Language, News and Volatility

Byström, Hans

2014

Document Version:
Other version

Link to publication

Citation for published version (APA):
Byström, H. (2014). *Language, News and Volatility*. (Working Paper / Department of Economics, School of Economics and Management, Lund University; No. 41). Department of Economics, Lund University.

Total number of authors:
1

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Working Paper 2014:41

Department of Economics
School of Economics and Management

Language, News and Volatility

Hans Byström

November 2014

LUND UNIVERSITY

Language, News and Volatility

HANS BYSTRÖM*

November 27, 2014

I use Google News ™ to study the relation between news volumes and stock market volatilities. More than nine million stock market-related news stories in English and (Mandarin) Chinese are collected and the dynamics of the news volume and the stock market volatility is compared in both the Anglophone world and the Sinophone world. I find that the stock market volatility and the number of publicly available global news stories are strongly linked to each other in both languages. Contemporaneous correlations between news and volatility are positive and highly significant, and regressions tell us that the directional link between news and volatility rather is *from* news *to* volatility than vice versa. In out-of-sample evaluations of volatility forecasts I find news volumes to improve forecasts, regardless of language. The relationship between news and volatility is weakest in mainland China and a possible reason for this is that Chinese retail investors do not read (traditional) news, neither in Chinese nor in English. The results suggest that news could be used in volatility-related financial applications such as GARCH-models or VIX-like fear indexes.

Keywords: news aggregator; news; language; volatility, stock market; Chinese; Mandarin; GARCH; VIX
JEL classification codes: G10; D80; C82

---

* Hans Byström is from the Department of Economics, Lund University, Box 7082, 22007 Lund, Sweden (hans.bystrom@nek.lu.se). Financial assistance from *The Marianne and Marcus Wallenberg Foundation* and *Handelsbankens Forskningsstiftelser* is gratefully acknowledged. The author thanks Google for making the news volume publicly available through Google News and Lu Liu, Mingfa Ding and Caren Yinxia Guo Nielsen for Chinese language assistance. All mistakes in this article are the author's.

In this paper I try to shed some new light on the old questions of whether the amount of publicly reported stock market-related news is linked to the volatility in the stock market and, if so, if it is news that causes volatility or vice versa. I use Google News ™, the news aggregator, to capture the actual month-to-month dynamics of the global news volume. To make the study more inclusive, I look at news written in the two most important global languages, i.e. English and Chinese.

Large movements in the stock market are often (ex post) explained as the market's reaction to sudden important news arrivals. At other times, however, markets move seemingly without any evidence of important news arriving. A comparison of two recent so-called "flash crashes" can be used to exemplify this. While the "Twitter Crash" of April 2013, when S&P 500 lost $120 billion in market value in seconds, was caused by fake tweets (i.e. news) about explosions at the White House, the "Flash Crash" of May 2010, when $1 trillion in market value temporarily was lost, is normally not considered to have been caused by the arrival of news. In other words, it is not obvious that price movements always are reactions to new information (news) arriving in the market.

The main contribution of this paper, compared to the typical study linking news and volatility, is its unique proxy for the total amount of (global) stock market related news in circulation. By using an automated web-based news aggregator, in my case Google News, I am able not only to collect a significant share of all globally available market-driving public information but, through the continuous data collection process, I am also able to capture the actual month-to-month dynamics of this news dissemination. That is, instead of merely looking at specific news events, I look at the dynamics of the overall flow of public information. Moreover, by focusing on the bulk of the relevant information flow (each month, I collect all available news stories where a generic phrase such as "stock market" is mentioned) I efficiently avoid any undue emphasize on news stories that, ex post, turn out to have had a significant effect on the volatility in our particular markets.

In total, I collect more than nine million stock market-related news stories published by major newspapers and other news sources worldwide over an eight-year long period. To put this amount of news into perspective it can be compared to the 120,000 Reuter's News Service news releases collected by Berry and Howe (1994), the 752,647 Wall Street Journal and Broadtape story headlines collected by Mitchell and Mulherin (1994) and the 129,737 Dow Jones and Reuters news announcements collected by Johnson and Marietta-Westberg (2004). Furthermore, to highlight the truly global nature of both news and stock markets I have collected news in both of the two major global languages, i.e. English and Chinese; eight million of the news stories are in English and one million are in Chinese (Mandarin). I believe this to be the first time anyone looks at news written in Chinese, i.e., arguably, the second-most important language in the world, in connection to market volatility. The English-Chinese language-pair is also particularly interesting in the light of the two languages' significant semantic and linguistic differences and due to the fact that few stock market participants actually read news both in English and in Chinese.

Stock return volatility varies widely across time. It also tends to be persistent and to exhibit so-called volatility clustering, where periods of high volatility are followed by high volatility and vice versa. Although it is something of a stylized fact that new information reaching the market, i.e. news, is the main contributor of this volatility, and although several studies have looked into the relationship between market volatility and news dissemination, the empirical evidence is not as strong as one would expect. In fact, even when a link between news and market movements is found, the strength of the link is often questioned. One of the first studies on news and volatility was French and Roll (1986) who compares the volatility in the US stock market during exchange trading hours and non-trading hours and concludes that the difference in the flow of information, particularly private information, explains the difference in volatility. In other words, they conclude that it is the variability in the flow of private information that explains most of the variability in volatility. Mitchell and Mulherin (1994), in turn, looks at public news and finds a positive and statistically significant, albeit weak, relationship between the variability (absolute return) in the US stock market and the number of public news announcements, measured as the daily number of story headlines reported by Dow Jones & Company (Wall Street Journal and Broadtape). Berry and Howe (1994), on the other hand, does not find that public information is statistically related to stock volatility in the US intraday stock market. Berry and Howe (1994) measures public information flow as the number of news releases by Reuter's News Services. The interest rate and foreign exchange markets also exhibit time-varying volatility and Ederington and Lee (1993) shows that the impact of scheduled macroeconomic news announcements has an immediate effect on prices and a long lasting effect on price volatility. In the stock market, an even longer lasting period of elevated volatility after announcements is found by Patell and Wolfson (1984). In a more recent study, Johnson and Marietta-Westberg (2004) finds that increases in idiosyncratic stock return volatility are positively related to increases in the amount of firm-specific public news. And in Byström (2009, 2011) I use the same Google News methodology of collecting news as in this study but across a much shorter sample and limiting the analysis to simple cross-correlations.[1] The results in Byström (2009,

[1] Byström (2009, 2011) looks at the link between news and volatility using Google News but otherwise those studies differ significantly from this study. In this paper I include news in Chinese (Mandarin) as well as in English, I look not only at contemporaneous correlations but focus instead primarily on regressions between current volatility and lagged news volumes, I run lead-lag regressions to tell in which direction information flows, I look at the volatility forecasting performance of news, I look at changes in addition to levels, I look at a time-period that is almost three times as long, I use monthly non-overlapping news volume observations rather than daily overlapping ones, I look at twelve major stock indexes in both the English-language dominated world and in the Chinese-language dominated part of the world and in a robustness section I investigate whether extreme news observations, the crisis period, missing observations or the exact wording of the news search string is driving the results.

2011), even if merely tentative, indicate a positive link between stock market volatility and news volumes.

A related strand of literature focuses on investor attention, rather than on news digestion. Here, another Google product, Google Trends ™, has been employed recently. In this literature, Google search frequencies (Search Volume Index (SVI)) are used as a proxy for investor attention. Dimpfl and Jank (2012) proxy investor attention with Google search frequencies and finds a strong correlation, using daily data, between the search query volume and US stock market volatility. They also show that search queries improve volatility forecasts. Similarly, Vozlyublennaia (2014) collects Google search frequencies on a weekly basis using Google Trends for a range of different financial markets, including stock markets, but finds a rather weak relationship between investor attention (Google search frequency) and volatility. Finally, Da, Engelberg and Gao (2011), although not explicitly focusing on volatility, finds a positive relationship between abnormal absolute returns in the US stock market and the Google Search Volume Index. Like the current paper, all these studies employ modern web-based tools, but rather than collecting news volumes they collect search frequencies.

Most studies on information flows and stock market reactions have wrestled with various data-related issues. Some studies have not been able to differentiate between good, bad and neutral news and many have not been able to measure the importance of a particular piece of news. Other studies isolate specific events and therefore lack a continuous measure of the number of available news stories. This is typically an issue when macroeconomic news is studied, and without a continuous sampling of the amount of news in circulation, the dynamics of news volumes cannot be studied. I believe that several of these issues are avoided in our research setup. First, since Google News allows the user to specify exactly which word strings to crawl, I am, by construction, able to group news according to sentiment (neutral news or bad news). The word strings in this particular paper are *stock market* and *stock market crash* in the US edition and *股市* (*stock market*) and *股市暴跌 / 股市崩盘* (approximately translated into English as *stock market collapse* and *stock market crash*) in the China edition. Second, by focusing on the *amount* of news rather than the mere *existence or not* of news I am automatically able to tell whether the underlying actions or events that shape the news are important or not. Third, through the continuous data collection process, I am able to capture the time-series dynamics of the news in circulation. Finally, the comprehensive nature of the Google News generated news database and the sheer number of news stories strengthens the results.

Since the news that I am collecting using Google News is global in nature, in the sense that it is written in two languages that together are read by a majority of the world's market participants, I have chosen to look at globally important stock markets and the volatility in the major stock indexes in these markets. Due to the dual-language focus of the research I have chosen half of the stock indexes from the English-speaking world and half from the Chinese-speaking world. The former are *MSCI World, S&P 500, DJIA, Nasdaq, Russell 2000* and *FTSE 100* and the latter are *Shanghai A, Shanghai B, Shenzhen A, Shenzhen B, Hang Seng China Enterprises* and *Hang Seng China-Affiliated Corporations*. In total, I look at the impact of Google News-generated global news volumes on twelve major stock indexes.

