부동산연구 : 제24집 제2호 2014. 6 (125~140)

인터넷 검색 활동과 주택 가격 및 거래량 간 동적 관계 분석*

A Dynamic Relationship Between Internet Search Activity, Housing Price, and Trading Volume

김대원** · 유정석***
Kim, Dai Won · Yu, Jung Suk

Abstract

In the housing market, the internet searching activity trail means collective thinking and represents purchase intent. Therefore, if we can capture the meaningful relationship between the internet searching activity, the housing price, and the housing trading volume, it would mean we can predict the future using the internet searching activity data as a index for the housing market. In this context, we conducted the empirical research to examine the dynamic relationship between the internet searching activity and the housing price and trading volume. Using the "NAVER Trend" data as a proxy for the internet searching activity, the apartment sale price index as a proxy for the housing price, and the apartment trading volume index as a proxy for the housing trading volume, we set up the panel data of 23 autonomous districts in Seoul form Jan. 2007 to Feb. 2014.

In results from the Arellano-Bond dynamic panel model, we found that the internet searching activity had an positive(+) effect on the housing price and trading volume with some time lags. From the results of IRFs and FEVDs derived from the panel VAR model analysis, we also found that the internet searching activity had the strongest effect on housing market at the first order lag and influences much more on the housing price rather than the trading volume. In addition, panel Granger causality test results showed that the internet searching activity and the housing price, the housing price and trading volume had mutual cyclic causalities each other.

Keywords ▶ internet searching activity, housing price and trading volume, arellano-bond dynamic panel model, panel VAR model, panel granger causality test

초 록

주택 시장에서 인터넷 검색 활동의 흔적은 집단적 사고이자 장래 주택 구매 수요의 반영을 의미한다. 따라서 인터넷 검색 활동과 주택 가격 및 거래량 간 유의미한 관계를 포착할 수 있다면, 이는 곧 인터넷 검색 활동이 주택 구매 수요의 대리 변수로서 미래 주택 시장을 예측할 수 있는 지표로 활용될 수 있음을 의미하게 된다. 이러한 맥락에서 본 연구는 인터넷 검색 활동과 주택 가격 및 거래량 간의 동적 관계를 확인하기 위하여 서울시 23개 행정구, 2007년 1월부터 2014년 2월까지를 연구 대상 및 범위로 설정하고, 인터넷 검색 활동의 대리 변수로서 "네이버 트렌드" 자료를, 주택 가격 및 거래량의 대리 변수로는 아파트매매가 지수 및 아파트매매거래량 지수 패널 자료를 사용하여 실증 분석을 실시하였다.

Arellano-Bond 동적 패널 모형 추정 결과, 인터넷 검색 활동은 일정 기간의 시차를 두고 주택 가격 및 거래량에 유의미한 양(+)의 영향을 미치는 것으로 확인되었다. 패널 VAR 모형을 통한 IRFs 및 FEDVs 분석 결과, 인터넷 검색 활동은 주택 가격 및 거래량에 1차 시차에서 가장 큰 양(+)의

* 이 논문은 2012년 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(NRF-2012S1A3A2033330).
** 단국대학교 일반대학원 도시계획및부동산학과 박사과정, 제1저자(dw2613@gmail.com)
*** 단국대학교 사회과학대학 도시계획부동산학부 조교수, 교신저자(jsyu@dankook.ac.kr)

영향을 미치며, 주택 거래량 보다는 주택 가격에 더 큰 비중으로 영향을 미침을 확인할 수 있었다. 패널 Granger 인과성 검정 결과, 인터넷 검색 활동과 주택가격, 주택 가격과 거래량은 상호 순환적 인과성을 갖는 것으로 나타났다.

주요단어▶ 인터넷 검색 활동, 주택가격 및 거래량, Arellano-Bond 동적 패널 모형, 패널 VAR 모형, 패널 Granger 인과성 검정

Ⅰ. 서 론

최근 국내 주택 시장에서 가장 두드러진 변화는 수요자 중심 마케팅의 부각이라고 할 수 있다. 이는 과거 공급하기만 하면 투자 붐을 일으키던 시기가 공급자 위주의 시장이었던 것에 비교할 때 매우 큰 변화라 할 수 있다. 미분양 물량의 증가로 인해 큰 손실을 경험한 주택 공급자들은 소비자들의 요구와 선호도를 반영하기 위해 많은 노력을 기울이고 있으며(김대원・유정석, 2014), 시장 분위기를 빠르게 파악하여 적절한 공급 시점을 찾기를 원하고 있다. 그야말로 국내 주택 시장은 수요자 중심 시장으로 변모하고 있는 것이다(이수욱, 2008). 이러한 시장 분위기를 반영이라도 하듯이 2011년 7월부터 국토연구원은 부동산 소비심리지수를 지역별로 조사하여 매월 발표하고 있으며, 학계에서도 주택 또는 부동산 시장과 관련하여 소비 심리에 대한 다수의 연구를 수행하면서 이러한 분위기의 확산을 입증하고 있다(최영걸 외, 2004; 정의철, 2010; 김대원・유정석, 2013a).

주택 시장의 분위기나 동향 파악은 미래의 주택 가격과 거래량을 예측하는 데 기초 정보를 제공하게 되며, 이를 통해 예측된 주택 가격과 거래량은 적정한 공급 시점, 가격, 물량 등을 결정하는 데 중요한 정보로 활용될 수 있다. 시장 동향을 파악하기 위한 방법으로는 주택 가격 또는 거래량의 변화 정도를 측정하여 시장 분위기를 역으로 추정하고, 이를 다시 미래의 가격 또는 거래량을 예측하는 데 활용하는 방법이 있다. 또한 상기 언급한 국토연구원의 부동산 소비심리지수처럼 주택 시장 참여자 또는 전문가들을 상대로 시장 분위기 예측 정도를 직접 묻는 질적 접근법도 있다. 물론, 이러한 방법들도 시장 동향을 파악하는 데 매우 유용한 정보를 제공할 수 있지만, 이러한 방법들은 자료를 수집하고 이를 정리하여 분석하는 데 상당 기간이 소요되므로 빠르게 변화하는 시장 분위기를 파악하기에는 다소 어려움이 따른다.

현재 시대를 일컬어 인터넷 시대(internet era)라고도 한다. 인터넷으로 인해 사람들의 생활에 많은 변화가 일어났기 때문이다. 사람들은 인터넷에서 정보를 얻기도 하고, 자신들이 알고 있는 정보를 인터넷을 통해 공유하기도 한다. 인터넷 상거래, 문서의 교류, 전문 정보의 공유, 인생 상담까지 모두 나열하기 어려울 정도로 인터넷은 우리의 삶에 지대한 영향을 미치면서 인류의 생활을 변화시키고 있다. Rangaswamy et al.(2009)은 사람들의 인터넷 검색 활동에 대하여 “집단적인 사고(thinking)”이자 “미래에 일어날 일에 대한 사고(thinking)”라고 해석하였다. 즉, 인터넷 검색 활동의 흔적을 추적하게 되면 다수의 사람이 현재 무엇을 생각하고 있는지를 파악할 수 있으며, 미래에 기대하고 있는 것이 무엇인지를 예측할 수 있다는 의미이다. 이러한 변화의 흐름은 주택 시장이라고 예외는 아니다. 주택을 중개하는 업소들은 매물 정보를 인터넷 포털 사이트 등에 올리고, 매도나 매수를 희망하는 소비자들은 인터넷을 통해 주변 시세를 파악하고 시장 분위기를 확인한다. 종종 실제와 일치하지 않거나 빠르게 갱신되지 않은 정보로 인해 인터넷 정보의 신뢰성이 의심받기도 하지만, 주택 시장에서 인터넷을 통한 정보 공유 및 탐색 작업은 이제 매우 보편적인 현상이다.

이러한 맥락에서 인터넷이 주택 시장 동향을 파악하기 위한 좋은 대안이 될 수 있을 것이란 생각을 하지 않을 수 없다. 사람들은 주택 매매에 앞서 시장 분위기에 대한 정보를 얻기 위해 가장 우선적으로 인터넷을 활용한다. 이들이 인터넷에서 얻고자

www.kci.go.kr

하는 주택에 대한 정보는 곧 미래에 그들이 소유하고자 하는 주택에 대한 정보일 것이다. 즉, 현재 사람들의 검색 활동은 미래에 대한 소망 또는 희망을 담은 활동이므로, 이는 곧 미래를 예측할 수 있는 양질의 정보로 활용이 가능할 것이다. 게다가 인터넷 검색 활동 정보는 검색 서비스를 제공하는 인터넷 포털 등의 데이터베이스에 고스란히 기록되므로 그 자료의 가공 및 활용이 거의 실시간으로 이루어질 수 있다는 장점도 갖고 있다. 따라서 이를 주택 시장에 적용하여 인터넷 검색 활동과 주택 가격 및 거래량 간의 관련성을 찾을 수 있다면, 인터넷 검색 활동 정보를 소비자의 선호도 및 심리 지수를 대용할 수 있는 지표로도 활용이 가능할 것이며, 이는 곧 미래의 주택 수요와 가격을 예측할 수 있는 효율적인 기초 자료로 활용 및 응용될 수 있을 것이다. 전술하였듯이, 인터넷 검색 활동은 컴퓨터나 모바일 장비 등 모두 실시간으로 자료 집계가 가능하므로 주택 시장을 예측하는 지표로서 인터넷 검색 활동 자료를 활용하게 될 경우 효율성이 매우 높을 것으로 판단된다.

