

주택금융연구

VOL. 10 NO. 1

1. 구조적 토픽모형(Structural Topic Model)을 활용한
주택연금 정책담론 분석: 2007~2026년 뉴스 보도의
토픽 변동과 매체별 프레임링 차이

이승형

2. 주택시장에서 가구의 차입금 상환부담과
추가적인 부채의 관계에 대한 연구

김종희

3. 트리 앙상블 기반 자동산정모형을 이용한 재고 주택가격
분포 관측: 전국 아파트 단지 전수 자료의 시계열 분석

홍정의

4. 아파트 거래량과 가격 변동성의 관계:
거래량 기반 가격 불확실성 지수의 구축과 활용

김천일

5. 주거 점유 형태 전이 행렬을 통해 본 주거 사다리의
분절 현상 분석

정경채 · 박정은

부록

- 주택금융연구의 내용을 인용할 때에는 반드시 “한국주택금융공사 주택금융연구 ○권 ○호 ○○페이지에서 인용”이라고 표시하여 주십시오.
- 주택금융연구에 대한 질의 또는 제안은 주택금융연구원(☎ 051-663-8178)으로 연락하여 주시기 바랍니다.
- 수록된 논문은 투고자의 개인 의견으로 한국주택금융공사의 견해와 무관합니다.

목 차

구조적 토픽모형(Structural Topic Model)을 활용한 주택연금 정책담론 분석: 2007~2026년 뉴스 보도의 토픽 변동과 매체별 프레이밍 차이 Analyzing Housing Pension Policy Discourse Using Structural Topic Models: Topic Dynamics and Media-Type Framing Differences in News Coverage, 2007-2026 이승형	5
주택시장에서 가구의 차입금 상환부담과 추가적인 부채의 관계에 대한 연구 The Relationship between Repayment Burden and Additional Debt in Households 김종희	35
트리 앙상블 기반 부동산정보형을 이용한 재고 주택가격 분포 관측: 전국 아파트 단지 전수 자료의 시계열 분석 A Tree Ensemble Approach to Estimating Census-Level House Price Distributions 홍정의	61
아파트 거래량과 가격 변동성의 관계: 거래량 기반 가격 불확실성 지수의 구축과 활용 Apartment Transaction Volume and Price Volatility: Construction and Application of a Transaction-Volume-Based Price Uncertainty Index 김천일	133
주거 점유 형태 전이 행렬을 통해 본 주거 사다리의 분절 현상 분석 The Paradox of Housing Finance Policy and the Broken Housing Ladder: Application of a Markov Chain Model and Structural Transition Sensitivity Analysis 정경채 · 박정은	165
부록	197

구조적 토픽모형(Structural Topic Model)을 활용한 주택연금 정책담론 분석: 2007~2026년 뉴스 보도의 토픽 변동과 매체별 프레이밍 차이

이승형*

요약

본 연구는 주택연금 정책담론을 구조적 토픽모형(structural topic model, STM)으로 분석한다. 빅카인즈의 2007년 7월~2026년 3월 주택연금 뉴스 12,799건에서 SimHash 중복 제거를 거쳐 12개 토픽을 도출하였다. 담론은 정책·가계부채(13.4%), 자산관리·은퇴설계(11.9%), 가입실적·통계(10.1%) 등으로 구성되었다. 이명박은 가입실적(18.0%), 문재인은 제도 설계(13.2%), 이재명은 고령가구·상속의향(10.5%)이 두드러졌고, 매체별로는 지역지가 가입실적, 종합일간지가 노후·세대 서사, 경제지가 은퇴설계에 집중하였다. 보증료 인하 직후 고령가구·상속의향이 +25.65%p 급증했으나 사후 표본 31건으로 잠정적이다. KNU와 KoELECTRA의 감성 분류 일치도는 49.0%로 정책 뉴스 도메인 갭이 확인되었다. 정부별 추정치는 인과가 아닌 기간별 기술 통계이다. 정책 담론에 STM을 적용한 분석 프레임워크를 제시한다.

핵심어 : 주택연금, 구조적 토픽모형(Structural Topic Model, STM), 정책담론, 미디어 프레이밍, 텍스트마이닝

I. 서론

한국주택금융공사가 2007년 7월 출시한 주택연금은 19년째 운영되고 있다. 2025년 12월 기준

* 이승형, 주저자, 동아대학교 교육학과 박사, homoludus4780@gmail.com

© Copyright 2026 Housing Finance Research Institute. 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.

누적 가입자는 15만 명을 돌파하였고 월평균 수령액은 127만 원에 달하나, 적격 대상 대비 가입률은 2%~2.5%로 정부 목표(3%)에 미달한다. 출시 이래 가입연령·가격상한·지급방식의 10차례 이상 개편에도 이 정체가 지속되는 가운데, 2026년에는 출시 이래 최대 규모의 개편이 진행 중이다. 제도 변천의 궤적과 2026년 개편의 세부 내용은 II.1절에서 상술한다. 이 시점은 그간 누적된 정책담론의 구조와 변동을 체계적으로 분석할 학술적 필요성을 제기한다.

주택연금의 정책적 중요성은 한국 사회의 인구구조 변화에서 비롯된다. 한국은 2024년 12월 초고령사회에 진입하였으며(65세 이상 1,024만 명, 전체의 20.0%; 행정안전부, 2024), 노인빈곤율은 40.4%로 OECD 평균(14.2%)의 약 3배에 달한다(OECD, 2023). 고령가구의 자산 구성은 부동산에 편중되어 소득은 부족하나 주거 자산은 보유한 자산형 빈곤(house rich, cash poor) 상태가 보편적이다. 주택연금은 이 구조적 불일치를 해소하는 핵심 정책 수단으로 설계되었으나, 출시 이후 가입률 정체가 지속되고 있다.

이 간극의 원인을 규명하려는 연구는 주로 가구 단위 설문조사나 행정 데이터를 활용한 로짓모형·DID 분석으로 가입 결정요인을 탐색하는 접근을 취해 왔다. 가입연령과 주택가격, 상속 의지 등이 주요 변수로 반복 확인되었으며(전병욱, 2022; 주택금융연구원, 2020), 행동경제학적 관점에서 최적의사결정 모델보다 임의적 판단이 지배한다는 점도 보고된 바 있다(전병욱, 2022; 주택금융연구원, 2020). 그러나 이러한 개인 수준의 분석은 주택연금에 대한 사회적 인식이 어떻게 형성·변화해 왔는지를 포착하지 못한다. 가입 결정은 진공 상태에서 이루어지지 않으며, 언론 보도를 통해 매개된 정보 환경이 잠재 가입자의 인식과 태도를 조건짓는다. 예컨대, 주택연금이 노후 빈곤 대응 수단으로 보도되는지 혹은 집값 하락 리스크를 수반하는 금융 계약으로 보도되는지에 따라 잠재 가입자가 수용하는 제도 이미지가 달라질 수 있다. 미디어 프레이밍 이론에 따르면, 동일한 정책이라도 어떤 측면이 부각되느냐에 따라 수용자의 태도와 행동이 달라진다(Entman, 1993; Scheufele, 2000). 그럼에도 주택연금 보도의 담론 구조를 19년 장기 시계열에 걸쳐 체계적으로 분석한 학술 연구는 국내 학술 데이터베이스(KCI·DBpia) 검색 결과 확인되지 않는다.

이 공백을 메우기 위해 본 연구는 구조적 토픽모형(structural topic model, STM)을 활용한다. STM은 대규모 텍스트 코퍼스에서 잠재 토픽을 추출하는 동시에, 문서 수준의 메타데이터(시간, 매체 특성)가 토픽 비중(topic prevalence)과 단어 분포에 미치는 영향을 추정할 수 있는 확률적 텍스트 분석 모형이다(Roberts et al., 2014). 국내외 정책 영역에서 STM의 활용은 점차 확장되고 있으며, 환경 영역에서는 미세먼지 재해 보도의 프레이밍 추출(이준웅·김성희, 2018), 보건 영역에서는 코로나 19 초기 온라인 정보 교환과 불안 확산 분석(Jo et al., 2020) 등이 대표적이다. 그러나 주택금융 분야에 STM이 적용된 사례는 학술 데이터베이스(KCI·DBpia·Web of Science·Scopus) 검색 결과 확인되지 않았다. 이 연구는 STM의 메타데이터 효과 추정 기능을 활용하여 다음 두 가지 연구 질문에 답한다.

RQ1. 주택연금 뉴스 담론은 어떤 토픽으로 구성되며, 시기별로 어떻게 변화하였는가?

RQ2. 정부 및 매체유형에 따라 각 토픽의 비중은 어떻게 달라지며, 이러한 차이는 정책담론의 어떤 분화 양상을 드러내는가?

이 학술지에서 텍스트마이닝 기반 담론 분석이 시도된 적이 없고, 주택연금을 미디어 프레이밍 관점에서 분석한 국내 학술연구도 확인되지 않는다는 점에서, 2007년 7월 제도 출시부터 2026년 3월 보증료 인하 직후까지 약 19년·6개 정부·10회 이상의 주요 제도 개편을 포괄하는 이 연구는 방법론·주제·시간 범위의 세 차원에서 기존 연구 공백을 보완한다. 디딤돌 대출 정책 효과 분석(김강석·박정은, 2025) 등 최근 호에서 방법론적 다양화가 시도되고 있으나, NLP 기반 정책담론 분석은 이 학술지에서 공백으로 남아 있다. 분석 대상 코퍼스는 빅카인즈에서 수집한 12,799건의 뉴스 보도이다. 다만 2026년 표본은 1월 1일부터 3월 19일까지의 부분년도(195건)이며, 보증료 인하(2026년 3월 1일) 직후 사건 분석은 사후 19일·31건이라는 표본 제약 아래 잠정적 결과로 해석한다.

이 분석의 결과는 정책담론의 구조적 변동 패턴을 가입률 정체의 거시적 맥락 위에 위치시키며, 주택금융 정책담론에 구조적 토픽모형을 적용한 분석 프레임워크를 제시함으로써 기존 계량경제학적 접근과 상호 보완적인 방법론적 확장에 기여한다. 이하에서 이론적 배경, 분석 설계, 결과, 함의를 순차적으로 논의한다.

II. 이론적 배경 및 선행연구

1. 주택연금 제도의 변천과 정책적 맥락

주택연금(역모기지, reverse mortgage)은 고령 주택 소유자가 자신의 주택을 담보로 제공하고 사망 시까지 매월 연금을 수령하는 금융상품이다. 한국에서는 한국주택금융공사가 2007년 7월 만 65세 이상, 시가 6억 원 이하 주택 소유자를 대상으로 출시하였다. 이후 제도는 점진적으로 확대되어 가입연령이 60세(2008년), 55세(2020년)로 하향되었고, 주택가격 상한도 9억 원(2009년), 12억 원(2023년)으로 상향되었다. 2014년에는 1.5억 원 이하 주택을 대상으로 우대형 주택연금이 도입되어 질적 다양화가 이루어졌다.

2026년에 이르러 제도는 출시 이래 가장 큰 규모의 개편을 맞이한다(금융위원회, 2026). 3월 1일부터 초기보증료가 1.5%에서 1.0%로 인하되었고, 월 수령액은 평균 3.13% 증가하였다. 연보증료는 재정 지속가능성을 위해 0.75%에서 0.95%로 소폭 인상되었다. 6월에는 자녀 승계 제도 신설,

실거주 요건 완화, 귀촌·인구감소지역 이주 예외 등이 시행될 예정이다. 자녀 승계 제도는 가입 장벽 조사에서 54.4%로 1순위를 차지하는 상속 의지 문제를 직접 겨냥한 정책 대응이라는 점에서 제도사적 전환점이 될 수 있다.

그러나 이러한 제도적 확대에도 불구하고 가입률은 적격 대상의 2%~2.5%에 머물고 있다. 선행연구는 이 간극의 원인을 개인 수준의 의사결정 변수에서 탐색해 왔다. 전병욱(2022)의 로짓모형 기반 가입 결정요인 분석에서는 가입연령, 주택가격, 자녀 유무, 금융이해력이 유의미한 설명변수로 보고되었으며, 행동경제학적 접근에서는 자산 편중과 상속 의지가 합리적 최적의사결정을 압도한다는 점이 지적되었다. 이호일 외(2025)는 고령가구의 자산구성이 삶의 만족도에 미치는 영향을 분석하여 부동산 자산 유동화의 필요성을 간접적으로 시사하였다. 이러한 연구는 미시적 의사결정 메커니즘을 밝히는 데 기여하였으나, 주택연금에 대한 사회적 담론의 형성과 변화를 포착하지는 못한다. 가입 결정이 언론 보도를 통해 매개된 정보 환경 속에서 이루어진다는 점을 감안하면, 정책담론 분석 프레임워크의 도입이 필요하다.

2. 미디어 프레이밍 이론과 토픽-프레임 매핑의 이론적 정당화

Entman(1993)에 따르면, 프레이밍은 (1) 문제 정의, (2) 인과적 해석, (3) 도덕적 평가, (4) 처방의 권고라는 네 가지 기능을 통해 특정 이슈의 일부 속성을 텍스트에서 두드러지게 만드는 과정이다. 이 정의는 프레이밍이 단순한 주제 선택을 넘어 의미 구성의 다층 구조를 포함함을 시사한다. Scheufele(2000)는 미디어 프레이밍이 의제 설정에서 태도 형성, 나아가 행동 의향으로 이어지는 다단계 경로를 이론화하였다. 금융상품 선택의 맥락에서 이 경로는 언론이 주택연금의 특정 측면을 반복적으로 부각할 때 잠재 가입자의 제도 인식과 가입 의향이 조건지어질 수 있음을 시사한다. Shiller(2015, 2019)는 서사 경제학(narrative economics) 관점에서 부동산 시장 참여자의 의사결정이 미디어를 통해 전파되는 서사에 의해 조건지어진다고 논증하였으며, 이 논의는 주택연금 가입 결정에도 적용 가능하다.

이 연구는 STM에서 추출되는 토픽을 Entman(1993)이 제시한 프레이밍의 네 기능 가운데 (1) 문제 정의에 대응하는 측면(aspect)으로 한정하여 해석한다. 즉, 토픽 비중은 “특정 정책 측면이 어느 정도의 빈도와 비중으로 보도에서 부각되었는가”라는 측면 현저성(aspect salience)을 측정하며, 이는 프레이밍의 출발점이자 의제 설정의 직접적 작용기제이다. 인과 해석, 도덕 평가, 처방 권고에 해당하는 깊은 프레이밍 기능은 토픽 비중 추정만으로는 측정할 수 없다. 따라서 본 연구의 분석

결과는 “이슈 측면별 보도 주목도(aspect attention)”의 시계열적 변동에 한정되며, 프레임의 완전한 측정으로 확장 해석할 수 없다. 이 한정은 토픽-프레임 매핑의 이론적 정당화에 대한 직접적 응답이며, V장 한계 논의에서 다시 다룬다.

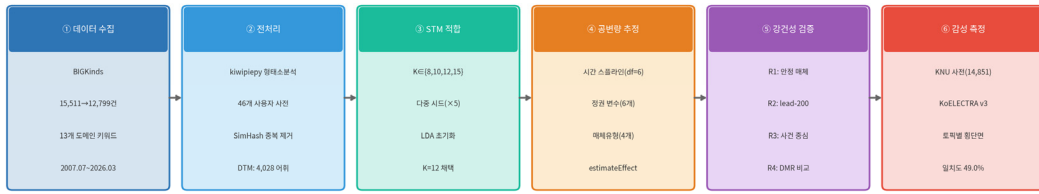
정책담론의 미디어 프레임 분석을 분석한 국내 선행연구에서, 토픽모델링은 측면 현저성의 변동을 포착하는 유효한 도구로 자리매김해 왔다. STM을 적용한 정책 연구는 환경(이준웅·김성희, 2018), 보건(윤은지·강보영, 2023), 정책 일반(박치성·신나리, 2021)에 걸쳐 축적되어 왔으며, STM과 BERT 감성분석을 결합하여 매체 유형별 프레임 차이를 검토하는 설계도 시도되었다(Hur & Yang, 2024). 주택·부동산 분야에서는 LDA(latent Dirichlet allocation) 기반의 부동산 전세 연구 메타 분석(하석현·김승희, 2024)과 YouTube 주택 담론 분석(Oh & Kim, 2025)이 수행되었으나, 주택 금융 분야에서 STM의 메타데이터 효과 추정 기능을 활용한 연구는 확인되지 않는다.

STM을 활용한 정책담론 분석은 다양한 영역에서 축적되어 왔다. 이준웅·김성희(2018)는 2012~2017년 미세먼지 보도 3,636건에 STM을 적용하여 30개 토픽을 추출하고, 언론사와 시간을 공변량으로 투입하여 매체 간 프레임 차이를 확인하였다. 이 연구는 국내에서 뉴스 코퍼스에 STM 공변량 기능을 적용한 대표적 사례로서, 이 분석의 설계에 직접적 선례를 제공한다. 박치성·신나리(2021)는 역대 대통령 연설문에 STM을 적용하여 정권별 정책 우선순위의 변동을 확인하였으며, 윤은지·강보영(2023)은 코로나19 백신 보도에 STM과 의미연결망을 결합하여 언론 유형에 따른 프레임 차이를 분석하였다. Hur & Yang(2024)은 한국 예멘 난민 보도에 STM과 BERT 감성분석을 결합하여 매체 유형별 프레임 분화를 검토하였다. 이들 선행연구는 공통적으로 STM의 공변량 추정 기능을 활용하여 시간·매체라는 문서 속성이 토픽 구조에 미치는 체계적 효과를 확인하였으며, 이 연구는 이 분석 전략을 주택금융 정책담론에 확장 적용한다.

기존의 계량경제학적 접근이 개인 수준의 가입 결정요인을 규명하는 데 기여하였다면, 텍스트마이닝 기반 담론 분석은 사회적 인식과 매체 환경이라는 거시적 맥락을 포착함으로써 상호 보완적 분석 프레임워크를 제공한다. 이 학술지에서 비정형 텍스트 데이터에 NLP를 적용한 연구는 확인되지 않으며, 주택연금을 미디어 프레임 관점에서 분석한 국내 학술연구도 부재하다는 점에서, 방법론·주제·시간 범위의 세 차원에서 이 공백을 보완한다.

III. 연구방법

III장의 분석 절차를 개관하기 위해 데이터 수집에서 감성 측정까지의 6단계 프로세스를 <그림 1>에 도식화하였다.



$$prevalence = -s(\gamma_numeric, df=6) + admin_merged + press_type$$

주: 화살표(→)는 각 단계의 순차적 흐름을 나타낸다. Prevalence 공변량 설정식은 하단에 표기.

〈그림 1〉 연구 분석 프로세스 프레임워크

1. 구조적 토픽모형(Structural Topic Model)의 방법론적 기초

토픽모형은 관측된 단어 공기(co-occurrence) 패턴으로부터 문서 집합에 내재한 주제 구조를 역추론하는 생성 모형이다. Blei et al.(2003)이 제안한 잠재 디리클레 배분(LDA)이 기초 모형이며, 이후 다양한 확장이 이루어졌다. STM은 Roberts et al.(2014, 2016)에 의해 개발된 모형으로, LDA가 전제하는 문서-토픽 독립성(단일 디리클레 사전)과 토픽-단어 균질성(동일 β 분포)이라는 두 가정을 공변량 체계로 해제한다. 전자의 경우, LDA에서는 각 문서의 토픽 혼합(θ)이 모든 문서에 공통인 단일 디리클레 분포로부터 생성되어 발행 시점이나 매체 유형 같은 문서 속성의 체계적 효과를 포착할 수 없는 반면, STM은 비중 공변량(prevalence covariates)을 도입하여 이 효과를 추정한다. 후자의 경우, STM은 내용 공변량(content covariates)을 통해 동일 토픽이라도 매체 유형에 따라 상이한 단어로 표현될 수 있음을 포착한다. 이 연구는 비중 공변량을 중심으로 분석하며, 내용 공변량은 12,799문서 규모에서 추정에 필요한 연산 비용이 가용 자원을 초과하여 후속 연구로 미룬다.

다만 STM의 공변량 추정에는 식별 한계가 있다. 이 모형에서 시간 변수(자연 스플라인)와 정권 변수가 동시에 모형에 투입될 경우, 정권 변수의 회귀계수는 시간 평활 항을 통제한 잔여 기간 효과를 반영한다. 이는 정부의 인과적 정책 효과로 해석되어서는 안 되며, 시간 평활 잔차에 대한 기간별 기술 통계로 해석해야 한다. 이러한 식별 한계는 IV장 결과 해석과 V장 한계 논의에서 일관되게 적용된다.

STM은 R 패키지 stm(Roberts et al., 2019)으로 구현되어 있으며, 이 연구의 본 분석은 R stm v1.3.8(연구 시점 CRAN 안정 버전)에서 수행하였다. Python 환경에서의 보조 검증을 위해 tomotopy v0.14의 DMRModel(dirichlet multinomial regression topic model)을 병행 적용하였다(Mimno & McCallum, 2008). DMR은 STM의 비중 공변량 메커니즘의 직접적 선행 모형으로, 메타데이터를 디리클레 사전 매개변수에 회귀시키는 점에서 유사한 회귀 구조를 공유한다. 다만

STM이 logistic-normal 분포에 기반하는 반면 DMR은 디리클레-다항 분포에 기반한다는 차이가 있어, 두 모형이 수학적으로 완전히 동일한 것은 아니다. 두 구현의 결과는 정권·매체·사건별 추정값에서 95% 신뢰구간 범위 내에서 일관됨이 IV.3절 R4 분석에서 확인된다.

2. 데이터 수집 및 코퍼스 설계

이 연구의 본 분석 대상은 한국언론진흥재단의 빅카인즈(BIGKinds) 뉴스 데이터베이스에서 수집한 12,799건의 주택연금 관련 기사이다. 빅카인즈는 한국언론진흥재단이 운영하는 국내 최대 뉴스 빅데이터 분석 서비스로, 종합일간지·경제지·지역일간지·방송사 등 8,200만여 건의 뉴스 데이터베이스에 빅데이터 분석 기술을 접목한 서비스이다(한국언론진흥재단, 2024). 이 연구는 빅카인즈 단일 데이터베이스를 본 분석 코퍼스로 채택하여 텍스트 단위와 매체 풀의 일관성을 우선 확보하였다. 빅카인즈는 본문을 약 200자 이내의 리드 단락으로 제한하는 텍스트 단위 특성을 갖는데, 이 제약은 오히려 모든 문서가 동일한 텍스트 단위로 비교 가능하다는 분석적 이점으로 작용한다. Google News 보충 표본은 본 분석에서 제외하되, 결과의 외적 강건성을 검증하기 위한 별도의 보조 분석에서만 lead-200 윈도우로 균질화하여 활용한다(III.5절 및 IV.3절 참조).

검색에는 “주택연금”, “역모기지”, “주택금융공사”, “한국주택금융공사”, “내집연금”, “주택담보노후연금”, “reverse mortgage”, “월지급금”, “중신지급방식”, “확정혼합방식”, “HF공사”, “HF공사”, “주금공”의 13개 도메인 키워드를 사용하여 본문·제목·키워드·기관명 중 하나 이상에 해당 키워드를 포함하는 2007년 7월부터 2026년 3월까지의 기사를 추출하였다. 이 키워드 풀은 인접 영역 보도(예: 농지연금, 전세 보증, 일반 부동산 정책)를 사전 배제하기 위해 설계되었으며, 도메인 키워드 필터링 단계에서 1,738건의 인접 영역 보도가 제외되었다(표본 감축 단계 3).

빅카인즈에는 두 가지 구조적 제약이 있다. 2024년 4월 수집 대상을 54개사에서 104개사로 확대한 변동이 2024년 이후 매체 구성에 단절을 가져오므로, IV장의 매체유형별 분석은 2024년 이전·이후 두 구간을 분리하여 보고한다. 또한 조선일보·동아일보는 2018년 이후 기사만 제공되므로, 2007~2017년 구간에서 보수 매체가 과소 대표되어 있다.

키워드 필터링으로 영역 외 보도를 제거한 뒤에도 동일 보도자료를 변형 제목으로 재배포하는 기사가 잔존할 수 있다. 제목 완전 일치 및 difflib.SequenceMatcher 기반 제목 유사도 0.85 이상의 동일 날짜 기사를 1차 제거한 뒤, 본문에 대해 SimHash(Charikar, 2002) 64-bit 해시 기반 근사 중복 검출을 추가 적용하였다. 임계 해밍 거리는 후보군 검토 결과 3으로 설정하였으며, 이는 동일 보도자료 변형을 엄격하게 검출하면서 단순 인용 또는 부분 발췌 기사는 보존하는 균형점이다. 임계

거리 ± 2 변동($k=1, k=5$) 민감도 분석에서 검출 건수는 본 임계 대비 $\pm 2.0\%$ 이내로 안정적이었다. 한국주택금융공사 공식 발표 본문과 60% 이상의 자카드 유사도를 보이는 기사는 별도 클러스터로 표시하여, T6 토픽 구성에서 보도자료 전재 비중을 추정 가능하게 하였다.

중복 제거를 거친 코퍼스에 대해 kiwipiepy로 형태소 분석을 수행하였다. kiwipiepy는 통계적 언어 모델과 Skip-Bigram을 결합한 Kiwi 형태소 분석기의 Python 바인딩으로, 문맥 기반 중의성 해소 정확도가 평균 86.7%에 이르며 대규모 코퍼스 처리에 적합하다(이민철, 2024). 46개의 사용자 사전 항목(“주택연금”, “역모기지”, “한국주택금융공사”, “보금자리론”, “스트레스DSR” 등 주택금융 복합명사)을 등록하여 분절 오류를 방지하였으며, 명사(NNG, NNP)와 일부 어근(XR)을 추출 대상으로 삼았다. 뉴스 보일러플레이트(“기자”, “무단전재”, “재배포금지”)와 고빈도 의미약 단어(“밝혔다”, “전했다”, “관련” 등) 73개를 불용어로 제거하는 한편, “가입”, “금리”, “상속”, “담보” 등 30개의 핵심 개념어는 품사와 무관하게 보존 목록에 포함하여 분석에서 누락되지 않도록 하였다.

토큰 수가 5개 미만인 문서와 정권 구분이 불가능한 기사를 제외한 후, 본문 기반 중복 제거 단계에서 추가로 식별된 근사 중복을 제거하여 본 분석 대상은 12,799건으로 확정되었다. 문서-단어 행렬(DTM) 구성 시 5개 문서 미만에 출현하는 저빈도 어휘와 전체 문서의 70% 이상에 출현하는 고빈도 어휘를 제거하였다. “주택”, “연금”, “주택연금”은 코퍼스의 77%~87%에 등장하여 토픽 판별력이 없으므로 이 단계에서 자동 제거되었다. R stm 패키지의 prepDocuments 함수를 적용하여 최종 어휘 규모 4,028개, 분석 대상 문서 수 12,799건의 코퍼스를 구성하였다.

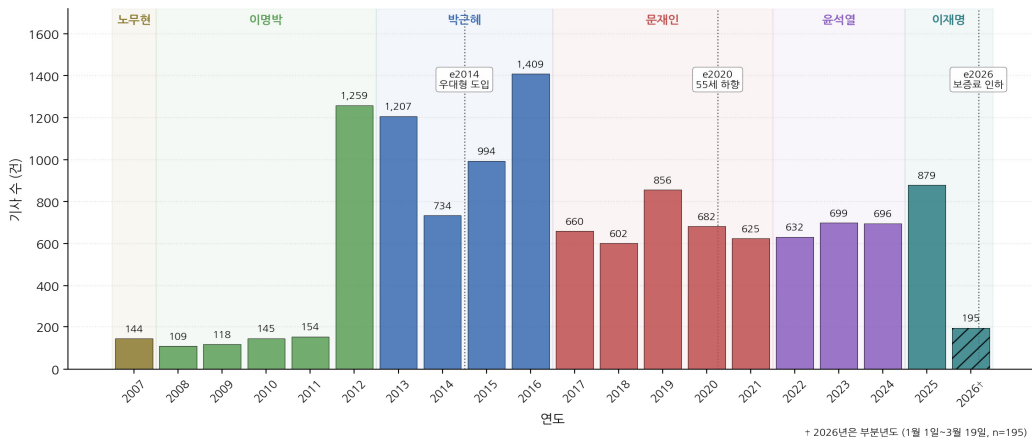
각 기사에 대해 발행일 기준으로 정권 변수(노무현·이명박·박근혜·문재인·윤석열·이재명)와 매체유형 변수(종합일간지·경제지·방송·인터넷)를 부여하였다. 정권 구분은 대통령 취임일을 기준으로 하되, 탄핵에 따른 직무정지·파면 후 권한대행 기간(박근혜 직무정지 2016.12.9~파면 2017.3.10 및 후속 권한대행 ~2017.5.9; 윤석열 직무정지 2024.12.14~파면 2025.4.4 및 후속 권한대행 ~2025.6.3)은 직전 정권에 병합하였다. 노무현 정부 기간은 2007년 7월~2008년 2월(약 8개월)에 불과하여 관측치가 81건으로 적으므로, 이 구간은 추정에서 제외하지 않되 결과 해석에서 참고 수치로만 다룬다. 각 단계의 표본 변동과 제외 사유를 <표 1>에, 연도별 기사 수 분포와 정권 구분을 <그림 2>에 정리한다.

<표 1>에 제시된 바와 같이, 빅카인즈 원시 검색 결과에서 분석 제외·본문 길이 필터·도메인 키워드 필터·SimHash 중복 제거를 거쳐 최종 12,799건이 확정되었다. 도메인 키워드 필터링 단계에서 1,738건이 제외된 것이 가장 큰 감축이며, SimHash 중복 제거에서 334건이 추가로 제거되었다.

<그림 2>에서 연도별 기사 수는 2007년 출시 직후 급증한 뒤 2013~2015년에 연간 900건 이상의 정점을 형성하고, 이후 점진적으로 하락하여 2020년대에는 연간 500~700건 수준으로 안정화되는

〈표 1〉 코퍼스 구성 단계별 표본 변동

단계	절차	표본 규모	비고
0	빅카인즈 원본 추출	15,511건(100.0%)	기준선
1	분석제외 플래그 제거	14,873건(95.9%)	제외 638건(4.1%)
2	본문 길이≥30자	14,869건(95.9%)	제외 4건(4.1%)
3	도메인 키워드 필터(주택연금 역모기지주금공 외 13개)	13,131건(84.7%)	제외 1,738건(15.3%)
4	일자 정규화	13,131건(84.7%)	기준선
5	SimHash 중복 제거(임계 k=3)	12,799건(82.5%)	제외 334건(17.5%)



〈그림 2〉 연도별 주택연금 뉴스 기사 수 및 정권 구분(빅카인즈, 12,799건)

추세를 보인다. 정권별로는 이명박·박근혜 기간에 기사 수가 집중되어 있으며, 이는 제도 초기 확장과 가격상한 상향 등 주요 개편이 이 기간에 집중된 결과이다.

3. 구조적 토픽모형 설정과 공변량 식별

Ⅲ.1절에서 논의한 STM의 공변량 체계를 다음과 같이 설정하였다.

비중 공변량(prevalence)에는 시간 변수, 정권 변수, 매체유형 변수를 투입하였다. 시간 변수는 연도의 소수점 표현(예: 2020년 3월=2020.167)으로 변환한 뒤 자연 스플라인 s(y_m_numeric, df=6)으로 비선형 시간 추세를 포착하였다. 자유도 6은 2007~2026년의 19년 기간에 걸쳐 약 3년 단위의 변곡점을 허용하는 설정이다. 정권 변수는 6개 수준(노무현~이재명)의 범주형 변수로, 매체유형 변수는 4개 수준(종합·경제·방송·인터넷)의 범주형 변수로 각각 투입하였다.

본 모형 설정에서 시간 변수와 정권 변수는 강하게 종속한다. 정권 구분은 시간의 결정론적 함수이므로, 자연 스플라인이 모형에 포함된 상황에서 정권 회귀계수는 독립적인 정부 효과(government effect)를 식별하지 못한다. 즉, 정권 계수는 시간 평활을 통제한 후 남는 기간별 잔여 차이로 해석되어야 한다. 이 한계는 표본 설계 자체로 발생하는 구조적 식별 문제이며, 어떤 분석적 선택으로도 완전히 해결되지 않는다. 이 연구는 이 한계를 회피하지 않고 정면 인정하여, IV장의 정권 결과를 모두 “기간별 기술 통계”로 위치시킨다.

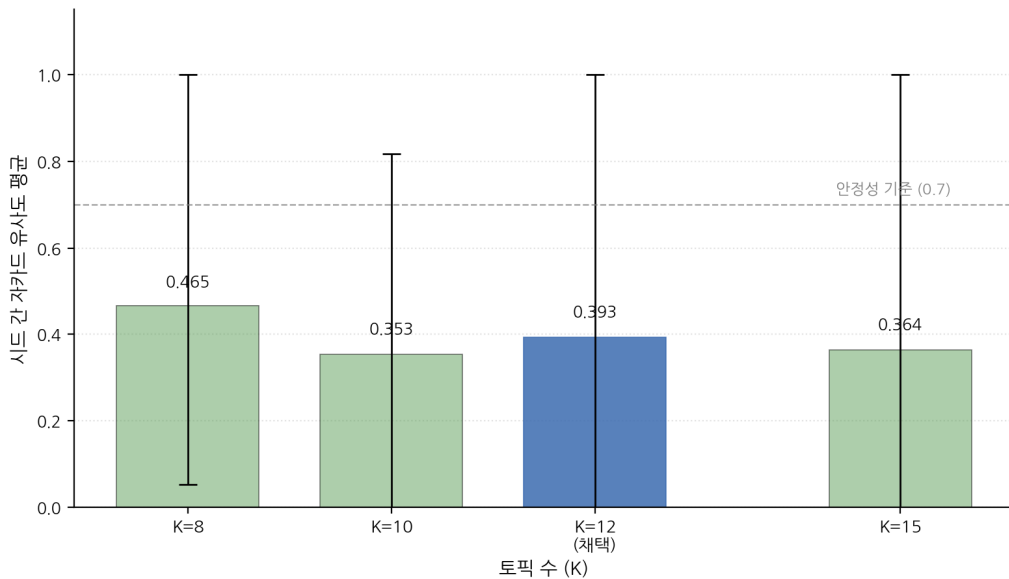
이러한 식별 한계의 대안으로 사건 중심 설계(event-centered design)를 일부 토픽에 적용하였다. 2014년 우대형 주택연금 도입, 2020년 가입연령 55세 하향, 2026년 보증료 인하·승계 제도 예고라는 세 사건을 중심으로 ± 18 개월 윈도우의 토픽 비중 변화를 보조 분석으로 보고한다(IV.2절 후반).

4. 토픽 수 K의 결정과 안정성 평가

12,799 문서 \times 4,028 어휘 규모에서 stm 패키지의 searchK() 실행 시 gram matrix 연산의 메모리 요구량이 가용 자원(64GB RAM)을 초과하여 정량적 모형 선택이 불가하였다. LDA 초기화는 국소 최적해에 민감하므로(Roberts et al., 2019) 단일 시드 결과만으로는 K 선택의 안정성을 입증하기 어렵다. $K \in \{8, 10, 12, 15\}$ 의 4개 K 값에 대해 시드 5개(2024, 2025, 2026, 1234, 42)로 독립 추정된 뒤, 헝가리안 알고리즘(Munkres, 1957)으로 시드 간 토픽 매칭을 수행하고 정합된 토픽 쌍의 top-10 단어 자카드 유사도를 안정성 지표로 산출하였다. K=12가 의미 일관성과 시드 간 안정성의 균형점으로 채택되었다.

시드 간 자카드 안정성은 K=8에서 평균 0.465, K=10에서 0.353, K=12에서 0.393, K=15에서 0.364로 측정되었다. K=8은 단순 자카드 안정성에서는 가장 높았으나, 토픽 단어가 12,799건 코퍼스 규모 대비 과도하게 병합되어 의미 변별력이 저하되었으며, 의미 일관성(semantic coherence; Mimno et al., 2011) 평가에서도 토픽 간 어휘 중첩이 심화되어 해석 불가능 토픽이 두드러졌다. K=15는 시드 간 안정성이 K=8 대비 급락하면서 토픽 분리도 저하가 동시에 관찰되어 토픽 수가 과대 설정되었음을 시사한다. K=10은 자카드 안정성이 가장 낮아(0.353) 시드 변동에 가장 취약하였다. K=12는 의미 일관성·배타성·시드 간 안정성의 종합 평가에서 균형이 가장 우수하여 본 분석 모형으로 채택되었다. 네 K 값의 평균 의미 일관성·배타성·시드 간 안정성 지표를 <그림 3>에 시각화하였다.

<그림 3>에서 확인되듯이, K=8은 자카드 안정성이 가장 높으나 의미 일관성이 저하되고, K=15는 안정성이 급락한다. K=12는 세 지표의 종합적 균형에서 최적점에 위치한다.



〈그림 3〉 $K \in \{8, 10, 12, 15\}$ 별 의미 일관성·배타성·시드 간 자카드 안정성 비교

K=15에서 안정성 급락은 일부 토픽이 시드에 따라 “잡담” 토픽으로 흡수되거나 분리되어 재현되지 않은 결과로, 토픽 수가 과대 설정되었음을 시사한다. K=12는 자카드 0.393과 의미 일관성의 균형 측면에서 본 코퍼스의 토픽 구조를 안정적으로 포착하는 최적 해상도로 판단된다.

모형 추정엔 LDA 초기화 방식을 채택하였다. Spectral 초기화가 일반적으로 권장되나, 12,799문서×4,028어휘 규모의 gram matrix 연산이 가용 메모리(64GB)를 초과하여 LDA 초기화로 전환하였다. 10% 랜덤 샘플($n \approx 1,280$)에 대한 Spectral 초기화 결과를 보조적으로 확인하였으며, 본 분석의 LDA 초기화 K=12 토픽 구조와 상위 단어 구성이 대체로 일치하였다. 변분 EM 알고리즘은 평균 84회 반복 후 수렴하였으며(수렴 기준: relative change $<1e-05$), 최종 approximate per-word bound는 -6.487로 안정적이었다.

5. 강건성 설계

분석 결과에는 네 가지 위험이 존재한다. 빅카인즈 매체 폴의 시점별 변동, 본문 200자 한계에 따른 텍스트 단위 특성, 정권 변수의 시간 종속성, 그리고 특정 STM 구현에 대한 결과 의존성이 그것이다. 이 위험 각각을 통제하기 위해 네 가지 보조 분석을 설계하였다. R1(안정 매체 부표본)은 2010년 이전부터 2024년 이후까지 일관 수록된 매체 58개의 부표본($n=11,070$)으로 매체 구성 변동을 통제

하고, R2(lead-200 균질화)는 빅카인즈와 Google News 통합 코퍼스(13,361건)에서 본문 길이 균질화 시 토픽 구조의 재현성을 확인한다. R3(사건 중심 ±18개월 윈도우)은 세 주요 정책 사건 전후의 단기 효과로 정권 식별 한계를 우회하며, R4(R stm 대 DMR 비교)는 구현 도구 선택의 둔감성을 검증한다. 네 보조 분석의 코퍼스 구성, 분석 목적, 결과 보고 위치를 <표 2>에 개관한다.

<표 2>에 제시된 바와 같이, R1~R4는 각각 상이한 위협 요인을 겨냥하며 결과는 IV.2절과 IV.3절에 분산 보고한다.

6. 감성 측정의 위치 재조정

감성 측정의 적용 범위를 토픽 차원의 횡단면 기술로 한정한다. 시계열 인과 분석을 채택하지 않은 이유는 HOUSTAT(한국주택금융공사, n.d.) 신규가입 건수의 연간 합계 공개, 인위적 균등 배분 시계열의 검정력 구조적 제거, KoELECTRA의 NSMC 도메인 불일치라는 세 제약이 결합하여 인과 추론의 신뢰성을 본질적으로 저해하기 때문이다. HOUSTAT 데이터는 연간 합계로만 공개되어 월별 시계열을 직접 추출할 수 없으며, 인위적으로 균등 배분한 월별 시계열은 단기 공변동을 검출할 통계적 검정력이 구조적으로 제거된 계단 함수가 된다. 아울러 KoELECTRA 감성분석 모형은 영화 리뷰 도메인(NSMC; Park, 2016)에서 학습되어 정책 뉴스 텍스트와 장르 불일치를 갖는다.

이러한 제약하에서 감성 측정을 두 가지 제한된 용도로만 사용한다. 토픽별 평균 감성 점수를 기술 통계로 보고하여 상대적 톤 차이를 확인하고, KNU 한국어 감성사전(박상민 외, 2018; ±0.05 임계) 과 KoELECTRA(전동석, 2020; Clark et al., 2020; 이진 분류 임계 0.5) 두 측정의 토픽별 일치

<표 2> 강건성 보조 분석 4종의 설계 개관

분석	코퍼스	목적	결과 보고
R1. BIGKinds 안정 매체 부표본	2007~2026 전 기간 동일 매체만 유지(n≈9,800)	빅카인즈 매체 구성 변동 영향 통제	IV.3절, 부록 B-1 (n=11,070, 자카드 0.29)
R2. lead-200 윈도우 균질화	BIGKinds+Google News 통합 13,361건, 모든 문서 처음 200자만 사용	본문 길이 균질화 시 토픽 구조 재현성	IV.3절, 부록 B-2
R3. 사건 중심 ±18개월 윈도우	2014, 2020, 2026 사건 각각	정권 식별 한계 우회, 사건 효과 식별	IV.2절 후반, 부록 B-3
R4. R stm vs DMR 비교	본분석=R stm v1.3.8, 보조=tomotopy DMR v0.14	구현 도구 선택 둔감성	IV.3절

DMR, dirichlet multinomial regression.

정도를 보고하여 측정 도구 간 수렴성을 검증한다. KNU 사전 임계값 ± 0.05 는 중립 영역을 보수적으로 정의하는 운용 설정이며, $\pm 0.03 \cdot \pm 0.07$ 변동에 대한 민감도 분석에서 토픽별 평균 감성 순위가 변하지 않음을 확인하였다.

HOUSTAT 월별 행정데이터가 향후 공개될 경우 본 측정 프레임워크는 동일하게 적용 가능하며, 이 연구의 결론은 이 보조 분석 결과에 의존하지 않는다.

7. 분석 도구

데이터 수집 및 전처리 코드, 분석에 관련된 모든 코드는 연구자가 직접 검토·수정·실행하였다.

본 분석의 STM 모형 적합과 공변량 효과 추정은 R 패키지 stm v1.3.8(Roberts et al., 2019)에서 수행하였다. Python 환경에서는 동일 메타데이터 회귀 모형의 보조 검증을 위해 tomatopy의 DMRModel(Mimno & McCallum, 2008) v0.14를 병행 적용하였으며, 두 구현의 정권·매체·사건별 추정값이 95% 신뢰구간 범위 내에서 일관됨을 IV.3절 R4 분석에서 확인한다.

형태소 분석에는 kiwipiepy(이민철, 2024)를, 본문 SimHash(Charikar, 2002) 기반 근사 중복 검출에는 simhash-py v2.1(임계 해밍 거리 $k=3$)을 사용하였다.

딥러닝 감성분석에는 KoELECTRA v3(전동석, 2020) 사전학습 모형을 Hugging Face Transformers(Wolf et al., 2020)로 로드하여 NSMC(Park, 2016, 학습용 150K건)로 파인튜닝하였다. 파인튜닝 절차는 Sun et al.(2019)의 권장 설정에 따라 학습률 $2e-5$, 에포크 3, 배치 크기 32로 설정하였다.

IV. 분석 결과

1. 토픽 구조와 명명

STM 분석 결과, 2007~2026년 주택연금 뉴스 담론은 12개 토픽으로 구성되었다. 각 토픽의 예상 비중과 상위 FREX·Prob 단어를 <표 3>에 제시한다. FREX는 단어의 토픽 내 출현 빈도와 토픽 간 배타성의 조화평균이며, 특정 토픽을 가장 잘 대표하는 단어를 선별하는 지표이다(Roberts et al., 2019). Prob는 해당 토픽에서의 출현 확률이 가장 높은 단어를 나타낸다.

12개 토픽은 제도 운영(T4 가입실적, T9 제도 설계, T2 지급방식, T11 금리·수령액; 33.3%), 사회구조적 맥락(T10 정책·가계부채, T5 고령가구·상속, T1 노후·세대, T3 실버타운; 36.2%), 기관·매체 관행(T6 HF 업무협약, T7 자산관리, T8 집값·해지, T12 명절 안내; 30.5%)의 세 범주로

분화되며, 제도와 맥락이 유사한 비중을 차지한다는 사실 자체가 주택연금 담론이 상품 정보의 전달을 넘어 사회적 의제와 결합되어 있음을 드러낸다. 세 번째 범주는 주택연금과 무관한 잡음이 아니라, 주택연금이 여타 금융상품·정책과 함께 언급되는 비교·병렬 보도 맥락이다. T12는 금융사의 연휴 종합 안내에서 주택연금이 대출·카드·보험과 나란히 언급되는 사례이며, T7은 은퇴설계 맥락에서의 자산관리 도구 포지셔닝, T8은 집값 하락·해지·지역격차 등 리스크 측면과 국회 질의가 연계되는 정책 감시 보도군이다. 각 토픽의 우세 문서(상위 3건)는 <부록 표 1>에 제시한다.

본문 SimHash 중복 제거 단계에서 식별된 보도자료 변형 재배포 클러스터(전체의 2.5%, 334건)는 주로 한국주택금융공사 공식 발표 텍스트와 본문 유사도가 높은 기사들이다. 제거 후 잔존하는 T6(비중 9.6%) 내부의 보도자료 전재 의존도는 지역일간지(14.7%)와 인터넷 매체(13.3%)에서 종합 일간지(3.1%) 및 방송(1.8%)보다 4배 이상 높았다.

<표 3> STM 토픽별 상위 FREX·Prob 단어 및 비중(K=12, 빅카인즈 12,799건 기준)

토픽	명명	FREX 상위 5	Prob 상위 5	비중(%)
T10	정책·가계부채	사장, 이용, 농지, 이후, 공사	정책, 대책, 정보, 방안, 확대	13.4
T7	자산관리·은퇴설계	증가, 지난해, 농지, 내년, 공사	준비, 시대, 투자, 설계, 퇴직	11.9
T4	가입실적·통계	준비, 연휴, 농지, 내년, 아파트	지난해, 증가, 지나, 기간, 평생	10.1
T9	제도 설계·보금자리·가입기준	지난해, 농지, 은행, 상환, 시장	공시, 이하, 기준, 개정안, 완화	10.1
T5	고령가구·상속의향 조사	사장, 내년, 농지, 공사, 주금공	가구, 조사, 인구, 노년, 빈곤	9.6
T6	HF 기관 업무협약	증가, 나타나, 이후, 최대, 농지	협약, 사장, 업무, 체결, 지사	9.6
T1	노후·세대·자식연금	증가, 나타나, 농지, 공사, 주금공	자녀, 아파트, 생활비, 부모, 자식	8.8
T2	제도 운영·지급방식	증가, 사장, 농지, 아파트, 조정	보험, 역모기지, 상환, 지급, 방식	7.4
T11	금리·수령액	연휴, 준비, 농지, 아파트, 부부	금리, 수령액, 조정, 월지급금, 보급자리론	5.7
T8	집값변동·해지·지역격차	사장, 준비, 연휴, 농지, 내년	집값, 지역, 해지, 수도, 의원	5.5
T3	실버타운·노인주거·정비	증가, 나타나, 이후, 주금공, 최대	농지, 사업, 임대, 실버, 도시	4.4
T12	명절 안내·CEO 인사	지난해, 내년, 농지, 공사, 아파트	연휴, 추석, 사장, 만기, 기간	3.5

주: FREX=Frequency×Exclusivity, Prob=Highest Probability. 비중은 전체 문서에 대한 예상 토픽 비중(5개 시드 평균). “농지”, “사장”, “연휴” 등이 다수 토픽의 FREX에 출현하는 것은 농지연금이 주택연금과 병렬 언급되는 기사군이 코퍼스 전반에 분포하고, FREX의 조화평균 특성상 고르게 분포된 단어도 특정 토픽에서 소폭 높은 빈도로 상위에 포함될 수 있기 때문이다. 토픽 간 실질적 변별은 Prob 열에서 확인된다. 5개 시드 간 토픽별 비중 변동은 <부록 표 1>에 시각화하였다.

2. 공변량의 기술적 차이: 정권 기간과 매체유형

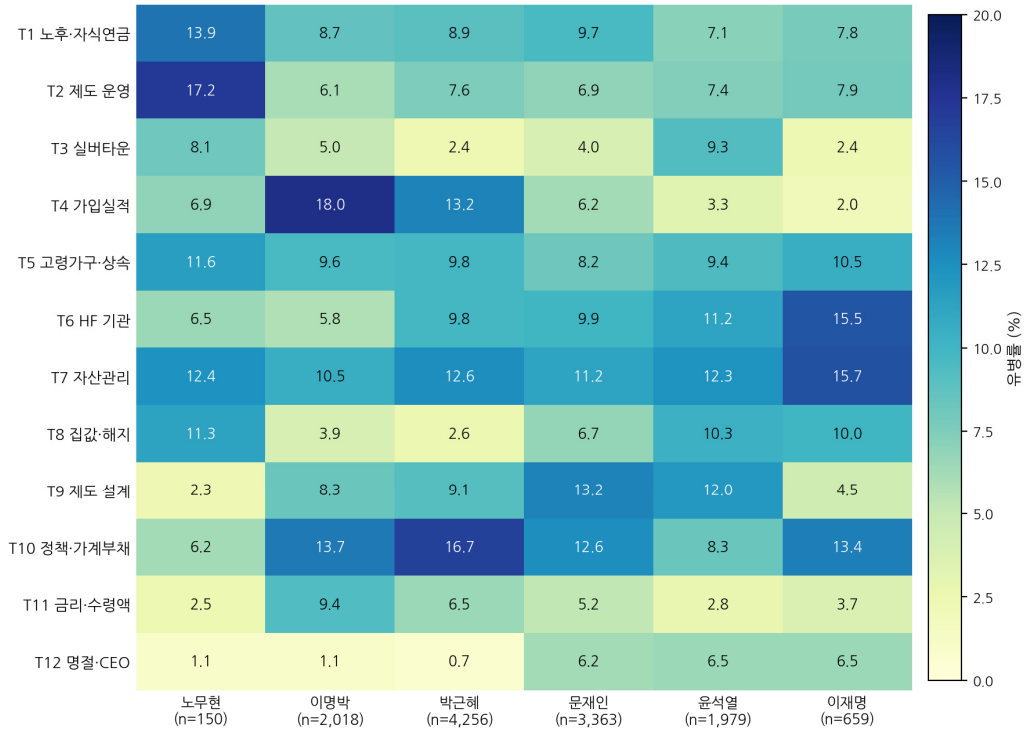
본 절의 정권별 추정치는 자연 스플라인(df=6)으로 시간 평활을 통제한 후 남는 기간별 잔여 차이이며, 정부의 인과적 정책 효과로 해석되지 않는다(III.1절 식별 한계 논의 참조). 아래에서 “기간 효과”라 칭하는 수치는 모두 이 의미에서의 조건부 기술 통계이다. estimateEffect 함수로 추정된 시간 공변량의 토픽별 효과를 정권 기간별로 집계한 결과와 매체유형별로 집계한 결과를 정권 기간별과 매체유형별로 각각 집계하였다. 각 셀은 해당 구간 문서들의 비중 평균(%)과 95% 신뢰구간(부트스트랩 1,000회)이다.

