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

Charles W. Calomiris and Harry Mamaysky

April 2017

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 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 efficacy of our approach for forecasting country-level market outcomes.

Calomiris: Columbia Business School, cc374@gsb.columbia.edu; Mamaysky: Columbia Business School, hm2646@columbia.edu. 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. We thank Kent Daniel, Robert Hodrick, Leif-Anders Thorsrud, Diego Garcia 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, Columbia University, and the University of Colorado for valuable comments.

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 U.S. 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 et al. 2008, Garcia 2013). This literature is still in its infancy, but already it has become apparent that it is possible to quantify salient aspects of word flow and link these measures to market outcomes. Moreover, the information contained in these textual analyses not only is associated with changes in returns and risk, but can provide incremental predictive value over and above other variables that are associated with changes in returns and risk (such as "value" or "momentum" measures, or measures of economic activity or changes in related asset prices, such as exchange rates or interest rates).

The promising early work in the literature linking textual analysis and stock returns has raised more questions that it has answered. Here is a list of nine important sets of questions that have been raised, and how our study addresses them:

First, and perhaps most importantly, *how* should one best measure news using word flow? One approach 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 in order to see which parts of word flow matter for market outcomes. An alternative approach would be 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 (for example, Baker, Bloom and Davis 2016). 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 by imposing discipline on the process by which text is analyzed. There are only a few atheoretical measures of word flow to consider in empirical work: frequency, sentiment, unusualness of words, and topical categories that are identified without imposing any a priori notions of how to group words. In contrast, there is no limit to the number of a priori mappings that could be constructed based on particular words or combinations of words. A potential advantage of the alternative a priori approach, however, is that it may cut through noise by searching in more relevant places for words or phrases that have likely importance.

We adopt an atheoretical approach, and cast a broad net to see which aspects of word flow are most relevant for market outcomes. We then compare that approach with one prominent a priori approach (the Baker, Bloom and Davis measure of uncertainty), which confirms our view that an atheoretical approach may have desirable properties.

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 pre-identified 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 contexts in which words appear (which we label "topics") is important to the meaning of word flow. In addition to sentiment, frequency, and context of word flow, other aspects may be relevant. Glasserman and Mamaysky (2016a) show that the "unusualness" of word strings

(also known as “entropy”) may have predictive power for market outcomes, especially when interacted with sentiment. Furthermore, as we show below, the effects of measured sentiment and unusualness may vary across topic categories, so it may be useful to allow frequency, sentiment and unusualness to affect market outcomes in a topic-specific manner.

Our empirical approach will include these various measures of text flow and their interactions, and explore their incremental information content relative to non-textual variables often included in asset pricing studies. We also include controls that capture differences over time related to the electoral cycle, which we thought might be relevant when measuring the effects of news on market outcomes.

Third, the patterns that link frequency, topics, sentiment and unusualness measures of word flow with market outcomes may vary over time. Regime changes over time ideally would be captured in a neutral manner (identifying switches endogenously using data patterns).

As a starting point, 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 further explore dynamic changes in coefficients using a rolling lasso 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? One could adopt an a priori approach to identifying topic areas, or alternatively, employ an atheoretical approach that divides words into topic groups based on certain statistical properties of the corpus. 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 et al. 2003) approaches, as we discuss further below. The Louvain method assigns salient words to mutually exclusive topic areas based on word co-occurrence (that is, with 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, should all market outcomes (i.e., both changes in expected returns and changes in risk) be related to each of the identified word flow measures, or should some word flow measures be more important for risk and other measures more important for expected returns?

Our approach allows all word flow measures to matter for all of the market outcome measures. Our findings suggest that when a word flow measure has positive content for expected returns it also has negative content for changes in risk (i.e. risk, however measured, will be lower). In other words, 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.

[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.

Sixth, one must decide which market measures to include in any analysis of the effects of word flow on market outcomes? Conceptually, the obvious focus, reflected in prior work, should be on expected returns and on risk. For each country, returns can be measured using the change in the value of the country stock market index. Risk, however, is a more subtle and problematic variable to measure. 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, however, generally are rejected, especially for emerging market (EM) countries, which exhibit pronounced momentum, and non-normality, both with respect to skew and kurtosis (see Bekaert et al. 1998, Karolyi 2015, and Ghysels et al. 2016).

Given those facts, to capture risk, in addition to using the standard deviation of returns, 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" (a focus on the standard deviation of returns treats them as identical).

Seventh, the existing literature focuses on individual U.S. companies or the U.S. stock market as a whole. Do empirical patterns that apply to individual company stocks or the aggregate index in the U.S. also apply to other countries? When analyzing countries' 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.[2] We divide countries into two groups: developed economies (DMs) and EMs, which we believe

[2] 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.

allows for important environment differences across countries while still preserving sufficient statistical power. On the one hand, statistical power can be enhanced by a panel approach. On the other hand, important differences across countries may be lost. As we have already noted, there is a great deal of evidence suggesting that returns processes differ between EMs and DMs. Furthermore, the amount of risk and the nature of the news that drives risk differ dramatically between EMs and DMs (Beim and Calomiris 2001, Karolyi 2015). Those categorical differences likely reflect fundamental differences in political contexts, which result in different 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 banking crises (Laeven and Valencia 2014). EMs suffer larger and more frequent major drawdowns of stock returns than EMs (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.

Specifically, our empirical study divides countries into two groups: DMs and EMs, and estimates the connections between a comprehensive set of word flow measures and market outcomes (*returns, sigma and drawdown*) for both samples, for the entire sample period (1998-2015) and for two sub-periods (April 1998 – February 2007, and March 2007-December 2015).

Finally, over what time frame should the information content of 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 (2016a) are an exception; using the U.S. stock market index, they find predictive power for risk several months following word flow. We examine both one month ahead and one year ahead predictions, and show that our country-level aggregates 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-year ahead return (*return)*,[3] 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

[3] Alternatively, we used one-month ahead measures of returns. Results were qualitatively similar but explanatory power was lower. When it's not clear from the context, we use *return* and $return^{12}$ to refer to one-month and one-year ahead returns respectively.

World Bank and the IMF, as detailed in the Data 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. See the Data Appendix for more details.

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 unusualness; and (iv) calculation of article-level sentiment, topic and unusualness measures. Here we present a high-level overview of the process. The Data 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[4] that are descriptive of either market or economic phenomena (pre-stemming examples include *barriers*, *currency*, *parliament*, *macroeconomist*, and *World Bank*). This list was derived as follows: We began with the index from Beim and 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 in an On-line Appendix.

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 non-overlapping 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 Data 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.[5] Figures 1 and 2 show the most frequently occurring words in each of our EM and DM topics.[6] 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

[4] Bigrams and trigrams (or 3-grams) are word phrases of length two and three respectively.
[5] 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 Data Appendix for more details.
[6] Figure 3 shows that the original Louvain clustering produced over 40 word groupings for each corpus, yet only 5 of these contained more than just a few words. We place words from the smaller groupings into the 5 large one for each corpus. This is discussed in greater detail in the Data Appendix.

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, sizeable. Our measure of word overlap is the Jaccard index, 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 analogue 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.[7] For example, in the emerging market corpus an article titled "Clinton says Putin can build strong, free Russia" is classified as being in

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

the *Govt* topic. A Portuguese language article (which was wrongly coded by our data source as an English 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 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' 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 considered restricting our sample of articles to those that were more narrowly focused on business, economics and politics topics, but we found that doing so slightly reduced the explanatory power of news for stock returns and risk, and so we retained 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 *unusualness* of a given article, following closely the methodology proposed in Glasserman and Mamaysky (2016a). An n-gram is a collection of *n* contiguous words.[8] We do not allow n-grams to cross sentence boundaries – so these are n-word phrases that appear entirely in a single sentence. Our measure of the

[8] The phrase "collection of *n* contiguous" is an example of a 4-gram.

unusualness of a given 4-gram is the probability of observing the 4th word in the phrase conditional on seeing the first 3 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 where the 3-word phrase that begins the 4-gram was not seen in the training corpus. In the Data Appendix, we discuss why this 1:10 rule is an appropriate choice.

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 Data 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 unusualness 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

$$H_j = -\sum_i p_i \log m_i \tag{2}$$

where p_i is the fraction of all 4-grams appearing in the article represented by the ith 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 unusualness 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

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

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. Sometimes, the lack of semantic context causes our sentiment classification to assign a value that we regard as inappropriate, 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.

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 non-overlapping 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

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

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".

We are interested in testing in this paper whether article decomposition matters. Does negative or positive sentiment matter more or less for forecasting future market outcomes when it occurs in news stories about governments than when it occurs in news stories about macroeconomics? And how do these distinctions change when we consider DM or EM economies? Much of our empirical analysis will be focused on these and related questions.

It should also be noted that a similar decomposition can be applied to the article level entropy measure H_j. However, we did not find this decomposition to be useful and in all our empirical results report article-level entropy. Following Glasserman and Mamaysky (2016a), we compute article level context-specific sentiment interacted with entropy as follows

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

SentEnt is intended to capture the possibility that extreme sentiment (either positive or negative) may have more information content for future market outcomes when it occurs in articles that appear unusual (i.e. have high entropy).

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 given by

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

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 measure.

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. Several features of the data are worth mentioning. Compared to DM, EM returns were higher (1.04% vs 0.65% per month), more

volatile (21.48% vs 19.20% annualized volatility), and subject to higher drawdowns (17.4% vs 15.3%). As reported by Ghysels et al. (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 vs 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 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 vs only 2.7% for DM). Finally, the average article level entropy for both corpora is roughly 2.45.

Figures 4 and 5 provide factor loadings and plots for each topic category related to the first two principal components for the 140 EM (5 series for 28 countries) and 120 DM (5 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 movements. 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. Interestingly, the second principal component 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 related

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 et al.'s (20008) and Hendershott et al.'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. Figures 6 and 7 plot cumulative abnormal returns (for EMs and DMs, respectively) for the 10 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.[9] We plot these abnormal returns separately for positive and negative news dates, along with standard error

[9] 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.

bands.[10] We also plot (in between the positive and negative top deciles) the results for the middle decile (45th-55th percentile) as a placebo.[11]

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 (dissimilar) topical category (*Macro* for EMs, and *Credit* for DMs). For both sets of countries, the shapes 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* and 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 et al. (2008) – their Figure 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

[10] 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.

[11] 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 non-trading days (e.g. weekends) or within 10 days of the start or end of the sample. We do not include such event days for the calculation of abnormal returns.

seems to hold for both EM (for *Mkt* and *Comms* topics) and DM (for *Mkt* and *Credit* topics) markets. In Tetlock et al. (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 post-event country level drift after negative news suggests that long/short news-based strategies may be more profitable at the country level than the individual stock level. Though tentative, this is potentially further evidence of the relative micro efficiency and macro inefficiency of markets (see Glasserman and Mamaysky 2016b 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 10 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

Turning to our regression analysis, in Tables 7-12, we report regressions employing country-month observations, divided into EM and DM samples, for our three dependent variables (*return*[12], *sigma* and *drawdown*).[12] We regress month *t* values of the dependent

[12] Regression results for one-month ahead returns are shown in the Supplementary Appendix.

variables on lagged (either *t-1* or *t-2*) values of our explanatory variables. In each table, we report nine different regressions, which consist of three regressions for each of three time periods. The three time periods are 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 (non-textual 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 in Table 5. We included indicator variables that capture electoral timing by dividing time periods into pre-election and post-election periods, as described in Section 2 and the Data Appendix.

In addition to the Baseline regression, for each time period, we report two additional regressions which 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 (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. See the Data Appendix for more details about our regression specifications.

Table 13 summarizes the results of all our panels (*return*, $return^{12}$, *sigma*, and *drawdown*) for the *Sent* specification over the entire sample.[13] The columns in Table 13 correspond to the dependent variable, and the rows correspond to the explanatory variables, with our textual variables toward the bottom of the table. 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).

We summarize our findings as follows. (1) Our findings with respect to Baseline variables' effects generally are consistent with prior studies and will not be commented on further here.[14] (2) Results differ across EMs and DMs (coefficient values are not consistent across the two groups of countries), 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 outcomes, differ in EMs and DMs (reflecting important differences in the political and economic environments, which are reflected in returns outcomes).

(3) 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,

[13] The Supplementary Appendix contains summary tables which include results from the early and late sub-periods.
[14] One interesting finding is that GDP growth is negatively associated with returns and positively associated with drawdown. We can think of many potential explanations for this finding. First, it may be that positive GDP growth raises the probability of contractionary monetary policy, which may be bad news for the stock market. Second, it is possible that GDP growth serves 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.

where good news increases *return* and reduces risk measured either by *sigma* or *drawdown*. In fact, we never observe a coefficient 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.

(4) The economic importance of word flow measures (incremental contribution to R-squared) tend 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 also 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, and the incremental contribution of word flow to *sigma* is small in DMs.

(5) In EMs, the economic importance of word flow measures is higher for all three 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 doubled the R-Squared for both $return^{12}$ and *drawdown* (from 0.07 to 0.13 and from 0.06 to 0.12). For the pre-crisis period, that difference between EMs and DMs is even greater: for EMs, R-squared 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.

The impacts of individual text flow measures on annual returns and drawdowns often are economically large. In DMs, text measures are only 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 3.7% higher return over the next year (the product of its standard deviation, 0.17, and its coefficient, 21.87). A standard deviation increase in *sMkt* is associated with a 5.2% increase in return, while a standard deviation increase in *sGovt* is associated with a 3.9% reduction in return.

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 in both subperiods. 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%.

With respect to the EM drawdown regressions (in Table 12), more variables are statistically significant, 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% 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 in the earlier subperiod.

(6) Somewhat contrary to the findings in Glasserman and Mamaysky (2016a) (which focused on financial corporations' and US stock market returns, rather than country returns around the world) we do not find that interacting sentiment measures with unusualness, the *SentEnt* specification,adds much explanatory power. Coefficient magnitudes and R-squared sometimes rise and sometimes fall across *Sent* and *SentEnt* specifications, but the changes tend to be small. *Entropy* does sometimes enter as a significant predictor of *drawdown* and *return*, but interacting sentiment with entropy adds little.

(7) Consistent with our principal component discussion in Section 2, we find important differences in coefficient values for our word flow measures over time – that is, differences between the pre-global financial crisis period and the post-global financial crisis period. Tables A1 and A2 in the Supplemental Appendix summarize our panel results by subperiod. For DMs, negative coefficients on *return* for *fGovt* and *sGovt* are a feature of the post-crisis sub-period, as is the positive coefficient for *return* for *sMkt*. For EMs, positive *fGovt* is associated with larger

drawdown in the earlier subsample, but not in the later. For EMs, the coefficient on *entropy* in the *return*^{12} regression is zero across the two sub-periods, 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-crisis 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.

(8) Topical context matters for the influence of frequency and sentiment. Coefficients for sentiment or frequency can be positive or negative, depending on the topic area, and depending on the 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 EMs. Clearly, there is something to be gained by considering the context in which positive or negative sentiment are expressed.

One interpretation of our findings on sentiment is that articles with positive sentiment about the government may reflect bad news that government actions are trying to address. The notion that positive sentiment in the context of government responses is reflecting adverse news events would also explain the post-crisis timing – especially for DMs – of those surprising coefficients.

3.2 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 U.S., Canada, Germany, the U.K., Italy, France, Spain, Netherlands, Japan, Australia and Ireland), and 6 EMs (Chile, China, South Korea, Brazil, Russia and India). Although these DM and EM sub-samples 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 sub-sample. Our sample time frame is from 1998 to 2015.