이에 본 연구는 지역별 주택 관련 키워드의 검색 횟수와 주택 가격 및 거래량의 관계를 살펴봄으로써 사람들의 인터넷 검색 활동이 주택 시장의 거래 또는 가격 변화와 연관성을 갖고 있는지를 확인하고자 하며, 나아가 인터넷 검색 활동이 주택 시장의 가격과 거래량을 예측할 수 있는 지표로 활용 가능한지에 대하여 살펴보고자 한다. 주택 구입 초기 시점에 대부분의 정보를 인터넷을 통해 구하고 있는 세태와 거의 실시간에 가깝게 자료가 수집 및 가공될 수 있다는 인터넷 정보의 특성을 감안할 때, 인터넷을 통한 주택 시장의 예측은 주택과 관련된 정책의 수립, 마케팅 시점 및 유형의 결정 등 주택 시장을 예측하고 의사 결정을 내리기 위한 유용한 도구가 될 수 있을 것이다.

이후 본 연구의 전체적인 구성은 다음과 같다. Ⅱ장에서는 인터넷 검색 활동을 통해 다양한 사회·경제 현상에 대한 예측 가능성을 연구한 선행 연구들을 살펴볼 것이며, Ⅲ장에서는 본 연구의 주요한 관심인 주택 가격 및 거래량과 인터넷 검색 활동 간의 관계를 살펴보기 위해 사용된 변수들과 분석 방법 및 절차에 대해 언급하고자 한다. Ⅳ장에서는 실증분석 결과를 근거로 인터넷 검색 활동과 주택 가격 및 거래량 간의 동적 관계를 살펴보고 활용 가능성을 논하고자 한다. 끝으로 Ⅴ장에서는 본 연구의 결론 및 시사점, 그리고 연구의 한계를 제시한다.

Ⅱ. 선행연구 고찰 및 연구의 차별성

인터넷 검색 활동 자료를 통해 경제 현상의 예측을 시도한 최초의 연구는 Ettredge et al.(2005)의 연구로 알려져 있다(Choi and Varian, 2012). Ettredge et al.(2005)은 인터넷 검색 활동 자료와 미국 실업률 간의 관계를 분석, 직업 관련 검색 횟수의 증가와 실업률 간에는 통계적으로 유의한 양(+)의 관계가 있음을 밝혔다.

Radinsky et al.(2008)은 인터넷 검색어를 통해 미래의 나타날 뉴스를 예측할 수 있는 기법을 제안하였는데, 이들은 구글 트렌드(Google Trends)[1] 자료를 사용하여 정해진 기간 동안 특정 검색어가 얼마나 많이 검색되었는지를 계산하여 미래에 해당 검색어가 뉴스 기사에 등장하게 될지를 예측할 수 있는 알고리즘을 제시하였다.

Suhoy(2009)는 이스라엘에 대한 구글 인사이트 검색어 자료와 실업률 간의 관계를 분석하였는데, 구직, 가정, 여행, 부동산, 음식, 미용, 개인생활 관련 검색어의 조회 횟수와 이스라엘 경제 지표와의 관련성에 대한 그랜저(Granger) 인과성 검정을 실시하였다. 분석 결과, 이스라엘 경제 지표를 예측함에 있어 구직 관련 검색어가 가장 높은 영향력을 미치고 있음을 제시하였다.

Preis et al.(2010)은 인터넷 검색과 금융 시장 유동성 간의 관계를 확인하기 위해 특정 회사 명칭의 인터넷 검색 횟수와 S&P 500의 거래량 자료를 주간 단위로 측정하여 양자 간의 관계를 실증 분석하였다. 분석 결과, 인터넷 검색 횟수의 증가가 주식 거래량의 증가에 영향을 주었다. 이와 유사하게

1) 2008년 8월부터 구글 인사이트(Google Insights for Search)라는 명칭의 서비스로 시작되어 2012년 11월 구글 트렌드(Google Trends, http://www.google.com/trends/)로 그 명칭을 변경하고 전 세계 언어에 대한 인터넷 검색 관련 통계를 제공하고 있다.

Goel et al.(2010)은 인터넷 검색 횟수가 사람들의 일상생활을 예측할 수 있는지를 확인하기 위해 특정 검색어 검색 횟수와 영화 수익률, 비디오 게임 판매량, Billboard Hot 100 차트 순위 등과의 관련성을 연구하여 유의미한 관련성을 밝혀냈다.

한편, Vosen and Schmidt(2011)는 개인들의 소비 정도를 예측하는 지표로서 전통적 설문 조사 기법과 구글 트렌드 자료 중 어떠한 것이 더욱 효과적인지를 확인하는 연구를 실시하였다. 분석 결과, 인터넷 검색 활동의 대용 변수인 구글 트렌드 자료가 전통적 설문조사 기법의 소비자 심리지수인 MCSI[2] 또는 CCI[3] 보다 개인 소비 증가율을 예측하는 데 효과적인 것으로 나타나, 인터넷 검색 활동이 소비자 심리를 보다 잘 대표한다고 주장하였다.

인터넷 검색 활동을 통해 미래 현상을 예측하기 위한 시도는 의학 분야에서도 활발하게 이루어지고 있는데, 의학 분야에서는 주로 특정 질병에 대한 인터넷 검색 활동의 증가와 실제 질병 발생 또는 관련 의료기관 방문 횟수와의 관련성을 찾는 연구들이 주류를 이룬다.[4]

주택과 인터넷 검색 활동 간의 관계를 확인하는 연구도 일부 존재하는데, 이 중 Baryla et al.(2000)의 연구가 인터넷 검색과 주택과의 관련성을 연구한 최초의 시도로 보인다. Baryla et al.(2000)의 연구는 시장 상황에 따라 소비자들의 검색 지속 기간이 달라지고 있음을 생존 분석(survival analysis)을 통해 검증하였다. 하지만, 이들의 연구는 상기에서 살펴본 선행 연구처럼 검색 활동이 주택 시장을 예측할 수 있는가를 살펴본 것이 아니라, 역으로 시장 상황이 검색 활동에 영향을 미치는지에 대하여 연구하였다는 점에서 차이가 있으며, 이는 인터넷 초창기 시절이라는 시대적 상황이 반영된 연구 주제 설정으로 보인다.

한편, Wu and Brynjolfsson(2013)은 구글 검색 횟수가 주택 가격과 거래량을 예측할 수 있는가를 측정하였으며, 주택 시장 역시 인터넷 검색 활동에 의해 예측이 가능함을 실증 분석 결과를 통해 시사하였다. 이들은 나아가 가까운 미래에 인터넷을 활용한 nanoeconomic 자료[5]가 다양한 시장, 경영, 소비자 의사 결정 분야의 예측 기법을 바꿔 놓을 것이라고 언급하였다. 또한 Beracha and Wintoki (2013)는 인터넷 검색 활동의 증가가 비정상적인 주택 가격 변화를 예측할 수 있다고 결론지었다.

국내 연구 중 인터넷 검색 활동과 사회・경제 현상 간의 관계를 연구하는 사례는 찾아보기 어렵다. 매우 소수의 연구만이 인터넷 검색 활동과 현상 간의 관계를 규명하고 있는데, 박정은 외(2009)는 상기 해외 선행연구에서 소개한 것과 유사하게 인터넷 검색과 유행성 독감 간의 관련성을 연구하여, 인터넷 검색이 유행성 독감을 감지할 수 있는 선행 자료로서 유용함을 언급하고 있다. 한편, 인터넷 검색 활동이라고 보기는 어렵지만, 인터넷 상에서 특정 게시 글을 조회한 횟수와 TV 프로그램의 인기도를 분석한 연구(김수도・조환규, 2012)가 수행된 바 있으며, 특정 인터넷 상점의 조회 수를 측정하여 실시간 구매 확률을 예측할 수 있는 방안을 제시한 연구(한송이 외, 2012)도 있으나, 국내 연구 중 인터넷 검색 동향과 사회・경제 현상 간의 관계를 밝히고 있는 연구는 아직까지 미비한 것으로 보인다.

상기한 바와 같이, 다수의 연구자들은 인터넷 검

2) MCSI(Michigan's Consumer Sentiment Index)는 미국 미시간 대학에서 매월 조사하여 발표하는 경제 전반에 대한 소비자 심리 지수이다. MCSI는 1966년 10월 지수를 100으로 설정하여 매월 알라스카와 하와이를 제외한 미국 전역을 대상으로 최소 500건 이상의 전화 설문조사를 통해 수집되는 자료를 사용하여 지수를 산출한다. 설문 조사 문항은 50여 개의 핵심 질문으로 구성되어 있고, 지수 산출 방법은 1940년대 미시간 대학의 George Katona 교수에 의해 고안되었다.

3) CCI(Conference Board's Consumer Confidence Index)는 미국 경제 조사 전문기관인 컨퍼런스 보드(Conference Board)에 의해 미국 내 5,000 여 가구를 대상으로 조사되는 경제, 경영에 대한 소비심리지수이다. CCI는 1985년 지수를 100으로 설정하여 문항별로 긍정과 부정을 5단계로 구분하여 답을 하는 방식으로 조사되며, 1967년부터 작성되기 시작하였다.

4) 구글 트렌드 서비스와 유사하게 구글은 전 세계 독감 확산 현황을 제공하고 있는데, 의학 분야에서의 연구는 주로 구글의 독감 트렌드(Google Flu Trends, http://www.google.org/flutrends/) 또는 블로그 활동 등이 질병 확산을 예측할 수 있겠는가에 대하여 연구하고 있으며, 상당부분 높은 상관관계를 갖는 것으로 밝히고 있다(Polgreen et al., 2008; Ginsberg et al., 2009; Corely et al., 2009; Pelat et al., 2009; Hulth et al., 2009; Brownstein et al., 2009; Valdivia and Monge-Corella, 2010).

5) Wu and Brynjolfsson(2013)은 인터넷 검색 자료가 실시간으로 집계되는 특징을 갖고 있어 매우 짧은 시간 단위로 예측치를 생성해 낼 수 있다는 의미로 nanoeconomic data란 용어를 사용하고 있다.

www.kci.go.kr

색 활동이 경제, 사회, 보건 등 다양한 분야의 미래를 예측할 수 있는가에 대하여 지대한 관심을 갖고 활발히 연구를 수행하고 있음을 알 수 있다. 이는 인터넷 검색 활동 자료가 개인들의 관심사를 모아 놓은 자료로서 그 자체로 집단적 사고를 보여주는 유용한 자료(Rangaswamy et al., 2009)이기 때문일 것이다.