To my knowledge, this is the first time a news aggregator is employed in the forecasting of financial market volatility and, furthermore, the first time relatively strong evidence is found that news volumes can actually predict stock market volatility. In addition to robust and significant positive contemporaneous correlations between the amount of news in circulation and the volatility in various major stock markets, lead-lag regressions tell us that the directional link between news and volatility rather is *from* news *to* volatility than vice versa. I also find evidence of news volumes predicting (one-month ahead) volatility. The latter finding is supported by significant and economically relevant news volume regression parameters and economically meaningful out-of-sample forecasting error reductions. The average impact of a one-standard deviation change in news volume on next month's stock index volatility is 11 basis points (0.11%). Albeit not large, the impact is economically meaningful when compared to the mean of the twelve stock indexes' unconditional volatility across the sample period (covering the very volatile credit- and euro-area sovereign crises) which is 123 basis points. Moreover, when I look at changes, rather than levels, the economic significance is even more significant (9 basis points versus a mean of 44 basis points). The out-of-sample mean absolute forecasting error (MAE), in turn, is on average reduced by 15% when lagged news volumes are added to past volatilities when predicting (one-month ahead) volatility. It should be stressed that no fine-tuning of the forecasts are made, the estimation window is for example set rather arbitrarily, and for some market/language/search string combinations the economic significance of including news volumes in the volatility prediction is much larger. In other words, there is scope for a more substantial forecasting improvement when allowing for systematic data mining. I conclude the empirical study with a robustness section where I find the results to be robust to the removal of extreme news volume observations, credit crisis observations and missing observations as well as to slight changes to the news volume collection process.

The rest of the paper is organized as follows. Section I describes the data and how the news collection has been done. Section II presents empirical evidence on the relation between news volumes and stock market volatilities and discusses some possible applications of the results. Section III summarizes the paper.

I. News Volume Collection and Data Description

News is written in hundreds of different languages, some read and understood by global audiences of millions of people, others read solely by locals or cognoscenti. At any point in time there are thousands of news pieces available to market participants and any of this news may affect the market in one way or another. In this paper, I focus on the volatility in the major stock markets and on how this volatility is related to the amount of news available. I therefore focus primarily on news that is likely to have an impact on stock market volatility. I have also chosen to focus on the two, arguably, most important global languages, English and Chinese. English is the lingua franca of today with up to a billion native and non-native speakers. Chinese, on the other hand, is the most commonly spoken native language in the world with around one billion native speakers. Compared to many other financially important language-pairs, such as French and English or Spanish and Portuguese, there is also very little overlap in the readership of English and Chinese news. In fact, speakers of one language often do not understand a single word in the other language. For us, this is of importance since this makes the English and Chinese Google News-generated news volumes more distinctive and more likely to have their own unique relation to market volatility.

I collect what I deem to be stock market-related news volumes using an English language edition (the US edition) as well as a Chinese language edition (the China edition) of the news aggregator Google News. This news aggregator makes it possible to collect a significant amount of the many thousands of available news pieces around the world selected and grouped by topic. By collecting, on a monthly basis, *the number of news stories* presented by Google News I get an estimate of the dynamics of the news volume, i.e. the dynamics of the overall flow of public information, rather than just snapshots of the volume around certain chosen events. The news volume is collected using two different editions of Google News as well as two different languages, English and Chinese (Mandarin).[2] The total number of separate news stories collected in this way over the time-period 2006 to 2014 is more than nine million, of which eight million are written in English and one million are written in Chinese.[3]

Google News is an automated news aggregator (computer-generated news service) that uses computer algorithms to collect, present and sort web news into categories. Google News aggregates news from more than 50,000 news sources worldwide. It then groups similar stories together, and displays them according to each reader's interests (Google, 2014). News stories are collected from news pages on the web with the geographical location of the news sources dependent on the edition. Google News includes news that appeared on any of the selected web pages during the past 30 days and since no human editors are used the political/ideological bias is minimized (Google, 2014). The actual news sources are probably not known outside the Googleplex (the Google HQ) but unverified rumors on the web claim that the largest contributors to the Google News flow are New York Post, Washington Post, Houston Chronicle and Bloomberg in the US, and The Guardian, BBC News and The Times in the UK.

In order to make the exercise feasible I focus solely on news pieces deemed relevant to stock market participants. I therefore limit my Google News search to the search strings *stock market* and *stock market crash* in the US edition and 股市 (*stock market*) and 股市暴跌 / 股市崩盘 (approximately translated into English as *stock market collapse* and *stock market crash*) in the

[2] A preliminary analysis indicates that the search results are very similar when the news data is collected from the Hong Kong edition of Google News instead of the China edition.

[3] In mid-May 2012, the number of news stories reported by Google News increases dramatically (most likely by the inclusion of additional news sources) and I have therefore chosen to adjust the numbers from June 2012 onwards. The numbers are normalized so that the first observation after the change is identical to the last observation before the change. Reported numbers are always normalized ones and they are therefore not directly comparable to current Google News volumes. The total number of news stories reported in this paper is therefore also under-reported. Some of the correlation analysis in this paper has been redone with data up until the change with roughly unchanged results.

Chinese edition.[4] Despite this limitation in news coverage I still manage to collect more than nine million separate news stories containing the word stock market over the eight-year period 2006 to 2014.[5]

The data, i.e. the amount of news publicly available worldwide over the last 30 days, is collected on a monthly basis from September 11, 2006 to September 1, 2014 for the English news and from November 1, 2010 to September 1, 2014 for the Chinese news. More exactly, the data is collected manually by the author every fourth Monday at approximately 9:00 a.m. Central European Time at the same location (the office of the author) and without being signed in to any Google Account. On a handful of occasions when the author was travelling on the scheduled Monday, the data collection was done on the following Tuesday or Wednesday or at another location.[6,7] On any particular day, the Google News aggregator collects data from the last 30-day period. This means that my definition of a month (as a 4-week or 28-day period) differs slightly from Google's definition of a month (a 30-day period). The difference is small, 2 days, and since the additional 2 days of news crawling always constitutes a weekend (when markets are closed and there is less market-related news available) four weeks ago I believe that it biases the results minimally. In any case, this discrepancy should probably bias the results *against* us finding a link between news volume and volatility.[8]

[4] The reason for choosing two different "pessimistic" search strings in Chinese is to control for any potential language-difference in the interpretation of the word "crash".

[5] Here, I assume that no news story is reprinted again at a later stage and that no two news stories are exactly identical.

[6] The "December" search result is sometimes missing due to Christmas and New Year's Eve. There are also some missing observations on other days randomly scattered throughout the years and these as well as the missing December observations are all replaced by the last available data point. In order to see whether these missing observations affect the results, a dummy is added to the regressions, and correlations are recalculated with the dates of the missing observations removed, with almost unchanged results.

[7] In order to make sure that the news volume is not platform-dependent the number of news stories was occasionally collected at several locations the same day (at randomly chosen days throughout the sample period) with very minor differences.

[8] According to the home page of Google News the news aggregator includes news articles that have been crawled within the last 30 days. However, a careful study of the search results sometimes reveals a few news stories that are

The monthly number of news stories varies not only across time but among the search strings used in the Google News search. While the more general search string *stock market* (in English) returns on average 89,000 separate news stories per month, the more narrow search string *stock market crash* (again, in English) returns on average 1,900 news stories per month. The pattern is similar for the Chinese-language news volumes where 股市 (*stock market*) on average returns 19,000 news stories per month and 股市崩盘 (*stock market crash*) on average returns 890 news stories per month (across a shorter time-period). The time-series variation of the *stock market crash* news volumes, in English and in Chinese, respectively, is graphically presented in Figures 1 and 2. Not surprisingly, the amount of stock market-related news in circulation was at its highest around the period of the Lehmann Brothers collapse. The number of news stories containing the words *stock market* was higher than 200,000 (at roughly 2 ½ times the monthly average) both in September and October in 2008. Two years earlier, in September and October 2006, the number of news stories was around 70,000 per month. The Chinese news was not collected at the time of the Lehman Brothers collapse and the peak in the number of news stories containing the word 股市 (*stock market*) was reached in the spring of 2012 when more than 30,000 news stories were released each month from March to June.

The stock market data, in turn, is collected for the same time-period as the news volumes, i.e. September 11, 2006 to September 1, 2014. The data is downloaded from Datastream and all the stock indexes are denominated in their home-currency. I include twelve different stock indexes in my analysis. Since my aim is to study the effect of language on the news-volatility link, I have chosen half of the indexes from the mainly English-speaking sphere, i.e. the US, the UK and the global community, and half from the mainly Chinese-speaking sphere, which I define as China

(a few days or weeks) older than 30 days. However, on these occasions the number of news stories that are older than 30 days have been found to be few compared to the total number of stories and therefore less likely to significantly including Hong Kong. As the stock index representing the global Anglophone community I have chosen the MSCI World stock index which includes securities from 23 developed stock markets around the world. From the US I have included the S&P 500-, DJIA-, Nasdaq- and Russell 2000 indexes. The main motivation behind including the last two indexes is their focus on small-cap stocks. It is possible that the news-volatility link is different for small stocks where the balance between small retail-investors and large institutional investors is different. The UK, finally, is represented by the FTSE 100 index which covers the 100 largest companies on the London Stock Exchange.

While most readers are familiar with the Anglophone world and its stock markets, the Chinese-speaking sphere and the stock markets dominated by Chinese-speakers probably needs some introduction. Since my focus is Mandarin, the main language spoken in China, I have only chosen stock indexes that contain Chinese stocks. The Chinese stock market is highly segmented with different markets aimed at different investors:

- *A-shares*: A-shares are RMB-denominated shares issued by domestic companies registered in mainland China and listed on the Shanghai Stock Exchange or the Shenzhen Stock Exchange. A-shares can only be purchased by domestic Chinese investors or holders of Qualified Foreign Institutional Investor (QFII) licenses.
- *B-shares*: B-shares are dollar-denominated shares issued by domestic companies registered in mainland China and listed on the Shanghai Stock Exchange (US$) or the Shenzhen Stock Exchange (HK$). B-shares can only be purchased by foreign investors or by domestic investors with foreign currency holdings, and capital controls restrict Chinese residents' ability to purchase B shares.

bias the results.

