Shock Waves of Political Risk on the Stock Market: The Case of Korean Companies in the U.S.*

YUNJUNG PAK | YONSEI UNIVERSITY
YOUNG-JIN KIM | YONSEI UNIVERSITY
MIN SONG | YONSEI UNIVERSITY
YONG-HAK KIM** | YONSEI UNIVERSITY

We introduce a method to examine the effects of the U.S. news sentiments *on Korean firms listed on the* New York Stock Exchange (NYSE) *when* North Korea *causes geo-political risks to the* South Korean *economy. Stock prices are evaluated by market factors and return to their fundamental values in the long term. Nevertheless, external shocks such as geo-political and international risks often cause stock volatility. Using a semi-supervised machine learning approach, we classify negative and positive news from five major newspapers in the U.S. to scrutinize the degree of North Korean risk and its influence on the stock prices of Korean firms listed on the NYSE. We find that news related to North Korea have an impact on the stock volatility in the U.S. and Korea. We could detect the direct impact of political risk posed by North Korea on the NYSE, but it was weaker than their indirect effects through the Korean stock Market.*

Keywords: *public news, asymmetric information, stock volatility, sentiment analysis, semi-supervised learning*

Introduction

The "Korea discount" refers to the phenomenon of South Korean firms' stock prices being undervalued because of continual investment risks posed by domestic and international factors. Researchers have cited the non-transparent governance of companies, the rigidity of the labor market, the high dependency on imports (Park 2005), and, specifically, the uncertainty of *North-South Korean relations*, as major factors contributing to the Korea discount (Kim 2011; Ahn et al. 2010; Moon 2007; Nam 2004). A report by S&P's Country Risk Assessment team, for instance, noted that "The South Korean economy is exposed to some security risks associated with the nuclear-capable [North Korea]. The security risks have been exacerbated in recent years by North Korea's nuclear program (S&P 2012.4.19)."

Researchers investigating the geo-political factors have focused on the direct effects of the North-South Korean relationship, and/or the mediating effects of the Korean media on the Korean stock market (Ahn et al. 2010; Moon 2006; Nam 2004). These studies, however, did not capture the essence of the Korea discount in the global market, where *foreign investors* evaluate the risks in the Korean peninsula. One notable exception was a study by Kim (2011), who analyzed the changing *net-demand of foreign investors* in the Korean stock market. Yet, this study also failed to show how the Korea discount operates in a global market such as the *New York Stock Exchange* (*NYSE*).

Since 1994, major Korean firms, starting with *Pohang Iron and Steel Company* (*POSCO*) and including *SK telecom* and *LG Display*, have enlisted in the *NYSE* so as to acquire direct foreign investment. Using these companies' daily stock price data, we are now able to investigate whether perceptions of geo-political risk absorbed from the media influences stock prices in a global market, and to measure the magnitude of this effect. In short, this paper aims to examine the impact of the news sentiments displayed by the U.S. press regarding North Korean political events (such as a nuclear bomb test) on the stock prices of South Korean firms listed in the *NYSE*. Needless to say, the macro-economic indices and financial status of individual firms are the main determinants of stock prices. Yet there exists a short-term synchronization between the risk perception of investors, and daily stock prices (Tetlock 2007; Groß-Klußmann and Hautsch 2011; Kim et al. 2012).

We collected articles from five major newspapers (*The Wall Street*

Journal, Financial Times, The Times, Washington Post, and *The New York Times*) containing the keyword "North Korea" in their headline over a period of twenty years, from 1993 to 2012. Incorporating an active learning algorithm using both human coders and automatic machine classifiers, we classified each of the articles into three categories; negative, neutral, and positive sentiment. This semi-supervised method, developed recently, enabled us to convert unstructured large text data into structured information systemically and objectively (Pang and Lee 2008; Bollen et al. 2011, Feldman et al 2011).

Theoretical Review

Political Risk and Stock Market

Stock market prices fluctuate in the event of changes that affect the fundamental values of firms (Fama 1970). However, as behavioral economists argue, stock prices also move according to the changing expectations of investors (Shiller 2003; Tetlock 2007; Antweiler and Frank 2004). In this context, political uncertainty has been studied as a significant external factor affecting investor expectations. Empirical studies have focused on the Middle East, where political and military conflicts have occurred frequently (Eldor and Melnick 2004; Arin et al. 2008; Zach 2003). Eldor and Melnick (2004), for example, found that suicide attacks had a strong negative effect on the Tel Aviv stock market between 1990 and 2003 and Zach (2003) has shown how the Israel's Tel Aviv stock market reacts to Israeli political events related to the peace process. He found that daily stock returns following political events were stronger than on days with no events. So despite some degree of difference, we can conclude that political risks have negative effects on national economies and the stock market of that country in particular.

However, the effects of geo-political factors could be nonlinear. Using the event study methodology, Guidolin and La Ferrara (2010) measured the markets' reaction to violent conflict. Interestingly, they found that the U.S. market, and sometimes even the markets of the UK, France, and Japan, positively reacted to international conflicts and shocks, and that these had an even more significant affects than national ones during the period from 1974 and 2004. The countries were related to the international consumption and transformation of raw materials, and some companies made their profits from economic activities related to the conflicts. In Turkey, Mehdian et al.

(2008) argued that political events caused rising stock indices regardless of whether the news is positive or negative. Investors set the stock prices below their fundamental values because unexpected information such as political risks raised the uncertainty and risk in the stock market, but they were subsequently able to recover this and more over a corrective period. In addition, recent studies suggest that the effects of political risks might be stronger in emerging markets or developing countries. Arin et al. (2008) and Bilson et al. (2001) found that political risks could influence both stock prices and stock volatility and that their impact was exaggerated in emerging markets. These findings showed that the effects of geo-political factors can vary according to international locations. Since the Korean peninsula is not only a high risk region due to the uncertainty of North-South relations, but has also been one of the most studied developing countries and emerging markets, it may be an interesting way to approach studying "the Korea Discount."