이러한 맥락에서 본 연구는 인터넷 검색 활동과 주택 시장 간의 관계를 살펴보고 국내 주택 시장에서 인터넷 검색 활동이 주택 시장의 미래를 예측할 수 있는 자료로 활용이 가능한지를 살펴보는 것을 목적으로 한다. 주택 시장과 관련된 인터넷 검색 활동은 미래에 사람들이 구입하고 싶은 주택 또는 거주하고 싶은 지역에 대한 사전적 정보 획득이 주된 목적이라 볼 수 있다. 즉, 사람들의 인터넷 검색은 그들의 구매 의사를 인터넷 검색에 투영한 것이라고 볼 수 있다. 관련 연구에 따르면, 구매 의사와 실제 구매 활동은 높은 상관관계를 갖고 있으며(Morwitz et al., 2007), 구매 의사는 미래의 수요를 예측하는 데 사용되기도 한다(Juster, 1966; Morrison, 1979). 따라서 인터넷 검색 활동이 사람들의 구매 의사를 반영하는 수단이라고 할 때, 인터넷 검색 활동은 미래 주택 수요를 예측할 수 있는 지표로서 활용될 수 있을 것이다.

이에 본 연구는 인터넷 검색 활동과 주택 시장 변화와의 관계를 확인하는 것을 목표로 삼으며, 시장 예측 지표로서의 인터넷 검색 활동 자료의 활용 가능성을 확인한다는 데에서는 상기 선행 연구들과 그 궤를 같이 하지만, 본 연구는 다음과 같은 점에서 선행 연구들과 차별적인 요소를 갖고 있다. 첫째, 본 연구는 국내 주택 시장에서 인터넷 검색 활동과 주택 가격 및 거래량 간의 동적 관계를 살펴보고자 하는데, 이는 아직까지 국내 주택시장 관련 연구에서 그 사례를 찾아보기 힘든 주제라는 점에서 차별성을 갖는다. 둘째, 본 연구는 인터넷 검색 활동과 주택 가격 및 거래량 간의 동적 관계를 확인함에 있어 다각적인 검토를 실시하고자 동적 패널 모형, 패널 VAR 모형 등 변수들 간 동적 관계를 확인할 수 있는 다양한 방법론적 접근 방법을 사용한다는 점에서도 차별성을 갖는다. 셋째, 인터넷 검색 활동에 대한 대용 변수로서 기존 선행 연구들에서는 주로 구글 트렌드 자료를 사용하였으나, 본 연구는 국내 주택 시장에 대한 연구로서 국내 상위권 검색 포털 서비스 업체인 네이버에서 제공하는 네이버 트렌드[6] 자료를 사용한다는 점에서 기존 해외 연구와 차별성을 갖는다.

Ⅲ. 연구의 범위 및 방법

1. 분석 자료

본 연구는 주택 가격 및 거래량의 대용 변수로서 아파트 가격 및 거래량 자료를 사용[7]하기로 하며, 공간적 범위는 서울시 23개 행정구[8]로 정하고, 시간적 범위는 2007년 1월부터 2014년 2월까지로 설정하고자 하는데, 이는 네이버 트렌드 자료가 검색어 통계를 제공하는 시점이 2007년 1월부터인 점을 감안한 것이다. 자료를 수집함에 있어 가격 자료는 국민은행에서 제공하는 아파트매매가격지수를 사용하고, 거래량 자료는 온나라 부동산 정보에서 제공하는 아파트매매 거래량 자료를 사용하고자 한다. 이 때, 아파트 가격 및 거래량 지수는 계절성을 통제하기 위해 X-12 승법 모형을 적용하여 행정구별로 계절 조정된 자료를 사용하였다.

6) 네이버 트렌드(http://trend.naver.com) 서비스는 국내 인터넷 포털 서비스업체인 네이버 통합검색에서 발생된 검색어 통계를 볼 수 있는 서비스로서 PC에서 검색된 통계는 2007년 1월부터 제공되고 있으며, 모바일 환경에서 검색된 통계는 2010년 1월부터 현재까지 주간 단위로 제공하고 있다.

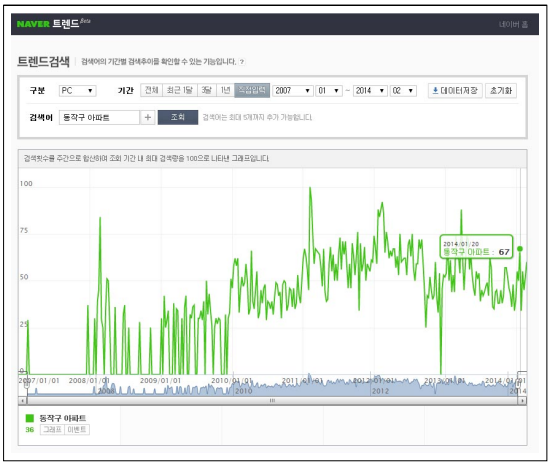
7) 통계청 발표 자료에 따르면, 2012년 현재 전국 전체 주택 유형 중 아파트가 차지하는 비율은 46.8%에 이르며, 서울시는 40.4%, 경기도는 53.6%, 인천시는 50.7%, 부산은 49.4% 등 전국적으로 아파트는 전체 주택 유형 중 절반정도의 수치를 보이고 있어 국내 주택 시장을 대표하기에는 충분해 보인다. 또한, 아파트는 단독주택 또는 다세대 주택 등에 비할 때 규격화 및 표준화가 잘 되어 있어 계량적 방법을 적용하기에 용이하다는 장점도 갖고 있다.

8) 서울시는 광역자치단체 중 가장 많은 전체 25개 행정구로 구성되어 있으며, 이 중 중구, 강서구를 제외한 나머지 행정구의 명칭은 다른 지역의 행정구와 중복되지 않는다. 따라서 검색어를 "강남구 아파트"라고 설정하여 인터넷 검색 횟수를 파악할 때, 이에 대한 통계치는 서울시 강남구 아파트의 통계치라고 확신할 수 있는 근거가 된다.

<표 1> 분석 자료 설명 및 기초 통계 (pooling data)

자 료 명	단위	출 처	관측수	평 균	표준편차	최소값	최대값
*PRICE** (아파트가격지수)	지수	KB 부동산통계	1978	103.77	4.64	84.63	117.32
*VOL** (아파트매매거래량지수)		온나라 부동산 정보		210.02	169.51	10.00	1841.00
*IS*** (인터넷검색지수)		네이버 트렌드		40.78	28.79	0.00	100.00

* *PRICE*, *VOL* 자료는 X-12 프로시저 승법모형을 적용하여 계절조정된 자료임
** *IS*는 네이버 트렌드에서 제공하는 주간 단위 지수를 월간 지수로 변환한 자료임



<그림 1> 네이버 트렌드 지역별 아파트 검색 횟수 조회(예시: 동작구 아파트)

한편, 웹로그 분석 전문 기업인 비즈스프링의 집계[9]에 따르면 2013년 국내 인터넷 포털 사이트의 점유율은 네이버가 78.56%, 다음(Daum)이 15.00%, 구글이 3.56%로 순으로 나타나 국내 인터넷 사용자들의 대부분이 검색 엔진으로 네이버를 사용하고 있음을 알 수 있다. 따라서 본 연구는 국내 주택 시장에 대한 인터넷 검색 활동을 측정함에 있어서 네이버 트렌드 자료를 사용하였다. 네이버 트렌드에서 행정구별 자료를 수집하는 방법은 검색어를 'OO구 아파트'로 입력하여 통계를 조회하는 방법을 사용하였다.[10]

다만, 네이버 트렌드 자료는 주간 단위로 검색 통계를 제공하고 있는데, 아파트 가격 및 거래량 자료는 월 단위로 자료가 제공되고 있어 분석을 위해 주기를 통일할 필요가 있으므로, 주간 지수로 제시되고 있는 네이버 자료를 식 (1)과 같이 월 단위 지수로 변환[11]하여 사용하였다.

$$\frac{\text{월별 주간 지수 합계}}{Max(\text{행정구별 월별 주간 지수 합계})} \times 100 \qquad (1)$$

상기와 같이 본 연구는 아파트 가격지수(*PRICE*), 아파트매매거래량지수(*VOL*), 인터넷 검색지수(*IS*)의 시계열 자료를 행정구별로 수집하여 패널 자료를 구성하였으며, 전체 자료에 대한 설명 및 기초 통계는 <표 1>에 제시된 바와 같다.

PRICE, *VOL*, 그리고 *IS* 자료는 분석을 위해 식 (2)와 같이 로그 차분하여 변화율을 산정하였다.[12] 이러한 변환을 통해 본 연구에서는 각 변수들의 변화율 간의 동적 관계를 살펴보았다.

$$APG_t = \ln\left(\frac{PRICE_t}{PRICE_{t-1}}\right) \qquad (2)$$

9) http://www.internettrend.co.kr/trendForward.tsp, 기간을 2013.1 ~ 2013.12로 설정하여 기간 내 평균치를 의미함.
10) 네이버 트렌드 통계는 PC 및 모바일 두 가지 통합 검색 방법으로 구분된 자료를 제공하고 있다. PC 검색 결과의 경우 2007년 1월부터 시계열 자료를 제공하고 있으나, 모바일 검색 결과의 경우에는 2010년 6월부터 자료를 제공하고 있기 때문에 본 연구에서는 충분한 시계열 자료를 확보하기 위해 PC 검색 자료를 사용하였다.
11) 네이버 트렌드에서 제공하는 통계는 실제 검색 횟수가 아니라 통합 검색의 검색 횟수를 0~100 수치로 환산하여 지수화한 자료이다. 이 때, 가장 많이 검색된 지점(주 단위)을 기준(100)으로 설정하여 나머지 기간의 검색 횟수를 상대값으로 환산한 것이다. 따라서 본 연구에서는 주간 단위로 제공되는 네이버 트렌드 자료를 월간 단위로 전환함에 있어 이와 동일한 방식으로 월간 지수 합계 최대값을 100으로 설정하여 나머지 기간을 상대값으로 환산하는 방식을 따랐다.
12) 시계열 자료의 변화율(growth rate)을 구함에 있어 일반적으로 로그차분 방식을 사용하게 되는데 이는 자연로그가 갖는 특성으로 인해 로그 변환하여 차분함으로써 원 자료의 변화율을 구할 수 있기 때문이다. 뿐만 아니라, 이러한 변환은 차분을 통해 이루어지기 때문에 시계열 자료의 추세를 제거하여 단위근 없는 시계열을 생성할 수 있다는 장점도 갖고 있다.

