주택도시금융연구
Journal of Housing and Urban Finance 2026;11(3):5-24
https://doi.org/10.38100/jhuf.2026.11.3.5
pISSN: 2508-3872 | eISSN: 2733-4139
https://www.ejhuf.org

건설경제 뉴스지수(CENI) 구축과 건설경기 설명력 분석*

Establishing a construction economic news index (CENI) and analyzing its explanatory power for construction business cycles in Korea

김갑진(Kab Jin Kim)[1,†]

Abstract

This study constructs a construction economic news index (CENI) using news big data and empirically examines its explanatory power for fluctuations in the Korean construction business cycle. Traditional analyses of construction activity rely on aggregated indicators such as orders, housing starts, and construction output, which are published with substantial time lags and have limitations in capturing real-time changes in market perceptions and risk signals. To address this limitation, this study develops a multi-layered CENI based on GDELT global news data, comprising a distress-risk intensity index and sentiment indices. Using monthly data from 2015 to 2025, dynamic regression models and business-cycle-dependent models are employed to assess whether the CENI provides additional information beyond conventional macroeconomic and financial variables. The results verify the CENI's statistically significant leading explanatory power for construction orders, even after controlling for traditional determinants such as interest rates and industrial production. In particular, its explanatory power is stronger during economic downturns and recession phases. However, its effects on housing starts and construction output are relatively weak, suggesting that news-based information primarily reflects changes in perceptions and decision-making at the early stages of the construction production process. Overall, the findings indicate that news big data can serve as a useful complementary indicator for monitoring construction business cycles and systemic risks.

Keywords: construction economic news index, news big data, construction business cycle, dynamic regression models, real estate news

Ⅰ. 서론

건설산업은 거시경제의 경기 흐름을 반영하는 동시에 경기변동을 증폭시키는 특성을 가진다. 특히 건설투자는 경제전반의 투자활동과 밀접하게 연계되어 있으며, 건설경기의 변동은 금융시장, 부동산시장 및 실물경제 전반에 영향을 미친다(김갑진, 2025a; Leamer, 2007). 이러한 이유로 건설경기 변화를 조기에 파악하고 경기 흐름을 적시에 판단하는 것은 정책당국과 시장 참여자 모두에게 중요한 과제라 할 수 있다.

그러나 건설경기를 판단하는 기존 통계지표들은 대부분 건설수주, 건설투자, 착공면적 등 집계 통계로서 통계 작성 및 발표 과정에서 상당한 시차가 존재한다. 이러한 지표는 시장 참여자들의 기대나 위험 인식을 충분히, 신속하게 반영하지 못한다는 점에서 한계가 있다(Barrie and Paulson, 1992; Leamer, 2007; OECD, 2019).

* 본고는 연구자 개인의 견해로 소속기관의 입장과는 무관함을 알려드립니다. 유익한 심사의견을 주신 심사위원님들께 감사드립니다.
1 건설공제조합 팀장(경제학박사), 서울시립대학교 도시과학대학 건축학부 강사(Korea Construction Guarantee Cooperative (KCG), University of Seoul, kabdol@hanmail.net)

최근에는 뉴스 기사, SNS, 보고서 등 텍스트 데이터를 활용하여 경제 상황을 분석하려는 연구가 활발히 이루어지고 있다. 뉴스는 경제주체들의 기대, 불확실성, 위험 인식을 비교적 즉각적으로 반영하는 고빈도 · 비정형 정보라는 점에서 기존 집계지표의 시차 문제를 보완할 수 있다. 실제로 뉴스에 내재된 감정이나 불확실성 지표가 실물경기와 금융시장의 변동성이나 투자 · 소비의 단기적 변화와 유의미한 관련을 갖는다는 점이 실증적으로 제시돼 왔다(Baker et al., 2016; Shapiro et al., 2022; Tetlock, 2007).

하지만 이러한 연구들은 주로 금융시장이나 거시경제 수준에 집중되어 있어 특정 산업의 뉴스지수를 구축하고 경기 분석에 활용한 연구는 상대적으로 제한적이다. 특히 건설산업은 경기 변동성이 크고 금융 및 부동산 시장과 밀접하게 연계되어 있음에도 불구하고 뉴스 데이터를 활용하여 건설경기에 대한 시장 인식을 계량적으로 분석한 연구는 충분히 이루어지지 않았다.

이러한 문제의식에서 출발하여 본 연구는 글로벌 뉴스 데이터베이스인 GDELT(Global Database of Events, Language and Tone)를 활용하여 건설산업 관련 뉴스 데이터를 수집하고 이를 기반으로 건설경제 뉴스지수(construction economic news index, 이하 CENI)를 구축한다. 또한 구축된 지수가 건설경기 변동을 설명하는데 어떠한 역할을 하는지 동태적 회귀모형을 통해 실증적으로 분석한다.

본 연구의 주요 기여는 다음과 같다. 첫째, 글로벌 뉴스 빅데이터를 활용하여 건설산업과 관련된 뉴스 정보를 체계적으로 수집하고 이를 기반으로 건설경제 뉴스지수를 구축하였다. 이는 전통적인 통계지표 중심의 건설경기 분석의 저변을 확대하는 새로운 시도로서 의의를 가진다.

둘째, 뉴스 텍스트에 나타난 위험 관련 정보와 정서 정보를 결합하여 뉴스지수의 활용을 다층적으로 적용할 수 있도록 하였다. 이를 통해 뉴스 데이터가 건설경기 분석에서 새로운 정보 변수로 활용될 가능성을 제시하였다.

셋째, 경기 하강국면에서 뉴스 기반 정보의 설명력이 어떻게 변화하는지 분석함으로써 건설경기 변동 과정에서 뉴스 정보의 역할을 검토하였다.

본 논문의 구성은 다음과 같다. 제Ⅱ장에서는 뉴스 기반 경제지표 연구와 건설경기 분석 연구 등 관련 선행연구를 검토하고 본 연구의 차별성을 제시한다. 제Ⅲ장에서는 뉴스지수 구축과 관련된 이론, 사례를 살피고 본고에 적용하는 CENI 구축과정을 설명한다. 제Ⅳ장에서는 실증분석 모형과 분석 방법을 제시하고 실증분석 결과를 제시한다. 마지막으로 제Ⅴ장에서는 연구 결과를 정리하고 정책적 시사점을 논의한다.

Ⅱ. 선행연구 고찰 및 이론적 배경

1. 선행연구 고찰

본 연구와 연관된 선행연구들은 크게 두 가지 흐름으로 구분할 수 있다. 하나는 뉴스텍스트와 같은 비정형 정성정보를 활용해 경기변동과 경제주체의 인식을 분석한 연구들로 전통적인 거시 집계지표가 포착하지 못하는 선행적 정보(content)의 존재 가능성을 다루어 왔다. 다른 하나는 거시 · 금융 변수에 기반하여 경기변동의 구조적 메커니즘을 분석한 연구들로 특히 금융 여건 변화가 주택 · 건설 · 부동산 부문을 매개로 실물 경기와 위험 지표로 전이되는 과정을 이론적 · 실증적으로 규명해 왔다.

1) 뉴스기반 정보를 활용한 경제분석 연구

뉴스 기반 정성지표 연구들은 경기변동을 설명하는 정보가 반드시 전통적인 거시 집계지표에 국한되지 않으며, 경제주체의 기대와 위험 인식이 반영된 정성적 정보 역시 중요한 설명력을 가질 수 있음을 제시해 왔다.

Tetlock(2007)은 뉴스 텍스트의 부정적 정서가 주식시장 수익률과 거래량에 영향을 미친다는 점을 실증하였다. 이는 뉴스 텍스트가 시장 참여자의 기대 형성에 영향을 미칠 수 있음을 보여준다.

Baker et al.(2016)은 신문 기사 빈도를 활용해 경제정

책 불확실성(economic policy uncertainty, 이하 EPU) 지수를 구축하고, EPU 지수 상승이 거시적으로 투자·산출·고용의 하락을 예고함을 제시하였다. 이러한 결과는 텍스트 기반 지표가 전통적인 거시 집계지표와는 다른 경로로 경제주체의 인식(예: 불확실성, 정서)을 포착함으로써 경기변동을 설명하는 추가적 정보를 제공할 수 있음을 시사한다.

최근에는 뉴스 텍스트가 갖는 이러한 거시경제 변동 예측력, 설명력에 대해 보다 직접적인 검증이 시도되고 있다. 대표적으로 Thorsrud(2020)은 주제별 뉴스텍스트를 거시 변수와 결합하여 거시경제 예측력을 개선할 수 있음을 보였다. 이는 텍스트 데이터가 전통적인 경제지표에 포함되지 않는 정보를 반영할 수 있음을 의미한다. 본 연구는 이러한 접근을 건설산업 경기 판단에 적용하여 뉴스 기반 정보가 건설경기 지표에 대해 갖는 선행적 설명력을 분석한다. Shapiro et al.(2022) 역시 뉴스에 나타난 기대와 서사가 소비·투자와 같은 실물 변수의 단기 변동과 밀접하게 연관되어 있음을 보여주며 뉴스가 경제주체의 인식을 반영하는 핵심 매개체임을 강조하였다.

국내에서는 송민채·신경식(2017)이 뉴스 기사 텍스트를 이용해 경기심리지수를 구축하고, 해당 지수가 GDP와 산업생산 등 주요 거시변수에 대해 통계적으로 유의한 예측력을 가지며 기존 설문 기반 심리지수보다 정보 반영 속도가 빠르다는 점을 제시한 바 있다. 이후 연구는 뉴스 주제·정서·핵심어 기반의 지수 구축과정을 거쳐 거시경제 지표와의 관련성 또는 경기 변동과의 연관성을 밝히는 방향으로 진화했다. 이성민·김성현(2019)은 국내 주요 언론 뉴스 텍스트에 대해 LDA 토픽모형과 감성분석을 적용한 텍스트 지수가 산업생산 등 실물경기 지표와 유의한 상관관계를 갖는다는 점을 확인하였다. 서범석(2022)은 뉴스 텍스트를 활용해 15개 주요 경제부문의 동향을 정량화한 경제지표를 개발하여 이를 관련 공식 통계와 비교한 결과 해당 뉴스지표가 공식 통계보다 0~9개월 선행하고, 이를 반영한 경기예측모형의 예측 정확도도 유의하게 향상됨을 보였다. 이는 뉴스 텍스트의 정량화가 실시간 경기동향 파악과 예측에 유용함을 시사하지만, 특정 의도가 반영된 뉴스와 텍스트 지표의 노이즈 통제의 필요성도 제시했다. 최재원·김도형(2021)은 뉴스 및 정책 문서 텍스트를 활용해 감성 사전 기반 경제심리지수를 구성하고 해당 지수가 경기국면별로 상이한 반응을 보이며, 특히 경기 전환기에 설명력이 강화된다는 점을 분석하였다.

