-0.198	0.272	0.290	0.387	0.388
$gdpdeflator_{t-1}$	3.392***	3.432***	3.067***	3.086***	-0.112	-0.122	-0.162	-0.171	-1.065***	-1.094***	-1.091***	-1.101***
cp_{t-1}	-0.226**	-0.231**	-0.215*	-0.216*	0.015	0.016	0.015	0.016	0.177***	0.181***	0.163***	0.164***
dcp_{t-1}	0.242	0.241	0.235	0.235	-0.035**	-0.035**	-0.035**	-0.035**	-0.145*	-0.144*	-0.138*	-0.138*
$rate_{t-1}$	-2.999**	-3.846***	-3.661***	-3.842***	0.741***	0.938***	0.711**	0.802**	2.611***	3.218***	2.978***	3.077***
$dexch_{t-1}$	0.255	0.224	0.194	0.188	-0.317*	-0.304*	-0.297*	-0.291*	-0.377*	-0.355	-0.261	-0.258
pre	-2.327	-2.241	-3.383	-3.371	-0.146	-0.164	-0.383	-0.388	2.377	2.316	2.595	2.589
post	-3.742	-3.549	-4.058	-4.018	-0.021	-0.065	-0.128	-0.147	1.836	1.698	1.779	1.758
EPU_{t-1}		-2.875*		-0.602		0.661*		0.299		2.061***		0.329
$entropy_{t-1}$			-2.225	-1.926			-1.094	-1.251			0.960	0.797
$artcount_{t-1}$			-2.016*	-2.001*			-0.193	-0.201			0.567	0.559
$sMkt_{t-1}$			9.425***	9.269***			-1.024	-0.952			-7.016***	-6.930***
$fMkt_{t-1}$			-6.609	-6.654			-0.381	-0.358			1.034	1.059
$sGovt_{t-1}$			-6.648***	-6.622***			1.105	1.095			6.147***	6.133***
$fGovt_{t-1}$			-14.702**	-14.701***			0.240	0.244			7.219***	7.218***
$sCorp_{t-1}$			-3.048*	-2.933*			-0.157	-0.215			1.052	0.990
$fCorp_{t-1}$			-10.806***	-10.714***			-0.314	-0.358			3.663***	3.613***
$sComms_{t-1}$			0.222	0.212			0.010	0.015			-0.670	-0.664
$fComms_{t-1}$			-1.116	-1.130			-0.596*	-0.588*			0.383	0.391
$sCredit_{t-1}$			0.841	0.685			-0.790	-0.711			-0.813	-0.728
$fCredit_{t-1}$			-2.795	-2.798			-0.319	-0.316			1.114	1.116
R2	0.266	0.275	0.344	0.344	0.481	0.483	0.491	0.491	0.299	0.31	0.403	0.403
start	Apr 1998	Apr 1998	May 1998	May 1998	Apr 1998	Apr 1998	May 1998	May 1998	Apr 1998	Apr 1998	May 1998	May 1998
end	Dec 2015	Dec 2015	Dec 2015	Dec 2015	Dec 2015	Dec 2015	Dec 2015	Dec 2015	Dec 2015	Dec 2015	Dec 2015	Dec 2015
Nobs	2063	2063	2056	2056	2063	2063	2056	2056	2063	2063	2056	2056
stderr	both	both	both	both	by time	by time	by time	by time	both	both	both	both

Table 15

Results of the panel regressions from Tables A4, 8, 10, and 12 modified to include the BBD policy uncertainty index as an explanatory variable. Results are shown for the full period regressions. Countries include: Brazil, Chile, China, India, Mexico, Russia, and South Korea. All text measures except *entropy* are normalized to unit variance. Variables (all of which are defined in Table 5) include realized volatility (*sigma*), monthly returns (*return*), negative portfolio of returns (*retmi*), ratio of five-year ago to current index level (*value*), year-over-year GDP growth (*gdp*), year-over-year inflation (*gdpdeflator*), private sector credit to GDP (*cp*), year-over-year change in *cp* (*dcp*), local currency interest rate (*rate*), percent US$ appreciation against local currency (*dexch*), pre- and postelection dummies (*pre* and *post*), country-month entropy (*entropy*), number of articles per country per month (*artcount*), and country-level article sentiment (*s*[*Topic*]) and frequency (*f*[*Topic*]) for *Topic* in markets, government, corporate sector, commodities, credit (DM), and macro (EM). All panels include country fixed effects, and standard errors are clustered either by time or by time and country (labeled "both"); the *stderr* row indicates the type of calculation. Coefficients labeled with "***", "**", and "*" are significant at the 1%, 5%, and 10% levels, respectively.