<표 2> 분석 자료 패널 기초 통계

자 료 명		평 균	표준편차	최소값	최대값	관측수
APG	overall	-0.0001	0.0055	-0.0346	0.0547	N = 1955
	between		0.0010	-0.0019	0.0017	n = 23
	within		0.0054	-0.0331	0.0532	T = 85
AVG	overall	0.0008	0.3972	-2.3590	1.5088	N = 1955
	between		0.0068	-0.0109	0.0147	n = 23
	within		0.3972	-2.3704	1.4974	T = 85
ISG*	overall	0.0208	0.4236	-1.8322	2.1777	N = 1547
	between		0.0235	-0.0634	0.0630	n = 23
	within		0.4232	-1.7480	2.2619	T-bar = 67.2609

* ISG는 패널 개체별로 지수가 0인 시점이 상이하므로, 평균 시계열 자료수를 통계치로 제시

<표 3> 패널 단위근 검정 결과

단위근 검정 (lag: 1)	APG	AVG	ISG
H_0: unit root (assumes common unit root process)			
Levin, Lin and Chu t-stat.	-3.636 (0.0001)	-31.639 (0.0000)	-19.362 (0.0000)
H_0: unit root (assumes individual unit root process)			
Im, Pesaran and Shin W-stat.	-8.979 (0.0000)	-33.198 (0.0000)	-34.956 (0.0000)
ADF Fisher Chi-square	171.276 (0.0000)	735.100 (0.0000)	663.115 (0.0000)
PP Fisher chi-square	287.763 (0.0000)	457.074 (0.0000)	808.556 (0.0000)
Data Information			
Cross-sections	23	23	23

주 : 괄호안 수치는 p-value를 의미(* p<0.1, ** p<0.05, *** p<0.01)

$$AVG_t = \ln\left(\frac{VOL_t}{VOL_{t-1}}\right)$$

$$ISG_t = \ln\left(\frac{IS_t}{IS_{t-1}}\right)$$

<표 2>는 식 (1)의 변환 과정을 통해 산출된 각 변수들의 패널 기초 통계량을 제시하고 있다.

모든 변수에서 패널 개체수는 23개이며, 시계열의 길이는 APG(아파트 가격지수 변동율), AVG(아파트매매 거래량지수 변동률)는 85개이며, ISG(인터넷 검색지수 변동률)는 일부 기간에서 지수가 0으로 나타난 구간은 로그 차분 과정에서 실측치로 간주되어 패널 평균 67.3개의 시계열을 갖는 불균형 패널 자료로 나타난다.

상기에서 산정된 APG, AVG, 그리고 ISG 변수를 사용하여 분석하기에 앞서 패널 자료들의 시계열적 정상성 여부를 확인하기 위하여 패널 단위근 검정을 실시하였으며, 그 결과는 <표 3>에 제시된 바와 같다. 패널 단위근 검정은 Levin, Lin and Chu 검정, Im, Pesaran and Shin 검정, ADF, PP 검정을 모두 사용하여 단위근의 존재 유무를 확인하였다.[13]

13) 본 연구에서 제시한 4가지 단위근 검정 방법 중 Levin, Lin and Chu 검정은 자료의 형태를 pooling 시계열 자료로 가정하여 검정하는 방법이다. 본 검정의 귀무가설은 각 패널별 시계열 자료가 단위근을 갖는다는 것이다(Levin et al., 2002). 나머지 3개의 단위근 검정 방법은 패널 개체의 이분산성을 가정하여 단위근을 검정하는 방법인데, 이 중 Im, Pesaran and Shin 검정은 평균 ADF

검정 결과, 모든 검정에서 1% 유의수준으로 패널 자료에 단위근이 존재한다라는 귀무가설을 기각하여 모든 변수는 안정적인 패널 자료임을 확인할 수 있다. 따라서 이후 실증 분석에서는 별도의 처리 없이 APG, AVG, 그리고 ISG 변수 자료를 그대로 사용하였다.

2. 분석 모형

1) Arellano-Bond 동적 패널 모형

Arellano and Bond(1991)는 패널 모형을 추정함에 있어 독립변수로 사용되는 시차 적용 종속변수와 오차항 간의 상관관계를 감소시킬 수 있는 방안으로 GMM(Generalized Method of Moments) 추정 방법을 제안하였다. 이에 본 연구에서는 종속변수로 설정되게 될 APG 및 AVG 변수에 자기 상관성이 존재하게 됨을 사전적으로 확인하였으므로 식 (3)과 같이 Arellano-Bond 동적 패널 모형을 설정하여 ISG와의 동적 관련성을 확인하였다.

$$APG_{it} = \sum_{j=1}^{p} \alpha_j APG_{i,t-j} + \sum_{m=1}^{4} \beta_m ISG_{i,t-m} + \nu_i + \epsilon_{it} \quad (3)$$

$$AVG_{it} = \sum_{k=1}^{q} \alpha_k AVG_{i,t-k} + \sum_{m=1}^{4} \beta_m ISG_{i,t-m} + \lambda_i + \omega_{it}$$

$$i = 1, \ldots, N \;\; t = 1, \ldots, T_i$$

식 (3)에서 i는 23개 횡단면 패널 자료를 의미하며, t는 2007년 1월부터 2014년 2월까지의 시계열 기간을 의미한다. 한편, ν_i 와 λ_i는 각 모형별 패널 수준 효과를 의미하며, ϵ_{it}와 ω_{it}는 i.i.d (independent and identically distributed) 특성을 갖는다.

상기 동적 패널 모형에 있어서 APG 와 AVG 는 각 모형별 종속 변수로 설정되며, ISG 는 두 개의 모형에서 모두 설명 변수로 대입된다. 또한 시차 j 및 k는 각 모형별 Arellano-Bond 자기 상관성 검정 결과에 따라 가장 적합한 시차로 결정하게 된다.

2) 패널 VAR 모형

동적 패널 모형 추정이 인터넷 검색 활동과 주택 가격 및 거래량 간의 동적 관계를 제시할 수는 있으나, 본 연구에서는 변수들 간의 동적 방향성과 인과성 등 보다 다각적인 검토를 위하여 패널 VAR 모형[14] 분석을 실시하였다. 본 연구에서의 패널 VAR 모형은 식 (4)와 같다.[15]

$$Y_{it} = C + \sum_{s=1}^{m} A_s Y_{i,t-s} + \eta_i + d_{c,t} + e_t \quad (4)$$

식 (4)에서 Y_{it}는 3개의 변수 벡터 {APG, AVG, ISG }를 의미하며, $d_{c,t}$는 행정구별 더미를 의미한다. 이 때, s는 4개월 간의 시차를 적용하여 변수 간 동적 관련성을 파악하기 위해 4로 설정하였다.

IRFs(Impulse Response Functions, 충격반응 분석)을 추정하기 위하여 본 연구에서는 패널 VAR 모형 계수로부터 IRFs 행렬을 사용하는 방식으로 신뢰 수준(confidence intervals)을 계산하였으며. IRFs의 표준 오차를 계산하여 몬테카를로 시뮬레이션 기법을 사용하여 신뢰수준을 생성하였다.

한편, FEVDs(Forecast Error Variance Decompositions, 분산분해분석)는 다른 변수들의 충격에 의하여 대상 변수가 시간의 흐름에 따른 분산의 변동 정도를 제시하는 분석 방법을 의미한다.

(Augmented Dickey-Fuller)에 근거하여 표준 T-bar 통계치를 사용하는 방법이며, Im et al.(2003)은 표준 T-bar 통계치를 사용하는 대신에 소규모 표본에서 더욱 효율적인 추정치를 제공하는 W-통계치를 사용할 것을 권하고 있어 본 연구에서는 W-통계치를 제시하고 있다. 한편, ADF와 PP 검정은 패널 단위근 검정에 있어 Fisher's chi square 통계치를 사용하고 있다(Maddala and Wu, 1999; Choi, 2001; Hadri, 2000).

14) 패널 VAR 모형의 경우 일반 시계열 자료의 VAR 모형에서처럼 모든 변수들을 외생 변수로 설정한다는 점에 있어서는 동일성을 갖으나, 패널 개체들의 이분산성을 통제한다는 점에서 차이가 있다(Love and Zicchino, 2006).

15) 패널 자료를 사용하여 VAR 분석을 실시하기 위해서는 모든 패널 개체들이 동일한 구조를 갖는다는 전제가 필요하게 된다. 이를 위해 식 (4)에는 패널 자료의 고정효과를 고려하기 위하여 η_i을 포함하고 있는데, 이 때 시차가 적용된 종속 변수의 사용으로 인하여 변수들 간 상관성이 발생되어 왜곡된 추정 결과가 도출될 수 있다. 이를 바로 잡기 위해서는 Helmert 절차(procedure)라는 전향 평균 차분(forward mean-differencing) 방법을 사용하여야 한다(Arellano and Bover, 1995).

