

이슈보고서 26-03

ISSUE
REPORT

자본시장 심리지수 시리즈 3

투자자 심리와 주식시장의 관계 고찰

김민기

투자자 심리와 주식시장의 관계 고찰

저자 김민기*

본 연구는 트랜스포머 기반 대형언어모형(LLM)을 국내 경제·증권 뉴스에 파인튜닝하여 구축한 자본시장 심리지수(CMSI)의 특징을 주식시장과 연계하여 분석한다.

2010년 1월부터 2025년 3월까지 산출한 지표는 해외 주식시장 심리지수, 국내 주식시장 수익률, 주식 파생상품시장에서 나타나는 투자자 공포 및 헤지 수요, 파생형 ETF에 대한 투기적 수요와 유의하게 연동되고, 주식 현물시장의 거래활동 및 IPO 관련 변수와도 높은 상관성을 보였다. 특히 CMSI는 개인투자자의 거래와 신규 상장 기업의 상장 프리미엄과 밀접하게 연결되어, 투자심리의 영향이 비정보거래자 및 정보 비대칭 하에서 확산되는 특성을 보여준다.

아울러 주식 수익률과 CMSI의 관계를 분석한 결과, 시장 전체 수익률 측면에서는 중기 예측력이 제한적이었으나, 단기적으로는 심리 변화와 같은 방향으로 시장수익률을 통계적으로 유의하게 예측하였다. 또한 개별주식 단위의 횡단면 분석에서는 낙관 국면에서의 과대평가, 비관 국면에서의 과소평가가 관찰되었으며, 이러한 패턴은 가치평가가 어려운 기업이나 투기적 성향이 강한 주식에서 더욱 뚜렷했다. 이는 본 심리지수가 개별주식 수익률의 횡단면 구조를 설명하는 데 유효한 정보를 담고 있는 변수가 될 수 있음을 시사한다.

종합하면, 본고는 새로운 방법론을 적용해 구축한 심리지표의 특성과 활용 가능성을 주식시장 관점에서 검증하였고, 그 결과 CMSI가 주식 가격의 변화와 수익률 구조를 설명하는 데 유용한 지표가 될 수 있음을 확인하였다. 비정형 데이터에 대한 활용과 이를 효과적으로 처리하여 정보를 추출하는 시도가 증가하는 가운데, 본 연구는 대형언어모형이 시장에 내재된 투자심리를 정교하게 계량화하는 데 효과적인 도구가 될 수 있음을 보여주는 사례가 될 것이다.

* 본고의 견해와 주장은 필자 개인의 것이며, 자본시장연구원의 공식적인 견해가 아님을 밝힙니다.

자본시장실 연구위원 김민기 (mkkim@kcmi.re.kr)

** 발행: 2026년 1월 5일

1. 논의 배경

자본시장에서 투자자 심리(investor sentiment)란 합리적 기대와 구별되는, 투자자의 감정적 상태와 편향된 기대가 집단적으로 형성된 상태를 의미한다. 전통적인 효율적 시장가설은 모든 가용 정보가 신속하고 정확하게 가격에 반영된다고 전제하지만, 실제 금융시장은 버블, 급락, 과잉 반응 등 합리적인 관점에서 설명하기 어려운 현상을 반복적으로 보여 왔다. 이러한 현상은 전통적 이론의 한계를 드러내었고, 이에 따라 투자자의 편향된 기대와 감정적 요인이 의사결정 과정에 체계적으로 영향을 미친다는 점을 강조하는 행동경제학(behavioral economics)이 확산되었다.

이후 다양한 실증 연구를 통해 투자자 심리가 단순히 투자 의사결정에 그치지 않고, 자산 가격의 형성과 변동에도 중요한 역할을 한다는 사실이 밝혀졌다. 그 결과 오늘날 연구와 실무 모두에서 투자자 심리가 시장의 움직임을 이해하고 예측하는 데 필수적인 변수로 자리매김하고 있다.

심리가 금융시장에 중대한 영향을 미친다는 점에서, 관측되지 않는 투자자 심리를 계량적으로 측정하려는 시도가 꾸준히 이어져 왔다. 초기 연구는 개별 시장 지표를 심리 수준의 대용치로 활용하거나(Brown & Cliff, 2004), 이를 종합해 시장 전반의 심리지수를 구축하였다(Baker & Wurgler, 2006). 이후에는 뉴스 기사, SNS, 검색 빈도 등 비정형 데이터가 활용되었으며(Tetlock, 2007; Bollen et al., 2011; Da et al., 2015), 이어 기계 학습 기법을 접목한 고도화된 분석도 활발히 진행되고 있다(Ke et al., 2019).

최근에는 자연어 처리(Natural Language Processing: NLP)에 특화된 트랜스포머(transformer) 기반 대형언어모형(Large Language Model: LLM)의 등장으로 심리지수 구축 방법론이 새로운 전환점을 맞이하였다. LLM은 방대한 텍스트 데이터를 효과적으로 처리해 단순 긍정·부정 판별을 넘어 문맥적 의미와 감정의 강도를 세밀하게 파악할 수 있다. 이를 통해 투자자 심리를 보다 정교하게 추정하고, 금융 변수와의 연관성을 한층 높은 수준에서 분석할 수 있게 되었다.

본 연구는 이러한 배경을 바탕으로, 대규모 뉴스 기사 데이터를 학습한 LLM 기반 노성호(2026)의 자본시장 심리지수(Capital Market Sentiment Index: CMSI)의 특성을 주식시장 관점에서 분석한다. 구체적으로, (i) CMSI와 주요 주식시장 변수 간의 연관성을 살펴보고, (ii) 투자자 심리가

주식 수익률에 어떤 영향을 미치는지를 규명한다. 이를 통해 심리지표의 속성을 보다 명확히 이해하고, 시장 심리가 주가 형성과 변동에 미치는 효과를 평가하고자 한다. 거시경제적 관점에서 CMSI와 경제 변수 간의 관계를 분석한 장보성(2026)의 연구와 달리, 본 연구는 주식시장이라는 미시적 영역에 초점을 맞추어 CMSI의 경제적 의미를 논의한다.

본 보고서의 구성은 다음과 같다. II 장에서는 주식시장에서 투자자 심리를 분석한 기존 연구와 심리지수 구축 방법론을 돌아보고, 본 연구의 차별성을 논의한다. III 장에서는 CMSI와 주요 주식 시장 변수 간의 관계를 분석해 심리지수의 속성을 파악한다. IV 장에서는 투자자 심리와 주식 수익률 간의 관계를 시계열 및 횡단면 분석을 통해 살펴본다. 마지막으로 V 장에서는 주요 분석 결과를 종합적으로 평가하며 보고서를 마무리한다.

II. 선행연구

투자자 심리와 주식시장의 관계를 탐구하는 연구는 전통적인 효율적 시장 가설(efficient market hypothesis)의 한계를 보완하는 중요한 학술적 축으로 자리잡고 있다. 초기 연구는 금융시장에서 관찰되는 자산 가격의 괴리(mispricing)가 시장 참여자의 심리 변동에 의해 체계적으로 설명될 수 있는지를 실증적으로 규명하는 데 주력하였다. 대표적으로 Brown & Cliff(2005)는 설문 기반 심리지표를 이용해 주식시장 전반의 가격 괴리가 투자자 심리와 밀접하게 연관되어 있음을 보였으며⁰¹, Baker & Wurgler(2006, 2007)는 투자자 심리의 영향이 특히 가치평가가 어렵거나(hard-to-value), 투기적 성격이 강하거나, 차익거래(arbitrage)가 제약된 종목에서 두드러진다는 사실을 제시하였다.⁰² 이러한 결과는 투자자의 심리적 요인이 단순한 단기적 잡음(noise)에 그치는 것이 아니라, 주가 형성 과정 전반에 구조적으로 작용하는 핵심 변수임을 시사한다.

이후 연구는 투자자 심리와 주식시장의 관계를 더욱 정밀하게 규명하기 위해 데이터와 방법론을 확장해 왔다. Tetlock(2007)은 월스트리트저널(WSJ) 금융 뉴스의 어조를 정량화하였고,

01 다우존스 지수의 내재가치(intrinsic value)를 추정하여 실제 시장가격과의 차이를 계산하고, 투자자 심리가 긍정적일수록 내재가치 대비 실제 가치가 과대평가되는 것을 실증하였다.

02 가치평가 및 차익거래가 어렵거나 투기적 요소가 강한 종목의 경우, 투자자 심리가 긍정(부정)적일수록 고평가(저평가)되는 경향이 강해지는 것으로 나타난다.

Bollen et al.(2011)은 트위터 등 소셜미디어에 나타난 대규모 이용자의 정서를 분석하였다. Da et al.(2015)은 경제 및 금융 관련 긍·부정 단어의 검색 빈도 차이를 활용해 투자자의 불안 수준을 측정하고 주식시장 수익률과 연결 지었다. 이처럼 비정형 데이터를 활용한 연구들은 설문이나 시장 기반 지표가 지니는 대표성의 한계를 보완한다는 점에서 의의가 있으나, 대부분 사전에 정의된 긍·부정 어휘 목록에 의존해 심리를 측정했다는 점에서 방법론적 유연성, 문맥 이해력 측면에 일정한 제약이 존재했다.⁰³

한편, 기계 학습과 자연어 처리(NLP) 기술의 발전은 투자자 심리 측정 방식에 획기적인 전환점을 마련하였다(Heston & Sinha, 2017; Ke et al., 2019).⁰⁴ 새로운 알고리즘은 기존의 어휘 기반 접근보다 더 높은 예측력을 보이는 심리지표 구축을 가능하게 했으며, 특히 트랜스포머 구조의 LLM은 텍스트의 문맥적 의미와 감정의 뉘앙스를 정밀하게 포착할 수 있는 도구로 자리매김했다. 이에 따라 LLM 기반 심리지표를 금융시장 분석에 적용하는 연구도 빠르게 확산되고 있다(Kim et al., 2024; Kirtac & Germano, 2025). 본 연구에서 활용하는 노성호(2026)의 자본시장 심리지수(CMSI) 역시 대규모 국내 경제·증권 뉴스를 학습한 BERT 기반 트랜스포머 모델을 통해 텍스트의 맥락적 감성 정보를 계량화한 결과물이다.