The Korea Discount

Numerous Empirical studies have focused on the North Korean risk as a factor in the under-valuing of Korean firms (Moon 2007; Ahn et al. 2010; Hwang and Kim 2004). As financial indices are sensitive to political changes and security problems (Arin et al. 2008), stock market reactions have been used to measure political uncertainties. The existence of the Korea discount depends on whether investors' decisions are synchronized with the perceived degree of risk emanating from North Korea or not. Two scenarios are possible: in the event of good news regarding North-South Korean relations, foreign investors tend to buy more Korean stocks as they perceive a relatively lower level of geo-political risks than usual; when the relations between the two countries become strained, investors tend to sell Korean stocks (Moon 2007; Ahn et al. 2010; Hwang and Kim 2004).

In this respect, Moon (2007) found a statistically significant difference between the effect of positive and negative news: when negative news were released, the Korea discount would increase, while it declined in the event of positive news over the 2003-2004 period. Another study (Ahn et al. 2010) found that news regarding North Korea had effects on the *KOSPI* and *KOSDAQ* (equivalents of *NYSE* and *NASDAQ*) between 1998 and 2009. Moreover, the authors found industry-specific impacts: the stock prices of firms related with North Korea moved up when positive news arose, while that of firms in the arms industry moved down when there were negative

news.

However, a few studies have questioned the very existence of the Korea Discount caused by North Korea, and have claimed that the Korea discount is exaggerated (Nam 2004; Park 2005; Kim 2011; Pyun and Huh 2014). For instance, Nam (2004) found a U-shaped curve: a quick recovery of stock indices after a sharp drop. The relationship between news and stock price, however, was not statistically significant when he picked sixty North Korean events of his own selection, and expanded the time horizon to cover the period 1990-2002. He interpreted this result as evidence of a learning effect, i.e., the investors learned the "false-alarm effects" of North Korean risks. Analyzing monthly data for the years 2000-2007, Kim (2011) also claimed that investors in the Korean stock market were not significantly affected by changes in the North-South Korea relationship. Pyun and Huh (2014) added that the effects of North Korea risk on the Korean stock market had significantly declined from 2004 to 2012, only explaining 0.2% of the stock volatility. Those studies suggest a "learning effect" in that as political risks caused by North Korea recurred with no significant effects on Korean economy, individual investors and even foreign institutional investors had reduced reflection rates of North Korea risks after 2000.

The contradictory results may arise from the different time periods and/ or differences in classification rules of news sentiments. Furthermore, previous studies have analyzed only the domestic stock market, not allowing us to determine whether the Korea discount exists in the global market. Kim's (2011) study may be a notable exception in that he scrutinized *foreign investors*' decisions in the Korean stock market. However, his study is also limited for the following reasons. Since he used the *Korea Peace Index (KOPI)* instead of the daily changes in the individual stock prices, the nature of the stock market as a short-term and firm-specific shock-absorbing system was overlooked. We will thus extend our scope to the global market and use daily stock price data for specific firms listed in the *NYSE*.

Media and Stock Market

Since it is hard to measure political changes in reality, researchers have focused on the relationships between the political events and media coverage of the issues as a proxy of public perception of the events. The expectations of investors are often affected by signals from bad or good news, resulting in stock price volatility (Mittermayer 2004; Tetlock 2007, 2010, 2011) Tetlock (2007) first introduced and pioneered in the research area studying the

relationships between the media and the stock market. He classified negative words from articles in the *Wall Street Journal* according to their definition in the Harvard-IV-04 dictionary, and showed that negative news lowered stock prices. The prices, however, went back to normal as time went on because the news themselves could not influence firms' fundamental value. In subsequent studies, Tetlock (2010, 2011; Tetlock et al. 2008) emphasized the role of the media by measuring the strength of negative words in the media. He demonstrated that investors are indeed affected by recent news under the condition of information asymmetry. Automatic text-mining methods have been effective in categorizing news sentiments, especially when a researcher has to deal with large number of news articles over a long period of time (Antweiler and Frank 2004). Groß-Klußmann and Hautsch (2011), for instance, used supervised learning to classify news sentiments in order to predict stock price fluctuations. Antweiler and Frank (2004) used a Naïve Bayesian Classifier to analyze 15 million internet entries on the stock message boards of about 45 companies and predicted *Dow Jones Industrial Index* and *Dow Jones Internet Index on Yahoo! Finance and Raging Bull*. The authors concluded that the general effects of the news sentiments were statistically significant but small. These supervised machine-learning approaches present the advantage of creating a new dictionary tailored to the researchers' purposes; they learn to distinguish the same word appearing in different contexts.

Text-Mining Approach

Despite the merits of supervised learning classifiers, it still demonstrates low classification reliability and a lack of accuracy. To deal with a *specific context* such as the news sentiments regarding the North Korean risk, unsupervised learning using a general dictionary such as the Harvard-IV-4 dictionary (Tetlock 2007), LIWC (Tumasjan et al. 2011), KrKwic (Park et al. 2011) has serious limitations. Moreover, as there was no preexisting training data dealing with North Korea supervised learning could also not be applied directly. We thus preferred to build a text-mining model.

Therefore, we had to develop our own semi-supervised learning methods in order to measure the North Korean risk represented on the U.S. news (Song et al. 2011). The semi-supervised approach adopted in this paper incorporates an active learning algorithm into the semi-supervised approach to sentiment classification. It has a comprehensive set of features, which are comprised of syntactic, semantic, and lexical properties. They were selected

by Natural Language Processing (NLP) techniques to boost the classification performance. After collecting a large dataset, we first trained the classification model with the training data of news sentiments curated by plural human coders, and fed them to the automatic machine classifiers to learn from human classification. Using text-mining methods to handle the large text documents such as news archives, we could measure a proxy of investor's perception patterns toward political risks.

Data and Analysis

Data

For the period from January 1 1993 to December 31 2012, we identified and downloaded 4,484 articles from five newspapers (*The New York Times, Financial Times, The Wall Street Journal, Washington Post,* and *The Times*) that contained the keyword "North Korea" in their headline. The criteria for choosing the newspapers included prominence and ideological diversity (Antweiler and Frank 2004; Tetlock 2011; Yun et al. 2012). We relied on news DB archive, *Factiva (http://global.factiva.com).*

The articles about North Korea appeared on 2,585 out of 7,670 days. Twenty-three news articles were published on October 18, 2002, the largest number in our data set, when the then U.S. Secretary of Defense, Donald Rumsfelt, confirmed the North Korean nuclear program. <Figure 1> summarizes our data collection and classification procedures of news sentiment. Our research strategy was to rely on human coders to classify the news sentiment first, and let computer algorithms to learn from human classification.