즉, FEVDs는 자신 또는 다른 변수들에 충격이 주어졌을 경우, 관심 대상인 변수가 시간의 흐름에 따라 받게 되는 상대적 영향 비중에 대한 정보를 제공하게 된다. 이에 본 연구에서는 FEVDs 기간을 2개월, 4개월, 6개월로 나누어 변수들이 받는 단기간 동안의 영향력 비중 변화를 살펴보았다.

3) 패널 Granger 인과성 검정

본 연구에서는 개별 패널 간의 이분산성을 고려하기 위하여 Dumitrescu-Hurlin(2012)이 고안한 방식으로 Granger 인과성 검정을 실시하였다. 즉, 23개 행정구 패널 개체별로 이분산성이 존재하는 것을 전제로 Granger 인과성 검정을 실시하였으며, 이로 인해 변수들 간의 인과성 방향을 확인할 수 있을 것으로 기대된다.

Ⅳ. 실증분석 결과

1. Arellano-Bond 동적 패널 모형 분석

인터넷 검색 활동과 주택 가격 및 거래량 간의 관계를 확인함에 있어, 인터넷 검색 활동이 주택 가격 및 거래량에 일정기간 시차를 두고 영향을 미칠 것을 가정하고 종속 변수로 설정되는 변수에 자기상관성(autocorrelation)이 존재하게 될 경우 왜곡된 추정이 될 수 있으므로 우선적으로 두 종속 변수에 대한 자기상관성을 점검하였다.

APG 및 AVG 자료에 대한 자기상관성 검정은 Cumby and Huizinga(1992)가 제안한 방법을 사용하였으며, 결과는 <표 4>에 제시된 바와 같다. <표 4>에서 왼쪽 부분의 검정은 시차별 범위 전체에 대한 자기상관성에 대한 검정이며, 오른쪽 검정은 특정 시차별 자기상관성에 대한 검정이다. 본 검정의 귀무가설은 시계열 자료가 시차 q의 이동평균을 가질 때 q의 값이 0(zero)인지 여부를 확인하여 자기상관성을 탐지하는 방식이다. 따라서 검정 결과 이동평균의 시차인 q가 0이라면 자기상관성을 갖지 않는 것을 의미하며, 0이 아니라면 자기상관성을 갖는다고 해석할 수 있다.

자기상관성 검정결과, APG 및 AVG 모두 자기상관성을 갖는 것으로 나타났으나, 특정 시차별 검정 결과는 다소 차이를 나타낸다. 이에 따라, 변수의 자기상관성 시차를 결정하기 위해 자기상관 시

<표 4> APG 및 AVG에 대한 Cumby-Huizinga 자기상관성 검정 결과

시차 범위에 대한 자기상관성 검정 (q = lags, H_0: q = 0, serially uncorrelated)				특정 시차별 자기상관성 검정 (q = lag, H_0: q=specified lag-1, serially uncorrelated)			
APG autocorrelation test							
lags	chi2	df	p-val	lag	chi2	df	p-val
1-1	1006.089	1	0.0000	1	1006.089	1	0.0000
1-2	1019.466	2	0.0000	2	315.624	1	0.0000
1-3	1019.598	3	0.0000	3	149.302	1	0.0000
1-4	1020.436	4	0.0000	4	74.529	1	0.0000
1-5	1020.710	5	0.0000	5	43.651	1	0.0000
AVG autocorrelation test							
1-1	100.011	1	0.0000	1	100.011	1	0.0000
1-2	105.751	2	0.0000	2	0.003	1	0.9560
1-3	111.886	3	0.0000	3	3.053	1	0.0806
1-4	112.04	4	0.0000	4	1.759	1	0.1847
1-5	135.069	5	0.0000	5	22.094	1	0.0000

주 : 1) H_0: disturbance is MA process up to order q
2) H_a: serial correlation present at specified lags > q

www.kci.go.kr

<표 5> Arellano-Bond 동적 패널 추정 결과(APG-ISG)

Arellano-Bond Dynamic Panel Estimation Results					
Variables	**lags: 1**	**lags: 2**	**lags: 3**	**lags: 4**	**lags: 5**
L1.APG	0.6627 (12.82)***	**0.6037 (9.79)*** **	0.6018 (9.71)***	0.6029 (9.68)***	0.6039 (9.72)***
L2.APG	-	**0.0864 (2.66)*** **	0.0717 (1.40)	0.7431 (1.41)	0.0735 (1.36)
L3.APG	-	-	0.0233 (0.61)	0.0493 (1.09)	0.0504 (1.12)
L4.APG	-	-	-	-0.0412 (-1.97)**	-0.0510 (-1.39)
L5.APG	-	-	-	-	0.0122 (0.38)
ISG	0.0011 (1.79)*	**0.0011 (1.77)***	0.0011 (1.76)*	0.0011 (1.79)*	0.0011 (1.76)*
L1.ISG	0.0024 (3.68)***	**0.0025 (3.60)*** **	0.0025 (3.55)***	0.0025 (3.56)***	0.0025 (3.58)***
L2.ISG	0.0011 (3.31)***	**0.0012 (3.40)*** **	0.0012 (3.36)***	0.0012 (3.31)***	0.0012 (3.38)***
L3.ISG	0.0007 (2.18)**	**0.0007 (2.07)** **	0.0007 (2.02)**	0.0007 (2.03)**	0.0007 (2.02)**
L4.ISG	0.0007 (3.10)***	**0.0007 (2.90)*** **	0.0007 (2.94)***	0.0006 (2.74)**	0.0006 (2.69)***
Cons.	-0.0004 (-4.19)***	**-0.0004 (-4.52)*** **	-0.0004 (-4.57)***	-0.0004 (-4.47)***	-0.0004 (-4.79)***
Arellano-Bond test for zero autocorrelation in first-differenced errors (H_0: no autocorrelation)					
order	**z (p-value)**	**z (p-value)**	**z (p-value)**	**z (p-value)**	**z (p-value)**
1	-3.24 (0.0012)***	**-3.22 (0.0013)*** **	-3.24 (0.0012)***	-3.24 (0.0012)***	-3.24 (0.0012)***
2	0.84 (-.4014)	**-0.10 (0.9228)**	0.64 (0.5238)	1.00 (0.3190)	1.04 (0.2994)

주 : Arellano-Bond Dynamic Panel Estimation 괄호안 수치는 z-statistics를 의미(* p<0.1, ** p<0.05, *** p<0.01)

차를 1-5차로 각각 적용한 결과를 도출하여 가장 적합한 모형을 선택하였으며, 우선 APG와 ISG 에 대한 Arellano-Bond 동적 패널 모형 추정 결과는 <표 5>에 제시된 바와 같다.

APG와 ISG에 대한 Arellano-Bond 동적 패널 모형 추정 결과, 자기상관 시차에 상관없이 설명변수인 각 시차별 ISG의 계수 및 통계적 유의도는 거의 비슷한 결과를 나타내지만, 각 시차별 종속 변수의 유의도는 상이하게 나타난다. <표 5> 하단에 제시된 Arellano-Bond 검정은 1차 차분 오차항에 대한 자기상관도를 보여주고 있는데, 모두 1계 차분일 경우에는 자기상관성이 나타났으나, 2계 차분에서는 자기상관성이 나타나지 않아 1~5차 자기상관 적용 모형 모두 시차적용에는 문제가 없는 것으로 나타났으나, 자기상관 시차가 3차 이상일 경우에는 시차 종속 변수의 통계적 유의도가 없는 것으로 나타났음을 알 수 있다. 따라서 모형 내 자유도를 최대한 확보하면서 모든 시차의 통계적 유의도를 가장 잘 보여줄 수 있는 모형은 <표 5>에서 음영으로 표시된 자기상관 시차 2차 모형이다. 자기상관 시차 2차인 모형을 살펴보면, ISG에 시차가 적용되지 않은 경우에는 통계적 유의도가 10% 수준으로 APG와의 관련성은 다소 낮지만 양(+)의 관련성을 갖는다. 반면, ISG 시차 1~4 변수는 모두 5% 또는 1% 수준에서 통계적 유의성을 갖으며, 모두 APG에 양(+)의 영향을 미치고 있는 것으로 확인된다. 즉, 행정구별 아파트에 대한 인터넷 검색 활동은 일정한 시차를 두고 아파트 가격 지수에 양(+)의 영향을 미치고 있다는 의미로서, 인터넷 검색 활동의 증가가 일정기간 경과 후 주택 가격의 상승으로 이어지는 것으로 해석된다.

