

# 포털 검색 강도가 주가 급락에 미치는 영향에 관한 연구

## The Effect of Portal Search Intensity on Stock Price Crash

김민수(Min-Su Kim)\*, 권혁준(Hyuk-Jun Kwon)\*\*

### 초 록

최근 주식의 수익률과 거래량을 설명하는 요인 중 하나로 투자자들의 관심이나 주식관련 정보 전파의 효율성 등이 중요하게 인식되고 있다. 또한 기업관련 정보가 투자자들에게 투명하게 전파되지 않을 때 기업 주가의 급락(crash) 위험을 증가시킨다는 연구 결과들이 축적되고 있다. 본 연구에서는 네이버 트렌드를 이용하여 포털에서의 검색 강도가 증가하는 것이 주식 수익률의 급락에 어떤 영향을 미치는지를 분석하였다. 다양한 주가급락 위험의 측정변수와 검색 강도 측정치를 이용한 분석에서 포털 검색강도가 상대적으로 높은 기업-연도에서 주가 급락의 위험이 감소하는 것으로 관찰되었다. 이러한 결과는 기업 관련 정보 전파가 투자자들에게 효율적으로 이루어지지 않을 때 미래의 주가급락을 초래한다는 논의와 일치하는 결과이다. 또한 이러한 결과는 분석에 발생가능한 내생성을 통제한 후에도 유의하게 성립하는 것으로 관찰되었다.

### ABSTRACT

Recent studies focus on the role of investor attention and transparency in stock-related information in explaining stock return and trading volume. Moreover, recent literatures predict that firm opacity will increase the likelihood of future stock price crashes. In this paper, we investigate, using Naver Trend, the relation between portal search intensity and stock price crash. Using various alternative measures of stock price crash risk and search intensity, we demonstrate that stocks with larger volume of portal search are less likely to experience stock price crashes. These results are consistent with our hypothesis that accumulated firm opacity cause future stock price crash. Finally, our results still hold even after we control for the potential effect of endogeneity in the regression specifications.

**키워드** : 포털 검색강도, 검색량 지수, 주가급락 위험, 음의 조건부 왜도, 투자자 관심  
Portal Search Intensity, Search Volume Index, Stock Price Crash Risk, Negative Skewness, Investor Attention

---

본 연구는 순천향대학교 학술연구비 지원으로 수행하였음.

\* First Author, Department of Economics and Finance, Soonchunhyang University(equili@sch.ac.kr)

\*\* Corresponding Author, Department of IT Finance Management, Soonchunhyang University  
(gloryever@sch.ac.kr)

Received: 2017-04-14, Review completed: 2017-05-02, Accepted: 2017-05-17

## 1. 서 론

전통적인 재무이론에서 주가의 급락(crash) 현상은 시장전반에 걸친 주가급락 현상 또는 시장미시적인 이론 등에 의해 설명되었으나, 최근에는 기업과 관련된 정보의 전파(diffusion of information)가 얼마나 효율적으로 일어나는가의 여부에 의해 설명하는 연구들이 축적되고 있다[22, 14, 13].

정보의 전파가 지연되는 기업들에서는 특히 부정적인 정보들이 주가에 반영되지 않게 되며, 이 경우 주가가 기업의 내재가치보다 과도하게 높게 평가되는 경우가 발생하게 된다. 그러나 이렇게 축적된 부정적 정보가 임계점에 도달하여 시장에 순식간에 유출되면, 과대평가된 주가가 조정되면서 주가의 급락이 발생하게 된다는 것이다[14]. 또한, 기업의 정보가 투자자들에게 효율적으로 제공되지 못하면, 이사회나 투자자들에 등의 감시기능도 제대로 작동하지 못하여 적절한 경영상 조치들이 실행되기 어려우며, 이러한 경영상의 비효율성은 주가의 급락이 발생할 때까지 계속될 것이다[7].

본 논문에서는 포털(portal) 사이트의 검색 강도를 이용하여 해당 기업에 대한 투자자의 관심(attention)의 정도를 측정하고, 이러한 검색 강도가 미래의 주가급락 위험에 어떤 영향을 미치는지를 분석한다. 특히 한국의 경우, 네이버가 포털 검색 시장에서 수년간 지배적인 시장점유율을 보여 왔으며, 주가, 거래량 및 회계 정보 등 다양한 금융정보를 투자자들에게 제공하고 있기 때문에, 네이버의 검색 강도가 투자자들의 관심을 측정하는 측정변수로 적절하다고 판단하였다. 네이버에서는 네이버 트랜

드 서비스를 통해, 2007년 이후 해당 단어의 주간 검색 강도(weekly search intensity)를 표준화하여 보고하고 있다.

기업-연도 수준에서 주가급락 위험(crash risk)을 분석하기 위해, 본 연구에서는 세 종류의 변수를 이용하였다. Chen, Hong, and Stein [8]에서는 기업의 주가급락 위험을 왼쪽으로 꼬리가 긴 분포, 즉 음의 왜도(negative skewness)를 갖는 분포를 갖는 것으로 정의하고, 이를 측정하기 위한 두 가지의 변수를 제시하였다. 본 연구에서도 이들의 방법론에 따라 (1) 주간 기업 고유수익률(firm-specific weekly return)의 음의 조건부 왜도, (2) 주간 수익률의 하락-상승 변동성(down-to-up volatility) 비율을 통해 주가급락 위험을 측정하였다. 또한, Jin and Myers[14], Hutton et al.[13], Kim et al.[17]에서 제시된 방법에 따라, (3) 표준정규 분포에서 하위 0.1% 이하의 수익률이 실현되는지 여부를 통해 주가급락을 측정하였다.

2007년부터 2014년까지 코스피(KOSPI) 시장에 상장된 제조업 기업을 대상으로 한 분석을 통해, 본 연구에서는 검색강도가 높아질수록 해당 연도의 주가급락 위험이 감소함을 관찰하였다. 이러한 실증연구 결과는 투자자들의 관심과 기업 정보의 투명성이 높아질수록 기업 정보의 전파속도가 빨라지고 기업 내부에 부정적 정보를 축적할 가능성이 낮아지게 되어, 미래의 주가급락 위험을 낮춘다는 가설과 부합하는 것이다. 이러한 결과는 주가급락과 검색 강도의 여러 측정변수에 대해 공통적으로 관찰되며, Chen et al.[8] 등에서 제시된 모든 통제변수들을 고려한 후에도 역시 성립한다. 한편, 포털에서의 검색 강도와 주가 급락의 관계에는 반대 방향의 인과관계(reverse cau-

sality), 즉 주가가 급락하여 검색 강도를 증가시킬 가능성이 존재한다. 이러한 인과관계로부터 발생하는 내생성(endogeneity)을 고려한 실증분석에서도, 검색강도는 주가급락 위험을 유의하게 감소시키는 것이 관찰되었다.