한국은행은 2021년 경제기사 기반 뉴스심리지수(news sentiment index, 이하 NSI)의 산출 및 공개 계획을 발표하며 해당 지수가 소비자심리지수(composite consumer sentiment index, 이하 CCSI), 선행종합지수 등 주요 실물·심리지표에 비해 약 1~2개월 선행하는 경향을 제시했다(서범석 외, 2022). 이러한 국내 연구들은 뉴스 기반 지표가 경기 인식을 포착하는 데 실질적인 정보 가치를 가질 수 있음을 보여준다.

기존 국내외 뉴스 지수 연구(서범석 외, 2022; Baker et al., 2016)는 주로 거시경제 전반이나 소비자 심리 포착에 집중되어 있다. 본 연구는 건설산업에 특화된 키워드(corpus)를 바탕으로 산업부실위험 필터링, 뉴스 내용의 정서를 결합한 이중적 지수체계를 구축한다. 이는 기존 통계의 시차를 극복할 뿐만 아니라, 특히 프로젝트파이낸싱(project finance, 이하 PF) 금융과 보증 구조가 결합되어 '시스템 위험'을 내재하는 한국 건설산업을 뉴스라는 고빈도 데이터를 통해 실시간 모니터링할 수 있는 틀을 제공한다는 점에서 학술적·정책적 기여를 찾을 수 있다.

2) 건설·부동산 경기변동과 관련 연구

한편 전통적인 거시·금융 변수에 기반한 연구들은 경기변동의 구조적 메커니즘을 설명하는 데 초점을 두어 왔다. Kiyotaki and Moore(1997)는 담보 가치와 신용 제약의 상호작용을 통해 금융 충격이 실물경제로 증폭되는 메커니즘을 이론적으로 제시하였다. Bernanke et al.(1999)는 금융 가속기 모형을 통해 금융 여건 변화가 경기변동의 진폭을 확대시키는 구조를 설명하였다. 이러한 연구들은 경기변동이 단순한 수요·공급 조정이 아니라 금융 환경을 매개로 비선형적으로 전이될 수 있음을 시사한다.

금융과 경기의 연관성은 특히 주택 및 건설 부문에서 뚜렷하게 조명되었다. Leamer(2007)는 주택 · 건설 투자가 경기 순환에서 차지하는 중심적 역할을 강조하며 금융 여건 악화가 건설 부문을 통해 실물 경기 하강으로 전이되는 경로를 분석하였다. Mian and Sufi(2014), Mian et al.(2017)은 부채 축적과 신용 수축이 경기 침체의 심화와 지속성에 중요한 역할을 한다는 점을 실증적으로 제시하며, 자산 · 신용과 연계된 산업에서 경기변동의 비대칭성이 강화됨을 보여주었다.

국내 연구들 역시 건설 · 부동산 부문에서 이러한 전이 구조를 실증적으로 분석해 왔다. 예컨대 한국은행은 금리 상승 및 신용 여건 악화가 부동산 및 건설 부문의 투자 위축과 PF 위험 확대와 연결될 수 있음을 지적하고 있다(한국은행, 2024a; 2024b). 김갑진(2025a)은 한국형 PF가 책임준공 등 신용보강 구조로 인해 비소구성이 약화되면서 경기 하강 국면에서 건설사 부도와 금융권 리스크가 결합된 시스템위험을 내재하고 있음을 지적한다. 황순주(2024)는 한국의 PF를 '갈라파고적 PF'로 규정하며, 저자본 · 고보증 구조와 책임준공 등 과도한 신용보강으로 인해 PF의 본래적 특성인 비소구성이 약화되고 사업 위험이 건설사와 금융권으로 전가되는 구조적 왜곡을 지적한다. 이러한 구조는 경기 하강 시 PF 부실이 개별 프로젝트 차원을 넘어 건설업 전반과 금융시스템으로 확산되는 경로를 형성하며, 결과적으로 한국의 PF를 시스템위험의 증폭 장치로 작동하게 만든다고 평가한다.

2. 건설경기변동의 파급 경로와 거시적 영향

2000년대 이후 건설 · 부동산 시장의 사업 구조와 금융조달 방식이 변화하면서 건설경기 변동 메커니즘 역시 과거와는 다른 양상을 보이고 있다. 한국의 건설산업은 분양 중심의 사업구조, PF에 대한 높은 의존도, 미분양 위험, 그리고 보증 · 신용공급 메커니즘이 결합된 구조를 지닌다. 이로 인해 건설경기를 포함한 실물 부문의 경기변동은 단순한 수요 · 공급 조정의 결과라기보다 금융환경 변화와 위험 인식이 결합된 연쇄적 · 비선형적 전이 메커니즘을 통해 나타나는 경향이 있다(Claessens and Kose, 2017; Kiyotaki and Moore, 1997; Leamer, 2007; Mian and Sufi, 2014; Mian et al., 2017).

이러한 구조 속에서 한국 건설경기는 두 가지 특징[1]을 보인다. 첫째, 건설투자의 금융(부채) 의존성이 높다는 점이다. PF 중심의 사업구조는 건설투자가 금리 수준, 금융기관의 대출 태도, 시장 참여자의 위험 인식과 같은 금융 환경 변화에 민감하게 반응하도록 만든다. 이에 따라 금융 여건이 완화되는 시기에 신규 개발사업과 건설투자가 빠르게 확대되는 반면, 금융 여건이 악화되면 자금조달 제약이 발생하면서 사업 지연이나 부실 위험이 동시에 확대되는 경향이 나타난다.

둘째, 건설경기 변동이 경기국면에 따라 비대칭적으로 축적되고 강화된다는 점이다. 건설 · 부동산 개발사업은 투자 규모가 크고, 사업 기간이 장기적이기 때문에 경기 확장기에는 신규 사업 확대와 함께 잠재적 위험이 누적되는 경향이 있다. 반면 경기 하강 국면에서는 분양 부진이나 자금조달 악화가 동시에 발생하면서 이전 확장기에 축적된 위험이 일정한 시차를 두고 집중적으로 현실화된다. 특히 분양 성과와 금융조달에 의존하는 사업 구조에서는 위험의 집중 현상이 더욱 뚜렷하게 나타나며, 경기 후퇴기에는 사업 중단이나 건설기업의 재무 위험이 급격히 확대되는 패턴이 관찰된다.

건설산업의 금융의존성 심화와 경기국면에 따른 위험 집중화 현상은 실물경제와 금융시스템에 동시에 영향을 미치는 파급 경로를 형성한다. 건설 · 부동산 부문에서 발생한 투자 위축이나 사업 부실은 건설기업과 금융기관의 건전성, 지역경제, 그리고 경제주체의 투자 심리에 영향을 미치며 거시경제 변동성을 확대시킬 수 있다. 따라서 건설경기 변동은 단순한 산업 경기의 문제가 아니라 금융 안정성과

1) 김갑진(2025a)의 연구를 요약정리 하였다.

밀접하게 연계된 거시경제적 현상으로 이해될 필요가 있으며, 경기 전환기에 대한 조기 인식과 위험 관리의 중요성이 함께 제기된다. 따라서 집계지표의 시차를 보완하고 뉴스 빅데이터에 반영된 시장 인식을 계량화하여 리스크를 선제적으로 모니터링할 수 있는 분석 틀이 요구된다.

Ⅲ. 건설경제 뉴스지수 구축

1. 뉴스 빅데이터 지수 구축 이론

최근 경제학에서는 텍스트를 계량 분석이 가능한 데이터로 취급하는 Text-as-Data 접근[2]이 확산되고 있다. Text-as-Data 접근에 따라 뉴스 텍스트를 경기 분석에 활용하기 위해서는 비정형 정보를 계량 분석이 가능한 시계열 지수로 변환하는 절차가 필요하다. 이러한 절차는 대체로 네 단계로 요약할 수 있다. 첫째, 분석 목적에 부합하는 뉴스 코퍼스(corpus)[3]를 구축하는 단계로 기사출처, 기간, 주제 범위를 설정한다. 둘째, 중복 기사 제거, 토큰화, 표제어 처리 등 전처리 과정을 통해 텍스트를 분석 가능한 형태로 정제한다. 셋째, 정서, 키워드, 주제 등 텍스트의 특징을 추출하고 이를 수치화한다. 마지막으로 이러한 정보를 일정 주기(일 · 월) 단위로 집계 · 표준화하여 시계열 지수로 변환하고 전통적인 경제 지표와의 관계를 통해 경제적 타당성을 검증한다.

한편 뉴스 텍스트를 지수화하는 방법은 크게 네 가지 범주로 구분할 수 있다. 첫째는 키워드 · 빈도 기반 접근으로 특정 경제 개념과 관련된 기사 또는 용어의 출현 빈도를 지수화하는 방식이다. Baker et al.(2016)의 EPU지수는 키워드 동시 출현 빈도를 활용한 대표적 사례이다. 이 접근은 해석이 직관적이라는 장점이 있으나 키워드 선정과 미디어 보도 행태 변화에 민감하다는 한계를 지닌다.

둘째는 감성(sentiment)기반 접근으로 뉴스 텍스트에 내재된 긍 · 부정 정서를 계량화하여 경기 인식의 방향성을 포착하는 방식이다. 한국은행의 NSI는 이러한 접근을 국내에 적용한 사례로 평가할 수 있다. 감성 기반 지수는 경기 전환 국면에서 시장 인식의 변화를 비교적 직관적으로 반영할 수 있다는 장점이 있다. 다만 단어의 사전적 의미와 실제 활용 의미가 다르거나 다중적 맥락으로 해석되는 경우 한계를 지닌다.

셋째는 주제(topic)기반 접근으로 뉴스가 다루는 이슈의 구성 자체를 시계열로 추적하는 방식이다. 토픽 모델은 문서를 여러 주제별로 분해하여 단순히 정서의 방향성뿐 아니라 경기변동과 연관된 위험 요인을 구조적으로 파악할 수 있게 한다. Thorsrud(2020)는 뉴스 토픽을 거시 변수와 결합해 뉴스를 잠재 요인으로 활용하는 분석 틀을 제시하였다. 이 접근은 산업별 · 이슈별 위험을 분리해 분석할 수 있다는 장점이 있으나 토픽 수 선택과 해석의 안정성 문제가 함께 제기된다.

넷째는 기계학습 및 딥러닝 기반 접근으로 단어의 단순 빈도를 넘어 문맥(context) 내에서 복합적인 의미 체계를 파악하는 방식이다. Shapiro et al.(2022)은 이러한 딥러닝 기법을 통해 뉴스 텍스트에서 경제적 불확실성과 경기 인식을 보다 정밀하게 추출하는 분석 틀을 제시하였다. 이 접근은 문맥의 다중적 의미를 포착하여 예측력을 극대화할 수 있다는 장점이 있으나, 모델의 복잡성에 따른 대규모 연산 비용과 분석 결과의 설명가능성(explainability) 확보 문제가 주요 과제로 제기된다.