한편, <표 6>은 AVG와 ISG 간의 Arellano-Bond 동적 패널 추정 결과를 제시하고 있다. AVG와 ISG 간의 Arellano-Bond 동적 패널 모형 추정 결과 또한 앞서 실시한 APG와 ISG 간 추정 결과와 유사하게 자기상관 시차의 증가와 상관없이 ISG 시차 변수의 계수 및 통계적 유의도는 거의 유사한 결과를 보여주고 있으나, 자기상관성에 대해 검정하고 있는 Arellano-Bond 검정 결과는 자기상관 계수가 5차일 경우에만 오차항의 2차 차분계수가 자기상관성을 갖지 않는 것으로 나타났으며, 이 때 모든 종속 변수에 시차가 적용된 설명 변수가 통계

www.kci.go.kr

<표 6> Arellano-Bond 동적 패널 추정 결과 (AVG- ISG)

Arellano-Bond Dynamic Panel Estimation Results					
Variables	**lags: 1**	**lags: 2**	**lags: 3**	**lags: 4**	**lags: 5**
L1.AVG	-0.3029 (-12.19)***	-0.3419 (-12.43)***	-0.3578 (-12.46)***	-0.3603 (-12.23)***	-0.3624 (-11.99)***
L2.AVG	-	-0.1272 (-5.91)***	-0.1739 (-6.62)***	-0.1770 (-6.10)***	-0.1937 (-6.11)***
L3.AVG	-	-	-0.1361 (-6.85)***	-0.1427 (-6.14)***	-0.1615 (-5.99)***
L4.AVG	-	-	-	-0.0195 (-0.94)	-0.0597 (-2.53)**
L5.AVG	-	-	-	-	-0.1223 (-6.25)***
ISG	0.0296 (0.84)	0.0435 (1.19)	0.0345 (0.91)	0.0360 (0.95)	0.0387 (1.02)
L1.ISG	0.1132 (3.69)***	0.1286 (4.06)***	0.1354 (4.29)***	0.1347 (4.25)***	0.1434 (4.59)***
L2.ISG	0.1224 (3.57)***	0.1358 (3.72)***	0.1465 (4.22)***	0.1474 (4.24)***	0.1421 (4.09)***
L3.ISG	0.0424 (-1.65)*	-0.0214 (-0.84)	-0.0119 (-0.44)	-0.0107 (-0.40)	-0.0069 (-0.24)
L4.ISG	0.0500 (2.73)***	0.0638 (3.52)***	0.0747 (4.12)***	0.0752 (4.13)***	0.0823 (4.33)***
Cons.	0.0026 (0.74)	0.0019 (0.50)	0.0021 (0.53)	0.0023 (0.58)	0.0034 (0.82)
Arellano-Bond test for zero autocorrelation in first-differenced errors (H_0: no autocorrelation)					
order	**z (p-value)**	**z (p-value)**	**z (p-value)**	**z (p-value)**	**z (p-value)**
1	-4.44 (0.0000)***	-4.44 (0.0000)***	-4.44 (0.0000)***	-4.44 (0.0000)***	-4.46 (0.0000)***
2	-3.25 (0.0012)***	3.63 (0.0003)***	-2.19 (0.0284)***	-2.06 (0.0392)**	-1.36 (0.1736)

주 : Arellano-Bond Dynamic Panel Estimation 괄호안 수치는 z-statistics를 의미(* p<0.1, ** p<0.05, *** p<0.01)

적 유의성을 갖는 것으로 나타나 가장 적합한 추정 모형은 <표 6>에 음영으로 표시되어 있는 자기상관 시차가 5차인 모형이다.

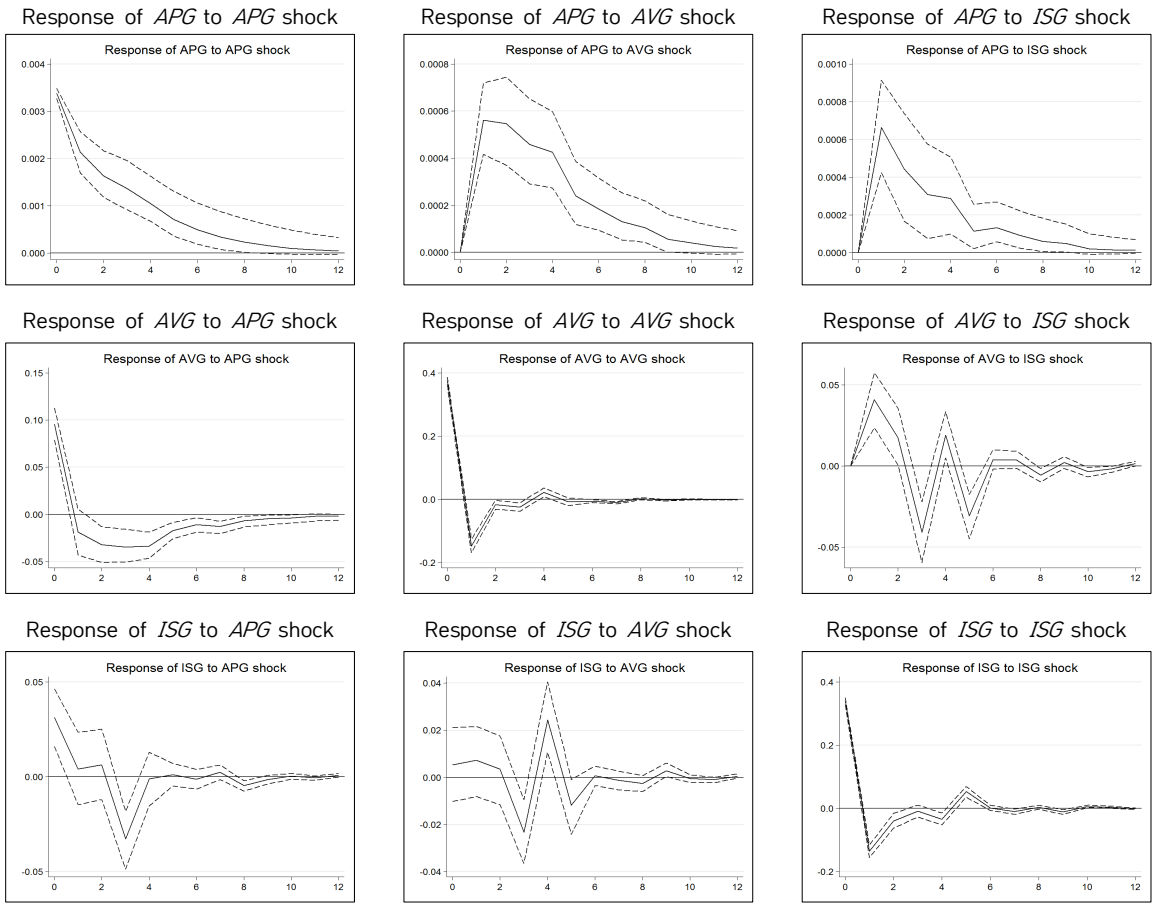
추정결과를 살펴보면, ISG의 1차 시차 및 3차 시차는 통계적 유의성을 갖지 않으며, 1차, 2차, 그리고 4차 시차에서 1% 수준으로 통계적 유의성을 갖는 것으로 나타났으며, 모두 부호는 양(+)으로 확인된다. 따라서 인터넷 검색 활동은 주택 가격에 대해서도 일정한 시차를 두고 영향력을 미치는 것으로 판단할 수 있는데, 본 연구 결과 1개월, 2개월, 그리고 4개월의 시차를 두고 양(+)의 영향을 미치는 것을 알 수 있다. 즉, 행정구별 아파트에 대한 인터넷 검색 활동이 증가하게 되면 1~4개월 정도의 시차를 두고 해당 행정구의 아파트매매 거래량이 증가할 수 있음을 본 분석 결과는 시사하고 있다.

2. 패널 VAR 모형 분석

1) IRFs(Impulse - Response Functions)

인터넷 검색활동과 주택 가격 및 거래량 간의 동적 관계를 다른 각도에서 살펴보기 위하여 본 연구는 패널 VAR 분석을 실시[16], IRFs(충격반응분석) 그래프를 통해 특정 변수에 충격이 주어졌을 경우 다른 변수들의 반응을 살펴보았다. <그림 2>는 APG, AVG, 그리고 ISG에 대한 IRFs 그래프를 제시하고 있으며, 중심 실선은 IRFs 선을 의미하며, 실선 주변 점선은 95% 신뢰 수준(confidence intervals)을 의미한다. 그래프 오른쪽 상단 첫 번째 그래프와 두 번째 그래프는 ISG에 1만큼 충격이 주어졌을 때, APG 및 AVG의 반응을 각각 보여주고 있다. 상기 Arellano-Bond 동적 패널 모형 추정

16) 본 연구는 패널 VAR 모형 분석을 위하여 Ryan Decker의 Stata 코드를 사용하였음을 밝힌다. Ryan Decker의 Stata 코드는 Inessa Love의 PVAR.ado 코드(Love and Zicchino, 2006)를 수정한 것으로서, 고정 효과와 시차가 적용된 종속 변수 간의 내재적 상관관계 문제를 해결하였으며, FEVDs 분석의 기간 설정 및 그래프 조정에 많은 유연성을 제공하고 있다(Fort et al., 2013).



<그림 2>*APG*, *AVG*, *ISG* 간 IRFs 그래프

결과에서와 동일하게 *ISG* 충격은 *APG*와 *AVG*에 양(+)의 영향력을 미치고 있으며, 특히 *APG*의 경우에는 시차 1차 시점에서 영향력이 크게 작용하고 있음을 알 수 있다. 반면, *AVG*의 경우, 양(+)의 관계와 음(-)의 관계가 반복적으로 나타나고 있으나, 양(+)의 영향 중 시차 1차 시점에서 영향력이 가장 큰 것을 확인할 수 있다. 즉, IRFs 그래프로만 볼 때, 인터넷 검색 활동이 주택 가격 및 거래량 증가에 가장 큰 영향력을 미치는 시점은 1개월 후로 볼 수 있다. 이러한 결과는 인터넷 검색 후 주택을 구입하러 현장에 나가 중개업자를 만나고 매물을 확인한 후 실제 주택을 구입하기까지 1개월이 소요되는 현실이 반영된 것으로 해석할 수 있다.

한편, 상기한 바와 반대로 주택 가격 또는 거래량이 인터넷 검색 활동에 영향을 줄 수도 있을 것이므로 이에 대한 IRFs 결과를 살펴보면, *APG* 및 *AVG* 모두 시차 1, 2 시점에 *ISG*에 양(+)의 영향을 주고 있음을 알 수 있다. 이는 주택 가격 또는 거래량이 인터넷 검색 활동에 영향을 미친다고 가정할 때, 주택 가격 또는 거래량이 인터넷 검색 활동에 1~2개월의 시차를 두고 영향을 미치고 있음을 시사하는 결과로 해석된다. 또한, *APG*와 *AVG* 간의 관계를 살펴보면, *AVG*의 충격은 전체 시점에 걸쳐 *APG*에 양(+)의 영향을 미치고 있는 반면, *APG*의 충격은 1개월 시차에서 *AVG*에 양(+)의 영향을 미치고 2개월부터는 음(-)의 관계로 전환되는 것을 볼 수 있다. 이러한 결과는 기존 주택 가격과 거래량에 대한 선행 연구에서의 결과와 유사함을 알 수 있다(김대원・유정석, 2013b).

www.kci.go.kr

2) FEVDs(Forecast Error Variance Decompositions)

<표 7>은 *APG*, *AVG*, 그리고 *ISG* 각각에 대한 FEVDs 분석 결과를 보여주고 있으며, 2개월, 4개월, 6개월 후로 구분된 경과시점에 따라 다른 변수에 의해 받는 상대적 영향의 정도를 제시하고 있다.