In Table 14 we show that our word measures are able to explain a substantial amount of the future variation in the BBD uncertainty measure (Table 14 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. High *Mkt* sentiment and frequency, high *Macro* frequency for EM and positive *Credit* sentiment for DM forecast lower future *EPUs*; whereas high *Govt* and *Corp* sentiment and frequency in DM, as well as higher article counts in EM, forecast higher future *EPUs*. Perhaps this is one channel through which positive sentiment and high frequency in the *Govt* and *Corp* topics forecast adverse market outcomes, as we saw in Table 13.

Of greater interest is the explanatory power of the BBD measure of economic policy uncertainty (EPU_{t-1}) for $return^{12}$, *sigma* and *drawdown*, both by itself and in regressions that include our word flow measures.[15] Tables 15 and 16 evaluate the incremental explanatory power of the BBD measure for those three variables. For each variable we report four regressions: a

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

Baseline regression that includes neither our word flow measures nor the BBD measure, a second regression that includes only the BBD measure, a third regression that includes only our word flow measures, and a fourth regression that includes both the BBD measure and our word flow measures.

For DMs, in the second regressions for each variable in Table 15, the BBD measure does exhibit incremental explanatory power, but the effects are small. R-squared for $return^{12}$ is increased by 0.008, 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 16, for $return^{12}$, *sigma* and *drawdown*, all show that the BBD measure adds almost no incremental explanatory power for all three variables relative to the Baseline. Only in the $return^{12}$ regression does the BBD measure enter significantly, and then only in the specification that includes all our text measures (adding the BBD measure to this specification results in a very low R-Squared increase of 0.008). 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

There are two important reasons to explore out-of-sample forecasting properties of our model. First, as the above discussion of changes in estimated coefficients reveals, we observe substantial variation over time in coefficient estimates reported in Tables 7-12, where the timing split was suggested by the principle components analysis reported in Figures 4-5. Given that variation, it is unclear whether a forward-looking application of our model would produce useful forecasts of market return and risk. Second the comparison of our model's performance with that of BDD in Tables 15 and 16 may not provide a fair indication of the relative usefulness of the two approaches for out-of-sample forecasting, especially given the fact that our model employs many textual analysis measures and their model employs only one – implying greater potential problems of over-fitting in our model.

When over-fitting 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 least absolute shrinkage and selection operator (lasso) regression, introduced by Tibshirani (1996)[16], to ameliorate this over-fitting problem. In our panel setting, we estimate rolling five-year regressions using the lasso objective function, which is given by

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

[16] See also Hastie and Qian (2016) for a description of the lasso and its implementation in the R package called *glmnet*. We use *glmnet* for our empirical work.

where N is the total number of observations in the regression, $y_{i,t}$is the response variable, and $x_{i,t-1}$is a vector of the predictors. In this specification, we normalize by first constructing demeaned y's and x's within each country grouping – so a constant in the above regression is not necessary – and then by setting the variance of y and the x's to 1. The choice of λ determines the penalty applied to the sum of the absolute values of coefficients. We use $\lambda = 1$, but our results are qualitatively similar for λ of 0.5 or 1.5.

A lasso regression performs both model selection – many of the β's can be set to zero – and shrinkage – the non-zero coefficient estimates are smaller than their OLS counterparts. A measure of the degree to which lasso coefficients are smaller than their OLS counterparts is the ratio $\|\beta\|_1/\|\beta^{OLS}\|_1$. In our empirical work this ratio (reported in the upper left-hand corners of Figures 8-13) ranges from close to zero, to as high as 70 percent, meaning that the lasso 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 70 percent as large as their OLS counterparts (for example, in some windows for forecasting 12 month ahead returns).

We investigate the out-of-sample properties of our model in Table 17. We compute root mean squared errors (RMSE) based on lasso coefficient estimates from a rolling five-year regression model for *return*, *return*^{12}, *sigma* and *drawdown*, separately for EMs and DMs.[17] We compare our model's performance (RMSE) relative to two benchmarks in Table 17: a *Naïve* model (with no forecasting variables except in-sample country fixed effects) and a *Base* model

[17] To make a forecast for month *t* we run a lasso regression of market outcomes on lagged explanatory variables – using the same panel specification as in Section 3 – from months *t-60* to *t-1*, and then use those coefficients and month *t-1* independent variables to make the month *t* forecast.

that uses only the non-textual measures in our model (analogously to our comparisons in Tables 7-12). Out-of-sample findings confirm our in-sample estimates. Our model (labeled *CM*) outperforms both the *Naïve* and *Base* models for *return*^{12} and *drawdown*, but adds little if any explanatory power to the *Base* model of *sigma*. Our model has little predictive power for the one month-ahead *return*.

Figures 8-13 show significant changes over time in the lasso coefficient estimates for the variables in our model, including the text measures. Coefficients 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 drawdowns is visible in the lasso estimates. When interpreting coefficient magnitudes it is important to bear in mind that multicollinearity (which, by construction, is not apparent in lasso estimates) leads to some non-comparability of coefficient magnitudes reported in Tables 7-12 and in Figures 8-13. The lasso results reinforce the message of Tables 7-12 that coefficients vary importantly over time, and the lasso results are broadly comparable to those in Tables 7-12. 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*, *sCorp* and *fCorp* tend to be bad news.

In Table 18, we apply our lasso out-of-sample testing to a comparison of our model and BBD (the *EPU* measure), based on RMSE. This table shows the RMSE using the *Naïve*, *Base*

and our model (*CM*), as described before. We also show the RMSE from adding *EPU* to the *Base* model (labeled *EPU)* and from adding *EPU* to our full model (labeled *Both*). Our model (*CM*) consistently out-performs the *EPU* model in out-of-sample tests, for both EMs and DMs, for one year ahead *return*, *drawdown*, and *sigma*, although neither model adds much to the *Base* model for explaining *sigma*. Neither model has incremental explanatory power for predicting one month-ahead *return*. The *Both* model, which introduces *EPU* to the *CM* model, does not improve RMSEs relative to *CM*. These results confirm the findings in Tables 15 and 16, and show that the superior forecasting ability of our model in comparison to BBD is not an artifact of in-sample over-fitting.

5. Conclusion

To the best of our knowledge, this is the first study of non-U.S. country-level stock returns and risks that relates news to future risk and return. 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, and we apply that approach to more than 50 countries over the time period 1998 to 2015. We find that it is possible to develop a parsimonious and flexible approach to mapping from news word flow to equity market risk and returns. We find that news contained in our text flow measures is more relevant for one year ahead returns and drawdowns than for one month ahead returns and volatility. Clearly, word flow has important medium-term predictive power for country-level stock returns. One interpretation of this finding of a delayed response 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 other 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 measures appear as negative news events.

Word flow measures tend to have greater incremental predictive power for understanding returns and risks in EMs, and this is particularly true for cumulative drawdowns. Those findings are consistent with the view that the nature and range of news events differs between DMs and EMs, which reflects differences in the political and economic environments of the two types of countries.

We compare the predictive power of our atheoretical approach to analyzing the word flow 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 a lasso regression analysis, which has two purposes: first, to investigate whether our model can be useful for out-of-sample forecasting despite the time variation in estimated coefficient parameters, and second, to ensure that our comparison with BBD does not reflect overfitting bias as a consequence of our larger number of text measures. The RMSEs for our model, when forecasting one year-ahead *return* and *drawdown,* are lower than for the Baseline model (which excludes textual analysis) or the BBD model.

We conclude that the meaning of news flow can be captured 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.

Data Appendix

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*[18], 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 which 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

[18] 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 2-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 ith column of the document term matrix, and $\|D_i\|$ is the Euclidean norm of the ith 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 larger 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 words 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 $(i,j)^{\text{th}}$ 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 intra-cluster 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.

Figure 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 4 year windows of our DM and EM corpora (recall the data set runs from 1996 to 2015). In each 4-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' which 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 4 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 sub-period.

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 – Latent Dirichlet Allocation (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 2 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 5 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 equation (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 which 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 3 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 equation (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 which 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 Figure A1 in the Supplementary 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 often these 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, *volatility*, 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 et al. (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 is obtained from International Financial Statistics of the International Monetary Fund. Annual data is 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 is obtained from International Financial Statistics of the International Monetary Fund. Annual data is

used only when quarterly data is 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 is 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 does not exist.
- **Monthly percent change in local currency exchange rate vs the US Dollar** (*dexch*): We obtain the monthly exchange rate data from Datastream. 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 depreciation. The US series is set to zero. This variable is truncated at $\pm 50\%$.

- **Pre-election dummy** (*pre*): The pre-election 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.
- **Post-election dummy** (*post*): The post-election 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 post-election month is labeled as a pre-election (but not a post-election) 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.

Finally, we investigate whether the non-normality 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 non-normality is unlikely to be problematic for our estimates.

References

Asness, C., T. Moskowitz, and L. Pedersen, 2013. Value and Momentum Everywhere. *The Journal of Finance*, 68 (3), 929-985.

Baker, S. R., Bloom, N., and Davis, S. J. 2016. Measuring Economic Policy Uncer-Tainty, *Quarterly Journal of Economics, forthcoming.*

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

Bekaert, G., Erb, C., Harvey, C. and Viskanta, T. 1998. Distributional Characteristics of Emerging Market Returns and Asset Allocation, *Journal of Portfolio Management*.

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

Blei, D, Ng, A. and Jordan, M., 2003, Latent Dirichlet Allocation. Journal of Machine Learning Research, 3 (4–5), 993–1022.

Blondel, V., J.-L. Guillaume, R. Lambiotte, and E. Lefebvre. 2008. Fast Unfolding of Communities in Large Networks, *Journal of Statistical Mechanics*, 10, 10008.

Calomiris, C. W., Love, I., and M. S. Martinez Peria. 2012. Global Returns' Sensitivities to Crisis Shocks, *Journal of International Money and Finance*, 30, 1-23.

Calvo, G. A., and Reinhart, C. M. 2002. Fear of Floating, *Quarterly Journal of Economics* 117 (2), 379-408.

Cruces, J. J., and Trebesch, C. 2013. Sovereign Defaults: The Price of Haircuts, *American Economic Journal: Macroeconomics*, 5 (3), 85-117.

Das, and Chen 2007. Yahoo! For Amazon: Sentiment Extraction from Small Talk on the Web, *Journal of Management Science*, 53(9) 1375-1388.

Fama, E., Fisher, L., Jensen, M. and Roll, R. 1969. The Adjustment of Stock Prices to New Information, *International Economic Review*, 10(1), 1-21.

Garcia, D. 2013. Sentiment during Recessions, *Journal of Finance*, 68(3), 1267-1300.

Glasserman, P., and Mamaysky, H. 2016a. Does Unusual News Forecast Market Stress? OFR Working Paper 16-04, April 20.

Glasserman, P., and Mamaysky, H. 2016b. Market Efficiency with Micro and Macro Information. Columbia Working Paper.

Ghysels, E., Plazzi, A., and Valkanov, R. 2016. Why Invest in Emerging Markets? The Role of Conditional Return Asymmetry, *Journal of Finance* 71 (5), 2145-2191.

Hastie, T. and J. Qian, 2016, Glmnet Vignette, working paper.

Hendershott, T., Livdan, D., and Schurhoff, N. 2015. Are Institutions Informed about News? *Journal of Financial Economics*, 117, 249-287.

Kaminsky, G., and Schmukler, S. 2008. Short-Run Pain, Long-Run Gain: The Effects of Financial Liberalization, *Review of Finance*, 12(2), 253-292.

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

La Porta, R., Lopez de Silanes, F., Shleifer, A., and Vishny, R. Law and Finance, *Journal of Political Economy*, 106 (6), 1113-1155.

Loughran, T., and McDonald, B. 2011. When Is a Liability not a Liability? Textual Analysis, Dictionaries, and 10-Ks, *Journal of Finance* 66(1), 35-65.

Meila, M, 2007. Comparing Clusterings – An Information Based Distance, *Journal of Multivariate Analysis, 98, 873-895.*

Netzer, O., R. Feldman, J. Goldenberg, and M. Fresko, 2012. Mine your own business: Marketstructure surveillance through text mining, *Marketing Science*, 31 (3), 1–23.

Newman, M.E.J. 2006. Modularity and Community Structure in Networks, *Proceedings of the National Academy of Sciences*, 103 (23), 8577–8582.

Newman, M.E.J. and M. Girvan. 2004. Finding and Evaluating Community Structure in Networks, *Physical Review E*, 69, 026113.

Roll, R. 1984. Orange Juice and Weather, *American Economic Review*, 74(5), 861-880.

Rule, A., J.-P. Cointet, and P. S. Bearman, 2015. Lexical shifts, substantive changes, and continuity in State of the Union discourse, 17902014, *Proceedings of the National Academy of Sciences*, 112 (35), 1790–2014.

Shiller, Robert J. 2017. Narrative economics. Presidential Address at the 129th Annual Meeting of the American Economic Association.

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., and 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.

Tibshirani, R., 1996. Regression Shrinkage and Selection via the lasso, *Journal of the Royal Statistical Society*, Series B, 58 (1), 267—88.

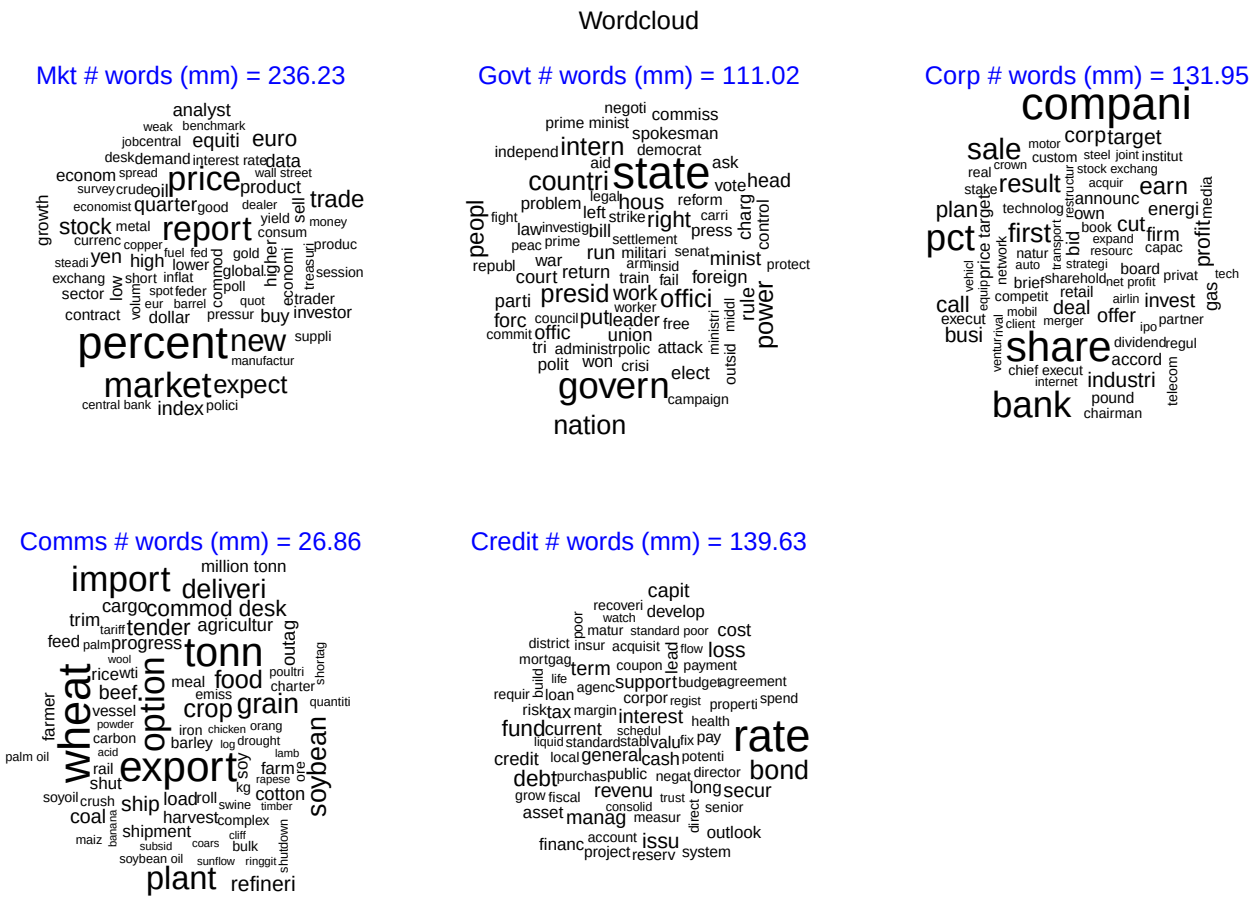

Figure 1: Word cloud plots for topics extracted from the developed markets corpus using the Louvain clustering algorithm. Each cluster shows the number of occurrences (in millions) of its words in the corpus.