- *H-shares*: H-shares are HK$-denominated shares issued by companies incorporated in mainland China but listed on the Hong Kong Stock Exchange. H-shares cannot be purchased by domestic Chinese investors.
- *Red Chip-shares*: Red Chip-shares are HK$-denominated shares issued by Chinese companies incorporated in Hong Kong and listed on the Hong Kong Stock Exchange. Red Chip-shares cannot be purchased by domestic Chinese investors.

While the A-share market is aimed *mainly at domestic investors*, the B-share market is aimed *predominantly at foreign investors* and the H-share and Red Chip-share markets are aimed *solely at foreign investors*. This gradual increase in segmentation facilitates the study of the effect of market participant and language on the news volume—volatility link.

I look at Chinese shares traded on three different exchanges; A- and B-shares traded on the Shanghai Stock Exchange, A- and B-shares traded on the Shenzhen Stock Exchange, and H- and Red Chip-shares traded on the Hong Kong Stock Exchange. The actual stock indexes are the *Shanghai SE A* Index, the *Shanghai SE B* Index, the *Shenzhen SE A* Index, the *Shenzhen SE B* Index, the *Hang Seng China Enterprises* Index and the *Hang Seng China-Affiliated Corporations* Index.

Each month, i.e. every fourth Monday, the past month's stock market volatility is calculated as the standard deviation of the daily stock index returns over the last four weeks so that the time-period for the volatility estimate matches the time-period for the news volume collection (except for the 2 days discussed above). The possibility of matching the news collection period (one month) with the volatility computation period (one month) is one huge advantage of using a monthly frequency in the analysis. Another advantage is that the exact time stamp of the news release is not required (a problem faced by several previous studies, for instance Dimpfl and Jank

(2012)) and that the time zone differences around the world, most notably between China and the US, have a minimal effect on the results.

II. The Relation between News Volumes and Stock Market Volatility

In efficient financial markets, price movements are the results of market participants reacting on market-related news. As a result, the more news that reaches the market over a certain time-period the higher the price volatility in the market is likely to be. In this study of stock markets worldwide I therefore expect the stock return volatility to be positively linked to the amount of stock market-related news worldwide. Indeed, an initial visual inspection of Figures 1 and 2, where monthly stock market volatilities and news volumes are presented on a monthly basis for English and Chinese news stories (smoothed using a quarter-of-a-year long window and normalized to start at one), motivates us to investigate this link further.

A. Correlation Analysis of News Volumes and Stock Market Volatility

To start with, I present simple contemporaneous correlations between news aggregator generated news volumes and stock market volatilities. I study twelve different stock market indexes and collect news in English as well as in Chinese. All variables are sampled on a monthly basis. In addition to *levels* I also look at *changes* in news volume and volatility. For the changes, I follow Mitchell and Mulherin (1994), who take differences from a multi-day moving average, by taking differences from a 12-month moving average of the past news volume and volatility, respectively. Just like Mitchell and Mulherin (1994), I take multi-period differences to avoid the loss of information around the occasional clustering of high news volumes and high volatility levels and to reduce the influence of possible month-of-the-year effects.

The correlation results are presented in Table I, which is divided into two parts, one for levels and one for changes. With very few exceptions, the correlation coefficients (based on the entire sample) among news volume- and volatility *levels* are large, positive and statistically significant. Most correlations lie in the 0.3 – 0.8 range and the only non-significant correlations are those involving the Chinese-language 股市崩盘 (*stock market crash*) news stories. These correlations are typically the lowest, regardless of stock market, and in the mainland China stock markets in Shanghai and Shenzhen the correlations are occasionally even slightly negative. Even if the negative correlations are small and not statistically significant, it is interesting that the link to Chinese news is found to be weakest in the Chinese stock market regardless of the actual Google News search string.

When I turn to *changes* rather than levels, the correlation coefficients are still largely positive and statistically significant. The statistical significance, i.e. the size, of the correlations is generally somewhat weaker for changes, however, and the correlations are again lowest when Chinese news or mainland China stock markets are involved. For changes, though, the link between mainland China stock markets and news is weak across the board, i.e. regardless of the language of the news. One possible explanation for the rather weak relationship between news and stock volatility in mainland China (i.e. excluding stocks traded in Hong Kong) is that these markets are dominated by investors that do not read, or at least do not trade on, (traditional) news. This hypothesis is strengthened by the fact that the correlation is particularly weak in the A-share market which is aimed *mainly* at domestic (retail) investors. In fact, the link between the volatility in the Chinese stock market and the amount of stock market-related news reaching the market participants is stronger in markets where there are fewer mainland Chinese retail investors. The link is strongest in the H-share and Red Chip-share markets in Hong Kong where Chinese retail investors are banned from trading. The link is somewhat weaker in the two B-share markets where few Chinese retail investors invest and it is weakest in the retail-dominated A-share markets.

To some extent, these findings on the segmentation and information processing in the Chinese stock market are in line with the limited amount of research that exists on the topic. In a previous study using Google News, Byström (2011) finds a strong link between English-language news volumes worldwide and the global stock market volatility, and a weaker link between the Chinese stock market and the same set of worldwide (English) news. Poon and Fung (2000) looks at how information flows among the various China- and Hong Kong-based markets and finds return and volatility spillover effects among securities listed on the different markets. The Red Chip market is found to process information faster than the other markets. The segmentation and information flows among Chinese stock markets has also been studied by Yang (2003) who finds the foreign investors who dominate the B-share market to be better informed than the domestic investors in the A-share market. And in a study relating news and stock market volatility in the segmented Chinese stock market, Su and Fleisher (1999) finds the volatility in the A-share market in the late 90s to be significantly higher than that in the B-share market and they try to explain this fact using arguments related to news flows and different investor bases.

Looking at the overall picture, however, the link between news and volatility is strong. Those months when a lot of news is released are typically also those months when the stock market volatility is high, and months with significant relative increases/decreases in news volumes (relative to the last twelve months' volume) are also months with a significant relative increase/decrease in market activity (i.e. volatility). Additional evidence of the strong association between news and volatility is given in Figures 1 and 2 where the relative size of the monthly movements in news volume is quite similar to the relative size of the movements in the MSCI

World volatility. That is, news volumes and stock volatilities are not only moving in the same direction at the same time but the actual size of the changes is similar as well.

B. *Regression Analysis of News Volumes and Stock Market Volatility*

The next step is to run univariate lead-lag OLS regressions between news volume and stock market volatility. The regressions allow us to assess the predictive ability of news volumes and to evaluate the economic significance of the inter-temporal news-volatility link in the stock market. I run two sets of regressions. In the first set, the dependent variable is the (monthly) volatility and the explanatory variables are (one-month) lagged news volume and (one-month) lagged volatility. In the second set, we reverse the regression and the dependent variable is the (monthly) news volume and the explanatory variables are (one-month) lagged volatility and (one-month) lagged news volume. To account for the possible influence of missing news volume values I include a dummy for the missing values in all reported regressions. The regressions include only one lag of volatility and news volume and the reason is that I look at monthly data; any empirical relationship found between stock market volatility this month and news released several months ago is likely to be spurious. Indeed, Tetlock (2007), who looks at daily news releases and includes lags up to five days, finds a reversal in the initial reaction to news already four or five days after the news release. And Vozlyublennaia (2014), who collects Google search queries on a weekly basis as a measure of investor attention, argues that her weak link between attention and volatility could be caused by her using weekly, rather than daily, data since the effect of attention on volatility could disappear already in days, rather than in weeks. In other words, if the same applies to news and to its effect on volatility, there is no reason to expect a link between stock market volatility this month and news released several months ago.

Tables II and III summarize the regression results. Since my focus is primarily on the news volume/volatility coefficients, the intercept in the regression equation and the missing value dummy coefficient (which is rarely statistically significant) are left out. Table II presents the results for the English-language news and Table III presents the results for news in Chinese. In the upper regressions labelled *News*, the news volume is the dependent variable and in the lower regressions labelled *Volatility*, the stock market volatility is the dependent variable. In both regressions, the first explanatory variable, β_1, is always the one-month lagged value of the other variable (i.e. volatility in the upper regression and news volume in the lower regression) and the second explanatory variable, β_2, is always the one-month lagged value of the dependent variable. In the analysis that follows I therefore focus on β_1. Results for changes follow after those for levels and *, ** and *** represent significance at the 10%, 5% and 1% level, respectively.

Overall, β_1 is often, but not always, positive. It is much more often positive in the *Volatility* regression (in all but five cases), i.e. lagged news volume is more likely to cause volatility than vice versa. For the *Volatility* regression, β_1 is statistically significant for roughly half the market/language/search string combinations in tables II and III, and there are 42 cases when news volume (Granger) causes volatility and volatility does not cause news volume, but only 7 cases when volatility causes news volume and news volume does not cause volatility. Interestingly, these latter seven instances are all found in mainland China when the news is in English. This is an indication of English-language news not reaching mainland Chinese stock market participants, or at least this news does not help in predicting volatility. Meanwhile, Chinese news (Granger) causes volatility in mainland China just as it does in most other stock markets. As for β_2, finally, it is well-known that volatility is persistent and that lagged volatility predicts future volatility and this is evident in my regressions in the (almost) unanimously significant β_2 coefficient. Interestingly, news volume seems to be as persistent as volatility.

As for the economic significance of the results, and here I focus solely on the impact of news on volatility, the size of the regression parameters tells us that, for levels, the average impact of a one-standard deviation change in news volume on next month's stock index volatility is 11 basis points (0.11%). Albeit not large, the impact is economically meaningful when compared to the mean of the twelve stock indexes' unconditional volatility across the sample period (covering the very volatile credit- and the euro-area sovereign crises) which is 123 basis points. Of course, for some market/language/search string combinations the economic significance of the regression parameters is much larger. For the Russell 2000 index of small US firms, for instance, the average impact of a one-standard deviation change in the volume of news containing the Chinese word "股市暴跌, *i.e.* stock market collapse", on next month's stock index volatility is 37 basis points. For changes, the average impact of a one-standard deviation change in news volume, compared to the 12-month average, on next month's change in stock index volatility is 9 basis points.[9] Although smaller, the impact for changes is actually more economically meaningful than that for levels when compared to the smaller mean of the monthly change in the stock indexes' unconditional volatility which is 44 basis points.