<표 7> *APG*, *AVG*, *ISG*에 대한 FEVDs 결과

경과 시점	2m	4m	6m
FEVDs of APG			
AVG	0.0189	0.0374	0.0444
ISG	0.0263	0.0331	0.0344
FEVDs of AVG			
APG	0.0547	0.0656	0.0723
ISG	0.0096	0.0204	0.0272
FEVDs of ISG			
APG	0.0074	0.0153	0.0148
AVG	0.0006	0.0046	0.0096

*APG*에 대한 FEVDs 분석 결과는 초기 2개월까지는 *ISG*의 영향력이 더 큰 것으로 나타났으나, 시간이 경과함에 따라 *AVG*의 영향력이 더욱 커지는 것을 확인할 수 있다. 반면, *AVG*에 대한 FEVDs 분석 결과, 기간 내내 *ISG* 보다는 *APG*가 더 큰 영향력을 미치고 있음을 알 수 있다. 따라서 인터넷 검색 활동의 증가는 단기적으로 주택 가격 상승에 높은 영향력을 주고 있음을 분석 결과는 시사하고 있다.

3. 패널 Granger 인과성 검정

<표 8>은 주택 가격과 거래량, 그리고 인터넷 검색 활동 간의 인과관계의 방향성 확인을 위한 패널 Granger 인과성 검정 결과를 제시하고 있다. 분석 결과, *ISG*는 *APG*에 대하여 모든 시차에서 인과성을 갖는 반면, *AVG*에 대하여는 모든 시차에서 인과성을 갖지 않는 것으로 나타났다. 또한, *APG*와 *AVG*는 모든 시차에서 상호 인과성을 갖는 것으로 확인되었다. 한편, *APG*와 *AVG*는 *ISG*에 대하여 최초 시차 1차와 시차 2차 시점에서는 인과성을 갖지 않다가 시차 3차 이후로 인과성을 갖는 것으로 나타났다.

패널 Granger 인과성 검정 결과를 종합적으로 판단할 때, 인터넷 검색 활동은 주택 가격에는 인과적 영향력을 미치고 있으나, 거래량에는 인과적 영향력을 미친다고 언급하기는 어려운 것으로 보이며, 가격과 거래량의 증가 또는 감소가 3개월 정도 이어지면 이는 인터넷 검색 활동에도 영향을 주고 있는 것으로 보인다. 또한 가격과 거래량은 상호 인과적 영향력을 미치고 있는 것으로 나타나고 있다. 따라서 패널 Granger 인과성 검정 결과만으로 볼 때, 인터넷 검색 활동은 주택 가격에 인과적 영향을 미치며, 주택 가격과 거래량은 서로 상호 인과성을 갖게 되고, 다시 가격과 거래량의 증가는 인터넷 검색 활동의 증가로 이어지게 되어 서로 순환적 인과성을 갖는 것으로 분석 결과는 보여 주고 있다.

<표 8> *APG*, *AVG*, *ISG* 에 패널 Granger 인과성 검정 결과

Variables	Null Hypothesis	W - Stat.			
		lags: 1	lags: 2	lags: 3	lags: 4
AVG	does not Granger cause *APG*	1.5659 (1.75)*	3.7163 (3.79)***	5.5593 (4.56)***	6.3391 (3.52)***
ISG		4.4584 (11.22)***	7.8950 (12.92)***	8.5351 (9.66)***	8.5510 (6.63)***
APG	does not Granger cause *AVG*	3.1976 (7.05)***	8.0794 (13.72)***	9.5824 (11.93)***	11.5002 (11.67)***
ISG		0.7337 (-0.96)	2.6279 (1.22)	3.5432 (0.75)	3.9029 (-0.42)
APG	does not Granger cause *ISG*	1.0662 (0.10)	2.6918 (1.37)	5.1576 (3.63)***	6.9410 (4.19)***
AVG		0.5891 (-1.42)	1.2634 (-1.81)*	2.3766 (3.63)***	3.2014 (4.19)***

주 : 괄호안 수치는 Zbar-statistics를 의미(* p<0.1, ** p<0.05, *** p<0.01)

www.kci.go.kr

Ⅴ. 결론 및 시사점

주택 시장에서의 인터넷 검색 활동은 잠재 수요자들이 희망하는 주택에 대한 정보를 사전적으로 획득하기 위한 목적으로 활용되는 수단으로서, 인터넷 검색 활동의 흔적은 집단적 사고이자 장래 구매 수요의 반영이라 할 수 있다. 이러한 맥락에서 인터넷 검색 활동과 주택 가격 및 거래량 간에 유의미한 관계가 확인된다면, 이는 곧 주택 시장에서의 인터넷 검색 활동이 주택 구매 수요의 대리 변수로서 미래 주택 시장을 예측할 수 있는 지표로 활용될 수 있음을 의미한다. 따라서 본 연구는 국내 최대 인터넷 점유율을 차지하고 있는 검색 엔진인 네이버에서 제공하는 네이버 트렌드 자료를 인터넷 검색 활동의 대리 변수로 사용하고, 아파트매매가 지수와 아파트매매거래량 지수를 각각 주택 가격 및 거래량의 대리 변수로 사용하여 변수들 간의 동적 관계를 분석하였다.

서울시 23개 행정구를 대상으로 구성된 패널 자료를 사용하여 변수들 간의 동적 관계를 다각적으로 검정하고자 Arellano-Bond 동적 패널 모형, 패널 VAR 모형, 그리고 패널 Granger 인과성 검정 순으로 실시하였다. Arellano-Bond 동적 패널 모형 분석 결과, 인터넷 검색 활동은 일정 기간의 시차를 두고 주택 가격 및 거래량의 양(+)의 영향을 미치고 있는 것으로 나타났으며, 인터넷 검색 활동이 주택 가격과 거래량에 영향을 미치는 시차에서는 경미한 차이가 존재함을 확인할 수 있었다. 한편, 패널 VAR 모형을 통한 IRFs 분석 결과에서도 인터넷 검색 활동은 주택 가격과 거래량에 양(+)의 영향을 주고 있는 것으로 확인되었으며, 주택 가격 및 거래량 모두에서 시차가 1차일 때 가장 큰 양(+)의 영향력을 미치는 것으로 나타나 인터넷 검색 활동이 주택 가격과 거래량에 단기적 양(+)의 영향력을 미치는 것을 확인할 수 있었다. 또한, FEVDs 분석 결과, 인터넷 검색 활동은 주택 거래량 보다는 가격에 더욱 큰 비중으로 영향을 주고 있는 것을 확인할 수 있었다. 변수들 간 인과성의 방향을 확인하기 위한 패널 Granger 인과성 검정 결과, 인터넷 검색 활동과 주택가격, 그리고 주택 가격과 거래량은 상호 인과성을 갖는 것으로 나타났으나, 인터넷 검색 활동은 주택 거래량에는 인과성을 갖지 않는 것으로 나타났다. 특이한 점은 주택 가격과 거래량이 초기에는 인터넷 검색 활동에 인과성을 갖지 않으나, 시차 3차 이후부터는 인과성을 갖는 것으로 나타났다. 따라서 인터넷 검색 활동은 주택 가격에 영향을 미치고, 주택 가격과 거래량은 상호 인과성을 갖으며, 일정 시간이 경과하면 주택 가격과 거래량이 다시 인터넷 검색 활동에 영향을 미치는 상호 순환적인 인과성을 갖고 있음을 확인할 수 있었다.

인터넷 검색 활동은 잠재 주택 수요자들의 구매 수요를 반영한 결과라 할 수 있으므로, 인터넷 검색 활동이 주택 가격 및 거래량에 유의미한 영향력을 미치고 있다는 사실은 인터넷 검색 활동을 주택 시장의 예측 지표로서 활용 가능함을 시사하는 것으로 보인다. 즉, 잠재 주택 수요자들은 그들이 희망하는 주택을 구입하기에 앞서 인터넷 검색을 실시할 것이고, 해당 주택에 대한 구매 활동을 시도할 것이다. 이러한 활동의 증가는 곧 주택 가격 및 거래량의 증가로 이어지게 될 것이므로, 결국 특정 지역 주택에 대한 인터넷 검색 활동의 증가는 해당 지역 주택 가격 및 거래량의 증가로 나타나게 될 것이다.

또한 본 연구에서의 결과는 인터넷 검색 활동이 주택 수요 예측과 관련되어 다양하게 응용 및 확장될 수 있음을 시사하기도 한다. 가령, 지역별 인터넷 검색 활동 자료를 사용하여 지역별 주택 관심도를 지수화하게 되면 지역별 단기간 주택 가격 및 거래량 예측 지수로써 활용이 가능할 것이다. 또한 인터넷 검색 활동과 지역별 소비심리지수 등과의 관련성이 입증된다면 인터넷 검색 활동을 소비심리지수의 대용변수로도 활용이 가능할 것이다.