Figure 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 words in the corpus.

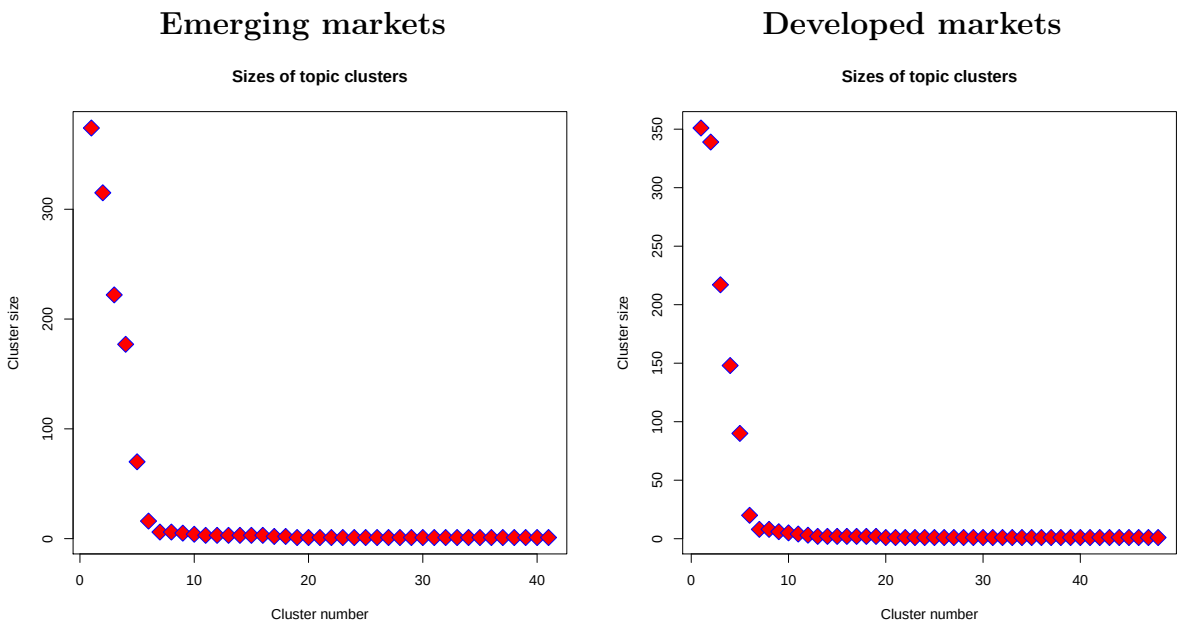

Figure 3: Number of words in original set of clusters identified by the Louvain algorithm for the emerging and developed markets corpora respectively.

Factors decomposition of news topic sentiment in emerging markets

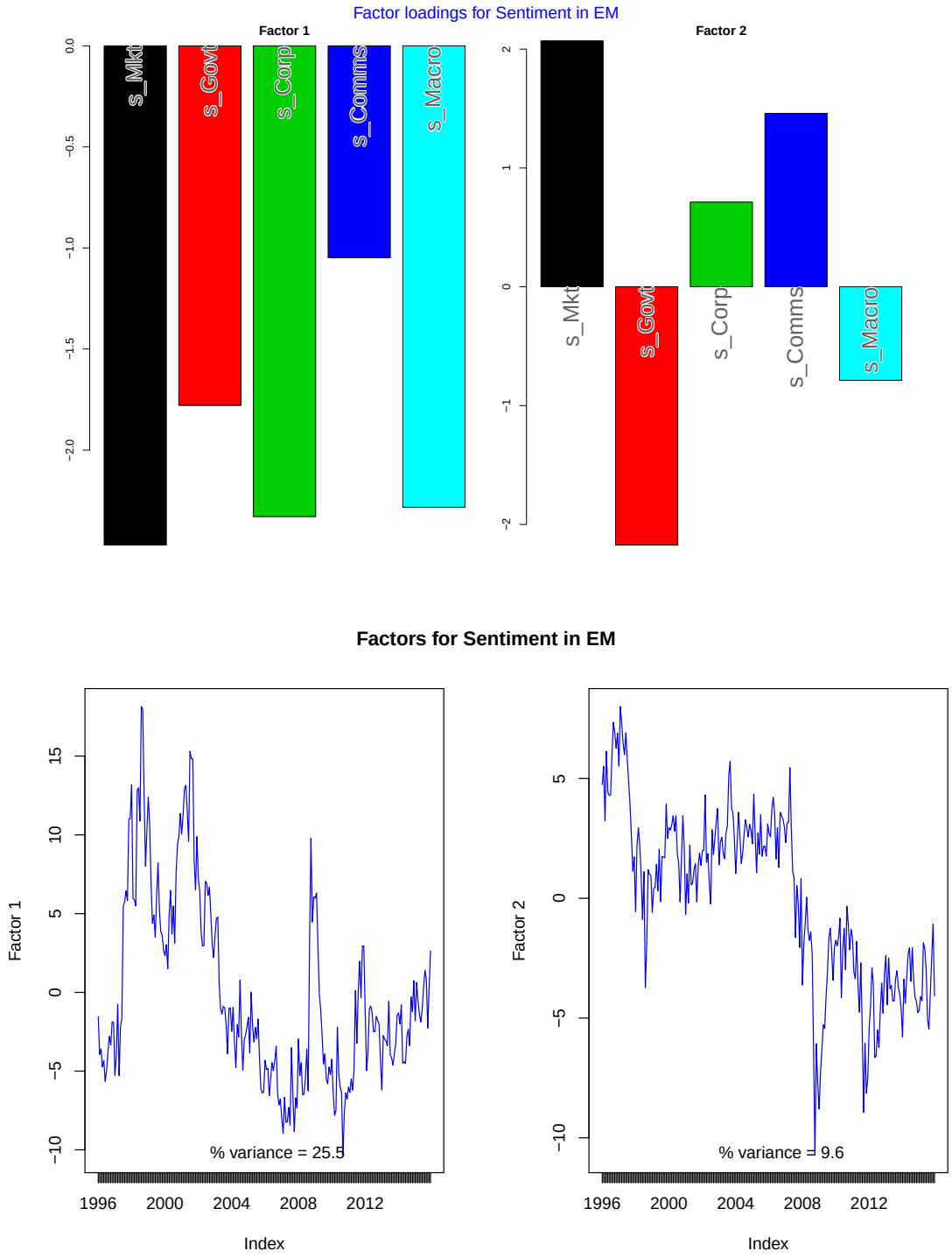

Figure 4: The figure shows the top factors from a principal components analysis of all country-topic sentiment series (i.e. $\#(countries) \times \#(topics)$) from emerging market countries. The top row shows the topic loadings of each factor. All country-topic sentiment series were normalized to unit variance.

Factors decomposition of news topic sentiment in developed markets

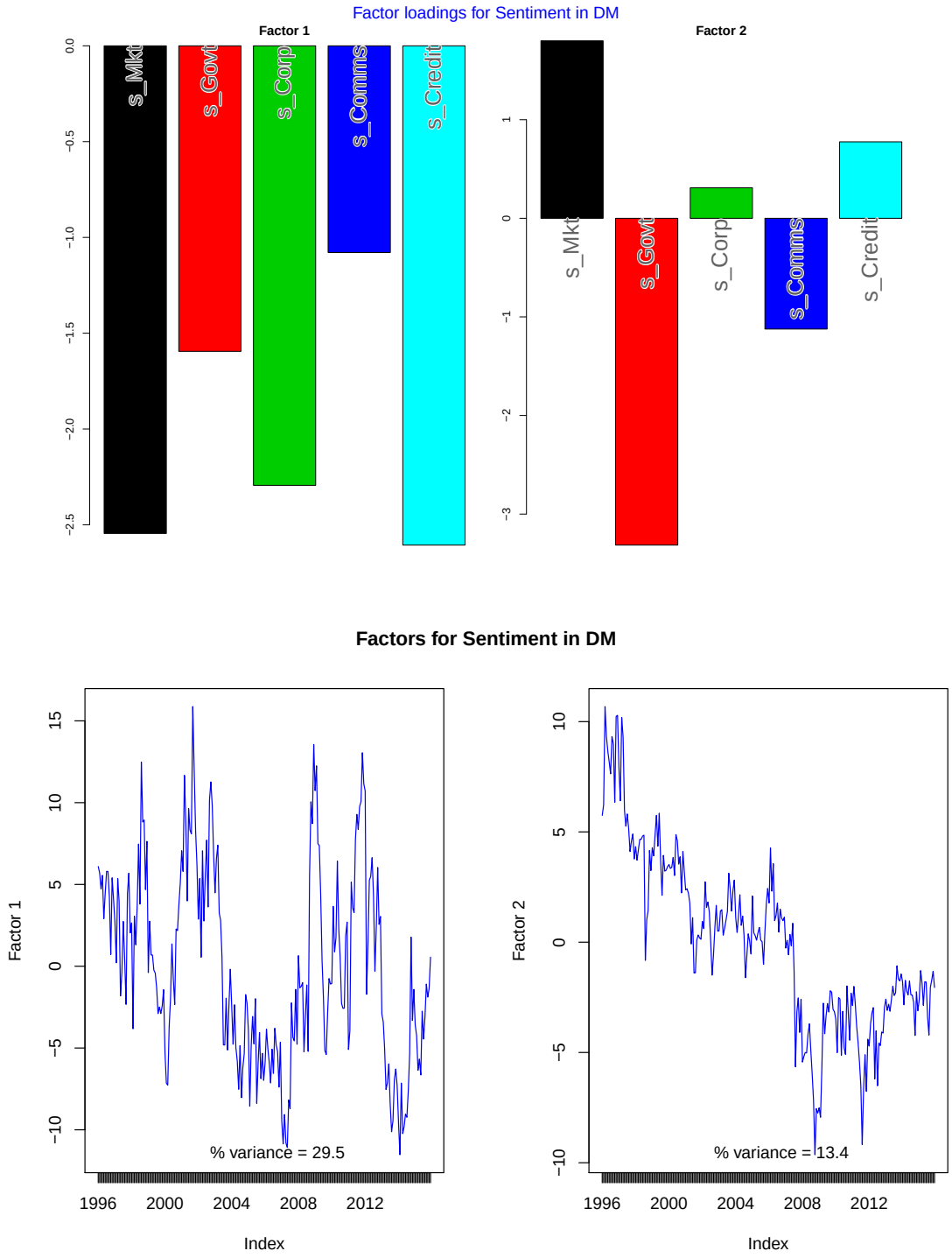

Figure 5: The figure shows the top factors from a principal components analysis of all country-topic sentiment series (i.e. $\#(countries) \times \#(topics)$) from developed market countries. The top row shows the topic loadings of each factor. All country-topic sentiment series were normalized to unit variance.

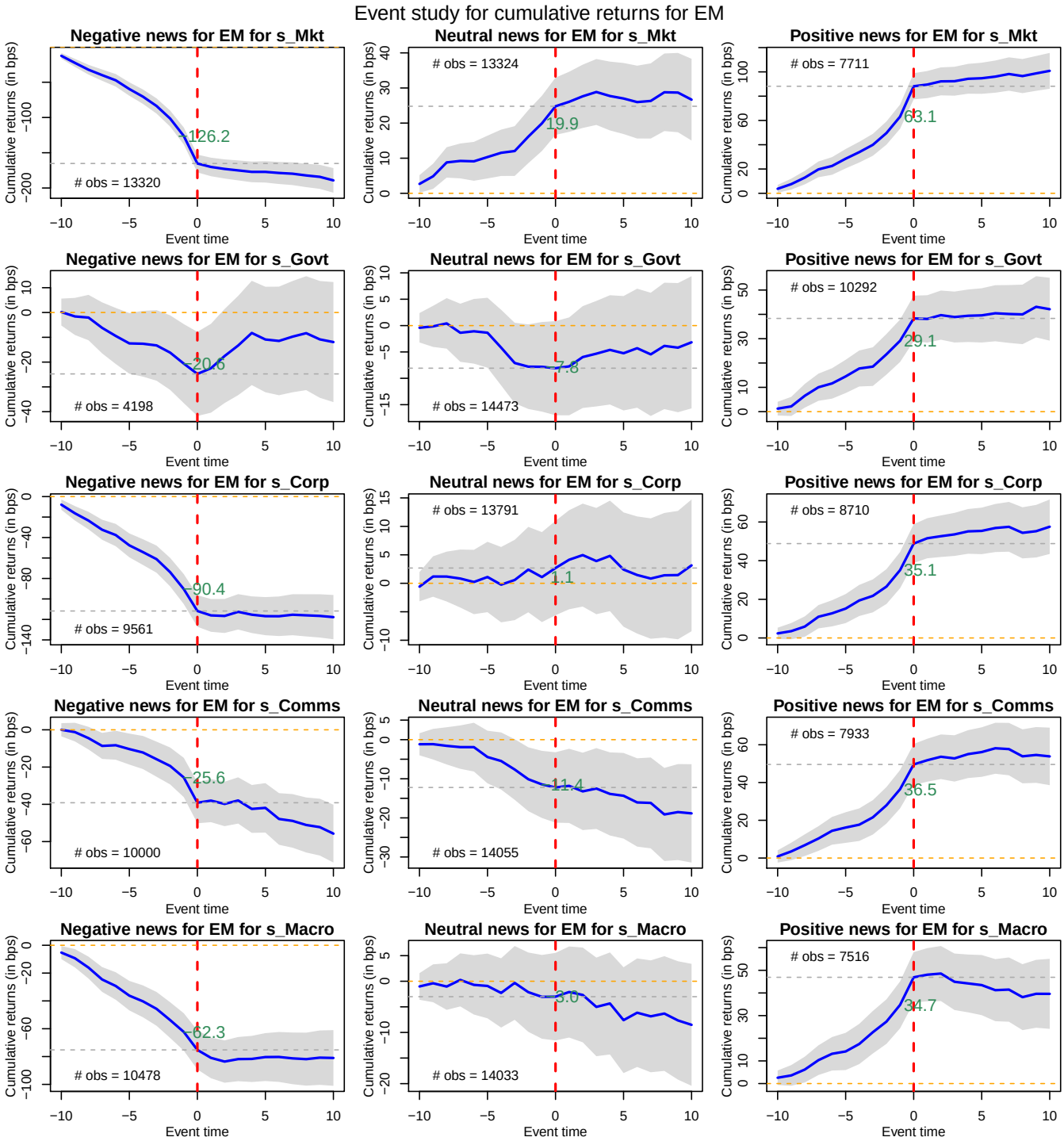

Figure 6: Event studies of cumulative abnormal returns on days that are in the bottom, middle (45%-55%) and top deciles by sentiment for each topic. Abnormal returns are the residuals from a regression of US dollar country index returns on a EM market index and a constant. Cumulative returns are shown in basis points, with two standard error bands. 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 Jan 1, 1996 to Dec 31, 2015.