Before using the machine classifiers, it was necessary to convert the content of the documents into unique features through text pre-processing.
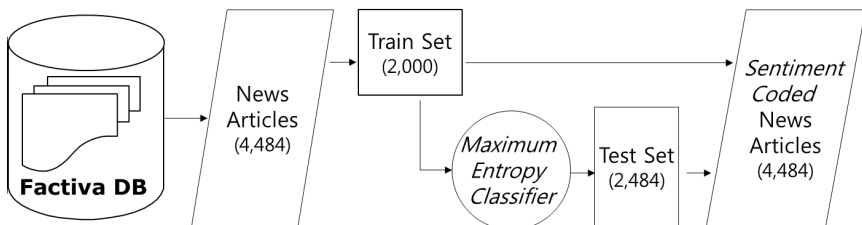

Fig. 1.—Research Design for Data Collection and Analysis Processes

This process was divided into three steps: a) feature extraction, b) feature selection, and c) document representation (Mittermayer, 2004). Feature extraction was the first pre-processing step generating "a bag of words" that described the Korea discount-related documents. The main objective of feature selection was to eliminate those features that provided few or less important items of information. At the document representation stage, the documents were represented in terms of the key features. Thus, the representation of a document was a feature vector of n elements, where n was the number of features remaining when the selection process was complete. In this paper, we used 30,347 unique features and over 233,057 words to build the machine models.

We selected 2,000 articles randomly from our data and had three human coders classify them. There were three possible values for news sentiment: negative, neutral, or positive. These became the training data set, which allow us to apply semi-supervised learning methods to process and classify the remaining 2,494 articles (which we call the test data). The training data consisted of all-consented data, 55.3% (1,105 articles) and two-coder consented, 44.3% (886 articles). These data provided basic features for making computational classifiers. To classify the extra data (2,484 articles), we used four different classifiers provided in *Mallet,* written in Java programming language (McCallum 2002): *Naïve Bayesian, Support Vector Machine, Decision Tree, and Maximum Entropy.* In order to evaluate them with precision, we used a 10-fold cross-evaluation, dividing 2,000 by 10 and
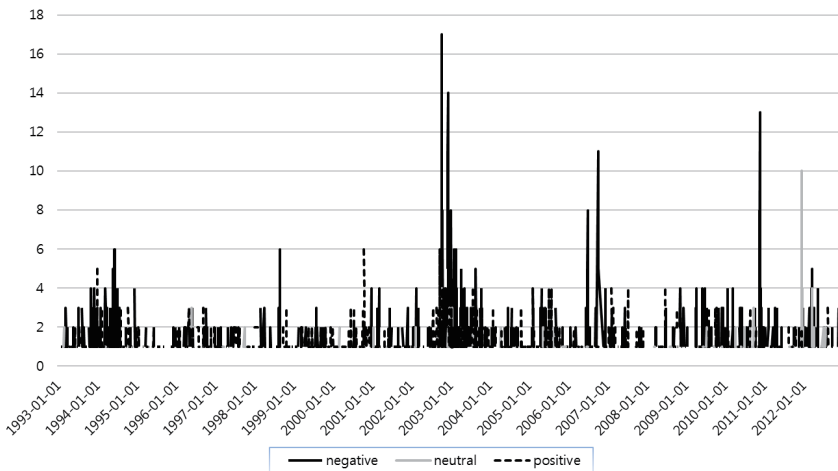

FIG. 2.—Distribution of News Sentiments

TABLE 1
Summary Statistics of 8 Korean Companies listed in NYSE

Company	Listed day	Obs.	Mean(%)	SD(%)	Min.(%)	Max.(%)
POSCO	1994.10.14	4579	0.117	3.019	-21.041	31.080
SK Telecom	1996.06.27	4149	0.069	2.963	-16.513	21.698
KT Corporation	1999.05.26	3415	0.031	2.514	-18.072	20.799
KB Financial Group	2001.11.14	2799	0.059	3.027	-19.097	23.969
Shinhan Financial Group	2003.09.17	2338	0.097	3.079	-15.132	22.123
Korea Electric Power Corporation	2003.10.01	2328	0.020	2.445	-14.631	23.228
Woori Financial Group	2003.10.01	2328	0.088	3.498	-20.215	33.490
LG Display	2004.07.22	2126	0.084	3.280	-17.327	37.906

using nine folds – excluding one – to evaluate the tenth's accuracy. This process was repeated 10 times and we used the mean of the 10 operations to eliminate contingency (Kim et al. 2013).

As a result, we were able to confirm that the *Maximum Entropy* classifier had the best performance. We therefore applied it to the remaining articles to be classified (the test data), resulting in 1,549 negative articles, 230 neutral articles, and 705 positive articles. <Figure 2> shows the distribution of the total number of articles from 1993 to 2012 classified by both humans and computers.

Our dependent variables were the stock prices of Korean companies listed in *NYSE*. As shown in <Table 1>, eight Korean firms are listed in the *NYSE*, starting from POSCO, enlisted on October 14 1993, to LG Display, enlisted on July 22 2004 (POSCO, SK Telecom, KT Corporation, Korea Electric Power Corporation, Woori Finance Holdings, Shinhan Financial Group Company Limited, KB Financial Group and LG Display). We downloaded the stock price data for 1993-2012 from *Yahoo! Finance*. The stock prices of parent companies in Korea were downloaded from *Data Guide (http://dataguide.co.kr)*.
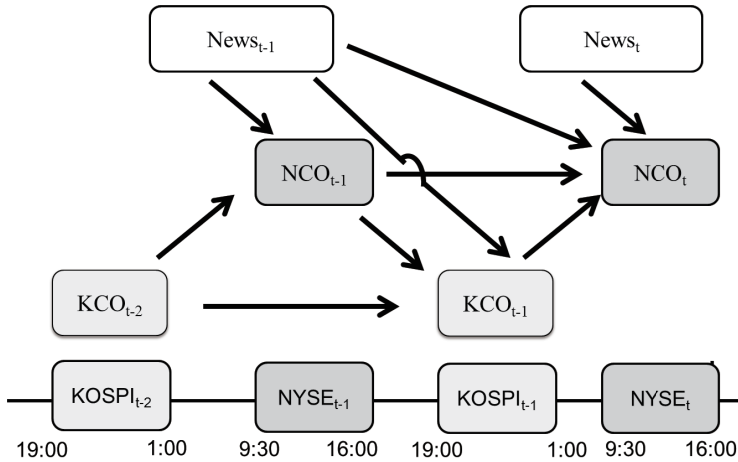
FIG. 3.—Path Analysis Model of North Korean News