국내에서도 시장 심리를 계량화하려는 시도가 점차 다양해지고 있다. 시장 지표를 활용한 연구(변진호·김근수, 2013; 박하연, 2024), 감성 사전 기반의 접근(송민채·신경식, 2017), 최신 NLP 기법을 적용한 연구(서범석 외, 2022) 등이 그것이다. 이러한 흐름 속에서 본 보고서는 트랜스포머 기반 LLM으로 구축된 심리지수를 주식시장 분석에 직접 적용함으로써, 국내 문헌에 새로운 시사점을 제공한다. 특히 기존 국내 선행연구(강장구 외, 2013; 변진호·김근수, 2013; 이효정, 2019; 박하연, 2024)와 달리, 최신 NLP 기반 심리지표를 주식시장 분석에 접목했다는 점에서 차별성을 갖는다.⁰⁵

03 이러한 접근을 사전 기반(lexicon-based) 감성 분석이라 하며, 사회과학 분야에서 널리 쓰이는 Harvard IV-4, Lasswell Value Dictionary와 같은 감성 사전에 방대한 비정형 텍스트를 대입해 심리 수준을 측정한다. 이 방식은 지표를 빠르게 산출할 수 있다는 장점이 있으나, 텍스트 전체의 문맥(context)을 충분히 반영하지 못한다는 한계가 있다.

04 신경망(neural network) 모형(Heston & Sinha, 2017), 지도 학습(supervised learning) 모형(Ke et al., 2019) 등을 뉴스 기사 자료에 적용하여 심리 수준을 계량화하였다.

05 심리지수와 주식시장의 관계를 분석한 국내 선행연구는 주로 시장 지표를 기반으로 한 투자자 심리 대응치를 이용하였다.

III. 자본시장 심리지수와 주식시장 변수의 관계

본 장에서는 자본시장 심리지수(CMSI)가 기존의 투자자 심리 관련 주식시장 변수와 어떠한 관계를 보이는지 살펴봄으로써, CMSI가 지니는 특성을 실증적으로 파악하고자 한다. 본 연구에서 사용한 CMSI는 노성호(2026)가 제시한 지수로, Kim et al.(2024)의 방법론을 확장하여 파인튜닝(fine-tuning)된 BERT 모델을 통해 장기간의 경제·증권 뉴스로부터 산출한 것이다. 해당 지수는 2010년 1월부터 2025년 3월까지 일별·월별 단위로 구축되어 있으며, 본 보고서의 분석에는 주로 월별 지수를 사용한다.⁰⁶

먼저 CMSI를 해외에서 널리 인용되는 대표적 투자자 심리지표와 비교해 보자. 해외 연구에서는 다양한 심리지수가 제시되어 왔으나, 가장 잘 알려진 지표는 Baker & Wurgler(2006)의 미국 주식시장 심리지수(이하 BW 지수)이다. BW 지수는 여섯 가지 시장 심리 대용치를 주성분 분석(PCA)으로 결합하여 구축되며, 월별 변동성을 완화하고 추세 변화를 포착하기 위해 각 변수의 12개월 이동평균을 사용한다. 본 연구 역시 이러한 방법론적 접근을 참고하였다.

CMSI의 12개월 이동평균을 BW 지수와 비교한 결과, 두 지수는 변동폭의 크기에는 차이가 있으나 전체 분석기간에서 유의미한 상관성을 보인다(<그림 III-1> A 참조). 특히 2015년 전후 CMSI와 BW 지수 모두 하락 추세를 보였는데, 이는 당시 글로벌 금융 불안 등 대외 요인의 영향이 반영된 것으로 풀이된다.⁰⁷ 이후 CMSI는 국내 주식시장이 반도체 슈퍼사이클을 중심으로 높은 수익률을 기록하는 과정에서 빠르게 회복하는 모습을 보인다(<그림 III-1>의 B 참조).

BW 지수는 코로나19 팬데믹 시기에 가장 큰 변동성을 기록했으며, 원지수(BW)뿐 아니라 거시 요인을 제거한 BW+에서도 동일한 패턴이 관찰된다. 이는 BW 지수의 구성 방식에서 기인한 특성으로 보인다.⁰⁸ CMSI 역시 팬데믹 직후 상승한 뒤 급락하는 모습을 보였으며, 2022년 이후에는 국내 투자심리의 회복세가 재차 강화되는 것으로 나타난다.⁰⁹

06 본 절에서 사용한 시장 변수의 정의와 산출 방식은 부록 1에 제시되어 있다.

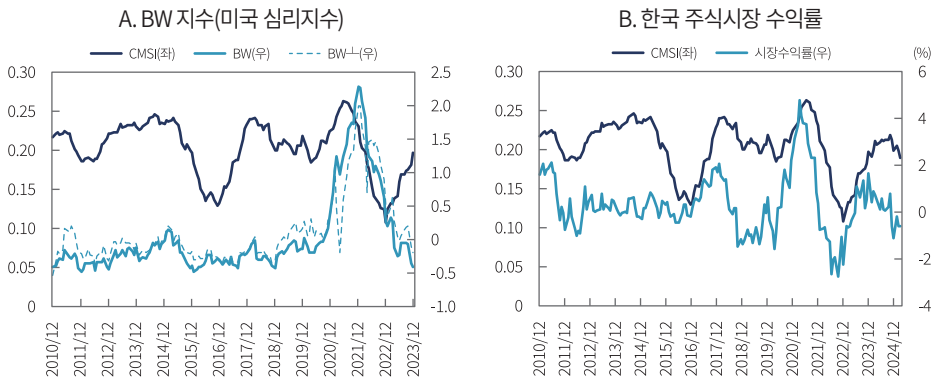
07 구체적으로 중국 주식시장 급락, 그리스 디폴트 위기, 미 연준의 양적완화 종료 후 첫 금리인상 가능성 등 투자심리가 약화되는 이벤트가 많이 발생했다.

08 코로나19 팬데믹(2020년 3월) 이후에 BW를 구성하는 여러 시장 지표들이 급격하게 변했고, 이러한 변동이 최소 12개월의 시차 동안 영향을 주기에 2022년까지 BW 지수의 큰 변화가 나타난 것으로 풀이된다.

09 BW 지수와 CMSI의 변동을 살펴보면, 코로나19 시기에는 글로벌 유동성 공급 확대와 정책 대응에 힘입어 투자심리가 빠르게 낙관적으로 전환된 것이 확인된다. 그러나 2022년 이후에는 급격한 인플레이션 압력, 이에 따른 주요국의 빠른 금리 인상,

또한 <그림 III-1>의 그림 B에서 확인되듯, CMSI는 국내 주식시장 수익률과 유의미한 공행성을 보인다. 이러한 관계는 2010년대 중반 이후 특히 뚜렷해졌으나, 그 이전 시기에는 다소 편차가 존재한다. 이는 뉴스 텍스트 기반 감성지표의 속성에 기인한 것으로¹⁰, 시기에 따라 뉴스가 반영하는 ‘기대와 전망’이 실제 주식시장의 ‘실현 수익률’과 반드시 동일한 방향으로 움직이지 않기 때문이다. 즉, 텍스트 기반 심리지표와 시장 성과 간에는 일정한 괴리가 발생할 수 있음을 알 수 있다.

<그림 III-1> CMSI와 해외 심리지수 및 한국 주식시장 수익률과의 관계



주 : 1) CMSI와 한국 주식시장 수익률은 과거 12개월 이동평균을 그림
 2) 그림 A의 BW+(점선)는 BW 지수 구성요소에 거시 요인을 통제한 심리지수
 3) 그림 B의 한국 주식시장 수익률은 코스피·코스닥 시가총액 가중평균 수익률 기준
 자료: 자본시장연구원, pages.stern.nyu.edu/~jwurgler/

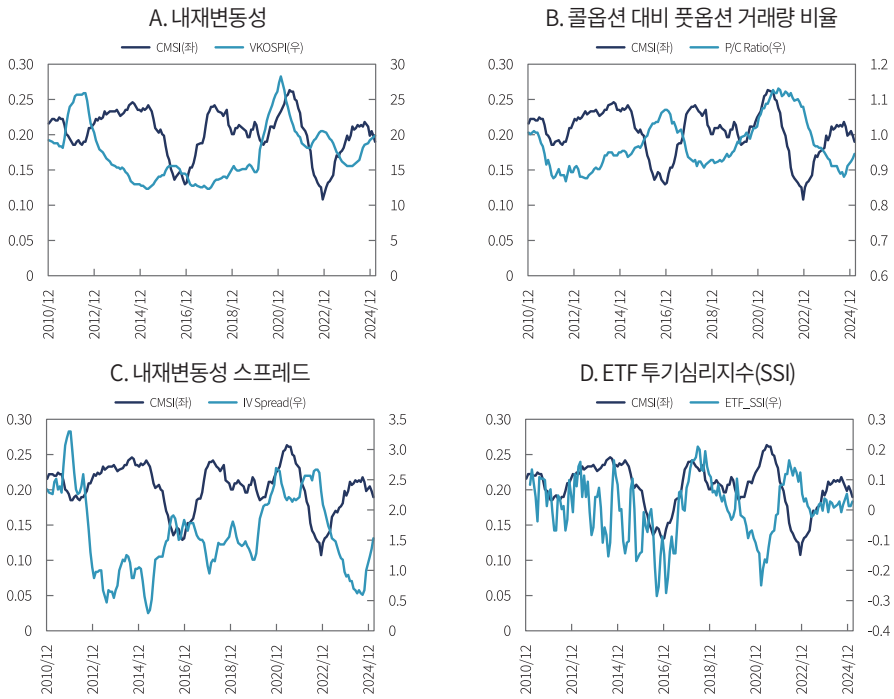
다음으로 주식 파생상품시장에서 추출한 간접적 투자자 심리 대응치와 CMSI의 관계를 살펴본다. 보고는 선행연구(Whaley, 2000; Bollen & Whaley, 2004; Brown & Cliff, 2004)를 참고하여 (i) 코스피200 지수 옵션의 내재변동성을 나타내는 VKOSPI 지수, (ii) 콜옵션 대비 풋옵션 거래량 비율(P/C Ratio), (iii) 풋옵션과 콜옵션 간 내재변동성 차이(IV Spread), 그리고 (iv) Davies(2022)의 방법론을 적용해 파생형 ETF의 초과수요 차이로 산출한 ETF 투기심리지수(Speculation Sentiment Index: SSI)를 분석한다. VKOSPI, P/C Ratio, IV Spread는 일반적으로 투자자 심리가 비관적일수록 상승하는 특성을 가지며, SSI는 투자자의 투기적 성향을 반영하는 지표다.