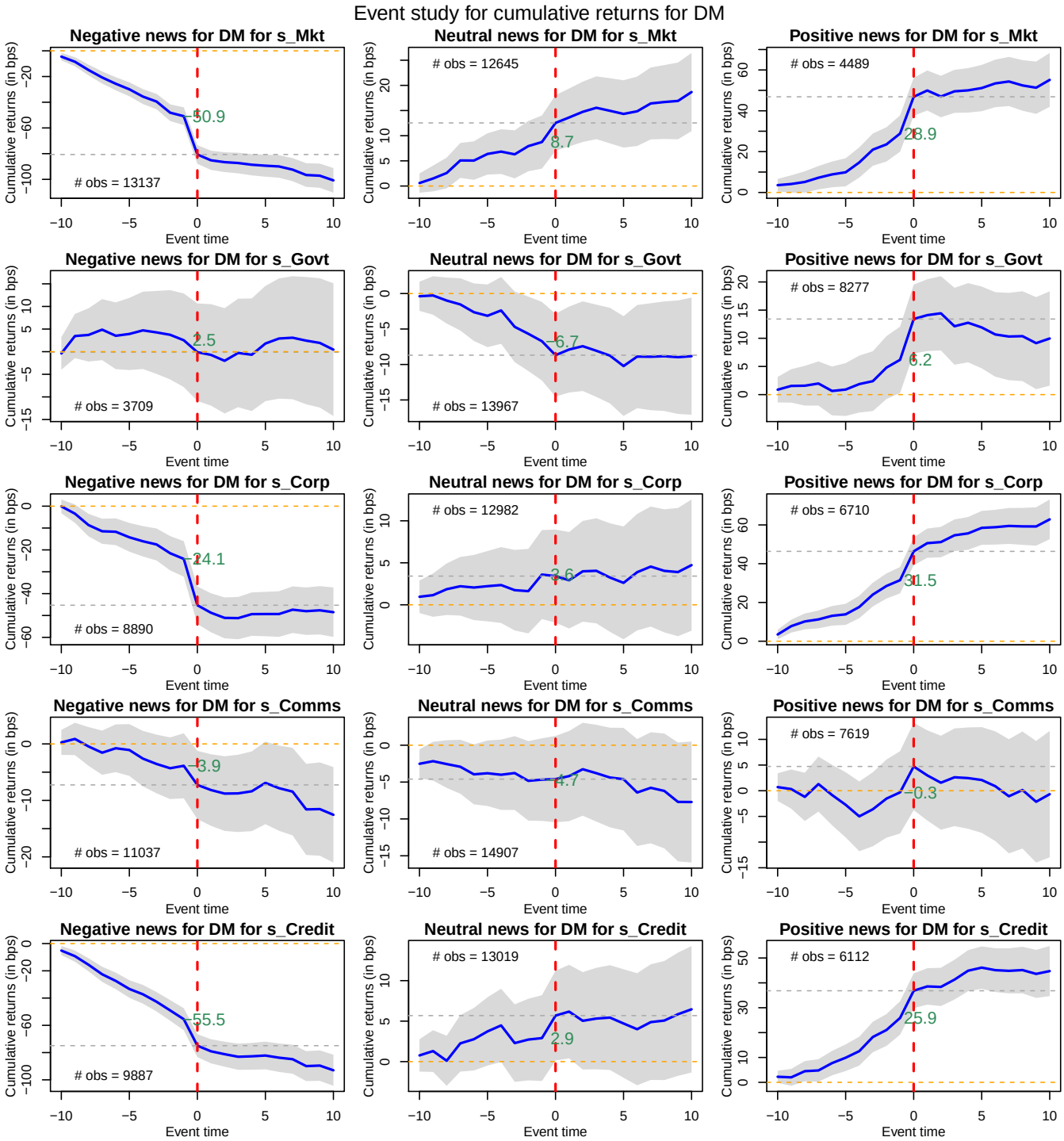

Figure 7: Event studies of cumulative abnormal returns on days that are in the bottom, middle (45%-55%) and top deciles by sentiment for each topic. Abnormal returns are the residuals from a regression of US dollar country index returns on a DM market index and a constant. Cumulative returns are shown in basis points, with two standard error bands. 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 Jan 1, 1996 to Dec 31, 2015.

Coefficient time series from lasso for 12 month returns for DM countries

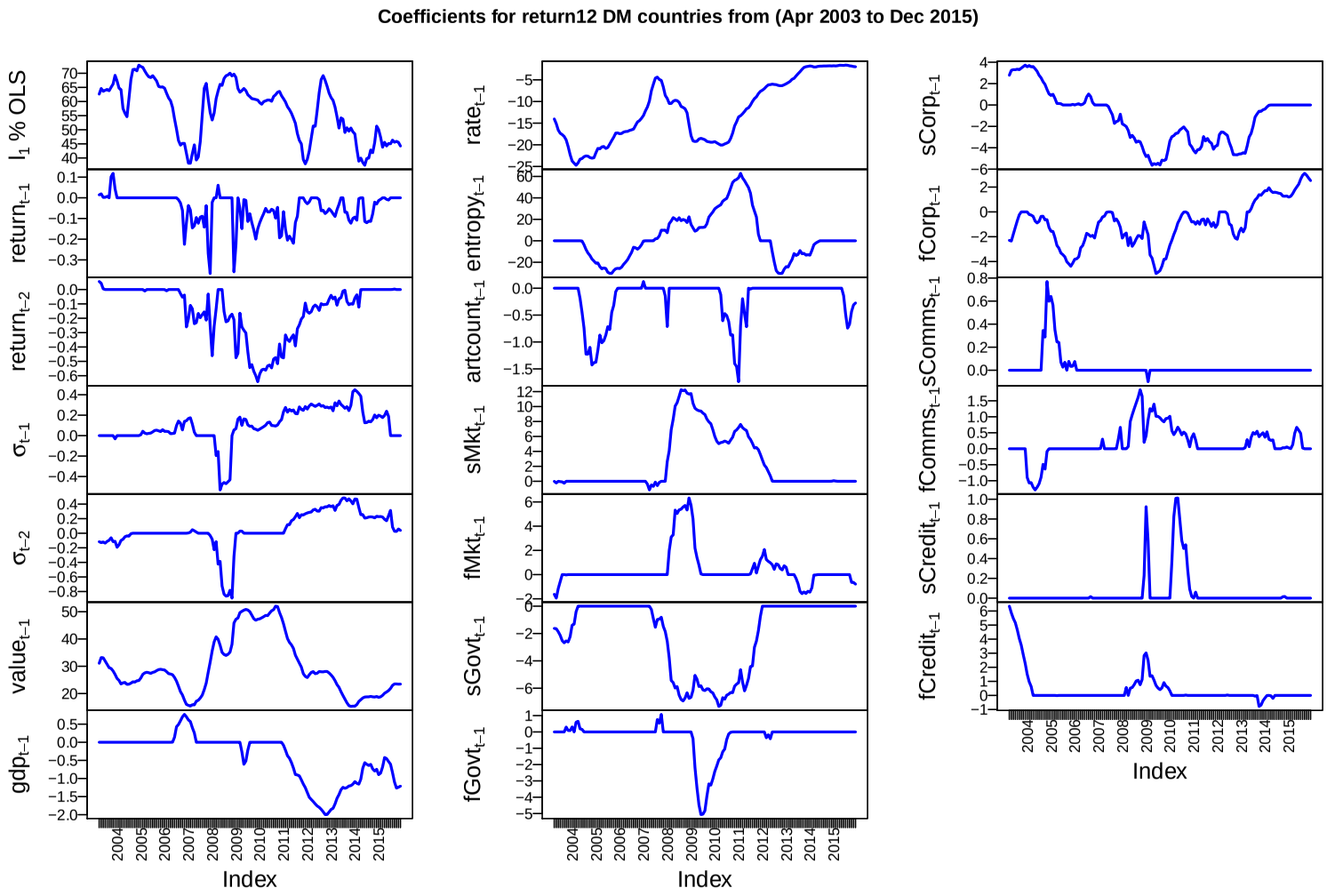

Figure 8: The charts show the time series of coefficient estimates from a rolling lasso regression used for the out-of-sample forecasting tests of 12 month returns for DM countries. The chart labeled "l_1 % of OLS" gives the ratio of the lasso coefficient l_1-norm to the OLS coefficient l_1-norm in every time period.

Coefficient time series from lasso for 12 month returns for EM countries

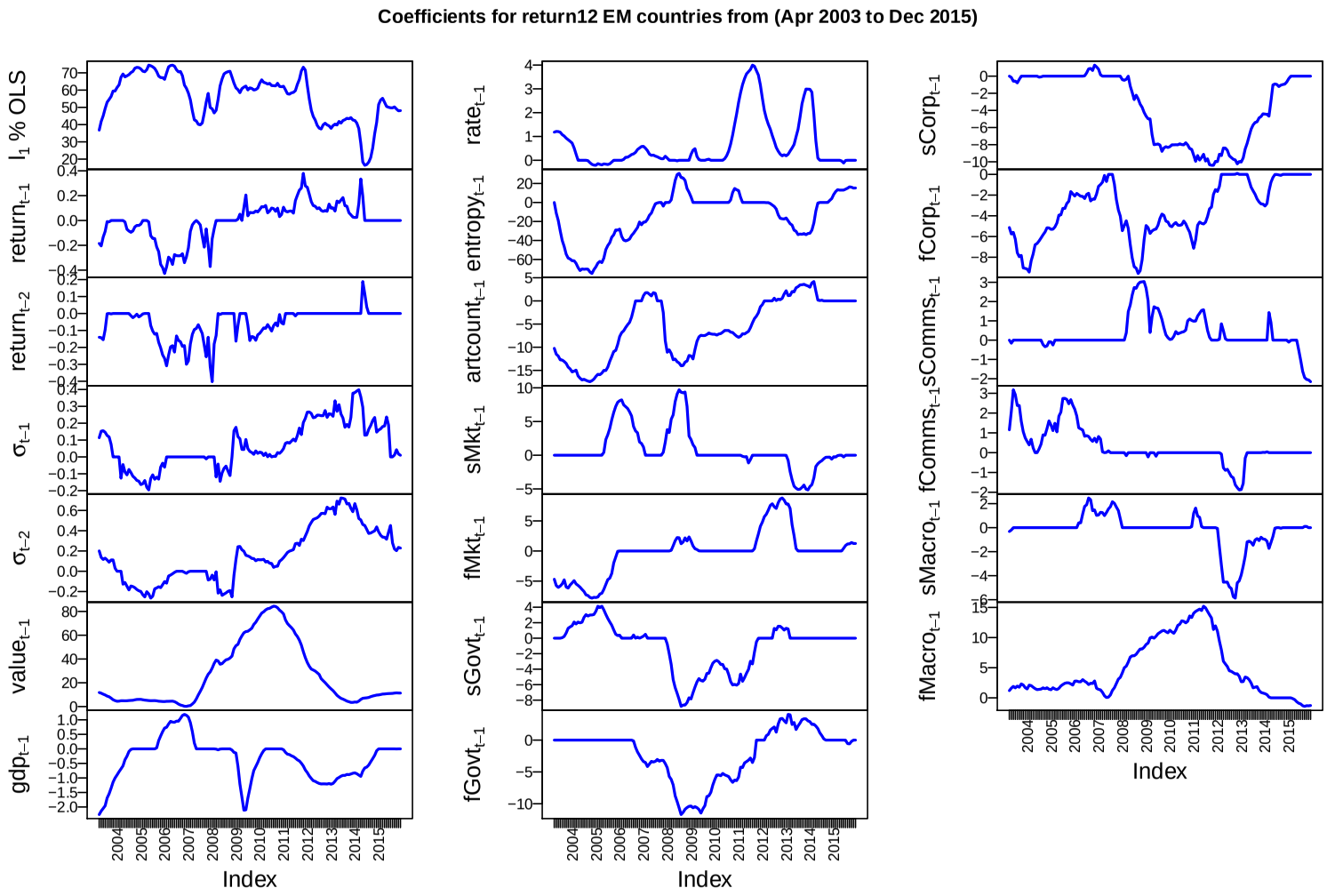

Figure 9: The charts show the time series of coefficient estimates from a rolling lasso regression used for the out-of-sample forecasting tests of 12 month returns for EM countries. The chart labeled "l_1 % of OLS" gives the ratio of the lasso coefficient l_1-norm to the OLS coefficient l_1-norm in every time period.

Coefficient time series from lasso for realized volatility for DM countries

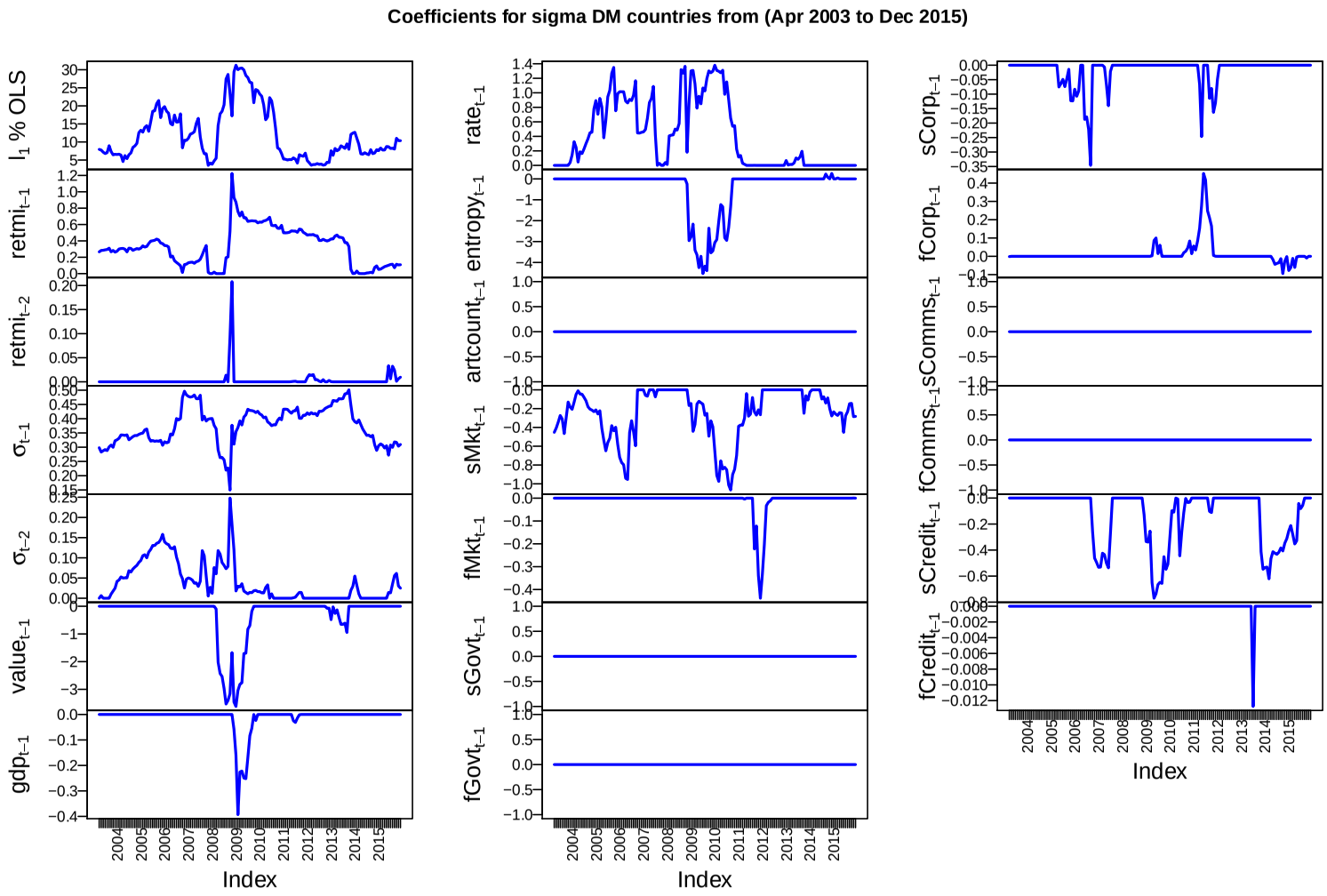

Figure 10: The charts show the time series of coefficient estimates from a rolling lasso regression used for the out-of-sample forecasting tests of realized volatility for DM countries. The chart labeled "l_1 % of OLS" gives the ratio of the lasso coefficient l_1-norm to the OLS coefficient l_1-norm in every time period.