〈표 1〉은 이상의 뉴스 지수 구축 방법론을 비교한 것이다. 각 방법론은 유의미한 정보를 제공함에도 불구하고 포

2) 'Text-as-Data 접근'은 텍스트를 단순한 질적 자료가 아니라, 경제주체의 인식 · 기대 · 서사(narratives)를 반영하는 관측 가능한 데이터로 간주하고 이를 계량 분석의 대상으로 활용하는 방법론을 의미한다. 이 접근에서 텍스트는 통계적 객체(statistical objects)로 변환되며, 적절한 전처리와 표현(representation), 차원 축소 및 검증 절차를 거칠 경우 전통적인 수치형 경제변수와 결합 분석이 가능하다. Gentzkow et al.(2019)는 Text-as-Data 접근을 경제학의 계량 방법론 체계 안에서 종합적으로 정리한 리뷰 논문으로, 텍스트 분석의 핵심은 알고리즘 자체가 아니라, 측정하고자 하는 경제적 개념에 대한 명확한 정의와 이를 계량화하는 설계에 있음을 강조한다.

3) 일반적으로 뉴스 코퍼스(news corpus)란 연구 목적에 따른 분석 대상의 경계를 규정하는 단계로 기사 출처와 기간, 주제 범위에 따라 선별 · 정제된 뉴스 텍스트(또는 기사 집합)를 의미한다.

〈표 1〉 뉴스지수 구축방법론 비교

구분	핵심개념	주요방법	대표연구 · 사례	장점	단점
키워드 · 빈도 기반	특정 경제 개념의 보도 강도	키워드 동시 출현 빈도 집계	Baker et al.(2016) – EPU	직관적 해석, 재현성 높음, 장기 시계열 구축 용이	키워드 선정 민감, 미디어 보도 행태 변화에 취약
감성(sentiment) 기반	경기 인식의 방향성	감성 사전, 지도학습 분류	Tetlock(2007); 한국은행 NSI	경기 전환 국면 포착 용이, 직관적 해석	위험의 성격 · 내용 구분 한계
주제(topic) 기반	이슈 · 위험의 구성 변화	LDA 등 토픽 모델	Thorsrud(2020)	위험 유형 분리 가능, 구조적 해석력	토픽 수 · 해석 안정성 문제
기계학습/딥러닝 기반	문맥 기반 의미 표현	지도학습, Transformer	Shapiro et al.(2022); 한국은행 NSI	문맥 · 복합 의미 반영, 예측력 우수	데이터 · 연산 비용, 설명가능성 제약

착 가능한 정보와 해석력 측면에서 상이한 장단점을 지닌다. 따라서 경기변동과 위험 전이의 복합적 과정을 단일 차원의 지표로 충분히 설명하기에는 한계가 있을 수 있다. 특히 금융 여건 변화와 위험 인식의 누적이 중요한 역할을 하는 건설산업의 경우 뉴스 정보의 다층적 특성을 반영할 수 있는 분석 틀이 요구된다.

2. 건설경제 뉴스지수 구축 과정

1)데이터 정보원 및 데이터 마커(Marker)설정

본 연구의 실증분석에 활용할 CENI는 글로벌 뉴스 데이터베이스인 GDELT Project에서 제공하는 뉴스 데이터[4]를 활용, 구축하였다. 분석 대상 뉴스는 2015년 1월 1일부터 2025년 12월 31일까지의 GDELT GKG 일별 데이터를 활용하였다. 뉴스 추출 시 'Korea', 'South Korea', 'Republic of Korea' 외에 'Seoul', 'Busan', 'Incheon'을 지리적 마커로 포함하였다. 이는 수도권 및 광역시 중심의 한국 건설시장 특수성을 반영하고, 글로벌 뉴스 매체가 지역명을 통해 한국 경제 이슈를 다루는 경향을 포착하기 위함이다. 지수 산출의 기저가 되는 분모(A)는 이상의 지역, 기간에 의한 일별 총 뉴스량으로 정의하였다.

2) 뉴스 코퍼스 및 정제 키워드 설정

수집된 한국 관련 뉴스 데이터로부터 건설경기와 직접적으로 관련된 정보를 선별하기 위해 건설뉴스 코퍼스(corpus)와 정제 키워드를 설정하였다. 건설뉴스 코퍼스는 전체 한국 관련 뉴스 중 건설 · 부동산 · 주택 뉴스만을 선별하는 단계로, 이를 위해 해당산업을 포괄적이며 중립적으로 설명하는 11개[5]단어를 필터링 키워드로 설정하였다. 한편 뉴스 데이터에는 동일 사건을 반복 보도하거나 해당 산업과 직접 관련되지 않은 기술적 용어가 포함될 수 있기 때문에 추가적인 정제 과정을 수행하였다. 예를 들어 'Build(건설)' 키워드가 IT 분야의 소프트웨어 빌드 시스템(build system)이나 반도체(chip, gpu) 등 건설 경기와 무관한 맥락에서 사용되는 경우 등을 방지하기 위해 제외 키워드(exclude list)[6]를 설정하였다. 이는 필터링 키워드가 최대한 코퍼스 범위내에서 유의성을 갖도록 하여 데이터의 순도(purity)를 제고하는 기능을 수행한다.

4) GDELT는 전 세계 주요 언론매체의 기사를 실시간으로 수집 · 구조화하여 제공하는 데이터베이스로 뉴스의 주제, 정서, 사건 유형 등에 대한 메타정보를 포함한다. GDELT는 날짜별로 GKG(global knowledge graph)계열 원시파일을 공개하며, 특히 일별 파일은 "YYYYMMDD.gkgcounts.csv.zip" 형태로 제공한다.

5) 건설경제 뉴스 코퍼스로 활용된 키워드는 1) 산업 및 시장을 포괄하는 HOUSING, REAL_ESTATE, PROPERTY, APARTMENT, CONSTRUCTION, 2) 도시개발 및 주택공급과 관련된 URBAN, REDEVELOPMENT, RECONSTRUCTION, 3) 주택시장 및 금융과 관련된 HOME_SALES, MORTGAE, RENT, LEASE 등 총 11개이다.

6) 배제키워드는 software, compiler, chip, gpu, ai model, build system으로 설정하였다.

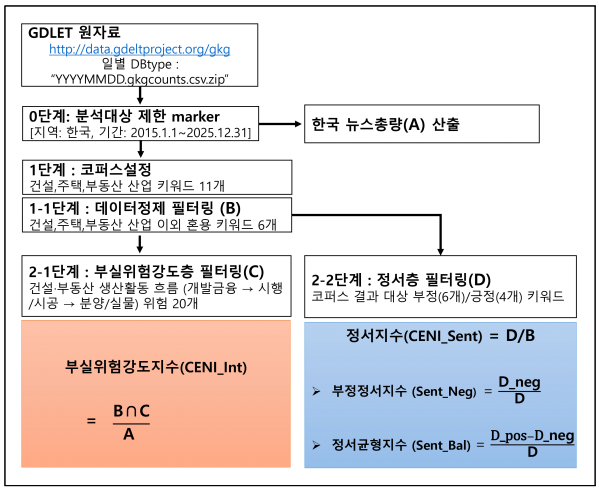
3) 이중 구조 건설경제 뉴스지수의 산출

앞에서 살펴본 바와 같이 기존의 뉴스 텍스트 기반 경기지수는 크게 키워드 빈도 기반 접근과 감성(sentiment) 기반 접근으로 구분되어 발전해 왔다. 전자가 특정 위험 또는 사건의 발생 강도를 포착하는 데 강점이 있는 반면, 후자는 경제주체의 기대와 심리적 방향성을 반영하는 데 강점이 있다. 본 연구는 건설경기 변동이 금융여건과 기대 변화가 결합된 비선형적 전이 구조를 갖는다는 점을 고려하여, 두 가지 정보를 분리하여 동시에 반영할 수 있는 이중 구조(부실강도 - 정서)를 채택하였다.

부실강도지수(CENI_Int)는 PF, 유동성 경색, 부도 등 건설경기 하방위험과 직접적으로 연관된 뉴스의 상대적 비중을 측정함으로써 '위험 신호의 집중도'를 계량화하였다. 반면 정서지수(CENI_Sent)는 뉴스 보도의 긍·부정 톤을 반영하여 시장 참여자들의 기대와 심리적 환경을 포착하는 지표로 설계하였다. 이러한 구성은 건설경기 변동이 단순한 실물 수급 변화가 아니라 금융제약, 기대 변화, 위험 인식의 상호작용을 통해 나타난다는 점을 반영한 것으로, 단일 차원의 뉴스지수보다 경기 전환 신호를 보다 입체적으로 포착하기 위한 설계이다.

구체적으로 CENI_Int는 뉴스 코퍼스 필터링을 거친 뉴스에 대해 건설·부동산 생산활동 흐름(개발금융 → 시행/시공 → 분양/실물)상 대두될 수 있는 부실 위험 관련 키워드 20개[7]로 필터링하여 산출하였다. CENI_Sent는 1단계 건설뉴스 코퍼스에 의해 도출된 뉴스를 대상으로 내용상 긍정적 또는 부정적 어휘의 출현 빈도를 계량화하여 산출하였다. 정서분석에 활용된 필터링 키워드는 부정, 긍정 각 6개와 4개[8]로 이들의 상대적 비중을 통해 시장 심리변화의 기준으로 활용한다.

CENI_Sent는 세부적으로 부정정서지수(SENT_Neg), 정서균형지수(SENT_Bal)로 구성되며 CENI_Int의 해석을 강화하거나 강건성 검증을 위한 보조 지표로 활용한다. 〈그림 1〉은 이상의 논의를 요약, 도식화한 것으로 지수 산출 과정 및 공식을 보여준다.



〈그림 1〉 건설경제 뉴스지수(construction economic news index, CENI)의 구성

Ⅳ. 실증분석

본 장에서는 건설 관련 뉴스 빅데이터로부터 구축한 CENI가 전통적인 건설경기 지표에 비해 어떠한 정보적 가치를 제공하는지를 실증분석한다. 실증분석을 관통하는 핵심질문은 다음의 두 가지이다. 첫째, 'CENI는 전통적인 거시경기 변수들을 통제한 이후에도 건설경기 변동을 설명하는 추가적 정보를 제공하는가', 둘째, '뉴스에 반영된 경기 인식과 위험 신호는 실제 건설경기 주기별로 유의한 차이를 보이는가'. 이러한 질문에 대한 답을 통해 본 연구는 CENI가 건설경기 판단과 위험 모니터링에 어떠한 역할을

7) 부실위험 키워드는 1) 금융관련 PROJECT_FINANCE, PF, LIQUIDITY, CREDIT, CREDIT_CRISIS, FINANCIAL_CRISIS, DELINQUENCY, 2) 시행, 시공사의 부실관련 DEFAULT, BANKRUPTCY, INSOLVENCY, WORKOUT, RESTRUCTURING, RESPONSIBLE_COMPLETION, COMPLETION_GUARANTEE, CONSTRUCTION_HALT, STOP_WORK, 3) 분양 및 실물관련 UNSOLD, OVERSUPPLY, SLUMP, PRICE_DROP 등 총 20개이다.