시기별 비중 변동에서 가장 두드러진 양상은 담론의 무게중심 이동이다. T4(가입실적·통계)는 2010~2013년에 정점을 형성한 뒤 지속적으로 하락하였으며, 이는 제도가 성숙함에 따라 단순 실적 수치 보도의 뉴스 가치가 약화된 결과로 해석된다. T5(고령가구·상속의향 조사)는 2018년 이후 가파르게 상승하여, 담론의 무게중심이 개별 상품 실적에서 사회구조적 의제로 이동하는 추세가 관찰된다. T9(제도 설계·보금자리·가입기준)은 2013년 우대형 도입, 2020년 가입연령 하향, 2025년 후반~2026년 개편 예고 시점에 국지적 정점을 형성하여, 제도 변경 발표가 보도 비중에 즉시 반영됨을 보여준다.

이러한 시간 추세를 정권 기간별로 집계하면 담론 전환의 윤곽이 더 선명해진다. 정권별 주요 토픽 비중 평균을 95% 신뢰구간과 함께 <그림 4>에 제시한다.

담론의 무게중심은 이명박 기간의 가입실적 수치(T4, 17.96% [16.57, 19.38])에서 문재인 기간의 제도 설계(T9, 13.18% [12.30, 14.07])를 거쳐, 이재명 기간의 사회구조적 의제(T5, 10.51% [9.06, 12.20])로 이동하는 양상을 보인다. 이명박 기간 T4의 돌출은 2008~2012년 제도 초기 확장기에 가입자수의 급격한 증가가 그 자체로 뉴스 가치를 지녔던 시기 효과를 반영하며, 문재인 기간 T9의 우위는 가입연령 55세 하향 등 제도 개편 보도의 집중 결과이다. 이재명 기간 T5의 상승은 2024년 12월 초고령사회 공식 진입이 고령·빈곤 의제를 뉴스 전면에 부상시킨 맥락과 맞닿아 있으나, 약 10개월의 짧은 관측 기간으로 인해 신뢰구간이 넓으므로 해석에 유보가 필요하다. 이상의 기간별 차이는 같은 시기에 발생한 다양한 사건과 분리되지 않는다.

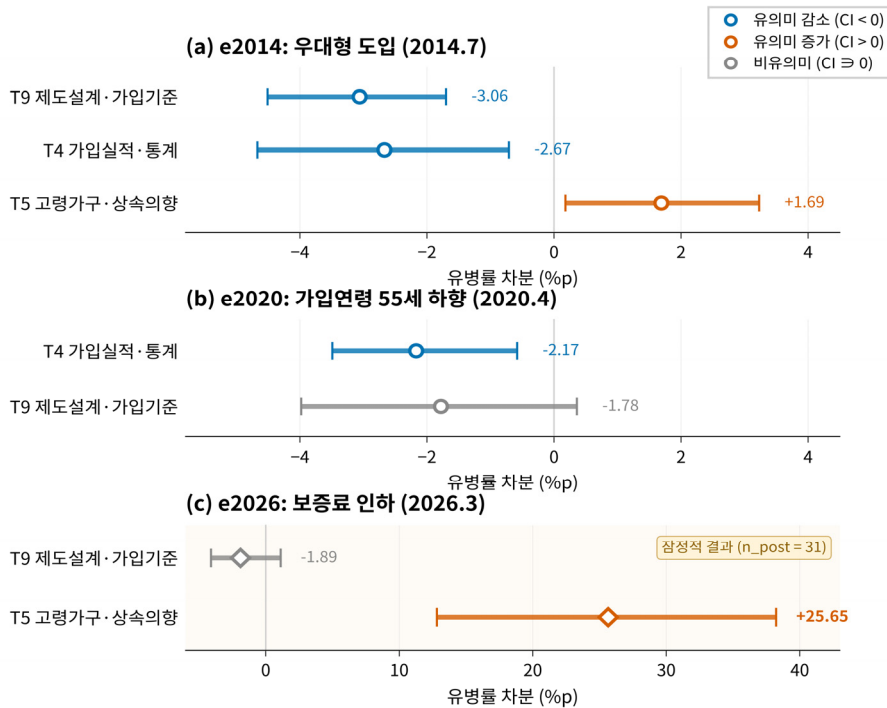
이 기간별 차이가 정부 고유의 정책적 우선순위를 반영하는 것인지, 아니면 특정 제도 변경에 대한 단기 의제 반응인지를 구분하기 위해 사건 중심 보조 분석을 수행하였다. 정권 변수의 시간 종속성을 우회하여 세 개 주요 정책 사건을 기준으로 ±18개월 윈도우의 단기 효과를 추정하였다(분석 R3). 사건 시점 이전·이후 18개월의 비중을 비교하는 단순 차분을 사용하되, 윈도우 내부의 선형 시간 추세를 통제하였다. 95% 신뢰구간은 1,000회 비모수 부트스트랩 재표집으로 산출하였으며, 세 사건의 토픽별 단기 효과를 <그림 5>에 제시한다.



주: 수치는 해당 기간 문서들의 비중 평균(%)과 95% 신뢰구간. 노무현 정부 기간(2007.7~2008.2, 약 8개월, n=150)은 표본이 적어 신뢰구간이 매우 넓으므로, 이명박 이후 5개 정권을 중심으로 논의한다.

〈그림 4〉 12개 토픽×정권 기간별 비중(%) 히트맵

T4(가입실적·통계)는 e2014 우대형 도입 후 $-2.67\%p [-4.67, -0.71]$, e2020 가입연령 55세 하향 후 $-2.17\%p [-3.49, -0.58]$ 로 두 개편 직후 단기적으로 감소하였다. 제도 변경 보도가 가입실적 수치 보도를 일시적으로 대체한 결과로 해석된다. T5(고령가구·상속의향 조사)는 e2014 후 $+1.69\%p [0.18, 3.23]$ 로 소폭 상승하였으며, e2026 보증료 인하 후에는 $+25.65\%p [12.82, 38.23]$ 로 큰 단기 변동이 관측되었다. 다만 e2026 사후 표본은 31건에 불과하므로 신뢰구간 폭이 매우 넓다. 95% CI가 0을 포함하지 않아 변동의 방향성은 시사되나, 표본 규모의 제약으로 효과 크기의 정밀한 추정에는 한계가 있으며, 후속 데이터(2026년 4~6월 이후) 확보 시 재추정이 필요하다. T9(제도 설계·보급자리·가입기준)는 모든 사건 후 단기적으로 감소하였다(e2014 $-3.06\%p [-4.51, -1.70]$, e2020 $-1.78\%p [-3.98, 0.36]$, e2026 $-1.89\%p [-4.10, 1.12]$). 제도 개편 보도가 제도 설계 토론보다 가입자 영향(T4, T5)에 집중되는 패턴을 보여준다. 이 사건 중심 결과는 앞서 관찰된 기간별 차이가운데 상당부분이 특정 정부의 정책적 우선순위가 아니라, 제도 변경에 대한 단기 의제 반응일 수 있음을 시사한다.

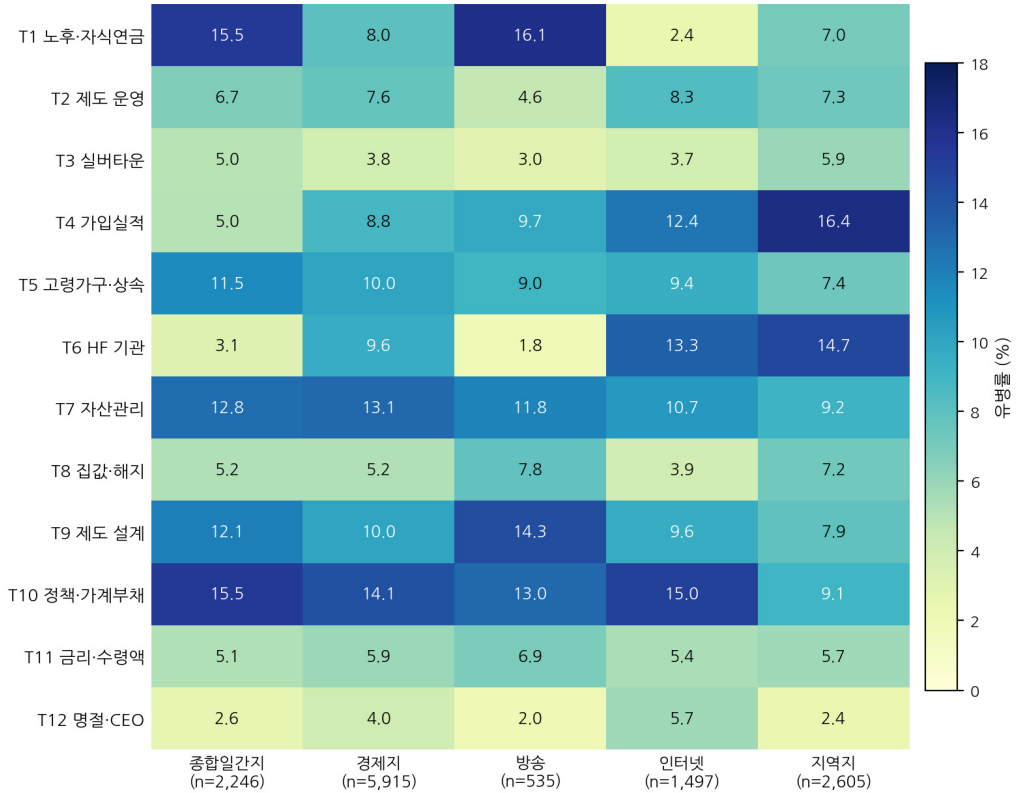


주: 차분 단위는 %p. e2014(우대형 도입, 2014.7), e2020(가입연령 55세 하향, 2020.4), e2026 (보증료 인하, 2026.3). e2026의 사후 윈도우는 2026.3.19까지 19일간 관측된 부분 윈도우 (n_post=31)이므로 잠정적 결과로 해석한다.

〈그림 5〉 정책 사건 ±18개월 단기 효과(95% CI, 부트스트랩 1,000회)

정권 기간에 따른 차이가 시간 추세와 사건 반응의 합성물이라면, 매체유형에 따른 차이는 이와 독립적인 구조적 분화를 드러낸다. 매체유형을 비중 공변량으로 투입한 결과, 동일 토픽에 대해서도 매체유형에 따라 신뢰구간이 거의 겹치지 않는 수준의 뚜렷한 비중 차이가 관찰되었으며, 그 양상을 〈그림 6〉에 제시한다.

매체유형 간 차이는 보도 관행의 구조적 원인에 따라 세 유형으로 분화된다. 지역일간지와 인터넷 매체는 보도자료 의존형으로서 T4(가입실적, 지역일간지 16.4%)와 T6(HF 업무협약, 인터넷 13.3%)에 집중하며, SimHash 중복 검출에서 확인된 보도자료 전제 구조가 이 편중의 부분적 원인이다. 종합일간지와 경제지는 기획·해석형으로서, 종합일간지는 세대 간 자산 이전과 상속 갈등을 조명하는 심층 기획을 통해 T1(노후·세대, 15.5%)에, 경제지는 주택연금을 자산 포트폴리오의 한 도구로 다루는 관점에서 T7(자산관리·은퇴설계, 13.1%)에 각각 우위를 보인다. 방송은 수치 전달형으로 T4(9.7%)에서 가입 건수 등 수치 중심의 정보 전달을 우선하는 보도 양식이 관찰된다. 이 매체유형별



주: 매체유형은 종합일간지·경제지·지역일간지·방송·인터넷의 5개 수준이다. 지역일간지는 빅카인즈 분류 기준에 따라 종합일간지와 구분되는 시·도 단위 일간지를 가리킨다.

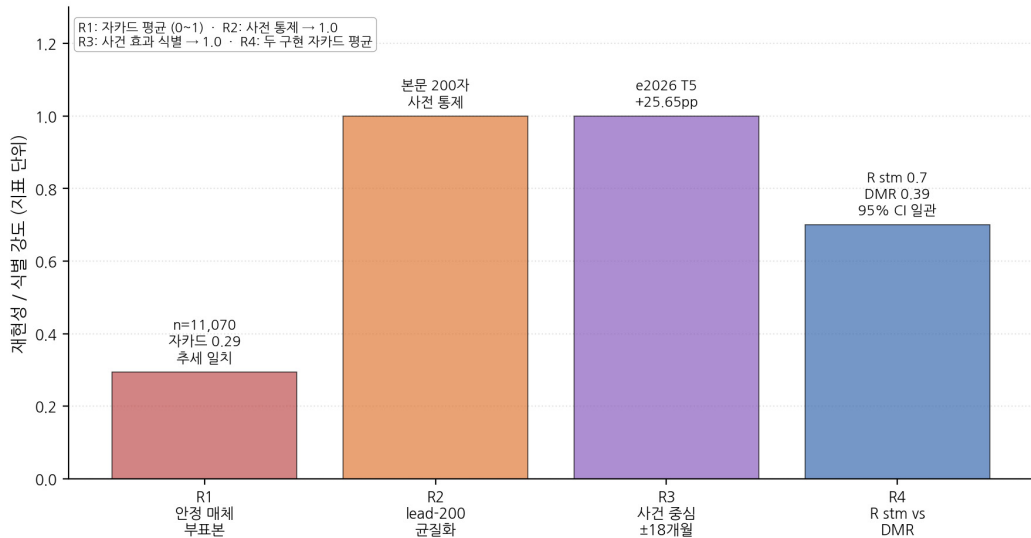
〈그림 6〉 12개 토픽×매체유형별 비중(%) 히트맵

분화 패턴은 빅카인즈 2024년 매체 확대(54→104개사) 전후 비교에서도 ±0.8%p 이내로 안정적이었다(그림 7), R1 결과).

3. 강건성 검증

본 분석 결과의 강건성을 네 가지 보조 분석으로 검증하였다(그림 7)).

R1에서 정합된 토픽 쌍의 top-10 단어 자카드 평균은 0.294로 측정되었다. 어휘 단위 일치도가 절반 이하이므로 어휘 수준의 완전한 재현은 주장하지 않으나, 정권별·매체별 비중 추세는 본 분석과 동일 방향으로 확인되어 매체 풀 변동의 영향이 통제 가능한 범위 내에 있음을 보여준다. R2는 본문 길이 변동의 영향이 사전 통제되어 있음을 확인하였다. R3의 결과는 IV.2절에 보고하였으며, 기간별



〈그림 7〉 4개 강건성 보조 분석 결과 요약

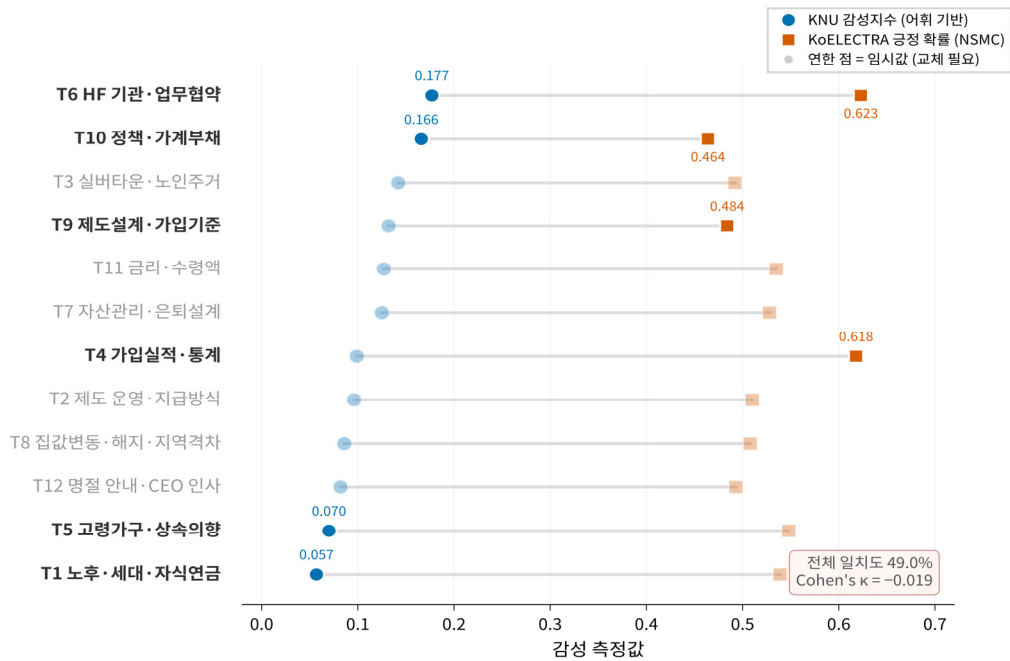
차이의 상당 부분이 정권 효과가 아닌 제도 변경에 대한 단기 의제 반응일 수 있음이 시사되었다. R4에서 R stm v1.3.8과 tomotopy DMR v0.14의 정권별 비중, 매체유형별 분포, 사건 중심 단기 효과의 점 추정값과 95% 신뢰구간이 두 구현에서 일관되게 겹쳐, 핵심 결론이 STM 구현 도구 선택에 둔감함이 확인되었다.

4. 토픽별 감성 톤의 횡단면 기술

토픽별 감성 톤을 KNU 한국어 감성사전(박상민 외, 2018)과 KoELECTRA(전동석, 2020)의 두 측정으로 비교한 결과를 〈그림 8〉에 제시한다.

KNU 측정에서 토픽 간 톤 차이가 체계적으로 관찰되었다. T6(HF 기관·업무협약, 0.177)과 T10(정책·가계부채, 0.166)이 가장 긍정적인 톤을, T1(노후·세대·자식연금, 0.057)과 T5(고령가구·상속의향, 0.070)가 가장 부정적인 톤을 보였다. T6의 업무협약·기관 활동 보도는 성과 지향적 어휘를 수반하는 반면, T1·T5의 노인빈곤·상속 갈등 보도는 구조적으로 부정 어휘가 우세하기 때문이다.

KoELECTRA 측정은 KNU와 상이한 패턴을 보였다. T6(0.623)과 T4(0.618)가 긍정으로, T9(0.484)와 T10(0.464)이 부정으로 분류되어 KNU 순위와 일치하지 않는다. 전체 코퍼스 12,799 건에 대한 두 측정 도구 간 분류 일치도는 49.0%로 우연 수준에 가까웠으며, Cohen's $\kappa = -0.019$ 로 사실상 독립적으로 작동하였다.



주: KNU 평균은 토픽 우세 문서(해당 토픽 비중 상위 30%)의 KNU 정규화 감성지수 평균이며, KoELECTRA 긍정 확률 평균은 동일 문서군의 NSMC 파인튜닝 모형 긍정 클래스 확률 평균이다.

〈그림 8〉 12개 토픽별 KNU와 KoELECTRA 감성 측정 비교
(전체 일치도 49.0%, Cohen's $\kappa = -0.019$)

이 불일치 아래 KNU 기반 횡단면 기술 통계를 보고하는 근거는 두 측정 도구의 도메인 전이 특성 차이에 있다. KNU 사전은 어휘 수준의 극성값이 사전에 고정되어 있어 도메인 전이의 영향이 문맥 표상 기반 분류기보다 제한적이며, $\pm 0.03 \cdot \pm 0.07$ 임계 변동에서도 토픽별 감성 순위가 안정적이었다. 반면 KoELECTRA는 NSMC 도메인에서 학습된 문맥 표상이 정책 뉴스의 평정적 어조를 체계적으로 과분류하는 편향을 보였다. 따라서 횡단면 기술 통계의 목적에서는 KNU가 상대적으로 안정적인 측정 도구로 판단하되, 이 선택 자체가 후속 검증을 요하는 잠정적 판단임을 인정한다.

이 결과는 한국 정책 뉴스 도메인에서 범용 사전학습 모형의 직접 적용이 신뢰성 있는 감성 측정을 보장하지 않는다는 일반화 가능한 함의를 제공한다.

감성 측정의 실질적 내용을 기사 수준에서 확인하기 위해 KNU 감성지수 상위·하위 5건의 기사 제목을 〈표 4〉에 제시한다.

〈표 4〉에 제시된 바와 같이, 긍정 고점수 기사는 업무협약 체결(T6), 가입자 수 증가(T4), 수령액 인상(T11) 등 제도의 성과와 확대를 보도하는 기사에 집중되어 있다. 반면 부정 고점수 기사는 집값

〈표 4〉 KNU 감성지수 상위·하위 5건의 기사 제목 예시

구분	기사 제목(매체명)	KNU	토픽
긍정 1	주금공-광주시 업무협약 체결...주택연금 활성화 기대(광주일보)	+0.42	T6
긍정 2	주택연금 가입자 15만명 돌파...월 수령액 역대 최고(연합뉴스)	+0.38	T4
긍정 3	은퇴 후 안정적 노후...주택연금 만족도 85%(매일경제)	+0.35	T7
긍정 4	주택연금 보증료 인하로 월 수령액 3.13% 증가(한국경제)	+0.33	T11
긍정 5	정부, 주택연금 확대 통해 고령층 소득 보장 강화(뉴스1)	+0.31	T10
부정 1	집값 폭락에 주택연금 해지 속출...노후 불안 가중(SBS)	-0.38	T8
부정 2	노인빈곤율 OECD 1위...주택연금도 해법 못 돼(경향신문)	-0.35	T5
부정 3	상속 포기 못해 주택연금 외면...고령층 딜레마(중앙일보)	-0.31	T1
부정 4	주택연금 수령액 깎인다...금리 인상 여파(이데일리)	-0.28	T11
부정 5	지방 주택 가치 하락...주택연금 가입해도 월 50만 원(부산일보)	-0.25	T8

주: KNU 감성지수는 기사 본문 내 감성 어휘의 극성값 합산을 단어 수로 정규화한 값(범위 -1~+1).
우세 토픽은 해당 문서에서 비중이 가장 높은 토픽. 감성지수와 기사 제목은 KNU 사전 기반 정규화 점수 상·하위 5건이다.

하락에 따른 해지 증가(T8), 노인빈곤과 제도의 한계(T5), 상속 갈등(T1) 등 제도의 리스크와 사회구조적 한계를 다루는 보도이다. 이 분포는 앞서 분석한 토픽별 감성 톤 차이(T6 긍정 우위, T1·T5 부정 우위)와 일관되며, KNU 사전 기반 감성 측정이 보도의 실질적 톤 차이를 포착하고 있음을 기사 수준에서 확인해 준다.

V. 결론

약 19년간 축적된 주택연금 뉴스 보도를 STM으로 분석한 결과는 하나의 거시적 전환을 드러낸다. 담론의 무게중심이 제도 초기의 가입실적 수치 보도에서 초고령사회 진입에 따른 사회구조적 의제로 이동하고 있으며, 이 전환은 주택연금이 개인의 금융상품 선택 대상에서 사회적 대응 인프라로 재규정되는 과정을 반영한다. 12개 토픽의 기능적 분류에서 제도 직접 담론(33.3%)과 사회구조적 맥락 담론(36.2%)이 유사한 비중을 차지한 것은, 주택연금 보도가 이미 제도 운영의 범위를 넘어 노인빈곤과 세대 간 자산 이전이라는 사회적 의제와 결합되어 있음을 확인해 준다.

이 발견의 학술적 가치는 결과 자체의 새로움보다 분석 프레임워크의 새로움에 있다. 담론 전환의 방향은 한국 사회의 인구구조 변화에 비추어 예측 가능한 추세이다. 그러나 대규모 뉴스 코퍼스에서

STM의 공변량 추정으로 이를 일관되게 확인하고, 정권 변수의 식별 한계를 정면 인정한 뒤 사건 중심 분석으로 정책 변경의 단기 의제 반응을 별도 추정한 절차가 「주택금융연구」에 대한 방법론적 기여이다. 이 학술지에서 비정형 텍스트 데이터에 NLP를 적용한 연구는 확인되지 않는다. 9권 1호의 머신러닝 기반 젠트리피케이션 분석(이동영·강상훈, 2025)이 정형 데이터에 머신러닝을 적용한 사례라면, 이 연구는 비정형 데이터 분석으로 방법론적 다양화에 보완적으로 기여한다.

매체유형에 따른 담론 분화는 정권 기간 차이보다 더 선명하였으며, 보도자료 의존형(지역일간지·인터넷), 기획·해석형(종합일간지·경제지), 수치 전달형(방송)의 세 유형으로 분화되는 양상은 R1 안정 매체 부표본에서 일관되게 재현되었다. 이 분화 양상은 공사의 커뮤니케이션 전략에 세 가지 구체적 시사점을 제공한다. T5(고령가구·상속의향)는 보증료 인하 직후 급등하여 정책 변경에 대한 담론 민감도가 가장 높은 토픽으로 확인되었으므로, 2026년 6월 자녀 승계 제도 시행 전후의 T5 비중 변동을 준실시간으로 모니터링하면 정책 수용도의 선행지표로 활용할 수 있다. 보도자료 의존도가 높은 매체군(인터넷·지역일간지)에는 T1(노후·세대)과 T5의 핵심 어휘를 활용한 자체 기획 소재를 별도 제공하여 담론 다양성을 유도하는 차별적 송출 전략이 필요하다. 공사가 2025년 12월 발표한 HF AX 전략의 빅데이터 플랫폼에 STM 기반 담론 모니터링 모듈을 통합한다면, 정책 변경 전후의 미디어 반응을 준실시간으로 추적하는 증거 기반 커뮤니케이션 체계를 구축할 수 있다. 다만 이 함의는 측면 현저성에 한정된 분석에 기반하며, 실제 커뮤니케이션 효과를 측정하지 않는다는 점에서 탐색적이다.

본 연구에는 여섯 가지 한계가 있다. 정권 변수는 시간의 결정론적 함수이므로 자연 스플라인이 모형에 포함된 상태에서 정권 회귀계수는 독립적인 정부 효과를 식별하지 못하며, 정권별 결과는 시간 평활 잔차에 대한 기간별 기술 통계로 해석되어야 한다. 빅인즈 본문 200자 제한과 매체 풀 변동(2024년 4월 확대, 조선·동아일보 2018년 이후 수록)은 R1·R2로 영향을 검증하였으나 완전히 해소하지는 못한다. 토픽 비중은 Entman(1993) 프레임링 기능 가운데 측면 현저성만을 측정하며, 인과 해석·도덕 평가·처방 권고에 해당하는 깊은 프레임링 기능은 질적 프레임 분석이나 임베딩 기반 의미 추출을 결합한 후속 설계를 요구한다. STM 내용 공변량을 통한 매체별 단어 분포 차이 추정은 연산 비용 제약으로 수행하지 못하였다. 감성 측정에서 확인된 KNU-KoELECTRA 간 분류 일치도 49.0%는 한국 정책 뉴스 도메인에서 범용 사전학습 모형의 직접 적용이 신뢰성 있는 감성 측정을 보장하지 않음을 정량적으로 보여주며, 정책 뉴스 특화 감성 코퍼스 구축과 도메인 적응 학습은 선결 과제로 남는다. FREX 상위에 “농지”, “사장” 등 범코퍼스 고빈도 토큰이 다수 토픽에 걸쳐 출현하나, 이들은 농지연금 비교 맥락과 기관장 보도 맥락에서 실질적 의미를 가지므로 불용어로 제거하지 않았으며, Prob 열에 의한 토픽 변별이 이 한계를 보완한다.

후속 연구 의제로는 HOUSTAT 월별 행정데이터 확보 시 보도량(뉴스 건수) 및 감성-가입 시계열 인과 분석, BERTopic·SBERT 등 임베딩 기반 토픽모형과의 결과 비교, 깊은 프레이밍 측정을 위한 질적·계량적 결합 분석, 보도자료 송출과 가입 의향 간 실험적 인과 식별이 우선적으로 요구된다.

이 연구가 남기는 것은 일회성 발견이 아니라 반복 적용 가능한 분석 프레임워크이다. STM 공변량 설계, SimHash 중복 제거, 사건 중심 강건성 검증, 이중 감성 측정의 결합은 보금자리론·전세보증·MBS 등 한국주택금융공사가 운영하는 전체 상품군에 동일하게 적용할 수 있다.

참고문헌

- 금융위원회. (2026). 100세 시대, 주택연금 보장을 확대하고 편의성을 제고하겠습니다 [보도자료]. Retrieved from <https://www.fsc.go.kr/no010101/86211>
- 김강석, 박정은. (2025). 저성장기 대응 주택 정책의 변화와 과제: 디딤돌 대출 정책의 변화가 신혼 및 다자녀 가구에 미치는 영향을 중심으로. *주택금융연구*, 9(1), 39-57.
- 박상민, 나철원, 최민성, 이다희, 온병원. (2018). Bi-LSTM 기반의 한국어 감성사전 구축 방안. *지능정보연구*, 24(4), 219-240.
- 박치성, 신나리. (2021). 토픽모델링을 활용한 역대 대통령의 정책변화 분석. *한국정책과학학회보*, 25(4), 1-33.
- 윤은지, 강보영. (2023). 구조적 토픽 모델링과 의미연결망 분석을 통한 코로나19 백신 보도 프레임 연구. *한국컴퓨터정보학회논문지*, 28(5), 129-153.
- 이동영, 강상훈. (2025). 다핵 상권 젠트리피케이션 위험도 및 결정요인 분석: 머신러닝 기반 예측 모델과 공간분석 기법을 적용하여. *주택금융연구*, 9(1), 159-199.
- 이민철. (2024). Kiwi: 통계적 언어 모델과 Skip-Bigram을 이용한 한국어 형태소 분석기 구현. *디지털인문학*, 1(1), 109-136.
- 이준용, 김성희. (2018). 미세먼지 재해 보도의 프레임 분석: 구조적 주제모형(Structural Topic Modeling)의 적용. *한국언론학보*, 62(4), 125-158.
- 이호일, 노승한, 이수민. (2025). 고령가구의 자산구성이 삶의 만족도에 미치는 영향: 부동산자산 소유와 주거 특성을 중심으로. *주택금융연구*, 9(1), 109-133.
- 전동석. (2020). KoELECTRA: pretrained ELECTRA model for Korean (version 3) [computer software]. Retrieved from <https://github.com/monologg/KoELECTRA>
- 전병욱. (2022). 주택연금의 미시적 가입 의사결정 요인의 분석. *주택금융연구*, 6(2), 99-122.
- 주택금융연구원. (2020). *2020년 주택연금 수요실태조사*. 부산: 한국주택금융공사.
- 하석현, 김승희. (2024). 텍스트 마이닝을 활용한 부동산 전세 연구 경향에 대한 메타분석: KCI 등재 학술지 게재 논문을 중심으로. *대한부동산학회지*, 42(2), 135-158.
- 한국언론진흥재단. (2024). 빅카인즈(BIGKinds). Retrieved from <https://www.bigkinds.or.kr/v2/news/index.do>
- 한국주택금융공사. (n.d.). 주택금융통계시스템 (HOUSTAT). Retrieved from <https://houstat.hf.go.kr>

- 행정안전부. (2024). 주민등록 인구 및 세대현황. Retrieved from <https://jumin.mois.go.kr>
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3, 993-1022.
- Charikar, M. S. (2002). Similarity estimation techniques from rounding algorithms. In *Proceedings of the 34th Annual ACM Symposium on Theory of Computing (STOC 2002)* (pp. 380-388). Montreal, QC: ACM.
- Clark, K., Luong, M. T., Le, Q. V., & Manning, C. D. (2020). ELECTRA: Pre-training text encoders as discriminators rather than generators. Retrieved from <https://doi.org/10.48550/arXiv.2003.10555>
- Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication*, 43(4), 51-58.
- Hur, J., & Yang, J. (2024). South Korean newspaper coverage of Yemeni refugees: Analysis of topics and sentiments using machine learning techniques. *Asian Journal of Communication*, 34(1), 57-72.
- Jo, W., Lee, J., Park, J., & Kim, Y. (2020). Online information exchange and anxiety spread in the early stage of the novel coronavirus (COVID-19) outbreak in South Korea: Structural topic model and network analysis. *Journal of Medical Internet Research*, 22(6), e19455.
- Mimno, D., & McCallum, A. (2008). Topic models conditioned on arbitrary features with Dirichlet-multinomial regression. In *Proceedings of the 24th Conference on Uncertainty in Artificial Intelligence (UAI 2008)* (pp. 411-418). Corvallis, OR: AUAI Press.
- Mimno, D., Wallach, H. M., Talley, E., Leenders, M., & McCallum, A. (2011). Optimizing semantic coherence in topic models. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 262-272). Stroudsburg, PA: Association for Computational Linguistics.
- Munkres, J. (1957). Algorithms for the assignment and transportation problems. *Journal of the Society for Industrial and Applied Mathematics*, 5(1), 32-38.
- Oh, J., & Kim, J. (2025). YouTube as a social listening tool: Mining housing discourse on YouTube. *Cities*, 166, 106279.

- Organization for Economic Co-operation and Development (OECD). (2023). *Pensions at a glance 2023: OECD and G20 indicators*. Paris, France: OECD.
- Park, E. (2016). Naver sentiment movie corpus (NSMC) [Data set]. Retrieved from <https://github.com/e9t/nsmc>
- Roberts, M. E., Stewart, B. M., & Airoldi, E. M. (2016). A model of text for experimentation in the social sciences. *Journal of the American Statistical Association*, *111*(515), 988-1003.
- Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). stm: An R package for structural topic models. *Journal of Statistical Software*, *91*(2), 1-40.
- Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., Albertson, B., & Rand, D. G. (2014). Structural topic models for open-ended survey responses. *American Journal of Political Science*, *58*(4), 1064-1082.
- Scheufele, D. A. (2000). Agenda-setting, priming, and framing revisited: Another look at cognitive effects of political communication. *Mass Communication and Society*, *3*(2-3), 297-316.
- Shiller, R. J. (2015). *Irrational exuberance* (3rd ed.). Princeton, NJ: Princeton University Press.
- Shiller, R. J. (2019). *Narrative economics: How stories go viral and drive major economic events*. Princeton, NJ: Princeton University Press.
- Sun, C., Qiu, X., Xu, Y., & Huang, X. (2019). How to fine-tune BERT for text classification. In *China National Conference on Chinese Computational Linguistics* (pp. 194-206). Cham, Switzerland: Springer.
- Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., Davison, J., Shleifer, S., von Platen, P., Ma, C., Jernite, Y., Plu, J., Xu, C., Le Scao, T., Gugger, S., ... Rush, A. M. (2020). Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations* (pp. 38-45). Stroudsburg, PA: Association for Computational Linguistics.

(논문 접수일: 2026.03.24. 수정논문 접수일: 2026.05.25. 논문 채택일: 2026.06.05.)

Analyzing Housing Pension Policy Discourse Using Structural Topic Models: Topic Dynamics and Media-Type Framing Differences in News Coverage, 2007–2026

Seung-hyung Lee *

Abstract

This study analyzes the long-term trajectory and differentiation of Korean housing pension (reverse mortgage) policy discourse using Structural Topic Modeling (STM). From 12,799 housing pension-related news articles collected from BIGKinds (July 2007–March 2026), a 12-topic solution was adopted after SimHash near-duplicate removal. The discourse centered on policy and household debt (13.4%), asset management and retirement planning (11.9%), and enrollment statistics (10.1%). Period-specific (time-smoothed) comparisons showed enrollment statistics predominating under the Lee Myung-bak administration (18.0%), institutional design under Moon Jae-in (13.2%), and elderly-household and bequest-intention surveys under Lee Jae-myung (10.5%). Media-type differentiation was pronounced: regional dailies emphasized enrollment figures, national dailies intergenerational narratives, and the business press retirement planning. Following the March 2026 guarantee-fee reduction, the elderly-household and bequest-intention topic surged by +25.65 percentage points, though this is provisional given a post-event sample of 31 articles. The KNU sentiment lexicon and KoELECTRA agreed on only 49.0% of classifications, confirming a domain gap in policy-news sentiment measurement. Government-period estimates represent descriptive statistics conditional on time-smoothing rather than causal effects. This study contributes a replicable STM-based framework for policy-discourse monitoring in housing finance.

Keywords : Housing Pension, Structural Topic Model (STM), Policy Discourse, Media Framing, Text Mining

* Seung-hyung Lee, Corresponding author, Ph.D., Department of Education, Dong-A University, homolodus4780@gmail.com

© Copyright 2026 Housing Finance Research Institute. 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.

부록

각 토픽의 내용적 타당성을 기사 수준에서 확인하기 위해, findThoughts() 함수로 추출한 토픽별 비중(θ) 상위 3건의 기사 제목을 <부록 표 1>에 제시한다.

<부록 표 1> 12개 토픽별 우세 문서 상위 3건의 기사 제목

토픽	명명	우세 문서 기사 제목 (매체명, 발행일)
T10	정책·가계부채	가계부채 종합대책 발표...주택연금 활성화 포함(경향, 2017.10) / 주택연금으로 노후 해결?...가계부채 리스크 경고(매경, 2018.03) / 정부, 주택시장 안정화 방안에 주택연금 확대(한경, 2019.12)
T7	자산관리·은퇴설계	100세 시대 은퇴설계...주택연금이 답이 될까(조선, 2019.05) / 퇴직 후 자산관리 전략, 주택연금·개인연금 비교(한경, 2020.08) / 은퇴 후 월 200만원 만들기...주택연금 활용법(매경, 2021.03)
T4	가입실적·통계	주택연금 가입자 10만 명 돌파(연합, 2021.02) / 지난해 주택연금 신규가입 1만 건(뉴시스, 2020.01) / 주택연금 월평균 수령액 127만 원(뉴스1, 2025.03)
T9	제도 설계·가입기준	주택연금 가입연령 55세로 하향(동아, 2020.04) / 주택가격 상한 12억 확대(서울경제, 2023.10) / 우대형 주택연금 도입(한겨레, 2014.07)
T5	고령가구·상속의향	노인 10명 중 4명 빈곤...주택연금 왜 안하나(KBS, 2024.12) / 상속 포기 못해 주택연금 망설이는 고령층(중앙, 2023.05) / 초고령사회 진입...주택연금 재조명(한국, 2025.01)
T6	HF 기관·업무협약	주금공, 광주시와 주택연금 활성화 업무협약(광주, 2019.06) / 주금공 사장, 지역 순회 주택연금 설명회(경남, 2020.09) / 주금공-신한은행 MOU(파이낸셜, 2022.11)
T1	노후·세대·자식연금	부모 집 물려받을까 vs 주택연금...세대 갈등(조선, 2022.04) / 자식연금 대신 주택연금...노후 자립(중앙, 2023.08) / 아파트 한 채가 유일한 노후자금(한겨레, 2021.06)
T2	제도 운영·지급방식	주택연금 종신 vs 확정기간...어느 방식이 유리(매경, 2020.02) / 역모기지 상환 방식 변경(서울경제, 2018.07) / 주택연금 보험료 구조 개편안(머니투데이, 2025.11)
T11	금리·수령액	기준금리 인상에 주택연금 수령액 줄어드내(한경, 2022.07) / 보금자리론 금리 인상...월지급금 영향(이데일리, 2023.01) / 2026년 수령액 3.13% 인상(연합, 2026.03)
T8	집값변동·해지·지역격차	집값 하락에 주택연금 해지 급증(SBS, 2023.03) / 수도권-지방 수령액 격차 확대(경향, 2024.05) / 국감서 주택연금 해지율 지적(뉴시스, 2023.10)
T3	실버타운·노인주거	실버타운 입주 시 주택연금 활용 가능할까(중앙, 2019.11) / 농촌 빈집 활용 노인주거...주택연금 연계(농어민, 2021.08) / 도시재생 사업지 주택연금 상담 급증(부산, 2020.03)
T12	명절 안내·CEO 인사	추석 연휴 주택연금 상담·신청 안내(뉴스1, 2023.09) / 주금공 신임 사장 취임...활성화 의지(연합, 2022.03) / 설 연휴 금융상품 만기 점검(파이낸셜, 2024.02)

주: findThoughts() 함수로 추출한 각 토픽 비중(θ) 상위 3건. 기사 제목은 STM findThoughts() 출력에서 토픽 비중(θ) 상위 3건을 선별하였다.

〈부록 표 1〉에서 확인되듯이, 각 토픽의 우세 문서 기사 제목은 〈표 3〉의 FREX·Prob 단어에 기반한 토픽 명명과 일관된다. T10(정책·가계부채)의 우세 문서는 가계부채 종합대책과 주택연금 활성화를 연계하는 정책 보도이며, T5(고령가구·상속의향)의 우세 문서는 노인빈곤과 상속 갈등을 다루는 사회구조적 보도이다. T12(명절 안내·CEO 인사)의 우세 문서는 연휴 안내와 기관장 인사 보도로서, 주택연금 고유의 정책 이슈가 아닌 기관 운영 맥락의 보도임을 확인해 준다.

주택시장에서 가구의 차입금 상환부담과 추가적인 부채의 관계에 대한 연구

김종희*

요약

본 연구는 지난 13년 동안의 마이크로 자료를 이용, 가구의 대출에 대한 원리금 상환부담과 추가적인 부채와의 관계, 그리고 한계대출을 이용한 과도한 대출수준의 영향에 대하여 분석하고 있다. 본 연구의 분석 결과는 다음과 같다. 첫째, 가구가 차입을 시도할 때 적용받는 금리가 상승하면 대출을 받지 않을 가능성이 높은 것으로 나타났으나, 청년층 가구의 경우 전체 가구의 경우와는 달리, 금리가 상승해도 대출을 받을 가능성이 높은 것으로 나타났다. 둘째, 가구의 소득으로 설명되지 않는 부채의 증가는 상환부담 가구가 될 가능성을 높이는 것으로 나타났으며, 특히 청년층 가구에게서 이러한 현상이 강하게 나타났다. 셋째, 가구는 자신의 한계대출 수준에 근접하여 실제 대출을 받는 것으로 나타났으며, 특히 청년층의 경우에는 한계수준을 상회하는 대출을 받고 있었다. 그리고 이와 같은 과도한 대출은 해당 가구의 추가적인 부채를 증가시키고 있었다.

핵심어 : 대출 상환부담, 금융규제, 추가적인 부채, 한계대출

1. 서론

주택을 점유한다는 것, 즉 주택의 수요는 서비스 측면에서의 주택을 점유하는 것과 같다. 주택을 구매한다는 것은 개별적이고 이질적인 주택 자체의 구입에 의미가 있기보다는 더 높은 수준의 주택 서비스를 얻고자 하는 것에 의미가 있으므로, 주택의 구매는 주택에 대한 수요이며, 이러한 주택의

* 김종희, 주저자, 전북대학교 경제학부 교수, jonghk5@jbnu.ac.kr

© Copyright 2026 Housing Finance Research Institute. 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.

수요는 주택의 서비스 수요라고 할 수 있다. 결국, 주택수요(housing demand)는 주택의 구매를 위한 가구의 지불능력은 물론 주거비용의 변화까지 고려된 경제적 개념인 것이다.

현 시점에서 한국의 주택시장에서 주택수요는 큰 변화가 발생하고 있다. 출산율의 급격한 저하와 노령인구의 지속적인 증가로 인하여 인구구조가 변화하고 있으며, 1~2인 가구를 중심으로 한 소형 가구의 비중이 늘어나며 가구형태의 변화도 급속히 진행되고 있다. 또한 노후주택의 대체수요도 증가하고 있으며, 부동산 가격의 상승으로 인하여 20~30대 청년층 가구의 대출을 이용한 주택구매의 증가, 즉 '영끌매수' 현상도 증가하고 있다. 이에 따라 향후 주택시장에서 신규 주택수요에는 많은 변화가 예상된다.

이와 같이 주택시장에서 주택수요의 점진적이며 또한 이례적인 변화가 발생하면서 기존의 소득, 가격, 경기상황 등의 주택수요 결정요인으로부터 최근에는 정책당국의 금융규제와 가구의 주택구매력에 대한 관심이 높아지고 있다.

주택은 경제주체가 소비할 수 있는 단일재화 중 가장 큰 금액을 차지하고 있으며, 이에 따라 주택의 가격과 이를 소비할 수 있는 경제주체의 주택구매여력은 주택수요에 가장 큰 영향을 미칠 수 있다. 즉 주택수요는 경제적 유인에 의해 의도된 주택구매능력을 의미한다고도 할 수 있다(변창흠·이희정, 2002). 그리고 이와 같은 가구의 주택구매력에 가장 큰 비중을 차지하고 있는 것은 금융기관으로부터의 차입, 즉 대출이다.

그러나 주택구매력의 확보를 위한 과도한 대출은 가계의 부채를 증가시키며 경제 전반에 대한 신용리스크를 가중시킨다. 한국의 가계부문에서 가처분소득 대비 부채의 비율은 2010년 158%에서 2018년 185.9%, 2021년에는 206.5%까지 상승하였으며, 2023년에는 186.5%를 기록하였다(김수현·황설웅, 2023). 가구의 금융기관으로부터의 차입은 이에 대한 상환부담을 발생시키며, 이는 추가적인 부채의 발생을 의미한다. 그리고 가구 감당할 수 있는 수준 이상의 대출은 상환부담을 가중시키며 추가적인 부채를 증가시키게 될 가능성을 높게 만든다.

이에 따라 차입을 시행한 가구의 차입규모와 가구가 감당할 수 있는 차입규모를 추정하여 비교하는 것은 큰 의미가 있을 수 있다. 이를 바탕으로 향후의 정책변화에 대하여 변화하는 주택수요를 예측하는 것은 분명 의미 있는 연구이다.

이와 같은 상황에서 다음과 같은 질문을 제기할 수 있다.

첫째, '가구가 안정적으로 자신의 부채를 통제할 수 있는 적정 대출수준, 즉 적정 원리금 상환수준은 어떻게 결정될 수 있는가?'이다. 이에 대해서는 각 가구의 소득, 자산, 부채는 물론 그 밖의 가구주의 특성과 더불어, 가구가 스스로 감당할 수 있는 대출수준, 즉 감내할 수 있는(tolerable) 상환규모를 파악해야 한다.

둘째, ‘정책당국의 주택시장에 대한 금융규제의 변화는 가구의 감당할 수 있는 차입규모에 어떠한 영향을 미치는가?’이다. 이에 대해서는 가구의 감당할 수 있는 차입규모, 즉 한계대출의 규모를 직접 파악해야 하며, 가구가 직접적으로 영향을 받는 정책당국의 규제의 정도 역시 수치화 되어야 한다.

셋째, ‘가구가 감내할 수 있는 대출수준 이상의 차입이 발생하였을 경우, 해당 가구가 통제할 수 있는 부채에는 어떠한 변화가 발생하는가?’이다. 이에 대해서는 가구의 현재의 부채수준은 물론, 장기간에 걸친 가구의 소득과 주택의 가격 요인으로는 설명할 수 없는 추가적인 부채의 수준을 파악하는 것이 중요하다.

이와 같은 질문을 검증하기 위하여 본 연구에서는 지난 13년간의 주택금융 및 보금자리론 실태조사 의 가구별 마이크로 자료를 이용, 가구가 대출로 인하여 발생할 수 있는 원리금 상환에 대한 부담의 정도를 추정한다. 그리고 가구의 상환부담에 해당 가구의 부채수준은 물론 추가적인 부채의 발생이 어떠한 작용을 하는지도 분석한다. 또한 정책당국의 금융규제의 정도를 적용하여 가구가 감당할 수 있는 대출수준, 즉 한계대출 수준을 추정, 해당 가구의 과도한 대출이 기존의 추가적인 부채와 상환부담 간의 관계를 어떻게 변화시킬 수 있는지를 분석한다. 특히 전체 가구로부터 청년층의 가구를 별도로 추출하여 주택시장에 새롭게 진입하는 청년층 가구주에게서 어떠한 차이점이 발생하는지를 분석한다.

가구의 원리금 상환부담과 추가적인 부채, 그리고 한계대출을 이용한 과도한 대출수준의 영향 등에 대한 본 연구의 분석과정을 통해 규제의 변화와 더불어 가구의 차입규모의 변화에 대한 의미 있는 판단이 가능해 질 수 있다. 특히 주택시장에 신규로 진입하는 상대적 금융제약층인 청년층 가구의 대출수준과 부채의 변화에 대한 의미 있는 시사점을 제공할 것으로 판단된다. 이 점이 본 연구의 연구목적이며, 기존의 연구들과의 차별성이다.

II. 기존문헌 검토

가구가 주택을 구매할 수 있는 능력, 즉 주택구매력에 대한 설명은 주택의 수요와 관련이 있다. 주택의 수요의 추정과 관련된 연구로는 인구의 연령구조에 초점을 맞춘 M-W 모형과 연령 및 시간효과를 동시에 고려한 APC 모형이 있다. 최근에는 소득과 가격의 탄력성을 이용한 2단계 추정방식인 Heckman 2단계가 사용되고 있다.

Ermisch et al.(1996)은 소득과 가격의 탄력성을 이용하여 주택수요를 추정하면서 OLS에서 선택적 편의가 있음을 밝히고 있으며, Chen & Jin(2014)은 Heckman 2단계를 사용, 자가 가구의 주택수요를 추정하면서 이와 같은 선택적 편의를 보정하고 있다. 국내의 연구에서는 김순용·박현수(2015),

윤주현·김혜승(2000) 등이 주택수요를 추정에서 자가와 차가를 구분하여 Heckman 2단계를 이용하였다. 본 연구는 금융기관으로부터의 차입으로 인한 수요의 변화, 즉 구매력의 변화를 설명할 수 있는 모형을 이용하며, 이는 소득과 가격의 탄력성을 이용한 방식과 관련이 있다.

가구의 감내 가능한 수준 이상의 차입규모, 즉 원리금 상환규모는 가구의 상환부담을 증가시키며, 상환부담의 증가는 해당 가구의 경제행위에 대한 변화를 유발할 수 있다. Zhang et al.(2025)에 의하면 가구의 상환부담의 증가는 가구소비의 행태에 유의미한 영향을 미치며, 특히 부정적인 효과를 가져 올 수 있다. Gelos et al.(2004)은 채무부담이 더 강한 가구일수록 정책당국의 정책변화에 더 민감하게 반응하게 됨을 밝혔다. 국내의 연구에서 전병욱·박나라(2022)는 가계부채가 증가할수록 주택 매수비율이 커지는 부채와 소득 간의 (+)의 유의성 관계를 언급하였다.

최근에는 가구가 가지고 있는 기존의 자산이나 부채의 요인과 더불어, 추가적인 자산이나 부채를 이용, 가구의 이분법적 특성을 분석하는 연구도 등장하고 있다. 심승규(2022)는 항상소득으로 설명되지 않는 추가적인 부채를 이용, 가구의 자가확률을 분석하고 있으며, 이준민 외(2022)는 가구자산의 잔차와 경상소득잔차를 이용, 주택면적과 점유형태를 분석하고 있다.

가구가 감당할 수 있는 대출수준, 즉 한계대출은 주택시장에 대한 정책당국의 규제의 변화에 크게 의존한다. 이에 따라 주택시장에서 정책당국에 의한 금융규제의 역할 역시 다양하게 연구되어 왔다. Hatchondo et al.(2014)은 주택가격의 변동에 따른 가계부채에서의 파산은 LTV(loan-to-value) 규제가 강화될수록 크게 증가하지 않음을 보였으며, 특히 Hatchondo et al.(2015)은 주택담보대출에서의 부실채권의 발생과 주택에 대한 수요에 미치는 영향으로서 LTV 규제를 들고 있다. Robinson & Yao(2015)는 경기변동의 진폭과 주기를 줄이는 데 효과적인 방법을 연구하면서 LTV 규제의 장단기 영향을 분석하고 있다.