이후 본 논문의 구성은 다음과 같다. 제2장에서는 포털 검색과 주가급락 위험에 관한 선행연구들을 요약하고 연구가설을 제시한다. 제3장에서는 분석 자료와 변수의 구성방법을 설명하고, 제4장에서는 실증분석 결과를 보고한다. 제5장에서는 실증분석 결과의 강건성을 검증하며, 제6장에서는 논문의 결론과 이후 연구 방향을 제시한다.

## 2. 연구 설계

### 2.1 관련 선행 연구

전통적인 재무이론은 효율적 시장가설(Efficient market hypothesis)에 근거하여, 이용 가능한 모든 관련 정보들이 시장가격에 반영되어 있다고 가정한다. 그러나 효율적 시장가설과는 달리, 투자자들의 관심이 주식 수익률이나 거래량 등에 미치는 효과가 존재한다는 이론적, 실증적 연구들이 진행되고 있다. 먼저, Merton[21] 등에 의해 제시된 ‘투자자 인지 가설(investor recognition hypothesis)’에서는 기업의 투명성(visibility)의 증가가 투자자에게 새로운 정보를 제공하여 주식 거래에 영향을 미친다고 주장한다. 한편, 행태이론에 기초한 최근 연구들에서는 공공의 관심(public attention)만으로도 주가가격을 변동시키기에 충분한 요소가 된다고 주장하는데, 이러한 논의를

‘가격압력 가설(price pressure hypothesis)’ 또는 ‘관심 가설(attention hypothesis)’이라 한다 [4]. 과거의 연구들에서는 투자자 공공의 관심을 측정할 방법이 부족하여 거래량의 변동 등으로 간접적으로 측정하였으나, 최근 주식 관련 인터넷 사이트, 각종 포털의 검색, 미디어에서의 보도 등에 대한 자료가 축적되고 데이터 마이닝의 발전으로 인해 투자자들의 관심을 직접적으로 측정하고 이러한 측정값들이 수익률이나 거래행태에 어떤 영향을 미치는지 분석하기 시작하였다. Barber and Odean[4]은 개인투자자들이 뉴스에 노출된 주식을 순매수(net-buy)하는 경향을 보고하였으며, Antweiler and Frank[2]에서는 야후 파이낸스(Yahoo! Finance)에 게시된 온라인 메시지들이 주식수익률과 거래량에 어떤 영향을 미치는지를 분석하였다. 한편 최근의 연구들에서는 특히 구글(Google)에서의 주간 검색 빈도를 표준화하여 발표하는 검색량 지수(Search Volume Index)를 통해 투자자들의 관심과 정보전파 속도를 측정하였다. Da et al.[9]와 Joseph et al.[15]은 미국 주식 식별코드(Ticker symbol)의 구글 검색 강도가 주식의 비정상 수익률(Abnormal return)을 증가시킨다는 결과를 보고하였다. Vlastakis and Markellos[25]는 주식 식별코드 대신 기업 이름을 사용한 연구에서도 같은 결과를 관찰하였으며, 뉴욕증권거래소와 나스닥의 주요 주식 수익률의 변동성과 거래량을 증가시키는 것을 관찰하였다. 미국 외의 시장에 대한 연구로, 독일 주식시장[3]과 일본 주식시장[24]을 대상으로 한 연구에서도, 구글 검색 강도를 통해 측정한 투자자의 관심 정도가 주식 거래량과 유동성을 증가시키는 것을 보고하였다. 국내 주식시장을 대상으로 한 연구에

서, Koo and Kim[19]은 검색량의 변화를 기반으로 한 투자전략이 주식수익률과 거래량을 유의하게 증가시키며, 이러한 효과는 중소기업에서 더 강하게 발생하는 것을 관찰하였다. 또한 Kim and Kim[16]은 한국의 네이버에서 언급 점유율이 높은 영화의 이메일이 유의미하게 높은 것을 관찰하기도 하였다. 이러한 연구들은 포털이나 관련 사이트에서의 검색 강도가 증가할수록, 기업의 수익과 주식의 수익률 및 거래량이 증가하며, 이러한 정보의 역할은 정보가 상대적으로 부족한 중소기업 등에서 더 두드러지게 나타난다는 것을 보여주는 연구들이라 할 것이다.

이상의 국내외 연구들은 시장 효율성의 검증 목적으로 포털 검색 강도가 주식 수익률과 거래량에 미치는 영향에 대해 주로 다루고 있다. 이에 비해 본 논문은 포털 검색강도로 나타난 투자자의 관심이 주가의 급락 위험, 즉 기업의 극단적 손실 위험에 미치는 영향을 분석한다. 주식수익률의 급락은 음(-)의 수익률에서 변동성이 커지는 비대칭적 변동성으로 정의될 수 있으며 이러한 현상의 설명을 위해 여러 이론적 또는 실증 연구들이 수행되어 왔다.

주가급락 현상에 대해 전통적으로 레버리지(leverage) 이론[5, 9], 변동성 환류(volatility feedback) 이론[23, 11], 확률적 버블(stochastic bubble) 이론[6]으로 설명되었다. 한편, Hong and Stein[12]은 투자자의 이질성(investor heterogeneity) 가정을 통해, 주식 가치에 대한 투자자들의 의견의 차이(difference of opinion)가 비대칭적 수익률 변동성을 발생시킬 수 있다는 이론을 제시하였으며, Chen et al.[8]에서는 이 이론을 실증적으로 분석하기도 하였다. 한편, Jin and Myers[14]의 연구 이후, 기업의

행태와 지배구조가 주가급락위험을 결정하는 중요한 요인으로 주목되고 있다. 즉, 기업의 불투명성으로 인해 주가에 반영되지 못한 부정적 정보가 갑작스럽게 시장에 유출될 경우 주가의 급격한 하락이 발생하게 된다는 것이다. 대표적인 연구로는 기업의 이익조정이 증가하거나[13, 20] 조세회피가 증가할수록[17, 18] 기업의 불투명성이 증가하여, 주가급락의 위험이 증가한다는 연구들이 발표되었다.