본 연구가 주택 시장 예측 지표로서 인터넷 검색 활동의 활용 가능성을 확인한 기초적 연구였다면, 추후 후속 연구에서는 이를 확장한 다양한 분석과 응용이 이루어질 수 있으리라 기대한다. 특히, 인터넷의 활용이 과거 PC를 통해서만 이루어졌던 것에 반해, 최근에는 모바일 기기를 통한 인터넷 활용이 증가하고 있는 추세인 점을 반영한다면 보다 확장적인 연구가 이루어질 수도 있을 것이다. 또한 인터넷 검색 활동 변수뿐만 아니라 주택 가격 및 거래

량에 영향을 줄 수 있는 다양한 변수를 추가하여 인터넷 검색 활동과 기타 변수들이 주택 시장에 미치는 영향을 동시에 확인하게 된다면, 주택 시장 예측 지표로서 인터넷 검색 활동에 대한 논의가 더욱 활발해질 수 있을 것이다.

▌참고문헌 ▐

1. 김대원 · 유정석, 2014, "패널분석을 이용한 지역별 미분양률 결정모형 개발 및 적용 연구", 주택연구 22(1), 한국주택학회: 149~178
2. _____, 2013a, "주택가격에 대한 심리적 태도가 주택 매매 거래량에 미치는 영향 분석", 주택연구 21(2), 한국주택학회: 73~92
3. _____, 2013b, "주택가격과 거래량 간 관계 형성에 매도-매수 선호시점이 미치는 영향 연구", 부동산학보 54, 한국부동산학회: 229~243
4. 김수도 · 조환규, 2012, "온라인 게시글의 조회수 분석을 통한 인기도 예측", 한국콘텐츠학회 논문집 12(4), 한국콘텐츠학회: 40~51
5. 박정은 · 홍창범 · 정진영 · 송정길, 2009, "검색어를 통한 유행성 독감 감지 및 예측 시스템", 한국인터넷정보학회 추계 학술대회 20, 한국인터넷정보학회: 205~210
6. 이수욱, 2008, "논단: 수요자 중심의 부동산시장 기반조성", 건설경제 55, 국토연구원: 59~68
7. 정의철, 2010, "소비자 심리가 주택시장에 미치는 영향 분석 - 주택매매가격을 중심으로", 부동산학연구 16(3), 한국부동산분석학회: 5~20
8. 최영걸 · 이창무 · 최막중, 2004, "서울시 주택시장에서 작동되는 가격기대심리에 관한 실증연구", 국토계획 39(2), 대한국토 · 도시계획학회: 131~141
9. 한송이 · 김종우 · 이홍주 · 오재훈, 2012, "인터넷 상점의 클릭스트림 데이터를 활용한 실시간 구매 확률 예측", *Entrue Jounal of Information Technology* 11(1), (주)엘지씨엔에스: 101~110
10. Arellano, Manuel, Stephen Bond, 1991, "Some Tests of Specification for panel data: Monte Carlo Evidence and an Application to Employment equations", *The Review of Economic Studies* 58(2), Oxford University Press: 277~297
11. Arellano, Manuel, Olympia Bover, 1995, "Another Look at the Instrumental Variable Estimation of Error-components Models", *Journal of Econometrics* 68(1), Elsevier: 29~51
12. Baryla, Edward A., Leonard V. Zumpano, Harold W. Elder, 2000, "An Investigation of Buyer Search in the Residential Real Estate Market under Different Market Conditions", *Journal of Real Estate Research* 20(1), American Real Estate Society: 75~91
13. Beracha, Eli, M. Babajide Wintoki, 2013, "Forecasting Residential Real Estate Price Changes from Online Search Activity", *Journal of Real Estate Research* 35(3), American Real Estate Society: 283~312
14. Brownstein, John S., Clark C. Freifeld, Lawrence C. Madoff, 2009, "Digital Disease Detection–Harnessing the Web for Public Health Surveillance", *New England Journal of Medicine* 360(21), Massachusetts Medical Society: 2153~2157
15. Choi, Hyunyoung and Hal Varian, 2012, "Predicting the Present with Google Trends", *Economic Record* 88(s1), Wiley: 2~9
16. Choi, In, 2001, "Unit Root Tests for Panel Data", *Journal of International Money and Finance* 20(2), Elsevier: 249~272
17. Corley, Courtney, Armin R. Mikler, Karan P. Singh, and Diane J. Cook, 2009, "Monitoring Influenza Trends through Mining Social Media", *BIOCOMP*, BIOCOMP: 340~346
18. Cumby, Robert E. and John Huizinga, 1992, "Testing the Autocorrelation Structure of Disturbances in Ordinary Least Squares and Instrumental Variables Regressions", *Econometrica* 60(1), Wiley: 185~195
19. Dumitrescu, Elena-Ivona, and Christophe Hurlin, 2012, "Testing for Granger Non-causality in Heterogeneous Panels", *Economic Modelling* 29(4), Elsevier: 1450~1460
20. Ettredge, Michael, John Gerdes, and Gilbert Karuga, 2005, "Using Web-based Search Data to Predict Macroeconomic Statistics", *Communications of the ACM* 48(11), Association for Computing Machinery: 87~92
21. Fort, Teresa C., John Haltiwanger, Ron S. Jarmin, and Miranda Javier, 2013, "How Firms Respond to Business Cycles: The Role of Firm Age and Firm Size", *IMF Economic Review* 61(3), Palgrave Macmillan: 520~559
22. Ginsberg, Jeremy, Matthew H. Mohebbi, Rajan S. Patel, Lynnette Brammer, Mark S. Smolinski, and Larry Brilliant, 2009, "Detecting Influenza Epidemics Using Search Engine Query Data", *Nature* 457(7232), Nature Publishing Group: 1012~1014
23. Goel, Sharad, Jake M. Hofman, Sébastien Lahaie, David M. Pennock, and Duncan J. Watts, 2010,

www.kci.go.kr

"Predicting Consumer Behavior with Web Search", *Proceedings of the National Academy of Sciences* 107(41), Stanford University's High Wire Press: 17486~17490

24. Hadri, Kaddour, 2000, "Testing for Stationarity in Heterogeneous Panel Data", *The Econometrics Journal* 3(2), Wiley: 148~161
25. Hulth, Anette, Gustaf Rydevik, and Annika Linde, 2009, "Web Queries as a Source for Syndromic Surveillance", *PloS one* 4(2), PLOS: e4378
26. Im, Kyung So, M. Hashem Pesaran, and Shin Yong-cheol, 2003, "Testing for Unit Roots in Heterogeneous Panels", *Journal of Econometrics* 115(1), Elsevier: 53~74
27. Juster, F. Thomas, 1966, "Consumer Buying Intentions and Purchase Probability: An Experiment in Survey Design", *Journal of the American Statistical Association* 61(315), Taylor & Francis: 658~696
28. Levin, Andrew, Chien-Fu Lin, and Chia-Shang James Chu, 2002, "Unit Root Tests in Panel Data: Asymptotic and Finite-sample Properties", *Journal of Econometrics* 108(1), Elsevier: 1~24
29. Love, Inessa and Lea Zicchino, 2006, "Financial Development and Dynamic Investment Behavior: Evidence from Panel VAR", *The Quarterly Review of Economics and Finance* 46(2), Elsevier: 190~210
30. Maddala, Gangadharrao S., and Shaowen Wu, 1999, "A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test", *Oxford Bulletin of Economics and statistics* 61(S1), Wiley: 631~652
31. Morrison, Donald G., 1979, "Purchase Intentions and Purchase Behavior", *The Journal of Marketing*, American Marketing Association: 65~74
32. Morwitz, Vicki G., Joel H. Steckel, and Alok Gupta, 2007, "When Do Purchase Intentions Predict Sales?", *International Journal of Forecasting* 23(3), Elsevier: 347~364
33. Pelat, Camille, Clement Turbelin, Avner Bar-Hen, Antoine Flahault, and Alain-Jacques Valleron, 2009, "More Diseases Tracked by Using Google Trends", *Emerging Infectious Diseases* 15(8), CDC: 1327
34. Polgreen, Philip M., Yiling Chen, David M. Pennock, Forrest D. Nelson, and Robert A. Weinstein, 2008, "Using Internet Searches for Influenza Surveillance", *Clinical Infectious Diseases* 47(11), Oxford Journals: 1443~1448
35. Preis, Tobias, Daniel Reith, and H. Eugene Stanley, 2010, "Complex Dynamics of Our Economic Life on Different Scales: Insights from Search Engine Query Data", *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 368(1933), Royahl Society Publishing: 5707~5719
36. Radinsky, Kira, Sagie Davidovich, and Shaul Markovitch, 2008, "Predicting the News of Tomorrow Using Patterns in Web Search Queries", *Proceedings of the 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology* 1, IEEE/WIC/ACM International Conference: 363~367
37. Rangaswamy, Arvind, C Lee Giles, Silvija Seres, 2009, "A Strategic Perspective on Search Engines: Thought Candies for Practitioners and Researchers", Journal of Interactive Marketing, 23(1), Elsevier: 49~60
38. Suhoy, Tanya, 2009, "Query Indices and a 2008 Downturn: Israeli Data", *Bank of Israel*, Bank of Israel: 1~32
39. Valdivia, Antonio and Susana Monge-Corella, 2010, "Diseases Tracked by Using Google Trends, Spain", *Emerging Infectious Diseases* 16(1), CDC: 168
40. Vosen, Simeon and Torsten Schmidt, 2011, "Forecasting Private Consumption: Survey-based Indicators vs. Google Trends", *Journal of Forecasting* 30(6), Wiley: 565~578
41. Wu, Lynn and Erik Brynjolfsson, 2013, "The Future of Prediction: How Google Searches Foreshadow Housing Prices and Sales", Economics of Digitization
42. 네이버 트렌드, http://trend.naver.com
43. BizSpring Inc., http://www.internettrend.co.kr
44. Google 트렌드, http://www.google.co.kr/trends/
45. Google Flu Trends, http://www.google.org/flutrends/

• 논문 접수일 : 2014. 4. 11
• 논문 수정일 : 2014. 6. 2
• 게재 확정일 : 2014. 6. 20

www.kci.go.kr