The regression results in this subsection reveal that the amount of news *this* month predicts the volatility *next* month. This is possibly an indication of news dissipating from news source to market participant rather slowly and it mirrors results found by Ederington and Lee (1993) in the interest rate and foreign exchange markets. Like the stock market, these two markets also exhibit time-varying volatility and Ederington and Lee (1993) shows that scheduled macroeconomic news announcements have an immediate effect on prices but a long lasting effect on price volatility. In the stock market, an even longer-lasting period of elevated volatility after

[9] When no lagged volatility is included in the regression, unreported regressions show that the average impact of a one-standard deviation change in news volume on next month's stock index volatility is 28 basis points and the

announcements is found by Patell and Wolfson (1984). This gradual digestion of news by the market could lie behind the predictability of volatility using news.

A recurring finding in this paper is that pessimistic (negative) news has a somewhat stronger connection to stock return volatility than neutral news. The correlations presented in Table I are generally larger for negative news, regardless of language, and the same holds for the regression coefficient, β_1, in Tables II and III. I have no explanation for this other than the possibility that risk-averse market participants are more prone to react to negative than to neutral news.

To conclude, the main finding in this subsection is that the inter-temporal link between news and volatility is statistically and economically significant and that it seems to be stronger in the direction *from* news to volatility than vice versa. A secondary finding is that the pattern is different in mainland China where there is less of a positive inter-temporal link between news and volatility and where, particularly English-language, news does not help in predicting volatility.

C. Out-of-sample Forecasting of Stock Market Volatility using News Volumes

The regressions in the previous subsection show that news causes volatility and that news volumes possibly could be used to predict future stock market volatility. While any forecasting-assessment based on the regressions above are mere indications based on in-sample evidence, in this subsection I try to assess whether news has true out-of-sample predictive abilities. Since I did not find volatility to cause news volume I focus solely on volatility prediction. I do this through a rolling window estimation of the OLS regression parameters in the previous subsection where only past information is used to predict future (one-month ahead) volatility. The sample is divided into two (essentially) equally long periods, one estimation period and one evaluation

---

average impact for changes is 16 basis points.

period.[10] The volatility is then forecasted in two ways; (i) based solely on past volatility (*Without news*) or (ii) based on past volatility together with past news volume (*With news*).

The forecasting performance is assessed using two different loss functions; the mean absolute error (MAE) and the quasi-likelihood (QL) loss function. The two loss functions are defined as

$$MAE = \left|\sigma_{realized,t+1} - \sigma_{forecast,t}\right| \quad (1)$$

and

$$QL = \frac{\sigma_{realized,t+1}}{\sigma_{forecast,t}} - \log\left(\frac{\sigma_{realized,t+1}}{\sigma_{forecast,t}}\right) - 1 \quad (2)$$

where $\sigma_{forecast,t}$ is the forecast at month *t* of the volatility in month *t+1* using information available up to and including month *t*.

Forecasting results, for levels as well as for changes, are presented in Tables IV and V.[11] While Table IV present results for English-language news, Table V presents the same results for news in Chinese. The smallest (best) forecasting error/loss is typed in bold and more often than not, the inclusion of news (*With news*) in the prediction improves the volatility forecast. For English news the forecast is improved for each and every stock index when the amount of news is acknowledged. This strong result holds both for levels and for changes. Meanwhile, the inclusion of Chinese-language news is less useful for the forecaster. Table V shows that Chinese news improves the volatility forecast in 80 out of 144 cases. This disappointing performance is mainly driven by the mainland China stock markets in Shanghai and Shenzhen. If I exclude these

[10] Somewhat arbitrarily, I have chosen to let the evaluation period for the English news-based forecasting start at the same date as the Chinese news collection starts, i.e. November 1, 2010. This leaves 51 months in the evaluation period and 54 months in the estimation period. The Chinese news-based forecasting starts on October 29, 2012 which leaves 25 months in the evaluation period and 26 months in the estimation period.

[11] The results for a third unreported loss function, mean squared errors (MSE), are very similar to those for the reported loss functions.

markets the inclusion of Chinese news improves the forecasting in 60 out of 96 cases, i.e. in almost two thirds of the cases.

Overall, the results point at news aggregators such as Google News potentially being useful in practice when investors or risk managers attempt to predict stock return volatilities.[12] The mean absolute forecasting error (MAE), for example is on average reduced by 15% when acknowledging lagged news volumes in the prediction. Furthermore, it should be stressed that no fine-tuning of the forecasts are made, the estimation window is for example set rather arbitrarily, and for some market/language/search string combinations the economic significance of including news volumes in the volatility prediction is much larger. For the Hong Kong H-share market (*Hang Seng China Enterprises*), for example, the mean absolute forecasting error when English "stock market crash" news volumes are included in the volatility prediction is more than halved (51% reduction). In other words, there is scope for a more substantial forecasting improvement when allowing for systematic data mining.

D. Robustness Analysis

In this subsection I investigate whether the results might be driven by certain extreme news volume observations, or perhaps by the credit crisis. I also revisit the missing observations and the possibility that they affect the results. Finally, the exact wording of the news search strings is slightly adjusted to test the sensitivity of the news-volatility link to the choice of search phrase.

[12] This is, of course, conditional upon the owner of the news aggregator permitting commercial use. In the case of Google News, the Terms of Use can be found on http://news.google.com.

I start by excluding certain observations. To save space, I focus on correlations[13] and on the MSCI World index, and exclude three groups of data (one by one); extreme observations, crisis observations and missing observations. First, I recalculate the correlations when all months with extreme news volumes are removed. There are, indeed, some months where the news volumes are much higher than the average month, but a comparison of the full-sample correlations in Table I and the reduced-sample correlations in Table VI, where the 20% most news-heavy months are removed, shows that extremes are not driving the results. Although the correlations are lower, overall, when the extremes are removed, they are still mostly positive and highly significant, both for English and Chinese news.

As a second robustness test I remove the months around the crisis when the news volume is clearly elevated (July 2008 to June 2009).[14] This time, the general correlation level is lower than for the full sample and in one case the correlation even turns negative. Despite this, it is unlikely that the crisis is the sole driver of the positive relationship between news and volatility found throughout this study. For instance, in the forecasting horse race above we saw that the inclusion of news in the prediction produces smaller forecasting errors across the evaluation period 2010-2014, i.e. *after* the crisis. Furthermore, unreported regression results where a dummy for the crisis is added shows only small changes in the regression parameters.

Third, finally, I compute correlations with all the missing observations removed. As we can see in Table V, the correlations are all essentially unchanged and I conclude that it is highly unlikely that the missing observations affect the results in any meaningful way. This was further

[13] Unreported regression results where the crisis and the extreme observations are treated with dummies, similar to how the missing values were treated with a dummy, are substantially similar to the correlation results presented in this subsection.

[14] The time-period of the Chinese-language news sample makes it impossible to test this for the Chinese-language news volumes.

demonstrated with the mostly statistically insignificant "missing value" dummy variable in the regressions above.

In addition to the exclusion of observations, I also test the robustness of the results to slight changes to the search strings in the news volume collection. First, there is the possibility that the automatic Google News search for, let's say, the word string *stock market crash* also includes news stories that have nothing to do with the stock market, perhaps stories of a *crash* in *stock* car racing or similar stories. To test for this I use the alternative search string "stock market crash", i.e. I put the phrase in quotes. In this way I limit my news collection to news stories that are more likely to deal with the probability or occurrence of actual stock market crashes. As can be seen in Table V, the results are almost unchanged when this alternative search string is used. The same holds for the inclusion of the word *global* in front of the search string *stock market crash* and for a third alternative way of expressing pessimism about the stock market in Chinese, using the search string 股市崩溃**,** i.e. *stock market slump*.

E. Applications

Above, I have shown that there is a close link between news and volatility. I now turn to two examples that serve to illustrate the potential use of news in volatility-related financial applications. First, I suggest that one could add a NEWS-factor to GARCH models. The time-varying (conditional) variance of stock returns is often modelled using autoregressive models where the conditional heteroskedasticity of the returns is taken into account. These so-called Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models were introduced by Bollerslev (1986) and capture the autoregressive nature of volatilities. Now, if news is able to predict volatility, perhaps the lagged news volume (*NEWS*) could be included in the GARCH-specification:

$$\sigma^2{}_t = \omega + \alpha \cdot \varepsilon^2{}_{t-1} + \beta \cdot \sigma^2{}_{t-1} + \delta \cdot NEWS^2{}_{t-1} \tag{3}$$

Here, I have chosen the GARCH (1,1) specification and, in the light of the previous results in this paper, such a NEWS-GARCH (1,1) model could potentially fit observed conditional volatilities better than other ARCH or GARCH models.

Another example of how news could be used in volatility-related financial applications is the design of alternative VIX-indexes, i.e. alternative fear gauges, based on the amount of news in circulation. The VIX index is backed out from options on the stocks in the S&P 500 US stock index and is often interpreted as the global fear indicator, despite its US-origin. Now, in other markets where there are no traded options one cannot construct VIX-indexes. Also, volatility is just one proxy for fear and another one could be the amount of fear-related news in circulation. By collecting the number of news stories containing the words *stock market crash* or a similar word-combination, perhaps using a news aggregator like in this paper, we have an alternative way of gauging the level of fear in the market. We have seen that the correlation between such news volumes and the stock market volatility is high (but not equal to one) and this opens up the possibility of constructing tailor-made fear indexes for specific markets, fear types, or groups of investors (perhaps using different languages). For example, if you are worried about bankruptcies in state-controlled firms in China, rather than the S&P 500 stock price corrections gauged by VIX, you could perhaps search for the number of news stories including the word *bankruptcy* in Chinese. Or, if you want to assess the difference in attitude to this particular risk between Chinese and non-Chinese investors, you could perhaps compare the number of stories written in English and in Chinese.