Kuttner & Shim(2013)은 금리를 이용한 규제방식 이외의 규제가 주택담보대출에 미치는 영향, 즉 LTV, DTI(debt-to-income) 규제의 영향을 분석하였으며, 이 중 DTI가 더 효과적임을 밝혔다. McDonald(2015)는 주택경기애 있어서 LTV와 DTI의 규제 강화시기와 완화시기의 효과를 비교분석하였다. 규제강화의 효과가 더 강하게 나타났으며, 특히 주택가격이 높을 때 이러한 효과가 더 큰 것으로 나타났다. Crowe et al.(2011)은 주택가격 상승률에 대한 LTV 규제의 수준의 영향을 연구하였으며, LTV 규제의 완화가 주택가격을 상승시키는 것으로 나타났다.

국내의 연구에서 이동규 외(2009)는 낮은 DTI 규제수준은 지역의 가계대출과 부동산 거래를 증가시키며, 해당지역의 주택거래는 DTI 규제수준이 높을 경우에 크게 감소함을 밝혔다. 김종희(2019)는 DTI의 강화와 같은 금융규제는 차주의 원리금 상환부담을 증가시킴으로써 주택시장에 변화를 유발할 수 있음을 밝혔다.

이와 같은 정책 규제는 가구의 상환부담에 영향을 미칠 수 있으며, 그 영향의 정도는 가구의 특성에 따라 다르게 나타날 수 있다. Linneman et al.(1997)에서는 가구의 자산제약과 소득제약 중에서 주택의 점유형태결정과 주택소유비율에 더 유의미한 영향을 미치는 것은 자산제약이라고 주장하였다. Quercia et al.(2003)은 대출제약은 상대적으로 취약계층의 주택소유에 유의미한 영향을 미침을 지적하였다. 국내의 연구에서 윤병우·최경욱(2021)은 사회적 취약계층에 해당하는 가구와 비은행대출과 신용대출, 그리고 만기 일시 상환대출을 보유한 가구에서 상대적으로 상환에 대한 위험성이 높게 나타남을 밝혔다.

이상과 같은 기존 연구들을 검토해 보면, 정책당국의 금융규제를 적용하는 방식에서 금융규제가 완화되었거나 강화되었던 시기에 더미를 주는 방식을 채택하고 있다. 또한 가구의 상환부담과 관련된 기존의 연구들에서도 대부분 상환부담이 발생하는 시점에 대한 기간의 특성에 초점이 맞추어져 있다. 즉 현재 원리금 상환부담을 느낀다고 응답한 가구에 더미를 주어 분석하는 방식이다. 그리고 추가적인 부채나 자산, 소득을 적용하는 방식에서는 장기간이 아닌 특정 연도에 대한 분석을 시도하고 있다.

그러나 가구의 상환부담과 추가적인 부채, 그리고 금융규제의 영향을 검증하기 위해서는 원리금 상환부담과 관계있는 가구의 적정 대출수준, 즉 한계대출 수준을 직접 추정할 필요가 있으며, 금융규제 역시 가구가 거주하는 지역과 다주택의 보유여부 등을 종합하여 규제수준을 가구별로 차등 산출할 필요가 있다. 또한 추가적인 부채 역시 항상소득 등의 영향이 제대로 반영될 수 있도록 장기간에 걸친 분석이 필요하다.

III. 모형 및 자료

본 연구는 다음과 같은 가설의 검증을 목적으로 하고 있다.

첫째, 주택시장의 주요 변수의 변화에 대한 민감도는 차입이 있는 가구와 없는 가구, 그리고 차입 가구 중에서도 상환부담이 있는 가구와 없는 가구에 따라 다르게 나타날 가능성이 크다.

둘째, 대출 차입에 대한 상환부담이 가증될수록 가구의 소득과 주택가격 요인으로 설명할 수 없는 추가적인 부채가 많아질 가능성이 크다.

셋째, 주택시장에 대한 정책당국의 금융규제는 가구가 감내할 수 있는 차입 수준, 즉 한계차입수준을 변화시키며, 이에 따라 해당 가구의 추가적인 부채 역시 많아질 가능성이 크다.

이를 검증하기 위하여 본 연구에서는 먼저, 가구의 대출 수요에 대한 결정요인을 추정한다. 대출에 대한 수요의 결정요인은 다음과 같은 Probit 모형을 이용하여 판단할 수 있다.

$$\Lambda \left(= \frac{P(Z_{it}, \alpha_i)}{1 - P(Z_{it}, \alpha_i)} \right) = \alpha_0 + \alpha_1 r_t + \alpha_2 p_t + \alpha_3 y_t + \epsilon_{it} \quad \langle \text{식 1} \rangle$$

$$Z_{it} = \begin{cases} 1 & \text{if } Z_{it}^* > 0 \\ 0 & \text{if } Z_{it}^* < 0 \end{cases}$$

여기에서 $\Lambda \left(= \frac{P(Z_{it}, \alpha_i)}{1 - P(Z_{it}, \alpha_i)} \right)$ 는 odds 비율로서 상대적인 대출 가능성을 나타낸다. 즉 $1 - P(Z_{it}, \alpha_i)$ 는 t 시점에 대출이 있는 가구를 의미하며, $P(Z_{it}, \alpha_i)$ 는 대출이 없는 가구이다. Z_{it}^* 는 관찰되지 않는 변수로서 실제 관찰되는 것은 Z_{it} 로서 대출 가구일 때 1, 비대출 가구일 때 0의 값을 갖게 된다.

기존의 연구에 의하면, 가구의 대출과 관련된 결정요인으로서, 금리(r_t)요인으로서는 금융권과 제2금융권의 주택담보대출 신규대출 금리를 가중평균하여 사용되며, 주택가격(p_t)에 대해서는 가구가 거주하고 있는 지역의 주택가격지수의 증가율이 사용되었다. 그리고 경제여건(y_t)은 역시 거주지역의 GRDP(gross regional domestic product), 즉 지역내 총생산의 증가율이 사용되었다(김중희, 2019).

그러나 개별 가구의 대출의 수요, 더 나아가서 대출과 관련된 상환부담에 대한 결정요인에는 해당 가구의 특징이 보다 직접적으로 반영될 필요가 있다. 이에 따라 본 연구에서는 앞선 <식 1>을 다음과 같이 응용하여 대출의 결정요인으로 분석한다.

$$A_{it} = \alpha_0 + \alpha_1 r_{it} + \alpha_2 \hat{p}_{it} + \alpha_3 \hat{y}_{it} + \epsilon_{it} \quad \langle \text{식 2} \rangle$$

위 식에서 A_{it} 는 현재 시점에서 대출의 유무에 대한 가구의 구분이다. 이때 대출은 주택담보대출과 전세자금대출을 모두 포함한다. 현재 시점에서의 대출수요에 대한 결정요인으로서 r_{it} 는 i 가구가 받은 대출에 대한 적용 금리이며, 변동금리의 경우에도 최초 대출 시점의 금리를 적용한다. 만약 대출을 받지 않는 가구라면 금융권과 제2금융권의 주택담보대출 신규대출 금리를 가중평균하여 사용한다.

\hat{p}_{it} 와 \hat{y}_{it} 는 주택가격과 가구소득의 요인으로서 장기간에 걸친 행태변화를 예측하기 위하여 장기적 특성을 반영할 수 있는 변수를 고려한다.

\hat{p}_{it} 는 i 가구가 거주하고 있는 주택의 상대가격이다. 주택의 가격으로는 실제가격(real price)과 상대가격(relative price)이 사용될 수 있다. 주택은 자가와 차가로 구분되며, 동일한 주택에서 주택 가격은 자가는 기회비용, 차가는 임대료로 정의된다. 따라서 두 비용을 동일하게 사용하여 주택가격을 산출할 필요가 있으며, 이와 같은 가격 방식은 상대가격으로서 임대가격 대비 구매가격이 해당된다. 주택시장의 장기균형에서는 차익거래를 통하여 주택 소유자의 기회비용이 동일한 주택에 대한 상대적 임대료와 같기 때문이다(Muellbauer, 2012). 주택의 상대가격은 다음과 같은 모형을 이용하여 추정한다.

$$\ln p_{it} = a + \text{barea}_{it} + \sum_{j=2}^4 D_j \text{type}_{it} + \sum_{j=2}^{17} D_j \text{region}_{it} + \epsilon_{it} \quad \langle \text{식 3} \rangle$$

이 모형은 헤도닉 모형(hedonic price model)으로서, 여기에서 $\ln p_{it}$ 는 가구가 거주하고 있는 현재 주택가격의 로그 값이다. area 는 현재 거주하고 있는 주택의 면적(m^2)이다. 그리고 $\sum_{j=2}^4 D_j \text{type}$ 는 거주하고 있는 주택의 유형으로서 단독/다가구, 아파트, 연립/다세대주택/빌라, 기타(주거용 오피스텔 등)에 따라 터미가 부여된다. $\sum_{j=2}^{17} D_j \text{region}$ 는 가구가 현재 거주하는 지역의 터미로서 서울 등 17개 광역도에 터미를 부여한 것이다. 따라서 위 식을 통하여 추정된 각 계수 값에 대하여 가구의 실제 자료를 적용하면 주택의 구매($\ln \hat{h}p_{it}$)와 임대가격($\ln \hat{r}p_{it}$)이 추정되며, 상대가격(\hat{p}_{it})은 이 두 가격의 비율에 의하여 산출된다.

\hat{y}_{it} 는 가구의 소득에 대한 로그 값이다. 이에 대해서는 가구의 현재의 소득이 사용될 수 있다. 그러나 주택은 내구재이며 장기간에 걸쳐 사용되므로 주택구매나 임차를 위한 소득으로서는 장기간에 대하여 가구가 얻을 수 있는 항상소득(permanent income)이 바람직하다. 현재 소득이 같다고 하더라도 미래의 기대소득이 높으면 현재의 소비가 커질 수 있으며, 반대로 기대소득이 낮아지면 현재의 소비가 줄어들 수 있기 때문이다(Tse & Raftery, 1999).

t 시점의 i 가구의 항상소득은 가구의 특성을 이용하여 다음과 같이 추정될 수 있다.

$$\ln y_{it} = a + \text{bage}_{it} + \text{cage}_{it}^2 + \text{dln}A_{it} + \text{esex}_{it} + \sum_{j=2}^4 D_j \text{job}_{it} + \epsilon_{i,t} \quad \langle \text{식 4} \rangle$$

여기에서 $\ln y_{it}$ 는 가구의 연간소득에 대한 로그 값이며, age_{it} 가구주의 연령을, 그리고 age_{it}^2 은

연령의 제곱으로서 시간효과를 나타낸다. $\ln A_{it}$ 는 총자산에 대한 로그 값이며, sex_{it} 는 가구주의 성별더미이다. $\sum_{j=2}^4 D_j job_{it}$ 는 가구주의 직업을 나타내는 더미이다. 기존의 연구에서는 가구주의 교육수준을 나타내는 더미가 주로 사용되었으나, 최근의 마이크로 자료의 응답설문에서는 가구주의 교육수준에 대한 질문이 없어지는 추세이다. 이에 따라 본 연구에서는 가구주의 직업을 사용하며, 생산직을 기준($j = 1$)으로 관리·전문직 2, 사무직 3, 판매직 4, 기타 기능직 등 5로 구성된다.

이와 같이 i 가구의 대출 수요에 대한 결정요인을 구성하면 가구가 직접적으로 적용받는 금리는 물론, 상대가격과 항상소득을 추정함에 있어, i 가구의 자산, 연령, 교육수준, 거주지의 면적과 유형 등의 특징이 모두 반영된다는 장점이 있다.

다음 단계는 가구의 대출에 대한 상환부담의 정도를 보다 세분시켜, 상환부담 가구의 특징을 분석하는 것이다. 즉 전체 가구를 다음과 같이 세 가지의 계층(class)으로 분류한다.

첫 번째 계층(A)은 현재 원리금 상환에 대한 부담을 가지고 있는 가구이며, 두 번째 계층(B)은 현재 원리금 상환에 대한 부담이 없는 가구다. 세 번째 계층(C)은 현재 대출이 없는 가구이다. 이때 대출에 대한 원리금 상환부담을 정의하는 것이 중요하다. 본 연구에서는 가구가 적정 수준의 원리금 상환규모라고 밝힌 금액을 실제 해당 가구가 받은 대출에 대한 원리금 상환규모와 비교한다. 그리고 실제 상환규모가 적정 상환규모보다 큰 가구를 원리금 상환부담이 있는 가구로, 실제 상환규모가 적정 상환규모보다 작거나 같은 가구를 원리금 상환부담이 없는 가구로 정의한다.

이와 같은 분류 하에 각 계층의 특성은 다음의 다항 프로빗(multinomial probit) 모형을 구성하여 결정할 수 있다.

$$\begin{aligned}
 A_{it}^A &= \beta_0 + \beta_1 r_{it} + \beta_2 \hat{p}_{it} + \beta_3 \hat{y}_{it} + \epsilon_{it} \\
 A_{it}^B &= \beta'_0 + \beta'_1 r_{it} + \beta'_2 \hat{p}_{it} + \beta'_3 \hat{y}_{it} + \epsilon_{it} && \langle \text{식 5} \rangle \\
 A_{it}^C &= 0
 \end{aligned}$$

여기에서 A_{it}^A 는 현재 원리금 상환부담이 있는 가구이며, A_{it}^B 는 현재 원리금 상환부담이 없는 가구이다. 그리고 A_{it}^C 는 대출이 없는 가구이다. 따라서 만약 $A_{it}^A > A_{it}^B$ 이고 $A_{it}^A > A_{it}^C$ 이면 현재 원리금 상환부담이 있는 가구가 선택되며, $A_{it}^B > A_{it}^A$ 이고 $A_{it}^B > A_{it}^C$ 이면, 현재 대출 → 원리금 상환부담이 없는 가구가 선택된다. 그리고 그 외의 경우에는 대출이 없는 가구가 선택된다.

다음으로, 가구의 원리금 상환부담에 영향을 미치게 되는 가구의 부채, 즉 추가적인 부채를 추정한다. 이때 가구의 추가적인 부채란 기존의 가구가 가지고 있는 부채로부터 해당 가구의 소득의 요인으로 설명할 수 없는 부분을 의미한다. 일반적으로 가구의 소득과 부채는 (+)의 상관관계를 갖는다. 소득이 많은 가구가 상환능력을 인정받아 더 많은 차입을 할 수 있는 것이다(심승규, 2022). 그러나 소득과 부채 간의 이와 같은 관계를 통해서는 가구의 상환부담과 부채 간의 관계를 정확하게 설명하는 데 한계가 있다. 가구의 상환부담과 부채 간의 보다 정확한 관계를 파악하기 위해서는 가구의 일시적인 소득을 제외한 항상소득과 다른 주요 특성변수를 이용, 이들로 설명되지 않는 부채의 잔차인 추가적인 부채를 고려해야 한다. 즉 추가적인 부채는 가구의 부채에서 항상소득의 기여분을 분리한 것이다. 이준민 외(2022)는 주택점유형태를 분석하면서 자산(소득)에서 소득(자산)의 기여분을 분리한 잔차를 사용한 바 있다.

본 연구에서는 이를 응용하여 다음의 식을 구성하여 가구의 부채의 변화를 분석한다.

$$D_{it} = \gamma_0 + \gamma_1 A_{it} + \gamma_2 \hat{y}_{it} + \gamma_3 job_{it} + \gamma_4 age_{it} + \gamma_4 age_{it}^2 + \epsilon_{it} \quad \langle \text{식 6-1} \rangle$$

$$\hat{D}_{it} = D_{it} - \hat{\gamma}_0 - \hat{\gamma}_1 A_{it} - \hat{\gamma}_2 \hat{y}_{it} - \hat{\gamma}_3 job_{it} - \hat{\gamma}_4 age_{it} - \hat{\gamma}_4 age_{it}^2 \quad \langle \text{식 6-2} \rangle$$

위 식에서 D_{it} 는 가구의 부채총액의 로그 값이며, A_{it} 는 총자산의 로그 값이다. \hat{y}_{it} 는 앞서 추정된 가구의 항상소득이며, job_{it} 역시 앞서 언급하였던 가구주의 교육수준에 대한 터미이다. age_{it} 는 연령이다. 따라서 <식 6-1>의 추정계수들을 활용하여, <식 6-2>와 같이 부채의 잔차를 추출, 이를 해당 가구의 설명되지 않는 부채(\hat{D}_{it})로 활용하는 것이다. 그리고 앞선 <식 5>를 다음과 같이 확장하여 상환부담 가구에서 부채가 차지하는 영향을 판단할 수 있게 된다.

$$A_{it}^A = \begin{cases} \delta_0 + \delta_1 D_{it} + \delta_2 r_{it} + \delta_3 \hat{p}_{it} + \delta_4 \hat{y}_{it} + \epsilon_{it} \\ \delta'_0 + \delta'_1 \hat{D}_{it} + \delta'_2 r_{it} + \delta'_3 \hat{p}_{it} + \delta'_4 \hat{y}_{it} + \epsilon_{it} \end{cases} \quad \langle \text{식 7} \rangle$$

위 식에서 A_{it}^A 는 앞선 <식 5>에서의 현재 원리금 상환부담이 있는 가구이다. 즉 상환부담 가구확률에 대하여 각각 가구의 부채총액(D_{it})과 부채잔차(\hat{D}_{it})를 사용하여 추정된 $\hat{\delta}_1$ 과 $\hat{\delta}'_1$ 의 계수 값의 유의성과 부호를 비교하여 추가적인 부채와 상환부담 가구가 될 가능성간의 관계를 판단하는 것이다. 마지막 단계는 가구의 상환부담과 추가적인 부채와의 관계에 대하여 정책당국의 금융규제가 미치

는 영향을 밝히는 것이다. 이를 위해서는 먼저, 가구의 과도한 대출을 정의하는 것이 중요하다. 본 연구에서의 가구의 과도한 대출의 규모는 해당 가구의 한계대출 수준으로부터 실제 대출수준이 차이가 나는 정도이며, 다음과 같이 정의된다.

$$O_{it} = l_{it} - \left(f \times Y \times \frac{(1+r)^n - 1}{r(1+r)^n} \right)_{it} \quad \langle \text{식 8} \rangle$$

여기에서 l_{it} 는 가구의 실제 대출수준이다. 그리고 가구의 한계 원리금 상환수준, 즉 한계 대출수준은 먼저, 원리금 균등분할 상환을 가정하고, 해당 가구의 대출금리 r 과 만기가 n 년인 대출의 상환금액 비율을 $\frac{r(1+r)^n}{(1+r)^n - 1}$ 과 같이 정의하고 여기에 해당 가구의 월 소득(Y)과 정책당국의 금융규제의 정도인 f 를 적용하여 산출된다.

이때, 정책당국은 금융규제의 정도인 f 를 통하여 가구의 한계대출 수준에 영향을 미칠 수 있게 된다. 주택시장에 대한 정책당국의 금융규제로서는 부채상환비율, 즉 DTI를 사용한다. 분석기간 동안의 주택시장에 대한 규제는 2011년부터 강화되었다가 2014년에 완화되었고, 2017년부터 2022년까지 강화가 지속되었으며, 2022년 말과 2023년 초부터 다시 완화되었다.

따라서 본 연구에서는 연도별, 가구가 거주하는 지역, 그리고 보유주택의 수 등을 고려하여 DTI(f)를 차등 적용한다. 즉 수도권에 거주하는 2주택 이상의 다주택 보유 가구에는 2018년 이후부터 50%의 f 를 적용하며, 서울, 경기, 세종 등의 조정대상지역의 거주 가구는 2014년부터 2016년까지 60%의 f 를 적용하고 2017년 이후에는 50%를 적용한다. 그리고 서울을 비롯하여 과천시와 성남시(분당구), 광명시, 하남시 등의 경기도 지역, 세종시, 대구시(수성구) 등의 투기과열지역의 거주 가구는 2018년 이후부터 40%의 f 를 적용한다.

이에 따라, 정책당국의 금융규제가 강화되면 DTI인 f 는 낮아지며, 결국 가구의 한계대출수준은 줄어들게 되는 구조이다. 그리고 다음과 같은 식을 구성하여 가구의 과도한 대출이 추가적인 부채를 발생시키는 경로(channel)를 파악한다.

$$\hat{D}_{it} = \eta_0 + \eta_1 O_{it} + \eta_2 r_{it} + \eta_3 \hat{p}_{it} + \eta_4 \hat{y}_{it} + \epsilon_{it} \quad \langle \text{식 9} \rangle$$

해당 가구의 과도한 대출(O_{it})로 인하여 추가적인 부채(\hat{D}_{it})가 증가하게 된다면, 위 식에서 추정된 $\hat{\eta}_1$ 의 계수 값은(+)를 보일 것이며($\hat{\eta}_1 > 0$), 이를 통하여 가구의 차입금 상환부담의 증가가 소득요인

으로는 설명할 수 없는 추가적인 부채를 증가시키게 됨을 확인할 수 있게 된다. 즉 위 식은 앞선 <식 4>에서 항상소득을 추정하고, 이를 이용하여 <식 6>에서 항상소득으로는 설명할 수 없는 추가적인 부채를 추정, 가구의 과도한 대출이 해당 가구의 일정 소득으로는 설명할 수 없는 부채에 어떠한 영향을 미치게 되는지를 밝히는 구조인 것이다.

본 연구에서는 주택금융공사에서 제공하는 ‘주택금융 및 보금자리론 실태조사’의 마이크로 자료를 이용한다. 즉 2011년부터 2023년까지의 65,403 가구(일반가구)를 대상으로 위와 같이 구성된 모형을 이용, 분석을 진행한다. 먼저, 다음의 <표 1>에는 전체 가구 중 차입가구의 분포가 나타나 있다.

전체 65,056가구 중 현재 대출을 받고 있는 가구는 23,614 가구로서 전체의 36.3% 수준이며, 이 중 주택담보대출 가구는 18,318가구로서 전체의 28.2%이며, 전세자금대출 가구는 5,322가구로서 전체의 8.2% 수준이다. 다음의 <표 2>에는 가구의 특성별 기초통계량이 나타나 있다.

<표 1> 차입 가구의 분포

	전체 가구	대출 가구	주택담보 대출 가구	전세자금 대출 가구
가구 수	65,056	23,614	18,318	5,322
비중(%)	100.0	36.3	28.2	8.2

<표 2> 주요 변수의 기초통계량

변수	전체 가구			청년층 가구		
	개체수	평균	표준편차	개체수	평균	표준편차
연령	65,043	49.0	13.5	13,601	31.6	4.5
총자산(만 원)	65,043	31,281.8	27,789.3	13,601	19,483.9	20,276.0
총부채(만 원)	64,646	4,933.2	6,608.1	13,601	4,969.0	6,383.5
주택 구매가격(만 원)	43,368	29,095.5	26,721.4	5,176	24,795.4	18,560.2
주택 임차가격(만 원)	41,131	9,882.6	8,971.5	11,664	9,676.4	9,259.7
대출금액(만 원)	23,576	9,199.6	10,234.6	4,916	8,125.8	6,682.9
대출금리(%)	23,583	3.81	0.94	13,601	3.96	0.88
월 소득(만 원)	65,043	396.9	265.3	13,601	352.9	244.5
월 상환 원리금(만 원)	18,250	59.5	48.6	2,515	62.5	50.3

주: 2011년~2023년의 평균 수치임.

분석대상 가구의 가구주의 평균 연령은 49세이며, 이들의 총자산은 평균 3억 1,282만 원, 총부채는 4,933만 원으로 나타났다. 가구가 현재 거주하고 있는 주택의 구매가격과 임대가격은 각각 2억 9,095만 원과 9,882만 원으로 나타났다. 그리고 가구가 현재 차입하고 있는 대출금액은 평균 9,199만 원이며, 평균 대출 금리는 3.81%로 나타났다. 한편, 가구의 월 소득은 약 397만 원이며, 대출에 대한 월 평균 원리금 상환금액은 약 60만원 수준으로 나타나, 월 소득의 15% 이상을 원리금을 상환하는 데 사용하고 있는 것으로 나타났다.

한편, 청년층(연령 40세 미만) 가구의 경우 평균 연령은 31.6세이며, 총자산은 1억 9,484만 원으로 전체 가구에 비하여 작은 반면, 부채의 경우에는 4,969만 원으로서 오히려 높은 것으로 나타났다. 거주하고 있는 주택의 구매와 임대가격, 그리고 총대출 금액은 모두 전체 가구에 비하여 작은 것으로 나타났으나, 대출금리는 3.96%로서 더 높게 적용받는 것으로 나타났다. 특히 전체 가구에 비하여, 청년층 가구의 월 소득은 353만 원으로서 낮은 반면, 대출에 대한 원리금 상환금액은 월 62.5만 원으로서 더 높게 나타났다. 즉 청년층 가구의 경우 월 소득의 18% 이상을 원리금 상환으로 사용하고 있는 것이다.

IV. 분석결과

1. 차입금 상환부담 가구의 결정요인

먼저, 대출(차입)가구의 특성을 분석한다. 이를 위해서는 가격요인과 소득요인을 추정해야 한다. 앞선 언급한대로, 대출수요에 대한 가격요인으로서의 가구 현재 거주하고 있거나 거주를 예상하고 있는 주택의 상대가격을 이용한다. 그리고 소득요인으로서의 가구의 항상소득을 이용한다.

다음의 <표 3>에는 앞선 <식 3>과 <식 4>를 이용, 주택가격의 결정요인과 가구의 소득의 결정을 분석한 결과가 나타나 있다.

주택의 가격 요인에 대한 분석결과, 주택의 구매가격과 임대가격 모두 유사한 결과가 나타나고 있다. 즉 가구가 현재 거주하고 있는 주택의 면적($area_{it}$)은 주택의 구매($lnhp_{it}$) 및 임대가격($lnrp_{it}$)에 모두 (+)의 영향을 미치고 있다. 가구가 거주하고 있는 주택의 유형의 경우, 단독주택($1type_{it}$) 대비 아파트($2type_{it}$)의 경우 주택 가격과의 (+)의 유의성이 가장 강하게 나타나고 있는 반면, 주거용 오피스텔($4type_{it}$)의 경우에는 단독주택 대비 (-)의 유의성을 보여, 상대적으로 가격이 낮은 것으로 나타났다. 즉 단독주택과 비교하여 상대적으로 매매가 대비 전세가가 낮은 것이며, 이는 주거용 오피

〈표 3〉 주택가격과 소득의 결정요인

	주택 구매가격	주택 임대가격	소득	
<i>Dep V.</i>	$lnhp_{it}$	$lnrp_{it}$	<i>Dep V.</i>	lny_{it}
$area_{it}$	0.014 (0.001) ^{***}	0.009 (0.001) ^{***}	age_{it}	0.058 (0.001) ^{***}
$2type_{it}$	0.511 (0.008) ^{***}	0.691 (0.009) ^{***}	age_{it}^2	-0.006 (0.001) ^{***}
$3type_{it}$	0.172 (0.011) ^{***}	0.155 (0.012) ^{***}	lnA_{it}	0.259 (0.002) ^{***}
$4type_{it}$	-0.103 (0.020) ^{***}	-0.233 (0.022) ^{***}	sex_{it}	0.010 (0.003) ^{***}
$\sum_{j=2}^{17} D_j region_{it}$	Y	Y	$2job_{it}$	0.022 (0.006) ^{***}
			$3job_{it}$	0.051 (0.006) ^{***}
			$4job_{it}$	-0.015 (0.007) ^{**}
			$5job_{it}$	-0.047 (0.006) ^{***}
<i>C</i>	1.143 (0.015) ^{***}	0.816 (0.016) ^{***}	<i>C</i>	1.686 (0.021) ^{***}
<i>Obs.</i>	64,193	64,193	<i>Obs.</i>	64,598
<i>R</i> ²	0.210	0.165	<i>R</i> ²	0.522

주: 1) 괄호 안은 standard error, ** 5% 수준, *** 1% 수준에서 각각 통계적 유의, *Obs.*(가구 수).

2) $lnhp_{it}$ (주택 구매가격의 로그값), $lnrp_{it}$ (주택 구입대가격의 로그값), $area_{it}$ (거주 주택면적), $2type_{it}$ (아파트), $3type_{it}$ (연립/다세대주택/빌라), $4type_{it}$ (주거용 오피스텔), lny_{it} (현재 소득의 로그값), age_{it} (가구주 연령), age_{it}^2 (연령의 제곱값), lnA_{it} (총자산의 로그값), sex_{it} (성별, 남=1, 여=0), $2job_{it}$ (관리, 전문직), $3job_{it}$ (사무직), $4job_{it}$ (판매직), $5job_{it}$ (기타 직종).

3) *region*은 17개 광역도에 대한 더미로서 계수추정 결과는 모든 행정구역에서 90% 신뢰 하에서 통계적으로 유의함.

4) lnA_{it} (총자산의 로그값)과 lny_{it} (소득) 간의 다중공선성 검증 결과(VIF)는 1.36으로서 다중공선성이 존재하지 않음.

VIF, variance inflation factor.

스텔의 경우 전세보다는 월세가 많이 통용되기 때문인 것으로 판단된다.

소득요인의 경우, 가구주의 연령(age_{it})은 현재 소득(lny_{it})과 (+)의 유의성을 보이고 있는 반면, 연령의 제곱값(age_{it}^2)은 (-)를 보이고 있다. 이는 가구주의 연령이 증가할수록 소득이 증가하지만,

일정 연령이 지나가면 소득은 감소한다는 의미이다. 성별(sex_{it})의 경우 남성 가구주의 소득이 상대적으로 높은 것으로 나타났으며, 직업의 경우에는 생산직($1job_{it}$)에 비하여 관리 및 전문직($2job_{it}$)과 사무직($3job_{it}$)의 소득은 높은 반면, 판매직($4job_{it}$)과 기타 직종($5job_{it}$)의 소득은 낮은 것으로 나타났다.

이와 같은 결과는 항상소득을 추정함에 있어 가구주의 교육수준을 고려한 김순용·박현수(2015), Chen & Jin(2014) 등의 연구결과와 비교하여 합리적인 것으로 판단된다. 즉 고학력인 가구주의 소득이 상대적으로 높게 나타난 기존의 연구결과와 비교하여 생산직 대비 관리 및 전문직, 그리고 사무직의 소득이 높게 나타나며, 판매직 등의 소득은 상대적으로 낮게 나타난 본 연구의 결과는 가구주의 교육수준을 묻지 않는 최근의 응답구조를 감안할 때, 항상소득의 추정에 무리가 없는 것으로 판단된다.

다음의 <표 4>에는 이상의 분석결과를 바탕으로 산출된 주택의 상대가격과 가구주의 항상소득이 나타나 있다.

가구가 현재 거주하고 있는 주택(동일 주택)에 대한 추정 구매가격과 임대가격은 각각 1억 9,710만 원과 1억 498만 원으로 나타났으며, 이에 따라 해당 주택에 대한 상대가격은 2.113 수준으로 나타났다. 즉 동일한 주택에 대한 임대가격 대비 구매가격은 약 2배 수준인 것이다. 그리고 산출된 가구주의 항상소득은 364만 원으로서 실제 월 평균 소득 397만 원보다는 낮은 것으로 나타났다.

이와 같이 추정된 가격과 소득요인을 활용하여 앞선 <식 2>를 이용, 차입가구의 특징을 살펴보았다. 다음의 <표 5>에는 이에 대한 분석결과가 나타나 있다.

전체 가구를 대상으로 분석한 결과, 금리(r_{it})와 가격요인(\hat{p}_{it})은 대출가구와 유의미한(-)관계가, 그리고 소득요인(\hat{y}_{it})은 유의미한(+)의 관계가 나타나고 있다. 즉 적용받는 금리가 상승하거나 주택 가격이 상승하면 대출을 받지 않을 가능성이 높으며, 소득이 증가하면 대출을 받을 가능성이 높음을 의미한다. 단 소득의 경우 계수 값이 크지 않다.

한편, 청년층 가구, 즉 전체 가구 대비 대출이 있는 청년층 가구를 대상으로 분석한 결과에서는 흥미로운 점이 발견된다. 전체 가구의 경우와는 달리, 금리와 소득요인의 유의성이 반대로 나타나고 있는 것이다. 즉 청년층 가구의 경우 금리가 상승해도 대출을 받을 가능성이 높으며, 소득이 증가하면 대출을 받지 않을 가능성이 높은 것으로 해석된다. 또한 주택가격과 대출과의(-)의 유의성의 정도도

<표 4> 상대가격과 항상소득의 추정결과

주택의 구매가격	주택의 임대가격	상대가격	실제소득	항상소득
19,710.7	10,498	2.113	396.9	364.0

주: 단위: 만 원. 2011년~2023년의 평균 수치임.

〈표 5〉 대출 가구의 특성(Probit 모형)

Classification	$\Lambda\left(\frac{P(Z_{it}, \alpha_i)}{1 - P(Z_{it}, \alpha_i)}\right)$	
	전체 가구 중 대출 가구	전체 가구 중 청년층 대출 가구
r_{it}	-0.062 (0.005) ^{***}	0.071 (0.006) ^{***}
\hat{p}_{it}	-0.041 (0.005) ^{***}	-0.017 (0.006) ^{***}
\hat{y}_{it}	0.002 (0.001) ^{***}	-0.011 (0.004) ^{***}
C	-1.095 (0.028) ^{***}	-0.632 (0.030) ^{***}
<i>Obs.</i>	65,015	65,015 ^a
$LR \chi^2$	-39,733.36	-32,833.85
$Prob > \chi^2$	0.000	0.000

주: 1) 2011~2023년의 전체 연도에 대한 분석결과임.

2) 괄호 안은 standard error, *** 1% 수준에서 각각 통계적 유의, *Obs.*(가구 수).

3) $\Lambda\left(\frac{P(Z_{it}, \alpha_i)}{1 - P(Z_{it}, \alpha_i)}\right)$ (현재 대출가구=1, 아니면 0), r_{it} (차입금리), \hat{p}_{it} (현재 거주 주택의 상
대가격), \hat{y}_{it} (항상소득).

4) ^a 전체 가구의 개체 수에서 청년층 가구의 특성을 나타내는 것이므로 전체 가구수와 표본은
동일함.

전체 가구에 비하면 상대적으로 크지 않은 것도 특징이다.

다음으로 대출 가구 중 상환부담을 느끼는 가구와 부담을 느끼지 않는 가구의 특성을 비교분석한다.
이를 위하여 가구가 응답한 적정 수준의 상환금을 실제 상환금과 비교, 상환부담을 느끼는 가구와
그렇지 않은 가구를 각각 추출한다.

이때, 가구의 적정 원리금 수준은 해당 가구의 응답에 따라 결정되는 구조로서 자기보고 편향(bias)
으로 인한 측정 편향(measurement bias)의 가능성이 존재한다. 그러나 ‘주택금융 및 보금자리론
실태조사’의 적정 수준의 상환금에 대한 응답가구는 실제 매달 대출 원리금을 상환하는 가구로서
소득 대비 원리금 상환에 대한 부담의 정도를 1에서 5까지의 단계로 먼저 응답한다. 그리고 실제
소득과 가계지출을 감안하여 차주 자신이 무리가 없는 수준의 원리금이라고 판단하는 금액을 밝히는
구조이다. 이에 따라 차주가 밝힌 적정 원리금 상환수준은 자신이 감내할 수 있는(tolerable) 상환수

준으로 판단할 수 있으며, 측정 편향의 가능성은 낮다고 할 수 있다.

다음의 <표 6>에는 이와 같이 추출된 가구의 분포가 나타나 있다.

전체 분석대상 가구 65,056 가구 중 현재 대출을 받고 있는 가구는 23,614 가구이며, 이 중 적정 원리금 상환수준을 밝힌 가구는 18,250 가구로서 차입 가구의 77.3%이다. 이들 가구가 밝힌 월 적정 원리금 상환수준은 약 54.4만 원이며, 이들의 실제 원리금 상환수준은 약 59.5만 원으로서 적정 수준보다 더 많이 지출하고 있는 것으로 나타났다. 그리고 실제 원리금 상환수준이 적정 상환수준보다 큰 가구, 즉 상환부담이 있는 가구는 10,429 가구로서 전체 응답 가구의 57.1% 수준으로 나타났다.

적정 원리금 상환수준을 밝힌 18,250 가구 중 청년층 가구는 2,515 가구이며, 적정 원리금은 54.6만 원으로서 전체 가구와 유사하다. 그러나 이들 가구의 실제 원리금은 62.5만 원으로서 전체 가구보다 더 많은 상환을 하고 있는 것으로 나타났다. 이에 따라 상환부담이 있는 가구의 비중도 60.2%(1,513 가구)로 더 높게 나타났다.

이와 같이 확인된 상환부담 수준을 바탕으로 앞선 <식 5>의 다항프로빗 모형을 이용, 상환부담 가구의 특징을 분석해 보았다. 다음의 <표 7>에는 이에 대한 분석결과가 나타나 있다.

전체 가구를 대상으로 한 분석에서, 금리요인(r_{it})은 상환부담이 없는 가구와는 (-)의 유의한 관계가, 그리고 상환부담이 있는 가구와는 (+)의 유의미한 관계가 나타나고 있다. 즉 대출 시 적용받는 금리의 상승은 상환부담이 없는 가구가 될 가능성을 줄이는 반면, 상환부담이 있는 가구가 될 가능성을 높이는 것이다. 주택가격의 요인(\hat{p}_{it})의 경우에도 마찬가지이다. 상환부담이 없는 가구와 있는 가구에게서 계수 값의 부호가 서로 반대로 나타나고 있다. 소득요인(\hat{y}_{it})의 경우에는 상환부담이 있는 가구에게서만 (+)의 유의성이 나타나고 있으나, 그 값은 크지 않다.

청년층 가구를 대상으로 한 분석에서는 전체 가구를 대상으로 한 분석결과와는 달리, 대출이 없는 청년층 가구에 비하여 대출이 있는 가구의 금리, 가격, 그리고 소득요인에 대한 민감도가 더 높게 나타나고 있다. 금리의 상승의 경우, 상환부담이 없는 가구가 될 가능성을 줄이며, 상환부담이 있는 가구가 될 가능성을 높이고 있다. 그리고 그 정도는 전체 가구의 경우에 비하여 더 강하다. 가격 요인의 경우에는 전체 가구의 경우와는 달리 상환부담이 있는 가구의 경우에서만 유의성이 나타난다. 그리고

<표 6> 상환 부담 가구의 분포

	실제 원리금(만 원)	적정원리금(만 원)	상환부담 가구 수
전체 가구(18,250)	59.5	54.4	10,429(57.1%)
청년층 가구(2,515)	62.5	54.6	1,513(60.2%)

주: 2011년~2023년의 평균 수치임.

〈표 7〉 가구의 대출상환 부담에 대한 결정요인: 다항프로빗 모형(multinomial Probit)

<i>Dep V.</i>	$\Lambda\left(= \frac{P(Z_{it}, \alpha_i)}{1 - P(Z_{it}, \alpha_i)}\right)$			
Basement	대출이 없는 가구 대비		대출이 없는 청년 가구 대비	
Classification	상환부담 없는 가구	상환부담 가구	상환부담 없는 가구	상환부담 가구
r_{it}	-0.061 (0.010) ^{***}	0.086 (0.009) ^{***}	-0.133 (0.027) ^{***}	0.150 (0.025) ^{***}
\hat{p}_{it}	-0.078 (0.010) ^{***}	0.061 (0.008) ^{***}	-0.006 (0.028)	0.199 (0.023) ^{***}
\hat{y}_{it}	0.004 (0.001) ^{***}	0.004 (0.001)	0.005 (0.002) ^{***}	-0.015 (0.002) ^{***}
C	-2.717 (0.051) ^{***}	-2.769 (0.049) ^{***}	-1.043 (0.161) ^{***}	-2.762 (0.145) ^{***}
<i>Obs.</i>	65,015		13,597	
$LR \chi^2$	-47,187.65		-7,347.02	
$Prob > \chi^2$	0.000		0.000	

주: 1) 2011~2023년의 전체 연도에 대한 분석결과임.

2) 괄호 안은 standard error, *** 1% 수준에서 각각 통계적 유의, *Obs.*(가구 수).

3) r_{it} (차입금리), \hat{p}_{it} (현재 거주 주택의 상대가격), \hat{y}_{it} (항상소득).

주택가격의 상승이 상환부담이 있는 가구가 될 가능성을 높이는 정도도 더 강한 것으로 나타난다. 소득의 경우에도 소득의 상승이 상환부담이 있는 가구가 될 가능성을 낮추는 모습이 발견되는 것도 청년층 가구의 특징이다.

이와 같은 분석결과는 채무부담이 더 강한 가구일수록 정책당국의 정책변화에 더 민감하게 반응하게 된다는 Gelos et al.(2004)의 연구결과와 비교하여, 위험가구일수록 금리와 주택가격 변화라는 보다 구체적인 요인변화에 민감하게 반응하게 됨을 시사한다. 특히 청년층에서 더 유의하게 나타나고 있는 분석결과는 Quericia et al.(2003)에서 지목한 금융제약계층이 주택시장에 신규로 진입하는 청년층 가구와 관계가 있음을 시사한다.

2. 가구의 추가적인 부채의 결정요인

다음으로 가구의 부채의 변화가 이와 같은 상환부담에 어떠한 영향을 미치는지를 파악하기 위하여

앞선 <식 6-1>과 <식 6-2>를 이용, 가구의 추가적인 부채를 추정해 보았다. 다음의 <표 8>에는 이에 대한 분석결과가 나타나 있다.

가구의 소득(Y_{it})과 자산(A_{it}), 그리고 직업(job_{it})은 모두 부채(D_{it})에 유의미한 (+)의 관계를 보이고 있다. 연령(age_{it})은 (-)의 관계를 나타내고 있다. 연령효과(age_{it}^2)의 유의성은 나타나지 않는다. 즉 특정 연령이 지나면 부채가 감소하는 현상 발생하지 않는 것이다.

앞선 <식 6-2>를 이용하여 추정된 부채의 잔차, 즉 가구의 추가적인 부채의 경우 대출이 없는 가구는 (-)를 보이며 추가적인 부채가 발생하지 않는다. 그러나 대출을 받은 가구의 경우 모두 추가적인 부채가 발생하는 것으로 나타났다. 특히 상환부담이 있는 가구의 경우 0.275로서 가장 높은 값을 보이고 있다.

다음의 <표 9>에는 앞선 <식 7>을 이용, 추출된 추가적인 부채가 상환부담 가구 확률에 어떠한 영향을 미치게 되는지 분석한 결과가 나타나 있다.

전체 65,056 가구 중 부채가 확인된 가구를 대상으로 분석한 결과, 가구의 부채총액(D_{it})의 증가는 원리금 상환의 부담가구가 될 확률과 유의미한 관계가 나타나지 않고 있다. 그러나 부채잔차인 추가적인 부채(\hat{D}_{it})의 증가는 상환부담 가구의 확률과 유의미한 (+)의 관계가 나타나고 있다. 즉 추가적인 부채가 증가할수록 상환부담 가구가 될 확률이 높아지는 것이다.

청년층 가구에서는 보다 흥미로운 점이 발견된다. 청년층 가구의 경우 부채총액(D_{it})과 상환부담 가구의 확률 간에서도 유의미한 (+)의 관계가 발견되며, 추가적인 부채(\hat{D}_{it})와의 관계에서도 (+)의

<표 8> 부채의 결정요인 회귀분석 결과

$$D_{it} = (3.966) + (0.346)A_{it} + (0.266)\hat{y}_{it} + (0.107)job_{it} + (-0.016)age_{it} + (-0.000)age_{it}^2 + \epsilon_{it}$$

(0.004)*** (0.007)*** (0.107)*** (0.002)*** (0.000)

$Obs. = 47,117$
 $R^2 = 0.337$
 $VIF = A_{it} (1.82), \hat{y}_{it} (1.77)$

<추가적인 부채>		
무대출 가구	상환부담 없는 가구	상환부담 있는 가구
-0.132	0.133	0.275

주: 괄호 안은 standard error, *** 1% 수준에서 각각 통계적 유의, $Obs.$ (가구 수), VIF (자산(A))과 향상소득(\hat{y})의 다중공선성 검증결과.

VIF, variance inflation factor.

〈표 9〉 가구의 부채와 상환부담 가구 특성과의 관계

Dep V.	$\Lambda\left(=\frac{P(Z_{it},\alpha_i)}{1-P(Z_{it},\alpha_i)}\right)$			
	전체 가구		청년층 가구	
Classification				
Method	부채 총액	부채 잔차	부채 총액	부채 잔차
D_{it}	0.061 (0.010)		0.155 (0.028) ^{***}	
\hat{D}_{it}		0.251 (0.010) ^{***}		0.553 (0.028) ^{***}
r_{it}	0.141 (0.007) ^{***}	0.106 (0.007) ^{***}	0.111 (0.019) ^{***}	0.081 (0.019) ^{***}
\hat{p}_{it}	0.015 (0.006) ^{***}	0.007 (0.006)	0.017 (0.016) ^{***}	0.056 (0.016) ^{***}
\hat{y}_{it}	0.001 (0.001) ^{***}	0.002 (0.001) ^{***}	-0.002 (0.001) ^{***}	-0.003 (0.001) ^{***}
C	-1.594 (0.087) ^{***}	-1.412 (0.041) ^{***}	-2.244 (0.245) ^{***}	-1.259 (0.114) ^{***}
$Obs.$	47,388	47,144	10,151	10,495
$LR\chi^2$	-22,369.24	-22,432.41	-3,674.69	-3,754.14
$Prob > \chi^2$	0.000	0.000	0.000	0.000

주: 1) 2011~2023년의 전체 연도에 대한 분석결과임.

2) 괄호 안은 standard error, *** 1% 수준에서 각각 통계적 유의, $Obs.$ (가구 수).

3) $\Lambda\left(=\frac{P(Z_{it},\alpha_i)}{1-P(Z_{it},\alpha_i)}\right)$ (상환부담 가구=1, 그렇지 않으면 0), D_{it} (부채 총액의 로그값), \hat{D}_{it} (부채 잔차), r_{it} (차입금리), \hat{p}_{it} (현재 거주 주택의 상대가격), \hat{y}_{it} (항상소득).

관계가 나타나고 있다. 그리고 그 정도는 전체 가구에 비하여 더 강하다. 즉 청년층 가구의 경우, 부채의 증가는 상환부담 가구가 될 가능성을 높이며, 특히 항상소득 등으로 설명할 수 없는 부채인 추가적인 부채의 발생은 상환부담 가구가 될 가능성을 더 크게 만드는 것으로 해석할 수 있다.

이와 같은 결과는 기존의 연구들에 비하여 보다 분명한 판단을 가능하게 한다. 심승규(2022)에 의하면 마이크로자료의 분석에서, 설명할 수 없는 부채의 증가는 자가 가구가 될 확률을 낮추는 것으로 나타났으며, 40세 이하 가구주에게서 그와 같은 경향은 더 강하게 나타났다. 이에 비하여 본 연구의 분석에서는 40세 이하 가구주(평균 연령 31.6세)에게서 설명할 수 없는 부채의 증가는 원리금 상환부

답이 가중되는 가구가 될 확률이 높아지는 결과를 보이는 것으로 나타났다. 또한 2020년 한 해 동안의 자료 분석을 통하여 자산 잔차나 소득 잔차가 자가 가구가 될 확률을 분석한 이준민 외(2022)의 연구결과와 비교하여 본 연구는 13년 동안의 마이크로 자료의 분석으로부터 부채잔차의 영향을 판단할 수 있다는 점에서 의미가 있다.

마지막 단계는 가구의 추가적인 부채가 발생하는 경로를 분석하는 것이다. 즉 가구의 과도한 대출이 추가적인 부채에 어떠한 영향을 미치게 되는지를 분석하는 것이다. 이를 위하여, 먼저, 앞선 <식 8>을 이용, 가구의 한계대출 수준을 추정한다. 다음의 <표 10>에는 이에 대한 추정결과가 나타나 있다.

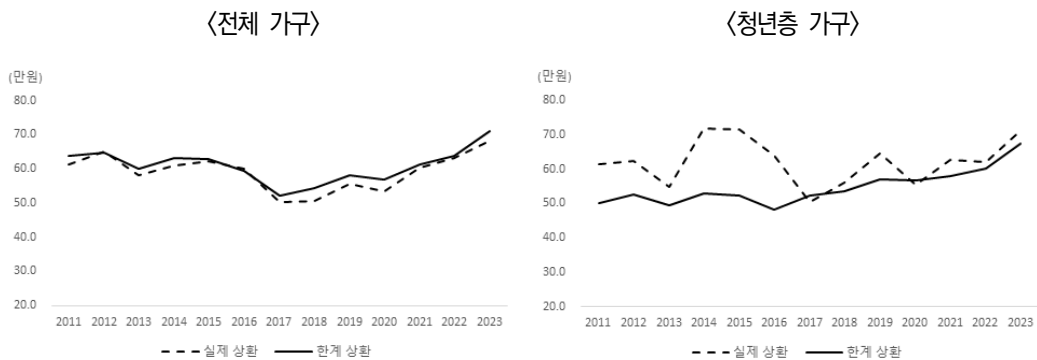
정책당국의 주택시장에 대한 금융규제의 정도, 즉 DTI를 각 가구별로 적용하여 산출된 한계 원리금 상환금액은 전체 차입가구의 경우 월 평균 61만 원 수준으로 나타났다. 그리고 이들 가구의 실제 원리금 상환수준은 약 59.5만 원으로서 한계수준에 거의 근접하는 것으로 나타났다. 청년층 가구의 경우, 산출된 한계 상환금액은 월 평균 54.7만 원으로서 전체 가구보다 더 낮은 것으로 나타났다. 앞서 확인하였듯이, 상대적으로 소득에 비하여 대출규모가 큰 것이기 때문인 것으로 판단된다. 실제 청년층 가구의 월 평균 상환금액은 62.5만 원으로서 전체 가구보다 더 높으며, 이들 가구들은 한계 상환금액을 크게 상회하는 것으로 나타났다.

다음의 <그림 1>에는 추정된 한계 상환금의 추이가 나타나 있다.

<표 10> 한계 상환 원리금 추정결과

	전체 가구		청년층 가구	
	실제 상환금액	한계 상환금액	실제 상환금액	한계 상환금액
단위: 만 원	59.5	61.0	62.5	54.7

주: 2011년~2023년의 평균 수치임.



<그림 1> 한계 상환 원리금의 추이

전체 가구의 경우 추정된 한계 상환금액은 실제 상환금액의 추이와 유사하다. 규제가 강화되었던 2017년 이후에는 실제 대출금액이 한계 대출금액에 미치지 못하는 것을 알 수 있다. 이와는 달리, 청년층 가구의 경우 거의 모든 연도에서 실제 대출금액은 한계 대출금액을 상회하는 것으로 나타난다. 특히 규제가 완화되었던 2014년부터 2016년 동안에는 실제 대출금액이 한계 대출금액을 크게 상회하는 것으로 나타난다. 청년층 가구의 경우 규제가 완화되면 과도한 차입을 시도할 유인이 강하게 나타나는 것으로 해석할 수 있다.

이와 같은 추정결과는 DTI를 직접 적용하면 가구의 원리금 상환부담이 더 가중된다는 의미로서 LTV나 DTI 규제가 가구의 특성 변수에 미치는 영향의 분석을 통하여 규제의 효과를 판단한 McDonald(2015), Robinson & Yao(2015) 등의 연구와 비교하여 가중되는 상환부담의 정도를 가구의 특성별로 직접 추정하였다는 점에서 의미가 있다.