Comparison to Baker, Bloom and Davis for emerging markets

	$return^{12m}$-1	$return^{12m}$-2	$return^{12m}$-3	$return^{12m}$-4	*sigma*-1	*sigma*-2	*sigma*-3	*sigma*-4	*drawdown*-1	*drawdown*-2	*drawdown*-3	*drawdown*-4
$sigma_{t-1}$	0.151	0.178	0.296***	0.329***	0.399***	0.401***	0.384***	0.384***	0.084	0.083	0.013	0.014
$sigma_{t-2}$	0.422**	0.423**	0.500***	0.500***	0.077	0.077	0.062	0.062	-0.016	-0.016	-0.056	-0.056
$return_{t-1}$	0.381	0.376*	0.229	0.240					-0.219***	-0.218***	-0.144	-0.143
$return_{t-2}$	0.159	0.127	0.056	0.027					-0.091	-0.089	-0.053	-0.054
$retmi_{t-1}$					0.738***	0.741***	0.761***	0.762***				
$retmi_{t-2}$					0.001	0.006	0.020	0.022				
$value_{t-1}$	12.427*	13.849*	16.731***	18.468***	-2.494***	-2.379***	-2.571***	-2.518***	-2.783	-2.831	-3.767**	-3.699**
gdp_{t-1}	-2.488	-2.531	-2.886*	-2.871*	-0.100	-0.103	-0.172	-0.171	0.847**	0.849**	0.945***	0.945***
$gdpdeflator_{t-1}$	0.899	0.933	0.144	0.156	0.190***	0.193***	0.210***	0.211***	0.150	0.149	0.309**	0.310**
cp_{t-1}	-0.584**	-0.523**	-0.346	-0.277	-0.093***	-0.088***	-0.157***	-0.155***	0.134	0.132*	-0.035	-0.032
dcp_{t-1}	-0.673**	-0.682*	-0.632**	-0.655**	0.106*	0.106*	0.107*	0.106*	0.368**	0.368**	0.372**	0.371**
$rate_{t-1}$	-1.366	-1.272	-1.300	-1.228	0.414***	0.422***	0.496***	0.498***	0.923**	0.919**	0.946**	0.949**
$dexch_{t-1}$	1.167**	1.226**	0.979*	1.028*	-0.308**	-0.304*	-0.248	-0.247	-0.541***	-0.543***	-0.450**	-0.448**
pre	-11.213	-10.777	-9.528	-9.019	0.888	0.924	0.819	0.835	4.273	4.258	3.651	3.672
post	-4.814	-4.711	-4.731	-4.741	0.907	0.919	0.897	0.897	1.701	1.697	1.572	1.571
EPU_{t-1}		-3.326		-3.849		-0.279		-0.122		0.112		-0.152
$entropy_{t-1}$			-6.820	-12.726			1.687	1.500			9.306	9.072
$artcount_{t-1}$			-4.460	-4.297			1.510**	1.516**			3.119	3.126
$sMkt_{t-1}$			-2.765	-3.325			1.495	1.479			0.695	0.672
$fMkt_{t-1}$			-0.531	-1.496			0.042	0.012			-0.731	-0.769
$sGovt_{t-1}$			-9.743	-9.908*			1.126*	1.121*			4.967***	4.961***
$fGovt_{t-1}$			-4.545	-5.004			0.572	0.558			2.160	2.141
$sCorp_{t-1}$			4.004	4.225			-0.749	-0.743			-0.349	-0.340
$fCorp_{t-1}$			-1.722	-1.992			0.228	0.219			-0.140	-0.151
$sComms_{t-1}$			0.069	0.074			-0.803	-0.803			-1.608	-1.607
$fComms_{t-1}$			8.955	8.865			-0.587	-0.590			-2.843	-2.847
$sMacro_{t-1}$			8.489	7.786			0.477	0.455			-3.264	-3.292
$fMacro_{t-1}$			10.010***	9.669***			0.091	0.080			-4.867***	-4.880***
R2	0.159	0.162	0.226	0.23	0.468	0.468	0.475	0.475	0.0998	0.0992	0.181	0.181
start	Apr 1998	Apr 1998	May 1998	May 1998	Apr 1998	Apr 1998	May 1998	May 1998	Apr 1998	Apr 1998	May 1998	May 1998
end	Dec 2015	Dec 2015	Dec 2015	Dec 2015	Dec 2015	Dec 2015	Dec 2015	Dec 2015	Dec 2015	Dec 2015	Dec 2015	Dec 2015
Nobs	1374	1374	1369	1369	1374	1374	1369	1369	1374	1374	1369	1369
stderr	both	both	both	both	by time	by time	by time	by time	both	both	both	both

Table 16

Panel A shows coefficients estimates from the Brusa, Ramadorai, and Verdelhan (2017) three-factor model estimated with monthly returns. Alphas are reported annualized in percent. T-statistics are shown in parentheses. Panel B shows t-tests of the differences of alphas between the CM and Base models. All standard errors in both panels are calculated using Newey-West with auto lag selection.

The three forecasting models are: Naive, which uses only in sample country fixed effects as the forecasting variables; Base, which includes lagged macroeconomic and lagged market variables as the regressors; and CM, which includes country specific article counts, entropy, sentiment, and frequency measures in addition to the variables from the Base model. These out-of-sample forecasts come from rolling elastic net regressions. The elastic net regressions are run over rolling 60-month windows, with weighting parameter, λ, chosen to minimize cross-validation error. We set $\alpha = 0.75$, which represents a 0.75 weight on the lasso penalty and a 0.25 weight on the ridge regression penalty function. The out-of-sample forecasts start in March 2003 and go to December 2015.

Panel A

Developed market strategy

Model	Alpha	Mkt.RF	fxcarry	fxusd
CM	8.816	0.443	-0.168	-0.413
	(1.438)	(2.702)	(-0.591)	(-2.756)
Base	6.809	0.570	-0.076	-0.395
	(1.380)	(4.286)	(-0.347)	(-2.784)
Naive	-2.765	0.666	0.123	-0.024
	(-0.678)	(4.635)	(0.749)	(-0.219)

Emerging market strategy

Model	Alpha	Mkt.RF	fxcarry	fxusd
CM	8.801	0.358	0.103	-0.298
	(1.960)	(2.621)	(0.591)	(-2.235)
Base	3.271	0.499	0.137	-0.318
	(0.892)	(4.677)	(1.158)	(-2.645)
Naive	2.347	0.529	0.334	-0.175
	(0.780)	(5.370)	(2.281)	(-1.766)

Panel B

Tests comparing alphas of CM and Base models

	Difference	T-test p-values	
Market	in alphas/yr	2-sided	1-sided
DM	2.01	0.082	0.041
EM	5.53	0.002	0.001