한편, 투자자 간 정보의 전파 또는 기업의 투명성이 해당 기업 주가의 급등(jump)에 미치는 영향에 대하여, Hutton et al.[13]이나 Lim and Kim[20] 등의 연구에서 분석되었다. 두 연구 모두에서 기업 정보의 불투명성은 주가의 급등에는 유의한 영향을 미치지 못하는 것으로 보고되었다. 이는 부정적 정보의 전파는 기업이 꺼리기 때문에 이러한 부정적 정보가 축적되었다 한꺼번에 유출될 경우 주가의 급락이 발생하는 반면, 긍정적 정보는 발생될 때마다 기업이 적극적으로 시장에 전파하기 때문에 주가 급등을 발생시키지 않는 것으로 해석된다.

## 2.2 가설의 설정

선행연구에서 살펴본 바와 같이, 최근의 연구들에서 포털에서의 검색 강도를 그 기업에 대한 투자자들의 관심의 대리변수(proxy variable)로 파악하여 측정하고 있으며, 이러한 투자자들의 관심은 그 기업에 대한 정보적 효율성을 향상시켜 투자자들 간 해당 기업에 대한 정보가 보다 정확하게 전달될 것이다. 또한 선행연구들에서는 기업 정보가 투자자들 사이에 잘 전파될수록 주가급락 위험을 감소된다는 사실을 이론적, 실증적으로 증명하였다.

본 연구에서는 이러한 논의들을 확장하여, 포털에서의 검색 강도로 측정되는 투자자들의 관심과 정보 전파의 효율성이 주식급락 위험을 유의하게 감소시키는지 실증적으로 분석한다. 또한 이는 투자자들의 관심 정도가 낮은 기업에서 특히 부정적 정보가 느리게 전파된다는 Hong et al.[12]의 연구결과를 주가급락 연구로 확장하는 분석이라 할 수 있다.

### 3. 분석자료 및 모형

#### 3.1 표본

본 연구의 표본기간은 2007년부터 2014년으로 하였으며, 표본기업은 코스피에 상장되어 있는 제조업체를 대상으로 하였다. 네이버의 검색 강도가 검색되는 2007년을 시작연도로 하였으며, 2015년 6월 이후 주가 상한 제한폭이 변경된 효과를 통제하기 위해 2014년까지 분석에 포함하였다. 실증분석을 위해 필요한 주가, 수익률, 거래량 및 회계 자료들은 FN-GUIDE를 이용하여 추출하였으며, 검색량 자료는 네이버를 통해 직접 수집하였다. 또한, 표본의 모든 기간 동안 수익률과 거래량이 보고된 기업만을 분석대상에 포함시켰는데, 이는 주식수익률과 검색량 모두에 비정상적인 영향을 미칠 수 가능성이 높은 상장, 증자, 상장폐지, 주식분할 등의 사건의 효과를 제거하기 위한 것이다.

#### 3.2 변수의 정의

##### 3.2.1 주가급락 위험(Crash Risk)

본 연구는 정보의 투명성과 투자자들의 관

심이 주가급락과 변동성에 미치는 영향을 분석하므로, 주식시장 전반의 영향을 제거한 기업고유(firm-specific) 주가수익률(weekly return)을 이용하였다. 이를 위해 먼저 다음의 식 (1)의 회귀식을 기업-연도별로 추정하였다.

$$r_{j,t} = \alpha_j + \beta_1 r_{m,t-2} + \beta_2 r_{m,t-1} + \beta_3 r_{m,t} + \beta_4 r_{m,t+1} + \beta_5 r_{m,t+2} + \epsilon_{jt} \quad (1)$$

식 (1)에서  $r_{j,t}$ 는 주식 j의 t기의 수익률을 의미하며,  $r_{m,t}$ 는 t기에서의 시장 수익률(market return)을 의미한다. 각각 2기간의 시장수익률의 선-후행(lead-lag)을 포함시킨 것은 비동차적 거래(non-synchronous trading)의 영향을 반영하기 위한 것이다. 기업 j의 고유수익률은 식 (1)의 잔차항( $\epsilon_{jt}$ )에 자연로그를 취하여,  $W_{j,t} = \ln(1 + \epsilon_{jt})$ 으로 측정한다. 본 연구에서는 이렇게 구한 기업고유 수익률을 이용하여 아래의 세 가지 방법으로 주가급락 위험을 측정하였다.

첫 번째 변수인 CRASH는 주가급락의 발생 여부를 측정하는 것이다. 이를 위해 먼저 각 기업-연도별로 주간수익률을 평균과 표준편차로 표준화(standardize)한 후, 이 표준화된 값이 -3.2보다 작은 값을 갖는 경우를 주가급락 주간(crash week)로 정의한다. 이때 -3.2는 표준정규분포에서 하위 0.1%에 해당하는 값을 구하기 위한 것이다. 본 논문에서는 Hutton et al.[13], Kim et al.[17]과 같이 주가급락 주간을 1회 이상 경험한 경우를 1로 하여, 주가급락의 첫 번째 변수인 CRASH를 구성한다.

주가급락 위험의 두 번째 변수는 Chen et al.[8]의 방법론을 따라, 음의 조건부 왜도를 이용한다. 여기서 (-)를 취한 것은 더 왼쪽으로

왜곡된(left-skewed) 수익률 분포를 가질수록 높은 값을 갖도록 조정한 것이다. 두 번째 변수인 NCSKEW는 아래의 식 (2)와 같이 나타나며,  $n$ 은 연도별 샘플의 수를 의미한다.

$$NCSKEW_{jt} = -[n(n-1)^{3/2} \sum W_{jt}^3] / [(n-1)(n-2)(\sum W_{jt}^2)^{3/2}] \quad (2)$$

주가급락의 세 번째 변수인 DUVOL 역시 Chen et al.[7]에서 제시된 것으로, 각 주식이 자신의 연간 평균수익률 이하인 주와 이상인 주로 구분한 후, 이 두 표본에서 각각 표준편차를 구한다. 이에 상승기간 대비 하락기간의 표준편차 비율에 로그를 취하여, 아래의 식 (3)과 같이 DUVOL을 계산한다.

$$DUVOL_{jt} = \log\left[\frac{(n_u - 1) \sum_{DOWN} W_{jt}^2}{(n_d - 1) \sum_{UP} W_{jt}^2}\right] \quad (3)$$

식 (3)에서  $n_u$ 와  $n_d$ 는 각각 상승기간과 하락기간의 수를 의미한다.