III. Conclusion

In this paper, I use Google News to study the link between news and stock market volatility. I collect more than nine million stock market-related news stories in both English and (Mandarin) Chinese using the news aggregator Google News. In this way I collect a much larger share of all available market-driving public information than previous studies do. Through the continuous data collection process I am also able to capture the actual month-to-month dynamics of the news dissemination. That is, instead of merely looking at specific news events, I look at the dynamics of the overall flow of public information. The number of news stories, i.e. the news volumes, is then compared to the volatility in stock markets in both the Anglophone world and the Sinophone world.

I find that news and volatility are strongly linked to each other, regardless of whether the news is in English or Chinese. Negative news has a somewhat stronger connection to stock return volatility than neutral news. Overall, contemporaneous correlations are positive and statistically significant, and regressions show that the directional link between news and volatility is *from* news *to* volatility rather than vice versa. That is, I find that more news this month leads to higher stock market volatility next month, and in a simple out-of-sample forecasting exercise I find news volumes to improve volatility forecasts, regardless of language. One possible explanation for this could be a gradual, rather than immediate, digestion of news by the market.

The relationship between news and volatility is different for stocks traded in mainland China, the only market where there is no evidence of English-language news volume causing volatility. One possible reason for this is that Chinese (retail) investors do not read (traditional) news, neither in Chinese nor in English. Or, at least, they do not trade on this news.

The relation between public news and market volatility is found to be robust to the exclusion of extreme news volume observations as well as the credit crisis. It is also robust to the exact wording of the news search strings. Furthermore, the comprehensive nature of the Google News generated news database strengthens the results overall.

Beyond the study of the relation between news volumes and stock market volatility, the paper concludes by illustrating the potential use of news in volatility-related financial applications such as GARCH-models and VIX-like fear indexes.

REFERENCES

Berry, D. B. and Howe, K.M., 1994. Public information arrival. *Journal of Finance*, 49, 1331–1346.

Bollerslev, T., 1986. Generalized autoregresive conditional heteroskedasticity. *Journal of Econometrics*, 31, 307–327.

Byström, H., 2009. News aggregators, volatility and the stock market. *Economics Bulletin*, 29, 2679-2688.

Byström, H., 2011. Does the Chinese stock market react to news? *Journal of the Asia Pacific Economy*, 16 (3), 448-455.

Da, Z., Engelberg, J. and Gao, P., 2011. In search of attention. *Journal of Finance*, 66, 1461–1499.

Dimpfl, T. and Jank, S., 2012. Can internet search queries help to predict stock market volatility? Working Paper, University of Tübingen.

Ederington, L. H. and Lee, J.H., 1993. How markets process information: news releases and volatility. *Journal of Finance*, 48, 1161–1191.

French, K. R. and Roll, R., 1986. Stock return variances: The arrival of information and the reaction of traders. *Journal of Financial Economics*, 17, 5–26.

Google., 2014. Google News Help Center. https://support.google.com/news

Johnson, W. C. and Marietta-Westberg, J., 2004. The effect of news on volatility: a study of IPOs. Working Paper, Michigan State University.

Mitchell, M. L. and Mulherin, J.H., 1994. The effect of public information on the stock market. *Journal of Finance*, 49, 923–950.

Patell, J. and Wolfson, M., 1984. The intraday speed of adjustment of stock prices to earnings and dividend announcements. *Journal of Financial Economics*, 13, 223-252.

Poon, W. and Fung, H-G., 2000. Red chips or H shares: which China-backed securities process information the fastest? *Journal of Multinational Financial Management*, 10, 315–343.

Su, D. and Fleisher, B.M., 1999. Why does return volatility differ in Chinese stock markets? *Pacific-Basin Finance Journal*, 7 (5), 557–586.

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

Vozlyublennaia, N., 2014. Investor attention, index performance, and return predictability. *Journal of Banking & Finance*, 41, 17–35.

Yang, J., 2003. Market segmentation and information asymmetry in Chinese stock markets: a VAR analysis, *The Financial Review*, 38, 591-609.

Table I
Correlations between News Volume and Volatility

In this Table I present News Volume – Volatility correlations for the twelve stock indexes when news volumes are collected by Google News is in English and Chinese, respectively. Results are presented both for levels and changes.

News in English (Levels)

	stock market	stock market crash
	ρ	ρ
MSCI	0.50^{***}	0.77^{***}
S&P 500	0.57^{***}	0.75^{***}
DJIA	0.57^{***}	0.77^{***}
NASDAQ	0.55^{***}	0.74^{***}
Russell 2000	0.48^{***}	0.66^{***}
FTSE 100	0.60^{***}	0.79^{***}
Shanghai A	0.59^{***}	0.34^{***}
Shanghai B	0.47^{***}	0.33^{***}
Shenzhen A	0.49^{***}	0.26^{***}
Shenzhen B	0.50^{***}	0.33^{***}
Hong Kong H	0.66^{***}	0.73^{***}
Hong Kong Red Chip	0.66^{***}	0.69^{***}

News in Chinese (Levels)

	股市 (stock market)	股市崩溃 (stock market “collapse”)	股市崩盘 (stock market “crash”)
	ρ	ρ	ρ
MSCI	0.47^{***}	0.75^{***}	0.22^{*}
S&P 500	0.34^{***}	0.71^{***}	0.26^{**}
DJIA	0.32^{**}	0.71^{***}	0.27^{***}
NASDAQ	0.33^{***}	0.67^{***}	0.25^{**}
Russell 2000	0.35^{***}	0.67^{***}	0.19^{*}
FTSE 100	0.45^{***}	0.70^{***}	0.18
Shanghai A	0.23^{**}	0.40^{***}	0.00
Shanghai B	0.33^{***}	0.42^{***}	-0.02
Shenzhen A	0.23^{**}	0.37^{***}	-0.06
Shenzhen B	0.37^{***}	0.54^{***}	-0.03
Hong Kong H	0.29^{**}	0.69^{***}	0.14
Hong Kong Red Chip	0.33^{***}	0.67^{***}	0.17

Table I - *Continued*

News in English (Changes)

	stock market	stock market crash
	ρ	ρ
MSCI	0.19**	0.71***
S&P 500	0.18**	0.62***
DJIA	0.16*	0.62***
NASDAQ	0.15*	0.61***
Russell 2000	0.17**	0.58***
FTSE 100	0.23***	0.62***
Shanghai A	0.09	0.17**
Shanghai B	0.03	0.18**
Shenzhen A	0.02	0.07
Shenzhen B	0.04	0.19**
Hong Kong H	0.21**	0.53***
Hong Kong Red Chip	0.27***	0.54***

News in Chinese (Changes)

	股市 (stock market)	股市崩溃 (stock market "collapse")	股市崩盘 (stock market "crash")
	ρ	ρ	ρ
MSCI	0.09	0.63***	0.37***
S&P 500	0.06	0.61***	0.35***
DJIA	0.06	0.60***	0.35***
NASDAQ	0.10	0.59***	0.35***
Russell 2000	0.05	0.58***	0.27**
FTSE 100	0.02	0.54***	0.28**
Shanghai A	-0.24**	0.26**	0.04
Shanghai B	0.01	0.27**	0.10
Shenzhen A	-0.08	0.21*	-0.01
Shenzhen B	-0.05	0.39***	0.11
Hong Kong H	-0.07	0.54***	0.23**
Hong Kong Red Chip	-0.02	0.55***	0.29**

Table II
Results of English-Language News Volume - Volatility Regressions

In this Table I present results from the English-language regressions. In the upper regressions (News) the news volume is the dependent variable and in the lower regressions (Volatility) the stock market volatility is the dependent variable. In both regressions, the first explanatory variable, β_1, is always the one-month lagged value of the other variable (i.e. volatility in the upper regression and news volume in the lower regression) and the second explanatory variable, β_2, is always the one-month lagged value of the dependent variable. Each regression has an unreported intercept and an unreported dummy for missing values (which is rarely significant). *, ** and *** represent significance at the 10%, 5% and 1% level, respectively. The regressions are based on 105 monthly observations from September 11, 2006 to September 1, 2014 and the results are presented both for levels and for changes.

News in English (Levels)

		stock market			stock market crash		
	Dependent Variable	β_1	β_2	$\hat{R}^2$	β_1	β_2	$\hat{R}^2$
MSCI	News	-0.11*	0.88***	0.70	0.12	0.42***	0.25
	Volatility	0.058	0.49***	0.62	0.11*	0.42***	0.62
S&P 500	News	-0.080	0.88***	0.70	0.22	0.35***	0.26
	Volatility	0.083	0.61***	0.65	0.14*	0.54***	0.65
DJIA	News	-0.090	0.88***	0.70	0.17	0.38***	0.25
	Volatility	0.086*	0.54***	0.64	0.14**	0.47***	0.64
NASDAQ	News	-0.066	0.87***	0.70	0.20	0.36***	0.26
	Volatility	0.083	0.57***	0.61	0.14**	0.49***	0.62
Russell 2000	News	-0.080	0.87***	0.70	0.092	0.45***	0.25
	Volatility	0.12**	0.69***	0.63	0.22***	0.58***	0.64
FTSE 100	News	-0.080	0.88***	0.70	0.186	0.37***	0.26
	Volatility	0.057	0.49***	0.55	0.080	0.45***	0.55
Shanghai A	News	0.10*	0.80***	0.70	0.22**	0.44***	0.29
	Volatility	0.064	0.50***	0.56	0.061	0.50***	0.56
Shanghai B	News	0.079	0.82***	0.70	0.24***	0.43***	0.29
	Volatility	0.10	0.56***	0.39	0.090	0.56***	0.39
Shenzhen A	News	0.083	0.82***	0.70	0.18**	0.47***	0.27
	Volatility	0.10*	0.40***	0.41	0.087	0.41***	0.41
Shenzhen B	News	0.066	0.83***	0.70	0.21**	0.44***	0.28
	Volatility	0.055	0.43***	0.38	0.055	0.42***	0.38
Hong Kong H	News	-0.049	0.87***	0.70	0.12	0.43***	0.25
	Volatility	0.13	0.71***	0.51	0.023	0.76***	0.50
Hong Kong Red Chip	News	-0.035	0.86***	0.70	0.15	0.41***	0.25
	Volatility	0.097	0.64***	0.52	0.010	0.67***	0.51