이와 같이 추정된 한계대출, 즉 한계 상환 원리금을 이용, 앞선 <식 9>와 같이 가구의 과도한 대출이 해당 가구의 부채에 어떠한 영향을 미치게 되는지 분석한다. 다음의 <표 11>에는 이에 대한 분석결과가 나타나 있다.

<표 11> 초과대출과 추가적인 부채와의 관계

Dep V.	D_{it} (부채 총액)		\hat{D}_{it} (부채 잔차)	
	전체 대출 가구	청년층 대출 가구	전체 대출 가구	청년층 대출 가구
Ol_{it}	0.034 (0.005)*	0.050 (0.015)***	0.198 (0.005)***	0.341 (0.014)***
r_{it}	0.011 (0.004)*	-0.020 (0.012)	0.070 (0.004)***	0.086 (0.011)***
\hat{p}_{it}	0.095 (0.004)***	0.111 (0.010)***	0.078 (0.003)***	0.061 (0.009)***
\hat{y}_{it}	0.031 (0.001)**	0.020 (0.001)***	0.019 (0.004)	-0.017 (0.001)
C	0.726 (0.027)***	0.796 (0.083)***	0.248 (0.026)***	0.427 (0.076)***
Obs.	17,504	2,437	17,504	2,437
R^2	0.280	0.174	0.090	0.097

주: 1) 2011~2023년의 전체 연도에 대한 분석결과임.

2) 괄호 안은 standard error, * 10% 수준, ** 5% 수준, *** 1% 수준에서 각각 통계적 유의, Obs.(가구 수).

3) D_{it} (부채 총액의 로그값), \hat{D}_{it} (부채 잔차), r_{it} (차입금리), \hat{p}_{it} (현재 거주 주택의 상대가격), \hat{y}_{it} (항상소득).

한계대출 수준을 상회하는 대출수준, 즉 가구의 과도한 차입(Ol_{it})은 특히 청년층 가구의 부채총액에 유의미한(+)의 영향을 미친다. 전체 대출 가구에서는 유의성이 강하지 않다. 이에 비하여 과도한 차입은 전체 가구와 청년층 가구 모두에서 부채잔차, 즉 추가적인 부채에(+)의 강한 유의성을 가지고 있다. 즉 전체 대출 가구에서도 과도한 대출은 소득 등으로 설명할 수 없는 추가적인 부채를 증가시키며, 특히 과도한 대출은 청년층 대출 가구의 추가적인 부채를 크게 증가시키는 것으로 나타나고 있다.

한편, 주요 변수 중에서는 금리의 인상이 모든 대출 가구에게서 추가적인 부채를 증가시키는 것으로 나타났으며, 이와는 반대로 항상소득의 경우 부채총액과는(+)의 유의성이 나타난 반면, 부채잔차와는 유의성이 없는 것으로 나타났다.

항상소득과 부채 간의 이와 같은 결과는 소득과 부채간의(+) 상관관계가 가질 수 있는 함의를 증명하고 있다는 점에서 의미가 있다. 심승규(2022)에 의하면, 소득과 부채간의(+)의 관계는 소득이 많은 가구가 상환능력을 인정받아 더 많은 차입을 할 수 있음을 의미한다. 그리고 차입을 통하여 자산을 취득한 가구는 그 자산으로부터 자산소득을 얻게 되어 소득이 더 증가하게 된다. 이는 주택점유 형태에 대한 소득효과와 부채효과의 구분을 어렵게 만들게 된다.

이와 같은 기존 연구를 감안하면 항상소득과 주요 변수를 이용해 설명되지 않은 추가적인 부채에 대한 분석의 필요성이 제기된다. 즉 위의 연구결과에서 전체 가구와 청년층 가구 모두에게서 항상소득과 부채총액 간에는 유의미한(+)의 관계가 발견된 반면, 항상소득과 부채잔차 간에는 유의성이 발견되지 않는 것은 이와 같은 점을 시사하고 있다고 할 수 있다.

V. 결론

본 연구는 2011년부터 2023년까지의 13년 동안의 주택금융공사의 '주택금융 및 보금자리론 실태조사'의 65,403 가구(일반가구)를 대상으로, 가구의 원리금 상환부담과 추가적인 부채와의 관계, 그리고 한계대출을 이용한 과도한 대출수준의 영향에 대하여 분석하였다.

본 연구의 실증분석 결과는 다음과 같다.

첫째, 가구가 대출을 시도할 때, 적용받는 금리가 상승하거나 주택가격이 상승하면 대출을 받지 않을 가능성이 높은 것으로 나타났다. 그러나 청년층 가구의 경우 전체 가구의 경우와는 달리, 금리가 상승해도 대출을 받을 가능성이 높은 것으로 나타났다.

둘째, 가구의 대출상환 부담과 관련하여, 대출 시 적용받는 금리의 상승은 상환부담이 없는 가구가 될 가능성을 줄이는 반면, 상환부담이 있는 가구가 될 가능성을 높이는 것으로 나타났다. 주택가격의 상승요인의 경우에도 동일하였다. 청년층 가구의 경우 가구의 금리, 가격, 그리고 소득요인에 대한

민감도가 더 높게 나타났다.

셋째, 가구의 부채총액의 증가는 원리금 상환가구가 될 확률과 유의미한 관계가 나타나지 않은 반면, 부채의 잔차, 즉 추가적인 부채가 증가할수록 상환부담 가구가 될 확률이 높아지는 것으로 나타났다. 그리고 이러한 경향은 청년층 가구에서 더 높게 나타났다.

넷째, 가구는 자신의 한계대출 수준에 근접하여 실제대출을 받는 것으로 나타났으며, 특히 청년층의 경우에는 한계수준을 상회하는 대출을 받고 있었다. 이와 같은 가구의 과도한 대출은 소득 등으로 설명할 수 없는 추가적인 부채를 증가시키며, 특히 청년층 대출 가구의 추가적인 부채를 더 크게 증가시키는 것으로 나타났다.

이와 같은 분석결과를 종합해 볼 때, 가구의 추가적인 부채는 대출을 차입한 가구에게서 발생할 가능성이 높으며, 추가적인 부채의 발생은 해당 가구의 대출 상환에 대한 부담을 더 높이는 것으로 판단된다. 그리고 청년층 가구의 경우 상대적으로 소득이 낮은 반면, 주택시장으로의 신규진입을 위하여 더 많은 차입을 시도하고 있기 때문에 추가적인 부채의 발생확률이 더 높으며, 이에 따라 더 높은 상환부담에 직면할 가능성이 높은 것으로 판단된다.

본 연구의 분석결과는 기존의 연구와 비교하여, DTI 규제와 같은 정책규제로 인하여 가중되는 가구의 상환부담의 정도를 가구별로 직접 추정하였다는 점과, 정책당국의 정책변화에 대한 가구의 민감도를 주택시장에 신규로 진입하는 청년층에 초점을 맞추었다는 점, 그리고 과도한 대출로 인한 가구의 상환부담의 증가를 향상소득 등으로 설명하지 못하는 추가적인 부채를 이용하여 장기간에 걸쳐 설명하고 있다는 점에서 학술적 의미가 있다.

2017년부터 강화되었던 주택시장에 대한 정책당국의 금융규제는 2022년 말과 2023년에 완화되었으며, 2025년부터 다시 강화되고 있다. 이러한 가운데 지난 13년 동안 주택시장에서 한계대출 이상의 대출을 차입하고 있는 가구의 비중도 높아지고 있다. 이에 따라 정책당국의 가계부채에 대한 관리와 접근은 가구의 적정수준의 상환규모의 파악에서부터 해당 가구의 상환부담을 인지하여, 가계 신용의 연체의 가능성을 사전에 적극적으로 관리할 필요성이 있다고 판단된다.

참고문헌

- 김수현, 황설웅. (2023). *우리나라의 가계부채와 소득불평등*. 서울: 한국은행.
- 김순용, 박현수. (2015). 소득 및 가격 탄력성을 이용한 지역 및 소득계층별 주택수요에 관한 연구. *서울도시연구*, 16(2), 71-86.
- 김종희. (2019). 금융규제가 차주의 주택구매 여력에 미치는 영향에 대한 연구. *경영경제연구*, 41(1), 111-137.
- 변창흠, 이희정. (2002). *서울시 주택수요 및 공급능력 추정에 관한 기초연구*. 서울: 서울시정개발연구원.
- 심승규. (2022). 연령대별 주택 점유형태 분석: 소득효과와 부채효과를 중심으로. *주택금융연구*, 4(1), 5-28.
- 윤병우, 최경욱. (2021). 부채상환에 대한 잠재 위험가구의 특성 분석. *시장경제연구*, 50(1), 41-69.
- 윤주현, 김혜승. (2000). 주택수요구조분석 및 전망에 관한 연구. *국토연구*, 29, 51-65.
- 이동규, 서인석, 박형준. (2009). 주택정책 규제수단으로서 DTI 규제정책의 효과 분석: System Dynamics를 활용한 시뮬레이션 분석. *한국정책학회보*, 18(4), 207-241.
- 이준민, 홍지훈, 심승규. (2022). 소득 및 자산에 따른 주거면적별 주택점유형태 분석. *부동산분석*, 8(1), 41-60.
- 전병욱, 박나라. (2022). 가계부채 및 세금부담이 무주택 가구의 주택 매수 의사결정에 미친 영향. *주택도시금융연구*, 7(1), 5-27.
- Chen, J., & Jin, M. (2014). Income elasticity of housing demand in China: Micro-data evidence from Shanghai. *Journal of Contemporary China*, 23(85), 68-84.
- Crowe, C., Dell'Arriccia, G., Igan, D., & Rabanal, P. (2011) *How to deal with real estate booms: Lessons from country experiences* (IMF Working Paper No. WP/11/91). Washington, DC: International Monetary Fund.
- Ermisch, J. F., Findlay, J., & Gibb, K. (1996) The price elasticity of housing demand in Britain: Issues of sample selection. *Journal of Housing Economics*, 5(1), 64-86.
- Gelos, G., Grinberg, F., Khan, S., Mancini-Griffoli, T., Narita, M., & Rawat, U. (2024). Has higher household indebtedness weakened monetary policy transmission? *IMF CEPR*, 20(1), 325-371.
- Hatchondo, J. C., Martinez, L., & Sánchez, J. M. (2014). *Mortgage defaults and prudential*

- regulations in a standard incomplete markets model* (Working Paper No. 2011-019A). St. Louis, MO: Federal Reserve Bank of St. Louis.
- Hatchondo, J. C., Marinez, L., & Sanchez, J. M. (2015). Mortgage defaults. *Journal of Monetary Economics*, 76, 173-190.
- Kuttner, K. N., & Shim, I. (2013). *Can non-interest rate policies stabilise housing markets? Evidence from a panel of 57 economies* (BIS Working Paper No. 433). Basel, Switzerland: Bank for International Settlements.
- Linneman, P. D., Megbolugbe, I. F., Wachter, S. M., & Cho, M. (1997), Do borrowing constraints change U.S. homeownership rates?, *Journal of Housing Economics*, 6, 318-333.
- McDonald, C. (2015). *When is macroprudential policy effective?* (BIS Working Paper No. 496). Basel, Switzerland: Bank for International Settlements.
- Muellbauer, J. (2012) *When is a housing market overheated enough to threaten stability?* (Economics Series Working Paper No. 623). Oxford, UK: University of Oxford.
- Quericia, R., McCarthy, G., & Wachter, S. (2003), The impacts of affordable lending efforts on homeownership rates, *Journal of Housing Economics*, 12(1), 29-59.
- Robinson, T., & Yao, F. (2015). Consequences of loan-to-value ratio policies for business and credit cycles. Retrieved from https://www.bsp.gov.ph/Media_And_Research/Events/2016/irc/downloads/papers/2016_BSP-IRC_10_paper.pdf?utm_source=chatgpt.com
- Tse, R. Y. C., & Raftery, J. (1999). Income elasticity of housing consumption in Hong Kong: A cointegration approach. *Journal of Property Research*, 16(2), 123-138.
- Zhang, Y., Yuan, Y., & Zhao, G. (2025). An empirical study on the impact of household debt level on consumption behavior: Based on the mechanism of the housing provident fund system and risk preference. *Finance Research Letters*, 72, 106507.

(논문 접수일: 2026.03.30. 수정논문 접수일: 2026.05.04. 논문 채택일: 2026.06.05.)

The Relationship between Repayment Burden and Additional Debt in Households

Jong-Hee Kim*

Abstract

This study analyzes the impact of excessive loans on the relationship between repayment burden and additional debt in households by focusing on marginal loans in the housing market. The following conclusions are drawn from the results of the empirical analysis. First, the rising interest rate decreases the incentive for households to borrow, except for the young households, despite rising rates. Second, the additional debt, which implies unexplainable debt from households' permanent income leads to an increase in the repayment burden, and this tendency is much stronger among young households. Finally, households' real loans are very close to their marginal loans; that is, their loan limit. Specifically, for young households, real loans exceeded marginal loans. Moreover, excessive loans increase the repayment burden by increasing household loans. Specifically, this tendency is much stronger in young households.

Keywords : Burden of Repayment, Financial Regulation, Additional Debt, Marginal Loan

* Jong-Hee Kim, Corresponding author, Professor, Department of Economics, Jeonbuk National University, jonghk5@jbnu.ac.kr

© Copyright 2026 Housing Finance Research Institute. 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.

트리 앙상블 기반 자동산정모형을 이용한 재고 주택가격 분포 관측: 전국 아파트 단지 전수 자료의 시계열 분석

홍정의*

요약

본 연구는 부스팅 트리 알고리즘(XGBoost)에 기반한 자동산정모형을 활용하여 추정 기반 재고 주택가격 관측을 시도한다. 본 연구는 2012년 3월 1일부터 2024년 4월 1일까지의 아파트 실거래 자료 6,118,730건과 전국 12,517,131호(주택 총조사 기준 전체 아파트 단지의 99.1%)의 개별 단지(평형별) 정보를 활용하여, 지역별 평균 가격 수준, 가격 분포, 소득 대비 주택가격 비율(price to income ratio, PIR) 등을 도출하였다. 결과는 평균 아파트 가격이 2017~2022년 가격상승기에 기존 공표가격지수에 비해 더 빨리 상승했다는 것과, 분석 기간 동안 지역 간 가격 격차와 지역 내 가격 분위수 간 격차가 동시에 확대되었음을 보여주었다.

핵심어 : XGBoost, 기계학습, 자동산정모형(Automated Valuation Model, AVM), 주택가격분포, PIR(Price to Income Ratio)

1. 서론

주택 시장에 대한 분석에서 가장 핵심이 되는 정보는 아마도 주택의 가격일 것이다. 하지만 주택 가격은 주택이라는 자산의 특성상 상당히 제한적인 형태로만 관측이 가능했다. 주식이나 채권과 같은 금융자산은 기본적으로 동질적이며 따라서 그에 대한 거래가 각 시점마다 빈번하게 발생하므로,

* 홍정의, 주저자, 한동대학교 경영경제학부 부교수, hwgh024@handong.edu

© Copyright 2026 Housing Finance Research Institute. 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.

개별 자산 가격의 변동성이 시계열 상 명확히 관측된다. 한편 모든 주택은 본질적으로 비동질적이며 개별 자산의 거래 빈도는 매우 낮다. 즉, 주택은 위치, 구조, 주변 환경과 같은 많은 특성들이 결합하여 개별 주택의 가격을 형성하며, 특정 시점에 거래가 이루어지는 주택은 전체 재고 주택의 극히 일부에 불과하다.

주택 시장의 이러한 특징이 "올해 서울의 평균 아파트 가격은 얼마인가?"와 같은 일견 평범해 보이는 질문에 대해서도 그 답변을 어렵게 만드는 이유가 된다. 만약 단순히 올해 거래된 아파트 가격의 평균을 답했다면, 그것은 올해 많이 거래된 아파트의 평균적 가격일 뿐 존재하는 전체 아파트의 실제 평균 가격이 아닐 것이다. 경제적 환경에 따라 많이 거래되는 아파트의 종류가 상당히 달라질 수 있음을 감안하면, 그러한 대답이 일반적으로 믿을만한 답변이 아닌 것은 이미 잘 알려져 있다. 이처럼, 주택 시장에서는 전체 자산 중 일부의 가격만이 관측되는 한계가 존재하며, 이를 극복하고 시장 전체의 가격 수준과 추이를 파악하기 위해서는 전통적으로 주택가격지수를 이용해왔다.

그러나 기존의 주택가격지수 역시 그 구성 방식에 내재하는 근본적 한계를 안고 있다. 가장 널리 사용되는 두 가지 접근인 반복매매지수(repeat-sales index)와 헤도닉 가격지수(hedonic price index)는 각각 다른 방식으로 주택의 질적 차이를 통제하고자 하지만, 양자 모두 해당 시점 거래된 주택만을 대상으로 한다는 공통적 한계를 갖는다. 반복매매지수는 동일 주택이 두 번 이상 거래된 경우의 가격 차이만을 이용하여 지수를 산출하므로 표본선택편의(sample selection bias)에 노출되며(Gatzlaff & Haurin, 1997), 헤도닉 가격지수는 함수 형태의 설정오류(functional form misspecification) 문제가 지적되어 왔다(Hill, 2013).

이보다 본질적으로, 주택가격지수는 모두 본질적으로 가격의 상대적 변화만을 측정하며, 가격의 절대적 수준에 대한 정보를 직접 제공하지 못한다. 예를 들어, 특정 연도 서울 지역의 주택가격지수가 110이라는 것은 기준 시점 대비 10% 상승했음을 알려줄 뿐, 실제 평균 가격의 수준 자체가 얼마인지에 대해서는 아무런 정보도 담고 있지 않다. 또한, 지수의 산출 과정에서 사용되는 표본이 고정되어 있으므로, 연구자가 원하는 기준에 따라 표본을 자유롭게 분할하여 하위 집단별 가격 추이나 가격 분포를 관측하는 것이 구조적으로 어렵다. 결과적으로, 기존의 주택가격지수는 시장의 가격 변화 방향과 크기에 대한 중요한 정보를 제공하지만, 시장 전체의 가격 수준, 분포 구조, 가격 불균등성, 소득 대비 가격 부담 등을 입체적으로 파악하는 데에는 근본적인 한계가 있었다.

만약 우리가 개별 주택들의 가격을 일정 수준 이상 정확하게 추정할 수 있다면, 이러한 한계를 상당 부분 극복할 수 있을 것이다. 특히, 거래된 주택뿐 아니라 거래되지 않은 주택을 포함한 전체 재고 주택의 가격을 추정할 수 있다면, 이는 거래 표본에 의존하는 전통적 가격 측정 방식을 보다 확장하는 관측 정보를 제공할 수 있다. 본 연구는 이러한 접근을 추정 기반 재고 주택 시장 관측으로 개념화한다.

이러한 전수적 가격 추정의 가능성은 대규모 주택가격 추정을 위한 자동산정모형(automated valuation model, AVМ)의 효율성 확장과 밀접한 관계가 있다. 주택가격 자동산정모형에 대한 논의는 꽤 오랜 역사를 가지고 있다(IAAO, 2018; Tekouabou et al., 2024). 그럼에도 불구하고 시장 상황을 파악하는 데에 있어서 자동산정모형은 지수산출과 같은 영역에서의 활용에는 비교적 관심을 받지 못해 왔다. 그 주요한 이유는 회귀 모형에 기반한 전통적 자동산정모형은 개별 자산 단위에서 그 예측의 정확성이 충분히 높지 않다는 근본적인 한계를 지니고 있었기 때문이다(Hill, 2013; Malpezzi, 2003). 이는 비록 전통적 모형이 이론적 기반과 결과의 직관적 해석 가능성을 제공함에도 불구하고, 직관적 함수 형태를 가정하는 모형의 특성 상, 현실의 복잡성을 충분히 포착하기 어려울 뿐 아니라, 모형의 설정 오류(misspecification)나 변수 간 내생성 문제에서도 자유롭지 못하다는 것과 관련되어 있다.

그러나 최근 들어, 기계학습 기반의 자동산정모형이 주택 가격 평가의 정확성과 안정성을 획기적으로 향상시킬 수 있다는 사실이 보고되고 있다(이인호 · 이경섭, 2020; 홍정의, 2020; Antipov & Pokryshevskaya, 2012; Hong & Kim, 2022; Hong et al., 2020; Pace & Hayunga, 2020). 특히, 개별 주택의 가격 예측에서 두드러진 성과를 보이고 있는 접근법은 랜덤포레스트나 XGBoost와 같이 의사결정나무(decision tree)의 앙상블을 활용한 모형이다. 전통적인 모형과 달리, 트리 앙상블에 기반한 모형은 함수 형태에 대한 사전 가정 없이 주어진 데이터를 통해 비선형적이고 복잡한 특성 간의 관계를 자동으로 학습할 수 있으며, 이는 개별 주택 가격의 추정 정확도를 크게 높이는 것으로 나타나고 있다. 이러한 기계학습 모형은 기존의 선형회귀 기반 모형에 비해 상대적으로 정확성이 높을 뿐만 아니라, 절대적인 예측 성능 면에 있어서도 상당한 수준의 정확성을 확보할 수 있는 것으로 드러나고 있다(Hong et al., 2020). 그 중에서도 특히 기존 모형의 약점을 반복적으로 보강하는 부스팅 트리 알고리즘(XGBoost)이 특히 높은 정확성과 안정성을 보이는 것으로 알려져 있다(Hong & Kim, 2022).

본 연구의 목적은 XGBoost 알고리즘에 기반한 자동산정모형(AVM)을 이용하여 전국 단위의 개별 아파트 단지(단지 내 각 평형별) 가격을 각 시점마다(분기별) 전수적으로 추정하고, 이를 통해 지역별 주택시장의 가격 수준, 분포, 그리고 그 시계열적 변화를 직접 관측할 수 있는 방법을 제시하는 것이다. 본 연구가 제시하는 것은 가격지수가 아니라, ¹⁾ AVМ의 정확성을 활용하여 시장 전체의 가격 분포를 시점별로 추정·관측하는 방법이며, 이를 통해 가격지수가 직접 제공하지 못하는 정보 차원에 접근한다.

이러한 작업을 수행하기 위해서는 실거래 자료뿐 아니라, 전국 단위의 개별 아파트 단지에 신뢰할 만한 정보가 필요하다. 본 연구에서는 한국부동산원으로부터 제공된 전국 개별 아파트 단지의 평형별

1) 전통적으로 가격지수는 시기별 거래 표본의 질적 차이뿐 아니라 재고 주택 자체의 시기별 변화(멸실 및 신규 입주)에 대한 보정과 가중 방식의 선택 등 별도의 방법론적 논의를 요구하나, 이는 본 연구의 초점은 아니다.

정보와 아파트 실거래 정보를 사용하였는데, 이는 현재 구축되어 있는 대규모 공동주택 정보 중 가장 신뢰도가 높고 포함 범위가 넓은 자료라고 할 수 있다. 본 연구의 자료는 2012년 3월 1일부터 2024년 4월 1일 사이의 아파트 실거래 정보 6,118,730건과 총 12,517,131호에 해당하는 개별 아파트 단지의 평형별 정보를 포함하고 있는데, 이는 우리나라에 존재하는 전체 아파트 세대의 약 99.1%(2023년 총주택 총조사 기준 우리나라 전체 아파트 12,631,608 호 중)에 해당하는 것이다. 따라서, 본 연구의 추정치는 우리나라의 아파트 시장 전체의 상태에 대한 실질적인 전수 가격 추정에 가까운 결과가 될 수 있다.²⁾

본 연구에서 실거래가 자료는 개별 아파트 가격의 시점별 추정을 위한 학습에 사용되었으며, 개별 단지의 평형별 정보는 추정의 대상으로서, 동일한 표본을 전 기간에 걸쳐 고정적으로 유지함으로써 표본 구성 변화의 영향을 배제하였다. 이렇게 추정된 아파트 단지별 가격을 바탕으로 본 연구는 주택 가격의 수준과 분포 및 그 추이를 탐색하였다. 첫째, 본 연구에서는 전국 시·도 등 지역별로 단위 가격(만 원/m²)의 수준을 찾고, 지역 간 및 지역 내 가격 분포의 추이를 추적하였다. 둘째, 지역별로 PIR(price to income ratio)을 산정하여 지역별 아파트 가격이 가구의 실제 구매력 대비 어느 정도의 수준이고 어떤 차이가 있으며 어떻게 변화하였는지를 구체적으로 제시한다.

본 연구가 제시하는 전수 추정 기반 시장 관측은 다음과 같은 이점을 갖는다. 첫째, 본 연구의 결과는 전통적인 가격지수와 달리 절대적 가격 수치에 대한 정보를 제공한다. 이는 서로 다른 지역의 가격 수준을 직접 비교할 수 있게 하며, 가격의 변화뿐 아니라 수준 자체에 대한 해석을 가능하게 한다. 둘째, 본 연구의 추정 결과는 실질적인 전수에 해당하는 표본을 통해 가격 추이를 관측하고 있으므로 반복매매지수의 표본선택편의 가능성(Gatzlaff & Haurin, 1997; Melser, 2023)에서 자유롭다. 셋째, 전수적 가격 추정 정보를 기반으로 연구자가 원하는 방식대로 표본을 분할하여 하위 집단별 가격 추이, 분위수, PIR, 지니계수 등 다양한 분석적 지표로 쉽게 확장될 수 있다. 이는 기존의 고정된 표본 바스켓에 의존하는 지수 체계에서는 구조적으로 불가능했던 것으로, 시장에 대한 정보 접근성을 질적으로 확장한다.

본 연구의 구성은 다음과 같다. II장은 본 연구에서 활용된 기계학습 기법인 의사결정나무의 앙상블 알고리즘에 대한 소개 및 관련 선행연구들을 검토하였다. III장에서는 본 연구에서 사용된 자료에 대한 개략적 정보를 소개한다. IV장에서는 개별 주택 가격에 대한 추정을 수행하기에 앞서, 먼저 본 연구에서 사용한 기계학습 기반 자동산정모형이 개별 아파트 가격을 얼마나 정확하고 안정적으로 추정하는지를 실거래가 자료를 통해 테스트하였다. V장에서는 지역별 아파트 평균 가격 추이, 가격 분포의 추이, 가격 불균등성, PIR의 추이를 제시한다. VI장에서는 결론 및 향후 연구 과제를 논의한다.

2) 특히, 거의 전수에 가까운 표본을 가지고 있다는 것은 추정의 일반적 편향성이 발생하지 않는 한 분석 대상에 대한 거의 실제에 가까운 통계량들을 얻을 수 있다는 것을 의미한다.

II. 모형 개요 및 선행연구 검토

전통적으로 주택가격의 자동산정모형으로 가장 널리 알려진 접근법은 헤도닉 가격 모형일 것이다. 헤도닉 모형은 Lancaster(1966)의 소비자 이론과 Rosen(1974)의 연구를 이론적 토대로 하여, 주택의 시장 가치를 크기, 구조, 주변 환경과 같이 그 효용을 구성하는 요소들에 대한 함수로 표현하는 방식이다. 이러한 접근 방식은 가치에 영향을 미치는 개별 특성과 주택 가격과의 명시적인 함수 관계를 사전에 설정하고(Sheppard, 1999), 주어진 자료를 통해 가정된 관계의 계수를 회귀분석을 통해 추정하는 방법으로 작동한다.

헤도닉 모형은 직관적이고, 계수에 대한 가설검정 등을 통해 쉽게 실증 분석에 적용될 수 있다는 특징으로 인해 학술 연구에서나 실용적으로나 널리 사용되어 왔다. 그러나 헤도닉 모형이 직관적이고 단순화된 함수 관계를 가정한다는 사실로 인해 개별 주택에 대한 정밀한 가격 추정을 수행하기에는 한계가 있다는 점도 알려져 왔다. 현실에서는 주택 가격이 개별 특성들 사이에 복잡한 상호작용과 비선형적 관계가 존재할 가능성이 크지만, 그 구체적 형태를 모두 연구자가 사전에 알거나 정형화된 식으로 반영할 수는 없기 때문이다(Malpezzi, 2003). 결과적으로 선형함수 가정에 기반한 헤도닉 모형은 모형식별문제에 노출된다(Ekeland et al., 2004). 이는 헤도닉 모형이 높은 해석 가능성을 제공하지만, 개별 주택 가격에 대한 예측의 정밀성에서는 근본적인 제약을 가지고 있음을 의미한다. 결과적으로 전통적인 자동산정모형의 활용은 일반적으로 제한적인 범위에서만 수행되어왔다.

이러한 맥락에서, 최근 데이터 분석 기술의 폭발적 발전과 함께, 기계학습에 기반한 자동산정모형의 활용 가능성이 주목되고 있다(Tekouabou et al., 2024). 그 중에서도 특히 의사결정나무(decision tree)의 앙상블 기법은 연구자가 주택 가격과 특성 간의 명시적인 함수 형태를 미리 가정할 필요가 없고, 주어진 데이터를 통해 비선형적이고 복잡한 특성 구조를 효과적으로 포착하여 주택 가격 예측의 정확성을 획기적으로 높일 수 있는 것으로 알려져 있다(홍정의, 2021; Antipov & Pokryshevskaya, 2012). 이러한 기계학습 모형은 전통적 모형과 달리 가격과 관련 변수 간의 내생적 관계를 직관적으로 해석하기에는 적합성이 떨어질 수 있지만, 그럼에도 자동산정모형의 주요 목표인 개별 주택 가격의 추정에 있어서는 기존 방법론의 정확성을 크게 상회하는 성능을 가질 수 있다. 다수의 선행 연구들은 의사결정나무의 앙상블을 통해 주택가격 예측의 정확성이 기존 전통적 모형에 비해 현저히 개선될 수 있음을 보고하고 있다(Hong et al., 2020; Kim & Hong, 2024; Pace & Hayunga, 2020; Park & Bae, 2015; Rico-Juan & de La Paz, 2021). 예를 들어, Hong et al.(2020)은 대표적인 앙상블 기법인 랜덤 포레스트를 이용하여 대한민국 서울의 강남 지역 아파트 가격을 분석하였는데, 기계학습 기반 모형의 예측 정확성이 평균백분율오차(mean absolute percentage error, MAPE) 기준으로

5% 내외인 것으로 나타나, 이러한 모형이 단순히 상대적으로 전통적 모형에 비해 나은 뿐 아니라, 절대적인 수준에서도 상당히 신뢰할 만한 예측을 수행할 수 있음을 보여준다.

의사결정나무는 특성 간의 명시적 함수 관계에 기초한 전통적 모형과 달리, 주어진 데이터를 특정 조건에 따라 반복적으로 나누어 가격의 차이를 가장 잘 설명할 수 있는 규칙을 자동적으로 탐색하는 알고리즘이다. 이는 마치 스무고개 놀이에서 세부적인 질문을 통해 점점 정답에 가까워지는 과정과 비슷한 형태로 작동하는데, 개별 주택의 특성에 따라 가장 유사한 그룹을 찾아 그 그룹의 가격 평균이나 중앙값을 예측값으로 삼는다. 특히 이 방식은 변수 간의 비선형 관계를 자연스럽게 포착할 수 있으며, 주택 가격에 영향을 미치는 위도와 경도 같은 변수도 별도의 변환 없이 직접 입력하여 유연하게 활용할 수 있다는 장점이 있다.

이러한 특성은 의사결정나무 기반 모형이 학습 자료에 분포의 양 극단에 위치한 비통상적 표본이 포함되어 있더라도 예측에서 상대적으로 강건하다는 점과도 연결된다. 의사결정나무는 분할 기준을 변수의 절대 수준이 아닌 순위(rank)에 따라 결정하므로, 일부 변수에 통상 범위를 크게 벗어난 값이 소수 존재하더라도 분할 지점의 선택이 그 값의 절대적 크기가 아닌 다른 표본들과의 상대적 순서에 의해 좌우되며, 모형 전체의 학습 결과가 소수의 극단값에 의해 크게 왜곡되지 않는다. 또한 의사결정나무의 예측값은 학습 자료에서 동일한 잎(leaf)에 도달한 표본들의 가격을 통계량(평균 또는 중앙값)으로 집계한 결과이므로, 예측 대상이 통상 범위를 벗어나는 입력값을 갖더라도 그 예측은 학습 표본 내 인접 그룹의 가격 범위 내로 자연스럽게 제약된다. 이러한 구조적 특성은 본 연구와 같이 전국 사실상 전수의 단지를 추정 대상으로 삼아 평균·분위수·지니계수·PIR 등의 집단 통계량을 산출하는 분석에서 특히 유리하게 작용한다. 즉, 추정 대상 표본의 일부에 비통상적 값이 포함되어 있더라도, 그러한 표본에 대한 추정치가 학습 자료의 인접 그룹 가격 범위 내로 제약되는 한, 전체 표본을 대상으로 집계한 결과 통계량에 미치는 영향은 제한적이다.

의사결정나무 단독으로 사용하면, 특정 훈련 자료에 지나치게 맞춰져 다른 자료에 적용할 때 성능이 떨어지는 과적합 문제가 발생할 수 있다. 이를 해결하기 위해 주택가격 예측 모형에서는 주로 다수의 의사결정나무를 결합하는 앙상블 기법이 주로 활용된다. 앙상블이란 여러 개의 의사결정나무를 동시에 또는 순차적으로 결합하여 보다 정확하고 안정적인 모형을 구성하는 방법을 말하는데, 이는 여러 개의 개별 모형의 예측 결과를 결합함으로써 개별 나무가 가진 한계나 오류를 감소시키고 보다 견고한 결과를 얻는 방식이다. 이러한 앙상블 기법은 크게 배깅(bagging)과 부스팅(boosting)으로 구분할 수 있다. 배깅 모형의 대표적인 예로는 Breiman(2001)에 의해 제시된 랜덤 포레스트가 있다. 이는 데이터를 서로 다른 방식으로 반복적으로 샘플링하여 독립적으로 훈련된 여러 나무의 평균 예측값을 사용하는 방식이다. 기계학습을 활용한 모형의 초기에 랜덤 포레스트를 활용한 연구가 다수 제시되었다. 예를 들어, Antipov & Pokryshevskaya(2012)는 러시아 상트페테르부르크의 아파트 2,848건

에 대한 데이터를 분석한 결과 랜덤 포레스트가 다른 여러 예측 모형에 비해 높은 예측 정확도를 나타낸다는 점을 확인하였다. 또한 Ho et al.(2021)은 홍콩에서 18년간 발생한 4만 건의 주택 거래 자료를 분석하여, 랜덤포레스트와 부스팅 기반의 모델들이 지지벡터기계(support vector machine, SVM)보다 더 예측 정확성이 높았음을 확인하였다. Yilmazer & Kocaman(2020)는 터키 앙카라의 복잡한 도시 구조를 지닌 마마크(Mamak) 지역을 사례로 선정하여 전통적 회귀모형과 랜덤 포레스트를 통한 대량평가의 결과를 비교하였다. 그 결과 랜덤 포레스트는 요인 변수가 복잡하게 형성되는 경우에도 그 구조를 자동으로 최적화하여 회귀모형보다 예측 정확성이 더 뛰어남(조정 R^2 0.734)을 확인하였다.

부스팅은 이전 단계에서 잘못 예측된 데이터 포인트를 다음 단계에서 보다 중점적으로 학습하여 점진적으로 성능을 높이는 방식이다. 배깅이 개별 나무를 독립적으로 생성하는 반면, 부스팅은 직전의 나무가 잘못 예측한 데이터에 초점을 맞추어 다음 나무가 이를 보완하도록 하는 원리이다. 이러한 방식으로 부스팅은 개별 나무의 민감성을 점진적으로 줄이고 전체 모형의 예측 정확성을 지속적으로 높여간다. 이러한 부스팅 방식의 대표적인 예가 최근 주택 가격 예측에서 뛰어난 성능을 보여주고 있는 것으로 알려진 XGBoost(eXtreme Gradient Boosting)이다. XGBoost는 순차적으로 새로운 의사결정나무를 추가하여 이전 단계의 예측 오류를 점진적으로 개선해나가는 방식으로 작동한다. 각 나무는 앞서 구성된 나무가 잘못 예측한 자료를 집중적으로 학습하도록 설계되어 전체적으로 예측력을 높여간다. XGBoost는 복잡한 비선형 관계를 효과적으로 탐색하여 데이터에 내재된 구조를 잘 포착할 뿐 아니라, 나무의 복잡성을 제어하는 정규화를 활용하여 과적합 문제를 보다 적극적으로 관리할 수 있다.

최근의 관련 연구들 역시 개별 주택가격 예측에 있어 XGBoost의 활용 가능성을 주목해왔다(Guliker et al., 2022; Hong & Kim, 2022; Ming et al., 2020). 예를 들어, Ming et al.(2020)은 XGBoost를 비롯한 부스팅 기반 모형들이 부동산 가격 예측에 탁월한 성능을 나타내며, 특히 가격에 영향을 미치는 복잡한 변수 간 관계를 보다 효과적으로 학습하여 전통적 모형과의 성능 차이가 명확하다는 것을 보여주었으며, Hong & Kim(2022)은 서울의 아파트 실거래가 자료를 기반으로 XGBoost를 비롯한 의사결정나무 모형들의 예측력이 평균백분율오차(MAPE) 기준으로 4% 내외에 불과해 실용적으로 활용이 가능한 수준임을 보여주었다.

관련 연구들은 해외뿐 아니라 국내에서도 빠르게 늘고 있는 추세이다. 예를 들어, 배성완·유정석(2018)은 의사결정나무 기반 모형인 랜덤 포레스트, 그래디언트 부스팅 회귀 트리를 비롯한 다양한 기계학습 기법을 활용하여 아파트 가격지수에 대한 예측 성능을 전통적 시계열 모형과 비교하였다. 분석 결과, 특히 의사결정나무 기반 모형들이 시계열 분석 모형에 비해 전반적으로 우수한 예측력을 나타냈으며, 시장이 급격히 변동하는 상황에서 추세 예측의 정확성이 더욱 두드러지게 향상되었음을 밝혀냈다. 이인호·이경섭(2020)은 서울의 아파트 실거래가 자료 71만건을 토대로 XGBoost, 랜덤 포레스트, LightGBM과 같은 의사결정나무 앙상블 기법의 설명력을 분석하였다. 분석 결과는 의사

결정나무 기반 알고리즘을 통해 전통적 헤도닉 모형에서 나타날 수 있는 다중공선성으로 인한 예측 왜곡을 극복할 수 있음을 보여준다. 홍정의(2021)는 주택시장의 복잡성을 랜덤 포레스트 모형이 어떻게 포착하는지에 대해 분석하고, 모형의 효율적 설계 방법을 제시하였으며, 홍정의(2020)는 의사결정나무 앙상블 모형 중 부스팅 기반의 세 알고리즘(XGBoost, LightGBM, CatBoost)을 통해 주택가격 대량 평가 모형을 구축하고, 성능과 특징을 비교 분석하였다. 논문은 2009년부터 2019년까지 서울에서 거래된 아파트 매매 데이터 620,617건을 통해 분석한 결과, 헤도닉 모형의 경우 시장 가격에 대한 예측의 평균 백분율 오차가 약 11.5% 내외인 반면, XGBoost, LightGBM, CatBoost는 각각 3.7%, 3.8%, 3.6%에 불과한 것으로 나타났다. 이는 의사결정나무 기반 모형의 예측력이 단순히 상대적인 측면(즉, 전통적인 헤도닉 모형에 비해)뿐 아니라, 절대적인 측면에서도 상당히 높게 나타나, 그 실용적 잠재력이 상당히 높다는 것을 보여준다.

이처럼 최근의 연구들은 의사결정나무 기반 모형의 성능과 활용 가능성에 대해 보여주고 있다. 이 시점에서 관련 연구는 단순히 예측 성능의 제시를 넘어, 높은 예측 정확성 등을 활용하여 어떻게 주택 시장을 보다 심층적으로 분석하고 이해하는 데에 활용될 수 있는지에 대한 탐구로 나아가는 추세이다. 초기 AVM 연구가 주로 개별 자산의 감정평가 정확도 향상에 집중되었다면(Antipov & Pokryshevskaya, 2012; Park & Bae, 2015), 최근에는 AVM이 산출하는 정확한 추정치를 시장 전체의 가격 동학을 관측하는 도구로 확장하려는 시도가 나타나고 있다. 예를 들어, Gallin et al.(2021)은 Zillow의 AVM을 활용하여 미국의 총 주택자산 가치를 추정하면서, AVM 기반의 재고 중심(stock-based) 측정이 기존의 반복매매지수나 설문 기반 자가추정보다 더 안정적인 결과를 제공할 수 있음을 보여준 바 있다. 이는 AVM이 단순한 개별 자산 감정평가의 도구를 넘어 시장 전체의 가격 수준을 측정하는 도구로 활용될 수 있음을 시사한다.

김이환 외(2022)는 네이버 부동산을 통해 수집한 아파트 정보에 랜덤포레스트와 인공지능망을 적용하여 서울시 아파트의 추정 매매가격을 산출한 뒤, 시점·자치구·권역별 시간더미 추정치를 활용하여 반기별 매매가격지수를 산정하였고, 그 지수를 통해 나타난 변동성이 KB지수나 한국부동산원 실거래가지수에 비해 가격 상승기에 더 크게 나타남을 보고함으로써 기계학습 기반 추정이 기존 지수의 평활화 한계를 보완할 수 있음을 시사하였다. 이소영·김경민(2025)은 LightGBM 기반 AVM을 통해 서울시 다세대주택 약 71,970개 단지(약 65만 호)의 월별 잠재가격을 추정한 후, 단지·기초구역·자치구·생활권 단위의 라스파이레스 가격지수를 산정함으로써 반복매매모형으로는 표본 한계 때문에 산출이 어려웠던 소지역 단위 가격지수의 가능성을 실증하였다.

두 연구는 본 연구와 유사한 방법론적 골격(AVM의 추정가격을 활용한 기중평균 산출)을 공유하나, 산출물의 형태와 분석의 범위에서 본 연구와 구별된다. 두 선행연구가 모두 가격지수의 산정과 그 평활화 한계의 보완을 1차적 목적으로 삼은 한편, 본 연구는 전국 단위의 전수적 재고 기반 추정을

목적으로 하고 있으며 이를 통한 재고 주택 가격 분포·하위집단별 통계량의 직접 관측에 초점을 맞춘다. 또한 자료의 측면에서 두 선행연구가 서울시에 한정된 단일 주택유형(아파트 또는 다세대주택)을 대상으로 한 반면, 본 연구는 전국 17개 시·도의 아파트 재고 주택 약 1,250만 호(전체의 99.1%)를 추정 대상으로 하여 분석 범위가 가장 넓다.

본 연구에서는 전국에 있는 개별 아파트 단지(단지 내 각 평형별)의 가격을 추정한 뒤, 지역별로 아파트 가격 수준과 분포, 그리고 소득 대비 주택가격 비율(PIR)이나 지니 계수와 같은 주요한 관련 지표들을 산출할 것이다. 이를 통해 본 연구에서 제시하는 접근을 통해 기존의 통계로는 직접 관측이 어려웠던 가격 변화의 추이 및 지역 간 차이 또는 지역 내 분포에 대한 현황을 효과적으로 관측할 수 있을 것으로 기대한다.

본 연구가 분석적 예시로 다루는 PIR과 지니계수는 산정 방식에 따라 결과가 상당히 달라질 수 있는 지표인 만큼, 그간의 관련 논의를 간략히 정리해둘 필요가 있다. PIR은 가계의 주택구입 부담을 측정하는 가장 대표적인 지표로, 국내에서는 KB국민은행, 한국주택금융공사, 국토교통부 주거실태 조사 등이 각기 다른 표본·통계지표·산정방식으로 PIR을 산출·공표하고 있는데, 그 결과 동일 시점에 대해서도 기관 간 수치가 상당한 차이를 보인다. 이창무 외(2012)는 PIR이 표본·지역 구분·통계 지표에 민감한 지표라는 점을 이론적·실증적으로 분석하고, 국제 비교를 위해서는 평균 기반·세전 정의를 채택하는 일관된 산정 방식이 요구됨을 제시하였다. 또한 권진웅·김남정(2020)은 서울 25개 자치구의 약 78만 건 실거래 자료를 활용하여 PIR을 자치구 단위로 산정한 후, 그 시계열 변동 패턴이 자치구별로 4개 유형으로 분화됨을 보였는데, 이는 미시 자료를 활용한 PIR 분석이 시·도 단위 평균보다 풍부한 정보를 산출할 수 있음을 시사한다.

한편 주택가격 분포의 불평등성에 대한 분석은 가구 자산 단위와 주택 가격 단위의 두 흐름으로 나누어 이해할 수 있다. 통계청의 가계금융복지조사를 활용한 분석들은 가구 단위 순자산 지니계수가 2017년 0.584에서 최근 0.625 수준까지 지속적으로 상승하였음을 보고하며, 이는 주택가격 상승이 가구 자산 불평등 심화에 결정적 역할을 했음을 의미한다(김지원·마강래, 2021; 송영호·마강래, 2025; 정의철 외, 2009). 한편 주택 가격 자체의 분포가 얼마나 불균등한지를 자산(주택) 단위에서 측정하려는 시도도 늘어나고 있다. Kang & Koo(2023)는 국토교통부 실거래 자료를 활용하여 서울 25개 자치구의 주택가격 지니계수를 산출하고, 이를 사회자본 수준과의 관계 분석에 활용하였다. 서재원·진장익(2021) 또한 주택가격 불평등이 계층이동 가능성에 대한 기대에 미치는 영향을 분석하였다.

이러한 관련 분석들은 필연적으로 이질적인 주택 가격과 소득 및 자산의 변화를 대규모로 추적하거나 그것을 대표하는 지표를 환산하는 작업을 수반하므로, 상당한 기술적 비용이 요구되었다. 본 연구에서 제시된 접근은 이러한 관련 분석을 기술적으로 용이하게 하며, 그 기술적 편의성을 바탕으로 보다 넓은 범위에 대해 접근할 수 있도록 한다. 본 연구는 이창무 외(2012)가 정리한 평균 기반·세전

정의의 따름으로써 국제비교성을 확보하면서, 동시에 17개 시·도 단위의 분기별 추정치를 일관된 방식으로 산출하며 지역 내 변화뿐 아니라 지역 간 비교를 가능하게 한다. 이처럼 본 연구는 주택 가격에 대한 보다 종합적이고 입체적인 정보를 제공함으로써 추후 다양한 후속 연구를 위한 토대가 될 수 있을 것으로 기대한다.

III. 분석 자료

본 연구에서는 전국 단위의 아파트 단지의 평형별 정보와 아파트 실거래 자료이다. 두 자료는 그 성격과 본 연구에서의 역할이 서로 구별된다. 단지 자료는 분석 시점 기준으로 전국에 존재하는 아파트 단지의 평형별 정보를 담고 있어 거래 여부와 무관하게 시장의 재고 상태 그 자체를 기술하며, 자료로 포착된 단지의 총 주택 호수는 12,517,131호로 2023년 인구주택총조사 기준 전국 아파트 12,631,608호의 약 99.1%에 해당하여 사실상 전국 아파트 재고의 거의 전수를 포괄한다. 한편 실거래 자료는 같은 단지들 가운데 2012년 3월 1일에서 2024년 4월 1일까지 약 12년의 관측 기간 동안 거래가 발생한 사례들의 가격 기록이며, 그 규모는 6,118,730건이다. 한 호의 주택이 해당 기간 동안 한 번도 거래되지 않을 수도, 여러 차례 거래될 수도 있으므로, 두 자료의 표본 수는 서로 단순한 부분집합 관계를 갖지 않는다. 본 연구의 분석 절차에서도 두 자료는 서로 다른 역할을 수행한다. 실거래 자료는 가격이 관측된 사례를 통해 주택 특성과 가격 사이의 관계를 학습시키는 데에 사용되며, 단지 자료는 학습된 모형의 추정 대상으로서, 모형이 단지 자료의 모든 단지·평형에 분기별로 적용됨으로써 거래 여부와 무관하게 전 단지의 시점별 추정 가격이 얻어진다.

본 연구의 단지 자료는 분석 시점을 기준으로 한 단일 시점의 단지·평형별 정보로 구성되며, 입지·구조 특성 또한 분석 시점 기준의 상태로 모든 추정 시점에 동일하게 적용된다. 즉, 분석 기간(2012년 3월~2024년 4월) 중에 발생한 지하철 노선의 개통, 학교의 신설·폐교, 재건축·재개발 추진단계의 변화 등은 분석 시점(2024년 4월 1일)의 상태로 통일되어 모형의 입력 변수로 사용된다. 이는 본 연구가 측정하고자 하는 대상이 시점에 따라 함께 변화한 입지·재고 특성을 반영한 시장 가격이 아니라, 동일하게 정의된 표본 바스켓에 대해 시점별 시장 환경이 부여한 가격 평가이기 때문이다. 이러한 의미에서 본 연구의 추정은, 표본 바스켓을 고정하여 산출하는 기존 가격지수의 확장된 형태에 가깝다. 이러한 설계는 본 연구가 산출하는 시점별 통계량(평균, 분위수, 지니계수 등)이 시점 간 표본 구성의 변화로부터 영향을 받지 않도록 함으로써, 전수적 추정에 기반하더라도 추정된 지표가 최대한 지수적 성격을 갖도록 가격 변화의 해석을 단순화하기 위한 것이다.

실거래 자료는 학습 기간으로 포함된 기간 중 발생한 모든 거래를 학습에 사용하였으며, 별도의

단지별 최소 거래 건수 기준은 설정하지 않았다. 즉 분석 기간 중 거래 빈도가 낮은 단지의 거래도 모두 학습에 포함되며, 단지·시점 단위의 거래 빈도 차이로 인한 표본의 불균형은 별도의 가중 조정 등을 적용하지 않은 채 모형의 학습에 반영하였다.

분석에서는 다양한 선행 연구들로부터 가격과의 관련성이 있는 것으로 알려진 총 27개의 변수가 사용되었다. 이들은 변수의 형식에 따라서는 수치형 변수와 범주형 변수로 나뉘며, 속성에 따른 분류로는 건물 특성과 단지 특성 및 입지적 특성으로 분류할 수 있다(〈표 1〉). 먼저, 건물 특성으로는 공급면적, 전용면적, 건축 구조, 층, 승강기의 유무, 복층의 유무, 1층 전용정원의 유무, 다락방의 유무가 포함된다. 이는 건축물 자체의 특성과 관련이 있다. 단지 특성으로는 건폐율, 용적률, 단지 내의 총 세대수, 대지 지분, 단지의 건설사 브랜드가 포함된다. 단지 특성은 광의의 건축적 특성에 해당하나, 개별 세대보다는 단지 전체에 포괄적으로 적용되는 특성들을 분류하였다. 입지적 특성은 가장 기본적인 입지 구분자인 위도와 경도, 배정되는 학교(초등학교, 중학교, 고등학교), 입지의 경사 수준, 용도지역, 가장 가까운 초등학교/중학교/고등학교/지하철로부터의 거리, 그리고 재개발 관련 상태를 나타내는 변수(사업유형, 안전진단등급, 추진단계)가 포함되었다.

〈표 2〉와 〈표 3〉은 이러한 변수들에 대한 기초 통계량을 보여주고 있다. 기초 통계량은 단지 자료와 실거래 자료 중, 단지 자료를 중심으로 계산하였는데 이는 실거래 자료가 모든 단지에 대한 정보가 아닌 해당 시점에 거래된 단지의 자료만을 담고 있을 뿐 아니라, 동일한 단지에 대한 중복된 거래 정보를 담고 있으므로 보다 거래가 빈번하게 이루어지는 단지의 정보를 더 많이 담고 있기 때문이다.