Coefficient time series from lasso for realized volatility for EM countries

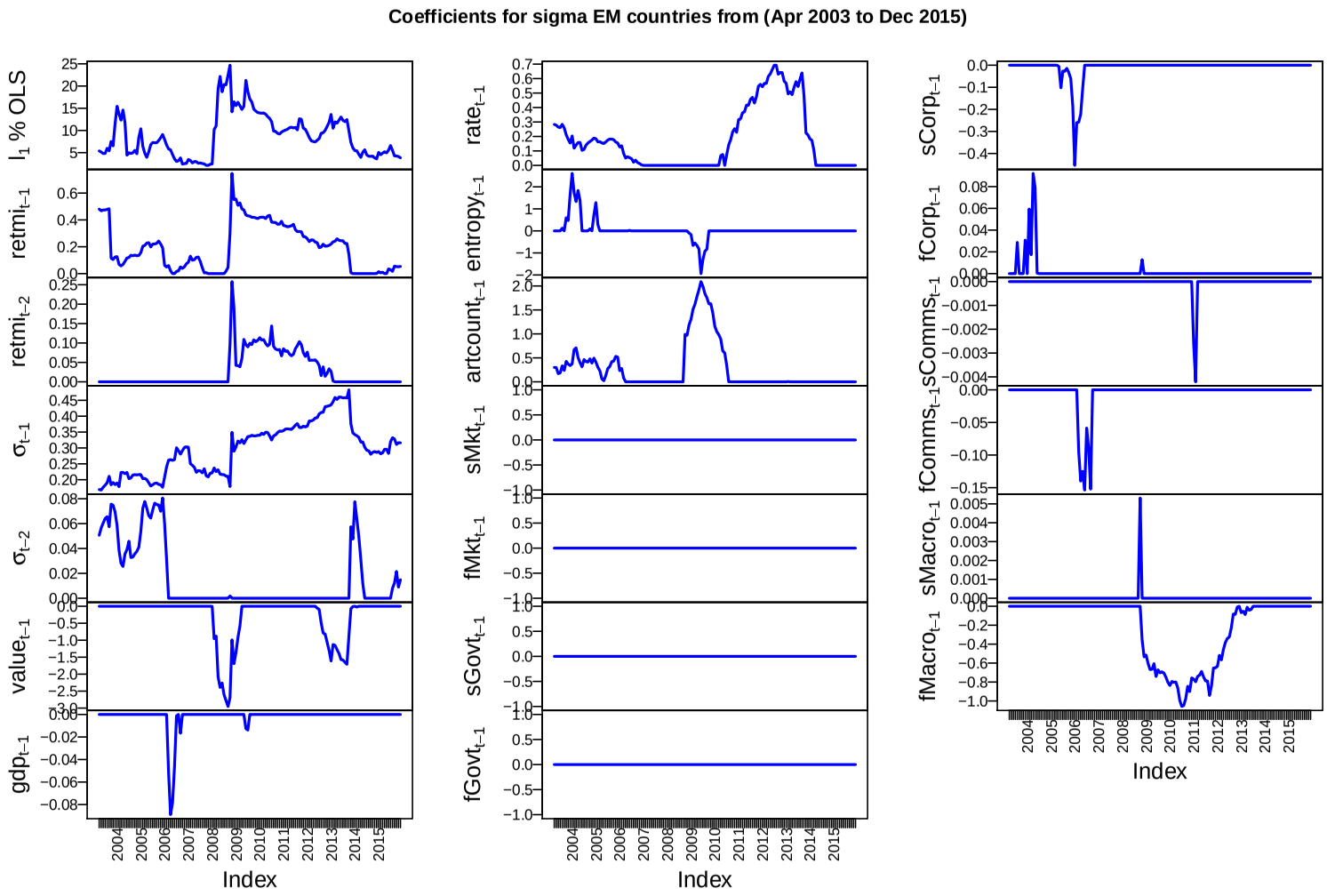

Figure 11: The charts show the time series of coefficient estimates from a rolling lasso regression used for the out-of-sample forecasting tests of realized volatility for EM countries. The chart labeled "l_1 % of OLS" gives the ratio of the lasso coefficient l_1-norm to the OLS coefficient l_1-norm in every time period.

Coefficient time series from lasso for next 12 month drawdown for DM countries

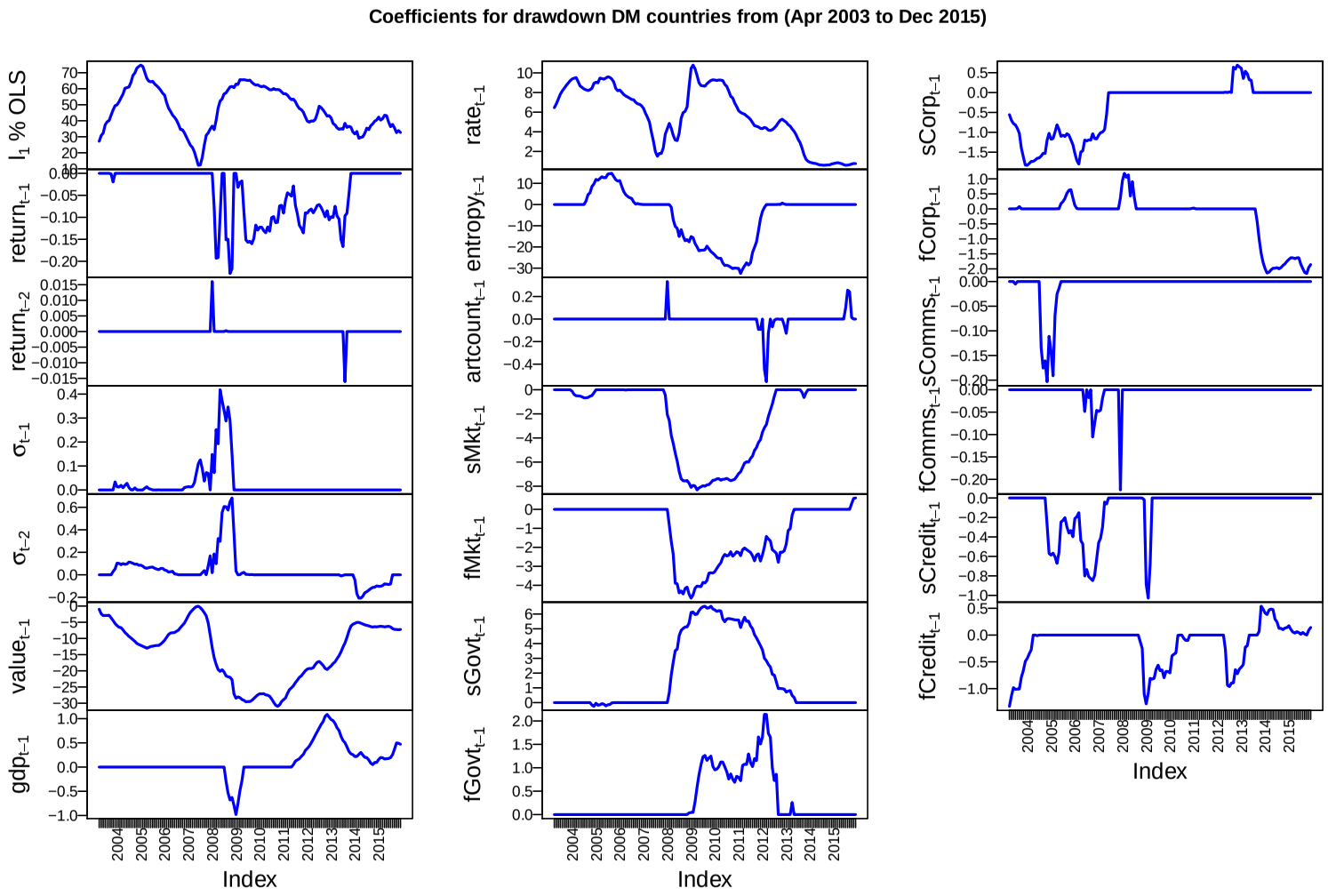

Figure 12: The charts show the time series of coefficient estimates from a rolling lasso regression used for the out-of-sample forecasting tests of next 12 month drawdown for DM countries. The chart labeled "l_1 % of OLS" gives the ratio of the lasso coefficient l_1-norm to the OLS coefficient l_1-norm in every time period.

Coefficient time series from lasso for next 12 month drawdown for EM countries

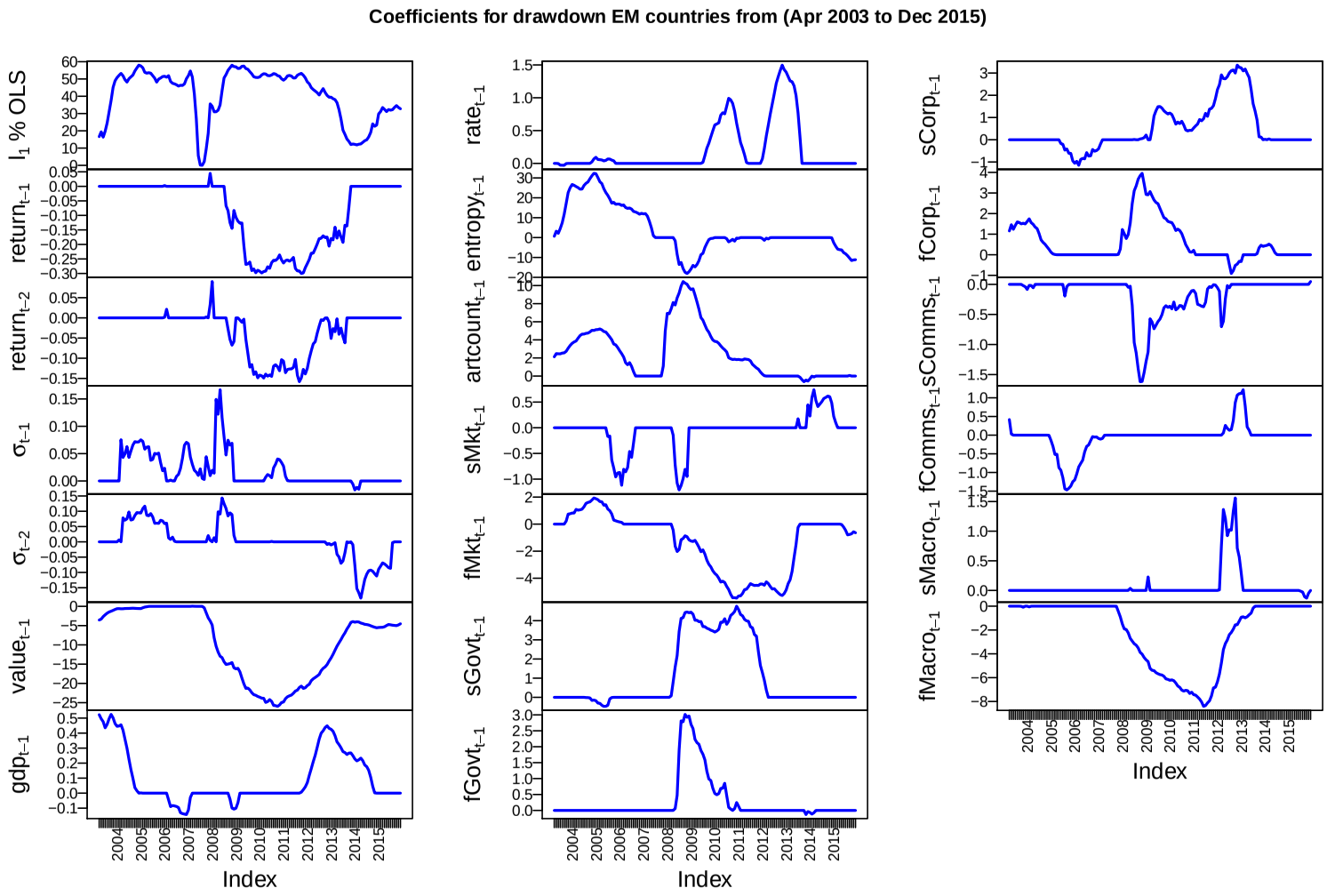

Figure 13: The charts show the time series of coefficient estimates from a rolling lasso regression used for the out-of-sample forecasting tests of next 12 month drawdown for EM countries. The chart labeled "l_1 % of OLS" gives the ratio of the lasso coefficient l_1-norm to the OLS coefficient l_1-norm in every time period.

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 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 as the market benchmark in several sections of the paper. Iceland 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.

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 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 $\#(A \cap B)/\#(A \cup B)$.

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 3: 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 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.

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
$drawdown^N$	Next N-month maximum drawdown in percent (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 5-10 year governemnt bond yields for DM
$dexch$	Change in value of US Dollar in terms of local currency (positive values are local currency depreciations), truncated at $\pm 50\%$
pre	Dummy variable set to 1 if month t is 6 or fewer months prior to an election
$post$	Dummy variable set to 1 if month t is 6 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 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*, *credit*, *commodities* or *corporate events*.

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	0.112	4803
$return^{12}$	8.983	27.341	-38.063	51.641		4792
$return^{24}$	18.783	42.977	-42.921	93.108		4578
$sigma$	19.196	11.427	7.470	41.669	0.619	4815
$drawdown^{12}$	15.305	15.465	-0.000	50.593		4817
$drawdown^{24}$	20.951	19.636	-0.000	61.803		4616
$retmi$	2.188	4.096	0.000	10.939	0.205	4803
$retpl$	2.835	3.914	0.000	10.783	0.050	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
$dexch$	0.024	7.823	-4.663	4.855	0.016	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 6: Data summary using country-month observations from April 1998 – December 2015 (the time frame for our panel regressions). For each variable, the table shows the mean, standard deviation, the 5th and 95th percentiles, the one-lag autocorrelation, and the number N of country-month observations. 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.