Table II - *Continued*

News in English (Changes)

	Dependent Variable	stock market β_1	stock market β_2	stock market $\hat{R}^2$	stock market crash β_1	stock market crash β_2	stock market crash $\hat{R}^2$
MSCI	News	-0.038	0.52***	0.26	0.049	0.39***	0.17
	Volatility	0.058	0.34***	0.39	0.11*	0.27***	0.40
S&P 500	News	-0.014	0.52***	0.26	0.071	0.38***	0.17
	Volatility	0.063	0.33***	0.35	0.13**	0.26***	0.37
DJIA	News	-0.028	0.52***	0.26	0.032	0.40***	0.17
	Volatility	0.060	0.33***	0.36	0.13**	0.26***	0.38
NASDAQ	News	-0.019	0.52***	0.26	0.056	0.39***	0.17
	Volatility	0.056	0.27***	0.31	0.13**	0.21***	0.34
Russell 2000	News	-0.007	0.52***	0.26	0.007	0.42***	0.16
	Volatility	0.08*	0.30***	0.37	0.14***	0.22***	0.40
FTSE 100	News	-0.011	0.52***	0.26	0.056	0.39***	0.17
	Volatility	0.044	0.28***	0.33	0.079	0.25***	0.34
Shanghai A	News	0.13*	0.50***	0.28	0.079	0.41***	0.17
	Volatility	0.026	0.11***	0.12	0.029	0.10***	0.12
Shanghai B	News	0.11	0.51***	0.27	0.14	0.40***	0.18
	Volatility	0.060	0.12***	0.09	0.035	0.12***	0.08
Shenzhen A	News	0.048	0.52***	0.26	0.048	0.42***	0.17
	Volatility	0.043	0.11***	0.12	0.017	0.10***	0.11
Shenzhen B	News	0.12	0.51***	0.27	0.10	0.41***	0.18
	Volatility	0.025	0.12***	0.11	0.022	0.11***	0.11
Hong Kong H	News	0.022	0.51***	0.26	-0.035	0.44***	0.17
	Volatility	0.043	0.24***	0.31	0.056	0.22***	0.31
Hong Kong Red Chip	News	0.031	0.51***	0.26	0.006	0.42***	0.16
	Volatility	0.051	0.23***	0.32	0.059	0.21***	0.32

Table III
Results of Chinese-Language News Volume - Volatility Regressions

In this Table I present results from the Chinese-language regressions. In the upper regressions (News) the news volume is the dependent variable and in the lower regressions (Volatility) the stock market volatility is the dependent variable. In both regressions, the first explanatory variable, β_1, is always the one-month lagged value of the other variable (i.e. volatility in the upper regression and news volume in the lower regression) and the second explanatory variable, β_2, is always the one-month lagged value of the dependent variable. Each regression has an unreported intercept and an unreported dummy for missing values (which is rarely significant). *, ** and *** represent significance at the 10%, 5% and 1% level, respectively. The regressions are based on 51 monthly observations from November 1, 2010 to September 1, 2014 and the results are presented both for levels and for changes.

		News in Chinese (Levels)								
		股市 (stock market)			股市崩溃 (stock market "collapse")			股市崩盘 (stock market "crash")		
	Dependent Variable	β_1	β_2	$\hat{R}^2$	β_1	β_2	$\hat{R}^2$	β_1	β_2	$\hat{R}^2$
MSCI	News	0.055	0.81^{***}	0.69	0.59^{***}	0.21	0.56	0.002	0.30^{**}	0.09
	Volatility	0.092^{**}	0.26^{***}	0.60	0.11^{**}	0.22^{***}	0.60	0.11^{**}	0.29^{***}	0.61
S&P 500	News	0.005	0.84^{***}	0.69	0.52^{***}	0.29^{**}	0.54	0.012	0.30^{**}	0.09
	Volatility	0.098^{*}	0.26^{***}	0.44	0.22^{***}	0.14^{**}	0.51	0.17^{***}	0.24^{***}	0.54
DJIA	News	-0.0002	0.84^{***}	0.69	0.51^{***}	0.30^{**}	0.53	0.013	0.30^{**}	0.09
	Volatility	0.084^{*}	0.23^{***}	0.45	0.17^{***}	0.14^{**}	0.51	0.15^{***}	0.22^{***}	0.54
NASDAQ	News	-0.002	0.84^{***}	0.69	0.49^{***}	0.33^{**}	0.54	0.008	0.30^{**}	0.09
	Volatility	0.12^{**}	0.23^{***}	0.36	0.24^{***}	0.11	0.45	0.20^{**}	0.22^{***}	0.47
Russell 2000	News	0.023	0.83^{***}	0.69	0.46^{***}	0.35^{**}	0.52	-0.054	0.31^{**}	0.09
	Volatility	0.14^{*}	0.33^{***}	0.39	0.37^{***}	0.13	0.55	0.25^{***}	0.34^{***}	0.51
FTSE 100	News	0.063	0.81^{***}	0.69	0.43^{***}	0.36^{**}	0.50	-0.001	0.30^{**}	0.09
	Volatility	0.12^{***}	0.21^{***}	0.55	0.14^{***}	0.16^{***}	0.53	0.076^{*}	0.25^{***}	0.50
Shanghai A	News	0.10	0.82^{***}	0.70	-0.051	0.69^{***}	0.40	-0.12	0.30^{**}	0.10
	Volatility	0.035	0.050	0.02	0.074^{*}	0.030	0.07	0.029	0.060	0.01
Shanghai B	News	0.10	0.81^{***}	0.70	0.14	0.60^{***}	0.42	-0.033	0.30^{**}	0.09
	Volatility	0.13^{*}	0.072	0.09	0.067	0.081	0.04	0.006	0.11^{*}	0.02
Shenzhen A	News	0.066	0.85^{***}	0.70	-0.13	0.71^{***}	0.42	-0.13	0.29^{**}	0.11
	Volatility	0.052	0.029	0.09	0.067	0.029	0.05	-0.016	0.039	0.00
Shenzhen B	News	0.092	0.81^{***}	0.70	0.19	0.56^{***}	0.43	-0.065	0.30^{**}	0.09
	Volatility	0.11^{*}	0.17^{***}	0.26	0.11^{*}	0.14^{**}	0.26	0.009	0.21^{***}	0.20
Hong Kong H	News	0.074	0.82^{***}	0.70	0.26^{*}	0.48^{***}	0.44	-0.075	0.31^{**}	0.10
	Volatility	0.082	0.33^{***}	0.43	0.27^{***}	0.16^{**}	0.57	0.084	0.34^{***}	0.43
Hong Kong Red Chip	News	0.071	0.82^{***}	0.70	0.24^{*}	0.50^{***}	0.44	-0.073	0.31^{**}	0.10
	Volatility	0.082	0.30^{**}	0.43	0.27^{***}	0.15^{**}	0.57	0.084	0.32^{***}	0.43

Table III - *Continued*

News in Chinese (Changes)

	Dependent Variable	股市 (stock market)			股市崩溃 (stock market “collapse”)			股市崩盘 (stock market “crash”)		
		β_1	β_2	$\hat{R}^2$	β_1	β_2	$\hat{R}^2$	β_1	β_2	$\hat{R}^2$
MSCI	News	0.025	0.61***	0.39	0.45***	0.23	0.36	0.18	0.009	0.00
	Volatility	0.039	0.25***	0.39	0.16***	0.16***	0.48	0.19***	0.19***	0.57
S&P 500	News	-0.020	0.61***	0.39	0.43***	0.25*	0.35	0.14	0.025	0.00
	Volatility	0.051	0.26***	0.27	0.29***	0.09	0.46	0.33***	0.15***	0.64
DJIA	News	-0.020	0.61***	0.39	0.41***	0.27*	0.34	0.14	0.027	0.00
	Volatility	0.049	0.26***	0.28	0.25***	0.11**	0.44	0.31***	0.16***	0.61
NASDAQ	News	-0.011	0.61***	0.39	0.44***	0.25*	0.36	0.13	0.028	0.00
	Volatility	0.059	0.20***	0.20	0.27***	0.06	0.40	0.31***	0.11**	0.57
Russell 2000	News	-0.005	0.61***	0.39	0.35**	0.31**	0.31	0.030	0.067	0.00
	Volatility	0.051	0.24***	0.22	0.33***	0.05	0.52	0.31***	0.15***	0.61
FTSE 100	News	0.026	0.61***	0.39	0.34**	0.34**	0.31	0.12	0.042	0.00
	Volatility	0.05	0.22***	0.35	0.14***	0.14***	0.43	0.17***	0.16***	0.52
Shanghai A	News	0.034	0.62***	0.39	-0.13	0.55***	0.24	-0.041	0.077	0.00
	Volatility	-0.074**	0.025	0.08	0.049	0.030	0.02	0.049	0.04	0.03
Shanghai B	News	0.076	0.61***	0.39	0.097	0.49***	0.24	0.10	0.065	-0.02
	Volatility	-0.039	0.090*	0.04	0.019	0.083	0.03	0.048	0.083*	0.04
Shenzhen A	News	0.038	0.62***	0.39	-0.19	0.56***	0.26	-0.081	0.075	-0.03
	Volatility	-0.058*	0.03	0.07	0.037	0.026	0.03	0.016	0.04	0.01
Shenzhen B	News	0.046	0.62***	0.39	0.14	0.46***	0.24	0.065	0.068	-0.03
	Volatility	-0.008	0.14***	0.14	0.069	0.11**	0.18	0.061	0.13***	0.17
Hong Kong H	News	0.037	0.62***	0.39	0.19	0.41***	0.25	0.019	0.071	-0.03
	Volatility	0.011	0.26***	0.44	0.18***	0.16**	0.59	0.14***	0.23***	0.57
Hong Kong Red Chip	News	0.048	0.61***	0.39	0.16	0.43***	0.24	0.021	0.069	-0.03
	Volatility	-0.004	0.23**	0.31	0.19***	0.12**	0.48	0.11**	0.20***	0.39

Table IV

Results of Volatility Forecasts using English-Language News Volumes

In this Table I present stock return volatility forecasting results for English-language news volumes. In total, 51 monthly forecasts are made in each category and all forecasts are based on regression parameters estimated using a rolling window of 54 monthly observations starting from November 1, 2010 to September 1, 2014. The numbers in the rows labelled (Without news) are the ones when the one-month lagged news volume is not included in the prediction and the numbers in the rows labelled (With news) are the ones when the one-month lagged news volume is included in addition to one-month lagged stock return volatility. MAE is the mean absolute error and QL is the QL loss function. The smallest (best) forecasting error/loss is typed in bold and the results are presented both for levels and for changes.