This figure depicts the theoretical paths of impact that news frequencies in the U.S. have on stock price changes of *KOSPI* and *NYSE*. The path directions are supported as time sequences based on the bottom time line. For example, if a nuclear test$_{t-2}$ by North Korea occurred, for a start, the shock would be absorbed in stock price changes of Korean firms listed in *KOSPI* (KCO_{t-2}). Then, the U.S. news ($News_{t-1}$) informed the U.S. investors of the nuclear test, and the influences and direction of the news could be recognized by the U.S. investors. Considering the investment atmosphere formed by the news sentiment ($News_{t-1}$) and the direction of Korean firms; stock price changes ($KCOt_{-2}$), the U.S. investors made their decisions and cause to stock price changes for Korean firms listed in *NYSE* (NCO_{t-1}). In addition, to control the auto-correlation of stock markets, we have to consider the volatility on the previous day as a control variable. Therefore, to cover the in/direct news about the event and the characteristics of stock markets, the reference point of this model should be the day after the events of North Korea.

NOTE.—+ $p < 0.10$, * $p < 0.05$, * $p < 0.01$, *** $p < 0.001$; News$_{t=}$ frequency of news on the day; News$_{t-1}$ = frequency of news on the previous day; NCO$_{t=}$ stock price changes of Korean firms listed in *NYSE* on the day; NCO$_{t-1}$ = stock price changes of Korean firms listed in *NYSE* on the previous day; KCO$_{t-1=}$ stock price changes of Korean firms listed in *KOSPI* on the previous day; KCO$_{t-2=}$ stock price changes of Korean firms listed in *KOSPI* on the day before yesterday

Path Analysis Models

It is hard to deal with the stock volatility because of high cross-auto-correlations in longitudinal data (Nelson, 1991; Campbell, Grossman, and Wang 1993). To show the simple causal relationship between the time-

varying data and the direct and indirect effects of the news, we designed a path analysis model as in Figure 3 (Wolfle 2003). Figure 3 depicts the theoretical paths of impact of news frequencies on the U.S. on stock price changes of *KOSPI* and *NYSE*. The path directions are supported as time sequences based on the bottom time line of Figure 3. To cover both the indirect and direct effects of the news about the event and the characteristics of stock markets, the reference point of this model should be the day after the events of North Korea. This path model is adequate to show specific relationships among time lagged variables (Byun 2008). We used Stata 11.0 to calculate the path coefficients and *p*-values.

The daily change of stock price is our focal dependent variable, where our independent variable is the frequency of negative news (or positive news). Regarding uncertainty in the market, researchers have shown a strong correlation between the volume and volatility in the stock market (Kim 2011; Tetlock 2007; Antweiler and Frank 2004). In order to test simple causal relationships, we picked only the movement of stock prices as our dependent variable, ignoring the volume of transactions. We calculated the price change index as follows:

$$\text{stock price change index} = \frac{\text{today's last stock price} - \text{yesterday's last stock price}}{\text{yesterday's last stock price}}.$$

The change index was normally distributed ranging from -49.544% to 33.491%. The volatility of the eight parent firms in the Korean stock market, *KOSPI*, was also normally distributed between -15% and 15%. This range is determined by the Korean "side-break" law.

To observe the impact of political risk on the stock price, we decided to follow the movement of stock prices and news sentiment for two days after the event that generated news. It enabled us to control for autocorrelations between daily stock prices and to measure the *net effect* of the news sentiment. We chose two exogenous variables: the news frequency of the day ($News_t$) and the news frequency of the previous day ($News_{t-1}$). Stock prices fluctuate according to fundamental market values, but sometimes they also reflect external market factors instantly. Isolating *net effect of news sentiment* on changing prices forced us to control the stock prices of the previous day in Korea and U.S. Yet, analyzing the movement of stock prices in two markets was not an easy task because of the time differences in the U.S. and Korea.

To explain the causal flows across time, we introduce the following illustration of event sequences on the *Eastern Time* scale at the bottom of

<Figure 3>. The causal chains of our path model are displayed above this time scale. When a nuclear bomb test was detected in North Korea, its impact might be felt almost immediately in the Korean stock market (KCO_{t-2}), if Korean investors respond to the shock. Yet, this event occurs during the night in *Eastern Time* because of the time zone difference. Therefore, the U.S. news begins to appear the following day (we call this variable, i.e., number of news articles, $News_{t-1}$). If investors in *NYSE* respond to these news sentiments, then their impacts will be reflected on the stock price change index (NCO_{t-1}). Stock prices in Korean market (KCO_{t-1}) may also be affected by the U.S. news coverage of the previous day ($News_{t-1}$), which in turn influences the stock prices of subsidiary companies in the U.S. (NCO_t) after eleven hours.

This path model allows us to measure the direct effect as well as the indirect effects of U.S. news sentiment mediated via Korean stock market after controlling auto-correlations. The stock prices of parent companies in Korea (*KOSPI*) and subsidiary companies listed in the U.S. (*NYSE*) are often synchronized as global economy has become more interdependent (Forbes and Rigobon 2002; Chen et al. 2002; Chung 2005; Mun and Brooks 2012). We can measure the strength of synchronization.

Findings

To compare the movement of stock prices in ordinary days with those when North Korean risk emerges or disappears, we run the path model including all days, i.e. days with no related news. This model of course does not have news variables, but it can be used as a base model. The results shown in <Figure 4> reveals two important things. First, synchronization of *NYSE* with *KOSPI* is positive and significant ($p = 0.581$ and 0.516 respectively), and inter-market lagged effect is much stronger than within-market lagged effect. More than 58% of variance of daily prices in *NYSE* is explained by the price of Korean market in the previous day. The lagged effect of *NYSE* (NCO_{t-1}) on Korean market (KCO_{t-1}) is much weaker with a partial correlation of 0.218.