8) 정서지수 산출에 활용된 부정키워드는 ANXIETY, FEAR, NEGATIVE, PESSIMISM, UNCERTAINTY, RISK이며 긍정키워드는 POSITIVE, OPTIMISM, CONFIDENCE, HOPE이다.

〈표 2〉 분석개요

구분	내용
분석목적	건설뉴스지수(CENI)가 건설생산지표에 전통적인 경기 · 금융변수 대비 추가적인 정보(content)를 제공하는지, 경기주기별 유의한 설명력을 보이는지 검증 – 핵심검증대상 (가설) 1) CENI는 기존 변수들을 통제한 이후에도 추가적 설명력을 제공하는가?(확장모형) 2) 그 정보력은 경기주기별로 유의한 차이를 보이는가?(경기모형)
분석기간	2015년 1월~2025년 12월(월별)
종속변수(Y)	건설수주(Y1), 건설착공(Y2), 건설기성(Y3)
통제변수(X)	기업여신평균금리, 광공업생산지수, 건설보증부도율, 건설공사비지수[1)]
설명변수	건설뉴스지수(CENI): 부실강도지수(CENI_Int), 정서지수(CENI_Sent)로 구성
기본모형	전통적인 건설경기 영향 요인만을 포함한 시차 모형 (Baseline Model) $dY_t = \alpha + \sum_{i=1}^{p}\beta_i dY_{t-i} + \sum_{j=1}^{q}\gamma_j X_{t-j} + \epsilon_t$
확장모형 A	기준 모형에 부실강도지수(CENI_Int: CENI)를 추가한 확장 모형 (Extended Model) 1. 뉴스지수 변수만 통제 $dY_t = \alpha + \sum_{i=1}^{p}\beta_i dY_{t-i} + \sum_{j=1}^{q}\gamma_j X_{t-j} + \sum_{j=1}^{q}\delta_j CENI_{t-j} + \epsilon_t$ 2. 뉴스지수 변수와 건설보증부도율 변수의 동시 통제 $dY_t = \alpha + \sum_{i=1}^{p}\beta_i dY_{t-i} + \sum_{j=1}^{q}\gamma_j X_{t-j} + \sum_{j=1}^{q}\delta_j CENI_{t-j} + \sum_{j=1}^{q}\theta_j dPDrate_{t-j} + \epsilon_t$
확장모형 B	뉴스지수 변수를 부실강도지수(CENI_Int: CENI), 정서지수(CENI_Sent:SENT), 양자의 상호작용항으로 분해 후 통제 $dY_t = \alpha + \sum_{i=1}^{p}\beta_i dY_{t-i} + \sum_{j=1}^{q}\gamma_j X_{t-j} + \sum_{j=1}^{q}\delta_j CENI_{t-j} + \sum_{j=1}^{q}\theta_j SENT_N_{t-j} + \sum_{j=1}^{q}\phi_j (CENI \times SENT_N)_{t-j} + \epsilon_t$
경기모형	분석기간을 경기주기(회복–호황–후퇴–침체)로 재편, 경기주기별 뉴스지수의 영향력 검증 $dY_t = \alpha + \sum_{i=1}^{p}\beta_i dY_{t-i} + \sum_{j=1}^{q}\gamma_j X_{t-j} + \sum_{k=1}^{4}\sum_{j=1}^{q}\delta_{k,j}(\ln_SENT_N_{t-j} \times D_k) + \epsilon_t$ where, D_k : 경기주기 더미, k=1(회복), k=2(호황), k=3(후퇴), k=4(침체)

주: 1) 거시 통제변수 후보 중 건설공사비 지수의 경우 분석기간 중 지속적, 급격한 상승(지수기준 86.52→132.7)을 보여 통제변수로 활용가능성에 불구하고 종속변수(수주, 착공, 기성)와 추세공선성을 감안하여 본 분석에서 보다 중립적인 광공업지수, 여신평균금리를 활용하였다.

할 수 있는지를 실증적으로 제시하고자 한다.

1. 분석개요 및 분석방법

〈표 2〉에서 보는 바와 같이 실증분석은 핵심질문의 성격에 맞추어 CENI의 경기 설명력을 파악하기 위해 세 가지 상호 보완적인 계량적 접근(기본모형, 확장모형, 경기모형) 으로 구성된다.

먼저 기본모형과 확장모형은 건설경기변동의 설명력을 평가하기 위한 동태적 단일방정식 회귀모형(dynamic regression models with lagged dependent variables, 이하 DRML)으로 설정한다. 기본모형이 전통적인 거시 · 금융 및 건설경기 영향요인만을 포함한 반면, 확장모형은 동일한 모형 구조에 CENI를 추가함으로써 CENI의 한계

적 설명력을 평가한다. 이들 모형은 건설경기 지표의 관성과 거시·금융 요인의 시차 효과를 통제한 상태에서 CENI가 제공하는 추가적 정보(content)를 조건부로 평가하는 데 목적이 있다.

이러한 접근은 모든 변수를 대칭적으로 내생화 하는 VAR(vector autoregression) 모형과 구별된다. VAR 모형은 시스템 전체의 동태적 움직임을 추정하는 데 유용하지만 개별 변수의 조건부 효과를 해석하기 어렵고, 변수와 시차가 증가할수록 표본 손실과 추정력이 불안정해지는 한계를 가진다. 반면 동태적 회귀모형은 종속변수와 설명변수를 명확히 구분함으로써 '기존 거시변수들을 통제한 이후에도 CENI가 추가적인 설명력을 갖는가'라는 연구 질문에 보다 직접적으로 대응할 수 있다. 이는 CENI가 전통 지표를 단순히 대리(proxy)하는 것이 아니라 독립적인 정보원으로 기능하는지를 검증하는 데 적합한 구조이다.

내생성 문제와 관련하여 동태적 단일방정식 회귀모형(dynamic single equation regression model, 이하 DRML)은 몇 가지 측면에서 동시적 상관과 역인과 가능성을 완화하는 특성을 갖는다. 우선 종속변수(건설경기 지표)의 시차항을 포함하여 건설경기 지표의 자기상관성을 명시적으로 통제함으로써 설명변수와 오차항 간의 상관 가능성을 줄인다. 또한 CENI는 동월 효과보다는 시차 형태로 모형에 포함되어 실물지표의 동시적 변동에 대한 즉각적인 반응보다는 선행적으로 축적·확산되는 경기 인식과 위험 신호를 반영하도록 하였다.

분석기간은 김갑진(2025b)의 확장성을 감안해 2015년 1월부터 2025년 12월까지로 설정한다. 이는 최근 건설·부동산 경기변동의 핵심적 국면을 포괄하는 기간으로, PF 구조 확산, 코로나19충격, 고금리 하에서의 건설경기 조정 국면 등을 포함한다.

2. 분석을 위한 사전검토

분석에 앞서 주요변수의 기초통계량과 시계열 변수의 안정성(stationary), 적정시차(lag)에 대해 〈표 3〉~〈표 5〉와 같이 점검하였다. 또 CENI와 건설실물지표 간의 선-후행성을 사전점검하기 위해 그랜저 인과분석을 〈표 6〉과 같이 시행하였다. 먼저 〈표 3〉에서 보는 바와 같이 주요변수별 기초통계량 분석결과 건설 실물지표인 수주, 착공, 기성액은 로그 변환 후 대체로 표준화된 분포를 보이며 안정적

〈표 3〉 변수별 기초통계량

변수명	N	평균	표준편차	최솟값	P50	최댓값	출처
부도율(PDrate;%)[1)]	132	3.029	0.903	1.7551	2.755	5.400	건설공제조합
건설기성(ln_C_cmp_m; 백만)	131	16.270	0.190	15.743	16.263	16.700	통계청 (건설업조사)
건설수주액(ln_C_cnt_m; 백만	131	16.411	0.324	15.508	16.377	17.333	
건설인허가(ln_C_pmt_ms; m^2)	131	16.362	0.233	15.711	16.389	16.953	
건설착공(ln_C_beg_ms; m^2)	131	16.105	0.305	15.215	16.158	16.713	
기업여신평균금리(CLoan_ir; %)	131	3.817	0.788	2.670	3.620	5.670	한국은행
광공업생산지수(IP_indx)	131	103.296	7.987	83.45	103.1	122.5	통계청
경기동행지수(BizIX_20_100)	131	101.976	7.975	87.200	100.500	114.200	한국은행
부실강도지수(CENI_Int)	132	0.017	0.013	0.004	0.011	0.060	GDELT; 일별뉴스데이타
부정정서지수(SENT_Neg_)	132	0.694	0.096	0.346	0.704	0.908	
균형정서지수(SENT_Bal)	132	−0.387	0.193	−0.817	−0.408	0.307	

주: 1) 건설공제조합의 부도율은 '보증대급금 및 실질손실 발생', '융자90일 이상 연체', '신용불량정보 발생', '당좌거래정지', '회생절차 신청' 등 사건을부도로 간주, 산출되었다.

〈표 4〉 단위근 검증결과

변수명		ADF(*p*값)		PP(*p*값)		비고
부도율 (PDrate; %)	수준	−0.553	0.2908	−0.461	0.9849	단위근
	차분	−2.579	0.0057***	−11.357	0.0000***	
건설기성액 (ln_C_cmp_m; 백만)	수준	−2.410	0.1390	−5.625	0.0000***	단위근
	차분	−6.816	0.0000***	−24.349	0.0000***	
건설수주액 (ln_C_cnt_m; 백만)	수준	−1.757	0.7253	−10.435	0.0000***	단위근
	차분	−4.438	0.0019***	−24.679	0.0000***	
건설인허가면적 (ln_C_pmt_ms; m^2)	수준	−0.860	0.1959	−5.625	0.0000***	단위근
	차분	−2.843	0.0027***	−24.349	0.0000***	
건설착공면적 (ln_C_beg_ms; m^2)	수준	−0.443	0.3293	−5.265	0.0000***	단위근
	차분	−3.035	0.0015***	−18.913	0.0000***	
기업여신평균금리 (CLoan_ir;%)	수준	−2.974	0.0018***	−1.287	0.6353	단위근
	차분	−2.034	0.0223**	−6.563	0.0000***	
경기동행지수 (BizIX_20_100)	수준	−1.775	0.7272	−2.204	0.2048	단위근
	차분	−3.029	0.0323**	−6.646	0.0000***	
광공업생산지수 (IP_indx)	수준	−2.472	0.3422	−10.010	0.0000***	단위근
	차분	−4.231	0.0040***	−25.437	0.0000***	
부실강도지수(CENI_Int)	수준	−2.105	0.0188**	−3.543	0.0069***	
부정정서지수(SENT_Neg)	수준	−2.856	0.0026***	−10.089	0.0000***	
균형정서지수(SENT_Bal)	수준	−2.856	0.0026***	−10.089	0.0000***	

주: *, **, ***은 각 10%, 5%, 1% 유의수준에서 귀무가설(단위근 존재)을 기각할 수 있음을 의미.
ADF, augmented Dickey–Fuller; PP, Phillips–Perron.