### 3.2.2 검색량 지수(Search Volume Index)

연구의 주된 변수인 검색량을 구하기 위해, 네이버의 주간 검색 지수인 네이버 트렌드(Naver Trend)를 이용하였다(<http://ca.datalab.naver.com/ca/step1.naver>). 네이버 트렌드는 검색량 지수라는 시계열 변수를 제공하는데, 이는 특정 키워드의 주간 검색 빈도를 보여주는 것으로, 해당 키워드의 평균으로 조정하여 0과 100 사이의 값을 갖도록 한 것이다. 본 연구에서는 코스피에 상장된 모든 기업을 대상으로 하여, 검색량 지수를 수집하였다.

기업명의 축약어를 이용하여 각 주식에 고유

식별기호(ticker symbol)를 부여한 미국의 경우와는 달리, 기업명을 검색 키워드로 할 경우, 두 가지 문제점이 발생할 가능성이 있다. 첫째는 기업명이 주식 투자 결정과 직접적으로 관계없는 이유, 즉 상품 정보, 채용, 기업 소속의 스포츠 팀 등의 이유 때문에 검색될 수 있다는 것이다. 두 번째는 기업명을 표기하는 방법이 여러 가지 존재할 수 있다는 것이다. 첫 번째 문제점과 관련하여, Vlastakis and Markellos[25]는 투자목적과 관련 없는 요소는 랜덤하거나 계절적 요인을 갖기 때문에 데이터의 적절한 조정으로 이러한 문제점을 경감시킬 수 있다고 주장한다. 즉, 기존의 연구들은 검색량이 수익률 또는 거래량에 단기적으로 어떤 영향을 미치는가에 대한 분석들이므로, 최근 여러 기간의 검색량 중위수 값을 구해 차감하거나, 지난 해 같은 시기의 검색량을 차감하여 투자와 무관한 검색량의 영향을 최소화하였다. 한편, 본 연구는 1년 중 검색량의 평균과 최대값을 사용하기 때문에 계절적 요인에는 영향을 받지 않을 것이며, 랜덤하게 발생하는 검색량의 미시적인 증감의 영향도 적을 것으로 판단할 수 있다. 두 번째 문제점과 관련하여, 네이버 트렌드에서는 복수의 키워드에 대한 검색량을 동시에 측정할 수 있는데, 여러 이름으로 검색될 것으로 예상되는 기업명을 함께 검색한 후, 이 검색량 지수를 평균한 값을 이용함으로써, 이러한 문제점을 제거하였다. 예를 들어, 엘지유플러스(거래소코드 032640)에 대해, “엘지유플러스”, “LG유플러스”, “엘지텔레콤”, “LG텔레콤” 등의 이름으로 검색한 후, 이 키워드에 대한 검색량 지수를 기간별로 평균하여, 해당 주식의 검색지수로 이용하였다.

이렇게 구성된 각 기업명에 대한 검색량 지수에 대해, 연도별로 검색량 지수의 평균값

(avgintense)과 각 연도에서의 검색량 지수의 최대값(maxintense)의 두 가지 변수를 구성하여, 투자자의 관심의 강도와 기업 정보전파의 효율성을 측정하였다.

### 3.2.3 통제변수

본 연구에서는 Chen et al.[8]의 연구 이후 주가 급락과 관련된 다양한 논의들을 모두 반영하여 통제변수를 구성하였다. 먼저, 12개월 전의 값을 차감하여 추세를 제거한 주식회전율(dturlover)을 계산하였으며, 기업-연도 별로 주식수익률이 변동성을 구하여 변수 sigma를 구성하였다. 또한 기업 수익률의 연간 평균값(avgret)을 분석에 포함시켰으며, 기업의 시가총액(mktcap), 시장-장부비율(mbratio), 총자산 대비 장기부채 값인 재무적 레버리지(lev), 자산이익률(roe) 등도 통제변수로 포함하였다. 마지막으로 이상의 변수들로 통제되지 않는 기업 특성을 반영하기 위해, 전년도의 음의 조건부 왜도(lncskew)를 분석에 포함하였다.

### 3.3 주요 변수의 요약 및 추세

표준화된 수익률이 -3.2 이하의 값을 갖게 되어, 주가의 급격한 하락을 보여주는 주가급락 여부(CRASH)의 발생빈도를 <Table 1>에서 보고하고 있다.

<Table 1> Frequency of Crash Weeks

Crashes in the firm-year	Numbers of Obs.	Percent of sample
0	4,585	88.51
1	585	11.29
2	10	0.19
Total	5,180	100

<Table 1>에서 보는 바와 같이, 분석대상인 5,180개의 전체 기업-연도 중, 주가급락을 경험하지 않은 기업-연도가 4,585개였으며, 주가급락을 경험한 기업-연도가 전체 표본의 11.49%인 585개였다. 이 중 585개의 경우는 1년 중 한 번의 주가급락을 경험하였으며, 10개 기업은 1년 중 두 번의 주가급락을 경험하였으며, 3개 이상의 주가급락을 경험한 기업은 존재하지 않았다.