Developed markets: Forecasting panel for next 12 month returns

	Base	Sent	SentEnt	Base	Sent	SentEnt	Base	Sent	SentEnt
$sigma_{t-1}$	0.118	0.176	0.183	0.157	0.116	0.112	0.293	0.334*	0.322*
$sigma_{t-2}$	0.053	0.095	0.097	-0.105	-0.048	-0.042	0.327*	0.377**	0.362*
$return_{t-1}$	0.166	0.033	0.029	0.211	0.108	0.106	-0.059	-0.183	-0.180
$return_{t-2}$	0.040	-0.056	-0.058	0.117	0.066	0.063	-0.072	-0.137	-0.136
$value_{t-1}$	21.315***	22.579***	22.680***	15.086***	13.254**	13.297**	25.646***	26.213***	26.299***
gdp_{t-1}	-0.484	-0.873	-0.919	0.216	0.299	0.255	-2.103***	-1.976***	-1.941***
$gdpdeflator_{t-1}$	0.883	0.490	0.468	0.566	0.517	0.477	0.032	0.104	0.112
cp_{t-1}	-0.278***	-0.223**	-0.227**	0.047	-0.010	-0.003	-0.304**	-0.297**	-0.302**
dcp_{t-1}	0.075	0.026	0.022	-0.180	-0.161	-0.161	0.268**	0.252**	0.257**
$rate_{t-1}$	-3.763***	-5.373***	-5.465***	-14.479***	-13.992***	-13.884***	-5.080***	-5.467***	-5.540***
$dexch_{t-1}$	0.038	0.024	0.026	-0.372**	-0.365**	-0.371**	0.060	0.054	0.054
pre	-0.956	-2.533	-2.585	-3.074	-3.013	-3.005	2.804	1.957	2.050
$post$	-0.475	-1.429	-1.477	-2.583	-2.930	-2.865	2.077	1.798	1.877
$entropy_{t-1}$		20.714	23.072		-23.902	-22.602		21.872*	21.303*
$artcount_{t-1}$		-0.586	-0.631		0.189	0.117		0.100	0.114
$sMkt_{t-1}$		2.765	3.017		-1.569	-2.315		5.235*	5.711**
$fMkt_{t-1}$		0.770	0.873		-4.118	-4.222		1.302	1.285
$sGovt_{t-1}$		-3.367	-3.554*		-0.203	-0.560		-3.896*	-2.792
$fGovt_{t-1}$		-3.768	-3.914		-3.370	-3.601		-6.289	-6.019
$sCorp_{t-1}$		-2.691	-1.775		3.085	3.777		-2.720	-2.916*
$fCorp_{t-1}$		-5.882*	-5.551*		-6.097	-5.965		-0.080	0.097
$sComms_{t-1}$		1.126	0.923		0.154	0.378		0.566	0.447
$fComms_{t-1}$		0.849	0.835		-1.491	-1.326		1.495	1.477
$sCredit_{t-1}$		3.703*	3.175		-1.083	-0.986		-0.852	-1.860
$fCredit_{t-1}$		0.123	0.116		0.741	0.716		-1.104	-1.312
$R2$	0.163	0.214	0.213	0.266	0.298	0.299	0.452	0.471	0.472
$start$	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 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. All text measures except *entropy* are normalized to unit variance. 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, and "***","**", and "*" indicate significance 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 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. All text measures except *entropy* are normalized to unit variance. 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, and "***","**", and "*" indicate significance at the 1%, 5% and 10% levels respectively.

Developed markets: Forecasting panel for volatility

	Base	Sent	SentEnt	Base	Sent	SentEnt	Base	Sent	SentEnt
$sigma_{t-1}$	0.423***	0.398***	0.397***	0.387***	0.348***	0.348***	0.403***	0.375***	0.374***
$sigma_{t-2}$	0.136**	0.120**	0.120**	0.164***	0.148***	0.148***	0.077	0.046	0.046
$retmi_{t-1}$	0.563***	0.526***	0.525***	0.512***	0.454***	0.459***	0.580***	0.555***	0.553***
$retmi_{t-2}$	-0.009	-0.026	-0.027	-0.075	-0.087	-0.084	0.077	0.101	0.098
$value_{t-1}$	-2.163***	-2.555***	-2.591***	-1.454	-2.440**	-2.425**	-2.435**	-2.497**	-2.532**
gdp_{t-1}	-0.183*	-0.135	-0.127	-0.100	0.013	0.016	-0.180	-0.148	-0.147
$gdpdeflator_{t-1}$	-0.019	0.019	0.022	0.046	0.116	0.117	0.181	0.159	0.156
cp_{t-1}	0.011	0.008	0.008	-0.026	-0.018	-0.017	0.030	0.033*	0.032*
dcp_{t-1}	-0.018	-0.012	-0.011	0.007	0.009	0.009	-0.048**	-0.048*	-0.045*
$rate_{t-1}$	0.796***	0.872***	0.889***	1.840***	1.832***	1.846***	0.815***	0.854**	0.870**
$dexch_{t-1}$	0.053*	0.053*	0.052*	0.001	0.010	0.008	0.057*	0.056*	0.056*
pre	0.052	0.010	0.021	-0.092	-0.050	-0.047	0.026	-0.160	-0.128
$post$	0.088	0.077	0.100	0.651*	0.791**	0.797**	-0.467	-0.519	-0.501
$entropy_{t-1}$		-2.107	-2.636		-0.040	-0.850		-1.787	-2.465
$artcount_{t-1}$		0.009	0.026		0.003	-0.002		-0.609	-0.581
$sMkt_{t-1}$		-0.818	-0.859		-1.525*	-1.428*		-0.484	-0.625
$fMkt_{t-1}$		-0.338	-0.331		-0.338	-0.299		-0.019	-0.064
$sGovt_{t-1}$		0.645	0.594		0.033	0.017		1.188	1.048
$fGovt_{t-1}$		0.325	0.297		0.060	-0.007		0.736	0.688
$sCorp_{t-1}$		0.093	0.082		0.070	0.052		-0.671	-0.615
$fCorp_{t-1}$		0.057	0.023		-0.431	-0.494		-0.411	-0.443
$sComms_{t-1}$		0.086	0.133		-0.120	-0.080		0.303	0.326
$fComms_{t-1}$		-0.224	-0.200		-0.175	-0.182		-0.292	-0.267
$sCredit_{t-1}$		-0.553	-0.593		0.066	-0.032		-0.380	-0.349
$fCredit_{t-1}$		-0.102	-0.122		0.313	0.289		-0.422	-0.428
$R2$	0.471	0.477	0.477	0.447	0.458	0.457	0.454	0.459	0.46
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 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. All text measures except *entropy* are normalized to unit variance. 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, and "***","**", and "*" indicate significance 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 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. All text measures except *entropy* are normalized to unit variance. 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, and "***","**", and "*" indicate significance at the 1%, 5% and 10% levels respectively.

Developed markets: Forecasting panel for drawdowns

	Base	Sent	SentEnt	Base	Sent	SentEnt	Base	Sent	SentEnt
$sigma_{t-1}$	0.114	0.028	0.024	0.074	0.034	0.037	0.016	-0.080	-0.074
$sigma_{t-2}$	0.157*	0.096	0.098	0.154**	0.098	0.096	0.064	-0.030	-0.018
$return_{t-1}$	-0.283***	-0.166*	-0.167*	-0.168*	-0.089	-0.092	-0.255**	-0.124	-0.126
$return_{t-2}$	-0.089	-0.019	-0.018	-0.026	-0.005	-0.004	-0.087	-0.035	-0.031
$value_{t-1}$	-12.351***	-13.909***	-13.980***	-5.980***	-6.590***	-6.568***	-15.914***	-16.823***	-16.983***
gdp_{t-1}	0.177	0.424	0.457*	-0.085	0.048	0.080	0.814**	0.808**	0.794**
$gdpdeflator_{t-1}$	-0.290	-0.084	-0.072	-0.355	-0.242	-0.218	0.377	0.165	0.158
cp_{t-1}	0.187***	0.156***	0.158***	-0.007	0.034	0.028	0.151**	0.147**	0.147**
dcp_{t-1}	-0.070	-0.038	-0.034	0.074	0.068	0.068	-0.143	-0.115	-0.113
$rate_{t-1}$	3.116***	3.940***	4.021***	7.668***	7.461***	7.392***	4.278***	4.514***	4.622***
$dexch_{t-1}$	-0.039	-0.025	-0.027	0.051	0.057	0.056	-0.036	-0.024	-0.025
pre	1.397	1.962	1.998	2.012	2.065	2.072	0.087	0.283	0.273
$post$	0.196	0.401	0.443	0.448	0.813	0.787	-0.444	-0.913	-0.954
$entropy_{t-1}$		-11.132*	-13.343**		11.432*	9.577		-11.476*	-12.288*
$artcount_{t-1}$		0.078	0.145		-0.064	-0.010		-0.912	-0.892
$sMkt_{t-1}$		-4.235***	-4.434***		-0.516	0.060		-6.819***	-7.337***
$fMkt_{t-1}$		-0.859	-0.882		0.690	0.817		-1.357	-1.413
$sGovt_{t-1}$		4.638***	4.366***		0.514	0.680		6.001***	4.928***
$fGovt_{t-1}$		3.942***	3.899***		0.968	1.056		5.583***	5.268***
$sCorp_{t-1}$		1.137	0.563		-1.540*	-2.022**		0.350	0.501
$fCorp_{t-1}$		2.621**	2.324**		1.457	1.341		-0.038	-0.286
$sComms_{t-1}$		-0.790	-0.512		-0.631	-0.655		0.261	0.393
$fComms_{t-1}$		-0.077	0.000		0.420	0.329		0.365	0.443
$sCredit_{t-1}$		-1.603	-1.156		-0.240	-0.384		1.669*	2.399***
$fCredit_{t-1}$		0.367	0.364		-0.389	-0.384		0.874	0.953
$R2$	0.26	0.32	0.32	0.403	0.45	0.453	0.385	0.444	0.444
$start$	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 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. All text measures except *entropy* are normalized to unit variance. 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, and "***","**", and "*" indicate significance 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 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. All text measures except *entropy* are normalized to unit variance. 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, and "***","**", and "*" indicate significance at the 1%, 5% and 10% levels respectively.

	Emerging markets				Developed markets			
	return	*return*^{12}	*sigma*	*drawdown*	*return*	*return*^{12}	*sigma*	*drawdown*
$sigma_{t-1}$		+	+				+	
$sigma_{t-2}$		+	+				+	
$return_{t-1}$	+		∅	-			∅	-
$return_{t-2}$			∅				∅	
$retmi_{t-1}$	∅	∅	+	∅	∅	∅	+	∅
$retmi_{t-2}$	∅	∅		∅	∅	∅		∅
$value_{t-1}$			-	-	+	+	-	-
gdp_{t-1}	-							
$gdpdeflator_{t-1}$								
cp_{t-1}	-	-	-		-	-		+
dcp_{t-1}			+					
$rate_{t-1}$		-	+	+	-	-	+	+
$dexch_{t-1}$			-	-			+	
pre								
$post$								
$entropy_{t-1}$					+			-
$artcount_{t-1}$		-	+	+	+			
$sMkt_{t-1}$					+			-
$fMkt_{t-1}$								
$sGovt_{t-1}$				+	-			+
$fGovt_{t-1}$	-		+	+				+
$sCorp_{t-1}$	-	-		+				
$fCorp_{t-1}$	-	-	+	+	-	-		+
$sComms_{t-1}$	+			-				
$fComms_{t-1}$								
$sMacro_{t-1}$								
$fMacro_{t-1}$	+	+		-				
$sCredit_{t-1}$						+		
$fCredit_{t-1}$								

Table 13: This table summarizes the coefficient estimates from all four panel specifications in the paper: *return*, *return*^{12}, *sigma*, and *drawdown*. A "+" ("-") indicates a positive (negative) coefficient estimate significant at the 10% lever or better; ∅ indicates the variable is not present in a given specification. The table shows results for the developed and emerging market regressions over the entire sample.

	Base EM	Full EM	Base DM	Full DM
gdp_{t-1}	-0.028***	-0.009	-0.081***	-0.029***
$gdpdeflator_{t-1}$	-0.005	-0.006	-0.030**	-0.015
cp_{t-1}	0.014***	0.010***	0.001	-0.001
dcp_{t-1}	-0.004	-0.005	-0.008***	-0.004***
$rate_{t-1}$	0.007	0.002	-0.188***	-0.226***
$dexch_{t-1}$	0.033***	0.016	-0.002	-0.003*
pre	0.131**	0.119*	0.024	0.036
$post$	0.211***	0.171**	0.162***	0.096*
$entropy_{t-1}$		-1.236***		-0.336
$artcount_{t-1}$		0.106***		0.009
$sMkt_{t-1}$		-0.164***		-0.331***
$fMkt_{t-1}$		-0.137**		0.034
$sGovt_{t-1}$		-0.069		0.137**
$fGovt_{t-1}$		-0.013		0.198**
$sCorp_{t-1}$		-0.063		0.165***
$fCorp_{t-1}$		0.049		0.194***
$sComms_{t-1}$		-0.036		-0.020
$fComms_{t-1}$		0.028		-0.006
$sMacro_{t-1}$		-0.059		
$fMacro_{t-1}$		-0.151***		
$sCredit_{t-1}$				-0.327***
$fCredit_{t-1}$				0.016
$R2$	0.118	0.227	0.136	0.336
$start$	Apr 1998	May 1998	Apr 1998	May 1998
end	Dec 2015	Dec 2015	Dec 2015	Dec 2015
$Nobs$	1220	1215	2249	2240
$stderr$	by time	by time	by time	by time