Now we turn to cases when the U.S. news media covered North Korean news more than three times. Analyzing the 208 events on which there appeared more than three negative articles, <Figure 5> shows only the statistically significant causal paths of negative news. The most striking finding is that the two markets are highly synchronized with path coefficient, $p = 0.705$ and p = 0.693 respectively, for two consecutive days. When the risk
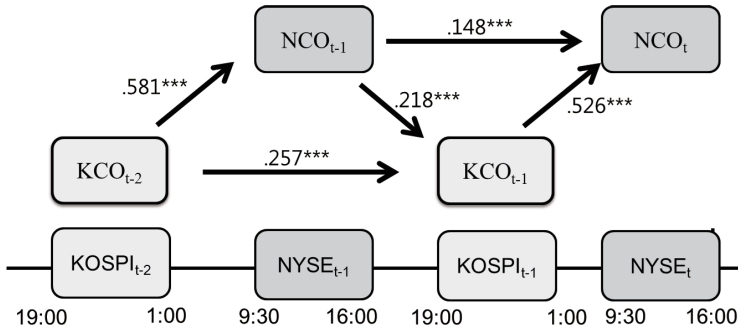
Fig. 4.—Path Analysis as a Base Model

The data for this figure comes from *KOSPI, NYSE*. The path analysis is based on all the data from January 1 1993 to December 31 2012.

NOTE.—+ $p < 0.10$, * $p < 0.05$, * $p < 0.01$, *** $p < 0.001$; NCO_t = stock price changes of Korean firms listed in *NYSE* on the day; NCO_{t-1} = stock price changes of Korean firms listed in *NYSE* on the previous day; KCO_{t-1} = stock price changes of Korean firms listed in KOSPI on the previous day; KCO_{t-2} = stock price changes of Korean firms listed in KOSPI on the day before yesterday

on Korean peninsula is enhanced, American stock prices moves closely with those of Korea, much closer than ordinary days. As expected, the negative news of the previous day ($News_{t-1}$) lowers the stock price of the previous day (NCO_{t-1}) with a path coefficient -0.11. It is surprising to note, however, that the effect of NCO_{t-1} on NCO_t is negative. It means that the impact of negative news on stock prices *after one day* turns out to be positive (-0.11 * -0.456 = 0.05). Why do the negative news lower the stock price immediately, but after one day, they make stock prices go up? To understand this counter intuitive result, we must inspect another causal chain operating through the Korean market.

Although previous researchers have focused on the effect of the North Korean risk on the Korean economy, little attention has been paid to this particular path, i.e., to the mechanism by which negative foreign news about North Korea can influence the general reputation of the Korean stock market and thereby change foreign investors' evaluation of the Korean firms listed in the foreign stock market. When geo-political risk in the Korea peninsula was reported by the U.S. news media, the stock prices of parent firms in Korea (KCO_{t-1}) plunged (path coefficient, $p = -0.237$), even after controlling for the previous day's stock price (KCO_{t-2}) that already had absorbed the shock firsthand. Stock prices of subsidiary companies in *NYSE* synchronized ($p =$
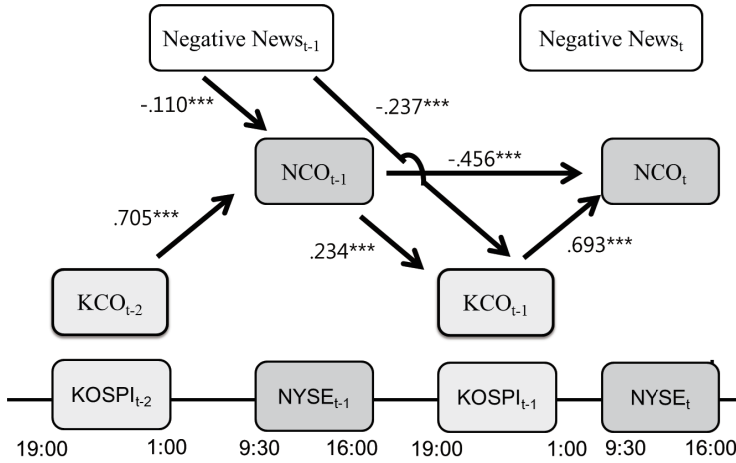
Fig. 5.—Path Analysis of Negative News

This figure data come from *KOSPI, NYSE*, and the negative U.S. news frequencies (more than three articles) about North Korea. The path analysis is based on 208 observations from January 1 1993 to December 31 2012.

NOTE.—+ $p < 0.10$, * $p < 0.05$, * $p < 0.01$, *** $p < 0.001$; Negative News$_t$ = frequency of negative news on the day; News$_{t-1}$ = frequency of negative news on the previous day; NCO$_t$ = stock price changes of Korean firms listed in *NYSE* on the day; NCO$_{t-1}$ = stock price changes of Korean firms listed in *NYSE* on the previous day; KCO$_{t-1}$ = stock price changes of Korean firms listed in KOSPI on the previous day; KCO$_{t-2}$ = stock price changes of Korean firms listed in KOSPI on the day before yesterday

0.693) with those in Korea. In other words, negative news lowered the U.S. stock prices drastically via the Korean market (-0.237 * 0.693 = -0.164). Therefore, the counter intuitive result mentioned above can be interpreted as a compensating mechanism; negative news lowered stock prices too much through Korean market, thus U.S. market had countervailing effect to restore fallen prices.