그리고 러시아-우크라이나 전쟁 발발과 같은 지정학적 충격이 연달아 발생하면서 글로벌 투자심리가 급격히 위축되는 흐름이 나타났다. CMSI 역시 이러한 전환점을 민감하게 반영하며 하락한 것으로 보인다.

10 2010년대 초중반은 글로벌 금융위기와 유로존 위기 이후 세계 경제가 점진적으로 회복세로 전환되던 시기였다. 이 과정에서 경기 반등 신호와 낙관적 전망이 누적되면서, LLM 기반 심리지표에는 긍정적 심리가 상대적으로 강하게 반영된 것으로 추정된다. 반면 같은 기간 한국 주식시장은 구조적 할인(discount) 요인 등에 의해 부진한 수익률을 보이고 있었다. 이로 인해 텍스트 기반 지표가 반영하는 심리 수준과 실제 시장의 성과 사이에 일정한 괴리가 발생할 것으로 풀이된다.

<그림 III-2>를 보면, CMSI는 이들 네 가지 주식 파생상품시장 변수와 일정 부분 상관성을 보이는 것으로 나타난다. 특히 코스피200 옵션 시장 변수와의 관계는 통계적으로 유의하며, 코로나 19 전후의 특수 국면을 제외하면 옵션 시장에서 비관적 심리가 강화될수록 CMSI가 하락하는 경향이 명확하게 나타난다.¹¹ ETF SSI 또한 유사한 움직임을 보이는데, SSI를 구성하는 주요 파생형 ETF가 코스피200을 추종하는 레버리지·인버스 상품으로 구성되어 있어, 일반 투자자의 방향성 투기 수요 변화가 CMSI의 변동과 일정 부분 연동되는 특성을 보인다. 한편, 옵션 시장 대응치와 마찬가지로 SSI 역시 코로나19 시기에는 CMSI의 상승과 달리 투기적 성향이 오히려 줄어드는 것으로 나타난다. 이는 당시 개인투자자의 단기 방향성 매매가 역추세적(contrarian) 성격이 강했기 때문으로 추정된다.¹²

<그림 III-2> CMSI와 주식 파생상품시장 변수의 관계

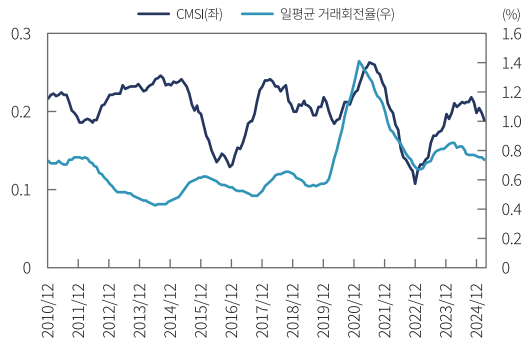


주 : 그래프는 원 시계열의 12개월 이동평균 기준
 자료: 자본시장연구원, KRX, DataGuide

- 11 코로나19 초기에는 옵션시장에서 비관적 심리가 급격히 강화되었으나, CMSI는 오히려 상승하는 모습을 보였다. 이는 팬데믹 이후 예상보다 빠르게 나타난 경기·시장의 회복 기대가 뉴스 텍스트에 강하게 반영된 반면, 실제 시장 참여자들은 하방 위험에 대비한 헤지 수요를 동시에 확대했던 결과로 해석된다.
- 12 코로나19 팬데믹 직후 주가 반등 국면에서 개인투자자의 인버스 상품 수요가 증가한 것으로 관찰된다(김준석·김민기, 2024). 해당 기간 CMSI 및 옵션 시장 심리지표 대응치의 변화를 고려하면, 당시 시장에는 상승 기대와 하방 리스크 회피가 공존하는 특수한 심리 구조가 있었음을 시사한다.

<그림 III-3>은 CMSI와 우리나라 주식시장 거래회전율의 관계를 시각화한 것이다. 그림에서 알 수 있듯이, 2010년대 초중반을 제외하면 CMSI는 거래회전율과 뚜렷한 양(+)의 상관성을 보인다. 최근으로 갈수록 투자심리가 상승할 때 시장 참여자의 거래가 활발해지는 경향이 강화되었으며, 특히 코로나19 국면에서 개인투자자의 대규모 시장 유입과 함께

<그림 III-3> 자본시장 심리지수와 주식시장 거래량



주 : 그래프는 원 시계열의 12개월 이동평균 기준
 자료: 자본시장연구원, DataGuide

이러한 패턴이 더욱 두드러졌다. 이는 심리가 낙관적일수록 과잉확신(overconfidence)이 강화되어 거래활동이 증가한다는 기존 연구의 설명과도 일치한다(Baker & Stein, 2004).

투자자 유형별로 보면, 개인 및 기관투자자의 거래회전율은 CMSI와 유의미한 양(+)의 상관성을 보인다(<그림 III-4>의 A1~A3 참조). 외국인투자자의 경우 시기에 따라 변동성이 존재하지만, 다른 요인을 통제한 회귀분석에서는 통계적 유의성은 낮더라도 거래회전율과 CMSI가 대체로 양(+)의 방향으로 연관된 것으로 나타난다.¹³

순매수를 통해 투자심리를 계량화했던 선행연구¹⁴를 참고하여, CMSI와 투자자 유형별 순매수의 관계를 살펴보면, 개인투자자와 외국인투자자의 순매수는 CMSI와 유의한 양(+)의 상관관계를 보이는 반면, 기관투자자는 음(-)의 상관관계를 보인다(<그림 III-4>의 B1~B3 참조). 이러한 결과는 투자자별 순매수에 영향을 미칠 수 있는 시장 변수를 통제해도 유지된다.¹⁵

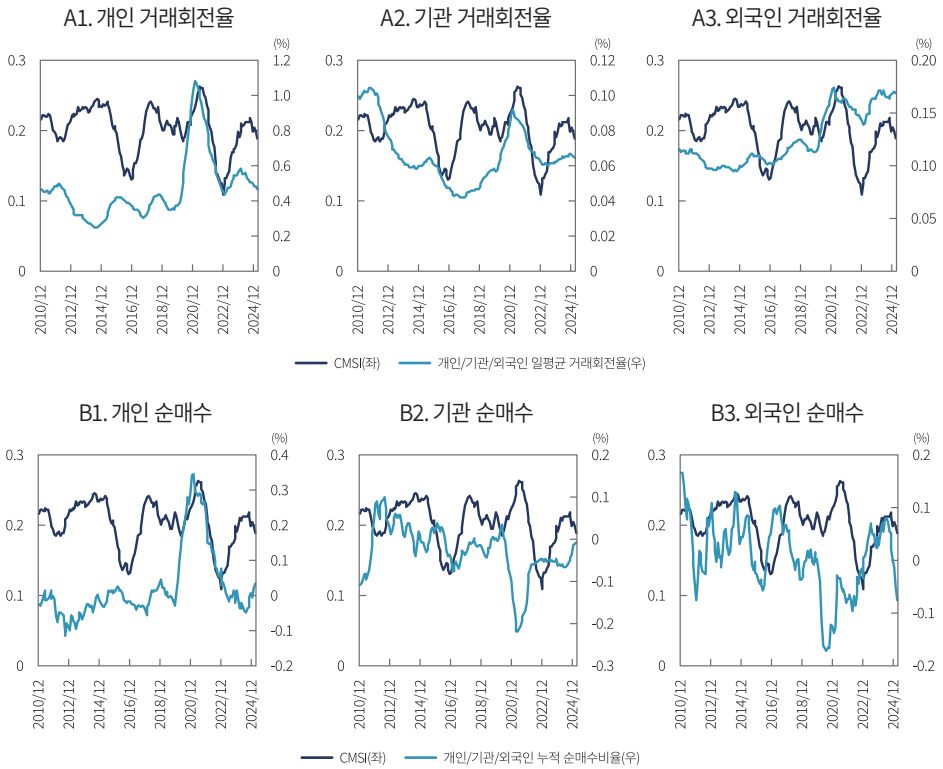
이러한 차별적 패턴은 투자자 유형별 심리 반응이 상이하다는 점을 시사한다. 개인투자자는 심리 변화에 가장 민감하게 반응하여, 투자심리가 낙관적인 시기에 거래와 순매수가 동시에 크게 증가하는 경향을 보인다. 외국인은 거래 규모의 변화는 제한적이지만 포지션 조절을 통해 심리 요인을 반영하는 것으로 해석된다. 반면 기관투자자는 심리 변화에 따라 거래는 늘리되 순매수는 오히려 줄이는 경향이 있는데, 이는 연기금 등 기관투자자가 시장 과열 국면에서 상대적으로 반대 포지션(contrarian stance)을 취하는 성향이 있음을 의미한다.

13 자세한 결과는 부록 2를 참조한다.

14 Barber et al.(2009)과 Kumar & Lee(2006)는 매수·매도 불균형을 통해 개별 투자자 수준의 심리를 측정한다.

15 자세한 결과는 부록 2를 참조한다.