Table 14: Regression of the BBD country-level policy uncertainty measures at time 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, 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. 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, and "***","**", and "*" indicate significance 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
$sigma_{t-1}$	0.107	0.170	0.297**	0.307**	0.432***	0.421***	0.386***	0.383***	0.077	0.033	-0.088	-0.094
$sigma_{t-2}$	0.034	0.054	0.161	0.162	0.101	0.096	0.058	0.057	0.177**	0.163**	0.043	0.043
$return_{t-1}$	0.082	0.075	-0.199	-0.195					-0.315***	-0.310***	-0.110	-0.112
$return_{t-2}$	0.003	-0.017	-0.137	-0.139					-0.128	-0.114	-0.034	-0.033
$retmi_{t-1}$					0.650***	0.638***	0.579***	0.578***				
$retmi_{t-2}$					0.007	0.000	-0.006	-0.007				
$value_{t-1}$	27.918***	30.337***	32.120***	32.614***	-2.784**	-3.310***	-3.788***	-3.969***	-15.749***	-17.447***	-19.845***	-20.138***
gdp_{t-1}	-1.185**	-1.205**	-1.350***	-1.346***	-0.282*	-0.278*	-0.209	-0.210	0.277	0.291	0.391	0.388
$gdpdeflator_{t-1}$	3.393***	3.406***	3.074***	3.091***	-0.106	-0.109	-0.163	-0.169	-1.061***	-1.070***	-1.096***	-1.106***
cp_{t-1}	-0.227**	-0.233**	-0.216*	-0.218*	0.016	0.018	0.018	0.018	0.179***	0.183***	0.165***	0.166***
dcp_{t-1}	0.243	0.240	0.235	0.235	-0.036**	-0.036**	-0.036**	-0.036**	-0.146*	-0.144*	-0.139*	-0.139*
$rate_{t-1}$	-3.044**	-3.887***	-3.708***	-3.936***	0.824***	1.010***	0.811**	0.895**	2.685***	3.277***	3.053***	3.188***
$dexch_{t-1}$	0.077***	0.068***	0.062***	0.060***	0.034	0.036	0.032	0.033	-0.038***	-0.032***	-0.030***	-0.029**
pre	-2.297	-2.187	-3.357	-3.340	-0.188	-0.211	-0.428	-0.434	2.325	2.248	2.559	2.549
post	-3.824	-3.420	-4.117	-4.013	0.042	-0.047	-0.080	-0.118	1.936	1.653	1.847	1.786
EPU_{t-1}		-2.798*		-0.761		0.618*		0.280		1.964**		0.451
$entropy_{t-1}$			-1.983	-1.685			-1.495	-1.609			0.515	0.338
$artcount_{t-1}$			-1.970*	-1.951*			-0.253	-0.260			0.509	0.498
$sMkt_{t-1}$			9.467***	9.314***			-1.059	-1.007			-7.082***	-6.991***
$fMkt_{t-1}$			-6.570	-6.611			-0.393	-0.377			0.984	1.009
$sGovt_{t-1}$			-6.704***	-6.645***			1.140	1.120			6.227***	6.192***
$fGovt_{t-1}$			-14.728**	-14.737**			0.285	0.292			7.276***	7.282***
$sCorp_{t-1}$			-3.006*	-2.883*			-0.264	-0.309			0.989	0.916
$fCorp_{t-1}$			-10.767***	-10.659***			-0.343	-0.381			3.623***	3.559***
$sComms_{t-1}$			0.236	0.222			-0.028	-0.023			-0.702	-0.694
$fComms_{t-1}$			-1.097	-1.121			-0.589*	-0.580*			0.360	0.375
$sCredit_{t-1}$			0.804	0.586			-0.723	-0.642			-0.751	-0.622
$fCredit_{t-1}$			-2.813	-2.830			-0.292	-0.285			1.133	1.143
R2	0.266	0.274	0.344	0.344	0.476	0.478	0.487	0.487	0.295	0.306	0.402	0.402
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 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 included: USA, Canada, Germany, UK, Italy, France, Spain, Netherlands, Japan, Australia, and Ireland. All text measures except *entropy* are normalized to unit variance. 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, and "***","**", and "*" indicate significance 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	0.195	0.309***	0.350***	0.417***	0.419***	0.399***	0.400***	0.101*	0.094*	0.028	0.026
$sigma_{t-2}$	0.435**	0.443**	0.497**	0.500***	0.074	0.074	0.063	0.063	-0.022	-0.023	-0.051	-0.051
$return_{t-1}$	0.405	0.389*	0.238	0.255					-0.194**	-0.192**	-0.132	-0.134
$return_{t-2}$	0.167	0.110	0.074	0.033					-0.068	-0.059	-0.036	-0.033
$retmi_{t-1}$					0.690***	0.693***	0.729***	0.730***				
$retmi_{t-2}$					-0.011	-0.006	0.015	0.017				
$value_{t-1}$	11.722	14.061	16.331***	18.782***	-2.497***	-2.398***	-2.498***	-2.438***	-1.798	-2.136	-2.909**	-3.059***
gdp_{t-1}	-2.678	-2.757	-3.000*	-3.000*	-0.119	-0.122	-0.212*	-0.213*	0.931**	0.942**	0.995***	0.995***
$gdpdeflator_{t-1}$	0.891	0.877	0.109	0.063	0.162**	0.161**	0.208**	0.207**	0.228	0.230	0.433***	0.436***
cp_{t-1}	-0.571***	-0.468**	-0.339	-0.251	-0.099***	-0.094***	-0.163***	-0.160***	0.148*	0.133**	-0.009	-0.015
dcp_{t-1}	-0.703**	-0.725**	-0.656**	-0.690**	0.099*	0.098*	0.097*	0.096*	0.346**	0.349**	0.359**	0.361**
$rate_{t-1}$	-1.379	-1.382	-1.229	-1.240	0.335**	0.335**	0.453***	0.453***	0.985	0.986	1.013**	1.014**
$dexch_{t-1}$	1.192**	1.304**	0.942	1.040*	-0.290*	-0.285*	-0.216	-0.213	-0.557***	-0.574***	-0.413*	-0.419*
pre	-17.788**	-17.439***	-15.631**	-15.315**	0.806	0.821	0.577	0.585	6.089**	6.038**	5.051**	5.031**
post	-7.162	-6.803	-7.522	-7.333	1.533	1.552	1.382	1.388	3.728	3.676	3.379	3.368
EPU_{t-1}		-5.599		-5.520*		-0.245		-0.140		0.810		0.337
$entropy_{t-1}$			-7.598	-15.561			2.093	1.891			15.733	16.219
$artcount_{t-1}$			-4.397	-3.974			1.636**	1.647**			3.429	3.404
$sMkt_{t-1}$			-1.859	-3.329			2.193**	2.159**			1.355	1.445
$fMkt_{t-1}$			2.053	0.320			-0.806	-0.850			-2.557	-2.451
$sGovt_{t-1}$			-6.064	-6.232			1.431*	1.426*			4.173***	4.183***
$fGovt_{t-1}$			-1.147	-1.635			0.025	0.012			0.170	0.200
$sCorp_{t-1}$			4.960	5.516			-0.932	-0.919			-1.341	-1.375
$fCorp_{t-1}$			-1.494	-1.713			-0.215	-0.221			-0.855	-0.842
$sComms_{t-1}$			-1.091	-1.086			-1.004*	-1.004*			-0.782	-0.783
$fComms_{t-1}$			11.862	11.625			-1.889*	-1.895*			-4.862	-4.848
$sMacro_{t-1}$			4.325	3.959			0.618	0.608			-2.060	-2.037
$fMacro_{t-1}$			10.272***	9.763***			-0.035	-0.048			-4.812***	-4.781***
R2	0.171	0.18	0.246	0.254	0.459	0.459	0.472	0.472	0.111	0.111	0.203	0.203
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	1161	1161	1157	1157	1161	1161	1157	1157	1161	1161	1157	1157
stderr	both	both	both	both	by time	by time	by time	by time	both	both	both	both

Table 16: 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 included: Brazil, Chile, China, India, Russia, and South Korea. All text measures except *entropy* are normalized to unit variance. 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, and "***","**", and "*" indicate significance at the 1%, 5% and 10% levels respectively.

Out of sample forecasting tests for developed markets sample

LHS	AvgNumObs	Naive	Base	CM
$return_t$	1272.38	6.67	6.78	6.78
$return_t^{12}$	1271.59	28.74	24.99	24.37
$sigma_t$	1272.38	11.54	8.89	8.89
$drawdown_t$	1272.38	17.57	14.92	14.32

Out of sample forecasting tests for emerging markets sample

LHS	AvgNumObs	Naive	Base	CM
$return_t$	1396.29	8.48	8.54	8.51
$return_t^{12}$	1396.29	42.17	41.59	39.45
$sigma_t$	1393.25	12.05	10.20	10.18
$drawdown_t$	1396.29	18.29	17.66	16.78

Table 17: The table shows the root mean squared errors from out of sample forecasting tests of different models for returns, volatilities, and drawdowns. The three 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 across the five topic areas in addition to the variables from the *Base* model. The forecast in month t comes from the model estimated in months $t-60$ though $t-1$. The forecasting model is estimated using a lasso regression for model selection and coefficient shrinkage. The column *AvgNumObs* is the average number of observations in each 60 month estimation window.

Out of sample forecasting tests for developed markets sample – Baker-Bloom-Davis subsample

LHS	AvgNumObs	Naive	Base	EPU	CM	Both
$return_t$	594.35	6.21	6.32	6.33	6.33	6.34
$return_t^{12}$	594.35	25.30	22.41	22.20	21.38	21.31
$sigma_t$	594.35	11.73	8.98	8.97	8.95	8.94
$drawdown_t$	594.35	16.63	14.12	13.88	12.86	12.88

Out of sample forecasting tests for emerging markets sample – Baker-Bloom-Davis subsample

LHS	AvgNumObs	Naive	Base	EPU	CM	Both
$return_t$	337.25	8.57	8.70	8.70	8.71	8.71
$return_t^{12}$	337.25	43.05	42.59	42.58	40.35	40.38
$sigma_t$	337.25	13.53	10.75	10.75	10.69	10.69
$drawdown_t$	337.25	17.94	17.53	17.57	16.49	16.51

Table 18: The table shows the root mean squared errors from out of sample forecasting tests of different models for returns, volatilities, and drawdowns. The five models are: *Naive* which uses only in sample country fixed effects as the forecasting variables; *Base* (or baseline) which includes lagged macroeconomic and lagged market variables as the regressors; *EPU* which includes the Baker-Bloom-Davis EPU measure in addition to the baseline model; *CM* which includes country specific article counts, entropy, sentiment and frequency measures across the five topic areas in addition to the baseline model; and *Both* which includes the EPU measure in addition to all the variables in *CM*. The forecast in month t comes from the model estimated in months $t-60$ though $t-1$. The forecasting model is estimated using a lasso regression for model selection and coefficient shrinkage. The column *AvgNumObs* is the average number of observations in each 60 month estimation window.

Supplementary Appendix

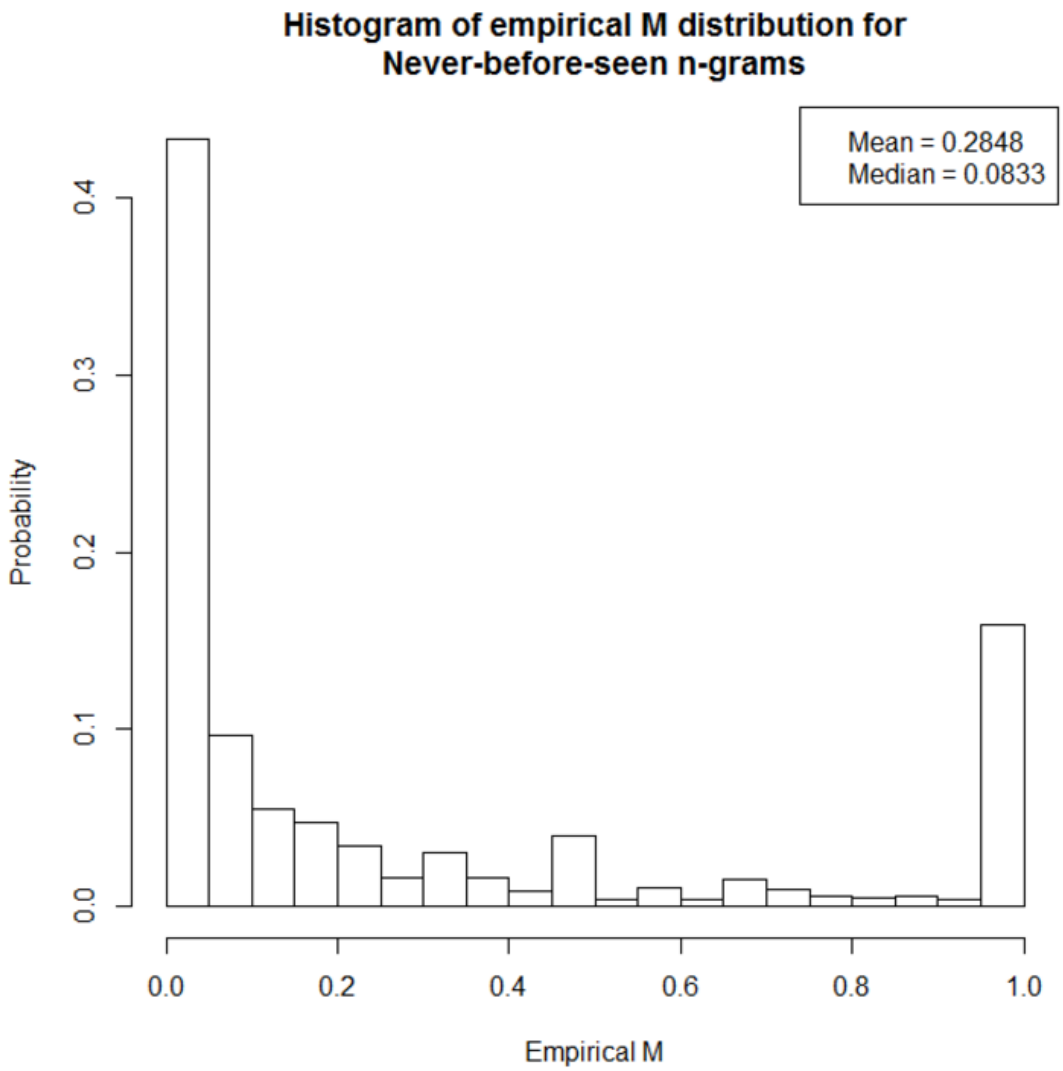

Figure A1: We observe 99.96 million never-before-seen n-grams from April 1998 through December of 2014. For a given never-before-seen n-gram in month t, we observe its empirical m in the subsample beginning in month t until the end of the full sample (December 2015). Collecting these m's across the 99.96 million never-before-seen n-grams we construct the empirical m distribution for a never-before-seen n-gram.

Coefficient time series from lasso for monthly returns for DM countries

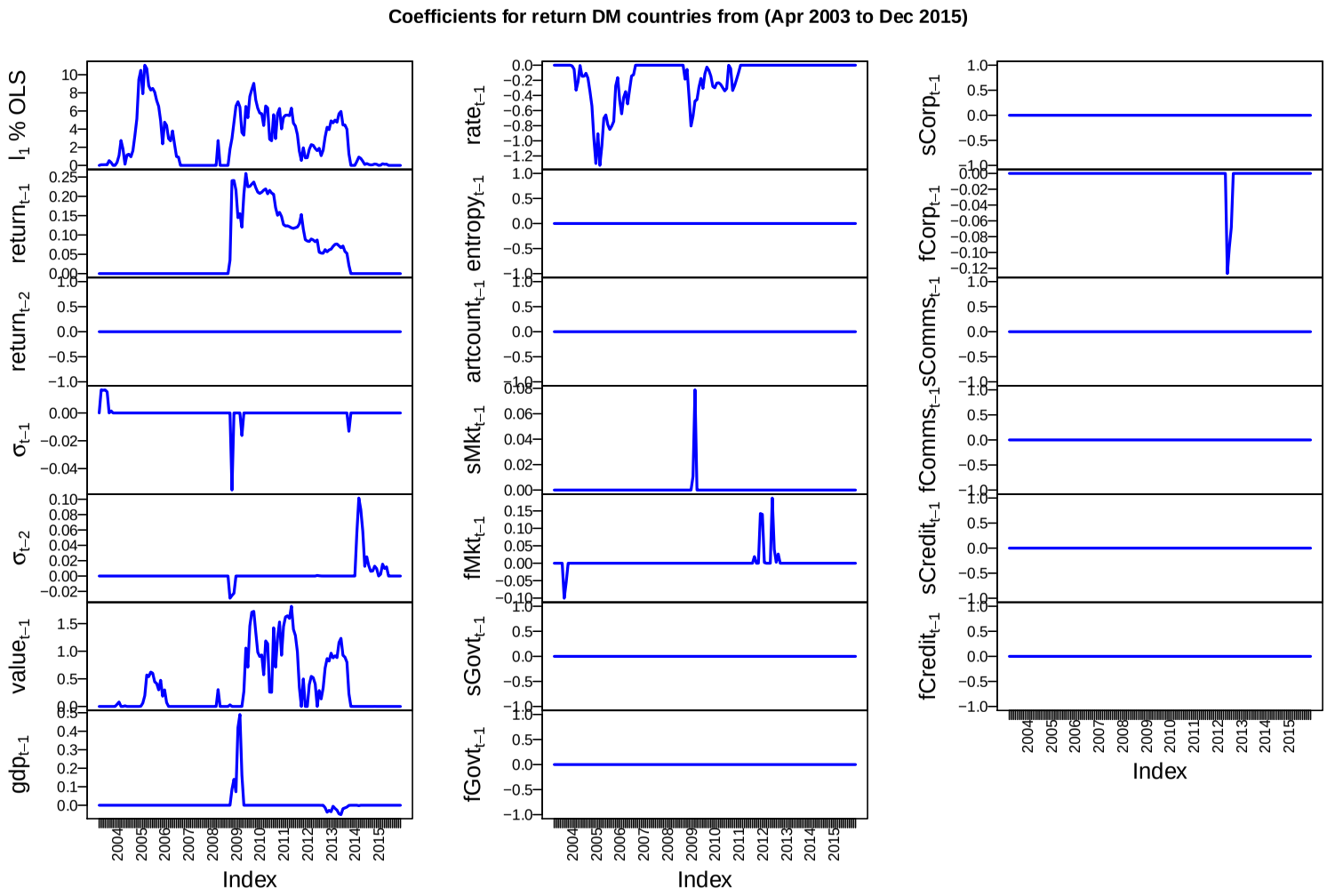

Figure A2: The charts show the time series of coefficient estimates from a rolling lasso regression used for the out-of-sample forecasting tests of monthly returns for DM countries. The chart labeled "l_1 % of OLS" gives the ratio of the lasso coefficient l_1-norm to the OLS coefficient l_1-norm in every time period.