〈표 5〉 종속변수별 시차 조합(*p*, *q*)에 대한 정보기준(AIC, BIC, HQIC) 상위 후보

종속변수	rank	*p*	*q*	N	log우도	AIC	BIC	HQIC
건설수주(Y1)	1	1	3	127	25.798	−35.596	−12.842	−26.351
건설수주(Y1)	2	11	2	119	44.444	−56.888	−12.422	−38.831
건설수주(Y1)	3	12	3	118	50.909	−63.819	−11.176	−42.444
건설착공(Y2)	1	12	1	118	171.585	−313.17	−271.61	−296.296
건설착공(Y2)	2	12	2	118	172.286	−310.571	−263.47	−291.447
건설착공(Y2)	3	12	3	118	173.527	−309.054	−256.41	−287.679
건설기성(Y3)	1	11	1	119	4.137	19.726	58.634	35.526
건설기성(Y3)	2	5	1	125	−10.318	36.635	59.262	45.827
건설기성(Y3)	3	12	1	118	4.112	21.776	63.337	38.651

인 시계열 분석의 토대를 형성하고 있다. 구체적으로 건설수주액(ln_C_cnt_m)의 표준편차가 타 지표 대비 높게 나타나는 것은 건설 경기의 초기 단계에서 시장 불확실성에 따른 변동성이 가장 민감하게 작용함을 시사한다. 뉴스기

〈표 6〉 주요변수 간 그랜저인과관계 분석결과

Equation	Excluded	chi2	*df*	Prob>chi2
dln_건설수주	dln_건설기성	25.67	3	0.000***
	dln_건설착공	35.45	3	0.000***
	D_부실강도지수	10.44	3	0.015**
	D_균형정서지수	16.57	3	0.001***
	ALL	61.57	12	0.000***
dln_건설기성	dln_건설수주	10.59	3	0.014**
	dln_건설착공	52.16	3	0.000***
	D_부실강도지수	2.62	3	0.454
	D_균형정서지수	12.53	3	0.006***
	ALL	67.96	12	0.000***
dln_건설착공	dln_건설수주	2.38	3	0.498
	dln_건설기성	47.53	3	0.000***
	D_부실강도지수	2.32	3	0.508
	D_균형정서지수	6.78	3	0.079*
	ALL	73.24	12	0.000***
D_부실강도지수	dln_건설수주	4.76	3	0.190
	dln_건설기성	5.10	3	0.165
	dln_건설착공	2.90	3	0.407
	D_균형정서지수	6.51	3	0.089*
	ALL	16.85	12	0.155
D_균형정서지수	dln_건설수주	4.94	3	0.176
	dln_건설기성	3.64	3	0.303
	dln_건설착공	13.31	3	0.004***
	D_균형정서지수	0.73	3	0.865
	ALL	21.24	12	0.047**

주: *, **, ***은 각 10%, 5%, 1% 유의수준에서 귀무가설을 기각할 수 있음을 의미.

반 변수의 경우 CENI_Int는 평균값 대비 최댓값이 높은 우측 꼬리가 긴 분포를 보이는데, 이는 특정 시기에 집중되는 건설 관련 부정적 뉴스나 부실 위험 신호가 지수에 효과적으로 반영되어 있음을 의미한다. 또 균형정서지수(SENT_Bal)는 평균이 음(-)의 값을 가지며 최솟값과 최댓값이 비대칭적 범위를 보이는데, 이는 분석기간 동안 건설 관련 뉴스가 전반적으로 부정 정서에 구조적으로 편향되어 있음을 의미한다. 동시에 표준편차가 비교적 크게 나타나는 점은 시기별로 정서가 크게 흔들리며, 심리적 국면 전환을 반영하는 변수로 해석할 수 있다. 안정성 검증과 관련하여 종속변수, 통제변수, 설명변수 후보군에 대하여 ADF(augmented Dickey-Fuller) 및 필립페론(Phillips-Perron, 이하 PP)검증을 실시한 바 〈표 4〉와 같이 주요 건설경기지표와 통제변수 군에서 단위근이 없다는 귀무가설을 기각할 수 없어 1계 차분 변수를 활용하였다.

시차(p, q)의 선택은 정보기준(AIC, BIC, HQIC)에만 기

계적으로 의존하지 않고 월별 자료의 특성, 경제적 전달시차, 잔차진단을 종합적으로 고려하여 결정하였다. 〈표 5〉에서 보는 바와 같이 정보기준 기반의 그리드 탐색에서는 일부 종속변수에서 비교적 긴 자기시차($p \approx 11$-12)가 선택되는 경향이 관찰되었으나 이는 월 더미를 사용하지 않은 설정에서 연주기(12개월) 성분이 자기 회귀항으로 흡수된 결과일 가능성이 있다. 이 경우 해석 가능성과 모형 단순성(parsimony)이 저하될 수 있으므로 종속변수 시차를 $p \approx 1$~6범위로 제한하여 재검토하였다. 그 결과 모델 구조의 안정성, 계수의 일관성, 그리고 Breusch-Godfrey 자기상관 검정 결과를[9] 종합할 때 p=5가 가장 균형적인 선택으로 판단되어 기준 시차로 채택하였다. 한편 금리 및 경기변수의 영향은 즉각적이라기보다 수개월에 걸쳐 전달되는 특성을 고려하여 $q \approx 1$~3범위에서 비교한 결과 설명력과 안정성이 균형을 이루는 q=3을 최종적으로 채택하였다. 월별 자료에서 잔존할 수 있는 자기상관 문제는 Newey-West(HAC) 표준오차(최대 시차 12)를 적용하여 보수적으로 보정하였다.

〈표 6〉은 본 분석의 주요 변수인 CENI와 건설 실물지표 변수 간 그랜저 인과관계 분석결과를 보여준다. 양 지수간 역인과관계, 동시성 문제에 대해 변수간 선-후행관계를 사전적으로 점검한 결과 CENI_Int와 SENT_Bal은 건설수주에 대해서만 유의한 선행적 설명력을 가지며, 착공 및 기성과 같은 생산 단계 변수에 대해서는 직접적인 인과관계가 확인되지 않았다. 이는 뉴스 기반 정보가 실물 생산에 직접적으로 영향을 미치기보다, 발주 및 계약 단계에서의 의사결정을 매개로 간접적으로 전이되는 구조를 시사한다. 특히 CENI_Int는 실물 변수로부터 역인과관계가 관측되지 않아 단순한 경기 반응 지표가 아니라 선행적 정보 변수로 기능하는 것으로 판단된다. 반면 SENT_Bal은 기성에도 일부 영향을 미치는 것으로 나타나, 구체적 위험 신호보다 심리적 환경이 생산단계까지 확산될 가능성을 보여준다.

3. 분석결과

1) 기본모형 분석결과

〈표 7〉은 건설생산활동을 나타내는 수주-착공-기성을 종속변수로 한 기본모형의 분석결과이다. 기본모형 분석결과 건설생산활동의 월별 변동은 공통적으로 강한 자기상관과 수개월에 걸친 조정 구조를 보인다. 이는 발주·공정·행정·현장 운영 등 건설 부문의 내생적 조정 메커니즘의 과정임을 알 수 있다.

건설수주증가율의 경우 자기조정 성향이 특히 강하게 나타나 단기 변동이 빠르게 되돌림(평균회귀)되는 특징을 보인다. 반면 거시 통제변수의 직접적인 영향은 제한적으로 관측되어 월별 수주 변동은 거시 환경의 즉각적 반영이라기보다 발주 의사결정 과정의 지연과 비선형성에 크게 좌우되는 것으로 해석된다. 이러한 특성은 수주가 정보·불확실성 요인의 영향을 상대적으로 크게 받을 가능성을 시사한다. 다만 거시통제변수의 영향이 전반적으로 제한적인 가운데 금리변화의 1기 시차가 양(+)의 방향으로 유의하게 나타난다. 이는 금리변화가 경기·정책 신호로 작용하면서 발주 의사결정에 영향을 주는 가운데 금리상승 국면에서 단기적으로 계약이 선반영되어 자금조달조건 악화 전 계약이 늘어날 수 있음을 유추할 수 있는 대목이다.

건설착공증가율은 수주에 비해 거시 실물경기와 연동성이 뚜렷하게 나타난다. 특히 산업생산지수의 시차항이 일관되게 음(-)의 유의성을 보이는 점은 착공이 단순한 경기확장 변수라기보다 거시적 자본배분과 정책적 경기조정 수단으로 작동함을 시사한다. 한국경제에서 제조업 중심의 실물경기가 확장 국면에 진입할 경우 투자와 유동성은 주로 설비·수출 부문으로 집중되며, 건설부문에 대한 유인은 상대적으로 약화되는 경향이 있다. 반대로 실물경기가 둔화되거나 조정 국면에 진입할 경우, 경기부양 수단으로서 건설투자가 적극적으로 동원되면서 착공이 증가하는 구조가 반복되어 왔다. 이러한 결과는 건설착공이 경기의 단순

9) BG 자기상관 검정결과 잔차 자기상관의 유의성은 종속변수별로 상이한 시차(수주:6, 착공:1, 기성4)에서 나타났으나, 전반적으로 $p \geq 6$ 이후 자기상관이 강화되는 경향이 관찰되었다. 이에 본 연구는 월별 자료의 해석 가능성과 모형 간 비교의 일관성을 고려하여 p=5를 공통 자기시차로 채택하였다.

〈표 7〉 건설경제지표의 기본모형 분석결과

구분	Y1=건설수주증가율			Y2=건설착공증가율			Y3=건설기성증가율		
	계수값	*t*값	*p*값	계수값	*t*값	*p*값	계수값	*t*값	*p*값
L1.Y	−0.843	−8.53	0.000***	−0.425	−5.27	0.000***	−0.878	−9.8	0.000***
L2.Y	−0.864	−8.42	0.000***	−0.240	−2.75	0.007***	−0.961	−8.46	0.000***
L3.Y	−0.624	−6.02	0.000***	−0.284	−3.11	0.002***	−0.548	−4.28	0.000***
L4.Y	−0.550	−8.23	0.000***	−0.272	−3.86	0.000***	−0.707	−8.63	0.000***
L5.Y	−0.487	−8.76	0.000***	−0.173	−2.79	0.006***	−0.611	−11.95	0.000***
L1.d_여신금리	0.534	2.09	0.039**	0.233	1.45	0.151	0.217	3.12	0.002***
L2.d_여신금리	−0.292	−1.22	0.226	−0.465	−2.75	0.007***	−0.050	−0.43	0.667
L3.d_여신금리	−0.245	−1.21	0.230	−0.179	−0.9	0.369	0.075	0.73	0.469
LD.광공업지수	0.004	0.55	0.583	−0.017	−5.57	0.000***	0.004	1.75	0.083*
L2D.광공업지수	0.008	1.27	0.208	−0.021	−6.83	0.000***	0.001	0.24	0.812
L3D.광공업지수	0.000	0.05	0.962	−0.013	−2.74	0.007***	−0.004	−1.68	0.096*
_cons	0.015	0.83	0.410	−0.004	−0.3	0.766	0.014	1.17	0.243

주: *, **, ***은 각 10%, 5%, 1% 유의수준에서 귀무가설을 기각할 수 있음을 의미.