먼저 수치형 변수에 대한 기초 통계량을 살펴보자. 아파트 개별 세대 및 단지 특성을 나타내는 주요 변수들이 비교적 넓은 분포 범위를 가지고 있는 것으로 나타났는데, 이는 본 연구에서 사용된 자료가 전국의 모든 단지를 대상으로 하고 있기 때문이다. 공급면적의 경우 전체 표본의 평균은 약 94.7m^2 이며, 중앙값은 89.7m^2 로 나타났다. 최소값은 9.3m^2 , 최대값은 $1,023.9\text{m}^2$ 에 이르며, 표준편차는 42.4m^2 로 상당히 높은 편이다. 이는 소형 원룸형 주택부터 대형 펜트하우스까지 다양한 규모의 주택이 대상으로 포함되어 있음을 보여준다. 전용면적 역시 평균은 76.4m^2 , 중앙값은 74.9m^2 이며, 최소값은 7.3m^2 , 최대값은 464.1m^2 로 공급면적과 유사한 분포 특성을 보인다. 이러한 면적 관련 변수들은 개별 세대의 가격 결정에 반영될 수 있는 핵심 요소로 알려져 있다. 단지 특성 변수 중 세대수는 평균 약 360세대이며, 중앙값은 136세대로 나타난다. 그러나 최대값은 9,510세대에 이르며, 표준편차 또한 521로 나타난다. 이 역시 자료에 포함된 아파트 단지가 수십 세대 규모의 소규모 단지부터 수천 세대를 포함한 대규모 단지까지 매우 다양한 형태로 존재한다는 점을 보여준다. 용적률은 평균 301.3%, 중앙값 245.1%이며, 최대 1,472%까지 관측되었다. 이는 고밀도 개발이 허용된 도심 지역 일부 단지에서 나타나는 특성으로, 전체적으로는 평균과 중앙값의 차이가 큰 비대칭 분포를 보인다. 건폐율 역시 평균이 37.0% 수준이며, 중앙값은 31.6%, 최대값은 98.4%로 나타났다.

〈표 1〉 변수의 목록

분류	변수명	형식	조작적 정의
건물특성	공급면적	수치형	전용면적과 공용면적의 합(m ²)
	전용면적	수치형	해당 세대만이 독점적으로 사용하는 면적(m ²)
	층	수치형	해당 세대가 위치한 층 (지하층은 음수로 표기, 예: 지하 1층=-1)
	승강기	범주형(0/1)	해당 단지의 승강기 설치 여부(더미: 1/0)
	복층여부	범주형(0/1)	해당 세대가 복층 구조인지 여부(더미: 1/0)
	1층 전용정원	범주형(0/1)	1층 세대에 부속된 전용정원의 유무(더미: 1/0)
	최상층 다락방	범주형(0/1)	최상층 세대에 부속된 다락방의 유무(더미: 1/0)
	구조	범주형(43범주)	건축 구조의 유형 (예: 철근콘크리트구조, 철골철근콘크리트구조 등)
단지특성	건폐율	수치형	단지 대지면적 대비 건축면적의 비율(%)
	용적률	수치형	단지 대지면적 대비 건축 연면적의 비율(%)
	세대수	수치형	단지 내 총 가구 수(호)
	대지지분	수치형	개별 주택에 귀속되는 단지 내 대지지분(m ²)
	브랜드	범주형(17범주)	단지의 건설사 브랜드(예: 일반, 힐스테이트 등)
입지특성	위도	수치형	단지 대표 좌표의 위도(도)
	경도	수치형	단지 대표 좌표의 경도(도)
	초등학교	범주형(4,082범주)	단지에 학군 배정된 초등학교 식별자
	중학교	범주형(2,398범주)	단지에 학군 배정된 중학교 식별자
	고등학교	범주형(939범주)	단지에 학군 배정된 고등학교 식별자
	고저	범주형(7범주)	단지 입지의 경사 수준(예: 급경사지, 완경사지 등)
	용도지역	범주형(20범주)	단지가 속한 「국토의 계획 및 이용에 관한 법률」상 용도지역(예: 제1종일반주거지역, 제2종일반주거지역 등)
	사업유형	범주형(8범주)	단지에 적용되는 정비사업의 유형(예: 주거환경개선, 공공재개발 등). 해당 없는 단지는 별도 해당없음 범주로 처리
	안전진단등급	범주형(6범주)	재건축 안전진단을 받은 단지에 부여된 등급(예: C, D 등). 미진단 단지는 별도 해당없음 범주로 처리
	추진단계	범주형(14범주)	정비사업의 추진 단계(예: 해당 없음, 정비구역지정 등)
	초등학교 거리	수치형	단지에서 가장 가까운 초등학교까지의 직선거리(m). 2km 초과 시 2km로 상한 처리
	중학교 거리	수치형	단지에서 가장 가까운 중학교까지의 직선거리(m). 2km 초과 시 2km로 상한 처리
	고등학교 거리	수치형	단지에서 가장 가까운 고등학교까지의 직선거리(m). 2km 초과 시 2km로 상한 처리
	지하철 거리	수치형	단지에서 가장 가까운 지하철까지의 직선거리(m). 2km 초과 시 2km로 상한 처리. 지하철이 부재한 지역에서는 모형을 제외

주 : 1) 지하철이 없는 지역의 경우, 지하철 거리 변수를 제외하고 모형을 학습 및 예측하였음.
 2) 모든 변수는 분석 시점 기준(2024.4.1.)의 단지-평균 상태를 기록한 값이며, 분석 기간 전체에 걸쳐 동일하게 적용된다.

〈표 2〉 기초 통계: 수치형 변수

	평균	최소값	중앙값	최대값	표준편차
건폐율	36.97	1.17	31.58	98.38	19.42
용적률	301.27	0.09	245.14	1,472	177.61
세대수	359.76	1	136	9,510	521.41
공급면적	94.66	9.26	89.69	1,023.91	42.43
전용면적	76.44	7.32	74.9	464.11	35.15
대지지분	39.39	0	37.25	31,257.4	75.34
층	11.83	-4	10	84	8.34
초등학교 거리	434.1	34	383	1,992	256.91
중학교 거리	613.02	40	546	1,999	360.41
고등학교 거리	606.89	0	574	1,998	508.67
지하철 거리	500.55	19	466	1,997	259.17
총 재고 표본수 (모든 평형)	1,128,355				

〈표 3〉 범주형 변수의 빈도

변수명	범주값	호수
승강기	N	718,762
	Y	11,798,369
복층여부	N	12,502,134
	Y	14,997
1층전용정원	무	12,516,740
	유	391
최상층다락방	무	12,506,782
	유	10,349
고저	고지	18,761
	급경사지	33,817
	없음	754
	완경사지	1,313,740
	저지	6,385
	평지	10,704,984
	미표기	438,690
사업유형	재개발(주택)	201,189
	재개발(도시환경)	18,657

〈표 3〉 계속

변수명	범주값	호수
사업유형	재건축(공동)	336,068
	재건축(단독)	36,381
	주거환경개선	43,426
	공공재개발	4,560
	공공재건축	270
	해당없음	11,876,580
안전진단등급	C	6,298
	D	218,512
	E	12,556
	F	4,703
	G	4,865
	해당없음	12,270,197

이외에도 수치형 변수로는 층수나 대지지분, 위도·경도 등 수치형들이 포함되어 있다.³⁾ 초등학교, 중학교, 고등학교, 지하철로부터의 거리는 최대 2,000미터까지만 계산되었으며 그 이상인 경우 최대 값으로 분류되었다. 또한, 분석대상 지역 중, 지하철이 존재하지 않는 지역에서는 모형의 학습 및 예측 과정에 해당변수(지하철역으로부터의 거리)를 제외하였다.

또한, 위도와 경도는 다른 수치형 변수와 달리 특정한 속성의 다소를 의미하는 정보라기보다는 위치값의 구분자 그 자체를 의미한다. 본 연구가 위도·경도를 별도의 변환 없이 입력 변수로 그대로 사용하는 것은 다음의 두 가지 근거에 따른 것이다. 첫째, 본 연구의 목적은 개별 변수의 가격 효과에

3) 언급된 것처럼 본 연구의 자료는 분포의 양 극단에 위치하는 비통상적 표본을 일부 포함하고 있다. 본 자료는 한국부동산원이 사전 검수를 거친 것이나, 자료 입력 단계의 정합성이 곧 모든 표본의 분석상 적합성을 보증하는 것은 아니므로, 본 연구에서는 비통상적 값을 갖는 표본에 대해 추가적인 검토를 수행하였다. 구체적으로, 층수, 공급·전용면적, 세대수, 대지지분 등 주요 변수의 분포에서 극단에 위치한 표본을 추출하여 단지명과 주소를 식별한 뒤, 공개 자료와의 대조를 통해 해당 단지의 실재 여부를 확인하였다. 다만 개별 단지의 변수가 입력 단계에서 얼마나 정확하게 코딩되었는지를 변수 단위로 완전히 검증하기 위해서는 사실상 현장 조사에 준하는 작업이 요구되며, 이는 본 연구의 범위를 넘어선다. 비통상적 값을 갖는 표본 중 일부에 입력 단계의 미세한 부정확성이 존재할 가능성은 100% 배제하기 어려우며, 이는 대규모 행정 자료에 기반한 연구가 일반적으로 갖는 한계의 하나로 이해될 수 있다. 한편 본 연구는 전국 사실상 전수의 아파트 단지를 추정 대상으로 삼고 있어, 통계적 기준이나 임의의 절단 기준을 적용하여 분포의 양 극단에 위치한 표본을 일률적으로 "이상치"로 분류·제외하는 처리는 전수 추정을 핵심 기여로 하는 본 연구의 설계와 상충되는 측면이 있다. 또한 II장에서 설명한 바와 같이 트리 앙상블 알고리즘은 분포의 양 극단에 위치한 표본에 대해 구조적으로 강건한 특성을 가지므로, 비통상적 표본이 결과 통계량에 미치는 영향은 제한적일 것으로 판단된다. 이상의 검토를 바탕으로, 본 연구에서는 비통상적 값을 갖는 표본을 별도로 분류·제외하지 않고 자료 전체를 그대로 분석에 활용하였다.

대한 해석이 아니라 시점별·단지별 가격 추정 그 자체에 있으며, 트리 앙상블 알고리즘은 위도·경도와 같이 변환되지 않은 좌표 자체를 입력으로 받아 입지 효과를 비선형적·국지적으로 포착하는 데에 효율적임이 이론적으로 알려져 있다. 둘째, 기존 연구들은 트리 앙상블 모형에서 위도·경도를 직접 입력으로 사용하는 방식이 일관적으로 가장 높은 예측 정확성을 제공함을 보고해 왔다(홍정의, 2020, 2021). 특히 위도·경도를 입력 변수로 사용한 모형에 행정구역 더미 등을 추가로 투입할 경우 예측력이 추가로 개선되지 않으며, 오히려 변수 수의 증가로 인한 효율성 저하가 발생할 수 있음이 보고된다(홍정의, 2021). 이러한 점들을 종합적으로 고려하여, 본 연구에서는 위도·경도를 직접 입력 변수로 사용하되 별도의 행정구역 단위 위치 변수는 추가하지 않는 방식을 채택하였다.

범주형 변수의 경우, 전반적으로 이진형 범주값을 가지며, 분포는 다소 불균형한 편이다. 예를 들어, 승강기 유무는 ‘있음’이 전체의 약 94%를 차지하고 있으며, ‘없음’은 약 6%에 불과하다. 이는 최근 수십 년간 건설된 아파트 대부분에 승강기가 설치되어 있는 현실을 반영한 결과이다. 한편, 복층 여부와 1층 전용정원의 경우, ‘있음’ 응답은 전체의 각각 약 1%와 0.1% 수준에 불과해 상당히 극히 드문 특성으로 확인되었다. 이는 대부분의 아파트가 단층 구조로 이루어져 있으며, 이러한 구조는 고급 주택이나 특정 단지에서 제한적으로 나타나는 선택적 구조임을 보여준다. 또한 초등학교·중학교·고등학교 변수는 분석 시점(2024년 4월 1일) 기준으로 해당 단지에 학군 배정된 학교의 학교명으로, 이는 입지 가치에 주요한 설명력을 가질 수 있는 것으로 알려져 있다(김세울·유선종, 2022). 이러한 학교 변수는 위도·경도와 상보적인 지리 식별자로서 모형에 포함되었다. 우리나라의 학군 배정 체계는 행정구역 단위보다 세밀한 공간적 구획을 자연적으로 형성하며, 그 경계는 연속적 변수인 위도·경도가 학습된 트리 구조 내에서 직접 포착하기 어려운 미시적 불연속성을 갖는다. 즉, 위도·경도가 연속적 입지 효과를 비선형적으로 포착하는 데에 효율적이라면, 학교 변수는 학군 경계에서 발생하는 단절적 입지 효과를 범주형으로 직접 식별하는 역할을 수행한다.⁴⁾⁵⁾

- 4) 이는 본 연구의 분석에 행정구역 더미를 추가하지 않은 결정과 모순되지 않는데, 행정구역 더미는 위도·경도가 이미 포착하는 광역 단위 입지 효과와 상당 부분 중복되는 반면, 학교 배정 단위는 행정구역보다 세밀하고 광역 좌표가 직접 포착하기 어려운 단절적 경계를 식별하기 때문이다.
- 5) 학교 변수는 위와 같이 4,082개에 이르는 범주 수를 가지나, 본 연구의 사용 맥락에서 범주 수 자체가 과적합 위험을 특별히 높인다고 보기는 어렵다. 본 연구가 사용하는 부스팅 트리 알고리즘은 분할 기준의 선택 단계에서 정규화 항을 통해 트리의 복잡성을 제어하며, 학습 표본의 규모(약 612만 건의 실거래) 대비 학교 변수의 범주 수는 범주당 평균 1,000건 이상의 학습 표본을 확보할 수 있는 수준이다. 또한 범주의 수가 많다는 것은 곧 학교 변수가 단지의 미시적 위치를 더 세밀하게 식별한다는 것을 의미하므로, 학습 표본의 밀도가 충분히 확보되는 한 미시적 식별력 자체가 모형의 일반화 성능을 저해할 이유는 없다. 본 변수의 시간적 자료 누설(temporal leakage) 가능성에 대해서도, 학교 배정 정보는 위에서 밝힌 바와 같이 분석 시점(2024년 4월 1일)의 상태로 전 시점에 동일하게 적용되므로, 추정 대상 시점의 정보를 사후적으로 사용하는 형태의 누설이 발생하지 않는 구조이다. 이러한 처리 방식의 적절성은 IV장의 평가 결과를 통해 실증적으로 뒷받침된다. 무작위 3-겹 교차검증 및 시점 기반 평가 모두에서 학습-평가 표본 간 성능의 일관성이 확인되며, 학교 변수에 기인한 과적합의 징후가 별도로 관측되지 않는다.

마지막으로, 본 연구에서는 특정 변수에 대한 결측치가 포함된 표본의 경우, 이를 제외하지 않고 해당 변수의 값에 '결측'에 해당하는 값을 할당한 뒤 그대로 모형의 학습에 사용하였다. 구체적으로, 범주형 변수의 결측치에는 '관측안됨'이라는 별도의 범주를 부여하였고, 수치형 변수의 결측치에는 해당 변수의 최소값보다 작은 특수값을 부여하여 결과적으로 XGBoost의 내장 결측 처리 방식과 동일한 분기가 이루어지도록 하였다. 다음 장에서는 예측 모형이 실제 가격을 얼마나 정확하고 안정적으로 추정하는지에 대한 평가 결과를 검토한다.

IV. 기계학습 기반 모형의 정확성 및 안정성 평가

1. 평가 개요 및 지표

본 연구는 XGBoost 알고리즘에 기반한 자동산정모형을 활용하여 개별 아파트의 가격을 추정하고, 이를 통해 전체 시장의 가격 분포와 관련 지표들을 도출하는 것을 목적으로 한다. 이러한 분석은 자동산정모형의 성능에 대한 신뢰를 전제로 이루어지며, 만약 모형의 예측 정확성이 일정 수준에 미치지 못할 경우, 해당 모형을 기반으로 한 추정 결과 역시 신뢰하기 어려울 것이다. 그러므로, 결과 자체를 제시하기에 앞서 본 연구에서 사용된 모형의 성능을 평가하는 것이 필수적이다.⁶⁾

따라서, 본 장에서는 자동산정모형을 통한 추정결과를 제시하기 이전에, 먼저 XGBoost 알고리즘에 기반한 자동산정 모형을 통해 분석하려는 자료에 대해 실제로 예측을 수행하고 그 정확성과 편향성을 검증한다. 본 연구에서는 이를 위해 3겹 교차 검증(three-fold cross-validation)을 실시하였다. 본 연구에서는 2012년 3월 1일부터 2024년 4월 1일 사이에 이루어진 아파트 실거래 자료를 각 지역별로 임의 분할 방식에 따라 3개의 하위 그룹으로 나누며, 각 반복에서 하나의 그룹을 평가 표본으로 설정하고 나머지 두 개를 학습 표본으로 사용하여 순차적으로 세 그룹 모두에 대한 평가를 수행하였다. 이와 같은 교차 검증은 모형의 일반화 능력을 보다 신뢰성 있게 판단할 수 있는 방법론으로 널리 사용된다. 자동산정모형을 활용한 주택 가격 예측 연구에서는 종종 시점을 기준으로 학습 표본과 평가 표본을 구분하기도 하지만, 본 연구는 과거 다양한 시점의 가격 분포를 재구성하려는 목적을 가지고 있으며, 이에 따라 전체 시점을 아우르는 임의 분할 방식이 실제 모형의 성능을 가장 잘 반영할 수 있는 평가 방식이라 판단된다.

6) 본 연구는 추정 모형의 변수 간 관계의 해석을 목적으로 하지 않으며, 모형의 추정치를 도구로 삼아 시장 전체의 가격 수준, 분포, 격차, 불균등성, 부담 등을 관찰하는 데에 분석의 초점을 둔다. 따라서 모형의 적절성에 대한 평가는 추정치의 정확성·편향성·안정성을 중심으로 이루어지며, 본 절에서는 이를 위한 평가 지표와 결과를 제시한다.

모형의 성능 평가는 다음과 같은 지표들을 통해 이루어졌다: 평균절대백분오차(MAPE), 산포계수(coefficient of dispersion, COD), 평균백분오차(mean percentage error, MPE), 오차 허용 구간 내 예측 비율(hit rate). 먼저, 가장 대표적인 오류 측정 지표인 MAPE는 예측값과 실제값 간의 차이를 백분율 기준으로 절대값 처리한 뒤 평균을 낸 지표로, 아래와 같이 표현된다.

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{\hat{p}_i - p_i}{p_i} \right| \quad \langle \text{식 1} \rangle$$

이때 n, \hat{p}_i, p_i 는 각각 표본수, i 번째 주택에 대한 예측값, i 번째 주택의 실제값이다. MAPE는 예측의 전반적인 정밀도를 쉽게 가늠하게 해준다. 이 지표는 개별 주택 단위에서 예측 오차가 얼마나 발생했는지를 직관적으로 보여주며, 가격 추정의 평균적 정확성을 가장 쉽게 파악할 수 있는 지표로 널리 활용된다.

한편 MAPE가 발생한 예측 오류의 평균적 수준을 측정하는 지표라면, COD와 MPE는 오류의 특성을 더 구체적으로 파악하는 데에 적합하다. 즉, 예측 오차는 비일관성(실제는 상대적으로 높는데 상대적으로 낮은 예측을 하는 경향)과 편향성(실제에 대해 일관적으로 더 높거나 낮은 값을 내는 경향)과 모두에 의해 발생할 수 있는데, COD는 비일관성(또는 불균형성)을 MPE는 편향성을 측정하는 데에 적합하다.

COD는 국제평가관협회(International Association of Assessing Officers, IAAO)가 대량평가 모형의 표준 평가 지표로 권고하는 산포 지표(IAAO, 2017)로, 추정가격과 실거래가의 비(sales ratio)가 그 중앙값으로부터 평균적으로 얼마나 벗어나 있는지를 백분율로 측정한다. 이때 COD를 계산하는 식은 아래와 같이 표현된다.

$$COD = \frac{100}{SR_m} \left[\frac{\sum_{i=1}^n |SR_i - SR_m|}{n} \right], \quad \langle \text{식 2} \rangle$$

이때 SR_i, SR_m, n 은 각각 i 번째 예측값과 실제값의 비율, 예측값-실제값 비율의 중앙값, 표본 크기를 의미한다. COD는 값이 작을수록 추정의 일관성이 높음을 의미하며, IAAO(2017)는 주거용 부동산의 경우 일반적으로 15 이하, 그 중에서도 비교적 동질적인 시장(예: 아파트)에서는 10 이하의 COD 값을 정확성-일관성이 양호한 수준으로 권고한다. MAPE가 평균적 오차의 절대적 크기를 측정한다면, COD는 sales ratio의 중앙값을 기준으로 한 산포의 정도를 측정하므로, 예측이 일관적으로 편향되어 있더라도(즉 모든 추정값이 일정한 비율로 과대 또는 과소 평가되더라도) 그 편향이 표본

전반에 일관적으로 나타나는 한 COD 값은 작게 유지될 수 있다는 특징이 있다. 즉, COD는 모형의 편향 여부와는 별개로 추정의 일관성·정밀도를 평가하는 데 적합하다.

한편 MPE(평균백분오차)는 예측값과 실제값 간의 차이를 백분율로 나타내되 그 부호를 유지한 채 평균을 계산하는 지표로 아래와 같이 표현된다.

$$MPE = \frac{100}{n} \sum_{i=1}^n \frac{\hat{p}_i - p_i}{p_i} \quad \langle \text{식 3} \rangle$$

이때 n, \hat{p}_i, p_i 는 각각 표본수, i 번째 주택에 대한 예측값, i 번째 주택의 실제값이다. MPE는 오차의 크기 보다는 예측값이 체계적으로 과소평가되었는지, 혹은 과대평가되었는지를 판단하는 데 사용될 수 있다.

특히 본 연구와 같이 시장의 평균 가격, 분위수, 지역별 통계량 등 집단 단위의 지표를 생성하려는 목적에서는, 개별 예측에 다소의 오차가 있더라도 MPE가 0에 가까우면 예측 오차들이 서로 상쇄되어 집단의 평균적 특성은 왜곡 없이 잘 반영될 수 있다. 예를 들어, 어떤 지역 내 아파트의 가격에 대한 개별 예측치들이 실제보다 다소 높거나 낮게 나온다고 하더라도, 만약 그 예측치의 편향성(MPE가 0에 가까우면)이 낮다면 예측 오류는 평균적으로 상쇄되어 지역의 평균치는 실제의 평균에 가깝게 관측될 수 있다. 반대로 MPE가 일정하게 뚜렷한 양의 값을 가진다면, 예측값은 일관되게 실제보다 높게 나타나며, 그 집단의 평균 가격이나 전체 분포의 중심이 실제보다 상향 왜곡되는 문제가 발생한다. 따라서 본 연구와 같이 개별주택에 대한 추정을 통해 지역별 평균이나 분포 등 통계량을 관측하려는 경우, 특히 예측모형의 MPE가 신뢰할만큼 낮아야 할 것이다. 마지막으로 hit rate는 예측값이 실제값과 비교하여 특정 허용 오차 범위($\pm 10\%$, $\pm 20\%$, $\pm 30\%$) 이내에 포함되는 비율을 측정하는 것이다. 이는 연구자가 허용 가능한 오차 기준에 따라 예측 결과의 수용 가능성을 판단할 수 있도록 하는 보조적 지표가 될 수 있다.

2. 자동산정모형의 정확성 및 불편향성 평가 결과

〈표 4〉는 전국 17개 행정구역(시·도)를 대상으로 실시한 XGBoost 기반 자동산정모형의 성능 평가 결과를 요약한 것이다. 아파트 실거래 자료를 3겹 교차 검증 방식으로 나누어 총 세 차례의 반복 실행을 진행하였으므로, 각 지역마다 3회의 결과가 나타나있다. 따라서 본 연구에서 제시된 성능평가의 결과는 모형이 특정 지역이나 시점이 아닌 일정 수준 이상의 성능을 일관적으로 유지할 수 있는지에 대한 관측 결과를 동시에 제공한다.

〈표 4〉 XGBoost 기반 AVM의 성능평가(3-fold 평가)

지역	분할	MAPE	COD	MPE	<10%	<20%	<30%
서울	1	4.36	4.36	0.51	91.88	98.66	99.57
	2	4.30	4.29	0.46	91.88	98.63	99.56
	3	4.29	4.29	0.48	91.96	98.68	99.57
부산	1	5.25	5.25	0.56	86.85	97.37	99.10
	2	5.27	5.27	0.48	86.75	97.40	99.12
	3	5.29	5.29	0.51	86.89	97.37	99.07
대구	1	4.57	4.57	0.39	90.19	98.26	99.41
	2	4.57	4.57	0.36	90.17	98.32	99.42
	3	4.56	4.56	0.36	90.32	98.34	99.41
인천	1	4.27	4.27	0.32	91.52	98.64	99.59
	2	4.28	4.28	0.31	91.57	98.62	99.58
	3	4.27	4.27	0.27	91.67	98.63	99.59
광주	1	5.04	5.04	0.41	87.82	97.90	99.27
	2	5.04	5.05	0.40	87.66	97.86	99.27
	3	5.07	5.07	0.44	87.56	97.84	99.23
대전	1	4.48	4.48	0.31	90.59	98.66	99.62
	2	4.49	4.49	0.37	90.57	98.62	99.57
	3	4.46	4.46	0.27	90.66	98.70	99.63
울산	1	5.27	5.27	0.47	86.47	97.43	99.21
	2	5.27	5.27	0.46	86.36	97.42	99.19
	3	5.37	5.37	0.58	86.06	97.20	99.06
세종	1	5.07	5.07	0.56	87.54	97.86	99.35
	2	5.06	5.06	0.45	87.87	97.90	99.27
	3	5.12	5.12	0.50	87.65	97.67	99.25
경기	1	3.95	3.95	0.27	93.25	99.03	99.69
	2	3.94	3.94	0.26	93.35	99.04	99.71
	3	3.97	3.97	0.30	93.24	99.02	99.70
강원	1	5.87	5.88	0.58	83.46	95.84	98.47
	2	5.83	5.83	0.64	83.57	95.89	98.50
	3	5.84	5.85	0.62	83.43	95.94	98.48
충북	1	5.72	5.73	0.55	84.72	96.22	98.49
	2	5.73	5.73	0.64	84.80	96.09	98.42
	3	5.72	5.72	0.59	84.91	96.21	98.53

〈표 4〉 계속

지역	분할	MAPE	COD	MPE	〈10%〉	〈20%〉	〈30%〉
충남	1	4.81	4.81	0.44	88.45	97.65	99.25
	2	4.82	4.82	0.37	88.42	97.68	99.19
	3	4.84	4.84	0.41	88.31	97.74	99.23
전북	1	5.96	5.97	0.55	82.84	96.18	98.70
	2	5.94	5.94	0.61	82.97	96.18	98.68
	3	5.96	5.96	0.64	82.97	96.16	98.71
전남	1	6.08	6.08	0.61	81.62	95.76	98.64
	2	6.06	6.06	0.65	81.89	95.83	98.58
	3	6.09	6.09	0.60	81.85	95.82	98.59
경북	1	6.67	6.67	0.76	79.97	94.54	97.94
	2	6.67	6.67	0.86	79.81	94.58	97.93
	3	6.73	6.73	0.85	79.81	94.57	97.86
경남	1	5.46	5.46	0.56	85.53	96.87	98.93
	2	5.44	5.44	0.51	85.58	96.99	99.03
	3	5.44	5.44	0.52	85.65	97.01	99.00
제주	1	7.33	7.33	1.29	77.31	93.15	97.18
	2	7.41	7.41	1.17	76.82	92.87	97.13
	3	7.24	7.24	0.89	77.46	93.24	97.32

주: 소수점 3자리 이하는 반올림되었다.

AVM, automated valuation model; MAPE, mean absolute percentage error; COD, coefficient of dispersion; MPE, mean percentage error.

먼저 평균절대백분오차(MAPE)를 기준으로 살펴보면, 수도권 3개 시·도(경기 3.95, 서울 4.32, 인천 4.27)와 대구(4.57), 대전(4.48), 충남(4.82) 등 6개 지역에서 4%대 또는 그 이하의 낮은 오차율이 관찰되었다. 그 외 광역시(부산, 광주, 울산, 세종)와 충북·경남·강원·전북 등 8개 지역은 5%대로 나타났으며, 전남(6.08), 경북(6.69), 제주(7.33)는 6%를 상회하였다. 비수도권 도지역, 특히 제주·경북·전남에서 상대적으로 큰 오차가 관측되는 것은 이들 지역의 시장 구조가 단지 간 이질성이 크고 거래 빈도가 낮아 학습 표본의 밀도가 낮은 데에서 비롯된 것으로 해석된다. 그럼에도 가장 오차율이 큰 제주조차 7%대에 머무르며 모든 지역에서 MAPE가 한 자릿수 범위 내에 머무른다는 점은, 모형의 예측 정밀도가 비교적 안정적인 수준에서 유지되고 있음을 보여준다.

이러한 경향은 산포계수(COD)에서도 유사하게 관측된다. 경기(3.95), 인천(4.27), 서울(4.31), 대전(4.48), 대구(4.57), 충남(4.82) 등 6개 지역은 COD가 5 미만으로 나타났으며, 부산·광주·세종·

울산·경남·충북·강원·전북 등 8개 지역은 5~6 수준이었다. 전남(6.08), 경북(6.69), 제주(7.33)의 3개 지역에서만 6을 다소 상회하였다. IAAO(2017)는 주거용 부동산에 대해 일반적으로 15 이하, 동질적 시장에서는 10 이하의 COD를 양호한 수준으로 권고하는데, 본 연구의 추정치는 가장 높은 제주(7.33)도 10 이하 수준을 큰 폭으로 하회하여 17개 시도 전체에서 IAAO 권고 기준을 충족한다.

또 한 가지 추정치의 활용 가능성을 높이는 지표인 평균백분오차(MPE)의 경우, 제주(1.12)를 제외한 16개 지역 모두에서 $\pm 1\%$ 이내로 수렴하였다. 이는 모형이 지역별로 체계적인 과대 또는 과소 예측 없이 거의 불편향적 예측을 수행함을 의미한다. 특히 경기(0.28), 인천(0.30), 대전(0.32), 대구(0.37)와 같이 거래량이 풍부한 수도권·광역시에서는 MPE가 0.3% 내외의 매우 낮은 수준에서 안정적으로 유지되었으며, 강원·전남·경북 등 비수도권 지역에서도 0.6%~0.8% 범위 내에서 예측 편향이 관리되었다. 유일하게 1%를 상회한 제주의 경우에도 절대 수준 자체는 1.12%에 머물러, 지역 단위 평균이나 분포 통계량의 산출에 실질적 왜곡을 초래할 정도의 편향은 아닌 것으로 평가된다. 이는 본 연구가 추구하는 '지역 단위 가격지표 산출'의 전제조건이 대체로 충족됨을 보여준다.

마지막으로 예측값과 실제값의 차이가 일정 수준 이하로 유지되는지를 평가하는 hit rate 분석에서는 지역 간 분포의 폭이 비교적 뚜렷하게 관측된다. $\pm 10\%$ 이내 hit rate는 수도권 3개 시도와 대구·대전에서 90% 이상(경기 93.3%, 서울 91.9%, 인천 91.6%, 대전 90.6%, 대구 90.2%)을 보였고, 그 외 광역시 및 충남·경남에서는 85%~89% 수준(부산 86.8%, 광주 87.7%, 울산 86.3%, 세종 87.7%, 충남 88.4%, 경남 85.6%)이었다. 반면 강원·충북·전북·전남 등 비수도권 도지역은 81%~85% 수준에 머물렀으며, 경북(79.9%)과 제주(77.2%)에서는 80% 미만의 가장 낮은 hit rate가 관측되었다. 허용 오차 기준을 완화하면 지역 간 격차는 빠르게 좁혀진다. $\pm 20\%$ 기준에서는 경북(94.6%)과 제주(93.1%)를 제외한 15개 지역에서 95% 이상이, $\pm 30\%$ 기준에서는 경북(97.9%)과 제주(97.2%)를 제외한 모든 지역에서 98.5% 이상이 포함된다. 이는 모형이 일부 비수도권 도지역에서 $\pm 10\%$ 라는 좁은 허용 오차 기준에 대해 다소 낮은 적중률을 보이긴 하나, 극단적인 예측 오차의 발생 자체는 전 지역에 걸쳐 매우 제한적임을 보여준다.

종합하면, 본 연구에서 사용된 XGBoost 기반 자동산정모형은 전국 17개 광역 시도 모두에서 IAAO(2017) 권고 기준을 충족하는 산포 수준과, 제주를 제외한 전 지역에서 $\pm 1\%$ 이내의 낮은 예측 편향을 일관적으로 보여주었다. 지역 간 정확성에서 일정한 격차가 존재하며(특히 거래 빈도가 낮고 단지 간 이질성이 큰 일부 비수도권 도지역에서 좁은 허용 오차 기준의 적중률이 다소 낮음) 이는 학습 표본의 밀도 차이로부터 비롯된 자연스러운 결과로 해석된다. 그럼에도 모든 지역에서 산포 및 편향성 기준이 일정 수준 이상의 안정성을 유지하고 있다는 점은, 본 모형의 추정이 개별 주택 가격 예측뿐 아니라 지역별 평균·분위수·분포 특성 등 집단 통계량 추정에도 활용 가능한 신뢰성을 갖춤을 시사한다.

3. 시점 기반 평가를 통한 강건성 검토

본 연구의 IV.2절에서 활용한 무작위 3-겹 교차검증은 본 연구의 분석 목적, 즉 과거 각 시점의 가격 분포를 재구성하기 위한 예측에 부합하는 평가 방식이다. 분석 모형은 시점 정보를 공변량으로 포함하고 있으며, 본 연구가 추정하고자 하는 가격은 학습 자료가 관측된 시점 범위 내(in-sample period)에 위치하는 가격이기 때문이다. 그러나 이러한 임의 분할 방식은 동일 시점의 거래가 학습 평가 표본에 동시에 포함될 수 있어, 성능 지표가 시점 누설(temporal leakage)에 의해 다소 낙관적으로 평가될 가능성을 완전히 배제하기는 어렵다. 따라서 본 절에서는 이러한 가능성에 대한 보완적 검토로서 시점 기반 평가(time-based holdout)를 추가로 수행한다.

평가 절차는 다음과 같다. 평가 시점은 1개월 단위로 설정하고, 학습 표본은 해당 평가 시점 직전 3년 간의 실거래 자료로 한정한다. 즉 평가 표본의 거래 시점은 학습 표본의 시점 범위와 중복되지 않으며, 이를 통해 시점 누설의 영향을 원천적으로 통제한다. 본 절에서는 분석 기간 전체에서 4개의 대표 평가 시점—2015년 1월, 2018년 1월, 2021년 1월, 2024년 1월—을 학습 기간과 동일한 3년 간격으로 선정하여 결과를 제시한다. 이 네 시점은 시장의 안정기, 상승 직후, 급등기 직후, 조정기를 각각 포괄하며, 다양한 시장 국면에서 모형 성능의 강건성을 함께 점검할 수 있게 한다.

〈표 5〉는 17개 시도별 시점 기반 평가 결과를 정리한 것이다. 결과를 요약하면 다음과 같다. 첫째, 17개 시도의 평균 MAPE는 2015년 1월 6.07%, 2018년 1월 6.24%, 2021년 1월 8.13%, 2024년 1월 6.61%로 나타났다. 이는 무작위 3-겹 교차검증 결과인 평균 4% 내외에 비해서는 다소 상승한 수준이나, IAAO(2017)가 동질적 시장에 대해 권고하는 산포 기준 등을 고려할 때 절대적 수준에서 여전히 양호한 범위 내에 위치한다.

둘째, 시점별 결과를 비교하면 2021년 1월의 평가 시점에서 성능이 가장 낮게 나타나는데, 이는 해당 시점이 2020~2021년의 급격한 가격 상승 구간에 위치한 것과 관련된다. 특히 울산(MAPE 11.34), 경북(9.12), 부산(9.10), 제주(9.74) 등 가격 변동성이 크게 나타난 일부 지역에서 9%~11% 대의 높은 MAPE가 관측된다. 즉, 학습 표본(2018년 1월~2020년 12월)이 후속 시기의 급격한 구조 변화를 사전적으로 반영하지 못하기 때문에 일시적인 성능 저하가 발생한 것으로 해석된다. 이는 시점을 가로지르는 외삽(extrapolation)이 모형의 일반화에 도전적인 과제임을 시사하나, 본 연구가 학습 시점 범위 내 시장의 가격 분포를 재구성하는 것을 목적으로 하고 있다는 점에서 본 연구의 분석 결과 자체의 신뢰성에 대한 직접적 위협으로 해석되지는 않는다. 한편 2024년 1월 시점의 부산(MAPE 10.17, MPE +4.14)에서도 비교적 큰 양의 편향이 관측되는데, 이는 학습 기간(2021년 1월~2023년 12월) 동안 부산 시장이 조정기에 들어선 후 평가 시점에서 일부 회복 또는 재상승의 신호가 발생하여 모형의 외삽이 어려웠던 결과로 해석된다.

〈표 5〉 XGBoost 기반 AVM의 성능평가(time-holdout 평가)

지역	시점	MAPE	COD	MPE	〈10%〉	〈20%〉	30%
서울	2015/01	4.00	3.84	-1.27	93.91	99.31	99.85
	2018/01	5.59	5.38	-1.80	85.09	97.95	99.32
	2021/01	6.51	6.07	-2.60	80.86	96.84	99.08
	2024/01	5.34	5.33	1.22	86.15	97.09	99.12
부산	2015/01	5.08	4.85	-1.94	88.58	98.43	99.55
	2018/01	6.35	6.35	-0.38	80.50	95.00	98.94
	2021/01	9.10	9.04	-1.70	65.62	91.64	96.91
	2024/01	10.17	10.19	4.14	77.58	92.92	95.47
대구	2015/01	6.04	5.31	-3.37	82.62	97.71	99.44
	2018/01	4.83	4.54	-2.01	89.51	99.08	99.72
	2021/01	7.60	7.42	-2.08	72.67	95.15	98.74
	2024/01	5.64	5.65	-0.04	84.46	97.65	99.59
인천	2015/01	4.56	4.34	-1.68	91.15	99.01	99.77
	2018/01	4.30	4.19	-1.26	92.75	98.97	99.76
	2021/01	6.14	5.69	-2.89	81.38	97.65	99.59
	2024/01	5.24	5.23	-0.44	86.82	98.10	99.64
광주	2015/01	6.41	5.95	-3.02	79.35	97.70	99.48
	2018/01	5.91	5.78	-1.16	85.40	97.29	98.41
	2021/01	7.04	7.01	-0.78	77.66	95.85	98.80
	2024/01	6.46	6.44	-1.30	79.63	96.56	99.63
대전	2015/01	4.76	4.53	-1.75	90.13	98.07	99.78
	2018/01	4.41	4.39	-0.74	91.38	99.19	99.94
	2021/01	7.01	6.48	-3.47	75.20	96.95	99.47
	2024/01	5.39	5.34	-1.17	85.70	98.42	99.61
울산	2015/01	6.02	5.54	-2.32	84.46	96.87	98.77
	2018/01	5.46	5.47	-0.46	86.30	98.72	99.74
	2021/01	11.34	11.06	-0.56	62.66	90.14	97.33
	2024/01	7.15	7.06	-1.24	78.91	95.09	97.91
세종	2015/01	5.62	5.49	-1.09	87.77	97.84	99.28
	2018/01	6.48	6.44	-2.19	80.71	94.64	98.93
	2021/01	8.28	8.12	-2.93	73.51	91.65	96.90
	2024/01	5.10	5.10	0.07	88.48	97.94	99.18
경기	2015/01	4.32	4.17	-1.62	92.02	98.89	99.78
	2018/01	5.02	4.87	-1.73	88.62	98.47	99.67
	2021/01	7.67	7.27	-3.35	72.03	94.89	99.04
	2024/01	4.96	4.96	-0.07	89.13	98.58	99.68

〈표 5〉 계속

지역	시점	MAPE	COD	MPE	〈10%	〈20%	30%
강원	2015/01	6.80	6.30	-3.18	78.57	95.54	98.77
	2018/01	6.41	6.29	-1.10	82.11	96.57	98.10
	2021/01	7.73	7.60	-1.96	72.75	94.60	98.08
	2024/01	6.94	6.95	-0.61	78.33	96.06	98.58
충북	2015/01	7.53	7.41	0.67	82.02	92.60	94.56
	2018/01	6.80	6.79	0.22	79.74	94.29	98.55
	2021/01	9.01	8.65	-2.55	66.58	93.00	97.01
	2024/01	5.95	5.88	-0.70	83.96	97.39	99.06
충남	2015/01	5.60	5.42	-1.72	86.03	96.80	99.26
	2018/01	6.35	6.19	-1.55	80.16	97.44	99.33
	2021/01	7.60	7.23	-2.94	73.91	94.68	98.41
	2024/01	6.69	6.68	-0.76	79.92	96.77	98.87
전북	2015/01	6.52	6.31	-1.92	79.59	95.98	98.90
	2018/01	6.19	6.00	-1.46	81.99	97.15	99.21
	2021/01	8.33	8.31	-0.82	69.28	93.31	98.09
	2024/01	7.29	7.31	-0.57	77.11	93.98	97.07
전남	2015/01	7.40	7.19	-1.76	75.05	93.44	97.57
	2018/01	7.59	7.51	-0.96	74.90	92.72	98.48
	2021/01	8.14	8.00	-2.22	70.88	93.01	98.53
	2024/01	8.34	8.37	0.17	73.30	93.13	97.00
경북	2015/01	7.48	7.32	-1.87	76.17	93.60	97.75
	2018/01	7.98	7.95	-1.16	72.72	93.59	98.09
	2021/01	9.12	8.78	-2.61	66.98	90.94	97.19
	2024/01	8.26	8.26	-0.83	71.86	92.72	97.37
경남	2015/01	5.15	4.91	-2.02	87.62	98.15	99.57
	2018/01	8.06	8.08	1.54	72.43	91.62	96.71
	2021/01	7.86	7.61	-1.95	73.06	94.22	98.29
	2024/01	6.62	6.60	-0.94	79.39	96.42	98.87
제주	2015/01	9.88	8.66	-5.51	53.73	91.07	98.36
	2018/01	8.35	8.30	-0.88	70.85	92.38	97.31
	2021/01	9.74	9.16	-3.97	60.05	88.94	98.24
	2024/01	6.83	6.76	-1.77	76.73	96.04	99.50

AVM, automated valuation model; MAPE, mean absolute percentage error; COD, coefficient of dispersion; MPE, mean percentage error.

셋째, 평균백분오차(MPE)의 절대값을 17개 시·도 평균으로 보면 2018년 1월(1.21%)과 2024년 1월(0.94%)에는 1% 안팎으로 낮게 유지되었으나, 2015년 1월(2.16%)과 2021년 1월(2.32%)에는 다소 상승하였다. 이는 시점 기반 평가에서는 무작위 교차검증에 비해 편향성이 다소 확대됨을 보여준다. 개별 지역·시점 단위에서도 $\pm 3\%$ 를 초과하는 편향이 8건(68건 중) 관측되며, 특히 제주 2015년 1월(-5.51%)과 부산 2024년 1월(+4.14%)에서 가장 큰 편향이 발생하였다. 그럼에도 $\pm 3\%$ 초과 사례의 대부분은 위에서 언급한 급격한 시장 국면 변화(2020~2021년 급등기, 2024년 부산 시장의 부분적 반전)에 집중되어 있으며, 여타 시점·지역에서는 편향이 $\pm 3\%$ 이내에서 관리된다.

넷째, 허용 오차 $\pm 30\%$ 이내의 적중률은 충북 2015년 1월(94.56%)과 부산 2024년 1월(95.47%) 두 조합을 제외한 모든 시점·지역 조합에서 96% 이상을 유지하며, 78%(53/68)에 해당하는 조합에서 98%를 상회한다. 이는 외삽 조건에서 일부 시점·지역의 평균 정확도가 다소 저하될지라도, 극단적 예측 오차의 발생 빈도 자체는 시점 기반 평가 조건에서도 제한적임을 보여준다(표 5).

종합하면, 시점 누설을 통제한 보다 엄격한 평가 조건에서도 본 모형은 시장의 다양한 국면에 걸쳐 일정 수준 이상의 정확성을 유지한다. 시점 기반 평가에서는 무작위 교차검증에 비해 편향성(MPE)과 평균 오차(MAPE)가 다소 확대되며, 이는 시장의 구조 변화 직후 시점(특히 2020~2021년 급등기 직후)에서 두드러진다. 그러나 이러한 성능 저하는 학습 시점 범위 밖으로의 외삽 조건에서 발생하는 자연스러운 현상이며, 본 연구가 IV.2절에서 채택한 평가 방식은 분석 목적에 부합한다는 판단을 보완적으로 뒷받침한다. 또한 $\pm 30\%$ 적중률 측면에서는 시점 기반 평가 조건에서도 압도적 다수의 조합이 96% 이상의 안정성을 유지함이 확인되어, 본 연구의 추정 결과가 시점 누설 가능성으로부터 비롯할 수 있는 낙관적 평가에 뚜렷하게 의존하지 않음을 시사한다. 이러한 결과를 바탕으로, 다음 장에서는 본 모형을 활용한 각 지역의 가격 수준 및 분포 특성을 비교·분석한다.

V. 분석 결과

1. 지역별 평균 가격의 추이

본 연구에서는 구축된 부동산정보형을 활용하여, 2012년 3월 1일부터 2024년 4월 1일까지 약 12년간 전국 및 17개 시·도 아파트 단지의 각 평형별 가격을 분기별로 추정하였으며, 이를 통해 전국 및 지역별 가중 평균(각 평형별 세대수를 기준한 가중평균)을 도출하였다. 이는 거래된 주택만의 가격 변화를 추적하는 전통적 지표와 달리, 동일하게 고정된 전수 표본의 가격을 각 시점마다 추정함으로써 시장 전체의 가격 수준을 직접 관측하는 접근이다. 이는 단순히 각 시점에 거래가 이루어진 주택의 가격이 아닌 모든 주택을 대상으로 한 실질적인 시장 전체의 평균가격을 도출한 데에 그 의미가 있다. 다만 이는 재고 주택의

특성 변화를 고려한 통계적 의미의 가격지수가 아니라 전수 추정에 기반한 가중평균값임을 미리 밝혀둔다.

먼저, <표 6>은 전국 및 지역별 평균 가격 수준을 보여준다. 2024년 1분기말 기준, 전국 17개 시도의 공급면적 기준 평균 아파트 가격은 약 470만 원/m²으로 나타난다. 서울이 약 1,121만 원/m²으로 가장 높은 평균가격을 기록했으며, 그 다음으로는 경기도(약 520만 원/m²)가 위치했다. 가장 평균 가격이 낮은 지역은 경북(약 181만 원/m²)이었으며 그 다음으로 낮은 지역은 전남(186만 원/m²)이었다. 전국 평균 아파트 가격은 2012년 1분기말 약 240만 원/m² 수준이었으나, 2024년 1분기말까지 전반적으로 상승했던 것으로 나타났다. 특히, 시기별 추이를 보면 2012년부터 2018년까지는 완만한 상승세를 보이다가 2018년 후반부터 2021년 상반기에 걸쳐 급등하고 2022년 이후 천천히 조정되는 형태를 보여준다.

<표 6> 아파트 평균가격(단위: 만 원/m²)

시점	전국	서울	부산	대구	인천	광주	대전	울산	세종	경기	강원	충북	충남	전북	전남	경북	경남	제주
2012/03	243	471	218	170	228	149	205	204	188	266	128	151	153	142	122	124	182	167
2012/06	240	464	216	171	228	151	202	206	190	262	127	153	155	141	123	125	178	167
2012/09	238	454	214	176	224	152	200	206	203	258	128	153	157	139	123	128	177	171
2012/12	237	453	213	177	224	152	201	204	217	256	128	153	158	138	122	132	176	170
2013/03	239	455	214	182	222	155	203	205	220	256	129	155	161	138	127	134	180	170
2013/06	241	459	214	188	224	157	203	205	224	258	128	154	163	137	125	137	179	172
2013/09	244	460	215	196	229	161	206	209	226	261	130	158	163	140	126	142	183	173
2013/12	247	465	218	205	231	163	208	212	238	264	131	159	166	140	125	145	185	177
2014/03	251	474	221	211	235	167	207	214	238	269	132	163	168	142	127	148	188	183
2014/06	252	474	224	215	232	167	208	218	238	270	133	166	169	141	127	150	190	190
2014/09	257	483	227	221	241	172	209	220	240	273	134	171	174	143	128	153	194	194
2014/12	261	490	230	230	245	178	208	226	241	278	136	175	176	143	128	157	197	203
2015/03	266	500	234	240	245	186	210	231	250	283	139	175	178	144	130	161	200	210
2015/06	273	514	240	253	253	192	211	236	261	290	142	178	179	146	131	164	205	225
2015/09	280	529	248	266	261	198	214	245	263	298	146	179	180	147	135	168	208	248
2015/12	283	540	254	257	264	198	213	252	269	301	150	176	181	147	137	166	209	266
2016/03	284	540	258	253	265	201	213	253	282	301	151	176	180	148	139	163	209	277
2016/06	287	553	261	249	268	200	213	254	287	306	154	175	179	149	140	160	208	287
2016/09	296	581	272	250	275	202	216	255	280	315	160	176	179	152	144	161	210	300
2016/12	301	589	286	253	283	202	220	256	291	320	165	174	179	154	147	161	210	307
2017/03	303	597	290	253	281	204	221	255	316	320	167	171	180	155	148	159	209	311
2017/06	308	620	293	256	281	207	223	254	324	326	171	174	179	156	151	158	206	314
2017/09	316	646	295	266	287	210	226	252	335	334	174	174	179	158	156	159	204	315
2017/12	320	674	293	270	289	211	227	249	321	335	173	174	178	159	158	159	202	316
2018/03	331	729	292	275	287	216	229	245	346	344	173	173	176	158	159	157	197	319
2018/06	333	746	287	280	286	220	230	237	348	347	171	173	174	158	162	154	195	314

〈표 6〉 계속

시점	전국	서울	부산	대구	인천	광주	대전	울산	세종	경기	강원	충북	충남	전북	전남	경북	경남	제주
2018/09	349	822	288	284	289	232	233	233	354	360	171	167	173	157	164	154	193	315
2018/12	350	816	286	289	288	242	243	230	356	365	169	166	175	157	165	151	191	314
2019/03	346	802	286	290	288	241	247	226	365	362	166	163	174	156	165	149	188	309
2019/06	350	823	285	290	289	240	254	226	377	363	163	161	173	154	166	149	186	310
2019/09	360	870	285	294	292	242	262	226	374	373	164	161	174	156	167	149	185	306
2019/12	376	926	299	300	299	242	279	234	393	389	164	163	178	157	169	150	189	308
2020/03	394	963	307	306	319	246	298	243	453	418	166	169	182	160	171	153	193	309
2020/06	400	966	310	308	337	249	311	246	473	428	169	177	184	161	174	154	194	310
2020/09	429	1,063	326	319	344	251	332	258	599	460	174	183	192	166	178	157	199	314
2020/12	452	1,098	372	353	358	263	348	291	614	486	179	191	200	176	183	166	215	323
2021/03	487	1,169	395	375	387	275	374	306	631	539	192	205	212	183	188	175	227	343
2021/06	505	1,207	408	378	423	281	388	307	617	565	198	215	218	189	189	180	232	366
2021/09	538	1,272	431	381	472	298	408	319	622	614	212	227	230	200	194	187	244	397
2021/12	544	1,280	433	376	485	317	411	320	597	618	224	236	238	208	199	192	252	404
2022/03	534	1,260	427	363	470	318	394	321	572	603	227	236	237	210	198	192	254	406
2022/06	531	1,255	425	355	459	320	388	317	562	598	233	236	237	212	198	192	255	413
2022/09	503	1,191	407	339	432	307	370	308	515	557	232	230	232	211	195	189	251	409
2022/12	470	1,105	382	325	395	293	347	294	487	516	221	220	221	201	187	183	237	395
2023/03	461	1,090	370	310	384	287	339	288	483	507	220	217	216	196	185	181	233	391
2023/06	468	1,110	372	310	388	287	346	290	498	517	220	218	216	196	186	181	234	394
2023/09	477	1,139	373	314	395	290	354	294	504	529	223	222	219	199	188	183	236	392
2023/12	472	1,125	369	314	393	287	353	293	494	521	222	223	218	197	186	182	234	390
2024/03	470	1,121	365	309	393	282	351	293	483	520	221	222	216	197	186	181	231	387

주: 소수점 이하는 반올림되었다.