<그림 III-4> 자본시장 심리지수와 투자자별 주식시장 거래활동



주 : 1) 그래프는 원 시계열의 12개월 이동평균 기준

2) 그림 A1~3(B1~3)은 투자자 유형별 일평균 거래회전율(누적 순매수비율)을 나타냄

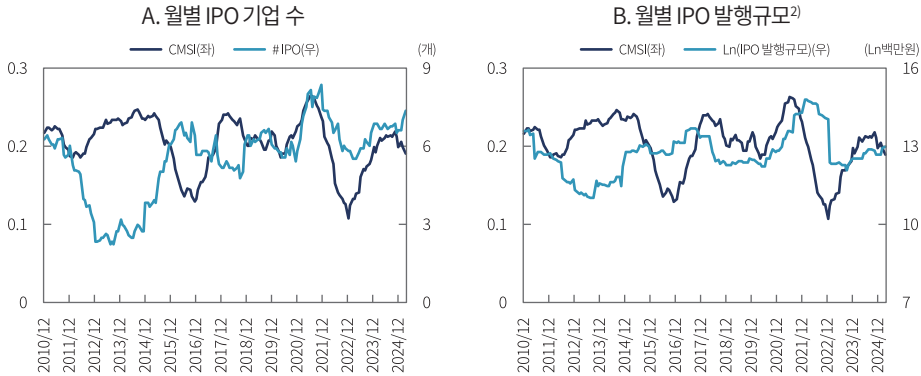
자료: 자본시장연구원, DataGuide

마지막으로 IPO 시장 활동 변수와 CMI의 관계를 살펴보자. 선행연구에 따르면 IPO 시장은 투자자 심리가 가장 직접적으로 반영되는 영역 중 하나로 평가된다. 투자심리가 고조된 시기에는 신규 상장 기업 수와 발행 규모가 확대되고(Lowry, 2003), 초기 투자자 수요가 과열되면서 상장 첫날 초과수익률이 높아지는 경향이 있다(Ritter, 1991; Derrien, 2005). 반대로 시장 심리가 위축되면 IPO 빈도와 초기 수익률이 모두 둔화되는 것으로 알려져 있다. 이러한 이유로 IPO 시장 변수는 오랜 기간 투자자 심리를 추정하는 대표적 대용지표로 활용되어 왔다(Baker & Wurgler, 2006).

<그림 III-5>는 CMI와 IPO 시장 규모 간의 관계를 보여준다. 분석 결과, 최근으로 올수록 CMI가 높은 국면에서 IPO 기업 수와 발행 규모가 동시에 증가하는 패턴이 뚜렷하게 나타난다. 이는 투자심리가 낙관적으로 형성되는 시기에 기업들이 자금조달 기회를 적극적으로 활용하며, 그 결과 신규 상장이 활발해진다는 점을 시사한다.

특히 IPO 시장 규모는 2018~2019년 이후 CMSI와의 연관성이 더욱 강화되는데, 이는 코로나19 전후 국면에서 풍부한 유동성과 개인투자자의 적극적 청약 수요가 결합되면서 투자심리가 양호한 시기에 상장하려는 기업의 경향이 보다 뚜렷해졌기 때문으로 풀이된다.

<그림 III-5> 자본시장 심리지수와 IPO 시장 규모

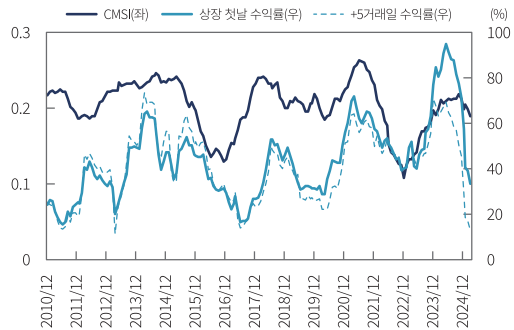


주 : 1) 그래프는 원 시계열의 12개월 이동평균 기준
 2) 월별 합산 신규공모금액(단위: 백만원)에 자연로그를 취한 값
 자료: 자본시장연구원, KRX, DataGuide

또한 CMSI와 상장 초기 수익률 간의 관계 역시 명확하게 관찰된다(<그림 III-6> 참조). CMSI가

높을수록 신규 상장 기업의 평균 초과 수익률¹⁶이 증가하는데, 이는 낙관적 투자심리가 상장 직후 수요 과열로 이어지기 때문으로 해석된다. 이러한 결과는 IPO 활동을 좌우할 수 있는 외부 요인을 통제한 이후에도 유지되며, 상장 후 일정 기간 수익률을 분석해도 투자심리에 따른 추가 흐름의 차이가 일정 부분 지속된다. 추가적인 회귀분석에서는 CMSI가 상위 분위일 때 상장 첫날 초과수익률이 하위 분위 대비 평균 약 20%p 높은 것으로 나타났다.¹⁷

<그림 III-6> 자본시장 심리지수와 상장 초기 초과수익률



주 : 1) 그래프는 원 시계열의 12개월 이동평균 기준
 2) 매월 평균 IPO 프리미엄을 산출하여 12개월 이동평균함
 자료: 자본시장연구원, DataGuide

16 상장 초기 초과수익률 산출 방식은 부록 1에 제시되어 있다.

17 자세한 결과는 부록 2를 참조한다.

IV. 자본시장 심리지수와 주식 수익률

앞서 상장 초기 수익률을 분석한 결과에서 알 수 있듯이, 투자자 심리는 주식 가격에 영향을 미칠 수 있다. 선행연구에서도 이러한 점을 보여주는데, 먼저 시계열 차원에서 투자자 심리의 변동이 단기적으로 시장 수익률에 대한 예측력을 갖는 것으로 보고된다(Brown & Cliff, 2004; Birru & Young, 2022). 낙관적 심리가 고조되면 단기 주가 상승이 나타나고, 반대로 비관적 심리가 확산되면 단기 하락이 발생하는 것이다. 다만 이러한 효과는 시간이 경과함에 따라 점차 약화되거나 소멸하는 경우가 많으며, 과도하게 주가가 움직일 경우 이후 반전(reversal) 현상이 나타나기도 한다.

다음으로 주식 수익률의 횡단면 측면에서는 투자자 심리가 종목별 위험-수익률 구조에 차별적인 영향을 미친다는 결과가 다수 보고되고 있다. Baker & Wurgler(2006, 2007)는 투자심리가 낙관적일수록 가치평가가 어려운 주식, 투기적 성격이 강한 주식, 차익거래가 제약된 주식이 더욱 과대평가(overpricing)되고, 이로 인해 이들 주식의 기대수익률이 체계적으로 낮아진다고 보고한다. 아울러 Stambaugh et al.(2012)은 주식시장의 대표적인 이상현상(anomaly)을 대상으로, 공매도 제약이 있는 상황에서 투자자 심리에 따른 과대평가가 두드러지고, 그 결과 이상현상이 더 크게 나타난다고 주장한다.¹⁸

최근 연구에서는 비정형 텍스트 데이터를 활용하여 심리를 정밀하게 계량화함으로써, 주식 수익률 분석에서 심리 요인의 중요성을 재확인하고 있다. 예를 들어, Tetlock(2007)은 뉴스 기사 어조의 시장 수익률과 거래량에 대한 예측력을 보고하였고, Da et al.(2015) 및 Birru & Young(2022)은 검색량 데이터로 측정한 투자자의 불안이 주식 수익률에 대한 단기 예측력을 가짐을 실증하였다. 본 연구에서 활용하는 CMSI 역시 이러한 맥락에서 LLM 기반 NLP 방법론을 적용하여 산출된 지표라는 점에서, 최신 기법으로 산출한 투자자 심리지표와 주식 수익률이 어떠한 관계를 갖는지 시계열 · 횡단면적 관점에서 검토한다.

18 이상현상(anomaly)은 전통적인 자산가격결정모형으로 설명되지 않는 체계적이고 반복적인 수익률 횡단면의 패턴을 말한다. 대표적으로 규모 효과(size effect), 모멘텀(momentum) 현상 등이 있다. Stambaugh et al.(2012)은 11개의 대표적인 현상을 중심으로, 공매도 제약에 영향을 크게 받는 매도 포지션(short leg)에 속한 주식 그룹이 투자자 심리가 낙관적일 때 과대평가에 따른 기대수익률 하락이 현저해진다고 보고한다.

1. 시장 수익률 분석

먼저 CMSI와 국내 주식시장 수익률 간의 시계열적 관계를 살펴본다. 분석에는 유가증권 시장 및 코스닥 시장의 일별, 월별 초과수익률¹⁹을 활용하며, CMSI의 시차(lagged variable)를 주요 독립 변수로 한 주식 수익률 예측 회귀분석을 실시한다.

분석 결과, 월별 주기로는 CMSI가 시장 수익률을 통계적으로 유의하게 예측하지 못하였다(<표 IV-1> 참조). 이는 월 단위로 집계된 지수 수준에서는 투자심리의 단기적 충격이 희석되거나 다른 요인에 의해 상쇄되었음을 시사한다. 동시점 시장 수익률과의 상관성은 매우 높지만($k = 0$), 이후 1~2개월 간의 시장 수익률에 대한 설명력은 유의하지 않다. 반면, 시차를 3개월로 늘릴 경우 통계적으로 유의한 음(-)의 예측력이 나타난다((1), (5)열). 이는 심리 변화에 따른 주가의 과도한 변화가 시차를 두고 조정되는 것으로 해석된다.