Now we turn to investigate the situation when the North Korean risk decreases. In case of positive news, causal effects are rather straightforward as in <Figure 6>. First, synchronization is much weaker than in case of negative news. When the North Korean risk is reduced, the partial correlation between Korean and *NYSE* stock prices is 0.547 (compared with 0.705 in case of negative news). The day after the favorable event occurred, it was 0.479 (compared with 0.693 in case of negative news). It is also interesting to note that the positive news did not have any mediating effects through the Korean
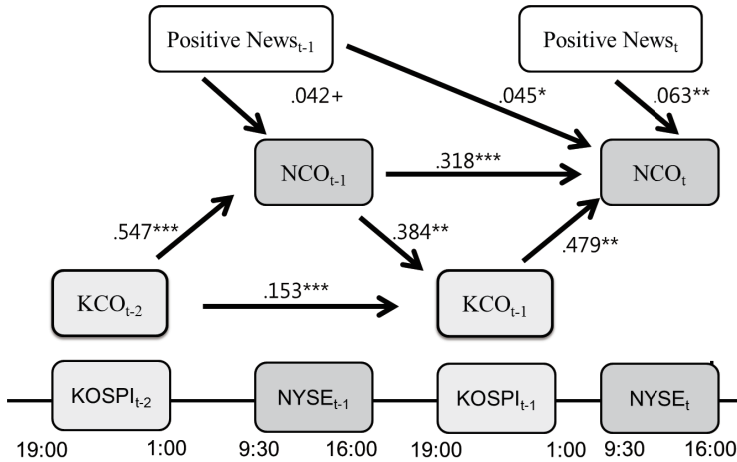
Fig. 6.—Path Analysis of Positive News

This figure data come from *KOSPI*, *NYSE*, and the positive U.S. news frequencies about North Korea. The path analysis is based on 1,175 observations from January 1 1993 to December 31 2012.

Note.—+ $p < 0.10$, * $p < 0.05$, * $p < 0.01$, *** $p < 0.001$; Positive News$_t$ = frequency of positive news on the day; Positive News$_{t-1}$ = frequency of positive news on the previous day; NCO$_t$ = stock price changes of Korean firms listed in *NYSE* on the day; NCO$_{t-1}$ = stock price changes of Korean firms listed in *NYSE* on the previous day; KCO$_{t-1}$ = stock price changes of Korean firms listed in *KOSPI* on the previous day; KCO$_{t-2}$ = stock price changes of Korean firms listed in *KOSPI* on the day before yesterday

market at all. Once favorable news appeared, their very slight effect lasts for two days ($p = 0.042$ and 0.045 respectively). This result may reflect that additional positive information and interpretations are added to the favorable news as time goes by. The positive news appearing in the following day (Positive News$_t$) also affect positively (NCO$_t$, $p = 0.063$), even after controlling the price of previous day.

In sum, firstly, the effects of positive news were small but statistically significant. Secondly, Korean stock prices are not influenced by U.S. news. Thirdly, synchronization of prices between Korea and U.S. is much weaker than that in case of negative news.

Information asymmetry might be the cause of higher synchronization when risk is enhanced. U.S. Investors who invest into foreign firms listed in *NYSE* may not have enough information about Korean firms. With the continuous expansion of individual investors in the stock market, these

investors' information might be limited as compared to that of institutional investors. Furthermore, they might be seeking short-term benefits, which are more sensitive to external shocks than those sought by institutional investors. Compared with domestic companies, information about foreign firms might be not only limited but also asymmetric as compared to that available to investors in the foreign firm's original region. In this situation, the previous day's shock on the stock price in that region could be helpful to decide the firm's value in *NYSE* as an essential signal. The negative news sentiments might be a good signal to devalue these firms when information is limited.

Conclusion

In this paper, we attempted to examine the existence and degree of the Korea discount phenomenon. We focused on the stock market to see if geo-political threats posed by North Korea influence the stock prices of Korean companies enlisted in *NYSE*. To do so, we collected a large number of news articles from five major U.S. newspapers related with North Korea, and then classified the U.S. news sentiments. The classification of news sentiment made us to develop a semi-supervised machine learning method, which classified negative, neutral and positive news.

We found that the news related to North Korea had direct and indirect effect on the stock prices on the *NYSE*. This new finding suggests that Korea discount exists in a global stock market. News sentiments have strong direct effects on the *KOSPI* but weak direct effects on the *NYSE*. Synchronization of stock prices in Korea and U.S. was much stronger when the risk culminated in the Korean peninsula. Yet it is interesting to note that the impact of negative news sentiments on the *NYSE* operated indirectly via the Korean stock market. We conclude that investors in the U.S. have recognized North Korea's geo-political risk to Korean firms over the twenty years we investigated.

References

Antweiler, Werner and Murray Z. Frank. 2004. "Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards." *The Journal of Finance*

59(3): 1259-1294.

Ahn, Hee-Joon, Seung-Pyo Jeon, and Jong-Bum Chay. 2010. "The Effects of the News Related to the North-South Korean Relationship on the Korean Stock Markets." *Journal of Korean Economic Analysis* 16(2): 199-238.

Ahn, Sung-Won and Sung-Bae Cho. 2010. "Stock Prediction Using News Text Mining and Time Series Analysis." *Proceedings of Korea Computer Congress* 37(2): 364-369.

Arin, K. Peren, Davide Ciferri, and Nicola Spagnolo. 2008. "The Price of Terror: The Effects of Terrorism on Stock Market Returns and Volatility." *Economics Letters* 101(3): 164-167.

Ban, Hyun, Nam-Ie Kim, and Hye-Jeong Noh. 2010. "Comparative Study of Coverage toward Korean Economy by Foreign and Domestic Press." *Korean Journal of Journalism & Communication Studies* 54(5): 397-422.

Bilson, Chris, Tim Brailsford, and Vince Hooper. 2001. "Selecting macroeconomic variables as explanatory factors of emerging stock market returns." *Pacific-Basin Finance Journal* 9(4): 401-426.

Bollen, Johan, Huina Mao, and Xiaojun Zeng. 2011. "Twitter Mood Predicts the Stock Market." *Journal of Computational Science* 2(1): 1-8.

Byun, Se-Il. 2008. "Path Analysis." *Planning and Policy* 59-67.

Campbell, John Y., Sanford J. Grossman, and Jiang Wang. 1993. "Trading Volume and Serial Correlation in Stock Returns." *The Quarterly Journal of Economics* 108(4): 905-939.

Chen, Gong-meng, Michael Firth, and Oliver Meng Rui. 2002. "Stock market linkages: evidence from Latin America." *Journal of Banking & Finance* 26(6): 1113-1141.

Chung, Young Chul. 2005. "Political Economy of the US Economic Sanctions against North Korea: Past, Present and Future." *Development and Society* 34(2): 217-238.