결과에 의하면, 해당 기간 동안 각 지역의 가격 변화에 상당한 차이가 있었음을 알 수 있다. 〈표 7〉은 〈표 6〉의 결과를 토대로 가격의 변화율을, 〈표 8〉은 최초 시점으로부터의 누적가격 변화율을 보여준다. 특히, 수도권 지역과 세종시에서는 2018년에서 2021년 사이의 가격 상승폭이 매우 컸다는 것을 알 수 있으며, 2022년 이후에는 대부분의 지역에서 조정 국면에 접어든 것으로 관찰된다. 2012년과 2024년 사이의 누적 상승률 상위 지역은 각각 서울(138%), 경기(95%), 세종(156%), 제주(131%)이며, 이들 지역은 공통적으로 인구 증가, 또는 수도권 연계성이 높은 특성을 가진다. 특히, 서울의 단위 면적 당 아파트 평균가격은 2012년 초 기준으로 470만 원이었으나, 2024년 초에는 약 1,121만 원에 달해 세종시를 제외하고는 해당 기간 가장 높은 누적성장률을 보여준다. 반면 하위 지역은 경남(27%), 충남(40%), 울산(43%), 경북(46%), 전남(52%)으로 나타났다.

〈표 7〉 아파트 평균가격 변화율(단위: %)

시점	전국	서울	부산	대구	인천	광주	대전	울산	세종	경기	강원	충북	충남	전북	전남	경북	경남	제주
2012/03	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2012/06	-0.91	-1.44	-0.92	0.82	-0.08	1.70	-1.58	0.98	1.20	-1.64	-0.64	1.71	0.96	-0.27	0.73	1.57	-1.83	-0.08
2012/09	-1.09	-2.15	-0.88	2.69	-1.66	0.54	-1.10	-0.17	6.62	-1.67	0.58	-0.24	1.33	-1.43	0.03	2.16	-0.65	2.19
2012/12	-0.25	-0.25	-0.64	0.63	0.05	-0.01	0.53	-0.88	6.91	-0.80	0.19	-0.02	0.51	-0.91	-0.16	2.61	-0.87	-0.67
2013/03	0.68	0.49	0.42	2.96	-0.91	1.80	1.27	0.54	1.55	0.11	0.99	1.12	2.07	0.18	3.66	1.85	2.18	0.03
2013/06	0.81	0.88	0.10	3.44	1.03	1.32	-0.10	0.11	1.61	0.91	-0.42	-0.17	1.03	-0.58	-1.39	2.48	-0.08	1.21
2013/09	1.17	0.23	0.68	4.08	1.93	2.82	1.33	1.95	0.91	1.08	1.29	2.32	0.23	1.59	0.98	3.51	2.03	0.93
2013/12	1.28	1.09	1.32	4.40	1.13	0.96	0.85	1.03	5.44	1.11	0.45	1.05	1.99	0.51	-0.89	1.83	1.00	2.00
2014/03	1.77	1.96	1.39	2.87	1.79	2.78	-0.13	1.12	-0.34	1.77	1.25	2.04	1.30	0.87	1.57	2.36	1.63	3.32
2014/06	0.47	-0.01	1.18	2.09	-1.58	-0.25	0.16	1.96	-0.01	0.62	0.30	2.17	0.61	-0.50	0.11	1.37	1.18	4.18
2014/09	1.80	1.93	1.53	3.03	4.18	2.79	0.86	0.96	1.02	1.15	1.24	2.60	2.62	1.46	0.06	1.89	2.05	1.93
2014/12	1.62	1.48	1.09	4.07	1.36	3.92	-0.70	2.83	0.28	1.53	0.92	2.62	1.38	-0.20	0.00	2.85	1.64	4.75
2015/03	1.84	1.92	2.00	3.94	0.07	4.17	0.90	2.10	3.80	1.81	2.26	-0.06	0.70	0.69	2.30	2.29	1.47	3.49
2015/06	2.66	2.79	2.53	5.59	3.24	3.20	0.46	2.18	4.32	2.70	2.49	1.46	0.57	1.64	0.11	1.72	2.22	7.04
2015/09	2.82	3.07	3.32	5.01	3.38	3.28	1.38	3.72	0.95	2.56	2.69	0.70	0.96	0.80	3.52	2.76	1.70	10.13
2015/12	1.01	1.90	2.19	-3.10	1.05	0.19	-0.46	2.69	2.34	1.03	2.94	-1.59	0.28	-0.18	1.59	-1.27	0.34	7.32
2016/03	0.25	0.10	1.57	-1.55	0.59	1.34	-0.07	0.49	4.75	0.21	0.53	0.05	-0.45	0.91	1.24	-2.08	0.14	4.05
2016/06	1.19	2.40	1.17	-1.64	1.07	-0.69	0.39	0.42	1.85	1.52	1.78	-0.45	-0.40	0.53	0.80	-1.48	-0.48	3.52
2016/09	3.02	5.11	4.45	0.35	2.40	1.21	1.39	0.57	-2.67	2.95	4.01	0.59	-0.32	1.97	2.48	0.44	0.78	4.50

〈표 7〉 계속

시점	전국	서울	부산	대구	인천	광주	대전	울산	세종	경기	강원	충북	충남	전북	전남	경북	경남	제주
2016/12	1.60	1.41	5.16	1.23	2.87	-0.02	1.52	0.41	3.97	1.46	3.32	-1.14	0.27	1.62	2.46	-0.12	0.11	2.37
2017/03	0.60	1.35	1.39	0.12	-0.52	1.08	0.52	-0.39	8.59	0.13	1.20	-1.62	0.25	0.54	0.52	-0.81	-0.28	1.45
2017/06	1.83	3.77	1.09	0.91	-0.02	1.27	1.20	-0.73	2.79	1.88	2.22	1.26	-0.46	0.61	2.17	-1.13	-1.61	0.79
2017/09	2.44	4.26	0.56	3.82	2.14	1.51	0.98	-0.69	3.38	2.32	1.76	0.36	0.30	1.07	2.99	0.78	-0.75	0.57
2017/12	1.41	4.36	-0.81	1.60	0.50	0.59	0.55	-1.23	-4.19	0.50	-0.39	-0.32	-0.57	0.58	1.76	0.08	-1.36	0.11
2018/03	3.29	8.11	-0.28	2.06	-0.47	2.09	0.86	-1.65	7.80	2.63	0.14	-0.44	-1.35	-0.15	0.67	-1.48	-2.11	0.96
2018/06	0.84	2.29	-1.58	1.55	-0.37	2.04	0.60	-3.04	0.61	0.89	-1.25	-0.14	-0.82	-0.35	1.86	-1.93	-1.17	-1.42
2018/09	4.67	10.23	0.36	1.74	1.14	5.69	1.22	-1.65	1.69	3.84	-0.08	-3.23	-0.69	-0.61	0.80	-0.08	-0.90	0.19
2018/12	0.18	-0.75	-0.75	1.50	-0.32	4.18	4.33	-1.33	0.60	1.14	-1.38	-0.46	0.97	0.10	0.70	-1.47	-1.13	-0.08
2019/03	-0.99	-1.73	-0.04	0.30	-0.12	-0.53	1.58	-2.00	2.29	-0.82	-1.94	-2.13	-0.62	-0.91	0.38	-1.77	-1.47	-1.66
2019/06	1.02	2.60	-0.42	0.06	0.22	-0.37	2.92	0.11	3.34	0.46	-1.71	-0.87	-0.20	-0.85	0.40	0.23	-0.95	0.16
2019/09	3.03	5.79	0.21	1.56	1.22	0.69	3.23	0.26	-0.63	2.72	0.68	0.03	0.51	0.85	0.48	0.21	-0.56	-1.15
2019/12	4.44	6.45	4.61	2.06	2.44	0.31	6.40	3.50	4.83	4.36	0.02	1.00	2.28	0.63	1.32	0.66	2.03	0.55
2020/03	4.75	3.95	2.76	1.75	6.51	1.64	6.93	3.62	15.31	7.35	1.27	3.57	2.10	2.23	0.86	1.53	1.96	0.22
2020/06	1.57	0.27	0.99	0.68	5.63	1.03	4.25	1.44	4.39	2.48	2.02	4.92	1.00	0.75	2.15	1.01	0.47	0.34
2020/09	7.26	10.08	5.32	3.72	2.21	0.87	6.63	4.63	26.70	7.36	2.95	3.28	4.55	2.85	2.25	1.66	2.87	1.45
2020/12	5.35	3.32	13.99	10.52	3.99	4.82	4.87	12.80	2.61	5.68	2.59	4.35	4.21	5.80	2.90	5.76	7.85	2.86
2021/03	7.67	6.48	6.26	6.32	8.22	4.60	7.53	5.17	2.66	10.95	7.45	7.26	5.60	4.34	2.53	5.72	5.59	6.12
2021/06	3.70	3.17	3.11	0.79	9.19	2.05	3.69	0.50	-2.17	4.82	3.16	4.82	3.17	3.22	0.40	2.73	2.00	6.72

〈표 7〉 계속

시점	전국	서울	부산	대구	인천	광주	대전	울산	세종	경기	강원	충북	충남	전북	전남	경북	경남	제주
2021/09	6.48	5.40	5.63	0.69	11.67	6.25	5.03	3.84	0.77	8.68	7.12	5.98	5.36	5.55	2.73	4.04	5.53	8.40
2021/12	1.07	0.67	0.54	-1.17	2.55	6.36	0.78	0.46	-4.01	0.64	5.30	3.89	3.34	4.19	2.43	2.63	3.32	1.88
2022/03	-1.76	-1.61	-1.35	-3.41	-3.00	0.22	-3.99	0.15	-4.09	-2.39	1.67	-0.29	-0.48	1.09	-0.26	-0.06	0.76	0.49
2022/06	-0.65	-0.39	-0.38	-2.22	-2.36	0.56	-1.56	-1.18	-1.85	-0.95	2.64	0.18	0.29	0.93	0.03	-0.15	0.37	1.61
2022/09	-5.11	-5.09	-4.30	-4.70	-5.86	-4.00	-4.58	-2.94	-8.36	-6.72	-0.59	-2.62	-2.17	-0.62	-1.58	-1.33	-1.88	-0.86
2022/12	-6.66	-7.17	-6.10	-4.14	-8.48	-4.64	-6.32	-4.67	-5.35	-7.41	-4.63	-4.12	-4.98	-4.81	-4.09	-3.19	-5.28	-3.46
2023/03	-1.87	-1.41	-3.21	-4.56	-2.99	-1.94	-2.44	-1.89	-0.78	-1.74	-0.48	-1.59	-2.17	-2.55	-0.85	-1.42	-1.75	-0.90
2023/06	1.44	1.85	0.58	0.17	1.17	-0.12	2.20	0.71	3.09	1.85	-0.18	0.40	0.25	-0.02	0.32	0.54	0.46	0.70
2023/09	2.06	2.65	0.21	1.26	1.91	1.24	2.43	1.48	1.18	2.35	1.30	1.80	1.21	1.63	0.97	0.92	0.82	-0.62
2023/12	-1.16	-1.28	-1.03	-0.08	-0.73	-1.32	-0.47	-0.34	-2.03	-1.53	-0.15	0.39	-0.61	-1.04	-0.82	-0.87	-0.85	-0.37
2024/03	-0.43	-0.33	-1.18	-1.49	0.04	-1.59	-0.44	-0.07	-2.33	-0.15	-0.56	-0.03	-0.90	0.05	-0.13	-0.30	-1.19	-0.87

주: 1) 직전 시점으로부터의 변화율(분기 변화율)을 의미한다.
 2) 소수점 2자리 이하는 반올림되었다.

〈표 8〉 아파트 평균가격 누적변화율(단위: %)

시점	전국	서울	부산	대구	인천	광주	대전	울산	세종	경기	강원	충북	충남	전북	전남	경북	경남	제주
2012/03	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2012/06	-0.9	-1.4	-0.9	0.8	-0.1	1.7	-1.6	1.0	1.2	-1.6	-0.6	1.7	1.0	-0.3	0.7	1.6	-1.8	-0.1
2012/09	-2.0	-3.6	-1.7	3.5	-1.7	2.2	-2.7	0.8	7.9	-3.3	-0.1	1.5	2.3	-1.7	0.8	3.8	-2.5	2.1
2012/12	-2.2	-3.8	-2.4	4.2	-1.7	2.2	-2.1	-0.1	15.4	-4.1	0.1	1.4	2.8	-2.6	0.6	6.5	-3.3	1.4
2013/03	-1.6	-3.3	-2.0	7.3	-2.6	4.1	-0.9	0.5	17.2	-3.9	1.1	2.6	4.9	-2.4	4.3	8.4	-1.2	1.4

〈표 8〉 계속

시점	전국	서울	부산	대구	인천	광주	대전	울산	세종	경기	강원	충북	충남	전북	전남	경북	경남	제주
2013 /06	-0.8	-2.5	-1.9	10.9	-1.6	5.4	-1.0	0.6	19.0	-3.1	0.7	2.4	6.0	-3.0	2.8	11.1	-1.3	2.7
2013 /09	0.4	-2.3	-1.2	15.5	0.3	8.4	0.3	2.5	20.1	-2.0	2.0	4.8	6.3	-1.4	3.8	15.0	0.7	3.6
2013 /12	1.7	-1.2	0.1	20.6	1.4	9.5	1.2	3.6	26.7	-0.9	2.4	5.9	8.4	-0.9	2.9	17.1	1.7	5.7
2014 /03	3.5	0.8	1.5	24.0	3.3	12.5	1.0	4.7	26.2	0.8	3.7	8.0	9.8	-0.1	4.5	19.9	3.4	9.2
2014 /06	4.0	0.7	2.7	26.6	1.6	12.2	1.2	6.8	26.2	1.4	4.0	10.4	10.5	-0.6	4.7	21.5	4.6	13.8
2014 /09	5.8	2.7	4.3	30.5	5.9	15.4	2.1	7.8	27.5	2.6	5.3	13.3	13.4	0.9	4.7	23.8	6.7	16.0
2014 /12	7.5	4.2	5.4	35.8	7.3	19.9	1.4	10.9	27.9	4.2	6.3	16.2	14.9	0.7	4.7	27.4	8.5	21.5
2015 /03	9.5	6.2	7.5	41.1	7.4	24.9	2.3	13.2	32.7	6.1	8.7	16.2	15.7	1.4	7.1	30.3	10.1	25.7
2015 /06	12.4	9.2	10.2	49.0	10.9	28.9	2.7	15.7	38.5	8.9	11.4	17.9	16.4	3.0	7.2	32.5	12.5	34.6
2015 /09	15.6	12.5	13.9	56.5	14.6	33.1	4.2	20.0	39.8	11.7	14.4	18.7	17.5	3.8	11.0	36.2	14.5	48.2
2015 /12	16.8	14.7	16.4	51.6	15.8	33.4	3.7	23.2	43.0	12.9	17.8	16.8	17.8	3.7	12.8	34.5	14.8	59.1
2016 /03	17.1	14.8	18.2	49.3	16.5	35.1	3.6	23.8	49.8	13.1	18.4	16.8	17.3	4.6	14.2	31.7	15.0	65.5
2016 /06	18.5	17.5	19.6	46.8	17.8	34.2	4.0	24.3	52.6	14.8	20.5	16.3	16.9	5.2	15.1	29.7	14.5	71.3
2016 /09	22.0	23.5	24.9	47.3	20.6	35.8	5.4	25.0	48.5	18.2	25.3	17.0	16.5	7.2	17.9	30.3	15.3	79.1
2016 /12	24.0	25.3	31.4	49.2	24.0	35.8	7.0	25.6	54.4	19.9	29.5	15.7	16.8	9.0	20.8	30.1	15.5	83.3
2017 /03	24.7	27.0	33.2	49.3	23.4	37.3	7.6	25.1	67.7	20.1	31.0	13.8	17.1	9.6	21.5	29.1	15.1	86.0
2017 /06	27.0	31.7	34.6	50.7	23.4	39.0	8.9	24.2	72.4	22.3	33.9	15.2	16.5	10.2	24.1	27.6	13.3	87.4
2017 /09	30.1	37.4	35.4	56.4	26.0	41.1	10.0	23.3	78.2	25.2	36.3	15.6	16.9	11.4	27.8	28.6	12.4	88.5
2017 /12	31.9	43.3	34.3	59.0	26.6	42.0	10.6	21.8	70.7	25.8	35.8	15.3	16.2	12.1	30.1	28.7	10.9	88.7

〈표 8〉 계속

시점	전국	서울	부산	대구	인천	광주	대전	울산	세종	경기	강원	충북	충남	전북	전남	경북	경남	제주
2018 /03	36.3	55.0	33.9	62.2	26.0	44.9	11.5	19.8	84.0	29.1	36.0	14.8	14.7	11.9	30.9	26.8	8.6	90.5
2018 /06	37.4	58.5	31.8	64.7	25.6	47.9	12.2	16.1	85.2	30.3	34.3	14.6	13.7	11.5	33.4	24.4	7.3	87.8
2018 /09	43.9	74.7	32.3	67.6	27.0	56.3	13.6	14.2	88.3	35.3	34.2	10.9	12.9	10.8	34.4	24.3	6.3	88.2
2018 /12	44.1	73.4	31.3	70.1	26.6	62.8	18.5	12.7	89.4	36.8	32.3	10.4	14.0	10.9	35.4	22.4	5.1	88.0
2019 /03	42.7	70.4	31.2	70.6	26.4	62.0	20.3	10.4	93.8	35.7	29.7	8.0	13.3	9.9	35.9	20.3	3.6	84.9
2019 /06	44.1	74.8	30.7	70.7	26.7	61.4	23.8	10.6	100.2	36.3	27.5	7.1	13.1	9.0	36.4	20.6	2.6	85.2
2019 /09	48.5	85.0	31.0	73.4	28.3	62.5	27.8	10.8	99.0	40.0	28.4	7.1	13.7	9.9	37.1	20.8	2.0	83.0
2019 /12	55.1	96.9	37.0	77.0	31.4	63.0	36.0	14.7	108.6	46.1	28.4	8.2	16.3	10.6	38.9	21.6	4.1	84.0
2020 /03	62.5	104.7	40.8	80.1	40.0	65.7	45.5	18.9	140.5	56.9	30.0	12.1	18.7	13.1	40.1	23.5	6.1	84.4
2020 /06	65.0	105.2	42.2	81.3	47.8	67.4	51.6	20.6	151.1	60.8	32.7	17.6	19.9	13.9	43.1	24.7	6.6	85.0
2020 /09	77.0	125.9	49.8	88.0	51.1	68.8	61.7	26.2	218.1	72.6	36.6	21.4	25.4	17.2	46.3	26.8	9.7	87.7
2020 /12	86.5	133.4	70.7	107.8	57.1	77.0	69.6	42.3	226.4	82.4	40.1	26.7	30.7	24.0	50.6	34.1	18.3	93.1
2021 /03	100.8	148.5	81.4	121.0	70.1	85.1	82.3	49.7	235.1	102.4	50.6	35.9	38.0	29.3	54.4	41.8	24.9	104.9
2021 /06	108.2	156.4	87.0	122.7	85.7	88.9	89.1	50.4	227.8	112.1	55.3	42.5	42.3	33.5	55.0	45.6	27.4	118.7
2021 /09	121.7	170.3	97.6	124.2	107.4	100.7	98.6	56.2	230.4	130.5	66.4	51.0	50.0	40.9	59.2	51.5	34.5	137.1
2021 /12	124.1	172.1	98.6	121.6	112.6	113.4	100.1	56.9	217.1	132.0	75.2	56.9	55.0	46.8	63.1	55.5	38.9	141.5
2022 /03	120.1	167.7	95.9	114.1	106.3	113.9	92.1	57.2	204.1	126.5	78.1	56.4	54.2	48.4	62.7	55.4	40.0	142.7
2022 /06	118.7	166.6	95.2	109.3	101.4	115.1	89.1	55.3	198.5	124.3	82.8	56.7	54.7	49.8	62.7	55.2	40.5	146.6
2022 /09	107.5	153.1	86.8	99.5	89.6	106.5	80.5	50.7	173.6	109.2	81.7	52.6	51.3	48.9	60.2	53.1	37.9	144.5

〈표 8〉 계속

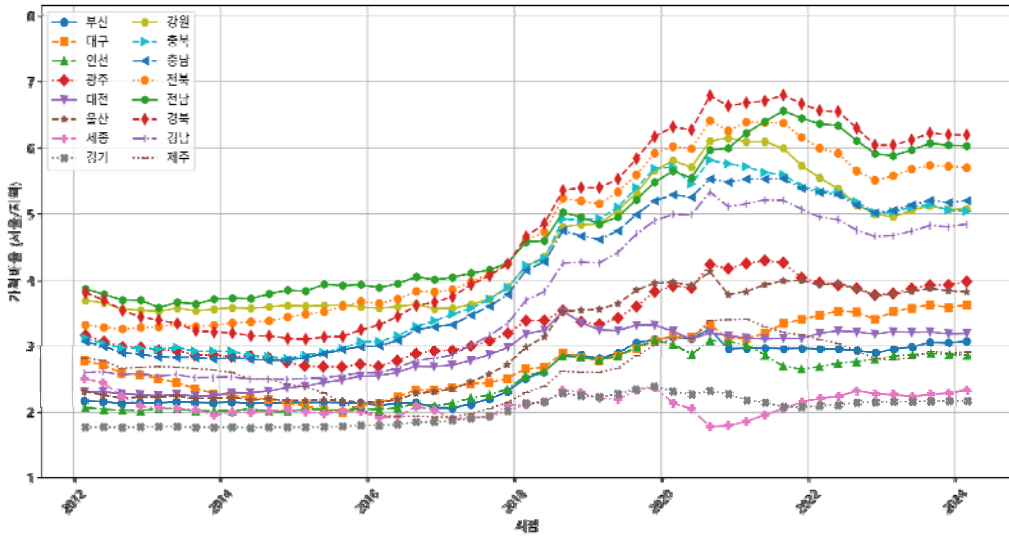
시점	전국	서울	부산	대구	인천	광주	대전	울산	세종	경기	강원	충북	충남	전북	전남	경북	경남	제주
2022/12	93.7	134.9	75.4	91.2	73.5	96.9	69.1	43.7	158.9	93.7	73.3	46.3	43.8	41.7	53.6	48.2	30.6	136.0
2023/03	90.0	131.6	69.8	82.5	68.3	93.1	65.0	41.0	156.9	90.4	72.5	44.0	40.7	38.1	52.3	46.1	28.3	133.9
2023/06	92.8	135.9	70.8	82.8	70.3	92.9	68.6	42.0	164.8	93.9	72.2	44.6	41.0	38.1	52.8	46.9	28.9	135.6
2023/09	96.7	142.1	71.1	85.1	73.6	95.3	72.7	44.1	168.0	98.5	74.4	47.2	42.7	40.3	54.3	48.3	30.0	134.1
2023/12	94.5	139.0	69.4	84.9	72.3	92.7	71.9	43.6	162.5	95.4	74.2	47.7	41.9	38.9	53.0	47.0	28.9	133.2
2024/03	93.6	138.2	67.4	82.2	72.3	89.6	71.1	43.5	156.4	95.1	73.2	47.7	40.6	39.0	52.8	46.5	27.3	131.2

주: 1) 누적변화율은 시계열의 최초시점(2012.3.1.)으로부터 각 시점까지의 누적된 변화율을 의미한다.
 2) 소수점 2자리 이하는 반올림되었다.

결과적으로, 이는 서울과 기타 지역 간의 절대적 가격 수준 격차가 매우 클 뿐 아니라 격차 자체가 지속적으로 확대되어왔음을 의미한다. 〈그림 1〉은 서울과 지역의 가격 차이(=서울가격/지역가격)가 어떻게 변화하는지를 시계열 그래프로 제시하고 있다. 그림을 보면, 2012년 초 기준으로 가장 높은 평균가격을 기록한 지역(서울)과 두 번째로 높은 지역(경기)의 배율은 1.76배 수준이었고, 가장 낮은 지역(전남)과의 배율은 3.86배 수준이었다. 그러나 2024년 3월에는 서울과 두 번째로 높은 지역(경기)와의 배율은 2.15배, 가장 낮은 지역(경북)과의 배율은 6.19배에 달하게 되었다. 이는 수도권, 특히 서울과 지역 간 가격의 격차 자체가 10여년간 큰 폭으로 상승했음을 직접적으로 보여준다.

본 연구의 추정 내용은 기존 관련 연구의 관찰 내용과도 부합한다. 예를 들어 허원제 외(2020)는 수도권과 비수도권 지역의 주택 가격 격차에 대한 분석에서, 2012년에서 2018년 이전에는 아파트 연평균 가격이 다소간 상승하는 가운데에서도 지역간 격차가 어느정도 유지되었으나, 2018년에 그 격차의 폭이 심화되기 시작했다는 것을 관측하였다. 본 연구의 추정치는 그 이후에도 격차의 심화가 3~4년간(2018년에서 2021년) 지속되었다는 것을 관측할 수 있도록 해준다.

다음으로, 서울 내 지역의 가격 수준을 살펴보자. 이는 〈표 9〉과 〈표 10〉에 제시되어 있다. 2024년 3월 기준으로 서울에서 가장 높은 평균 가격을 기록한 지역은 강남구로, 평균 1,634만 원/m²에 달하였다. 이는 서울 전체 평균인 1,121만 원/m²을 크게 상회하며, 가장 낮은 평균가격을 기록한 도봉구(약 524만 원/m²)와 비교하면 3.12배에 달하는 수준이다. 강남구 다음으로는 용산구(1,454만 원/m²), 서초구(1,391만 원/m²), 송파구(1,053만 원/m²) 등의 가격 수준이 높게 나타났다.



주: 가격배율은 서울의 단위 면적(1m²) 당 아파트 가격을 지역의 단위 면적(1m²) 당 아파트 가격으로 나눈 값을 의미한다.

<그림 1> 지역 대비 서울 가격배율의 추이

<표 9> 서울 평균 아파트 가격: 강남구~동작구(단위: 만 원/m²)

시점	강남구	강동구	강북구	강서구	관악구	광진구	구로구	금천구	노원구	도봉구	동대문구	동작구
2012/03	704	386	299	329	380	441	309	294	327	287	333	418
2012/06	693	378	302	321	381	438	308	298	320	283	334	414
2012/09	674	376	294	319	372	431	304	297	316	277	327	408
2012/12	670	378	295	321	375	431	301	299	311	276	329	404
2013/03	689	380	291	323	378	428	305	301	312	278	327	402
2013/06	672	381	302	325	377	433	306	298	312	277	326	405
2013/09	677	384	315	324	376	438	307	304	316	281	330	399
2013/12	689	386	307	328	382	438	311	308	316	284	336	407
2014/03	698	393	323	332	385	442	309	309	320	285	340	410
2014/06	701	390	324	333	387	448	312	305	319	285	340	412
2014/09	718	400	329	336	389	443	311	315	324	291	343	415
2014/12	725	404	329	340	392	445	317	312	330	288	349	419
2015/03	739	410	327	351	396	453	316	317	333	292	350	423
2015/06	762	421	329	364	405	458	323	319	341	297	364	437
2015/09	782	429	337	377	413	466	331	329	349	302	376	447
2015/12	802	439	344	383	423	469	333	342	353	314	379	450

〈표 9〉 계속

시점	강남구	강동구	강북구	강서구	관악구	광진구	구로구	금천구	노원구	도봉구	동대문구	동작구
2016/03	801	438	343	387	423	478	334	336	351	311	390	464
2016/06	822	440	346	398	433	475	338	346	354	316	389	467
2016/09	866	454	358	413	443	492	345	344	367	328	403	484
2016/12	876	460	361	420	449	503	353	352	372	329	411	494
2017/03	885	460	361	423	445	505	354	356	376	330	414	492
2017/06	915	479	362	434	456	511	357	369	384	333	425	512
2017/09	950	512	370	449	462	527	364	378	398	343	439	529
2017/12	996	518	378	459	473	539	370	381	403	349	442	545
2018/03	1,071	543	384	473	481	573	374	385	408	354	465	582
2018/06	1,088	557	415	485	488	588	383	397	413	359	485	599
2018/09	1,172	593	432	526	515	615	405	418	447	377	510	650
2018/12	1,154	592	440	529	518	617	415	422	463	386	524	647
2019/03	1,129	586	436	522	515	617	412	425	459	382	520	630
2019/06	1,175	598	426	525	529	630	418	430	470	387	529	653
2019/09	1,249	624	442	543	542	661	432	442	487	394	547	685
2019/12	1,317	651	452	562	562	698	447	462	513	407	565	735
2020/03	1,320	675	471	581	581	722	458	482	539	420	596	750
2020/06	1,325	683	475	585	589	722	472	494	556	425	605	746
2020/09	1,418	741	517	622	631	781	505	522	628	472	650	805
2020/12	1,441	780	540	648	641	808	533	536	656	492	667	839
2021/03	1,491	819	547	675	666	843	568	565	704	525	707	876
2021/06	1,556	842	559	689	678	870	598	586	738	553	725	899
2021/09	1,627	872	597	739	705	910	627	622	782	589	748	963
2021/12	1,668	882	607	745	708	907	634	621	777	585	746	975
2022/03	1,654	884	601	744	708	919	631	623	763	577	749	967
2022/06	1,664	870	614	736	709	921	635	620	770	582	754	988
2022/09	1,657	835	588	717	691	902	613	606	715	550	743	943
2022/12	1,571	779	577	672	664	841	590	576	680	530	699	883
2023/03	1,588	787	577	671	666	848	574	575	667	530	699	883
2023/06	1,615	789	580	680	670	867	577	572	668	532	700	890
2023/09	1,641	806	596	692	681	900	580	582	671	536	712	915
2023/12	1,636	798	585	690	669	872	573	575	656	526	706	911
2024/03	1,634	795	562	679	669	882	567	565	653	524	712	904

주: 소수점 이하는 반올림되었다.

〈표 10〉 서울 평균 아파트 가격: 마포구~중랑구(단위: 만 원/m²)

시점	마포구	서대문구	서초구	성동구	성북구	송파구	양천구	영등포구	용산구	은평구	종로구	중구	중랑구
2012/03	433	335	592	514	340	482	365	441	547	330	427	419	324
2012/06	428	329	590	516	339	478	360	438	555	329	419	411	319
2012/09	423	330	579	503	327	471	356	425	561	324	421	427	313
2012/12	425	327	578	502	328	471	355	425	564	324	416	421	318
2013/03	425	333	579	507	331	468	355	424	553	323	416	415	314
2013/06	432	330	586	508	333	475	360	431	552	329	415	421	320
2013/09	431	334	583	514	333	475	360	431	568	330	419	410	320
2013/12	440	333	591	516	338	477	362	434	570	332	420	430	316
2014/03	442	340	602	521	340	487	367	442	576	337	429	415	328
2014/06	449	348	609	524	340	485	366	438	574	338	430	422	326
2014/09	455	349	618	529	343	497	369	444	588	344	442	429	323
2014/12	456	356	633	530	354	502	375	447	591	347	458	438	330
2015/03	464	356	648	526	356	502	384	455	600	350	455	455	331
2015/06	479	368	659	546	367	511	392	457	623	357	463	452	339
2015/09	494	374	675	560	382	530	401	475	631	362	472	483	346
2015/12	505	383	691	571	390	536	408	478	643	372	481	485	348
2016/03	505	383	696	578	385	542	408	482	642	371	486	484	344
2016/06	512	391	710	595	395	549	412	494	656	380	496	492	352
2016/09	530	403	731	628	401	569	433	524	670	395	513	517	359
2016/12	546	419	743	647	411	577	439	535	690	402	523	521	363
2017/03	555	419	750	658	409	580	441	540	701	406	527	543	366
2017/06	572	434	789	680	414	608	453	562	722	414	529	547	368
2017/09	583	451	804	699	427	628	463	570	742	425	546	549	378
2017/12	595	460	834	729	441	657	476	598	785	429	549	569	387
2018/03	632	477	879	787	452	707	493	630	816	443	579	594	394
2018/06	652	500	898	823	470	710	503	649	848	452	588	593	401
2018/09	705	535	964	886	509	755	546	721	912	485	608	612	435
2018/12	703	545	954	892	528	750	550	718	908	491	623	617	440
2019/03	707	539	955	879	520	735	543	717	923	488	624	611	449
2019/06	717	548	975	902	523	761	551	734	935	496	622	622	455
2019/09	754	574	1,033	986	550	805	579	780	971	505	650	649	471
2019/12	785	603	1,079	1,042	578	849	620	820	1,040	520	662	681	483
2020/03	813	617	1,093	1,085	611	877	646	848	1,081	540	669	700	508
2020/06	819	628	1,095	1,096	627	887	648	856	1,105	550	697	715	521

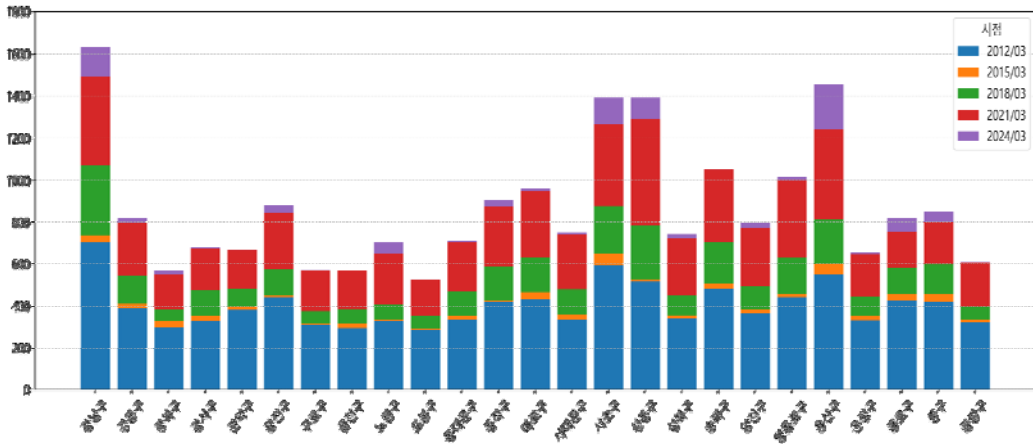
〈표 10〉 계속

시점	마포구	서대문구	서초구	성동구	성북구	송파구	양천구	영등포구	용산구	은평구	종로구	중구	중랑구
2020/09	876	677	1,183	1,179	684	962	704	923	1,166	587	713	752	568
2020/12	914	709	1,219	1,212	709	1,008	741	950	1,199	618	735	779	585
2021/03	947	744	1,267	1,291	745	1,050	774	1,002	1,240	647	755	800	605
2021/06	968	782	1,299	1,325	764	1,081	802	1,035	1,285	661	777	816	630
2021/09	1,020	809	1,368	1,377	809	1,123	837	1,072	1,374	698	801	847	667
2021/12	1,024	811	1,390	1,405	813	1,133	844	1,090	1,394	700	809	879	668
2022/03	1,013	803	1,413	1,416	790	1,137	844	1,077	1,411	698	805	878	665
2022/06	1,026	797	1,435	1,428	790	1,137	848	1,085	1,438	700	811	874	672
2022/09	977	787	1,400	1,400	755	1,092	824	1,054	1,404	683	808	857	646
2022/12	956	757	1,365	1,336	710	1,018	784	991	1,413	660	805	870	625
2023/03	948	745	1,364	1,314	712	1,022	781	982	1,392	649	805	866	613
2023/06	953	749	1,378	1,327	710	1,043	790	1,008	1,398	653	813	863	620
2023/09	985	767	1,417	1,372	730	1,075	810	1,035	1,418	660	830	846	626
2023/12	964	756	1,403	1,363	728	1,063	802	1,030	1,410	655	818	851	601
2024/03	963	751	1,391	1,391	723	1,053	795	1,017	1,454	655	819	851	612

주: 소수점 이하는 반올림되었다.

서울 내 지역의 가격 변동 양상을 살펴보면, 전체적으로는 완만한 상승세가 2012년부터 2017년까지 이어졌으며, 2018년부터 2021년 상반기까지는 모든 자치구에서 동반적으로 급등세를 보였다. 그러나 2022년 이후로는 대부분의 자치구에서 가격 상승세가 꺾이거나 조정 국면에 진입하였으며, 일부 지역에서는 소폭 하락세도 나타났다. 다만, 이러한 조정기의 가격 변화는 고가 지역일수록 완만한 양상을 보였고, 중저가 지역에서 더 두드러진 조정이 일어난 것으로 관측된다. 가격 상승이 두드러진 곳은 주로 고가 지역으로 강남구, 서초구, 성동구, 용산구로 각각 2012년 3월부터 2024년 3월까지 약 132%, 134%, 170%, 165% 상승한 것으로 나타난다. 반면, 강북구(88%), 구로구(83%), 도봉구(82%), 금천구(92%), 종로구(92%)의 경우 상대적으로 가격 상승의 속도가 완만했다. 결과적으로, 해당 기간 동안 지역 내 가격 격차 역시 심화된 것으로 나타난다(〈그림 2〉). 강남구와 비교할 때, 금천구와의 배율은 2012년 3월 2.39배에서 2024년 3월 2.89로, 강북구와의 배율은 2012년 3월 2.35배에서 2024년 3월 2.90배로 확대된 것으로 나타났다.

본 연구에서는 재고 주택의 표본을 측정 시점에 고정하고 있으므로, 전수적 추정이라는 것을 제외하면 고정 바스켓 기반 가격 지수와 유사한 성격의 지표라고 볼 수 있다. 그러므로 본 연구의 추정 결과를 기존의 한국부동산원 매매가격지수 및 KB부동산 매매가격지수와 같은 기존 공표 지수와 정량적으로 비교해 보는 것도 본 추정의 특성을 이해하는 데에 도움이 될 수 있다. 〈그림 3〉은 세 시계열을 18개

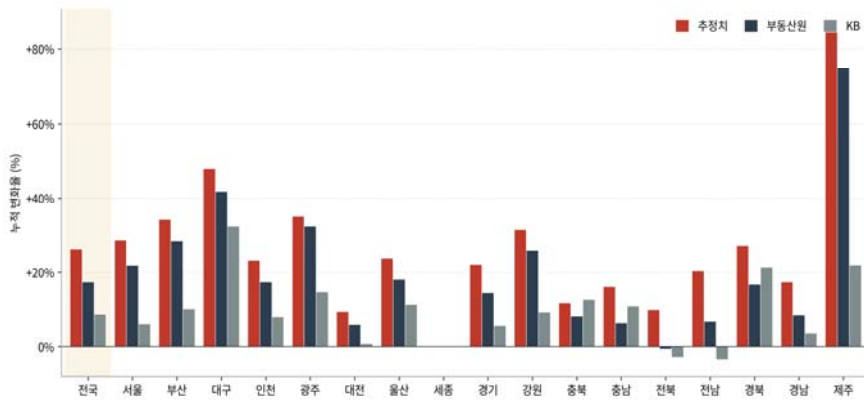


주: 그래프에서 바(bar)의 높이는 각 지역의 단위 면적(1m²) 당 평균 아파트 가격의 시점별 값을 의미한다.

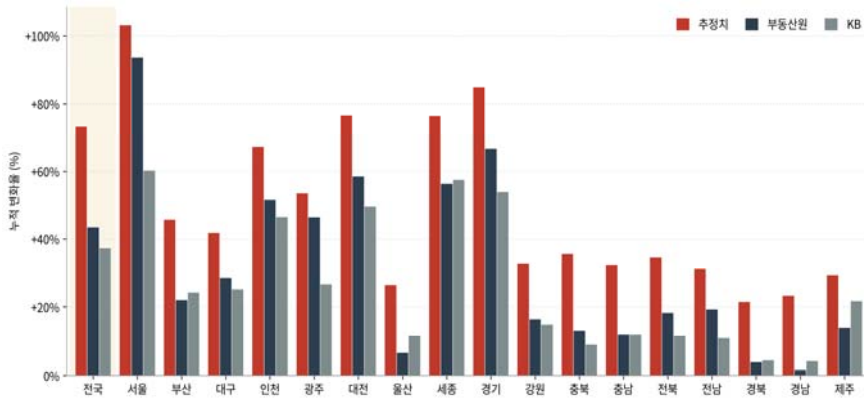
<그림 2> 서울 지역별 아파트 평균가격의 변화

지역(전국 및 17개 시·도)별로 2012년 1분기=100으로 재지수화한 후, 분석 기간을 회복기(2012년 2분기~2017년 1분기), 폭등기(2017년 2분기~2022년 1분기), 조정기(2022년 2분기~2024년 1분기)의 세 구간으로 구분하여 각 구간의 누적 변화율을 비교한 결과이다(<그림 3>).

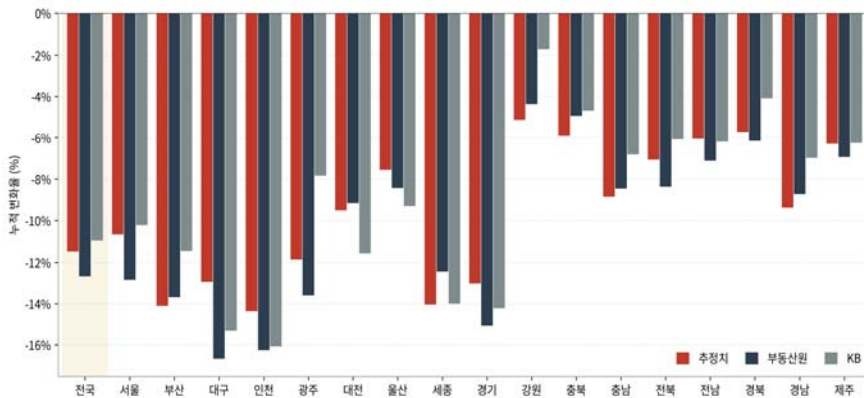
본 연구의 추정치를 통한 지수와 다른 두 가격지수의 누적 상승률, 분기 변화율의 표준편차로 측정 한 변동성, 그리고 변화율 상관계수를 비교할 수 있다. 비교 결과의 주요 양상은 다음과 같다. 첫째, 본 연구 추정치의 누적 상승률은 두 공표 지수에 비해 일관되게 큰 폭으로 나타난다. 분석 기간 전체에 걸친 전국 단위의 누적 상승률은 본 연구 추정치 93%, 한국부동산원 45%, KB 34%이며, 서울의 경우 본 연구 138%, 한국부동산원 107%, KB 54%로 그 차이가 더욱 두드러진다. 다만 이러한 격차는 분석 기간 전반에 균일하게 발생하기보다는 가격 상승 국면에 집중되는 경향이 관측된다. 회복기에는 격차가 비교적 작은 편이며(전국 추정치 +26%, 한국부동산원 +17%, KB +9%), 폭등기에 격차가 가장 크게 벌어진다(전국 추정치 +73%, 한국부동산원 +44%, KB +37%; 서울 추정치 +103%, 한국 부동산원 +94%, KB +60%). 반면 조정기에서는 세 지수가 모두 -10% 안팎의 비슷한 폭으로 하락하 되, 두 공표 지수의 하락폭이 본 연구 추정치보다 다소 크게 나타난다(전국 추정치 -11.5%, 한국부동산원 -12.7%, KB -11.0%). 둘째, 시기적 추세의 정합성 면에서는 세 시계열이 매우 높은 일관성을 공유한다. 셋째, 가격 정점 시점은 본 연구 추정치(전국·서울 모두 2021년 4분기)와 한국부동산원 (2021년 3분기)이 거의 동일한 반면, KB지수는 2022년 2분기로 다소 늦게 형성된다. 이는 KB지수의 평활 산출 구조와 일관되는 결과이다.



(A) 2012/2분기~2017/1분기 누적변화율



(B) 2017/2분기~2022/1분기 누적변화율



(C) 2022/2분기~2024/1분기 누적변화율

주: 각 구간의 시작 시점 대비 끝 시점의 누적 변화율(%)을 계산하였다. 세종시는 1구간에서 자료 결측으로 제외되었다. 부동산원은 아파트 실거래가 매매가격지수를, KB는 주택가격동향의 월간 아파트 매매가격 지수를 사용하였다.

<그림 3> 추정치와 기존 가격지수의 비교

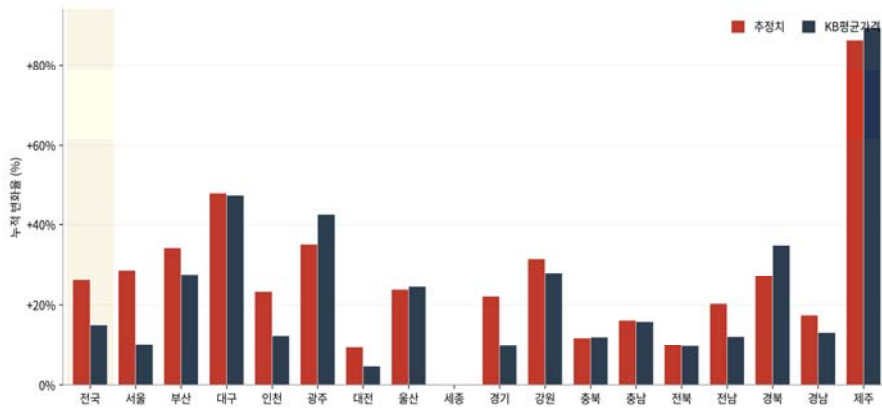
비교 결과를 요약하면, 본 연구의 추정치는 시기적 추세의 방향과 시점에 있어서 기존 공표 지표와 매우 일관된 시장 신호를 공유하면서도, 동일 자산의 가격 변화 폭은 보다 분명하게 드러내는 결과를 보인다. 이러한 차이는 부분적으로는 본 연구가 분석 기간 종료 시점 기준 고정 재고 표본의 가격을 동일 자산 바스켓에 대해 시점별로 추정하는 데에 비해, 두 공표 지표는 표본의 갱신·가중·평활화 과정을 거치는 산출 구조를 가지는 데에서 비롯한 것으로 해석될 수도 있다. 다만 가격 상승기와 조정기 사이에 관측되는 격차의 비대칭적 양상은 이러한 산출 구조의 차이만으로 온전히 설명되지 않는 요소의 가능성에 대해 검토의 여지가 있을 것으로 보인다.

본 연구의 추정치를 보다 입체적으로 평가하기 위해, KB부동산의 m²당 평균 매매가격(전용면적 기준) 자료와도 추가로 비교하였다. KB의 평균 가격 자료는 본 연구의 추정치와 마찬가지로 가격의 절대적 수준 정보를 담고 있다는 공통점을 가지나, 본 연구가 측정하는 면적 기준(공급면적)과 KB가 측정하는 면적 기준(전용면적)이 다르므로 절대 수준의 직접 비교는 적절하지 않다. 이에 두 시계열을 모두 2012년 1분기 = 100으로 재지수화한 후, 앞서와 동일한 세 구간 구분에 따라 각 구간의 누적 변화율을 비교하였다. <그림 4>는 그 결과를 18개 지역에 대해 보여준다.

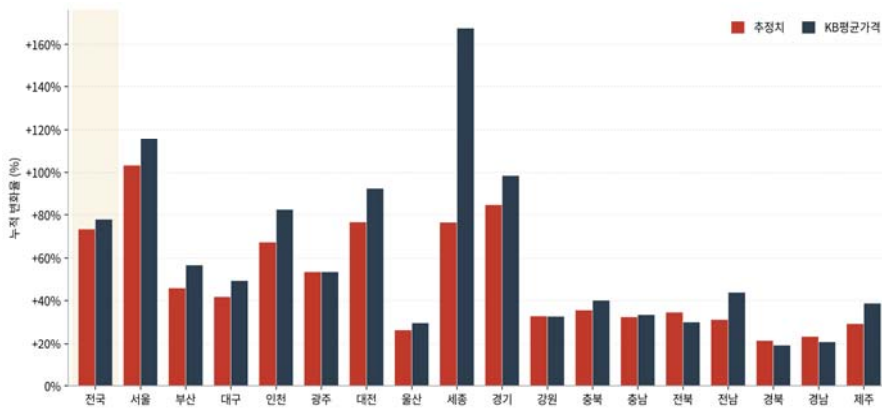
평균가격과의 누적 상승률 비교 결과는 앞서의 KB지수와와의 비교에서 관측된 양상과는 다소 다른 모습을 보인다. KB지수와와의 비교에서는 본 연구의 추정치가 일관되게 더 큰 누적 상승률을 보였으나, KB평균가격과의 비교에서는 18개 지역 중 15개 지역에서 오히려 KB평균가격의 누적 상승률이 본 연구보다 더 크게 나타난다. 특히 세종(추정치 115.6% vs KB평균 209.2%), 제주(131.7% vs 165.2%), 광주(89.3% vs 114.7%) 등에서 그 차이가 두드러진다. 반면 서울(138.0% vs 122.8%), 경기(95.5% vs 89.9%), 전국(93.4% vs 84.0%)에서는 본 연구의 추정치가 여전히 더 큰 누적 상승률을 보인다. 이러한 차이를 구간별로 살펴보면, KB평균가격의 우세는 주로 폭등기에 집중되어 나타나 는 반면(전국 추정치 +73% vs KB평균 +78%; 세종 추정치 +77% vs KB평균 +168%; 인천 추정치 +67% vs KB평균 +82%), 조정기에서는 두 시계열이 전국 단위에서 -11.5%로 거의 동일한 폭으로 하락하며 일부 지역에서는 오히려 본 연구 추정치의 하락폭이 다소 크게 관측된다. 즉, KB평균가격이 본 연구 추정치를 상회하는 양상은 가격 상승 국면에서 명확히 드러나지만, 가격 조정 국면에서는 그러한 우세가 관측되지 않는다.