<Table 2>는 각 기업의 주간 수익률과 시장수익률을 주가급락 발생 여부에 따라 분류한 후 기술통계량을 구하여, 주가급락 시기(CRASH = 1인 표본)와 일반적인 시기(CRASH = 0인 표본)간의 수익률을 비교분석하였다. 전체 656,077개의 기업-주간 표본 중, 약 0.19%인 1,245개가 주가급락을 경험한 주간으로 보고되는데, 이러한 기간 동안 주식의 주간수익률은 평균적으로 -26.08%로 보고되어, 한 주 동안 극단적인 주가 하락이 발생했음을 보여준다. 이 시기에는 시장 수익률도 -6.06% 정도로 포함되어, 시장 전반적인 수익률의 저하가 개별 주식의 수익률 급락에

<Table 2> Return in Crash Weeks(CRASH = 1) Versus all other Weeks(CRASH = 0)

CRASH	Freq.	firm returns			Market returns		
		Mean	Median	Std. Dev.	Mean	Median	Std. Dev.
1	1245	-26.0916	-23.08	15.66004	-6.05514	-5.11	7.590874
0	654,812	0.413695	0.00	9.054845	0.238914	0.32	3.822532
Total	656,057	0.363395	0.00	9.144961	0.22697	0.32	3.842957

〈Table 3〉 Yearly Distribution of Variables

year	Freq.	CRASH	NCSKEW	DUVOL	avgintense	maxintense
2007	620	0.109677	-1.3546	0.555481	16.26186	39.33177
2008	633	0.344392	-1.28106	0.646962	15.77914	36.54186
2009	641	0.056162	-1.31338	0.560732	18.45025	45.12652
2010	635	0.070866	-1.01751	0.583437	18.80086	48.29874
2011	652	0.07362	-1.18163	0.574105	18.00745	46.43681
2012	666	0.105105	-1.02484	0.595945	17.03413	44.02808
2013	665	0.061654	-0.89625	0.591777	15.89069	42.94105
2014	668	0.103293	-0.96869	0.620065	16.66646	45.13443
Total	5180	0.114865	-1.12641	0.591272	17.10846	43.51691

영향을 미친 것으로 판단된다. 추가급락이 발생하지 않은 일반적인 기간들의 경우는, 주식수익률과 시장수익률의 평균이 각각 0.41%와 0.24%로 보고된다. 주목할 만한 또 한 가지의 사실은 추가급락이 발생한 시기의 표준편차가 일반적인 경우에 비해 유의하게 크다는 것이다. 이는 추가급락이 발생할 경우, 주식수익률의 변동성이 매우 크다는 것을 의미하는데, 이는 추가급락이 발생할 경우 수익률의 분포가 왼쪽으로 긴 꼬리를 갖는 형태를 갖게 되면서 극단적으로

낮은 수익률들이 발생하여 변동성이 유의하게 높게 나오는 것으로 판단된다.

〈Table 3〉은 추가급락 위험의 세 변수인 CRASH, NCSKEW, DUVOL과 네이버 검색강도의 두 변수에 대한 연도별 추세를 보여주고 있다. 추가급락의 발생을 보여주는 CRASH의 경우, 금융위기 기간인 2007년에서 2009년에 가장 큰 값을 보이고 있으며, DUVOL의 경우에도 2008년에 가장 큰 값을 갖는 것으로 관찰된다. 이에 비해, 네이버 검색강도의 두 변수는 2007

〈Table 4〉 Descriptive Statistics

Variable	Obs.	Mean	Std. Dev.	Q1	Median	Q3
CRASH	5180	0.114865	0.31889	0	0	0
NCSKEW	5180	-1.12641	0.86719	-1.71762	-1.40417	-0.79533
DUVOL	5180	0.591272	0.38435	0.334278	0.58401	0.85174
avgintense	5180	17.10846	13.53332	6.461321	14.27612	24.8057
maxintense	5180	43.51691	32.62836	15.7	37	69
dturnover	5180	-0.00466	0.418405	-0.05482	-0.00427	0.03344
avgret	5180	0.243621	1.081509	-0.318	0.197404	0.78543
sigma	5180	6.544737	3.406447	4.31693	5.762982	7.864766
mktcap	5180	1.246662	6.594087	0.050993	0.120469	0.414054
mbratio	5180	1.302308	4.194941	0.54782	0.85548	1.383606
lev	5180	0.448253	0.251139	0.27446	0.442984	0.588951
roa	5180	0.013865	0.150308	0.001803	0.029188	0.065109
lncskew	5180	-1.15003	0.853465	-1.72736	-1.43075	-0.82279



년과 2008년에 가장 작은 값을 갖는 것으로 관찰되는데, 이는 본 연구의 가설인 투자자들의 높은 관심과 정보의 투명성이 주가급락의 위험을 낮춘다는 논의와 부합되는 결과이다.

<Table 4>는 분석에 포함된 변수들의 요약 통계량을 보고하고 있다. 앞서 살펴본 바와 같이 1년 동안 1회 이상의 주가급락을 경험하는 표본이 전체의 11.49%에 이르는 것으로 나타난다. 하락기간과 상승기간의 수익률의 변동성 비율

을 보여주는 DUVOL이 (+)의 값을 갖는다는 것은, 평균보다 낮은 수익률을 보이는 하락기간 동안의 수익률 변동성이 상승기간에 비해 높다는 것으로, 본 연구에서 고려하고 있는 비대칭적인 수익률 변동성을 보여주는 결과이다.

한편, 네이버 검색강도를 나타내는 두 변수들(avgintense, maxintense)은 0과 100 사이의 값을 갖는 변수들인데, 평균적으로는 각각 17.11과 43.52의 값을 갖는 것으로 나타났으며, 분포

<Table 5> Pairwise Correlation

	CRASH	NCSKEW	DUVOL	avg intense	max intense	dturnover	avgret
CRASH	1						
NCSKEW	0.0524*	1					
DUVOL	0.0669*	0.6431*	1				
avgintense	-0.0141	0.0562*	0.0466*	1			
maxintense	-0.0172	-0.0422*	-0.0458*	0.7222*	1		
dturnover	-0.0424*	-0.0908*	-0.0358*	0.0993*	0.1956*	1	
avgret	-0.2072*	-0.1098*	-0.0632*	0.0632*	0.1160*	0.2513*	1
sigma	0.0003	-0.4311*	-0.1670*	-0.0367*	0.1189*	0.3343*	0.0803*
mktcap	-0.0331*	0.0789*	0.0794*	0.0802*	0.0415*	-0.0012	0
mbratio	-0.0264	-0.0508*	-0.025	0.0091	0.0253	-0.0177	-0.0187
lev	0.0133	-0.1491*	-0.0453*	-0.0775*	-0.01	0.0416*	-0.2012*
roa	-0.0226	0.1220*	0.0866*	0.0826*	0.0082	-0.0803*	0.2401*
lncskew	0.0128	0.1705*	0.0125	0.0750*	0.0409*	0.0703*	0.0429*

	sigma	mktcap	mbratio	lev	roa	lncskew
CRASH						
NCSKEW						
DUVOL						
avgintense						
maxintense						
dturnover						
avgret						
sigma	1					
mktcap	-0.0773*	1				
mbratio	0.1232*	0.0255	1			
lev	0.3446*	-0.0444*	0.0706*	1		
roa	-0.3642*	0.0788*	-0.1518*	-0.3985*	1	
lncskew	-0.2446*	0.0620*	-0.0476*	-0.1321*	0.1010*	1

를 보면 (+)의 왜도를 갖는 것으로 관찰되었다. 이는 표본기간 중 대체로는 검색강도가 그다지 높지 않다는 것을 의미하며, 특정시점에 일시적으로 매우 높은 검색 강도를 갖기 때문인 것으로 해석된다.