한 동행·후행 지표가 아니라, 제조업 경기의 대체적 조정변수로 기능하고 있음을 보여준다.

건설기성증가율은 세 지표 중 가장 강한 관성을 보이며 단기 외생 변수보다는 이미 결정된 공정과 투입 경로에 따라 점진적으로 조정되는 특성이 두드러진다. 건설기성이 누적 공정의 결과물이라는 구조적 성격을 통계적으로 확인해 주는 결과이다. 따라서 기성 변동은 단기 거시 충격보다는 중·단기적인 정보 축적과 기대 변화가 누적될 때 영향을 받을 가능성이 크다.

요컨대 기본모형은 건설생산활동이 모두 단기 거시변수의 즉각적 충격보다는 자기상관과 조정 과정을 중심으로 움직인다는 점을 보여준다. 동시에 금리변화는 단계별로 상이한 방식으로 관측된다. 수주와 기성에서는 단기(1기)에서 유의한 반응이 나타나고 착공에서는 2기 시차에서 유의한 제약 효과(음의 영향)가 확인된다. 이는 건설경기가 단일한 반응함수가 아니라, 수주(계약)-착공(투자결정)-기성(실현/집행)의 서로 다른 시간구조 속에서 금융여건이 차별적으로 작동함을 시사한다. 이러한 결과는 이후 확장모형에서 CENI와 같은 정보·불확실성 변수가 추가될 경우 그 효과가 단순한 통제변수 보강이 아니라 추가적 설명력으로 평가될 수 있는 기반을 제공한다.

2) 확장모형 A 분석결과: 부실강도지수의 설명력 검증

〈표 8〉은 자기상관(내생적 조정)이 강한 건설생산지표에 뉴스기반 CENI_Int 신호가 추가 설명력을 갖는지를 검정한 확장모형 A1의 분석결과이다. 먼저 건설수주증가율에 대해 CENI_Int는 통계적으로 의미 있는 패턴을 보였다. 2기 전 CENI_Int가 수주증가율을 큰 폭으로 감소시킨 이래 1기 전 수주증가율을 이전 감소폭의 상당부분을 되돌리는 것으로 관측된다. 이 결과는 단순히 부실 뉴스가 증가하면 수주가 감소한다는 직선적 인과라기보다 발주 의사결정이 불확실성 신호에 노출될 때 물량·타이밍이 재배치되며 변동성이 커지고, 지연된 조정이 이어지는 현상으로 볼 수 있다. 수주는 착공·기성에 비해 계약 조정의 재량이 크고 발주·수주 타이밍이 정책·금융·심리 요인에 의해 좌우되기 쉽다. 뉴스기반 CENI_Int는 건설생산 과정 중 수주가 갖는 이 같은 특성에 부합하여 수주 단계의 변동을 매개하는 신

〈표 8〉 건설경제지표의 확장모형 A1(부실강도지수포함) 분석결과

구분	Y1=건설수주			Y2=건설착공			Y3=건설기성		
	계수값	*t*값	*p*값	계수값	*t*값	*p*값	계수값	*t*값	*p*값
L1.Y	−0.827	−8.320	0.000***	−0.428	−5.430	0.000***	−0.917	−11.310	0.000***
L2.Y	−0.850	−8.320	0.000***	−0.241	−2.660	0.009***	−1.029	−11.040	0.000***
L3.Y	−0.627	−5.830	0.000***	−0.281	−3.060	0.003***	−0.630	−6.710	0.000***
L4.Y	−0.553	−8.290	0.000***	−0.269	−3.500	0.001***	−0.768	−10.160	0.000***
L5.Y	−0.500	−9.240	0.000***	−0.172	−2.780	0.006***	−0.650	−12.530	0.000***
L1.d_여신금리	0.535	2.200	0.030**	0.233	1.430	0.156	0.222	3.640	0.000***
L2.d_여신금리	−0.301	−1.090	0.277	−0.463	−2.720	0.008***	−0.015	−0.130	0.901
L3.d_여신금리	−0.233	−1.160	0.249	−0.176	−0.880	0.380	0.093	1.010	0.313
LD.광공업지수	0.003	0.480	0.631	−0.016	−5.270	0.000***	0.004	1.570	0.120
L2D.광공업지수	0.008	1.140	0.256	−0.020	−6.780	0.000***	0.001	0.300	0.762
L3D.광공업지수	0.001	0.100	0.918	−0.013	−2.640	0.009***	−0.004	−1.530	0.129
L1.부실강도지수	7.209	2.140	0.034**	0.332	0.140	0.886	2.202	1.920	0.057*
L2.부실강도지수	−9.157	−2.350	0.021**	0.949	0.210	0.833	−1.054	−0.570	0.572
L3.부실강도지수	2.428	0.890	0.374	−0.842	−0.260	0.792	1.430	1.150	0.251
_cons	0.008	0.220	0.828	−0.010	−0.520	0.603	−0.028	−1.950	0.054

주: *, **, ***은 각 10%, 5%, 1% 유의수준에서 귀무가설을 기각할 수 있음을 의미.

호가 될 수 있음이 확인된다. 반면 금리나 산업생산과 같은 전통적 거시 통제변수는 뚜렷한 영향이 제한적이어서 단기 수주 변동이 거시 충격의 기계적 전이보다 정보 · 기대 · 불확실성 경로(뉴스 채널)에 더 민감할 여지를 남긴다.

건설착공증가율에 대한 뉴스지수의 유의성은 전반적으로 확인되지 않는다. 착공은 수주보다 한 단계 후행하며 실제로 인허가 · PF조달 · 현장준비 등 의사결정의 제약을 강하게 받는다. 이 단계에서는 뉴스 신호가 존재하더라도 그것이 착공 변동으로 전이되기보다 행정 · 금융 · 기술적 제약과 일정 관리에 의해 흡수될 가능성이 크다. 실제로 착공에서 산업생산 변화는 강한 유의성을 보인 반면, 뉴스지수는 유의하지 않아 착공의 단기 변동은 정보 충격보다 실물 국면과 실행 제약(경기 · 자금 · 제도)에 더 직접적으로 묶여 있음을 보여준다.

건설기성증가율에 대한 뉴스지수의 영향은 대체로 강건하지 않은 가운데 1기 전 시차에서 경계적 수준에 그친다. 기성은 이미 결정된 공정의 누적 결과로 단기적으로 공정이 급격히 변동되기 어렵고 뉴스가 확산되더라도 계약 해지 · 중단 · 속도 조절로 이어지기까지 상당한 마찰과 시간이 필요하다.

확장모형 A1의 분석결과는 뉴스기반 CENI_Int가 건설생산 전반을 일률적으로 설명하기보다 발주 · 수주 단계(의사결정 단계)의 변동을 포착하는 설명력이 상대적으로 두드러진다는 점을 알려준다.

확장모형 A1에서 확인된 뉴스기반 CENI_Int의 설명력은 뉴스지수의 독자적 성분에 의한 것인가? 아니면 단순히 건설업 부실 등 유사성격의 변수를 대리한 것인가? 만약 뉴스지수가 단순히 유사성격의 변수를 대리하는 신호라면 유사변수와 동시에 통제했을 때 뉴스지수의 유의성은 급격히 약화되거나 소멸할 가능성이 크다. 〈표 9〉는 이를 확인하기 위해 뉴스기반 CENI_Int와 유사한 성격의 통제변수로 건설보증부도율을 추가하여 뉴스지수 설명력을 점검한 확

〈표 9〉 건설경제지표의 확장모형 A2(부실강도지수, 건설보증부도율 포함) 분석결과

구분	Y1=건설수주			Y2=건설착공			Y3=건설기성		
	계수값	*t*값	*p*값	계수값	*t*값	*p*값	계수값	*t*값	*p*값
L1.Y	−0.817	−8.950	0.000***	−0.432	−5.380	0.000***	−0.912	−10.640	0.000***
L2.Y	−0.853	−7.230	0.000***	−0.241	−2.630	0.010***	−1.023	−10.590	0.000***
L3.Y	−0.623	−5.500	0.000***	−0.309	−2.820	0.006***	−0.622	−5.680	0.000***
L4.Y	−0.582	−9.300	0.000***	−0.261	−3.490	0.001***	−0.769	−9.370	0.000***
L5.Y	−0.506	−11.720	0.000***	−0.180	−3.190	0.002***	−0.652	−11.880	0.000***
L1.d_여신금리	0.454	2.010	0.047**	0.206	1.190	0.235	0.218	3.560	0.001***
L2.d_여신금리	−0.305	−1.050	0.297	−0.425	−2.250	0.026**	−0.015	−0.130	0.900
L3.d_여신금리	−0.183	−0.790	0.429	−0.194	−0.940	0.350	0.094	1.010	0.315
LD.광공업지수	0.003	0.470	0.642	−0.016	−4.500	0.000***	0.004	1.550	0.124
L2D.광공업지수	0.007	0.940	0.352	−0.020	−5.880	0.000***	0.001	0.270	0.790
L3D.광공업지수	0.001	0.130	0.898	−0.012	−2.330	0.021**	−0.004	−1.460	0.147
L1.부실강도지수	7.169	2.130	0.036**	0.008	0.000	0.997	2.194	1.860	0.065*
L2.부실강도지수	−10.510	−2.610	0.010***	1.336	0.300	0.762	−1.150	−0.590	0.555
L3.부실강도지수	2.293	0.810	0.422	−1.353	−0.440	0.662	1.365	1.050	0.298
LD.부도율	−23.864	−1.020	0.308	3.773	0.210	0.834	−3.094	−0.300	0.768
LD2.부도율	−3.961	−0.180	0.858	−21.991	−1.180	0.240	−2.238	−0.260	0.793
LD3.부도율	−32.999	−1.670	0.098*	1.988	0.130	0.900	−1.045	−0.270	0.785
_cons	0.041	1.380	0.170	−0.001	−0.070	0.943	−0.024	−1.880	0.063

주: *, **, ***은 각 10%, 5%, 1% 유의수준에서 귀무가설을 기각할 수 있음을 의미.

장모형 A2의 분석결과이다.

〈표 9〉에서 보듯이 건설수주증가율에 대해서 부도율 변화의 자기시차를 추가하더라도 뉴스지수의 시차효과 패턴이 크게 훼손되지 않는다. 동시에 부도율 변화 자체는 통상적으로 유의성이 약하거나 경계적 수준에 머무른다. 이는 뉴스지수가 단순히 부도율 정보를 대신 담고 있는 것이 아니라, 부도율 변화와는 다른 의사결정 단계의 정보(불확실성 · 리스크 인식 · 시장 심리)로서 수주증가율 변동에 독자적으로 작용하고 있음을 시사한다.