Coefficient time series from lasso for monthly returns for EM countries

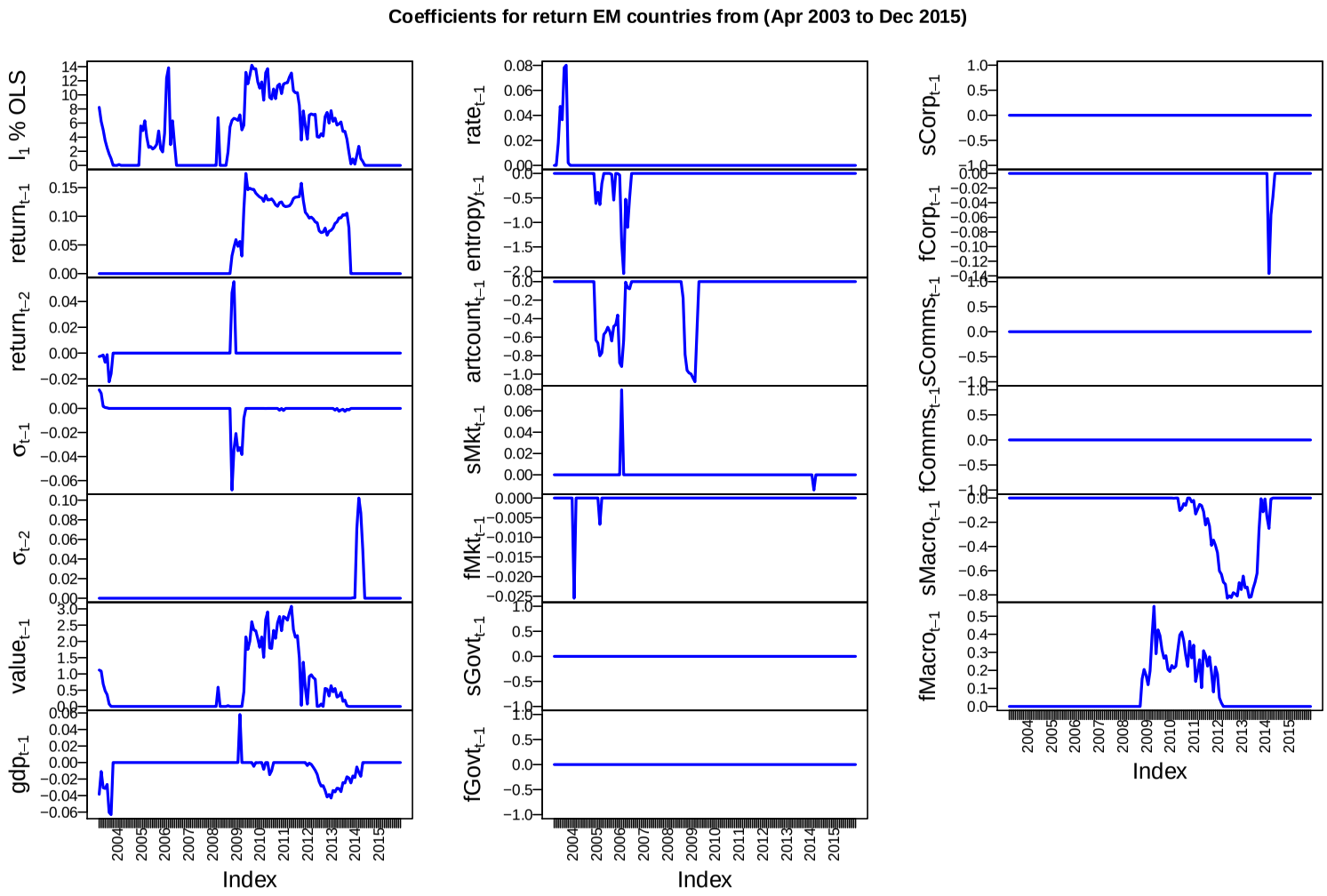

Figure A3: The charts show the time series of coefficient estimates from a rolling lasso regression used for the out-of-sample forecasting tests of monthly returns for EM countries. The chart labeled "l_1 % of OLS" gives the ratio of the lasso coefficient l_1-norm to the OLS coefficient l_1-norm in every time period.

Summary of coefficients for emerging market panels

Time window	98-15 *return*	98-15 $return^{12}$	98-15 *sigma*	98-15 *drawdown*	98-07 *return*	98-07 $return^{12}$	98-07 *sigma*	98-07 *drawdown*	07-15 *return*	07-15 $return^{12}$	07-15 *sigma*	07-15 *drawdown*
$sigma_{t-1}$		+	+				+			+	+	
$sigma_{t-2}$		+	+				+		+	+		-
$return_{t-1}$	+		∅	-			∅			+	∅	-
$return_{t-2}$			∅				∅				∅	-
$retmi_{t-1}$	∅	∅	+	∅	∅	∅	+	∅	∅	∅	+	∅
$retmi_{t-2}$	∅	∅		∅	∅	∅		∅	∅	∅		∅
$value_{t-1}$			-	-	+	+	-	-			-	-
gdp_{t-1}	-				-	-		+	-	-		+
$gdpdeflator_{t-1}$					-				-			+
cp_{t-1}	-	-	-								-	
dcp_{t-1}			+								+	
$rate_{t-1}$		-	+	+			+		-		+	+
$dexch_{t-1}$			-	-			-			+		-
pre					-							
post					-					-		
$entropy_{t-1}$					-		+	+	+			-
$artcount_{t-1}$		-	+	+	-	-		+			+	+
$sMkt_{t-1}$							+					
$fMkt_{t-1}$					-	-	+	+		+		-
$sGovt_{t-1}$				+								
$fGovt_{t-1}$	-		+	+	-	-		+				
$sCorp_{t-1}$	-	-		+					-	-		+
$fCorp_{t-1}$	-	-	+	+		-	+	+				
$sComms_{t-1}$	+			-				-	+			
$fComms_{t-1}$									+			
$sMacro_{t-1}$												
$fMacro_{t-1}$	+	+		-					+		-	-

Table A1: This table summarizes the coefficient estimates from all four panel specifications in the paper: *return*, $return^{12}$, *sigma*, and *drawdown*. A "+" ("-") indicates a positive (negative) coefficient estimate significant at the 10% lever or better; ∅ indicates the variable is not present in a given specification. The table shows results for the emerging market regressions over the entire sample, as well as the two sub-periods.

Summary of coefficients for developed market panels

Time window	98-15 *return*	98-15 $return^{12}$	98-15 *sigma*	98-15 *drawdown*	98-07 *return*	98-07 $return^{12}$	98-07 *sigma*	98-07 *drawdown*	07-15 *return*	07-15 $return^{12}$	07-15 *sigma*	07-15 *drawdown*
$sigma_{t-1}$			+		+		+			+	+	
$sigma_{t-2}$			+				+		+	+		
$return_{t-1}$			∅	-	+		∅				∅	
$return_{t-2}$			∅				∅				∅	
$retmi_{t-1}$	∅	∅	+	∅	∅	∅	+	∅	∅	∅	+	∅
$retmi_{t-2}$	∅	∅		∅	∅	∅		∅	∅	∅		∅
$value_{t-1}$	+	+	-	-		+	-	-	+	+	-	-
gdp_{t-1}										-		+
$gdpdeflator_{t-1}$												
cp_{t-1}	-	-		+					-	-	+	+
dcp_{t-1}					-				+	+	-	
$rate_{t-1}$	-	-	+	+	-	-	+	+	-	-	+	+
$dexch_{t-1}$			+			-					+	
pre												
post							+					
$entropy_{t-1}$	+			-				+	+	+		-
$artcount_{t-1}$	+											
$sMkt_{t-1}$	+			-			-		+	+		-
$fMkt_{t-1}$												
$sGovt_{t-1}$	-			+					-	-		+
$fGovt_{t-1}$				+					-			+
$sCorp_{t-1}$								-	+			
$fCorp_{t-1}$	-	-		+								
$sComms_{t-1}$												
$fComms_{t-1}$												
$sCredit_{t-1}$		+										+
$fCredit_{t-1}$					+							

Table A2: This table summarizes the coefficient estimates from all four panel specifications in the paper: *return*, $return^{12}$, *sigma*, and *drawdown*. A "+" ("-") indicates a positive (negative) coefficient estimate significant at the 10% lever or better; ∅ indicates the variable is not present in a given specification. The table shows results for the developed market regressions over the entire sample, as well as the two sub-periods.

Developed markets: Forecasting panel for returns

	Base	Sent	SentEnt	Base	Sent	SentEnt	Base	Sent	SentEnt
$sigma_{t-1}$	-0.035	-0.022	-0.020	0.103**	0.103**	0.101**	-0.115	-0.090	-0.087
$sigma_{t-2}$	0.025	0.040	0.041	-0.086	-0.070	-0.070	0.114	0.140*	0.140*
$return_{t-1}$	0.113*	0.078	0.078	0.128**	0.096*	0.097*	0.066	0.025	0.024
$return_{t-2}$	-0.016	-0.034	-0.034	-0.049	-0.060	-0.059	0.005	-0.020	-0.023
$value_{t-1}$	1.591**	1.992***	2.011***	1.110	1.365	1.321	1.985**	2.529***	2.574***
gdp_{t-1}	-0.053	-0.125	-0.133	0.040	-0.042	-0.041	-0.140	-0.174	-0.179
$gdpdeflator_{t-1}$	0.034	-0.035	-0.038	0.123	0.060	0.060	-0.126	-0.145	-0.139
cp_{t-1}	-0.024*	-0.023**	-0.022**	0.001	0.001	0.002	-0.039**	-0.051***	-0.048***
dcp_{t-1}	0.004	0.001	0.000	-0.021**	-0.019*	-0.019*	0.040**	0.041**	0.037*
$rate_{t-1}$	-0.331**	-0.592***	-0.610***	-1.015*	-1.217**	-1.184**	-0.383*	-0.659***	-0.689***
$dexch_{t-1}$	0.001	-0.006	-0.006	-0.031	-0.042	-0.040	0.011	0.005	0.006
pre	0.169	-0.010	-0.025	-0.055	0.021	0.016	0.439	0.069	0.035
$post$	0.185	0.089	0.082	-0.124	-0.122	-0.107	0.440	0.314	0.321
$entropy_{t-1}$		3.060*	3.653**		-3.279	-3.154		6.457**	7.165***
$artcount_{t-1}$		0.448*	0.420*		0.506	0.492		0.544	0.524
$sMkt_{t-1}$		0.844**	0.897**		-0.093	-0.236		1.666**	1.831***
$fMkt_{t-1}$		0.311	0.307		-0.044	-0.088		0.379	0.411
$sGovt_{t-1}$		-0.978**	-0.956**		0.016	-0.009		-1.591***	-1.449***
$fGovt_{t-1}$		-0.754	-0.774		0.656	0.640		-1.952***	-1.911***
$sCorp_{t-1}$		0.271	0.367		0.460	0.418		0.864*	0.904*
$fCorp_{t-1}$		-0.477*	-0.440*		0.393	0.359		-0.472	-0.411
$sComms_{t-1}$		0.237	0.192		0.323	0.346		0.066	-0.049
$fComms_{t-1}$		0.210	0.204		0.156	0.196		0.288	0.205
$sCredit_{t-1}$		0.080	0.016		0.004	0.122		-0.702	-0.810
$fCredit_{t-1}$		0.050	0.051		0.573**	0.573**		-0.358	-0.339
$R2$	0.028	0.0462	0.0472	0.034	0.0482	0.0477	0.0588	0.0927	0.0946
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 A3: Panel regressions for developed market 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. All text measures except *entropy* are normalized to unit variance. 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, and "***","**", and "*" indicate significance at the 1%, 5% and 10% levels respectively.

Emerging markets: Forecasting panel for returns

	Base	Sent	SentEnt	Base	Sent	SentEnt	Base	Sent	SentEnt
$sigma_{t-1}$	-0.034	-0.025	-0.026	0.017	0.058	0.057	-0.074	-0.074	-0.076
$sigma_{t-2}$	0.039	0.046	0.045	-0.020	0.001	0.001	0.111**	0.111**	0.110**
$return_{t-1}$	0.114***	0.096**	0.099**	0.073*	0.043	0.046	0.123*	0.101	0.104
$return_{t-2}$	-0.006	-0.008	-0.007	-0.051	-0.060	-0.058	0.037	0.031	0.033
$value_{t-1}$	0.227	0.242	0.261	0.471	1.145***	1.167***	0.165	-0.106	-0.116
gdp_{t-1}	-0.091	-0.106*	-0.107*	-0.037	-0.180*	-0.178*	-0.160*	-0.132*	-0.133*
$gdpdeflator_{t-1}$	-0.036	-0.048	-0.048	-0.012	-0.117**	-0.115**	-0.117**	-0.112**	-0.115**
cp_{t-1}	-0.035**	-0.024*	-0.024*	-0.022	-0.011	-0.010	-0.021	-0.028	-0.028
dcp_{t-1}	-0.009	-0.004	-0.004	0.024	0.005	0.009	-0.029	-0.022	-0.027
$rate_{t-1}$	-0.008	-0.036	-0.034	-0.035	0.086	0.086	-0.221	-0.312**	-0.305*
$dexch_{t-1}$	0.068	0.071	0.069	0.110	0.124	0.122	0.044	0.030	0.027
pre	-0.063	-0.062	-0.066	-0.493	-0.866*	-0.891*	0.077	0.097	0.110
$post$	-0.557	-0.542	-0.562	-0.732	-0.978*	-1.012*	-0.474	-0.359	-0.393
$entropy_{t-1}$		2.691	2.800		-10.044**	-9.559**		7.508**	7.373**
$artcount_{t-1}$		-0.229	-0.232		-1.197*	-1.208*		0.007	0.039
$sMkt_{t-1}$		0.396	0.190		0.141	-0.217		0.186	0.039
$fMkt_{t-1}$		-0.469	-0.522		-1.562***	-1.661***		0.511	0.429
$sGovt_{t-1}$		-0.188	-0.162		-0.314	-0.365		0.005	0.001
$fGovt_{t-1}$		-0.918*	-0.872*		-1.328*	-1.304*		0.129	0.075
$sCorp_{t-1}$		-0.809*	-0.675*		0.089	0.128		-1.014**	-0.845*
$fCorp_{t-1}$		-1.031***	-0.977***		-0.766	-0.707		-0.452	-0.418
$sComms_{t-1}$		0.840***	0.813***		0.397	0.535		1.279***	1.156***
$fComms_{t-1}$		0.287	0.267		-0.079	0.033		0.833**	0.716**
$sMacro_{t-1}$		0.132	0.197		0.524	0.710		-0.704	-0.610
$fMacro_{t-1}$		0.409*	0.448*		0.247	0.343		0.565*	0.623**
$R2$	0.0144	0.0242	0.0236	-0.00486	0.0179	0.0188	0.0404	0.062	0.0597
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	by time	by time	by time	by time	by time	by time	by time	by time	by time

Table A4: Panel regressions for emerging market 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. All text measures except *entropy* are normalized to unit variance. 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, and "***","**", and "*" indicate significance at the 1%, 5% and 10% levels respectively.