		News in English (Levels)			
		stock market		stock market crash	
		MAE $(\cdot 10^{-5})$	*QL*	*MAE* $(\cdot 10^{-5})$	*QL*
MSCI	Without news	2.13	0.0244	2.13	0.0244
	With news	**1.80**	**0.0183**	**1.56**	**0.0142**
S&P 500	Without news	2.91	0.0322	2.91	0.0322
	With news	**2.42**	**0.0235**	**1.77**	**0.0135**
DJIA	Without news	2.45	0.0249	2.45	0.0249
	With news	**2.13**	**0.0195**	**1.40**	**0.0093**
NASDAQ	Without news	4.37	0.0646	4.37	0.0646
	With news	**3.74**	**0.0505**	**3.51**	**0.0456**
Russell 2000	Without news	4.69	0.0470	4.69	0.0470
	With news	**4.00**	**0.0363**	**2.88**	**0.0207**
FTSE 100	Without news	2.46	0.0247	2.46	0.0247
	With news	**1.88**	**0.0154**	**1.66**	**0.0123**
Shanghai A	Without news	3.78	0.0225	3.78	0.0225
	With news	**2.82**	**0.0134**	**2.95**	**0.0146**
Shanghai B	Without news	6.37	0.0650	6.37	0.0650
	With news	**4.83**	**0.0418**	**5.59**	**0.0529**
Shenzhen A	Without news	6.90	0.0483	6.90	0.0483
	With news	**6.09**	**0.0393**	**6.39**	**0.0426**
Shenzhen B	Without news	6.46	0.1678	6.46	0.1678
	With news	**5.96**	**0.1505**	**5.91**	**0.1487**
Hong Kong H	Without news	2.81	0.0111	2.81	0.0111
	With news	**2.39**	**0.0082**	**1.38**	**0.0029**
Hong Kong Red Chip	Without news	2.96	0.0166	2.96	0.0166
	With news	**2.38**	**0.0112**	**2.02**	**0.0084**

Table IV - *Continued*

News in English (changes)

		stock market		stock market crash	
		MAE $(\cdot 10^{-3})$	*QL*	*MAE* $(\cdot 10^{-3})$	*QL*
MSCI	Without news	1.24	0.1056	1.24	0.1056
	With news	**1.15**	**0.0863**	**0.77**	**0.0320**
S&P 500	Without news	1.91	0.2377	1.91	0.2377
	With news	**1.81**	**0.2019**	**1.08**	**0.0507**
DJIA	Without news	1.95	0.2602	1.95	0.2602
	With news	**1.85**	**0.2202**	**1.21**	**0.0679**
NASDAQ	Without news	3.92	0.3364	3.92	0.3364
	With news	**3.83**	**0.3126**	**3.66**	**0.2711**
Russell 2000	Without news	3.25	0.3327	3.25	0.3327
	With news	**3.16**	**0.3038**	**2.68**	**0.1841**
FTSE 100	Without news	1.08	0.0901	1.08	0.0901
	With news	**1.04**	**0.0814**	**0.39**	**0.0083**
Shanghai A	Without news	0.69	0.0865	0.69	0.0865
	With news	**0.66**	**0.0749**	**0.35**	**0.0164**
Shanghai B	Without news	1.14	0.1074	1.14	0.1074
	With news	**1.01**	**0.0774**	**0.96**	**0.0676**
Shenzhen A	Without news	2.78	0.7350	2.78	0.7350
	With news	**2.66**	**0.6132**	**2.62**	**0.5785**
Shenzhen B	Without news	4.83	0.2569	4.83	0.2569
	With news	**4.78**	**0.2481**	**4.70**	**0.2362**
Hong Kong H	Without news	0.96	0.0767	0.96	0.0767
	With news	**0.91**	**0.0665**	**0.37**	**0.0083**
Hong Kong Red Chip	Without news	1.14	0.1092	1.14	0.1092
	With news	**1.10**	**0.1003**	**0.75**	**0.0379**

Table V
Results of Volatility Forecasts using Chinese-Language News Volumes

In this Table I present stock return volatility forecasting results for Chinese-language news volumes. In total, 25 monthly forecasts are made in each category and all forecasts are based on regression parameters estimated using a rolling window of 26 monthly observations starting from October 29, 2012 to September 1, 2014. The numbers in the rows labelled (Without news) are the ones when the one-month lagged news volume is not included in the prediction and the numbers in the rows labelled (With news) are the ones when the one-month lagged news volume is included in addition to one-month lagged stock return volatility. MAE is the mean absolute error and QL is the QL loss function. The smallest (best) forecasting error/loss is typed in bold and the results are presented both for levels and for changes.

News in Chinese (Levels)

		股市 (stock market)		股市崩溃 (stock market "collapse")		股市崩盘 (stock market "crash")	
		MAE $(\cdot 10^{-5})$	*QL*	*MAE* $(\cdot 10^{-5})$	*QL*	*MAE* $(\cdot 10^{-5})$	*QL*
MSCI	Without news	5.82	0.0400	5.82	0.0400	5.82	0.0400
	With news	**4.91**	**0.0301**	**5.47**	**0.0361**	**4.71**	**0.0280**
S&P 500	Without news	7.49	0.0473	7.49	0.0473	**7.49**	**0.0473**
	With news	**6.82**	**0.0405**	**6.56**	**0.0380**	7.73	0.0497
DJIA	Without news	5.90	0.0329	5.90	0.0329	**5.90**	**0.0329**
	With news	**5.31**	**0.0276**	**5.30**	**0.0275**	6.35	0.0372
NASDAQ	Without news	13.6	0.1210	13.6	0.1210	**13.6**	**0.1210**
	With news	**12.7**	**0.1104**	**12.7**	**0.110**	14.9	0.1375
Russell 2000	Without news	**13.7**	**0.0823**	13.7	0.0823	13.7	0.0823
	With news	14.6	0.0905	**13.3**	**0.0787**	**13.1**	**0.0769**
FTSE 100	Without news	5.90	0.0325	**5.90**	**0.0325**	5.90	0.0325
	With news	**4.59**	**0.0212**	6.53	0.0386	**3.06**	**0.0103**
Shanghai A	Without news	**7.61**	**0.0221**	7.61	0.0221	**7.61**	**0.0221**
	With news	7.70	0.0225	**7.04**	**0.0192**	7.87	0.0234
Shanghai B	Without news	**7.55**	**0.0267**	**7.55**	**0.0267**	7.55	0.0267
	With news	7.72	0.0277	7.98	0.0294	**6.44**	**0.0203**
Shenzhen A	Without news	**10.9**	**0.0318**	**10.9**	**0.0318**	10.9	0.0318
	With news	11.2	0.0330	11.0	0.0322	**8.70**	**0.0215**
Shenzhen B	Without news	9.10	0.0998	**9.10**	**0.0998**	9.10	0.0998
	With news	**8.42**	**0.0889**	9.17	0.1009	**4.79**	**0.0898**
Hong Kong H	Without news	10.6	0.0326	10.6	0.0326	10.6	0.0326
	With news	**9.64**	**0.0276**	**10.1**	**0.0297**	**10.6**	**0.0322**
Hong Kong Red Chip	Without news	**7.65**	**0.0252**	7.65	0.0252	7.65	0.0252
	With news	8.08	0.0276	**6.69**	**0.0199**	**4.94**	**0.0116**

Table V - *Continued*

News in Chinese (Changes)

		股市 (stock market)		股市崩溃 (stock market "collapse")		股市崩盘 (stock market "crash")	
		MAE $(\cdot 10^{-3})$	*QL*	*MAE* $(\cdot 10^{-3})$	*QL*	*MAE* $(\cdot 10^{-3})$	*QL*
MSCI	Without news	1.99	0.0571	**1.99**	**0.0571**	1.99	0.0571
	With news	**1.86**	**0.0480**	2.52	0.1043	**1.77**	**0.0425**
S&P 500	Without news	5.19	0.6378	**5.19**	**0.6378**	**5.19**	**0.6378**
	With news	**5.17**	**0.6245**	5.23	0.6542	6.10	1.3033
DJIA	Without news	5.21	0.6697	**5.21**	**0.6697**	**5.21**	**0.6697**
	With news	**5.15**	**0.6381**	5.35	0.7447	6.02	1.2770
NASDAQ	Without news	13.0	3.2917	13.0	3.2917	**13.0**	**3.2917**
	With news	**13.0**	**3.2835**	**12.9**	**3.1819**	14.0	7.0598
Russell 2000	Without news	**9.88**	**1.8001**	9.88	1.8001	9.88	1.8001
	With news	9.93	1.8603	**9.50**	**1.4474**	**9.21**	**1.2339**
FTSE 100	Without news	2.28	0.0983	**2.28**	**0.0983**	2.28	0.0983
	With news	**2.22**	**0.0919**	3.06	0.2260	**0.50**	**0.0031**
Shanghai A	Without news	3.17	1.2796	**3.17**	**1.2796**	**3.17**	**1.2796**
	With news	**3.13**	**1.2003**	3.65	2.9913	4.23	19.926
Shanghai B	Without news	0.91	0.0114	**0.91**	**0.0114**	0.91	0.0114
	With news	**0.56**	**0.0040**	1.60	0.0416	**0.62**	**0.0050**
Shenzhen A	Without news	**6.06**	**0.9781**	**6.06**	**0.9781**	6.06	0.9781
	With news	6.74	1.6757	6.57	1.4629	**5.36**	**0.5843**
Shenzhen B	Without news	**8.53**	**0.1660**	**8.53**	**0.1660**	8.53	0.1660
	With news	9.25	0.2118	9.01	0.1955	**6.46**	**0.0783**
Hong Kong H	Without news	**3.06**	**0.2712**	**3.06**	**0.2712**	**3.06**	**0.2712**
	With news	3.35	0.3636	3.54	0.4415	3.47	0.4124
Hong Kong Red Chip	Without news	**2.52**	**0.1358**	2.52	0.1358	2.52	0.1358
	With news	2.58	0.1458	**2.25**	**0.1003**	**1.13**	**0.0187**

Table VI
Correlations between News Volume and Volatility: Robustness

As a robustness test, in this Table I present News Volume – Volatility correlations for the MSCI World stock index when news volumes are collected by Google News in English and Chinese, respectively. In the first part of the Table I test the robustness of the results in the paper to the removal of extreme observations, crisis observations and missing observations and in the second part I test the robustness of the results to slight changes to the news volume collection process. Results are presented both for levels and for changes.