착공 및 기성증가율에 대해 부도율 변화를 추가로 통제해도 뉴스지수의 비유의성이 지속된다. 이는 두 가지 방식으로 해석할 수 있다. 첫째, 착공 · 기성 단계는 구조적으로 실행 제약과 공정 누적의 관성이 강해 뉴스나 부도율과 같은 위험 신호가 존재하더라도 단기 변동으로 전환되기까지 지연이 크다. 둘째, 뉴스지수가 실제로 영향을 미치더라도 그 영향은 비선형적(임계치) 또는 국면 의존적으로 나타날 가능성이 있다. 요컨대 부도율로 추가 통제한 확장모형 A2의 분석결과는 뉴스지수가 유사성격의 대리변수의 성격과는 다른 독자적인 성분으로 작용할 수 있음을 시사한다.

3) 확장모형 B 분석결과: 건설뉴스지수의 분해[부실강도, 부정정서, 상호작용]

본 절에서는 뉴스효과를 CENI_Int와 SENT_Neg으로 분해[10]하고 양자의 상호작용을 통해 뉴스충격의 비선형 구조를 분석한다. 〈표 10〉은 이를 추정한 확장모형 B의 분석결과이다. 건설수주의 경우 CENI_Int는 2개월 시차에서 유

〈표 10〉 확장모형 B(건설뉴스지수의 분해) 분석결과

구분	Y1=건설수주			Y2=건설착공			Y3=건설기성		
	계수값	*t*값	*p*값	계수값	*t*값	*p*값	계수값	*t*값	*p*값
L1.	−0.735	−7.670	0.000***	−0.386	−5.160	0.000***	−0.873	−10.210	0.000***
L2.	−0.821	−8.070	0.000***	−0.170	−1.650	0.101	−0.989	−9.900	0.000***
L3.	−0.618	−5.710	0.000***	−0.297	−3.010	0.003***	−0.643	−7.530	0.000***
L4.	−0.535	−8.210	0.000***	−0.296	−3.450	0.001***	−0.784	−10.730	0.000***
L5.	−0.484	−9.090	0.000***	−0.186	−2.880	0.005***	−0.642	−12.130	0.000***
L1.d_여신금리	0.508	2.080	0.040**	0.189	1.010	0.315	0.231	3.840	0.000***
L2.d_여신금리	−0.209	−0.680	0.500	−0.322	−1.990	0.049**	0.002	0.020	0.982
L3.d_여신금리	−0.245	−1.310	0.193	−0.195	−1.010	0.314	0.099	1.120	0.267
LD.광공업지수	0.002	0.210	0.831	−0.018	−4.710	0.000***	0.003	1.120	0.264
L2D.광공업지수	0.006	0.860	0.393	−0.023	−6.280	0.000***	0.000	−0.060	0.954
L3D.광공업지수	0.001	0.200	0.845	−0.011	−2.180	0.032**	−0.004	−1.550	0.124
L1.부실강도지수	27.079	1.270	0.208	−11.590	−0.720	0.475	−1.100	−0.210	0.838
L2.부실강도지수	−50.064	−2.010	0.047**	−30.682	−1.990	0.049**	−21.965	−3.170	0.002***
L3.부실강도지수	−6.422	−0.240	0.812	9.110	0.510	0.611	−1.320	−0.240	0.807
L1.부정정서지수	−0.153	−0.330	0.746	−0.312	−1.090	0.280	−0.154	−1.460	0.148
L2.부정정서지수	−0.170	−0.310	0.755	−0.176	−0.460	0.644	−0.280	−2.120	0.036**
L3.부정정서지수	−0.415	−0.990	0.323	−0.229	−0.680	0.496	−0.030	−0.210	0.831
L1.뉴스지수상호작용	−28.191	−0.930	0.357	19.060	0.790	0.432	4.349	0.570	0.568
L2.뉴스지수상호작용	57.605	1.560	0.121	43.558	2.310	0.023**	30.348	3.200	0.002***
L3.뉴스지수상호작용	13.561	0.330	0.742	−14.106	−0.590	0.558	4.070	0.510	0.610
_cons	0.523	0.820	0.415	0.490	1.560	0.121	0.299	1.810	0.074*

주: *, **, ***은 각 10%, 5%, 1% 유의수준에서 귀무가설을 기각할 수 있음을 의미.

의한 음(−)의 효과를 보이는 반면, 부정정서 및 상호작용항은 통계적으로 유의하지 않았다. 이는 수주가 뉴스의 정서적 성격보다 정보 노출의 양적 확대, 즉 보도 강도 자체에 더 민감하게 반응하는 초기 단계의 의사결정 변수임을 시사한다.

반면 착공과 기성에서는 강도와 부정정서의 결합효과가 뚜렷하게 관측된다. 착공의 경우 뉴스강도는 2개월 시차에서 위축 효과를 보이나, 부정정서비중과의 상호작용항이 양(+)의 값을 갖는 것으로 나타났다.

착공과 기성에서 상호작용항이 양(+)의 값을 갖는 것은 정서지수가 음(−)의 값을 중심으로 분포한다는 점을 고려할 때, 부정적 정서가 심화될수록 CENI_Int의 한계효과가 체감되는 구조로 해석할 수 있다. 이는 부정적 뉴스환경이 이미 형성된 상황에서는 추가적인 부실 뉴스가 새로운 정보로 작용하기보다 기존 기대에 내재되어 해석되기 때문으

10) 뉴스강도(Intensity)는 부정 뉴스의 절대적 발생량을 기반으로 한 사건 중심 지표로 정보 충격의 크기를 측정하는 반면, 뉴스정서(Tone)는 부정 뉴스의 상대적 비중 또는 긍·부정 균형을 통해 동일한 정보에 대한 시장의 해석 방향을 포착하는 지표이다. 뉴스강도가 '얼마나 많은 문제가 발생했는가'를, 뉴스정서는 '그 문제가 얼마나 비관적으로 인식되는가'를 각각 반영한다. 따라서 부정 뉴스가 많이 발생하더라도 긍정 보도가 병존하면 강도는 크지만 정서는 중립적일 수 있으며, 반대로 부정 보도만 존재할 경우 강도는 작지만 정서는 극단적으로 부정적일 수 있다.

로 볼 수 있다.

종합하면, 뉴스효과는 단순한 선형 충격이 아니라 강도와 정서 톤의 결합을 통해 단계별로 상이하게 작동하며 특히 착공과 기성 단계에서 뉴스 충격의 한계효과가 부정정서 환경에 따라 조정되는 비선형 구조가 존재함을 알 수 있다.

4) 경기모형 분석결과: 경기주기별 뉴스지수의 설명력 비교

〈표 11〉은 건설뉴스지수의 경기주기별 효과를 분석한 결과이다. 분석기간(2015~2025)의 건설경기변동을 4개의 경기국면(회복-호황-후퇴-침체)으로 분류[11]하여 각 국면별 뉴스지수의 설명력을 SENT_Neg로 확인하였다. 분

〈표 11〉 경기국면별 뉴스지수 효과 분석 결과

구분	Y1=건설수주			Y2=건설착공			Y3=건설기성		
	계수값	t값	p값	계수값	t값	p값	계수값	t값	p값
L1.	−0.880	−9.910	0.000***	−0.399	−4.060	0.000***	−0.993	−11.950	0.000***
L2.	−0.883	−8.420	0.000***	−0.254	−2.780	0.007***	−1.065	−11.830	0.000***
L3.	−0.685	−7.680	0.000***	−0.298	−2.740	0.007***	−0.695	−6.870	0.000***
L4.	−0.620	−8.940	0.000***	−0.318	−4.350	0.000***	−0.785	−10.290	0.000***
L5.	−0.483	−7.220	0.000***	−0.214	−2.790	0.006	−0.696	−11.180	0.000***
L1.d_여신금리	0.641	2.560	0.012**	0.248	1.280	0.204	0.109	1.240	0.219
L2.d_여신금리	−0.237	−1.050	0.298	−0.239	−1.080	0.284	0.027	0.260	0.799
L3.d_여신금리	−0.260	−1.030	0.305	−0.323	−1.670	0.099*	−0.005	−0.070	0.946
LD.광공업지수	0.003	0.390	0.696	−0.018	−5.910	0.000***	0.004	1.700	0.093*
L2D.광공업지수	0.005	0.820	0.414	−0.022	−6.640	0.000***	0.001	0.390	0.700
L3D.광공업지수	0.001	0.120	0.903	−0.012	−2.320	0.022**	−0.003	−1.320	0.188
L1부정정서_Recov	−0.136	−2.430	0.017**	−0.043	−0.690	0.492	−0.024	−1.100	0.274
L2부정정서_Recov	0.068	0.900	0.370	0.121	4.100	0.000***	0.016	0.710	0.480
L3부정정서_Recov	0.076	1.230	0.222	−0.056	−1.410	0.162	0.021	1.310	0.194
L1부정정서_Boom	−0.145	−3.650	0.000***	−0.044	−0.750	0.454	−0.010	−0.570	0.573
L2부정정서_Boom	0.045	0.580	0.563	0.075	2.530	0.013**	0.003	0.160	0.876
L3부정정서_Boom	0.111	2.070	0.041**	−0.010	−0.270	0.791	0.018	1.020	0.312
L1부정정서_Reces	−0.194	−4.850	0.000***	−0.083	−1.500	0.137	−0.037	−2.080	0.040**
L2부정정서_Reces	0.062	0.710	0.482	0.133	5.290	0.000***	0.005	0.280	0.781
L3부정정서_Reces	0.139	2.230	0.028**	−0.031	−0.790	0.431	0.029	1.750	0.083*
L1부정정서_Downtn	−0.120	−2.360	0.020**	−0.064	−1.100	0.273	−0.033	−2.040	0.044**
L2부정정서_Downtn	0.053	0.720	0.475	0.162	6.240	0.000***	0.024	1.210	0.228
L3부정정서_Downtn	0.084	1.110	0.271	−0.076	−1.940	0.055*	0.013	0.650	0.520
_cons	−0.032	−0.100	0.922	−0.113	−0.710	0.477	−0.015	−0.310	0.758

주: *, **, ***은 각 10%, 5%, 1% 유의수준에서 귀무가설을 기각할 수 있음을 의미.

11) 본 분석에 활용한 분석기간 중 경기주기별 기간은 김갑진(2025b)의 연구에서 활용한 건설경기 분석(건설생산지표별 HP필터 분해)을 업데이트하여 다음과 같이 적용하였다. 회복기(2015.3분기~2016.2분기, 2021.2분기~2022.3분기), 호황기(2016.3분기~2017.4분기, 2022.4분기~2024.1분기), 후퇴기(2018.1분기~2019.3분기, 2024.2분기~2025.4분기), 불황기(2015.1분기~2015.2분기, 2019.4분기~2021.1분기).

석결과 건설수주의 경우 전체 경기주기에 걸쳐 부정뉴스의 1개월 시차 효과가 일관되게 유의한 음(-)의 값을 보였다. 경기전환 초기국면에 감소한 수주는 이후 3기 시차까지 회복되는 패턴을 보였다. 이는 확장모형의 결과에서 본 것처럼 건설수주가 단기적인 심리 및 정보 충격에 반응하는 선행적 성격의 지표임을 재확인하는 결과이다.