<표 IV-1> 심리지수의 주식시장 수익률 예측력: 월별 시계열 분석 결과

| 종속변수 | 종속변수 시차별 CMSI 회귀계수 | | | | | |
|--------------|--------------------|-----------|-----------|-----------|-----------|-----------|
| | 코스피 수익률 | | 코스닥 수익률 | | 합산 수익률 | |
| 시차(k) | (1) | (2) | (3) | (4) | (5) | (6) |
| $k = 0$ | 0.0319*** | 0.0208*** | 0.0361*** | 0.0252*** | 0.0327*** | 0.0216*** |
| $k = 1$ | -0.0075 | -0.0045 | -0.0018 | 0.0014 | -0.0065 | -0.0035 |
| $k = [1, 2]$ | -0.0099 | -0.0040 | 0.0014 | 0.0032 | -0.0078 | -0.0018 |
| $k = [1, 3]$ | -0.0154** | -0.0072 | -0.0010 | -0.0068 | -0.0140* | -0.0057 |
| 시차종속변수 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| 통제변수 | | ✓ | | ✓ | | ✓ |

주 : 1) 아래 회귀분석을 수행하여 $CMSI_t$ 의 추정계수(β_1)를 표에 보고

$$MKT_{t+k} = \alpha + \beta_1 CMSI_t + \beta_2 VKOSPI_t + \beta_3 SP500_t + \gamma_0 MKT_t + \gamma_1 MKT_{t-1} + \varepsilon_t$$

- MKT_{t+k} = 월별 미래 시장 수익률; $CMSI_t$ = 월별 자본시장 심리지수; 통제 변수 t = 월별 VKOSPI 지수 및 S&P500 수익률 (회귀 모형에 포함되면 ✓ 표시)

- 시차종속변수 = 동시점 및 1개월 전 시장 수익률 (통제할 경우 ✓ 표시)

2) ***, **, *은 각각 1%, 5%, 10% 유의수준에서 통계적으로 유의함을 나타냄

자료: 자본시장연구원, DataGuide

월별 회귀분석 결과와 달리, 일별 주기로 시계열 간격을 좁혀보면, CMSI는 단기적으로 시장 수익률을 양(+)의 방향으로 예측하는 것으로 나타난다(<표 IV-2> 참조). 심리지수가 상승하는

19 통안증권 금리를 무위험 이자율로 가정하여 이를 차감한 시장 수익률을 사용한다.

시기에는 그 직후 며칠간($k = 4\sim 5$ 거래일) 주식시장이 상승했다. 다만 그 효과는 경제적으로 크지 않으며, 이후 빠르게 소멸한다. CMI의 예측력이 며칠 내 급속히 약화되고, 일정 기간 이후에는 오히려 음(-)의 부호로 전환되나 통계적으로는 유의하지 않다. 이는 시장 전반에 걸쳐 투자 심리의 변동이 주가의 과대 또는 과소평가를 일으키고, 점차 조정 과정을 거치면서 초기 충격이 이후 반전되는 것으로 판단된다. 이는 투자 심리 변화에 따른 단기 과잉반응 및 이후 조정 패턴과 일관되며, 과거 선진 주식시장을 다양한 심리지표로 분석한 선행연구(Brown & Cliff, 2004; Tetlock, 2007; Da et al., 2015; Birru & Young, 2022)의 결과와 질적으로 유사하다.²⁰

<표 IV-2> 심리지수의 주식시장 수익률 예측력: 일별 시계열 분석 결과

| 종속변수 시차별 CMI 회귀계수 | | | | | | |
|-------------------|-----------|-----------|-----------|-----------|-----------|-----------|
| 종속변수 | 코스피 수익률 | | 코스닥 수익률 | | 합산 수익률 | |
| 시차(k) | (1) | (2) | (3) | (4) | (5) | (6) |
| $k = 0$ | 0.0080*** | 0.0077*** | 0.0090*** | 0.0093*** | 0.0082*** | 0.0080*** |
| $k = 1$ | 0.0009*** | 0.0010*** | 0.0009*** | 0.0009*** | 0.0010*** | 0.0011*** |
| $k = [1, 2]$ | 0.0007* | 0.0006 | 0.0012*** | 0.0009* | 0.0007* | 0.0007* |
| $k = [1, 3]$ | 0.0007 | 0.0009* | 0.0012** | 0.0011* | 0.0007 | 0.0009* |
| $k = [1, 4]$ | 0.0009* | 0.0010* | 0.0015** | 0.0011* | 0.0010* | 0.0011* |
| $k = [1, 5]$ | 0.0004 | 0.0009 | 0.0008* | 0.0011 | 0.0004 | 0.0008 |
| $k = [1, 6]$ | 0.0006 | 0.0011 | 0.0005 | 0.0009 | 0.0006 | 0.0011 |
| $k = [1, 7]$ | 0.0002 | 0.0005 | 0.0005 | 0.0008 | 0.0002 | 0.0005 |
| $k = [1, 8]$ | 0.0006 | 0.0012 | 0.0006 | 0.0012 | 0.0005 | 0.0012 |
| $k = [1, 9]$ | 0.0002 | 0.0007 | 0.0003 | 0.0008 | 0.0001 | 0.0007 |
| $k = [1, 10]$ | -0.0002 | 0.0006 | -0.0001 | 0.0007 | -0.0003 | 0.0005 |
| $k = [1, 11]$ | -0.0006 | -0.0000 | 0.0000 | 0.0006 | -0.0006 | -0.0000 |
| $k = [1, 12]$ | -0.0010 | -0.0004 | -0.0005 | 0.0001 | -0.0011 | -0.0005 |
| 시차종속변수 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| 통제변수 | | ✓ | | ✓ | | ✓ |

주 : 1) 아래 회귀분석을 수행하여 CMI의 추정계수(β_i)를 표에 보고

$$MKT_{t+k} = \alpha + \beta_1 CMI_t + \beta_2 VKOSPI_t + \beta_3 SP500_t + \gamma_0 MKT_t + \gamma_1 MKT_{t-1} + \dots + \gamma_5 MKT_{t-5} + \varepsilon_t$$

- MKT_{t+k} = 일별 미래 시장 수익률; CMI_t = 일별 자본시장 심리지수; 통제 변수 ε_t = 일별 VKOSPI 지수 및 S&P500 수익률 (회귀 모형에 포함되면 ✓ 표시)

- 시차종속변수 = 동시점 및 1~5거래일 전 시장 수익률 (통제할 경우 ✓ 표시)

2) ***, **, *은 각각 1%, 5%, 10% 유의수준에서 통계적으로 유의함을 나타냄

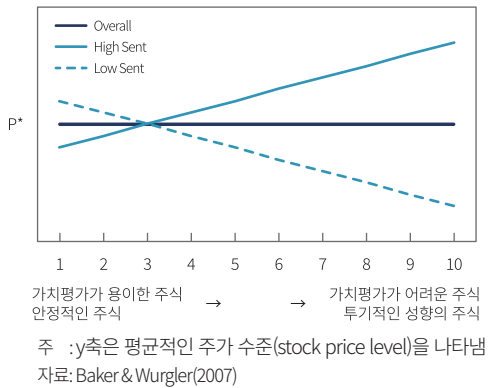
자료: 자본시장연구원, DataGuide

20 시계열 관점에서 투자자 심리와 거시·금융변수 간 영향 관계는 장보성(2026)의 연구를 참조하길 바란다.

2. 주식 수익률의 횡단면 분석

다음으로 CMSI와 국내 개별 상장종목 수익률 횡단면의 관계를 살펴보자. Baker&Wurgler(2006, 2007)와 이후 축적된 문헌은 투자자 심리의 영향이 모든 종목에 균질하게 나타나지 않고, 기업 특성에 따라 차별적으로 작용하는 것을 보여주었다. 이를 그림으로 표현하면 <그림 IV-1>과 같은데, 가치평가가 어렵거나(hard-to-value) 투기적인 성향이 짙은 종목일수록 투자 심리가 낙관적일 때(High sentiment) 주가가 더욱 상승하고, 반대로 투자자 심리가 비관적인 시기(Low sentiment)에는 반대 현상이 나타난다. 이후 심리의 영향이 축소되고 주가가 본질가치(p^*)로 회귀하면서, 투자자 심리는 미래 주식 수익률에 영향을 끼치게 된다. 이를 주식의 기대수익률 관점에서 정리해보면, 시장 심리가 긍정적(부정적)일 때 주가의 과대(과소)평가가 발생하고, 결과적으로 이러한 종목의 기대수익률이 감소(증가)한다. 그리고 이러한 현상은 가치평가가 어렵거나 투기적인 종목에서 현저하게 관찰된다.

<그림 IV-1> 주식 특성에 따른 투자자 심리의 영향 차이



국내 주식시장에서도 투자자 심리의 변동에 따라 이러한 현상이 체계적으로 나타나는지를 검증하기 위해, 본 연구는 월별 CMSI의 중위수를 기준으로 분석기간을 낙관적 국면과 비관적 국면으로 구분한다. 이후 종목의 특성에 따라 기대수익률²¹이 어떻게 달라지는지를 살펴본다.

특히 Baker & Wurgler(2006)의 접근을 따라, 본 연구에서는 가치평가가 어려운 주식(hard-to-value stocks)을 ① 규모가 작거나, ② 업력이 짧거나, ③ 장부가-시장가 비율이 낮거나²², ④ 재무성과가 부진하거나, ⑤ R&D 비중 및 ⑥ 무형자산 비중이 높은 기업으로 정의한다. CMSI로 구분된 각 시기별로 개별 종목의 가치평가 난이도를 나타내는 대용치²³를 산출하고, 이를 10개의

21 여기서 기대수익률을 심리지수에 따라 구분된 시기가 지난 익월의 실현수익률로 정의한다. 수익률은 1절과 같이 무위험 이자율을 차감한 초과수익률을 기준으로 한다.

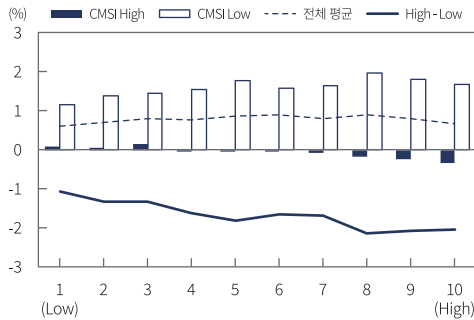
22 장부가-시장가 비율은 PBR(주가순자산비율)의 역수 개념으로, 해당 비율이 낮을수록 가치주(value stocks)보다 성장주(growth stocks)에 가까운 것을 의미한다.