News in English (Levels)

		stock market	stock market crash
		ρ	ρ
MSCI	Excluding Extremes	0.18^{**}	0.35^{***}
	Excluding Crisis	-0.04	0.53^{***}
	Excluding Missing	0.47^{***}	0.79^{***}

News in English (Changes)

		stock market	stock market crash
		ρ	ρ
MSCI	Excluding Extremes	0.20^{**}	0.50^{***}
	Excluding Crisis	0.11	0.50^{***}
	Excluding Missing	0.18^{**}	0.72^{***}

News in Chinese (Levels)

		股市 (stock market)	股市崩溃 (stock market “collapse”)	股市崩盘 (stock market “crash”)
		ρ	ρ	ρ
MSCI	Excluding Extremes	0.44^{***}	0.57^{***}	0.07
MSCI	Excluding Missing	0.46^{***}	0.77^{***}	0.26^{**}

News in Chinese (Changes)

		股市 (stock market)	股市崩溃 (stock market “collapse”)	股市崩盘 (stock market “crash”)
		ρ	ρ	ρ
MSCI	Excluding Extremes	0.05	0.27^{**}	0.24^{**}
MSCI	Excluding Missing	0.06	0.65^{**}	0.40^{***}

Table VI - *Continued*

	Levels		
	"stock market crash"	global stock market crash	股市暴跌 (stock market "slump")
	ρ	ρ	ρ
MSCI	0.67^{***}	0.71^{***}	0.54^{***}
S&P 500	0.66^{***}	0.69^{***}	0.49^{***}
DJIA	0.68^{***}	0.71^{***}	0.51^{***}
NASDAQ	0.64^{***}	0.66^{***}	0.46^{***}
Russell 2000	0.56^{***}	0.60^{***}	0.41^{***}
FTSE 100	0.71^{***}	0.75^{***}	0.55^{***}
Shanghai A	0.29^{***}	0.30^{***}	0.30^{**}
Shanghai B	0.31^{***}	0.32^{***}	0.33^{***}
Shenzhen A	0.21^{**}	0.23^{***}	0.27^{**}
Shenzhen B	0.31^{***}	0.33^{***}	0.40^{***}
Hong Kong H	0.69^{***}	0.72^{***}	0.38^{***}
Hong Kong Red Chip	0.65^{***}	0.69^{***}	0.47^{***}

	Changes		
	"stock market crash"	global stock market crash	股市暴跌 (stock market "slump")
	ρ	ρ	ρ
MSCI	0.60^{***}	0.73^{***}	0.48^{***}
S&P 500	0.51^{***}	0.65^{***}	0.41^{***}
DJIA	0.50^{***}	0.65^{***}	0.42^{***}
NASDAQ	0.49^{***}	0.61^{***}	0.37^{***}
Russell 2000	0.48^{***}	0.59^{***}	0.34^{***}
FTSE 100	0.50^{***}	0.68^{***}	0.43^{***}
Shanghai A	0.11	0.16^{*}	0.26^{**}
Shanghai B	0.18^{**}	0.20^{**}	0.24^{***}
Shenzhen A	0.03	0.07	0.25^{**}
Shenzhen B	0.18^{**}	0.25^{***}	0.33^{***}
Hong Kong H	0.43^{***}	0.61^{***}	0.35^{***}
Hong Kong Red Chip	0.45^{***}	0.62^{***}	0.48^{***}

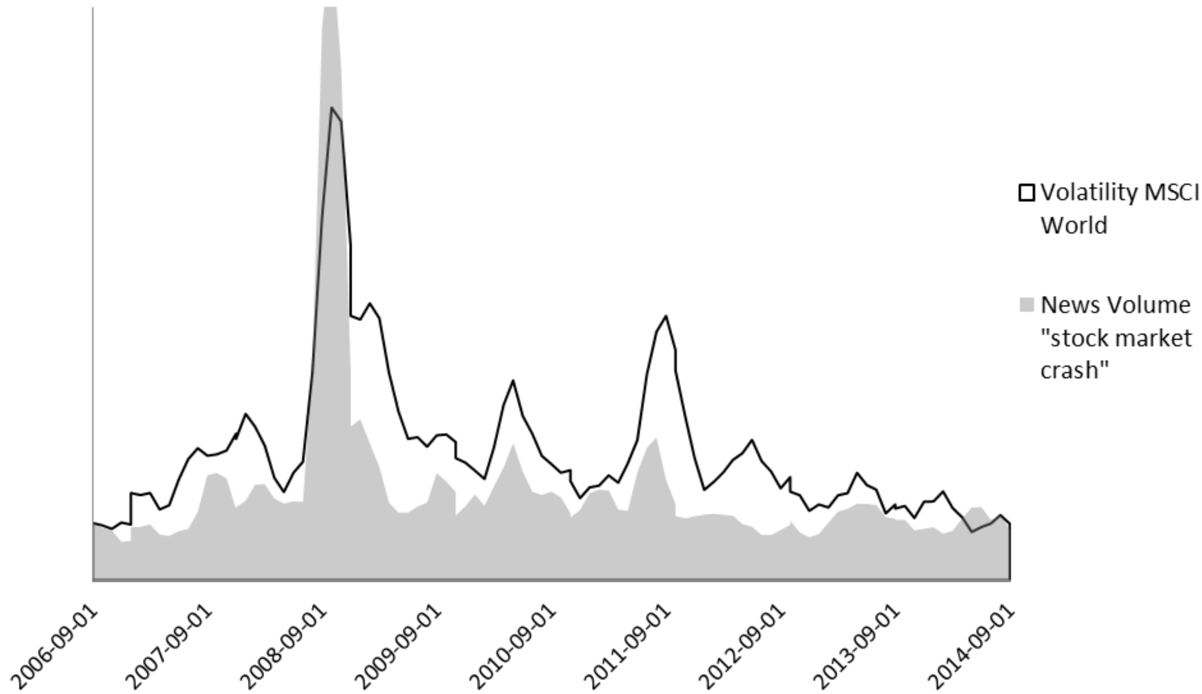

Figure 1. English-language news volume and MSCI World stock market volatility. This graph shows the English-language (US edition) Google News volumes for the search string "stock market crash" together with the MSCI World stock return volatility. Both the news volume and the stock volatility are normalized to start at one and are sampled on a monthly basis but smoothed using a three-month (quarterly) smoothing window.

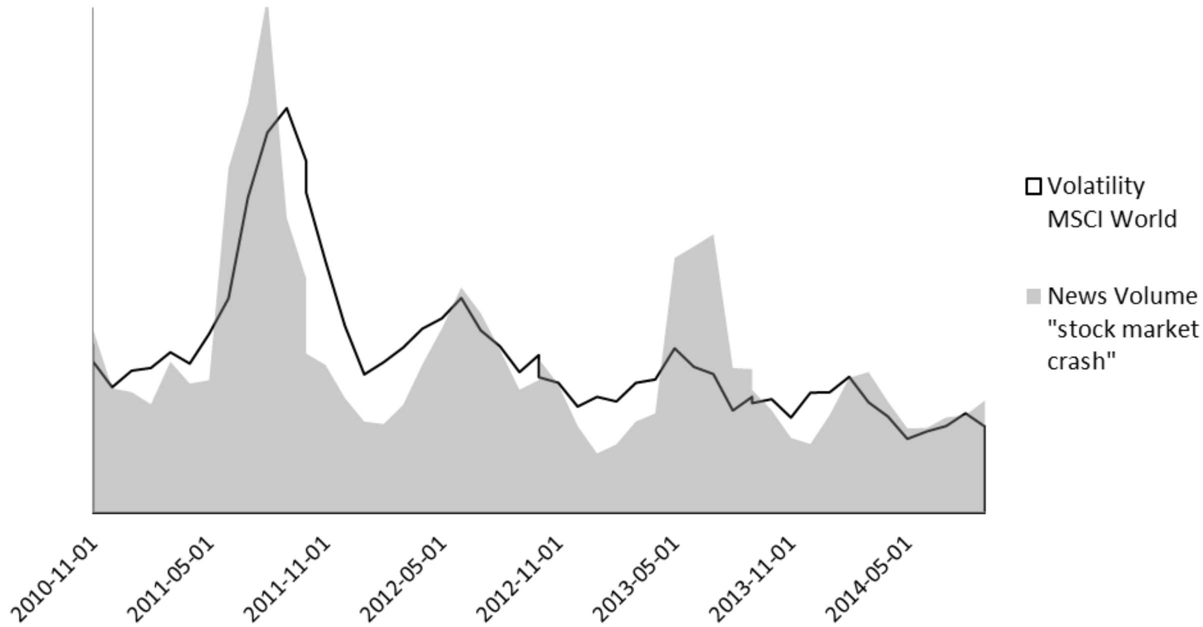

Figure 2. Chinese-language news volume and MSCI World stock market volatility. This graph shows the Chinese-language (China edition) Google News volumes for the search string "stock market crash" together with the MSCI World stock return volatility. Both the news volume and the stock volatility are normalized to start at one and are sampled on a monthly basis but smoothed using a three-month (quarterly) smoothing window.