<Table 5>는 변수들 간의 피어슨 상관계수를 보고하고 있다. 표에서 \*는 두 변수간의 상관관계가 5% 수준에서 유의함을 의미한다. <Table 5>에서 보는 바와 같이, 추가급락 위험을 측정하는 세 변수 CRASH, NCSKEW, DUVOL은 서로 유의한 (+)의 상관관계가 관찰된다. 네이버 검색 강도를 나타내는 두 변수들을

보면, 검색량의 1년 중 최대값인 maxintense가 특히 NCSKEW와 DUVOL 두 종속변수와 유의한 (-)의 상관관계를 갖는 것으로 관찰된다.

#### 4. 실증분석 결과

위의 추세 및 상관분석 등을 통해 네이버 검색 강도와 추가급락 위험 간의 관계를 분석했으나, 이러한 분석들은 다른 변수들의 영향을 통제하지 못하기 때문에 회귀분석의 결과를 제시한다.

<Table 6>은 추가급락 위험의 세 가지 변

<Table 6> Regression Results

VARIABLES	(1) CRASH	(2) CRASH	(3) NCSKEW	(4) NCSKEW	(5) DUVOL	(6) DUVOL
avgintense	-0.0223*** [-3.212]		-0.0036** [-2.454]		-0.0006 [-0.778]	
maxintense		-0.0021 [-1.072]		-0.0011*** [-2.919]		-0.0009*** [-4.125]
dturnover	-0.2229 [-1.306]	-0.2273 [-1.350]	0.1631*** [6.353]	0.1656*** [6.520]	0.0455*** [2.679]	0.0507*** [3.040]
avgret	-0.8977*** [-13.675]	-0.9102*** [-13.664]	-0.0608*** [-6.170]	-0.0601*** [-6.075]	-0.0259*** [-3.734]	-0.0246*** [-3.574]
sigma	0.1083*** [4.622]	0.0994*** [4.192]	-0.1102*** [-18.602]	-0.1094*** [-18.277]	-0.0181*** [-6.024]	-0.0160*** [-5.278]
mktcap	-0.0659 [-0.808]	-0.0784 [-0.962]	0.0052* [1.736]	0.0055* [1.813]	0.0019 [1.027]	0.0022 [1.117]
mbratio	-0.0210 [-0.959]	-0.0173 [-0.784]	-0.0012 [-0.886]	-0.0011 [-0.854]	-0.0037*** [-4.308]	-0.0036*** [-4.181]
lev	-1.0393*** [-2.999]	-1.0591*** [-3.070]	0.0865 [1.143]	0.0862 [1.148]	-0.0307 [-0.629]	-0.0358 [-0.738]
roa	-0.4774 [-0.913]	-0.5144 [-0.976]	-0.1113 [-1.308]	-0.1301 [-1.535]	0.0300 [0.511]	0.0197 [0.339]
lnskew	0.0205 [0.331]	0.0326 [0.530]	-0.0953*** [-6.108]	-0.0935*** [-6.064]	-0.0599*** [-9.088]	-0.0602*** [-9.309]
Constant			-0.4799*** [-8.401]	-0.4962*** [-9.114]	0.6728*** [21.749]	0.6878*** [23.418]
Obs.	3,075	3,075	5,180	5,180	5,180	5,180
R-squared			0.136	0.136	0.037	0.041

수에 대한 고정효과(fixed effect) 회귀 결과를 보여준다. 분석에서 고정효과 모형을 사용한 것은 기업의 추가급락 위험에 대해 기업 고유의 요인들이 작용하기 때문에 이러한 요인들을 통제하기 위한 것이다. 또한 Breusch-Pagan의 이분산성 테스트와 하우스만(Huasman) 테스트 결과 역시, 고정효과 모형의 사용을 지지한다. <Table 6>의 결과에서 괄호안의 숫자들은 이분산성을 고려한 z(또는 t) 통계량들을 의미하며, \*\*\*, \*\*, \*은 각각 추정된 계수값이 1%, 5%, 10% 수준에서 유의함을 의미한다.

첫 번째 종속변수인 CRASH의 경우, 추가급락의 발생 여부에 따라 1 또는 0의 값을 갖는 변수이므로 로짓(logit) 회귀분석을 이용하여 분석하였다. 또한 로짓 모형은 고정효과 모형에서도 추정 가능한 회귀분석 방법이다. 로짓 분석 결과를 보면, 네이버 검색의 연간 평균이 증가할수록 추가급락 발생이 유의하게 감소하는 것이 관찰된다.

나머지 두 개의 종속변수인 NCSKEW와 DUVOL에 대한 고정효과 분석결과에서도 네이버 검색 강도의 연간 평균이나 최대값이 증가할수록 유의하게 추가급락 위험이 감소하는 것으로 나타난다.

이상의 결과들은 모두 네이버 검색 강도로 측정된 투자자들의 관심이나 주식 관련 정보의 전파가 증가할수록 주식의 급격한 하락 위험이 감소할 것이라는 본 연구의 가설과 부합하는 것이다. 한편, 다른 변수들에 대한 회귀 분석 결과를 보면, 기존의 이론들과는 일치하지 않는 것으로 해석된다. 즉, 부채비율이 여러 모형에서 추가급락위험에 유의한 영향을 보이지 않는 것은 레버리지 가설을 지지하지

않는다는 것을 의미하며, 평균수익률이 추가급락위험을 유의하게 감소시킨다는 결과는 확률적 버블 가설의 예측과 상반되는 것이다. 한편, Chen et al.[8]에서는 추세를 제거한 주식 회전율을 이용하여 투자자들 간의 의견 차이를 측정하였는데, <Table 6>의 고정효과 분석 결과에서도 투자자 의견 차이가 추가급락 위험을 증가시키는 것으로 관찰되었다.