23 상술한 기업 특성별로 상장기업의 순위(rank)를 매기고, 이를 방향에 맞게 정렬하여 평균한 값을 사용한다. 자세한 정의는 부록 1에 제시되어 있다.

포트폴리오 그룹으로 분류한 뒤 그룹별 평균 기대수익률을 분석하였다.

그 결과, <그림 IV-2>에서 보듯이 투자자 심리가 낮은 시기(CMSI Low)에서 반대의 시기(CMSI High)보다 평균 기대수익률이 높게 나타났다. 또한 국면별 수익률의 차이(실선)는 가치평가가 어려운 주식일수록 더욱 확대되는 것으로 확인되었다(그룹1: -1.1%, 그룹10: -2.0%). 구체적으로, 대형주보다 소형주, 안정적인 현금흐름을 보유한 기업보다 성장주 또는 수익성이 낮은 기업, 업력이 짧은 기업, 무형자산 비중이 높은 기업에서 이러한 패턴이 강해진다.²⁴

<그림 IV-2> 주식 특성에 따른 투자자 심리의 영향: 가치평가 난이도



→ 가치평가가 어려운(hard-to-value) 주식 →
 주 : 상장주식을 가치평가 난이도로 정렬한 10개 그룹의 국면별 평균 수익률(막대 그래프), 분석기간 수익률 평균(점선), 국면 간 평균 수익률 차이(실선)를 그림
 자료: 자본시장연구원, DataGuide

다음으로, 주식의 투기적 성향에 따라 투자자 심리의 차별적 효과를 검토한다. 선행연구(Han & Kumar, 2013; 김민기 · 김준석, 2022)를 참고하여, 본 연구는 ‘복권형 주식(lottery-like stocks)’을 투기적 성향이 강한 종목으로 분류한다. 복권형 주식은 ① 주가 수준이 작고(penny), ② 변동성이 크며, ③ 수익률 분포의 왜도가 크거나, ④ 수익률의 극단치가 높은 주식으로 정의된다.²⁵

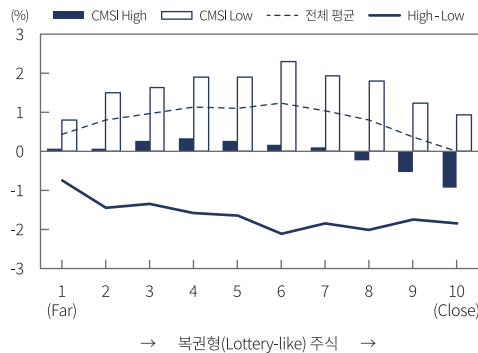
분석기간 전체 상장주식을 이러한 특성에 따라 10개 그룹으로 분류한 결과, 앞선 분석과 마찬가지로 투기적 성향이 강할수록 심리지수 변동에 따른 기대수익률의 차이(실선)가 크게 나타났다(<그림 IV-3> 참조). 흥미로운 점은 이러한 차이가 주로 심리지수가 낮은 시기(하얀 막대)보다

24 자세한 결과는 부록 2를 참고한다.

25 자세한 정의는 부록 1을 참고한다.

높은 시기(검은 막대)에서 더욱 뚜렷하게 나타난다는 사실이다.²⁶ 즉, 낙관적 투자심리가 형성된 국면에서 복권형 주식의 과대평가가 확대되고, 이후 기대수익률이 더욱 낮아지는 경향이 관찰된다.²⁷

<그림 IV-3> 주식 특성에 따른 투자자 심리의 영향: 투기적 성향



주 : 상장주식을 복권(투기적) 성향으로 정렬한 10개 그룹의 국면별 평균 수익률(막대 그래프), 분석기간 수익률 평균(점선), 국면 간 평균 수익률 차이(실선)를 그림
 자료: 자본시장연구원, DataGuide

마지막으로 투자심리의 국면별 차이가 주식 수익률의 횡단면에도 나타나지는지를, 신규 상장 기업의 상장 초기 프리미엄을 통해 살펴보자. 앞서 확인한 바와 같이 정보 비대칭이 크거나 가치평가가 어려운 기업일수록 투자심리에 따른 주가 영향이 확대되는 경향이 있다. 이러한 특징은 IPO 기업에서도 뚜렷하게 나타날 수 있는데, 신규 상장 기업은 정보 공개의 역사가 짧고 가치평가 기준이 불완전해 투자자 심리의 영향을 크게 받을 수 있기 때문이다.

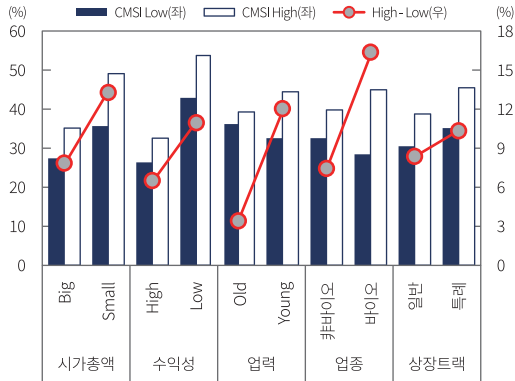
이러한 맥락에서 신규 상장 기업의 특성에 따라 투자심리의 영향을 구분해 보고자 하였다. 구체적으로 상장 당시 기업의 규모(시가총액), 수익성(ROE), 업력, 바이오 섹터 및 특례상장 여부로 신규 상장 기업을 두 집단으로 분류하고²⁸, 각 그룹에 속한 기업의 상장 초기 5거래일 동안의 평균 초과수익률을 상장 당시 CMSI 수준에 따라 나눠 산출해 보았다.

26 심리지수가 낮은 시기(CMSI Low)에서 그룹1과 그룹10의 기대수익률 차이는 -0.1%인 반면, 심리지수가 높은 시기(CMSI High)에서는 기대수익률의 차이가 -1.0%로 유의하다.

27 <그림 IV-2, 3>에 나타난 CMSI 국면별 주식 수익률 횡단면 패턴의 차이는 공통 위험요인을 통제해도 통계적으로, 경제적으로 유의하게 나타난다.

28 상장 당시 기업의 시가총액 규모가 작을수록, 이익 미실현 기업 등 수익성이 낮거나, 업력이 짧을수록, 바이오 섹터 또는 특례상장 트랙에 속하는 경우, 가치평가가 어려운 기업으로 분류한다.

<그림 III-4> 투자자 심리에 따른 상장 초기 수익률 변화



주 : 일별 CMSI로 상장일 전 2주간의 일평균 심리를 측정하고, 기업별 상장 전 심리 수준의 중위값 기준 상위·하위로 High 및 Low 집단을 구분
 자료: 자본시장연구원, KRX, DataGuide

<그림 III-4>는 이러한 분석 결과를 요약하고 있는데, 심리지수가 높게 형성된 시기에 상장된 기업일수록 상장 초기 수익률이 평균적으로 높지만, 심리 수준에 따른 차이(High-Low)는 규모가 작고 업력이 짧거나, 이익실현 규모가 작은 기업, 바이오 산업 및 특례상장 기업일수록 그 크기가 현저해진다. 즉, 기업가치평가의 불확실성이 큰 기업에서 투자자 심리에 따른 상장 직후 주가 영향이 강해지는 것을 알 수 있다(Lowry et al., 2010).

종합하면, CMSI는 국내 주식시장에서 기업 특성에 따라 차별적인 영향을 미치는 것으로 나타난다. 특히 가치평가가 어렵거나 투기적 성향이 강한 종목일수록 낙관적 심리 국면에서 과대평가가 두드러지며, 이후 기대수익률이 상대적으로 낮아지는 패턴이 관찰된다. 또한, 반대로 비관적인 국면에서는 가치평가가 어려운 주식이 더욱 과소평가되어 이후 주식 수익률이 높아진다. 한편, IPO 기업의 상장 직후 프리미엄도 가치평가가 어려운 기업일수록 투자심리에 따른 영향이 확대되는 것으로 분석된다. 이러한 결과는 투자자 심리가 주식의 가격 형성과 기대수익률에 중요한 역할을 한다는 점을 시사한다.

V. 결론 및 시사점

본고는 대형언어모형(LLM) 기반 자연어 처리(NLP) 기법으로 산출한 노성호(2026)의 자본시장 심리지수의 속성을 살펴보고, 주식 수익률과의 관계를 시계열·횡단면 차원에서 분석하였다. 본고의 분석 결과의 경제적 함의는 다음과 같다.

먼저, 이번에 구축한 자본시장 심리지수는 주식 파생상품시장에서 관찰되는 투자자 공포 및 헤지 수요, 파생형 ETF에 대한 투기적 수요와 일정 부분 공행하고, 주식시장 거래활동 및 IPO 시장 변수와도 유의미한 상관성을 보인다. 이는 자본시장 심리지수가 단순한 감성지표를 넘어, 시장 참여자의 기대 형성과 위험 선호의 변화를 종합적으로 포착한다는 점을 보여준다. 특히 자본시장 심리지수는 개인투자자의 거래활동과 정보 비대칭이 큰 신규상장 기업의 주가와 더욱 밀접하게 연동되어, 투자심리의 파급 경로가 비정보거래자 및 가치평가가 어려운 기업군을 중심으로 작동하는 것을 시사한다.

아울러 주식 수익률과의 관계를 실증분석한 결과, 자본시장 심리지수는 월별 주기에서 유의미한 예측력을 보이지 않지만, 일별 주기에서는 일시적인 양(+)의 예측력을 보였다. 이는 투자자 심리가 일시적으로 주가의 과대평가를 유발하는 것을 의미한다. 다만, 이는 경제적으로 유의미한 수준의 영향은 아닌 것으로 분석된다. 또한 기대수익률의 횡단면을 분석한 결과, 투자심리가 낙관적일 때 가치평가가 어려운 주식, 투기적 성향이 강한 주식을 중심으로 과대평가가 관찰되고, 반대로 비관적인 국면에서는 과소평가의 패턴이 나타나며, 이러한 결과는 경제적으로도 유의하다. 이는 투자자 심리지수가 단순히 시장 전체 수익률의 단기 예측력을 넘어, 주식 수익률의 횡단면 구조를 설명하는 데 유용한 정보를 담고 있음을 보여준다.