이러한 누적 상승률 패턴의 비대칭성은 두 시계열의 산출 구조 차이에 기인하는 것으로 보인다. 본 연구는 재고 표본을 고정한 가운데 동일 자산 바스켓의 가격 변화를 측정하는 데에 비해, KB평균가격은 시점별로 다른 표본의 평균 가격을 산출하므로 주로 거래되는 표본의 변화(신축 단지의 추가, 거래 활성화 지역으로의 표본 비중 이동 등)가 가격 변화에 함께 반영된다. 즉, KB평균가격이 본 연구 추정치보다 큰 누적 상승률을 보이는 지역에서는 분석 기간 중 신규 공급된 단지의 가격 수준이

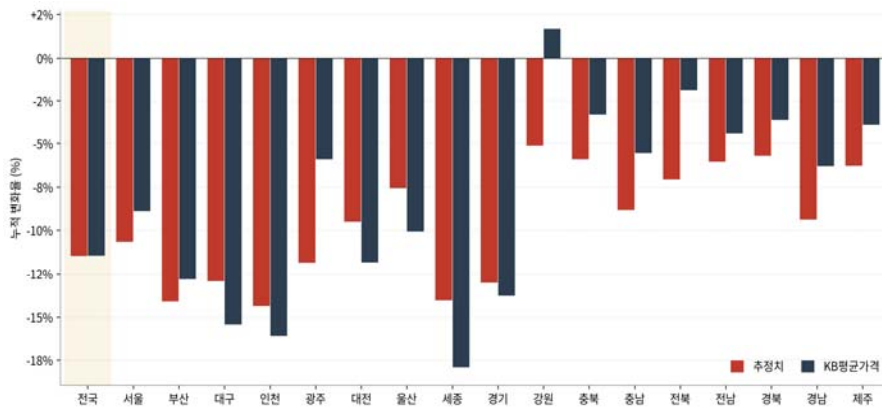
트리 앙상블 기반 부동산정보형을 이용한 재고 주택가격 분포 관측: 전국 아파트 단지 전수 자료의 시계열 분석



(A) 2012/2분기~2017/1분기 누적변화율



(B) 2017/2분기~2022/1분기 누적변화율



(C) 2022/2분기~2024/1분기 누적변화율

주: 각 구간의 시작 시점 대비 끝 시점의 누적 변화율(%)을 계산하였다. 세종시는 1구간에서 자료 결측으로 제외되었다. KB평균가격은 m²당 매매평균가격을 의미한다.

〈그림 4〉 추정치와 평균 거래가격(지수화)의 비교

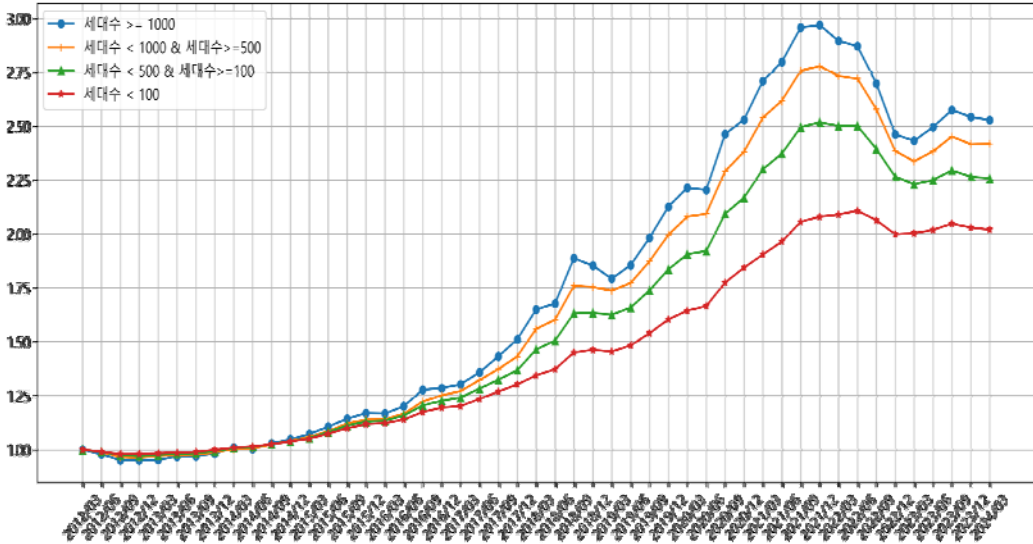
기존 단지에 비해 상대적으로 높아 표본 변화 효과가 평균을 끌어올렸을 가능성이, 반대로 본 연구 추정치가 더 큰 누적 상승률을 보이는 지역에서는 동일 자산 바스켓 자체의 가격 변동 폭이 표본 갱신 효과를 상회할 만큼 컸을 가능성이 시사된다. 이는 두 추정치가 서로 다른 질문에 답하는 도구라는 점에서 상보적인 정보를 제공함을 시사한다.

정리하면, 본 연구의 추정치는 기존 주택가격지수와 비교에서는 평활화로 인한 변동성 축소에서 적시성을, KB 평균가격과의 비교에서는 표본 구성 변화로부터 자유로운 동일 자산 바스켓 기반의 일관성을 각각 확인시켜 준다. 이러한 양 측면의 비교는 본 연구의 추정 결과가 기존 자료들이 갖는 산출 구조상의 두 가지 한계인 평활화로 인한 적시성 저하 및 표본 구성 변화로 인한 일관성 저하를 동시에 통제된 분석 도구로서 기능함을 정량적으로 뒷받침한다. 다만 두 비교에서 공통적으로 관측되는, 가격 상승 국면에서 격차가 두드러지고 조정 국면에서는 그러한 격차가 악화되는 비대칭적 양상은 산출 구조의 차이만으로 전부 설명되지 않을 가능성을 시사하는 측면이 있으며, 이에 대한 보다 면밀한 검토는 향후 과제로 남는다.

본 연구에서 제시된 방식의 장점은 분석 목적에 따라 하위 집단의 가격지수를 자유롭게 편성할 수 있다는 점이다. 전수적 자료에 기반하여 개별 주택 단위까지 가격이 추정되어 있으므로, 지역, 면적, 준공연도, 세대수 등 임의의 기준에 따라 부분 표본을 구성하고 그에 대응하는 가격지수를 즉시 도출할 수 있다. 이는 기존의 공표 가격지수가 제공하지 못하는 세분화된 시장 진단을 가능하게 한다는 점에서 본 방식의 활용 가치를 잘 보여준다.

그 예시로 단지 규모에 따른 가격 변동의 차이를 살펴보았다. 구체적으로, 분석 기점인 2012년 3월을 기준 시점(=1)으로 하여 서울 지역 아파트를 단지 세대수에 따라 100세대 이하, 100~500세대, 500~1,000세대, 1,000세대 이상의 네 집단으로 구분한 뒤 각 집단별 가격 변화 추이를 비교 관찰하였다. <그림 5>에서 확인할 수 있듯이, 분석 기간 동안 서울 지역의 아파트 가격이 전반적으로 상승하는 가운데서도 특히 단지 규모가 큰 아파트일수록 가격 상승 폭이 더욱 가파르게 나타났다. 즉, 세대수가 큰 집단일수록 누적 상승률이 높았으며, 집단 간 격차는 시간이 경과함에 따라 점진적으로 확대되는 양상을 보였다.

일반적으로 단지의 세대수가 크다는 것은 자산으로서의 유동성, 즉 향후 거래 성사 가능성이 높고 거래 비용이 상대적으로 낮음을 의미한다. 따라서 <그림 5>가 보여주는 결과는, 주택가격 상승기에 들어서면서 시장이 대규모 단지가 갖는 높은 유동성과 그로부터 함의되는 자산적 가치를 점차 더 높게 평가하는 방향으로 변화해 왔음을 시사한다. 이는 동일한 시기, 동일한 지역 내의 아파트라 하더라도 단지 규모라는 미시적 속성에 따라 가격 형성 메커니즘이 차별적으로 작동할 수 있음을 보여주는 결과이며, 본 연구에서 구축한 전수적 가격추정은 이러한 세분화된 시장 동학을 포착하는데 활용될 수 있음을 동시에 보여준다.



주: 각 지수는 2012/1분기를 기준으로 지수화되었다.

〈그림 5〉 단지 크기별 아파트 가격 추이(서울)

2. 가격 분포의 추이

이번에는 분위값을 통해 가격의 분포를 관찰해보자. 〈표 11〉은 전국의 아파트 전체를 대상으로 그 가격의 분포 추이를 관측한 결과이다. 먼저 확인할 수 있는 것은, 해당기간 동안 평균가격 뿐 아니라 가격의 분포 구조 자체가 변화해 왔다는 것이다.

2012년 초 기준으로 가격 분포의 하단부인 10% 분위수는 m^2 당 116만 원 내외 수준에 형성되어 있었으며, 중앙값(50% 분위수)은 208만 원/ m^2 , 90% 분위수는 396만 원/ m^2 에 머물렀다. 이후 2014에서 2019년 사이에는 완만한 상승세가 이어졌으며, 전국 각 분위수 모두에서 가격 수준이 점진적으로 상승하였다. 그러나 눈에 띄는 변화는 2020년을 전후로 급격하게 나타났다. 이 시기부터 특히 상위 분위수의 상승 폭이 하위 분위수보다 뚜렷하게 크다는 점이 확인되었다. 예를 들어, 90% 분위수의 경우 2020년 중반 이후 빠르게 1,000만 원/ m^2 을 상회하였고, 이후 2021년에는 1,082만 원/ m^2 에 육박하는 수준으로 치솟았다. 하지만 같은 기간 하위 10% 분위수와 25% 분위수는 다소 더디게 상승하여, 분포의 상하위 간 격차가 점차 확대되는 구조가 나타났다. 다시 말해, 특정 고가 지역(예: 서울, 세종, 일부 수도권 신도시 등)의 가격 급등이 전체 가격 분포 상위 분위수의 빠른 상승을 견인하였으며, 상대적으로 저가 주택이 분포한 지역은 상승폭이 제한되었음을 시사한다. 2022년 이후의 분포 변화

〈표 11〉 아파트 가격 분포: 전국(단위: 만 원/m²)

시점	5% 분위수	10% 분위수	25% 분위수	중앙값(50%)	75% 분위수	90% 분위수	95% 분위수
2012/03	96	116	155	208	289	396	501
2012/06	98	117	155	207	283	390	494
2012/09	98	118	156	206	280	382	482
2012/12	98	118	156	206	278	380	483
2013/03	100	120	159	208	279	382	484
2013/06	101	121	159	209	282	384	488
2013/09	103	123	162	212	285	388	492
2013/12	103	124	164	215	289	393	495
2014/03	105	125	167	219	294	399	504
2014/06	105	126	168	219	295	402	504
2014/09	106	127	171	224	299	408	518
2014/12	107	129	174	227	305	416	524
2015/03	109	131	176	231	310	426	532
2015/06	109	133	181	237	320	438	550
2015/09	111	136	185	244	329	453	568
2015/12	111	136	186	245	332	460	575
2016/03	112	137	186	246	333	462	574
2016/06	112	138	186	248	337	471	587
2016/09	112	139	189	252	346	490	616
2016/12	113	140	190	256	354	501	625
2017/03	113	140	191	257	357	505	632
2017/06	113	140	192	258	361	520	657
2017/09	113	141	192	261	370	542	691
2017/12	112	140	192	261	371	553	715
2018/03	111	138	191	261	376	584	789
2018/06	108	137	188	261	382	604	812
2018/09	107	134	187	262	396	670	889
2018/12	105	133	187	263	406	670	883
2019/03	103	130	184	263	408	660	865
2019/06	101	127	183	263	414	673	884
2019/09	100	127	183	266	419	702	939
2019/12	100	127	185	272	436	747	1,001
2020/03	100	127	187	280	477	795	1,053
2020/06	100	128	190	290	491	805	1,059

〈표 11〉 계속

시점	5% 분위수	10% 분위수	25% 분위수	중앙값(50%)	75% 분위수	90% 분위수	95% 분위수
2020/09	101	129	194	302	533	891	1,155
2020/12	102	132	201	325	569	931	1,206
2021/03	105	136	212	356	620	1,005	1,292
2021/06	107	140	220	375	645	1,033	1,330
2021/09	111	148	234	406	688	1,088	1,408
2021/12	114	153	244	415	689	1,082	1,398
2022/03	116	154	245	410	667	1,059	1,369
2022/06	118	157	247	405	657	1,049	1,355
2022/09	119	157	243	386	611	982	1,269
2022/12	117	153	232	358	569	907	1,176
2023/03	116	151	226	347	559	899	1,162
2023/06	116	151	226	349	571	913	1,189
2023/09	116	151	226	354	583	933	1,223
2023/12	116	150	225	352	575	919	1,203
2024/03	115	148	222	350	572	918	1,205

주: 소수점 이하는 반올림되었다.

에서도 또 하나의 중요한 특징이 나타난다. 전체 가격 수준이 다소 조정 국면에 들어간 것으로 보이지만, 분포 상단(90%, 95%)은 여전히 높은 수준을 유지하고 있는 반면, 하단 분위수에서는 가격이 다소 하락하거나 정체되는 모습을 보이고 있다. 이는 가격 조정기에도 고가 지역 아파트의 절대 가격 수준이 유지되는 경향이 있으며, 시장 내에서 가격 하방의 탄력성이 낮음을 의미할 수 있다. 종합적으로 이러한 변화는 전국적으로 일정부분 공통적인 가격 변화 추세를 공유하는 가운데에서도, 중상위 이상 가격대 아파트의 상대적 상승률 및 하방경직성이 더 크다는 것을 보여주는 결과라 할 수 있다. 물론 주의할 점은, 전국 단위의 가격 분포는 지역 간 격차를 그대로 반영한다는 것이다. 그러므로 위와 같은 비대칭성의 확대는 지역간 격차의 확대의 재확인이라고도 볼 수 있을 것이다.

시도간 격차 확대를 제외한 아파트 가격 분포 추이를 관측하려면, 지역 내 가격 분포를 살펴봐야 할 것이다. 다만 모든 지역을 상세히 기술하면 지면을 지나치게 차지하게 되므로, 본 연구에서는 서울의 경우에 대해서 자세히 기술하려 한다. 〈표 12〉는 서울 내 아파트 가격 분포를 보여준다. 서울은 전국에서 가장 높은 가격 수준을 유지하고 있는 지역으로, 가격 분포의 상단이 특히 빠르게 팽창하는 특징이 있다. 2012년 초 기준으로, 서울 아파트 가격의 하단부(10% 분위수)는 303만 원/m², 중앙값은 410만 원/m², 90% 분위수는 765만 원/m² 수준이었다. 2014년부터 2019년까지의 기간에는

〈표 12〉 아파트 가격 분포: 서울(단위: 만 원/m²)

시점	5% 분위수	10% 분위수	25% 분위수	중앙값(50%)	75% 분위수	90% 분위수	95% 분위수
2012/03	283	303	342	410	534	765	879
2012/06	278	297	335	403	524	756	860
2012/09	273	291	332	398	516	730	830
2012/12	271	290	331	395	517	726	830
2013/03	271	291	332	396	516	738	838
2013/06	272	292	333	399	523	740	844
2013/09	274	294	335	401	523	739	849
2013/12	275	296	339	405	530	748	852
2014/03	279	301	342	410	539	765	879
2014/06	278	299	342	413	539	766	885
2014/09	283	303	347	418	549	790	902
2014/12	286	308	352	425	557	787	916
2015/03	292	312	357	433	565	811	953
2015/06	297	322	366	442	584	836	984
2015/09	306	330	379	460	603	849	1,006
2015/12	311	335	384	468	612	877	1,031
2016/03	311	336	386	470	613	870	1,022
2016/06	316	341	392	479	624	893	1,061
2016/09	326	352	406	500	661	967	1,154
2016/12	329	355	414	512	667	972	1,154
2017/03	331	360	416	515	677	981	1,182
2017/06	337	368	426	532	702	1,035	1,238
2017/09	345	378	443	558	732	1,086	1,319
2017/12	350	384	449	572	776	1,161	1,428
2018/03	355	392	464	606	848	1,308	1,577
2018/06	360	399	479	624	880	1,303	1,590
2018/09	379	430	523	696	967	1,435	1,767
2018/12	395	443	539	701	955	1,379	1,676
2019/03	393	444	537	692	942	1,351	1,613
2019/06	398	450	544	710	967	1,405	1,709
2019/09	409	461	560	743	1,033	1,509	1,866
2019/12	420	477	586	779	1,105	1,654	2,017
2020/03	442	506	624	827	1,161	1,674	1,997
2020/06	447	516	637	839	1,156	1,658	1,968
2020/09	508	581	713	918	1,265	1,824	2,163
2020/12	529	602	741	960	1,311	1,854	2,189

〈표 12〉 계속

시점	5% 분위수	10% 분위수	25% 분위수	중앙값(50%)	75% 분위수	90% 분위수	95% 분위수
2021/03	560	642	794	1,026	1,389	1,984	2,325
2021/06	578	666	821	1,062	1,432	2,028	2,393
2021/09	608	705	878	1,112	1,490	2,154	2,508
2021/12	612	709	878	1,112	1,498	2,168	2,575
2022/03	611	700	863	1,090	1,461	2,146	2,559
2022/06	618	698	857	1,082	1,447	2,108	2,566
2022/09	590	665	803	1,016	1,374	2,013	2,457
2022/12	564	631	751	938	1,271	1,869	2,237
2023/03	553	616	732	929	1,252	1,865	2,246
2023/06	556	618	737	937	1,276	1,939	2,322
2023/09	559	622	745	959	1,319	2,010	2,402
2023/12	556	616	736	947	1,289	1,989	2,371
2024/03	554	612	731	940	1,289	1,976	2,362

주: 소수점 이하는 반올림되었다.

하단부부터 상단부에 이르기까지 전 구간에 걸쳐 뚜렷한 상승이 지속되었다. 예를 들어, 중앙값은 2012년 410만 원에서 2019년 말에는 640만 원 수준까지 약 56% 상승했으며, 90% 분위수도 같은 기간 약 750만 원에서 1,127만 원/m²로 대폭 상승하였다. 그러나 가장 가파른 상승은 2020년에서 2021년 사이에 집중적으로 발생하였다. 서울의 중앙값은 2020년 초 680만 원에서 2021년 말에는 922만 원을 돌파하였고, 90% 분위수는 같은 기간 동안 1,200만 원을 넘어서며 고가 아파트의 폭발적인 상승세가 있었음을 확인할 수 있다.

눈여겨볼 점은, 서울의 경우 가격 상승이 단지 평균 가격의 상승에 그치지 않고, 분포의 상단이 상대적으로 더 빠르게 위로 늘어나는 상단 팽창형 구조를 나타낸다는 것이다. 이는 특히 95% 분위수의 변화를 통해 더욱 뚜렷하게 확인된다. 2012년 약 880만 원/m²이었던 95% 분위수는 2022년에는 약 1,600만 원/m²에 도달하였으며, 이는 서울 내에서도 상위 5% 아파트의 가격 상승률이 다른 계층 대비 월등히 높았음을 의미한다. 중위값 역시 2020년 초 827만 원에서 2021년 말까지 1,490만 원으로 크게 팽창하였다. 그러나 같은 기간 하위 10% 분위수는 약 300만 원에서 450만 원 수준으로의 상승에 그쳤다. 요약하면, 전반적인 가격 상승의 기초 하에서도 중위값 이상의 고가 아파트의 가격 상승이 더 컸다는 것을 보여준다. 또한, 2022년 이후의 조정 국면에서도 서울의 가격 분포는 일정한 특징을 유지하고 있다. 전체 분위수 중 하위 분위수(예: 10%, 25%)에서는 미미한 조정 또는 정체 양상이 나타난 반면, 상위 분위수(90%, 95%)에서는 조정폭이 제한적이며, 여전히 고점에 근접한 가격 수준을 유지하고 있다. 이는 전국 분포에서

관찰된 바와 마찬가지로 고가 아파트의 하방 경직성이 더 강하게 나타나고 있음을 보여준다.

흥미로운 것은 이러한 경향이 전국적으로 공통적으로 관찰된다는 것이다. <표 13>은 각 지역별 가격 분포를 토대로, 각 지역별 가격의 75분위수와 25분위수의 비율이 어떻게 변하는 지를 보여주고 있다. 결과는, 비록 지역마다 정도의 차이는 있으나, 상위 분위 가격 아파트와 하위 또는 중앙값의 비율이 2018년 즈음을 기점으로 2021년까지 거의 모든 지역에서 벌어졌다는 것을 보여준다. 예를 들어, 부산의 경우 2012년 3월에는 75분위수와 25분위수의 비율은 1.37배 수준이었으나 2021년 9월을 기준으로 2.16배까지 상승한 것을 확인할 수 있다. 이러한 경향은 대부분의 다른 지역에서도 나타난다. 2022년 이후 분위수간 격차는 다소 진정되는 추세이나, 대부분의 지역에서 2018년 이전 수준을 회복하지는 못하고 있다. 이는 아파트 가격의 양극화가 지역 간에만 발생하는 것이 아니라 지역 내에서도 심화되어 왔으며, 전국에서 공통적으로 나타났음을 보여준다.

<표 13> 아파트 가격 분위수 비율 75~25(단위: 배)

시점	서울	부산	대구	인천	광주	대전	울산	세종	경기	강원	충북	충남	전북	전남	경북	경남	제주
2012/03	1.56	1.37	1.37	1.30	1.33	1.45	1.45	1.28	1.51	1.55	1.67	1.53	1.60	1.50	1.59	1.49	1.34
2012/06	1.56	1.37	1.36	1.31	1.32	1.44	1.42	1.23	1.47	1.52	1.63	1.54	1.57	1.48	1.58	1.47	1.39
2012/09	1.56	1.37	1.36	1.33	1.33	1.44	1.38	1.19	1.47	1.52	1.61	1.53	1.57	1.49	1.55	1.45	1.36
2012/12	1.56	1.37	1.38	1.29	1.34	1.45	1.40	1.12	1.48	1.56	1.61	1.54	1.59	1.46	1.57	1.46	1.38
2013/03	1.55	1.38	1.34	1.31	1.33	1.46	1.40	1.15	1.46	1.53	1.59	1.51	1.57	1.45	1.54	1.44	1.47
2013/06	1.57	1.40	1.37	1.27	1.35	1.45	1.41	1.14	1.49	1.51	1.58	1.52	1.58	1.44	1.54	1.49	1.42
2013/09	1.56	1.38	1.33	1.33	1.34	1.44	1.40	1.11	1.49	1.49	1.56	1.50	1.61	1.44	1.55	1.47	1.45
2013/12	1.56	1.39	1.31	1.35	1.35	1.46	1.41	1.11	1.49	1.51	1.55	1.50	1.61	1.45	1.57	1.49	1.39
2014/03	1.58	1.40	1.32	1.31	1.33	1.44	1.39	1.13	1.50	1.50	1.54	1.51	1.61	1.48	1.58	1.51	1.46
2014/06	1.57	1.40	1.32	1.36	1.34	1.43	1.39	1.14	1.52	1.49	1.61	1.53	1.60	1.49	1.57	1.52	1.42
2014/09	1.58	1.41	1.31	1.31	1.33	1.40	1.38	1.14	1.52	1.48	1.60	1.53	1.61	1.49	1.60	1.55	1.38
2014/12	1.58	1.42	1.31	1.34	1.36	1.44	1.38	1.14	1.53	1.50	1.65	1.51	1.62	1.49	1.62	1.54	1.40

〈표 13〉 계속

시점	서울	부산	대구	인천	광주	대전	울산	세종	경기	강원	충북	충남	전북	전남	경북	경남	제주
2015/03	1.58	1.42	1.33	1.32	1.37	1.43	1.38	1.10	1.54	1.47	1.62	1.53	1.65	1.50	1.64	1.53	1.40
2015/06	1.60	1.45	1.37	1.30	1.35	1.45	1.40	1.12	1.54	1.49	1.64	1.52	1.65	1.51	1.65	1.52	1.42
2015/09	1.59	1.46	1.39	1.31	1.37	1.41	1.41	1.12	1.52	1.50	1.58	1.55	1.68	1.51	1.67	1.51	1.43
2015/12	1.59	1.48	1.39	1.32	1.37	1.41	1.42	1.14	1.51	1.51	1.61	1.56	1.69	1.52	1.68	1.52	1.43
2016/03	1.59	1.49	1.35	1.33	1.36	1.45	1.42	1.21	1.50	1.47	1.61	1.55	1.72	1.53	1.64	1.52	1.40
2016/06	1.59	1.47	1.36	1.29	1.36	1.44	1.44	1.22	1.52	1.48	1.58	1.54	1.72	1.54	1.67	1.50	1.34
2016/09	1.63	1.51	1.37	1.30	1.38	1.47	1.43	1.27	1.51	1.49	1.61	1.57	1.72	1.57	1.69	1.51	1.33
2016/12	1.61	1.55	1.42	1.32	1.37	1.44	1.45	1.27	1.53	1.53	1.58	1.55	1.73	1.62	1.70	1.51	1.33
2017/03	1.63	1.54	1.42	1.33	1.40	1.48	1.45	1.25	1.53	1.54	1.53	1.55	1.75	1.58	1.74	1.52	1.36
2017/06	1.65	1.55	1.42	1.32	1.42	1.48	1.44	1.26	1.54	1.54	1.60	1.56	1.74	1.57	1.78	1.52	1.39
2017/09	1.65	1.55	1.45	1.34	1.42	1.46	1.45	1.28	1.56	1.56	1.63	1.56	1.75	1.61	1.82	1.52	1.40
2017/12	1.73	1.55	1.46	1.32	1.41	1.49	1.47	1.31	1.57	1.58	1.64	1.58	1.75	1.64	1.83	1.53	1.41
2018/03	1.83	1.54	1.47	1.32	1.46	1.50	1.51	1.30	1.61	1.57	1.66	1.62	1.78	1.60	1.91	1.54	1.44
2018/06	1.84	1.57	1.52	1.34	1.52	1.53	1.52	1.34	1.66	1.59	1.67	1.63	1.76	1.64	1.92	1.59	1.46
2018/09	1.85	1.57	1.55	1.34	1.51	1.55	1.54	1.43	1.73	1.62	1.74	1.67	1.75	1.65	1.97	1.61	1.42
2018/12	1.77	1.57	1.56	1.34	1.58	1.60	1.56	1.42	1.79	1.63	1.76	1.72	1.81	1.65	1.97	1.61	1.44
2019/03	1.75	1.59	1.55	1.39	1.65	1.66	1.61	1.44	1.80	1.67	1.79	1.71	1.78	1.67	2.02	1.66	1.45
2019/06	1.78	1.61	1.57	1.37	1.65	1.67	1.63	1.49	1.81	1.71	1.80	1.76	1.80	1.64	2.07	1.69	1.47
2019/09	1.85	1.63	1.56	1.40	1.67	1.78	1.66	1.55	1.85	1.72	1.82	1.78	1.81	1.74	2.10	1.72	1.47

〈표 13〉 계속

시점	서울	부산	대구	인천	광주	대전	울산	세종	경기	강원	충북	충남	전북	전남	경북	경남	제주
2019 /12	1.89	1.69	1.54	1.43	1.68	1.84	1.71	1.62	1.90	1.72	1.81	1.80	1.86	1.70	2.11	1.77	1.47
2020 /03	1.86	1.74	1.58	1.50	1.68	1.88	1.74	1.58	2.03	1.77	1.85	1.84	1.86	1.70	2.12	1.81	1.47
2020 /06	1.81	1.75	1.57	1.53	1.70	1.85	1.75	1.49	2.00	1.77	1.95	1.84	1.87	1.71	2.19	1.81	1.47
2020 /09	1.77	1.81	1.59	1.54	1.71	1.87	1.83	1.40	2.03	1.80	1.99	1.93	1.88	1.78	2.21	1.84	1.47
2020 /12	1.77	2.07	1.68	1.59	1.80	1.82	2.03	1.42	2.02	1.81	2.04	1.97	1.91	1.78	2.21	1.91	1.49
2021 /03	1.75	2.12	1.64	1.60	1.82	1.93	2.09	1.45	1.98	1.85	2.12	2.02	2.00	1.81	2.24	1.96	1.57
2021 /06	1.74	2.16	1.64	1.55	1.87	1.89	2.02	1.41	1.91	1.88	2.21	2.03	2.02	1.81	2.22	1.97	1.68
2021 /09	1.70	2.16	1.62	1.54	1.92	1.86	2.07	1.36	1.84	1.92	2.19	2.04	2.03	1.85	2.22	2.02	1.77
2021 /12	1.71	2.09	1.62	1.52	1.90	1.87	2.03	1.36	1.80	1.97	2.16	2.01	2.12	1.89	2.20	2.07	1.80
2022 /03	1.69	2.07	1.57	1.46	1.85	1.78	2.04	1.34	1.75	1.96	2.14	2.02	2.04	1.87	2.21	2.05	1.83
2022 /06	1.69	2.04	1.56	1.41	1.83	1.75	2.04	1.32	1.75	1.95	2.09	1.99	2.05	1.91	2.16	2.02	1.86
2022 /09	1.71	1.97	1.56	1.42	1.78	1.67	2.00	1.31	1.72	1.89	2.00	1.94	2.03	1.87	2.13	2.00	1.84
2022 /12	1.69	1.89	1.56	1.39	1.76	1.65	1.94	1.33	1.74	1.86	1.94	1.91	2.00	1.85	2.11	1.92	1.75
2023 /03	1.71	1.88	1.57	1.43	1.75	1.69	1.89	1.41	1.77	1.82	1.90	1.90	1.95	1.84	2.06	1.91	1.76
2023 /06	1.73	1.88	1.61	1.46	1.75	1.69	1.91	1.40	1.81	1.82	1.93	1.90	1.97	1.84	2.09	1.90	1.75
2023 /09	1.77	1.92	1.65	1.50	1.74	1.73	1.95	1.42	1.81	1.85	1.89	1.88	1.97	1.86	2.07	1.92	1.71
2023 /12	1.75	1.91	1.67	1.51	1.72	1.78	1.95	1.40	1.80	1.83	1.92	1.91	2.00	1.84	2.08	1.95	1.71
2024 /03	1.76	1.92	1.66	1.53	1.73	1.75	1.95	1.41	1.82	1.84	1.93	1.87	1.96	1.85	2.07	1.97	1.74

주: 소수점 세자리 이하는 반올림되었다.

3. 분석적 활용 예시 1: 가격 불균등성 지표의 산출

앞에서는 각 지역별 가격 분위수 및 분위수 간의 배율을 확인함으로써 가격 분포의 추이를 확인하였다. 이번에는 직접적인 불균등성 지표인 지니계수를 관측해볼 수 있다. 지니계수는 소득·자산·주택가격처럼 비음수(0 이상)로 측정되는 연속형 변수의 분포가 '완전평등(모든 관측치가 동일)' 상태에서 얼마나 이탈했는지를 0에서 1 사이의 단일 지표로 정량화한 값이다. 값이 0에 가까우면 평등, 1에 가까우면 불평등이 심함을 뜻하며, 관측치 전체를 쌍대 비교해 얻은 절대격차의 평균을 전체 평균으로 정규화한 수준으로 이해할 수 있다. 이때 지니계수의 산식은 아래와 같다.

$$Gini = \frac{1}{2n^2\mu} \sum_{i=1}^n \sum_{j=1}^n |x_i - x_j|, \quad \langle \text{식 4} \rangle$$

이때 n, x_i, μ 는 표본수, i 번째 관측값, 평균을 의미한다. 지니계수는 다음과 같은 특징을 갖는다. 첫째, 지니계수는 척도불변성을 가진다. 이는 모든 관측값에 동일한 비율을 곱하더라도 지니계수의 값은 변하지 않는다는 것을 의미하며, 이를 통해 단순한 물가 상승이나 가격 수준의 전반적인 이동과는 무관하게 분포의 불균등성만을 측정할 수 있게 된다. 둘째, 지니계수는 값의 범위가 0에서 1 사이로 제한되기 때문에 시점 간 또는 지역 간의 비교가 직관적으로 가능하다. 값이 0에 가까울수록 완전한 평등 상태에, 1에 가까울수록 극심한 불균등 상태에 해당한다. 마지막으로, 지니계수는 계산 효율성 측면에서도 장점을 가진다. 누적비율(로렌츠 곡선)의 면적을 활용하면 비교적 간단한 정렬과 합산만으로 계산이 가능하므로, 대규모 데이터를 다룰 때에도 실용적으로 사용할 수 있다. 이러한 이유로 지니계수는 불균등의 정도를 하나의 수치로 요약하여 시기별 또는 지역 간 비교 분석을 수행할 수 있는 대표적인 지표로 널리 활용된다. 따라서 앞서 살펴본 분위수 기반 지표가 분포의 위치·폭 변화를 보여줬다면, 지니계수는 전체 분포의 불균등성 수준을 단일 지표로 요약해 지역 간·시점 간 불평등 구조를 비교할 수 있다는 장점이 있다.

본 연구에서는 단지별 아파트 가격의 추정치를 통해 지역별 지니계수의 변화를 산출하였다. 본 연구에서 추정된 지니계수는 아파트 단지·평형 단위의 m^2 당 가격을 관측 단위로 산출한 것으로, 자산(주택)의 단위 가격 분포에 대한 불균등성 지표이다. 이는 가구 단위 소득 또는 가구 단위 자산을 관측 대상으로 하는 통상의 지니계수와 분석 단위가 상이하며, 가구의 주택 보유 수, 거주 가구원 수, 가구 소득 등은 반영하지 않는다는 것을 염두에 두어야 한다. 즉, 본 결과는 "주택의 단위 가격 분포가 얼마나 분산되어 있는가"에 대한 측정으로 이해되어야 하며, "가계 자산의 불균등성"이나 "주거 부담의 불균등성"과는 구별되어야 한다.

〈표 14〉는 각 지역에 대해 아파트 단위 면적 당 가격의 지니계수를 도출한 결과이다. 당연한 결과이지만, 전국을 대상으로 하는 경우 각 지역별 지니계수보다 더 높은 값(즉, 더 불균등한 결과)이 도출된다. 표를 보면, 앞 장에서 얻었던 것과 유사한 합의의 결과를 확인할 수 있다. 결과를 요약하면 다음과 같다. 먼저, 지역 간 불균등성의 수준은 전반적으로 크게 상승한다. 전국을 대상으로 지니계수를 추정하는 경우, 2012년 3월 기준 0.283이지만, 2024년 3월에는 0.397까지 상승한다. 둘째, 지역 내 불균등성의 수준도 2018년을 기점으로 2021년까지 뚜렷하게 상승한다. 예를 들어, 경기도의 경우 2012년 3월 기준 지니계수는 0.174에 불과했으나, 2021년 9월이 되면 0.259까지 상승한다. 이러한 경향은, 지역에 따라 정도의 차이는 있으나 전국의 각 지역에서 공통적으로 발견된다. 셋째, 2022년 이후 지역에 따라서 약간의 불균등성 하락이 관측되는 곳이 있다. 하지만 서울의 경우 2022년 이후에도 큰 조정없이 불균등성이 커지는 추세를 보인다. 결론적으로 지니계수의 추정결과 역시 앞의 가격 분포의 관측을 통해 확인할 수 있었던 것과 유사한, 즉 가격의 양극화가 전국적으로도 지역 내에서도 전반적으로 확대되었음을 의미하는 결과를 얻을 수 있다.

본 연구의 가격 지니계수 추정치를 기존의 관련 추정치와 비교해 두는 것은 결과 해석에 도움이 된다. 우선 분석 단위가 다른 통계청의 가구 단위 순자산 지니계수와의 비교가 가능하다. 통계청 가계금융복지조사에 따르면 가구 순자산 지니계수는 2017년 0.584에서 2021년 약 0.603으로 상승하였으며(통계청, 2018, 2022), 이러한 가구 자산 불평등 심화의 주된 원동력은 주택가격 상승으로 보고된다. 본 연구가 추정한 전국 단위 가격 지니계수의 동일 기간 상승 폭(2017년 0.29 내외 → 2021년 0.41 수준)은 가구 단위 지표의 상승 추세와 방향 및 시점에 있어 일관된 신호를 보여, 본 연구의 추정 결과가 거시적 자산 불평등 흐름과 모순되지 않음을 시사한다. 다만 두 지표의 분석 단위(가구 자산과 주택 가격)와 자산 범위(전체 자산과 아파트 단위 가격)가 다르므로 절대 수준의 일대일 비교는 적절하지 않으며, 본 연구의 결과는 가구 단위 지표를 보완하는 자산(주택) 단위 정보로 해석되어야 한다.

〈표 14〉 아파트 가격 분위수 비율 90~50(단위: 배)

시점	서울	부산	대구	인천	광주	대전	울산	세종	경기	강원	충북	충남	전북	전남	경북	경남	제주
2012/03	1.30	1.17	1.16	1.16	1.15	1.20	1.18	1.09	1.22	1.21	1.28	1.21	1.25	1.20	1.22	1.18	1.16
2012/06	1.30	1.17	1.18	1.15	1.14	1.20	1.17	1.07	1.21	1.20	1.27	1.22	1.26	1.21	1.21	1.18	1.20
2012/09	1.30	1.17	1.17	1.16	1.15	1.20	1.15	1.04	1.21	1.21	1.25	1.23	1.27	1.19	1.21	1.17	1.17
2012/12	1.31	1.17	1.17	1.15	1.16	1.20	1.16	1.04	1.22	1.22	1.24	1.21	1.26	1.18	1.19	1.18	1.17

〈표 14〉 계속

시점	서울	부산	대구	인천	광주	대전	울산	세종	경기	강원	충북	충남	전북	전남	경북	경남	제주
2013/03	1.30	1.17	1.15	1.15	1.15	1.19	1.15	1.05	1.21	1.22	1.23	1.21	1.26	1.18	1.20	1.18	1.21
2013/06	1.31	1.17	1.17	1.13	1.16	1.19	1.14	1.04	1.22	1.21	1.22	1.22	1.27	1.20	1.20	1.20	1.20
2013/09	1.30	1.17	1.15	1.18	1.17	1.18	1.15	1.06	1.22	1.21	1.22	1.20	1.29	1.18	1.19	1.19	1.20
2013/12	1.31	1.17	1.14	1.18	1.16	1.19	1.17	1.04	1.23	1.22	1.21	1.21	1.27	1.19	1.19	1.20	1.19
2014/03	1.32	1.17	1.15	1.16	1.14	1.18	1.15	1.06	1.23	1.21	1.22	1.19	1.27	1.22	1.20	1.20	1.22
2014/06	1.31	1.16	1.15	1.17	1.15	1.19	1.15	1.05	1.24	1.22	1.24	1.21	1.28	1.22	1.19	1.21	1.18
2014/09	1.31	1.17	1.14	1.15	1.14	1.17	1.15	1.05	1.23	1.20	1.26	1.20	1.27	1.20	1.22	1.21	1.17
2014/12	1.31	1.18	1.15	1.17	1.17	1.18	1.15	1.05	1.24	1.20	1.28	1.18	1.28	1.21	1.23	1.19	1.18
2015/03	1.30	1.17	1.16	1.16	1.16	1.19	1.16	1.04	1.24	1.19	1.27	1.19	1.30	1.21	1.24	1.20	1.19
2015/06	1.32	1.18	1.19	1.16	1.14	1.18	1.17	1.06	1.23	1.19	1.24	1.20	1.31	1.22	1.25	1.19	1.20
2015/09	1.31	1.19	1.19	1.15	1.16	1.17	1.18	1.06	1.22	1.18	1.20	1.21	1.32	1.21	1.25	1.17	1.21
2015/12	1.31	1.20	1.21	1.16	1.16	1.16	1.19	1.07	1.22	1.19	1.24	1.22	1.30	1.21	1.26	1.19	1.19
2016/03	1.30	1.20	1.18	1.15	1.16	1.19	1.19	1.11	1.22	1.18	1.23	1.21	1.30	1.21	1.24	1.19	1.18
2016/06	1.30	1.20	1.19	1.14	1.17	1.18	1.16	1.12	1.22	1.19	1.23	1.22	1.33	1.21	1.25	1.18	1.15
2016/09	1.32	1.21	1.19	1.16	1.18	1.17	1.16	1.10	1.21	1.19	1.21	1.22	1.31	1.22	1.26	1.19	1.14
2016/12	1.30	1.23	1.21	1.16	1.17	1.18	1.16	1.11	1.22	1.22	1.19	1.21	1.33	1.24	1.26	1.19	1.15
2017/03	1.32	1.23	1.21	1.16	1.18	1.19	1.17	1.10	1.22	1.20	1.18	1.22	1.32	1.25	1.25	1.19	1.14
2017/06	1.32	1.24	1.21	1.17	1.19	1.19	1.17	1.12	1.23	1.18	1.21	1.23	1.31	1.24	1.28	1.20	1.17
2017/09	1.31	1.25	1.21	1.17	1.18	1.18	1.18	1.12	1.23	1.18	1.21	1.23	1.30	1.24	1.29	1.21	1.20
2017/12	1.36	1.23	1.23	1.16	1.19	1.20	1.20	1.12	1.24	1.20	1.21	1.21	1.30	1.25	1.29	1.19	1.20

〈표 14〉 계속

시점	서울	부산	대구	인천	광주	대전	울산	세종	경기	강원	충북	충남	전북	전남	경북	경남	제주
2018/03	1.40	1.24	1.23	1.16	1.21	1.21	1.21	1.11	1.27	1.21	1.22	1.24	1.30	1.23	1.33	1.21	1.20
2018/06	1.41	1.24	1.24	1.17	1.23	1.22	1.21	1.12	1.29	1.22	1.24	1.24	1.26	1.26	1.33	1.24	1.22
2018/09	1.39	1.25	1.24	1.18	1.24	1.23	1.23	1.19	1.32	1.23	1.28	1.26	1.29	1.25	1.35	1.24	1.21
2018/12	1.36	1.24	1.26	1.18	1.29	1.28	1.23	1.18	1.34	1.25	1.28	1.27	1.29	1.24	1.37	1.23	1.20
2019/03	1.36	1.25	1.26	1.19	1.31	1.31	1.25	1.16	1.34	1.27	1.31	1.27	1.29	1.27	1.37	1.24	1.22
2019/06	1.36	1.26	1.26	1.19	1.31	1.32	1.26	1.16	1.35	1.28	1.28	1.27	1.31	1.28	1.38	1.28	1.23
2019/09	1.39	1.26	1.26	1.20	1.32	1.36	1.25	1.17	1.37	1.29	1.29	1.28	1.31	1.30	1.39	1.28	1.24
2019/12	1.42	1.30	1.24	1.20	1.33	1.38	1.26	1.23	1.40	1.30	1.30	1.30	1.33	1.29	1.40	1.30	1.22
2020/03	1.40	1.32	1.26	1.26	1.33	1.37	1.27	1.24	1.43	1.31	1.30	1.32	1.37	1.28	1.40	1.32	1.24
2020/06	1.38	1.32	1.26	1.26	1.34	1.35	1.26	1.20	1.41	1.31	1.36	1.33	1.36	1.29	1.42	1.33	1.23
2020/09	1.38	1.34	1.28	1.26	1.35	1.36	1.32	1.18	1.42	1.33	1.37	1.35	1.36	1.32	1.44	1.34	1.22
2020/12	1.37	1.46	1.32	1.27	1.39	1.34	1.38	1.19	1.39	1.32	1.38	1.36	1.38	1.32	1.44	1.37	1.24
2021/03	1.35	1.46	1.30	1.27	1.40	1.40	1.39	1.17	1.39	1.32	1.40	1.38	1.43	1.34	1.47	1.39	1.27
2021/06	1.35	1.48	1.29	1.24	1.41	1.36	1.37	1.15	1.37	1.32	1.44	1.39	1.43	1.33	1.47	1.37	1.30
2021/09	1.34	1.49	1.29	1.24	1.42	1.35	1.41	1.15	1.34	1.36	1.43	1.39	1.44	1.35	1.46	1.41	1.33
2021/12	1.35	1.46	1.29	1.24	1.40	1.37	1.37	1.16	1.32	1.36	1.38	1.39	1.46	1.38	1.44	1.43	1.32
2022/03	1.34	1.44	1.29	1.21	1.37	1.34	1.37	1.16	1.30	1.37	1.38	1.36	1.44	1.39	1.45	1.43	1.36
2022/06	1.34	1.43	1.28	1.18	1.37	1.32	1.36	1.14	1.30	1.36	1.37	1.39	1.43	1.40	1.44	1.40	1.33
2022/09	1.35	1.40	1.28	1.19	1.37	1.28	1.36	1.12	1.31	1.34	1.34	1.36	1.40	1.41	1.42	1.38	1.33
2022/12	1.35	1.36	1.28	1.18	1.36	1.26	1.32	1.13	1.31	1.31	1.33	1.35	1.41	1.38	1.41	1.36	1.26

〈표 14〉 계속

시점	서울	부산	대구	인천	광주	대전	울산	세종	경기	강원	충북	충남	전북	전남	경북	경남	제주
2023/03	1.35	1.36	1.30	1.20	1.36	1.28	1.33	1.16	1.34	1.31	1.31	1.36	1.37	1.38	1.39	1.34	1.31
2023/06	1.36	1.37	1.34	1.21	1.35	1.28	1.35	1.16	1.35	1.31	1.32	1.35	1.40	1.37	1.40	1.34	1.28
2023/09	1.38	1.40	1.36	1.24	1.34	1.29	1.36	1.15	1.35	1.32	1.32	1.36	1.40	1.39	1.38	1.35	1.28
2023/12	1.36	1.40	1.36	1.24	1.36	1.31	1.36	1.15	1.34	1.32	1.32	1.37	1.40	1.38	1.41	1.37	1.29
2024/03	1.37	1.38	1.36	1.24	1.36	1.28	1.36	1.15	1.35	1.32	1.35	1.36	1.40	1.37	1.41	1.38	1.33

주: 소수점 세자리 이하는 반올림되었다.

한편, 본 연구의 가격 지니계수 추정치는 동일한 분석 도구를 활용한 선행 연구의 결과와도 그 위치를 비교해 둘 수 있다. 서재원·진장익(2021)은 2011~2019년 국토교통부 실거래가 자료를 활용하여 서울 25개 자치구별 주택가격 지니계수를 산출하고, 이를 계층이동 가능성에 대한 인식과 결합한 분석을 수행하였다. 본 연구의 V.3절은 동일한 도구(주택가격 지니계수)를 (i) 자치구가 아닌 17개 시도 단위로 확장하고, (ii) 분석 기간을 2012~2024년의 13년 시계열로 연장하며, (iii) 거래 표본이 아닌 전수 추정 표본에 적용한다는 점에서 선행 연구의 자치구 단위·단일 시점 분석을 일관된 방식으로 일반화한다. 자치구라는 작은 공간 단위 내부의 가격 분산만을 반영하는 선행 연구의 지니계수와, 시도 안의 모든 단위·평형을 포함하는 본 연구의 지니계수는 측정 대상이 구조적으로 다르므로 절대 수준의 직접 비교는 적절하지 않지만, 두 결과는 모두 분석 기간 동안 주택가격 분포의 불균등성이 확대되어 왔다는 일관된 결론을 공유한다.

4. 분석적 활용 예시 2: 아파트 가격 부담(Price to Income Ratio)의 산출

앞서 확인했듯이, 우리나라 아파트 시장에서는 지난 10여 년간 전반적인 가격 상승과 함께 지역 간 및 지역 내 격차의 확대가 동시에 나타났다. 그러나 주택 가격의 수준을 확인한 것만으로는 그것이 시장의 여력 대비 얼마나 높은 수준인지를 구체적으로 파악하기는 어려울 것이다. 이런 경우, 가계의 소득 대비 주택 가격의 상대적 비율인 PIR은 매우 유용한 정보가 될 수 있다. PIR은 일정 지역의 평균 또는 중위 주택가격을 해당 지역 가구의 연평균 소득으로 나눈 값으로, 가계의 지불 여력 대비 주택가격 (즉, 평균적 가구가 자기 지역의 평균적 주택을 구매하기 위해 몇 년치 소득을 모아야 하는지)을 나타낸다.

PIR에 대해 일정 수준 이상 과학적으로 추정하려면 기본적으로 전체 분포를 고려한 평균 주택 가격(또는 중위가격)과 소득 수준의 추정이 필요하므로, 지금까지는 이에 대한 추정이 다소 산발적으로 이루어져 왔다. 하지만 개별 주택에 대한 가격 추정을 하는 경우, PIR을 광범위하게 도출하기 용이할 뿐 아니라 그것이 시계열적으로 어떻게 변화했는지 또는 지역 내에서 PIR의 분포가 어떻게 되었는지 등을 보다 쉽게 관측할 수 있다. 이에 본 절에서는 각 지역의 아파트 시장 PIR의 추이를 먼저 살펴보고, 마지막으로 아파트 가격 분포를 고려한 PIR의 분포를 살펴보려 한다.⁷⁾

PIR의 계산을 위해서는 평균적 가구 크기와 주택 크기, 소득 수준에 대한 가정이 필요하다. 본 연구에서는 평균적 가구 크기에 대한 보정을 위해 2023년 인구 총조사 기준 평균 가구원수(2.2명)을 적용하였으며, 평균적 주택 크기로는 전국 아파트 공급면적의 평균(94.6m²)을 사용하였다. 또한, 1인당 소득 수준은 통계청의 지역별 개인소득을 적용하였다. 본 연구에서는 이를 통해 지역별 평균 가구 소득과 평균 크기의 아파트 가격과의 비율을 도출하였다.⁸⁾ 이때 소득은 평균 기반·세전 정의를 따른다. 즉, 분자에는 본 연구가 추정된 지역별 m²당 평균 가격에 평균 아파트 면적을 곱한 평균 주택가격을, 분모에는 통계청 지역소득의 1인당 개인소득(세전 기준)에 평균 가구원수를 곱한 평균 가구소득을 사용한다. 이러한 산정 방식은 World Bank Housing Indicators Programme의 표준 PIR 정의 및 이창무 외(2012)가 국제비교에 적용한 평균 기반 정의와 일치한다.

〈표 15〉는 전국 및 각 지역의 아파트 가격 PIR 추이를 보여준다. 전국 단위의 PIR은 2015년 6.14에서 시작해, 2019년까지 6.19로 완만한 상승세를 보이다가 2020년과 2021년을 거치면서 9.40으로 급등하였다. 이는 앞서 확인할 수 있었던 전국적인 주택 가격 급등기와 맞물려 있다. 즉, 이 같은 추이는 2020년에서 2021년 사이 주택가격이 소득에 비해 지나치게 빠르게 상승했다는 사실을 명확히 보여준다. 이후 2022년과 2023년에는 각각 8.77, 7.91로 소폭 하락했지만, 이는 2020년 이전에 비하여 여전히 뚜렷이 높은 수준이다. PIR의 절대적 수준에 대한 해석은 제반 경제 환경에 따라 다를 수 있으나 일반적으로 9 이상의 PIR이 대단히 낮은 주택 접근성을 의미한다는 점에서 이는 우리나라의 아파트 가격이 가구소득

7) 본 연구에서 제시하는 PIR은 아파트만을 대상으로 하므로, 시장 전체의 주택 부담 수준을 대표 하는 지표로 해석될 수는 없을 것이다. 또한, PIR의 산출은 기준 소득에 따라 달라질 수 있으므로, 본 연구에서 제시되는 결과는 하나의 추정치로 제시된다. 그럼에도 본 절의 결과는 기계학습 기반 전수 추정 자료가 PIR과 같은 시장 부담 지표의 산출로 어떻게 용이하게 확장될 수 있는지에 대한 예시로 의미를 가질 수 있을 것이다.

8) 본 연구는 평균 가구원수와 평균 공급면적에 대해 전국 평균값(각각 2.2명, 94.6m²)을 모든 지역에 동일하게 적용하였다. 이는 본 PIR이 측정하고자 하는 바를 "전국 평균에 해당하는 표준적 가구가 지역별 평균 아파트를 구매할 경우의 부담"으로 명확히 정의하기 위함이다. 만약 가구원수와 평균 면적을 지역별로 달리 적용할 경우, 산출된 PIR의 지역 간 차이에는 가격·소득의 격차에서 비롯되는 부분과 가구·주택 구조의 지역별 차이에서 비롯되는 부분이 동시에 혼재되어 지표의 의미가 모호해지는 측면이 있다. 본 연구는 가격·소득의 차이에 초점을 맞추어 비교를 단순화하기 위해 인구·구조적 모수는 고정하는 방식을 채택하였다.