## 5. 강건성 검증

회귀분석에 대한 강건성 검증(robustness check)으로, 본 연구의 실증분석에서 제기될 수 있는 역의 상관관계(reverse causality) 문제를 제거한 분석을 시행하였다. 즉, 본 연구에서는 포털 검색 강도를 통해 측정된 투자자들의 관심이 추가급락에 미치는 영향을 분석하였는데, 추가급락이 발생하기 때문에 투자자들의 관심이 증가하는 역의 상관관계가 존재할 수 있으며, 이 경우 내생성(endogeneity)의 문제가 발생하게 된다. 회귀모형에 내생성이 존재할 경우 일치추정량을 얻을 수 없는 문제가 발생하므로, 이러한 문제를 해결하기 위한 방법으로 동적 패널 모형(dynamic panel model)을 이용하였다. 동적 패널 모형은 내생성을 제거하기 위한 도구변수(instrument variable)로 종속변수의 과거값(lagged value)를 이용하는 모형으로, 이 연구에서는 특히 아렐라노-본드 [2]의 GMM 모형을 이용하였다. 또한 도구변수로는 각 종속변수들의 1계 과거값들을 이용했는데, 이는 고정효과 모형에서 NCSKEW의 과거값을 통제변수로 이용한 것과 같은 의미이다.

〈Table 7〉 Dynamic Panel Regression

	(1)	(2)	(3)	(4)
VARIABLES	NCSKEW	NCSKEW	DUVOL	DUVOL
avgintense	-0.0060** [-2.417]		-0.0019 [-1.577]	
maxintense		-0.0012** [-2.006]		-0.0010*** [-3.723]
dtumover	0.1591*** [4.131]	0.1558*** [4.043]	0.0500*** [2.699]	0.0538*** [2.919]
avgret	-0.0623*** [-4.055]	-0.0629*** [-4.090]	-0.0403*** [-5.434]	-0.0381*** [-5.158]
sigma	-0.1091*** [-15.876]	-0.1089*** [-15.755]	-0.0160*** [-4.771]	-0.0143*** [-4.256]
mktcap	0.0119 [1.235]	0.0115 [1.185]	0.0038 [0.818]	0.0044 [0.948]
mbratio	0.0048 [0.613]	0.0046 [0.586]	-0.0028 [-0.746]	-0.0029 [-0.761]
lev	0.1672 [0.977]	0.1666 [0.973]	0.0292 [0.352]	0.0218 [0.265]
roa	-0.0946 [-0.554]	-0.1131 [-0.660]	0.0910 [1.107]	0.0682 [0.831]
L.ncskewness	-0.0120 [-0.496]	-0.0094 [-0.386]		
L.duval			-0.0103 [-0.405]	-0.0165 [-0.650]
Constant	-0.3832*** [-4.091]	-0.4324*** [-4.854]	0.7288*** [15.619]	0.7366*** [16.703]
Obs.	3,723	3,723	3,723	3,723

〈Table 7〉은 동적 패널 GMM 모형을 이용한 분석결과를 보고한다. 결과를 보면, 내생성을 통제한 후에도, 네이버 검색 강도의 두 변수들이 추가급락 위험을 유의하게 감소시키는 것이 관찰되었다. 한편, 이 모형들에서 도구변수로 사용된 NCSKEW와 DUVOL의 과거값들은 유의한 영향을 미치는 못하는 것으로 관찰되는데, 기업들의 고정효과를 이미 충분히 고려하여 과거값들에 포함된 설명력이 약화되는 것으로 판단된다.

## 6. 결론 및 향후 연구 방안

최근의 주식수익률과 위험에 대한 연구들에서는 투자자들의 관심과 정보 전파의 효율성의 역할이 더욱 강조되고 있다. 본 연구에서는 한국 포털에서 가장 큰 지배력을 갖는 네이버의 검색 강도를 이용하여 투자자들의 관심과 정보 전파 속도를 측정하고, 이러한 변수들이 추가급락 위험을 감소시키는 것을 관찰하였다. 이는 투자자들의 관심이 높아질수록, 기업의

부정적 정보들이 주식시장에 신속하게 전파되어 주가의 극단적인 하락 위험이 감소하는 것으로 해석할 수 있다.

본 논문의 공헌점은 다음과 같이 요약할 수 있다. 첫 번째, 기존의 국내외 연구들에서는 포털 검색 강도가 주식수익률에 미치는 영향을 중심으로 진행되어 자본시장의 효율성에 관한 시사점을 도출하려 하였다. 그러나 본 연구에서는 주식수익률 자체가 아니라 수익률의 분포 형태, 특히 수익률의 극단적 하락 위험에 대해 분석하였다. 주가급락과 같은 비대칭적 변동성 (asymmetric volatility)에서 발생하는 위험은 투자다각화를 통해 분산시키기 어렵다는 점을 고려할 때, 주가급락에 대한 분석은 투자자와 경영자 모두에게 중요한 문제이다. 두 번째, 과거의 국내 연구들에서는 기업 정보의 투명성을 측정하기 위해 이익조정(earnings management)의 크기[20]나 조세회피의 크기와 변동성[18] 등 회계 정보를 이용하였는데, 본 연구에서는 포털에서의 검색강도를 통해 보다 직접적이고 직관적인 변수로 측정하여 수익률 분포에 미치는 영향을 분석할 수 있었다는 점이다.

이전의 연구들에서는 특정 기업에 대한 투자자들의 관심이나 감정을 측정하기 어려웠기 때문에 이러한 변수들을 주식 거래량 등을 통해 간접적으로 측정할 수밖에 없었다. 본 연구에서는 투자자들의 관심을 포털에서의 검색 강도를 통해 보다 직접적으로 측정하려 노력하였으나, 투자 또는 기업분석과 무관한 검색의 영향을 완전히 제거할 수 없는 한계 또한 존재한다. 또한 포털에서 검색 강도를 발표한 기간이 길지 않기 때문에, 이러한 요인들이 주식수익률에 미치는 영향에 대한 통계적 강건

성이나 예측가능성을 제고하기 위해 지속적인 연구가 필요할 것으로 판단된다.