본 연구에서 사용한 심리지수는 BERT 기반 트랜스포머를 국내 경제·증권 뉴스에 재학습, 파인 튜닝하여 산출한 지표이다. 본고는 새로운 방법론을 통해 계량화한 심리지표인 자본시장 심리지수의 특징을 파악하고, 이러한 심리지표가 갖는 경제적 의미를 주식시장과 연계하여 살펴보았다. 그 결과 자본시장 심리지수는 단기적인 시장 변화의 심리적 동인을 포착하는 동시에, 개별 주식의 위험-수익률 구조를 설명하는 유의미한 지표로 활용 가능하다는 점을 확인하였다. 이는 자본시장 심리지수가 단순한 보조지표를 넘어, 주식시장의 미시·거시적 현상을 이해하는 핵심 변수가 될 수 있음을 뒷받침한다.

참고문헌

- 강장구 · 권경윤 · 심명화, 2013, 개인투자자의 투자심리와 주식수익률, 『재무관리연구』 30(3), 35-68.
- 김민기 · 김준석, 2022, 『국내 개인투자자의 행태적 편익과 거래행태』, 자본시장연구원 연구보고서 22-02.
- 김준석 · 김민기, 2024, 『ETF 시장의 개인투자자』, 자본시장연구원 연구보고서 24-02.
- 노성호, 2026, 『자본시장 심리지수 시리즈 1: 자본시장 심리지수의 구축과 활용』, 자본시장연구원 이슈보고서 26-01.
- 박하연, 2024, 국내 주식시장 투기심리지수와 시장 수익률 사이의 관계, 『한국증권학회지』 53(1), 49-102.
- 변진호 · 김근수, 2013, 주식시장 투자 심리지수의 유용성, 『재무관리연구』 30(4), 225-248.
- 서범석 · 이영환 · 조형배, 2022, 『기계학습을 이용한 뉴스심리지수(NSI)의 작성과 활용』, 한국은행 국민계정리뷰 22-01.
- 송민채 · 신경식, 2017, 뉴스기사를 이용한 소비자의 경기심리지수 생성, 『지능정보연구』 23(3), 1-27.
- 이효정, 2019, 투자심리지수와 수익률 예측력에 관한 횡단면분석, 『재무관리연구』 36(4), 139-167.
- 장보성, 2026, 『자본시장 심리지수 시리즈 2: 거시 · 금융변수와의 관계와 유용성』, 자본시장연구원 이슈보고서 26-02.
- Baker, M., Stein, J.C., 2004, Market liquidity as a sentiment indicator, *Journal of Financial Markets* 7(3), 271-299.
- Baker, M., Wurgler, J., 2006, Investor sentiment and the cross-section of stock returns, *The Journal of Finance* 61(4), 1645-1680.
- Baker, M., Wurgler, J., 2007, Investor sentiment in the stock market, *Journal of Economic Perspective* 21(2), 129-151.

- Barber, B.M., Lee, Y.T., Liu, Y.J., Odean, T., 2009, Just how much do individual investors lose by trading? *Review of Financial Studies* 22(2), 609-632.
- Birru, J., Young, T., 2022, Sentiment and uncertainty, *Journal of Financial Economics* 146(3), 1148-1169.
- Bollen, J., Mao, H., Zeng, X., 2011, Twitter mood predicts the stock market, *Journal of Computational Science* 2(1), 1-8.
- Bollen, N.P., Whaley, R.E., 2004, Does net buying pressure affect the shape of implied volatility functions? *The Journal of Finance* 59(2), 711-753.
- Brown, G.W., Cliff, M.T., 2004, Investor sentiment and the near-term stock market, *Journal of Empirical Finance* 11(1), 1-27.
- Brown, G.W., Cliff, M.T., 2005, Investor sentiment and asset valuation, *Journal of Business* 78(2), 405-440.
- Da, Z., Engelberg, J., Gao, P., 2015, The sum of all FEARS investor sentiment and asset prices, *Review of Financial Studies* 28(1), 1-32.
- Davies, S.W., 2022, Speculation sentiment, *Journal of Financial and Quantitative Analysis* 57(7), 2485-2515.
- Derrien, F., 2005, IPO pricing in “hot” market conditions: Who leaves money on the table? *The Journal of Finance* 60(1), 487-521.
- Han, B., Kumar, A., 2013, Speculative retail trading and asset prices, *Journal of Financial and Quantitative Analysis* 48(2), 377-404.
- Heston, S.L., Sinha, N.R., 2017, News vs. sentiment: Predicting stock returns from news stories, *Financial Analysts Journal* 73(3), 67-83.
- Ke, Z.T., Kelly, B.T., Xiu, D., 2019, Predicting returns with text data, NBER Working Paper.
- Kim, S., Choi, Y., Jeon, J.H., Lu, Y., 2024, Sentiment matters in stock market: Construction of sentiment index using machine learning, *Journal of Economic Theory and Econometrics* 35(4), 87-112.

- Kirtac, K., Germano, G., 2025, Large language models in finance: Estimating financial sentiment for stock prediction, Working Paper.
- Kumar, A., Lee, C.M., 2006, Retail investor sentiment and return comovements, *The Journal of Finance* 61(5), 2451-2486.
- Lowry, M., 2003, Why does IPO volume fluctuate so much? *Journal of Financial Economics* 67(1), 3-40.
- Lowry, M., Officer, M.S., Schwert, G.W., 2010, The variability of IPO initial returns, *The Journal of Finance* 65(2), 425-465.
- Ritter, J.R., 1991, The long-run performance of initial public offerings, *The Journal of Finance* 46(1), 3-27.
- Stambaugh, R.F., Yu, J., Yuan, Y., 2012, The short of it: Investor sentiment and anomalies, *Journal of Financial Economics* 104(2), 288-302.
- Tetlock, P.C., 2007, Giving content to investor sentiment: The role of media in the stock market, *The Journal of Finance* 62(3), 1139-1168.
- Whaley, R.E., 2000, The investor fear gauge, *Journal of Portfolio Management* 26(3), 12-17.

<부록 1> 변수 정의

<부록 표 1> 주요 변수 정의

| 변수명 | 정의 |
|-------------------------|--|
| A. 투자자 심리 관련 주식시장 지표 | |
| 자본시장 심리지수 (CMSI) | <p>국내 경제·증권 뉴스 기사 자료를 학습한 LLM을 통하여 추출된 월별·일별 심리 (sentiment) 지표(노성호, 2026)</p> <ul style="list-style-type: none"> 자료 기간: 2010.01 ~ 2025.03 |
| Baker-Wurgler(BW) 심리지수 | <p>월별 폐쇄형 펀드 할인율, 거래회전율, IPO 기업 수, IPO 상장일 첫날 수익률, 주식 발행/(주식+채권 발행) 비율, 배당 프리미엄 시계열의 첫 번째 주성분(Baker & Wurgler, 2006)</p> <ul style="list-style-type: none"> 노이즈가 있는 변수는 12개월 이동평균 또는 합산 지표를 사용하고 모든 변수를 표준화 거래회전율은 HFT의 증가로 인해 지수를 구성하는 성분 변수에서 제외 BW¹는 각 성분을 주요 거시경제 지표에 대해 회귀분석한 잔차(residual)로 재산출한 지수 |
| 내재변동성 (VKOSPI) | KRX가 산출하는 KOSPI200 옵션의 내재변동성 지수 |
| Put/Call 비율 (P/C Ratio) | 코스피200 옵션의 풋옵션 거래량/콜옵션 거래량 비율 |
| 내재변동성 스프레드 (IV Spread) | 코스피200 풋옵션의 평균 내재변동성 - 콜옵션의 평균 내재변동성 |
| ETF 투기심리지수 (ETF SSI) | <p>코스피200 레버리지·인버스 ETF의 발행주식 수 변화율 차이(Davies, 2022)</p> $ETF\ SSI_t = \Delta_t^{2X} - \Delta_t^{-X}$ <ul style="list-style-type: none"> $\Delta_t^{2X}(\Delta_t^{-1X})$는 레버리지(인버스) ETF의 발행주식 수 변화율 보유·거래규모가 가장 큰 KODEX200 레버리지·인버스 자료를 사용하여 계산 |
| 거래회전율 (Turnover) | <p>일별 주식시장 전체 거래대금/시가총액 비율을 기간별로 평균</p> <ul style="list-style-type: none"> 투자자 유형별(개인, 기관, 외국인투자자)로 산출 유형별 산출 시 증권사 및 기타법인은 제외 |
| (누적) 순매수비율 | <p>일별 합산 순매수 금액/시가총액 비율을 기간별로 누적</p> <ul style="list-style-type: none"> 투자자 유형별(개인, 기관, 외국인투자자)로 산출 유형별 산출 시 증권사 및 기타법인은 제외 |
| IPO 기업의 상장 후 초과수익률 | <p>공모가 대비 상장 후 종가 수익률 - 동기간 상장시장 수익률</p> <ul style="list-style-type: none"> 동기간 상장시장 수익률은 공모가격이 결정된 시점부터 계산 상장시장은 유가, 코스닥 시장으로 구분 |