착공은 수주와 달리 즉각적인 반응보다는 2개월 시차에서 경기주기별로 일관된 반응이 관측된다. 이는 착공이 이루어지기까지 인허가, 자금조달, 시공준비 등 제도적 · 기술적 절차가 필요하여 뉴스충격이 착공지표에 반영되기까지 구조적인 시차가 존재함을 의미한다.

마지막으로 기성의 경우 부정뉴스 효과는 주로 후퇴기와 불황기에서만 단기적으로 유의하게 나타났다. 이는 평상시에는 진행 중 공정의 관성이 기성의 단기 변동성을 완충하지만 경기침체 국면에서 자금 제약과 위험관리 강화로 인해 뉴스충격이 실물생산으로 보다 직접적으로 전이됨을 추정할 수 있다. 결론적으로 경기주기별 뉴스효과는 경기 전반에 균질하게 작동하지 않으며 뚜렷한 비대칭성을 보였는데 경기상승 국면에서 뉴스의 설명력이 제한적인 반면, 하강 국면에서 뉴스 강도 및 SENT_Neg가 건설경기에 유의한 영향을 미치는 것으로 나타났다.

Ⅴ. 결론

본 연구는 뉴스 빅데이터를 기반으로 빈도와 감성기반의 CENI를 구축하고 건설실물 지표에 대해 유의한 설명력을 갖는지 검증하였다. 분석 결과 부실강도지수(CENI_Int)와 정서지수(CENI_Sent)는 건설수주에 대해 통계적으로 유의한 인과관계를 보이는데 반해, 착공 및 기성과 같은 생산 단계 변수에는 직접적인 영향이 제한적인 것으로 나타났다. 이는 건설경기 변동이 뉴스 기반 위험 인식과 기대 변화 → 발주 및 계약 단계 → 생산 단계로 이어지는 전이 구조를 갖는다는 점을 시사한다.

한편 뉴스지수는 경기주기별로 뚜렷한 비대칭성을 보였는데 경기상승 국면에서 뉴스의 설명력이 제한적인 반면, 하강 국면에서 뉴스 강도 및 뉴스의 부정정서가 건설경기에 유의한 영향을 미치는 것으로 나타났다. 이는 건설경기가 불황기에 들어설수록 실물 변수보다 정보 및 심리 변수에 대한 민감도가 증가함을 의미하며, 건설경기의 구조적 불안정성이 정보 채널을 통해 확대될 수 있음을 시사한다.

이러한 결과는 건설경기 모니터링 체계가 기존의 실물지표 중심 접근에서 벗어나, 뉴스 기반 지표를 결합한 이중적(dual-layer) 모니터링 체계로 확장될 필요가 있음을 보여준다. 전통적인 건설수주, 착공, 기성 지표는 경기의 방향성과 규모를 비교적 정확히 반영하지만 공표 시차와 누적적 성격으로 인해 경기 전환 신호를 신속히 포착하는 데 한계를 지닌다. 반면, 뉴스지수는 시장 참여자의 기대와 위험 인식이 반영된 고빈도 정보로서 경기 전환 이전 단계에서 형성되는 신호를 포착할 수 있다.

따라서 정책적으로는 두 지표를 대체 관계가 아닌 보완적 정보 체계로 통합할 필요가 있다. 구체적으로는 (1) 뉴스지수를 활용한 조기경보 단계(early warning layer)와, (2) 실물지표를 활용한 확인 및 집행 단계(confirming layer)를 구분하여 운용하는 것을 고려할 수 있다. 예를 들어, 부실강도지수(CENI_Int)의 급등이나 정서지수(CENI_Sent)의 급격한 악화가 관측될 경우 이를 경기 하방 위험의 초기 신호로 인식하고, 이후 일정 시차를 두고 나타나는 수주 · 착공 · 기성 지표의 변화를 통해 그 현실화 여부를 점검하는 방식이다. 이러한 단계적 접근은 경기 판단의 시차를 줄이는 동시에 과잉 반응의 위험을 완화할 수 있다.

금융기관 및 보증기관의 활용 측면에서도 뉴스지수는 유용한 보조 지표로 기능할 수 있다. 건설 · 부동산 프로젝트는 금융조달과 기대 형성에 크게 의존하는 특성을 가지므로, 부실 · 위험 관련 뉴스의 증가는 단순한 정보의 축적을 넘어 신용위험 확대의 선행 신호로 해석될 수 있다. 이에 따라 금융기관은 여신 심사 및 포트폴리오 관리 과정에서 뉴스지수를 활용하여 특정 시기 또는 특정 부문에서의 위험 집중도를 사전에 점검할 수 있으며, 보증기관 역시 보증공급의 속도와 규모를 조정하는 데 참고지표로 활용할 수

있다. 특히 뉴스지수는 기존의 재무지표 기반 리스크 관리 체계를 일정부분 보완하여 시의성 있는 위험 감지에도 기여할 수 있을 것으로 사료된다.

다만 뉴스지수는 본질적으로 비정형 정보에 기반한 지표로서 과잉 반응 또는 일시적 노이즈를 포함할 가능성이 있다. 따라서 단일 지표에 의존하기보다 실물지표 및 금융지표와 결합한 종합적 판단 체계 내에서 활용하는 것이 필요하다. 이러한 점에서 본 연구는 뉴스 기반 지표가 기존 건설경기 분석 틀을 대체하기보다 이를 보완하는 선행적 정보 레이어로 기능할 수 있음을 제시한다는 데 의의가 있다. 향후 연구에서 뉴스를 보다 입체적으로 분해하고, 건설사업 유형별 영향력을 분석하는 등 연구의 깊이를 제고할 필요성이 있다.

참고문헌

김갑진. (2025a). 건설·부동산 PF위험의 시스템위험 전이 현상 분석과 정책대안. *금융공학연구*, *24*(1), 173-203.

김갑진. (2025b). 경기상태별 부동산 개발사업의 사고율 특성과 건설·부동산 경제에 미치는 영향분. *주택도시금융연구*, *10*(1), 5-30.

서범석. (2022.5.16). *뉴스 텍스트를 이용한 경기 예측: 경제 부문별 텍스트 지표의 작성과 활용*. 한국은행. https://www.bok.or.kr/portal/bbs/P0002353/view.do?nttId=10070554&menuNo=200433

서범석, 이영환, 조형배. (2022). 기계학습을 이용한 뉴스심리지수(NSI)의 작성과 활용. 한국은행 (편저). *국민계정리뷰*.

송민채, 신경식. (2017). 뉴스기사를 이용한 소비자의 경기심리지수 생성. *지능정보연구*, *23*(3), 1-27.

이성민, 김성현. (2019). 텍스트 마이닝을 활용한 경기동향 분석. *응용경제*, *21*(3), 1-27.

최재원, 김도형. (2021). 텍스트 기반 경제심리지수의 구성과 경기변동 설명력. *한국경제학보*, *69*(1), 91-123.

한국은행. (2024a). *금융안정보고서(2024년 6월)*. 한국은행.

한국은행. (2024b). *금융안정보고서(2024년 12월)*. 한국은행.

황순주. (2024). 갈라파고스적 부동산 PF, 근본적 구조개선 필요. *KDI FOCUS 제134호*. 한국개발연구원 (KDI).

Baker, S. R., Bloom, N. & Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, *131*(4), 1593-1636.

Barrie, D. S., & Paulson, B. C. (1992). *Professional construction management: Including CM, design-construct, and general contracting*. McGraw-Hill.

Bernanke, B. S., Gertler, M., & Gilchrist, S. (1999). The financial accelerator in a quantitative business cycle framework. In Taylor J. B., & Woodford M. (Eds.), *Handbook of Macroeconomics* (Vol. 1C, pp. 1341-1393), Elsevier.

Claessens, S., & Kose, M. A. (2017). Asset prices and macroeconomic outcomes: A survey. (BIS Working Paper No. 676), Bank for International Settlements.

Gentzkow, M., Kelly, B., & Taddy, M. (2019). Text as data. *Journal of Economic Literature*, *57*(3), 535-574.

Kiyotaki, N., & Moore, J. (1997). Credit cycles. *Journal of Political Economy, 105*(2), 211-248.

Leamer, E. E. (2007). *Housing is the business cycle* (Working Paper 13428). National Bureau of Economic Research (NBER).

Mian, A. & Sufi, A. (2014). *House of debt: How they*

(and you) caused the great recession, and how we can prevent it from happening again, University of Chicago Press.

Mian, A., Sufi, A., & Verner, E. (2017). Household debt and business cycles worldwide. *The Quarterly Journal of Economics*, *132*(4), 1755-1817.

OECD. (2019). *OECD business and finance outlook 2019: Strengthening trust in business*. OECD Publishing, Paris. https://doi.org/10.1787/af784794-en

Shapiro, A. H., Sudhof, M., & Wilson, D. J. (2022). Measuring news sentiment. *Journal of Econometrics*, *228*(2), 221-243.

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.

Thorsrud, L. A. (2020). Words are the new Numbers: A new approach to forecasting with text data. *Journal of Business & Economic Statistics*, *38*(2), 393-409.

논문접수일: 2026.02.03
논문심사일: 2026.04.27
게재확정일: 2026.07.29

국문 요약

본 연구는 건설 관련 뉴스 빅데이터를 활용하여 건설경제 뉴스지수(construction economic news index, 이하 CENI)를 구축하고, 해당 지수가 전통적인 건설경기 지표에 대해 갖는 정보력을 실증적으로 분석한다. 기존 건설경기 분석은 수주, 착공, 기성 등 월·분기 단위 집계지표에 의존함으로써 경기 전환 국면에서 발생하는 인식 변화와 위험 신호를 적시에 포착하는 데 한계를 지닌다. 이에 본 연구는 GDELT 글로벌 뉴스 데이터를 활용하여 부실강도지수와 정서지수로 구성된 이중 CENI를 설계하고, 2015~2025년 월별 자료를 이용해 동태적 회귀모형과 경기국면별 모형을 통해 분석하였다. 분석결과, CENI는 기존 거시·금융 변수를 통제한 이후에도 건설수주 변동에 대해 유의한 선행적 설명력을 보였으며, 특히 경기 후퇴 및 침체 국면에서 그 정보력이 강화되는 것으로 나타났다. 반면 착공과 기성에 대해서는 뉴스지수의 효과가 제한적으로 나타나, 뉴스 기반 정보가 주로 발주·의사결정 단계에서 경기 인식을 반영하는 특성을 지님을 확인하였다. 본 연구는 뉴스 빅데이터가 건설경기 판단과 위험 모니터링을 위한 보완적 지표로 활용될 수 있음을 시사한다.

주제어: 건설경제 뉴스지수(CENI), 뉴스 빅데이터, 건설경기, 동태적 회귀모형(DRML), 부동산뉴스

© Copyright 2026 Korea Housing & Urban Guarantee Corporation. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.