Eldor, Rafi and Rafi Melnick. 2004. "Financial markets and terrorism." *European Journal of Political Economy* 20(2): 367-386.

Fama, Eugene. 1970. "Efficient Capital Markets: A Review of Theory and Empirical Work." *Journal of Finance* 25(2): 383-417.

Feldman, Ronen, Benjamin Rosenfeld, Roy Bar-Haim, and Moshe Fresko. 2011. "The Stock Sonar—Sentiment Analysis of Stocks Based on a Hybrid Approach." In *Twenty-Third IAAI Conference.*

Forbes, Kristin and Roberto Rigobon. 2002. "No Contagion, Only Interdependence: Measuring Stock Market Comovements." *The journal of Finance* 57(5): 2223-2261.

Guidolin, Massimo and Eliana La Ferrara. 2010. "The Economic Effects of Violent Conflict: Evidence from Asset Market Reactions." *Journal of Peace Research* 47(6): 671-684.

Groß-Klußmann, Axel and Nikolaus Hautsch. 2011. "When Machines Read the News: Using Automated Text Analytics to Quantify High Frequency News-

Implied Market Reactions." *Journal of Empirical Finance* 18(2): 321-340.

Hwang, Sun-Ung and Hyun-Suk Kim. 2004. "Firm Valuation and Korea Discount". *Korean Journal of Financial Studies* 33(1): 139-170.

Kim, Chi-Wook. 2011. "Inter-Korean Relations and "Korea Discount": An Analysis of Foreign Investors' Stock Trading." *Journal of Peace and Unification Studies* 3(1): 219-252.

Kim, Meen-Chul, Kyu-Seung Shim, Nam-Ci Han, Yee-Un Kim, and Min Song. 2013. "Automatic Classification of Malicious Usage on Twitter." *Journal of the Korean Society for Library and Information Scienc* 47(1): 269-286.

Kim, Yoo-Sin, Nam-Gyu Kim, and Seong-Ryoul Jeong. 2012. "Stock-Index Invest Model Using News Big Data Opinion Mining." *Journal of Intelligent Information Systems* 18(2): 143-156.

McCallum, Andrew Kachites. 2002. *MALLET: A Machine Learning for Language Toolkit* (http://mallet.cs.umass.edu).

Mittermayer, Marc-Andre. 2004. "Forecasting Intraday Stock Price Trends with Text Mining Techniques." In *Proceedings of the Proceedings of the 37th Annual Hawaii International Conference on System Sciences. (HICSS'04)-Track 3-Volume 3* (pp. 30064-2). IEEE Computer Society.

Mehdian, Seyed, Tevfik Nas, and Mark J. Perry. 2008. "An Examination of Investor Reaction to Unexpected Political and Economic Events in Turkey." *Global Finance Journal* 18(3): 337-350.

Moon, Seung-Dong. 2007. "A Study on Risk Premium of Korea Division: Korea Discount." Doctorate Thesis, Dongguk University.

Mun, Melissa and Robert Brooks. 2012. "The Roles of News and Volatility in Stock Market Correlations during the Global Financial Crisis." *Emerging Markets Review* 13(1): 1-7.

Nam, Sung-Wook. 2004. "North Korean Variable`s Influence on the Change of the Index of Stocks and Policy Suggestions." *Journal of Northeast Asian Economic Studies* 16(1): 73-106.

Nelson, Daniel B. 1991. "Conditional Heteroskedasticity in Asset Returns: A New Approach." *Econometrica: Journal of the Econometric Society* 59(2): 347-370.

Pang, Bo and Lillian Lee. 2008. "Opinion Mining and Sentiment Analysis." *Foundations and Trends in Information Retrieval* 2(1-2): 1-135.

Park, Hyun-Soo. 2005. "Korea Discount and Corporate Governance Structure." *SERI Issue Paper.* SERI.

Park, Se-Jung, Yon-Soo Lim, Seven Sams, Sang-Me Nam, and Han-Woo Park. 2011. "Networked Politics on Cyworld: The Text and Sentiment of Korean Political Profiles." *Social Science Computer Review* 29(3): 288-299.

Pyun, Ju-Hyun and In Huh. 2014. "Responses of South Korea Financial Market to North Korea Nuclear Risk." *The Journal of Northeast Asian Economic Studies* 26(4): 175-196.

Shiller, Robert J. 2003. "From Efficient Markets Theory to Behavioral Finance."

Journal of Economic Perspectives 17(1): 83-104.

Song, Min, Hwan-Jo Yu, and Wook-Shin Han. 2011. "Combining Active Learning and Semi-Supervised Learning Techniques to Extract Protein Interaction Sentences." *BMC Bioinformatics, 12* (Suppl 12), S4.

Tetlock, Paul C. 2007. "Giving Content to Investor Sentiment: The Role of Media in the Stock Market." *The Journal of Finance* 62(3): 1139-1168.

_____. 2010. "Does Public Financial News Resolve Asymmetric Information?" *Review of Financial Studies* 23(9): 3520-3557.

_____. 2011. "All the News That's Fit to Reprint: Do Investors React to Stale Information?" *Review of Financial Studies* 24(5): 1481-1512.

Tetlock, Paul C., Maytal Saar-Tsechansky, and Sofus Macskassy. 2008. "More Than Words: Quantifying Language to Measure Firms' Fundamentals." *The Journal of Finance* 63(3): 1437-1467

Tumasjan, Andranik, Timm O. Sprenger, Philipp G. Sandner, and Isabell M. Welpe. 2011. "Election Forecasts With Twitter: How 140 Characters Reflect the Political Landscape." *Social Science Computer Review* 29(4): 402-418.

Wolfle, Lee M. 2003. "The Introduction of Path Analysis to the Social Sciences, and Some Emergent Themes: An Annotated Bibliography." *Structural Equation Modeling* 10(1): 1-34.

Yun, Sun-Jin, Dowan Ku, Nyun-Bae Park, and Joon Han. 2012. "A Comparative Analysis of South Korean Newspaper Coverage on Climate Change." *Development and Society* 41(2): 201-228.

Zach, T. 2003. "Political Events and the Stock Market: Evidence from Israel." *International Journal of Business* 8(3): 243-266.