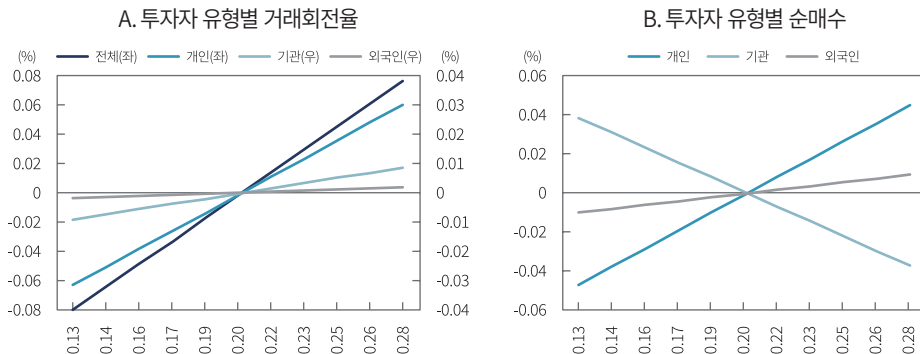
<부록 표 1> 주요 변수 정의(계속)

| 변수명 | 정의 |
|-------------|--|
| B. 기업 특성 변수 | |
| 가치평가 난이도 | 매월 아래 6가지 특장의 평균 순위(rank) (Baker & Wurgler, 2006) <ul style="list-style-type: none"> • 규모(시가총액), 업력, 장부가-시장가 비율, ROE, R&D 집약도, 유형자산 비중 • 규모, 업력, 장부가-시장가 비율, ROE, 유형자산 비중은 작을수록 높은 순위, R&D 집약도는 클수록 높은 순위 |
| 장부가-시장가 비율 | 지배주주지분/시가총액 |
| ROE | 지배주주순이익/지배주주지분 |
| 유형자산 비중 | 유형자산/총자산 |
| R&D 집약도 | 연구개발비/매출액 |
| 복권(투기적) 성향 | 월별 아래 4가지 특장의 평균 순위(rank) (Han & Kumar, 2013; 김민기·김준석, 2022) <ul style="list-style-type: none"> • 주가, 고유변동성, 고유왜도, 일별 최대 수익률 • 주가는 작을수록 높은 순위, 그 외 3개 변수는 클수록 높은 순위 |
| 고유변동성 | 과거 3개월간 공통 위험요인으로 설명되지 않는 일별 주식 수익률 잔차의 표준편차 |
| 고유왜도 | 과거 3개월간 공통 위험요인으로 설명되지 않는 일별 주식 수익률 잔차의 왜도 |
| 일별 최대 수익률 | 과거 1개월 내 최대 일별 수익률 |

<부록 2> CMSI와 주식시장 거래활동 및 IPO 변수의 관계

<부록 그림 1>은 CMSI와 주식시장 거래활동 변수와의 관계를 회귀분석 결과를 통해 시각화한 것으로, CMSI가 하위 10%에서 상위 10%로 이동할 때 투자자 유형별 일평균 거래회전율(그림 A)과 월별 순매수비율(그림 B)이 분석기간 시계열 평균 대비 얼마나 변하는지 나타낸 것이다. 모두 그 관계가 통계적으로 유의하며, 개인투자자의 경우 CMSI가 상위 10% 수준일 때 일평균 거래회전율이 0.08%p 증가하는 반면, 기관투자자는 0.01%p, 외국인투자자는 0.002%p 증가하는데 그친다(시계열 평균은 개인 0.47%, 기관 0.07%, 외국인 0.13%). 또한, CMSI가 상위 10%인 시기에는 개인과 외국인 투자자의 월별 순매수 비율이 각각 0.05%p, 0.01%p 상승하고(시계열 평균은 개인 0.02%, 외국인 -0.03%), 반면 기관투자자는 순매도하는 것으로 나타난다.

<부록 그림 1> 자본시장 심리지수와 주식시장 거래활동의 관계



주 : 1) 아래 회귀분석으로 도출한 자본시장 심리와 주식시장 거래활동의 관계를 도식화

$$\text{거래활동}_i = \alpha + \beta \text{CMSI}_i + \gamma \text{통제변수}_i + \varepsilon_i$$

- 거래활동_i = 월별 일평균 거래회전율, 월별 누적 순매수비율; CMSI_i = 월별 자본시장 심리지수;

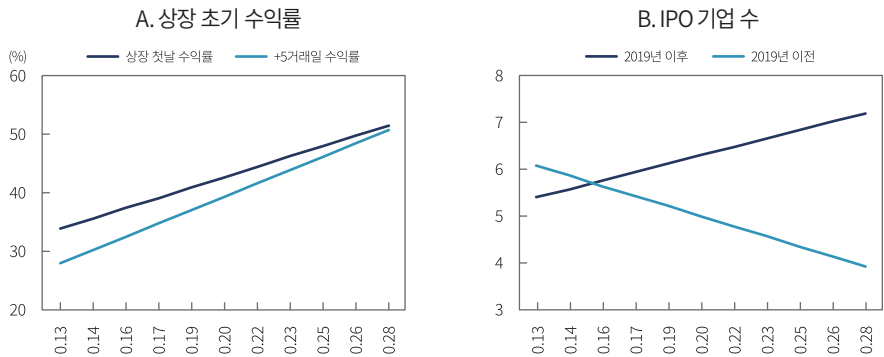
통제변수_i = VKOSPI 지수, 주식시장 합산 수익률

2) 심리지수(x축)가 하위 10%부터 상위 10%까지 변할 때 거래활동 지표의 변화를 그림

자료: 자본시장연구원, KRX, DataGuide

<부록 그림 2>는 CMSI와 IPO 시장 변수(평균 상장초기 수익률, IPO 기업 수)의 관계를 회귀분석한 결과를 그림으로 나타낸 것이다. 주식시장 수익률 및 변동성과 같이 IPO 활동에 영향을 줄 수 있는 외부 요인을 통제해도 투자자 심리와 상장 첫날 초과수익률의 관계는 양(+)의 값으로 유의하다(그림 A). 한편, 신규 상장 기업 수로 대표되는 IPO 시장 규모는 최근 분석기간(2019년 이후)에서 심리지수와 더 밀접한 연관성을 보이는 것으로 분석된다(그림 B).

<부록 그림 2> 자본시장 심리지수와 IPO 시장 변수의 관계



주 : 1) 아래 회귀분석으로 추정한 자본시장 심리와 IPO 시장 변수 간 관계를 도식화

$$IPO \text{ 시장 변수}_i = \alpha + \beta CMSI_i + \gamma \text{통제 변수}_i + \epsilon_i$$

- IPO 시장 변수_i = 상장 초기(1일, 5일) 수익률, IPO 기업 수; CMSI_i = 월별 자본시장 심리지수;

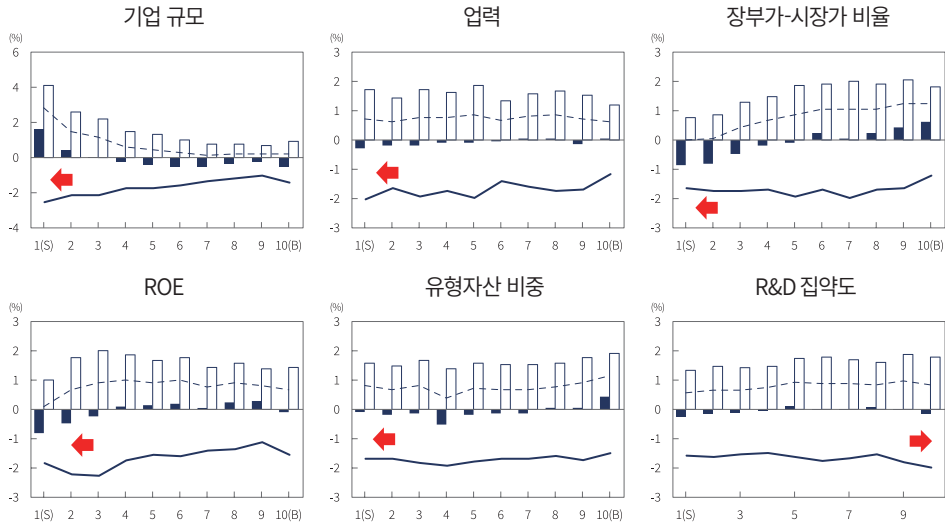
통제 변수_i = 과거 3개월 평균 VKOSPI 지수, 주식시장 누적수익률

2) 심리지수(x축)가 하위 10%부터 상위 10%까지 변할 때 IPO 시장 변수의 변화를 그림

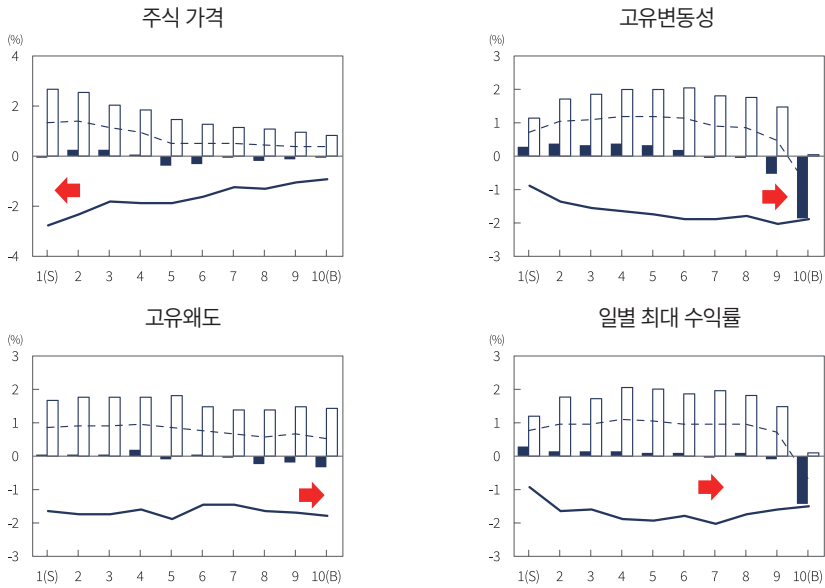
자료: 자본시장연구원, KRX, DataGuide

<부록 3> 개별 기업 특성별 수익률 주식 횡단면 분석 결과

<부록 그림 3> 가치평가 난이도 세부 특성별 분석 결과



<부록 그림 4> 투기적(복권) 성향 세부 특성별 분석 결과



주 : 1) 상장주식을 여러 특징으로 분류한 10개 그룹의 국면별 평균 수익률(막대 그래프), 분석기간 수익률 평균(점선), 국면 간 평균 수익률 차이(실선)를 그림

2) 하얀(남색) 막대 그래프는 각각 투자자 심리(CMS)가 비관(낙관)적인 국면을 의미

자료: 자본시장연구원, DataGuide