현재 국내외의 여러 인터넷 포털과 웹 사이트에서 주식과 기업에 대한 분석과 정보 공유가 활발해지고 있다. 또한 최근의 연구들에서는 데이터 마이닝 등을 통한 투자자들의 관심과 감정을 측정하거나 신문 등 미디어에 노출되는 정도 등을 측정하여 이들을 주식 시장에 연결시키는 분석들이 증가하고 있다. 향후 이러한 연구들이 더욱 확장될 것으로 예상되며, 정보의 전달이나 미디어에서의 노출 등이 정보 보유 수준이 다른 여러 투자자 집단의 거래 행태에 어떤 영향을 미치는지 분석하는 것으로 연구를 확장할 수 있을 것으로 기대된다.

---

## References

---

- [1] Allerano, M. and Bond, S., "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations," *Review of Economic studies*, Vol. 58, pp. 277-297, 1991.
- [2] Antweiler, W. and Frank, M. Z., "Is all that talks just noise? The information content of internet stock message boards," *Journal of Finance*, Vol. 59, pp. 1259-1294, 2004.
- [3] Bank, M., Larch, M., and Peter, G., "Google Search Volume and its influence on liquidity and returns of German Stocks," *Financial Markets and Portfolio Management*, Vol. 25, pp. 239-264, 2011.

- [4] Barber, B. M. and Odean, T., "All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors," *Review of Financial Studies*, Vol. 21, pp. 785-818, 2008.
- [5] Black, F., "Studies of stock price volatility changes," *Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economical Statistics Section*, pp. 177-181, 1976.
- [6] Blanchard, O. J. and Watson, M. W., "Bubbles, rational expectations, and financial Markets," In Wachtel, P. (Ed.), *Crisis in Economics and Financial Structure*, Lexington Books, Ma, pp. 295-315, 1982.
- [7] Bleck, A. and Liu, X., "Market Transparency and the Accounting Regime," *Journal of Accounting Research*, Vol. 45, pp. 229-256, 2007.
- [8] Chen, J., Hong, H., and Stein, J., "Forecasting crashes: Trading volume, past returns, and conditional skewness in stock prices," *Journal of Financial Economics*, Vol. 61, pp. 345-381, 2001.
- [9] Christie, A. A., "A stochastic behavior of common stock variances-value, leverage and interest rate effect," *Journal of Financial Economics*, Vol. 10, pp. 407-432, 1982.
- [10] Da, Z., Engelberg, J., and Gao, P., "In search of attention," *Journal of Finance*, Vol. 66, pp. 1461-1499, 2011.
- [11] French, K. R., Schwert, G. W., and Stambaugh, R. F., "Expected stock returns and volatility," *Journal of Financial Economics*, Vol. 19, pp. 3-29, 1987.
- [12] Hong, H. and Stein, J. C., "Differences of opinion, short-sales constraints, and market crashes," *Review of Financial Studies*, Vol. 16, pp. 487-525, 2003.
- [13] Hutton, A. P., Marcus, A. J., and Tehranian, H., "Opaque financial reports,  $R^2$ , and crash risk," *Journal of Financial Economics*, Vol. 94, pp. 67-86, 2009.
- [14] Jin, L. and Myers, C. S., " $R^2$  around the world: New theory and new tests," *Journal of Financial Economics*, Vol. 79, pp. 257-292, 2006.
- [15] Joseph, K., Wintoki, M. B., and Zhang, Z., "Forecasting Abnormal Stock Returns and Trading Volume using Investor Sentiment: Evidence from Online Search," *International Journal of Forecast*, Vol. 27, pp. 1116-1127, 2011.
- [16] Kim, J. and Kim, J., "Relationship between Internet Buzz Share and Market Share: Movie Ticket Case," *The Journal of Society for e-Business Studies*, Vol. 18, No. 2, pp. 241-255, 2015.
- [17] Kim, J. B., Li, Y., and Zhang, L., "Corporate tax avoidance and stock price crash risk: Firm-level analysis," *Journal of Financial Economics*, Vol. 100, pp. 639-662, 2011.
- [18] Kim, M. and Lim, H., "Tax avoidance, information quality and crash risk," *The Korean Journal of Financial Management*, Vol. 32, No. 1, pp. 35-68, 2015.
- [19] Koo, P. and Kim, M., "A Study on the Relationship between Internet Search

- Trends and Company's Stock Price and Trading Volume," *The Journal of Society for e-Business Studies*, Vol. 20, No. 2, pp. 1-14, 2015.
- [20] Lim, H. and Kim, M., "Corporate opacity and crash risk: Evidence from Korean listed firms," *Journal of Money & Finance*, Vol. 28, No. 3, pp. 35-68, 2014.
- [21] Merton, R., "A Simple Model of Capital Market Equilibrium with Incomplete Information," *Journal of Finance*, Vol. 42, No. 3, pp. 483-510, 1987.
- [22] Morck, R., Yeung, B., and Yu, W., "The information content of stock markets: why do emerging markets have synchronous stock price movement?," *Journal of Financial Economics*, Vol. 58, pp. 215-260, 2000.
- [23] Pindyck, R. S., "Risk, inflation, and stock market," *American Economic Review*, Vol. 83, pp. 334-351, 1984.
- [24] Takeda, F. and Wakao, T., "Google Search Intensity and Its Relationship with Returns and Trading Volume of Japanese Stocks," *Pacific-Basin Finance Journal*, Vol. 27, pp. 1-18, 2014.
- [25] Vlastakis, N. and Markellos, R. N., "Information demand and stock market volatility," *Journal of Banking and Finance*, Vol. 366, pp. 1808-1821, 2012.

## 저 자 소개



김민수  
1998년  
2000년  
2015년  
2016년~현재  
관심분야

(E-mail: equili@sch.ac.kr)  
고려대학교 경제학과 (경제학사)  
고려대학교 경제학과 (경제학석사)  
고려대학교 경영학과 (경영학박사)  
순천향대학교 경제금융학과 조교수  
정보검색, 데이터마이닝, 가상화폐, 행태재무



권혁준  
1997년  
2003년  
2013년  
2015년~현재  
관심분야

(E-mail: gloryever@sch.ac.kr)  
Virginia Commonwealth University (경영학사)  
연세대학교 경영학과 (경영학석사)  
연세대학교 정보대학원 (정보시스템 박사)  
순천향대학교 IT금융경영학과 조교수  
Blockchain, 가상화폐, 가상현실, UX, 의료정보시스템