YUNJUNG PAK is an MA student of sociology at Yonsei University. His research utilizes computational innovations in text analysis and social network analysis to understand social change and its influences on the real world. *Address*: Social Development Institute, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 120-749. [*E-mail*: hermetic@yonsei.ac.kr]

YOUNG-JIN KIM is a PhD in sociology at Yonsei University. Social networks and new digital media are the two threads that weave the fabric of his research interests. His research interests lie in new media, media convergence, scientific communication, and network analysis. Currently, his primary research focuses on the temporal and spatial constraints on digital behaviors using mobile log data. *Address*: Social Development Institute, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 120-749. [*E-mail*: yj5412@yonsei.ac.kr]

MIN SONG, PhD, is an associate professor of library and information science at

Yonsei University. His recent work in Text Mining addresses automatic database selection, entity and relation extraction, high speed document filtering, algorithms that learn a person's information needs from experience, automatic analysis of gathered information. He is currently interested in applying Text Mining algorithms to Bioinformatics and Social Media. *Address*: Department of Library and Information Sciences, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 120-749, Korea. [*E-mail*: min.song@yonsei.ac.kr]

YONG-HAK KIM, PhD, is a professor of sociology at Yonsei University. His research interests lie in the area of information society, the social network theories, comparative sociology, and methodology of the social sciences. He has published many books including Social Structure and Action in Social Theories, Methodological Issues of Comparative Sociology, and the Theories of Social Network. *Address*: Department of Sociology, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 120-749, Korea. [*E-mail*: yhakim@yonsei.ac.kr]

Appendix

The independent variables in this study were developed from a corpus of media coverage using a semi-supervised learning method to classify U.S. news sentiments. The semi-supervised learning method was defined as a text-mining method that distinguished the whole data into training data and test data, categorized the sentiments of the training data through human coders, and then classified automatically the test data through machine classifiers learning the training data's categorization. The training data in this study consists of 2,000 articles randomly sampled from the whole 4,484 articles. The human coders categorized the training data's sentiments in that

COPYRIGHT	THE WALL STREET JOURNAL 11 October 2006
LENGTH	1449
HEADLINE	North Korea's neighbors fear possibility of collapse --- Beijing, Seoul worry harsh sanctions risk upsetting balance
AUTHOR	Gordon Fairclough in Shanghai and Evan Ramstad
REGION	SEOUL, SOUTH KOREA
BODY	Beneath the condemnations and diplomatic maneuvering over North Korea's declared atomic test lies a fundamental fear: That the fall of dictator Kim Jong Il may be more dangerous and destabilizing than leaving him in charge of a nuclear-armed state. As the United Nations Security Council weighs its response to Pyongyang's detonation of a powerful but unspecified explosive Monday morning, member states are balancing their desire to act forcefully to prevent the spread of atomic weapons with the possibility that harsh sanctions could help hasten the end of Mr. Kim's Stalinist government. The prospect of a North Korean regime change cheers many in Washington, but for the country's neighbors, especially China and South Korea, the risks are deeply unsettling and more immediate than the eventual threat of proliferation -- providing a powerful incentive for them to maintain the status quo. …

FIG. 7.—Sample Press Release

TABLE 2
20 Highest-Ranked Terms Per News Sentiments

Negative	Neutral	Positive
North Korea	North Korea	North Korea
the U.S.	the U.S.	the U.S.
nuclear	nuclear	talk
say	China	nuclear
talk	talk	aid
missile	leader	deal
new	new	say
UN	say	China
China	world	agree
Japan	Kim	Pyongyang
Pyongyang	Iran	offer
arm	Seoul	Japan
sanction	time	open
warn	urge	food
aid	Pyongyang	plan
Seoul	Bush	Seoul
threat	end	World
plan	visit	visit
test	deal	envoy
Bush	no	new

they focused on whether the headline includes negative or positive words, and what they feel in general after reading the body. When an article only described an event or held back from giving an opinion about it, the human coders regarded the article as a neutral sentiment. As previously noted, the training data consisted of the data that all three human coders agreed on the classification of (55.3%, or 1,105 articles) and the data two-coders agreed upon (44.3%, or 886 articles). The words patterns of training data provided basic features for making computational classifiers.

Machine learning algorithm regards a document as vector values of numerous words features. As shown in <Figure 7>, an article of the Wall Street Journal was randomly selected as a training data assigned to be learned by the machine classifier. Three human coders classified this article as a negative tone article including the negative words in this article such as "fear", "harsh", and "sanction". Since the machine classifier we used learned these words as elements of negative sentiments, the machine classifier could categorize unlabeled articles in the test data set according to these word

vectors. <Table 2> showed the 20 highest ranked words in each sentiment: the positive sentiment contained hopeful or peaceful words such as "talk", "agree", "offer", "open" or "envoy"; the neutral sentiment contained factual or comparative words such "leader", "talk" or "Iran"; and, finally, the negative sentiment contained oppressive words such as "missile", "arm", "sanction", "warn", "threat". As shown in <Table 2>, it is interesting that there were a few words overlapping with each sentiment such as "North Korea", "the U.S", and "nuclear". These seemingly paradoxical results reflected the relational characteristics of semi-supervised learning algorithm in that the machine classifier learned what the human coders classified. In other words, the machine classifier classified the news sentiment, depending on not only the words frequency, but also the co-occurrence patterns of the words.

To classify the extra test data (2,484 articles), we used four different classifiers provided in *Mallet,* written in Java programming language (McCallum 2002): *Naïve Bayesian, Support Vector Machine, Decision Tree, and Maximum Entropy*. In order to evaluate them with precision, we used a 10-fold cross-evaluation, dividing 2,000 by 10 and using nine folds – excluding one – to evaluate the tenth's accuracy. This process was repeated 10 times and we used the mean of the 10 operations to eliminate contingency (Kim et al. 2013). As a result, we were able to confirm that the *Maximum Entropy* classifier had the best performance. We therefore applied it to the remaining articles to be classified (the test data) resulting in 1,549 negative articles, 230 neutral articles, and 705 positive articles. In sum, the semi-supervised learning methods we used could automatically classify the unlabeled news regarding North Korea into positive, neutral, negative sentiments, based on not only the statistical distribution of the words, but also the social context of those words that human coders perceived.