〈표 15〉 아파트 가격 분위수 비율 95~50(단위: 배)

시점	서울	부산	대구	인천	광주	대전	울산	세종	경기	강원	충북	충남	전북	전남	경북	경남	제주
2012/03	2.15	1.44	1.47	1.51	1.41	1.45	1.47	1.12	1.69	1.49	1.56	1.49	1.59	1.51	1.51	1.59	1.68
2012/06	2.13	1.45	1.49	1.57	1.40	1.47	1.45	1.10	1.69	1.47	1.54	1.53	1.59	1.50	1.53	1.62	1.67
2012/09	2.09	1.45	1.46	1.56	1.39	1.49	1.44	1.09	1.64	1.44	1.52	1.52	1.60	1.49	1.49	1.61	1.63
2012/12	2.10	1.44	1.46	1.58	1.39	1.48	1.46	1.11	1.66	1.45	1.55	1.49	1.60	1.53	1.45	1.57	1.67
2013/03	2.11	1.46	1.43	1.61	1.39	1.45	1.42	1.13	1.64	1.43	1.52	1.48	1.60	1.49	1.47	1.61	1.65
2013/06	2.11	1.46	1.42	1.59	1.42	1.45	1.43	1.12	1.65	1.44	1.52	1.49	1.62	1.51	1.46	1.63	1.65
2013/09	2.12	1.46	1.39	1.59	1.42	1.45	1.42	1.13	1.66	1.42	1.48	1.47	1.61	1.49	1.46	1.58	1.59
2013/12	2.10	1.46	1.38	1.60	1.41	1.46	1.42	1.13	1.65	1.45	1.48	1.48	1.60	1.52	1.46	1.60	1.56
2014/03	2.14	1.45	1.41	1.56	1.39	1.45	1.43	1.24	1.68	1.46	1.52	1.47	1.62	1.56	1.47	1.58	1.63
2014/06	2.14	1.47	1.43	1.53	1.42	1.48	1.43	1.18	1.71	1.46	1.56	1.48	1.63	1.59	1.47	1.60	1.59
2014/09	2.16	1.48	1.42	1.58	1.40	1.47	1.46	1.17	1.74	1.44	1.60	1.47	1.63	1.57	1.49	1.60	1.60
2014/12	2.16	1.52	1.45	1.60	1.47	1.47	1.43	1.16	1.75	1.44	1.62	1.45	1.64	1.60	1.50	1.58	1.57
2015/03	2.20	1.51	1.52	1.54	1.47	1.52	1.46	1.15	1.78	1.43	1.59	1.44	1.64	1.59	1.50	1.60	1.70
2015/06	2.23	1.55	1.52	1.52	1.46	1.52	1.47	1.20	1.77	1.42	1.48	1.46	1.69	1.65	1.53	1.59	1.79
2015/09	2.19	1.55	1.54	1.50	1.48	1.53	1.48	1.20	1.76	1.43	1.48	1.48	1.68	1.62	1.51	1.55	1.78
2015/12	2.21	1.56	1.57	1.48	1.45	1.52	1.47	1.20	1.76	1.41	1.49	1.45	1.67	1.63	1.53	1.55	1.73
2016/03	2.17	1.59	1.54	1.50	1.47	1.55	1.48	1.25	1.75	1.45	1.47	1.47	1.67	1.65	1.51	1.56	1.77
2016/06	2.22	1.59	1.59	1.48	1.48	1.55	1.45	1.28	1.76	1.47	1.50	1.48	1.68	1.66	1.52	1.54	1.73
2016/09	2.31	1.62	1.60	1.46	1.47	1.57	1.45	1.22	1.77	1.47	1.45	1.49	1.69	1.64	1.53	1.54	1.72

〈표 15〉 계속

시점	서울	부산	대구	인천	광주	대전	울산	세종	경기	강원	충북	충남	전북	전남	경북	경남	제주
2016/12	2.25	1.68	1.61	1.49	1.48	1.54	1.46	1.22	1.77	1.49	1.47	1.53	1.72	1.64	1.54	1.55	1.72
2017/03	2.30	1.68	1.63	1.49	1.50	1.56	1.47	1.24	1.77	1.46	1.46	1.53	1.72	1.67	1.57	1.51	1.62
2017/06	2.33	1.71	1.62	1.50	1.52	1.58	1.48	1.28	1.83	1.44	1.50	1.53	1.71	1.67	1.60	1.52	1.66
2017/09	2.37	1.73	1.67	1.49	1.51	1.60	1.50	1.29	1.87	1.41	1.53	1.56	1.69	1.79	1.64	1.53	1.67
2017/12	2.49	1.71	1.72	1.49	1.54	1.62	1.50	1.31	1.91	1.44	1.55	1.54	1.70	1.80	1.62	1.51	1.68
2018/03	2.60	1.75	1.77	1.50	1.65	1.64	1.50	1.30	2.10	1.46	1.56	1.59	1.73	1.78	1.67	1.57	1.67
2018/06	2.55	1.76	1.82	1.54	1.74	1.69	1.53	1.36	2.17	1.48	1.57	1.62	1.72	1.77	1.71	1.66	1.73
2018/09	2.54	1.82	1.88	1.54	1.84	1.76	1.59	1.44	2.28	1.54	1.66	1.65	1.78	1.76	1.76	1.68	1.74
2018/12	2.39	1.81	1.89	1.56	1.83	2.01	1.59	1.43	2.22	1.58	1.65	1.68	1.80	1.77	1.77	1.69	1.73
2019/03	2.33	1.84	1.94	1.58	1.85	2.05	1.66	1.39	2.19	1.65	1.73	1.76	1.87	1.84	1.78	1.73	1.73
2019/06	2.41	1.87	1.93	1.62	1.84	2.16	1.64	1.40	2.21	1.69	1.76	1.75	1.94	1.85	1.80	1.76	1.74
2019/09	2.51	1.88	1.91	1.63	1.85	2.00	1.67	1.46	2.30	1.72	1.79	1.77	1.95	1.89	1.82	1.78	1.78
2019/12	2.59	2.06	1.88	1.71	1.88	2.03	1.69	1.55	2.43	1.79	1.86	1.92	1.99	1.88	1.89	1.84	1.77
2020/03	2.41	2.21	1.91	1.86	1.89	2.02	1.81	1.48	2.40	1.82	1.87	1.96	2.23	1.97	1.95	1.88	1.78
2020/06	2.35	2.25	1.90	1.80	1.91	1.92	1.85	1.45	2.29	1.87	2.02	2.01	2.26	2.00	1.95	1.93	1.80
2020/09	2.36	2.43	1.96	1.79	1.94	1.99	1.98	1.38	2.38	1.91	2.09	2.10	2.43	2.15	1.96	2.02	1.78
2020/12	2.28	2.52	2.19	1.73	2.09	2.06	2.19	1.39	2.31	1.98	2.15	2.18	2.61	2.15	2.10	2.19	1.85
2021/03	2.27	2.47	2.06	1.83	2.17	2.07	2.12	1.35	2.29	2.08	2.39	2.36	2.74	2.23	2.21	2.30	1.96
2021/06	2.25	2.48	2.03	1.72	2.13	1.90	2.02	1.36	2.20	2.11	2.27	2.38	2.66	2.18	2.21	2.25	2.02
2021/09	2.26	2.46	2.03	1.79	2.09	1.82	2.08	1.34	2.16	2.18	2.19	2.23	2.65	2.20	2.20	2.29	2.19

〈표 15〉 계속

시점	서울	부산	대구	인천	광주	대전	울산	세종	경기	강원	충북	충남	전북	전남	경북	경남	제주
2021/12	2.32	2.39	2.04	1.76	2.06	1.81	2.00	1.38	2.08	2.14	2.13	2.17	2.60	2.21	2.18	2.27	2.15
2022/03	2.35	2.30	2.04	1.74	2.02	1.79	1.92	1.37	2.08	2.12	2.06	2.07	2.61	2.20	2.13	2.26	2.08
2022/06	2.37	2.33	1.97	1.75	1.99	1.83	1.87	1.42	2.07	2.06	2.03	2.03	2.58	2.17	2.14	2.20	1.99
2022/09	2.42	2.20	1.94	1.63	1.92	1.78	1.85	1.43	2.06	1.98	1.93	1.96	2.38	2.12	2.03	2.13	1.94
2022/12	2.38	2.21	2.01	1.60	1.88	1.74	1.79	1.42	2.05	1.96	1.89	1.94	2.31	1.99	1.99	2.02	1.89
2023/03	2.42	2.17	2.01	1.68	1.91	1.80	1.83	1.43	2.12	1.95	1.86	1.96	2.25	2.03	1.94	1.97	1.91
2023/06	2.48	2.21	2.05	1.73	1.90	1.83	1.89	1.44	2.16	1.93	1.94	1.98	2.38	2.07	2.01	2.00	1.89
2023/09	2.51	2.23	2.09	1.75	1.94	1.97	1.93	1.42	2.19	1.92	2.01	2.01	2.46	2.14	2.00	2.04	1.86
2023/12	2.50	2.26	2.08	1.72	1.98	1.82	1.89	1.43	2.16	1.95	2.00	2.04	2.45	2.13	2.05	2.05	1.87
2024/03	2.51	2.29	2.10	1.72	1.94	1.80	1.95	1.45	2.17	1.93	2.02	2.05	2.48	2.12	2.09	2.05	1.89

주: 소수점 세자리 이하는 반올림되었다.

에 비해 평균적으로 상당히 높은 수준이라는 것을 의미할 수 있다.⁹⁾ 물론 이는 서울을 포함한 평균 가격으로부터 도출된 것이며 서울과 기타 지역의 격차가 상당하다는 것을 감안해야 할 것이다.

지역별로 보면, 가장 수준이 높고 해당기간 상승을 보여준 지역은 단연 서울이다. 2015년 서울의 PIR은 9.81로 이미 전국 평균을 크게 상회하고 있었는데, 이후 2018년에는 13.36, 2020년 16.64, 그리고 2021년에는 무려 19.53까지 치솟았다. 이는 평균적 가구 소득을 전혀 지출하지 않고 약 20년 치를 모아야 평균 수준의 아파트를 구매할 수 있다는 것을 의미하여, 가계의 구매력 대비 아파트 가격의 괴리가 서울에서 얼마나 심화되었는지를 구체적으로 보여준다. 2022년과 2023년에는 PIR 수치가 각각 18.35, 16.34로 다소 완화되었지만, 절대 수준은 여전히 전국에서 가장 높은 상태이다.

9) PIR의 절대적 수준에 대한 해석은 제반 경제 환경에 따라 다를 수 있으나, Demographia International Housing Affordability 보고서(Cox, 2024)의 분류에 따르면 PIR 9 이상은 "Impossibly Unaffordable"에 해당하는 가장 심각한 주택 접근성 단계로 분류된다. 다만 동 분류 체계는 중위 기반 PIR을 전제로 하므로, 하나의 참고 기준으로 이해되어야 한다.

경기도와 인천 역시 수도권 일원으로 PIR이 함께 상승하는 경향을 보였다. 경기도의 경우 2015년 6.59에서 시작해 2021년에는 10.46까지 상승하였다. 이는 서울보다는 낮지만 전국 평균보다는 높은 수치이며, 수도권 내 비서울 지역 역시 주택가격 부담이 심화되었음을 시사한다. 인천 또한 같은 시기 6.12에서 8.24로 상승한 후 최근에는 6.80으로 다소 조정된 모습을 보인다. 지방 대도시의 PIR은 수도권에 비해서는 상대적으로 낮은 수준에서 유지되었으나, 정점인 2021년에는 부산(8.02), 대구(7.31), 세종(7.58), 제주(7.58)와 같이 상당한 수준에 이르기도 했다. 한편, 충청, 호남, 영남권 대부분의 비수도권 지역은 수도권, 그리고 지방 대도시에 비해서도 PIR 수준이 확연히 낮았음을 확인할 수 있다. 예컨대 충북의 PIR은 2021년에도 5.80, 전남은 5.37, 경북은 5.06 등으로 유지되었다. 그러므로 PIR은 아파트 가격 자체에 비해서도 더 큰 지역간 차이를 보여준다.

이를 동일 시기에 대해 다른 산정방식을 택한 PIR 추정결과와 비교할 수 있다. 국토교통부 주거실태 조사는 자가주택의 중위가격을 중위가구소득으로 나누는 방식으로 PIR을 산출하며, 2022년 기준 전국 6.3배, 수도권 9.3배, 광역시 등 6.8배, 도지역 4.3배로 보고된다(국토교통부, 2023). KB국민은행은 주택가격과 가계소득을 각각 5분위로 분할한 5×5 PIR 매트릭스를 산출하며, 그 중 3분위(중위) 가격을 3분위(중위) 소득으로 나눈 대표 수치로서 2020년 3월 기준 전국 5.22배를 보고하였다(KB국민은행, 2020). 한국주택금융공사는 자사 주택담보대출 이용자를 대상으로 하여 2020년 1분기 기준 전국 4.45배, 서울 5.22배의 PIR을 산출하였다(한국주택금융공사, 2020).

이러한 수준 격차는 본 연구의 추정 자체의 편향이 아니라 산정방식의 구조적 차이로 설명된다. 첫째, 본 연구는 평균 기반 PIR을 채택하였는데, 평균은 분포 상단의 고가 자산에 의해 끌어올려지므로 동일 시장에 대해서도 중위 기반 PIR보다 큰 값으로 산출된다. 이는 Cox(2024)가 평균 기반 PIR이 분포 상단에 의해 왜곡되기 쉽기 때문에 정책 평가용으로는 중위 기반 median multiple을 권장한다고 지적한 점과도 부합한다. 둘째, 본 연구는 분석 대상을 아파트로 한정하였는데, 우리나라에서 아파트는 비아파트 주택 유형(다세대·연립·단독)에 비해 상대적으로 가격대가 높으므로 주거실태 조사의 전체 자가주택 기준 PIR보다 큰 값이 나타난다. 셋째, 본 연구의 서울 PIR이 특히 높게 나타나는 것은 서울이 시도 단위 비교에서 가장 가격대가 높을 뿐 아니라 분포 상단이 빠르게 팽창해 평균이 더 큰 폭으로 끌어올려지는 상단 팽창형 분포를 갖고 있기 때문이다.

이러한 차이를 통제하고 시기적 추세의 정합성을 살펴보면, 본 연구의 PIR과 공표 PIR은 매우 일관된 함의를 공유한다. 주거실태조사 기준 수도권 PIR은 2021년 9.5배에서 2022년 9.3배, 2023년 8.5배로 조정 패턴을 보이며(국토교통부, 2024), 이는 본 연구가 추정한 수도권 PIR의 정점 시점(2021년) 및 이후 조정 추이와 일치한다. 즉 본 연구의 PIR은 절대 수준의 측면에서는 산정방식의 차이로 인해 공표 수치보다 높게 산출되지만, 변동의 방향·시점·지역 간 순서에 있어서는 공표 수치와 일관된 결과를 제공한다(〈표 16〉, 〈표 17〉).

〈표 16〉 단위면적당 가격의 지니계수(단지·평형 기준, 세대수 가중)

시점	전국	서울	부산	대구	인천	광주	대전	울산	세종	경기	강원	충북	충남	전북	전남	경북	경남	제주
2012/03	0.282	0.204	0.137	0.141	0.119	0.125	0.151	0.148	0.115	0.174	0.168	0.182	0.170	0.182	0.158	0.178	0.174	0.157
2012/06	0.278	0.205	0.135	0.139	0.126	0.123	0.152	0.143	0.112	0.172	0.165	0.180	0.169	0.177	0.153	0.177	0.174	0.168
2012/09	0.272	0.201	0.136	0.134	0.128	0.122	0.153	0.140	0.112	0.167	0.163	0.178	0.167	0.174	0.154	0.173	0.172	0.153
2012/12	0.271	0.202	0.135	0.140	0.126	0.124	0.154	0.144	0.107	0.167	0.166	0.179	0.169	0.179	0.153	0.174	0.169	0.159
2013/03	0.268	0.204	0.137	0.133	0.129	0.123	0.154	0.144	0.117	0.165	0.161	0.174	0.165	0.175	0.151	0.170	0.171	0.162
2013/06	0.270	0.205	0.138	0.127	0.125	0.127	0.152	0.145	0.115	0.170	0.161	0.173	0.165	0.176	0.151	0.170	0.175	0.159
2013/09	0.266	0.203	0.139	0.120	0.129	0.125	0.151	0.143	0.113	0.171	0.159	0.167	0.164	0.178	0.152	0.171	0.172	0.156
2013/12	0.267	0.203	0.141	0.117	0.135	0.126	0.153	0.140	0.119	0.171	0.160	0.169	0.163	0.179	0.152	0.174	0.175	0.149
2014/03	0.268	0.208	0.142	0.120	0.128	0.124	0.150	0.139	0.134	0.173	0.160	0.171	0.164	0.182	0.159	0.175	0.176	0.154
2014/06	0.268	0.208	0.144	0.120	0.127	0.126	0.148	0.141	0.123	0.176	0.160	0.178	0.167	0.183	0.165	0.177	0.178	0.150
2014/09	0.269	0.212	0.145	0.120	0.128	0.126	0.147	0.142	0.126	0.179	0.156	0.180	0.168	0.182	0.166	0.179	0.180	0.150
2014/12	0.269	0.212	0.149	0.124	0.131	0.134	0.149	0.140	0.125	0.181	0.160	0.185	0.166	0.184	0.166	0.181	0.181	0.152
2015/03	0.270	0.214	0.151	0.132	0.127	0.135	0.151	0.142	0.091	0.183	0.156	0.181	0.165	0.186	0.169	0.182	0.181	0.159
2015/06	0.273	0.215	0.157	0.139	0.122	0.135	0.153	0.143	0.099	0.185	0.161	0.177	0.168	0.190	0.173	0.185	0.181	0.167
2015/09	0.274	0.213	0.162	0.144	0.124	0.138	0.154	0.143	0.100	0.184	0.161	0.175	0.169	0.192	0.174	0.186	0.179	0.170
2015/12	0.277	0.216	0.164	0.143	0.121	0.135	0.155	0.143	0.105	0.184	0.162	0.173	0.170	0.193	0.179	0.187	0.178	0.170
2016/03	0.276	0.214	0.167	0.138	0.125	0.134	0.156	0.145	0.117	0.182	0.162	0.173	0.171	0.195	0.182	0.184	0.176	0.175
2016/06	0.281	0.219	0.168	0.141	0.121	0.132	0.157	0.145	0.121	0.184	0.163	0.174	0.171	0.197	0.186	0.186	0.174	0.163
2016/09	0.289	0.226	0.174	0.144	0.120	0.136	0.158	0.145	0.156	0.186	0.166	0.174	0.174	0.199	0.186	0.191	0.175	0.167
2016/12	0.291	0.224	0.183	0.152	0.125	0.134	0.159	0.149	0.158	0.188	0.172	0.173	0.175	0.201	0.191	0.193	0.174	0.164
2017/03	0.293	0.227	0.183	0.155	0.127	0.140	0.160	0.149	0.126	0.188	0.171	0.170	0.176	0.202	0.194	0.199	0.173	0.164

〈표 16〉 계속

시점	전국	서울	부산	대구	인천	광주	대전	울산	세종	경기	강원	충북	충남	전북	전남	경북	경남	제주
2017/06	0.301	0.234	0.185	0.156	0.125	0.143	0.162	0.150	0.133	0.193	0.171	0.177	0.178	0.204	0.195	0.206	0.172	0.166
2017/09	0.308	0.235	0.183	0.166	0.128	0.146	0.164	0.154	0.137	0.200	0.170	0.183	0.181	0.203	0.202	0.210	0.172	0.170
2017/12	0.318	0.248	0.183	0.173	0.128	0.150	0.168	0.157	0.167	0.204	0.175	0.185	0.183	0.208	0.207	0.214	0.173	0.167
2018/03	0.338	0.267	0.185	0.181	0.128	0.161	0.174	0.163	0.147	0.221	0.174	0.188	0.188	0.211	0.205	0.222	0.180	0.173
2018/06	0.345	0.263	0.188	0.190	0.134	0.176	0.179	0.170	0.161	0.230	0.176	0.190	0.194	0.213	0.205	0.228	0.189	0.172
2018/09	0.367	0.266	0.192	0.199	0.136	0.189	0.191	0.175	0.169	0.247	0.183	0.198	0.198	0.217	0.207	0.234	0.192	0.174
2018/12	0.366	0.252	0.194	0.201	0.137	0.195	0.212	0.180	0.169	0.249	0.188	0.203	0.210	0.221	0.208	0.237	0.196	0.175
2019/03	0.365	0.247	0.199	0.207	0.143	0.201	0.219	0.193	0.175	0.247	0.195	0.209	0.216	0.225	0.213	0.242	0.203	0.175
2019/06	0.372	0.253	0.205	0.208	0.145	0.200	0.224	0.194	0.177	0.250	0.201	0.217	0.221	0.230	0.215	0.248	0.210	0.179
2019/09	0.385	0.263	0.209	0.206	0.149	0.201	0.228	0.199	0.187	0.261	0.203	0.221	0.226	0.236	0.220	0.253	0.215	0.182
2019/12	0.397	0.271	0.228	0.206	0.158	0.206	0.239	0.208	0.200	0.273	0.212	0.226	0.240	0.242	0.220	0.258	0.225	0.183
2020/03	0.404	0.260	0.242	0.209	0.179	0.207	0.245	0.222	0.205	0.282	0.218	0.235	0.252	0.258	0.228	0.266	0.234	0.188
2020/06	0.400	0.254	0.248	0.207	0.181	0.210	0.238	0.225	0.202	0.276	0.226	0.253	0.251	0.261	0.233	0.271	0.239	0.188
2020/09	0.415	0.249	0.265	0.216	0.181	0.213	0.248	0.242	0.188	0.286	0.232	0.258	0.267	0.272	0.245	0.272	0.251	0.187
2020/12	0.413	0.244	0.292	0.236	0.183	0.230	0.251	0.274	0.187	0.282	0.236	0.271	0.276	0.292	0.251	0.286	0.276	0.198
2021/03	0.414	0.243	0.294	0.224	0.193	0.240	0.255	0.272	0.177	0.280	0.253	0.291	0.292	0.297	0.257	0.293	0.283	0.220
2021/06	0.412	0.242	0.295	0.221	0.183	0.242	0.248	0.266	0.175	0.268	0.260	0.290	0.289	0.297	0.260	0.291	0.282	0.238
2021/09	0.410	0.241	0.295	0.217	0.187	0.245	0.242	0.269	0.166	0.259	0.265	0.288	0.285	0.299	0.257	0.292	0.287	0.253
2021/12	0.404	0.244	0.289	0.216	0.181	0.242	0.237	0.262	0.160	0.249	0.266	0.282	0.280	0.296	0.261	0.288	0.288	0.252
2022/03	0.400	0.244	0.282	0.210	0.170	0.237	0.227	0.256	0.156	0.245	0.267	0.275	0.274	0.292	0.258	0.285	0.283	0.250
2022/06	0.398	0.246	0.279	0.206	0.164	0.233	0.224	0.251	0.154	0.244	0.262	0.271	0.267	0.289	0.258	0.279	0.277	0.249
2022/09	0.391	0.248	0.267	0.203	0.157	0.223	0.216	0.244	0.153	0.237	0.253	0.258	0.257	0.280	0.250	0.272	0.271	0.234

〈표 16〉 계속

시점	전국	서울	부산	대구	인천	광주	대전	울산	세종	경기	강원	충북	충남	전북	전남	경북	경남	제주
2022/12	0.385	0.245	0.262	0.206	0.150	0.218	0.211	0.236	0.159	0.238	0.247	0.247	0.249	0.273	0.240	0.264	0.255	0.230
2023/03	0.388	0.248	0.257	0.209	0.158	0.216	0.213	0.235	0.167	0.244	0.242	0.245	0.249	0.265	0.244	0.259	0.254	0.225
2023/06	0.393	0.254	0.260	0.215	0.167	0.219	0.214	0.241	0.169	0.250	0.241	0.249	0.251	0.270	0.245	0.263	0.256	0.225
2023/09	0.398	0.260	0.264	0.222	0.173	0.221	0.225	0.249	0.172	0.255	0.243	0.254	0.254	0.275	0.249	0.267	0.262	0.222
2023/12	0.396	0.260	0.264	0.222	0.174	0.220	0.224	0.249	0.170	0.251	0.242	0.254	0.254	0.277	0.251	0.264	0.263	0.215
2024/03	0.397	0.261	0.265	0.225	0.176	0.220	0.221	0.251	0.172	0.253	0.240	0.257	0.257	0.280	0.250	0.265	0.263	0.219

주: 소수점 네자리 이하는 반올림되었다.

〈표 17〉 지역별 PIR 추이(단위: 배)

	2015	2016	2017	2018	2019	2020	2021	2022	2023
전국	6.14	6.37	6.60	6.88	6.91	7.82	9.40	8.77	7.91
서울	9.81	10.41	11.38	13.36	14.02	16.64	19.53	18.35	16.34
부산	5.64	6.21	6.57	6.19	5.89	6.51	8.02	7.44	6.64
대구	6.08	5.85	5.92	6.10	6.08	6.40	7.31	6.34	5.65
인천	6.12	6.30	6.34	6.03	5.84	6.54	8.24	7.74	6.80
광주	4.46	4.59	4.59	4.73	4.77	4.77	5.33	5.28	4.88
대전	4.65	4.65	4.68	4.62	4.88	5.80	6.86	6.16	5.65
울산	4.53	4.79	4.72	4.37	4.07	4.52	5.24	4.92	4.46
세종	4.81	5.42	6.15	6.63	6.90	9.43	10.46	8.88	8.19
경기	6.59	6.79	6.92	7.08	7.09	8.24	10.48	9.69	8.67
강원	3.62	3.84	4.02	3.82	3.49	3.47	4.02	4.26	4.08
충북	4.35	4.22	4.00	3.70	3.36	3.55	4.15	4.10	3.83
충남	4.35	4.32	4.20	3.89	3.74	3.83	4.38	4.30	3.95
전북	3.58	3.67	3.69	3.51	3.30	3.32	3.81	3.79	3.50
전남	3.33	3.43	3.58	3.61	3.53	3.55	3.71	3.52	3.26
경북	4.07	3.93	3.77	3.51	3.27	3.28	3.71	3.62	3.41
경남	4.99	4.95	4.76	4.34	4.06	4.18	4.83	4.78	4.43
제주	5.68	6.87	7.21	6.99	6.63	6.48	7.58	7.76	7.36

주: 1) 평균적 크기의 아파트(공급면적 기준 94.6m²)와 가구원수(2.2명)가 가정되었다.

2) 소수점 2자리 이하는 반올림되었다.

PIR, price to income ratio.

VI. 결론

본 연구는 대표적 의사결정나무 기반 알고리즘인 XGBoost를 활용하여, 전국 아파트 단지(단지 내 개별 평형별)의 가격을 전수적으로 추정하고, 이를 통해 시장 전체의 가격 수준과 분포를 직접 관측하는 방법을 제시하였다(부록 표 1)). 본 연구는 기계학습 기반 AVM이 어떻게 전통적 가격지수와 다른 차원의 시장 정보(절대적 가격 수준, 분포 구조, 하위 집단별 통계량 등)에 직접 접근하는 도구로 활용될 수 있는지를 보여준다.

본 연구의 추정 결과를 요약하면 크게 다음과 같다. 첫째, 2024년 기준 전국 아파트의 평균 가격은 470만 원/m²이며 서울의 경우 1,121만 원 수준인 것으로 나타났다. 둘째, 전국의 아파트 가격은 2018년부터 2021년까지 큰 폭으로 상승하였다가 2022년 이후 약한 조정 상태로 전환하였다. 이때 조정폭은 서울 및 수도권에 비해 지방도시가 더 큰 것으로 나타났다. 셋째, 서울과 지역 간의 가격 격차는 해당 기간 동안 크게 상승하였다. 넷째, 지역 간 가격 격차 뿐 아니라 각 지역 내에서의 가격 격차 역시 전반적으로 상승하였다. 다섯째, PIR 역시 전반적으로 상승하였으며, 특히 서울 지역에서 두드러지게 높아졌음을(2021년 기준 19.53) 확인할 수 있다.

본 연구의 기여는 최근 빠르게 발전하는 기계학습 기법의 활용을 통해 시장 전체의 가격 분포를 직접 관측하는 경로를 제안하여 최초로 시연한 데에 있다. 이 접근은 거래 표본 의존성에서 자유로우며, 가격의 절대적 수준을 제공하고, 연구자가 원하는 기준에 따라 표본을 분할하여 다양한 하위 집단별 통계량을 자유롭게 생성할 수 있다는 점에서 기존의 거래 표본 기반 정보 체계와 보완적이다.

또한, 본 연구에서 제시된 추정 기반 재고 주택가격 관측 기법은 평균이나 분포, 관련 지표의 생성 뿐 아니라 패널 분석, 이층차분 설계, 표본 균형을 고려한 선형회귀, 편집 가능한 주택가격지수 등 다양한 분석으로 확장될 여지가 많다는 것을 고려하면, 추후 풍부한 후속 연구들이 이루어질 것을 기대할 수 있을 것으로 보인다. 예를 들어, 아파트 거래 자료를 통해 헤도닉 모형을 추정하는 경우, 크기가 크고 거래가 자주 일어나는 단지 또는 지역의 거래 표본이 더 많이 포함되어 있기 때문에 표본 불균형으로 인한 편향적 추정의 가능성에 노출되어 있다. 그러나, 전체 거래 자료가 아닌 분석하고자 하는 지역의 개별 단지에 대한 신뢰할 만한 추정치를 표본으로 사용하는 경우 그러한 표본 불균형의 가능성을 줄일 수 있을 것이다.

본 연구의 한계는 다음과 같다. 첫째, 본 연구는 특정한 하나의 가설보다는 시장의 전반적 상태를 관찰하고 재고 기반 주택가격 관측의 가능성을 제시하는 데에 초점이 맞춰져 있으므로, 그 자체로서는 분석적 해석을 제공하지 않는다는 한계가 있을 수 있다. 본 연구에서 관측된 가격 격차의 확대, PIR의 상승 등에 대해 본 연구는 해석의 단서를 제공할 뿐 인과적 결론을 제시하지는 않는다.

둘째, 본 연구에서는 전수적 추정을 통한 가격지수 산정 방법론 자체를 심도 깊게 다루지 않았다. 가격지수의 산정에는 시기별 재고 주택의 변화(멸실 및 신규 입주)에 대한 보정과 가중 방식 선택의 정당화라는 별도의 논의가 요구된다. 본 연구의 단지 자료는 분석 시점 기준의 단일 시점 정보로 고정되어 사용되었으므로, 분석 기간 중에 발생한 입지 속성의 변화(지하철 노선의 개통, 학교의 신설·폐교, 재건축 추진단계의 변화 등) 자체가 시장 가격에 미친 효과를 분리하여 관측하는 데에는 본 연구의 설계가 적합하지 않다. 향후 연구에서는 본 연구의 전수 추정치 위에 적절한 가중·보정 체계를 결합하여 정식 가격지수를 구성하는 방향으로 확장될 수 있다.

셋째, 본 연구가 활용한 기계학습 기반 AVM 자체에도 잠재적 한계가 있을 수 있다. 거래 빈도가 잦은 주택에 대한 과적합 가능성, 급격한 가격 변동 시기의 예측력 저하, 학습 자료가 희소한 지역에서의 정확도 감소 등이 그 예이다. 본 연구의 IV장 평가 결과는 모든 시·도에서 모형이 평균적으로 거의 불편향에 가까운 정확성을 보임을 확인하였으나, 특정 시점·지역에서의 국지적 편향 가능성을 완전히 배제하지는 못한다.

넷째, 본 연구의 IV장 평가에서 관측된 17개 시·도 간 모형 성능의 격차는 추정 결과의 지역별 해석에 함의를 가지므로 별도의 논의가 필요하다. $\pm 10\%$ 허용 오차 기준의 적중률을 보면 수도권 3개 시·도와 대구·대전은 90% 이상에 위치하는 반면, 강원·충북·전북·전남 등 비수도권 도지역은 81%~85% 수준, 경북과 제주는 80%를 다소 하회하는 수준에 머문다. 이러한 지역 간 격차는 (i) 거래 빈도의 차이로 인한 학습 표본 밀도의 차이, (ii) 비수도권 일부 지역의 단지 규모 분포가 소규모 단지에 치우쳐 단지 간 이질성이 상대적으로 큰 점, (iii) 해당 지역의 주택 재고에서 아파트가 차지하는 비중이 상대적으로 작아 아파트 시장의 가격 형성 구조가 수도권·광역시 대비 표준화 정도가 낮을 가능성 등이 복합적으로 작용한 결과로 추정 가능하다. 다만 IV.2절에서 확인한 바와 같이 모든 지역에서 IAAO(2017)의 동질적 시장 권고 기준($COD \leq 10$)을 충족하며 평균백분오차(MPE)가 제주를 제외한 모든 지역에서 $\pm 1\%$ 이내에서 관리되는 점은, 지역별 평균이나 분포 등 집단 단위 통계량 산출에 대해서는 모든 지역에서 일정 수준 이상의 신뢰성이 확보됨을 의미한다. 따라서 본 연구의 추정 결과를 정책적 목적으로 활용할 때에는 수도권 및 광역시에 대해서는 비교적 정밀한 정량적 해석이 가능한 반면, 비수도권 도지역, 특히 경북·제주 등에 대해서는 결과 통계량의 방향성과 시계열적 추이는 신뢰할 수 있되 절대 수준의 미세한 차이에 대한 해석에는 보다 신중한 접근이 요구된다.

이상의 한계에도 불구하고, 본 연구는 기계학습 기반 AVM의 활용 영역을 단순 가격 예측을 넘어 시장 전체의 가격 분포·불균등성·접근성에 대한 직접 관측으로 확장하였다는 점에서 의의를 갖는다. 이러한 확장은 전통적 가격지수가 제공하지 못하던 정보 차원을 학술 및 정책 영역에 새롭게 더해 줄 수 있다는 점에서 후속 연구의 토대로 활용될 수 있을 것이다. 향후에는 본 연구의 전수 추정치

위에 적절한 가중·보정 체계를 결합하여 정식 가격지수를 구성하는 방법론적 발전, 비아파트 주택 유형 및 보다 미시적 공간 단위로의 확장, 그리고 본 연구가 산출한 분포·격차·PIR 지표를 활용한 정책·인과 분석 등 다양한 방향의 후속 연구가 이어질 수 있을 것으로 기대된다.

참고문헌

- 국토교통부. (2023). *2022년도 주거실태조사 결과*. 세종: 국토교통부.
- 국토교통부. (2024). *2023년도 주거실태조사 결과*. 세종: 국토교통부.
- 권진웅, 김남정 (2020). 서울시 자치구별 아파트 Price-to-Income Ratio(PIR) 변동 특성에 관한 연구. *주택도시금융연구*, 5(2), 55-72.
- 김세울, 유선종. (2022). 교육환경과 아파트 전월세가격 상관성에 관한 실증연구: 서울시 송파구를 중심으로. *주택도시연구*, 12(1), 1-18.
- 김이환, 김형준, 류두진, 조훈. (2022). 기계학습 방법론을 활용한 아파트 매매가격지수 연구. *부동산 분석*, 8(3), 1-29.
- 김지원, 마강래. (2021). 거주지에 따른 자산 격차에 관한 연구: 주택 보유여부를 중심으로. *대한부동산학회지*, 39(2), 5-28.
- 배성완, 유정석. (2018). 머신 러닝 방법과 시계열 분석 모형을 이용한 부동산 가격지수 예측. *주택연구*, 26(1), 107-133.
- 서재원, 진장익. (2021). 주택가격의 불평등은 계층이동 가능성의 기대감을 낮추는가? *국토계획*, 56(7), 141-156.
- 송영호, 마강래. (2025). 거주지역과 자산 축적 간의 관계에 관한 연구: 주택 보유와 자산 분위에 따른 차이를 중심으로. *국토계획*, 60(1), 113-126.
- 이소영, 김정민. (2025). 머신러닝을 활용한 다세대주택 매매가격지수 산정에 관한 연구: 서울시 소지역 단위를 중심으로. *주택도시금융연구*, 10(2), 93-127.
- 이인호, 이경섭. (2020). 트리 기반 앙상블 방법을 활용한 자동 평가 모형 개발 및 평가: 서울특별시 주거용 아파트를 사례로. *한국데이터정보과학회지*, 31(2), 375-389.
- 이창무, 김현아, 조만. (2012). 소득대비 주택가격 비율(PIR)의 산정방식 및 그 수준에 대한 국제비교. *주택연구*, 20(4), 5-25.
- 정의철, 김진욱, 하두나. (2009). 부동산소득이 소득불평등에 미치는 영향 분석. *주택연구*, 17(2), 5-28.
- 통계청. (2018). *2017년 가계금융복지조사 결과*. 대전: 통계청.
- 통계청. (2022). *2022년 가계금융복지조사 결과*. 대전: 통계청.
- 한국주택금융공사. (2020). *2020년 2분기 주택금융 및 유통화증권 통계: 2020년 1분기 기준*. 부산: 한국주택금융공사.
- 허원제, 오경수, 신미정, 유동영. (2020). 수도권·비수도권 간의 주택가격 격차 요인 분해에 관한

- 연구. *한국지방세연구원 기본연구보고서*, 2019(6), 1-91.
- 홍정의. (2020). 기계학습 알고리즘을 이용한 주택가격감정 시스템의 구축 및 평가: XGBoost, LightGBM, CatBoost 알고리즘에 기반하여. *주택금융연구*, 4, 33-64.
- 홍정의. (2021). 랜덤 포레스트 알고리즘을 통한 주택 대량평가모형 연구. *부동산분석*, 7(1), 1-28.
- KB국민은행. (2020). *KB주택가격동향 보고서: 2020년 5월*. 서울: KB국민은행.
- Antipov, E. A., & Pokryshevskaya, E. B. (2012). Mass appraisal of residential apartments: An application of random forest for valuation and a CART-based approach for model diagnostics. *Expert Systems with Applications*, 39(2), 1772-1778.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32.
- Cox, W. (2024). *Demographia international housing affordability: 2024 edition*. Orange County, CA: Newgeography.
- Ekeland, I., Heckman, J. J., & Nesheim, L. (2004). Identification and estimation of hedonic models. *Journal of Political Economy*, 112(S1), S60-S109.
- Gallin, J., Molloy, R., Nielsen, E., Smith, P., & Sommer, K. (2021). Measuring aggregate housing wealth: New insights from machine learning. *Journal of Housing Economics*, 51, 101734.
- Gatzlaff, D. H., & Haurin, D. R. (1997). Sample selection bias and repeat-sales index estimates. *The Journal of Real Estate Finance and Economics*, 14, 33-50.
- Guliker, E., Folmer, E., & van Sinderen, M. (2022). Spatial determinants of real estate appraisals in the Netherlands: A machine learning approach. *ISPRS International Journal of Geo-Information*, 11(2), 125.
- Hill, R. J. (2013). *Hedonic price indexes for housing* (OECD Statistics Working Paper No. 2011/01). Paris, France: Organization for Economic Co-operation and Development.
- Ho, W. K. O., Tang, B. S., & Wong, S. W. (2021). Predicting property prices with machine learning algorithms. *Journal of Property Research*, 38(1), 48-70.
- Hong, J., Choi, H., & Kim, W. (2020). A house price valuation based on the random forest approach: The mass appraisal of residential property in South Korea. *International Journal of Strategic Property Management*, 24(3), 140-152.
- Hong, J., & Kim, W. (2022). Combination of machine learning-based automatic valuation models for residential properties in South Korea. *International Journal*

of Strategic Property Management, 26(5), 362-384.

International Association of Assessing Officers (IAAO). (2017). *Standard on ratio studies*. Kansas City, MO: IAAO.

International Association of Assessing Officers (IAAO). (2018). *Standard on mass appraisal of real property*. Kansas City, MO: IAAO.

Kang, S., & Koo, J. H. (2023). Exploring social capital level in regions with large and increasing wealth inequality: Lesson from Seoul, South Korea. *Social Indicators Research*, 168(1), 165-183.

Kim, W., & Hong, J. (2024). Stacked ensemble model for the automatic valuation of residential properties in South Korea: A case study on Jeju Island. *Land*, 13(9), 1436.

Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of Political Economy*, 74(2), 132-157.

Malpezzi, S. (2003). Hedonic pricing models: A selective and applied review. In T. O'Sullivan, & K. Gibb (Eds.), *Housing economics and public policy* (pp. 67-89). Oxford, UK: Blackwell Science.

Melser, D. (2023). Selection bias in housing price indexes: The characteristics repeat sales approach. *Oxford Bulletin of Economics and Statistics*, 85(3), 623-637.

Ming, Y., Zhang, J., Qi, J., Liao, T., Wang, M., & Zhang, L. (2020). Prediction and analysis of Chengdu housing rent based on XGBoost algorithm. In: *ICBDT 20: Proceedings of the 3rd International Conference on Big Data Technologies* (pp. 1-5). New York, NY: Association for Computing Machinery.

Pace, R. K., & Hayunga, D. (2020). Examining the information content of residuals from hedonic and spatial models using trees and forests. *Journal of Real Estate Finance and Economics*, 60(1), 170-180.

Park, B., & Bae, J. K. (2015). Using machine learning algorithms for housing price prediction: The case of Fairfax County, Virginia housing data. *Expert Systems with Applications*, 42(6), 2928-2934.

Rico-Juan, J. R., & de La Paz, P. T. (2021). Machine learning with explainability or spatial hedonics tools? An analysis of the asking prices in the housing market in Alicante, Spain. *Expert Systems with Applications*, 171, 114590.

- Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, 82(1), 34-55.
- Sheppard, S. (1999). Hedonic analysis of housing markets. In P. Cheshire, & E. S. Mills (Eds.), *Handbook of regional and urban economics* (pp. 1595-1635). Amsterdam, The Netherlands: Elsevier.
- Tekouabou, S. C. K., Gherghina, Ş. C., Kameni, E. D., Filali, Y., & Idrissi Gartoumi, K. (2024). AI-based on machine learning methods for urban real estate prediction: A systematic survey. *Archives of Computational Methods in Engineering*, 31(2), 1079-1095.
- Yilmazer, S., & Kocaman, S. (2020). A mass appraisal assessment study using machine learning based on multiple regression and random forest. *Land Use Policy*, 99, 104889.

(논문 접수일: 2026.03.16. 수정논문 접수일: 2026.05.27. 논문 채택일: 2026.06.05.)

A Tree Ensemble Approach to Estimating Census-Level House Price Distributions

Jengei Hong*

Abstract

The sample includes 6,118,730 apartment transactions from March 2012 to April 2024 and complex- and floor-plan-level information on 12,517,131 housing units nationwide (99.1% of all apartment units in the Population and Housing Census). The study estimates quarterly prices for virtually the entire apartment stock in Korea. From these estimates, regional average prices, price distributions, and the price-to-income ratio (PIR) are derived. Our estimates show that average apartment prices rose more rapidly than the official housing price index during the 2017–2022 price boom, and that, over the analysis period, both cross-regional price gaps and within-region gaps across price quantiles widened simultaneously.

Keywords : XGBoost, Machine Learning, Automated Valuation Model (AVM), Housing Price Distribution, Price to Income Ratio (PIR)

* Jengei Hong, Corresponding author, Associate Professor, School of Management and Economics, Handong Global University, hwgh024@handong.edu

© Copyright 2026 Housing Finance Research Institute. 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.

부록

〈부록 표 1〉 XGBoost 모형 하이퍼 파라미터 및 난수 설정

구분	파라미터	값	설명
하이퍼파라미터	n_estimators	400	부스팅 라운드 수(트리 개수)
	learning_rate	0.06	학습률(각 트리의 기여도)
	max_depth	20	개별 트리의 최대 깊이
	gamma	10	추가 분할에 필요한 최소 손실 감소
	reg_alpha	1.5	가중치에 대한 L1 정규화 계수
	reg_lambda	1.2	가중치에 대한 L2 정규화 계수
	colsample_bytree	0.5	각 트리 생성 시 사용하는 변수의 비율
난수 시드	seed	10	재현성 확보를 위한 난수 시드

주: 위에 명시되지 않은 파라미터는 XGBoost의 기본값(default)을 사용함.

아파트 거래량과 가격 변동성의 관계: 거래량 기반 가격 불확실성 지수의 구축과 활용

김천일*

요약

주택시장에서 관측되는 가격 변동성에는 가치의 실제 변동과 거래 회소성에서 오는 표본 잡음이 혼재되어 있으나, 기존 지표는 이 두 성분을 구분하지 못한다. 본 연구는 헤도닉 가격모형 잔차의 이동 구간 표준편차 및 거래 건수를 반영한 거래량 기반 가격 불확실성 지수(price uncertainty index, PUI)를 구축하였다. PUI와 기존 변동성 지수(realized volatility, RV)의 상관계수는 0.38에 불과하여 두 지표가 상이한 정보를 담고 있음을 확인하였다. 거래량의 변동성 설명력은 가격 잡음의 시변성을 경유하여 작동하였으며, PUI는 RV가 포착하지 못하는 미래 거래량 변화를 유의하게 설명하였다. 호황기·경색기 비대칭 분석에서는 거래가 충분한 시기에는 추가 거래가 변동성에 영향을 주지 않지만 경색기에는 변동성 증폭 효과가 강하게 나타났으며, 이 비대칭적 반응은 PUI에서 더 뚜렷하게 관찰되었다.

핵심어 : 주택가격 변동성, 가격 불확실성 지수, 거래량, 헤도닉 가격모형, 거래량-변동성 비대칭성

1. 서론

주택가격 변동성은 가계 자산의 안정성, 금융기관의 담보가치 평가, 정부의 시장 안정화 정책에 이르기까지 광범위한 경제적 의사결정에 영향을 미친다. 변동성이 높은 시장에서는 가계가 주거 이동과 자산 배분의 최적 시점을 판단하기 어려워지고, 금융기관은 주택담보대출의 리스크를 과대

* 김천일, 주저자, 강남대학교 부동산건설학부 조교수, ckim@kangnam.ac.kr

© Copyright 2026 Housing Finance Research Institute. 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.

또는 과소 평가할 수 있으며, 정책 수립 기관은 시장 개입의 적정 시점과 강도를 결정하는 데 어려움을 겪는다. 이러한 이유로 주택가격 변동성을 어떻게 정확히 측정하고 해석할 것인가는 실무와 정책 양면에서 중요한 사안으로 인식되어 왔다.

그러나 주택시장에서 관측되는 가격 변동성에는 근본적인 측정 문제가 존재한다. 주택은 주식이나 채권과 달리 가격이 거래 시점에만 관측된다. 연속적으로 호가가 형성되는 금융자산과 달리, 주택의 가격은 매수자와 매도자가 탐색과 협상을 거쳐 거래가 성사되는 순간에만 드러난다. 이는 관측되는 가격 변동 내에 두 가지 이질적인 성분이 혼재되어 있음을 의미한다. 하나는 해당 지역 주택의 근본적 가치(fundamental value)가 실제로 변동하는 것이고, 다른 하나는 거래가 드물어 소수의 거래로부터 가격 수준을 추정할 때 발생하는 표본 잡음(sampling noise)이다.

같은 수준의 가치 변동이 있더라도, 분기에 수천 건이 거래되는 시장에서의 관측 변동성과 수십 건만 거래되는 시장에서의 관측 변동성은 질적으로 다르다. 전자는 풍부한 거래 정보가 가격 잡음을 평활화하여 관측 변동성이 실제 가치 변동에 가깝게 되지만, 후자는 소수 거래의 특이성(급매, 프리미엄 거래, 특수한 거래 조건 등)이 가격지수에 과도하게 반영되어 관측 변동성이 부풀려진다. 문제는 기존의 변동성 지표가 이 두 원천을 구분하지 못한다는 것이다. 관측 변동성이 높다는 사실만으로는 실제로 시장이 불안정한 것인지, 단순히 거래가 적어서 가격 신호가 거칠게 보이는 것인지를 판단할 수 없다.

거시·정책 충격이 한국 주택시장의 거래량과 가격에 미치는 영향이 크다는 점은 잘 알려져 있으며, 금리, 세제, 대출 규제 등 외생적 정책 변수가 시장 변동의 중요한 인자로 작용해 왔다. 다만 본 연구는 이러한 충격의 원천을 식별하는 데 목적이 있는 것이 아니라, 어떤 원인인든 그 결과로 나타나는 거래 희소성과 관측 변동성 사이의 측정 문제에 초점을 둔다. 외생적 충격으로 거래가 위축되든 내생적 피드백으로 거래가 줄어들든, 가격 신호에 표본 잡음이 침투하는 메커니즘은 유사하게 작동하기 때문이다.

이러한 성질에 주목하여 본 연구는 주택시장의 관측 변동성에서 거래 희소성이 만들어내는 잡음 성분을 분리·측정하는 새로운 지표를 제안하였다. 헤도닉 가격모형을 추정한 후 모형으로부터 잔차를 추출하여 지역(시/구)별 분기별 가격 잡음의 크기를 이동 구간(롤링 윈도우, rolling window)에 따라 산출하고, 이를 해당 기간의 거래건수 제곱근으로 나눈 「거래량 기반 가격 불확실성 지수(price uncertainty index, PUI)」를 구축하였다. PUI의 분자는 개별 거래에 내재된 가격 잡음의 크기를 나타내고, 분모는 그 잡음을 평활화할 수 있는 거래의 수를 포함한다. 가격 잡음이 크거나, 거래가 적거나, 두 사건이 동시에 발생하는 경우 PUI가 상승한다. 이는 해당 시장의 가격 정보가 거칠게 시현(示現)되고 있음을 뜻한다.

실증 분석을 통해 검증하고자 하는 연구 질문은 다음과 같다. 본 연구는 다음의 세 가지 질문에 답하고자 한다. 첫째, 거래량과 관측 변동성(가격지수 수익률의 이동 구간 표준편차, realized volatility, RV) 사이에는 어떠한 관계가 존재하며, 이 관계는 가격 잡음의 시변성을 어떻게 경유하는가? 둘째, 기존 RV에 PUI를 추가하면 미래 가격 변동 및 거래량 변동에 대한 설명력이 증가하는가? 셋째, 거래량-변동성 관계는 시장 국면(호황기 vs 경색기)에 따라 비대칭적이며, 이 비대칭적 반응은 RV와 PUI에서 어떻게 다르게 나타나는가? 이러한 질문에 답하기 위해 아파트 실거래가 데이터를 활용하여 가격 불확실성 지수를 산출하고, 시/구의 분기별 패널 데이터를 구축하여, 거래량 및 불확실성 지수에 대해 패널 회귀, 설명력 검증, 호황기·경색기 비대칭 분석을 수행하였다.

본 논문의 구성은 다음과 같다. 제2장에서는 탐색-매칭 구조, 금융 미시구조, 거래량-변동성 관계에 관한 선행연구를 검토하고 본 연구의 차별성을 제시한다. 제3장에서는 거래 회소성이 관측 변동성을 내생적으로 결정하는 이론적 틀을 전개한다. 제4장에서는 분석에 활용된 데이터, 헤도닉 모형, PUI 구축 방법, 패널데이터 구축 절차, 실증 분석 절차를 서술한다. 제5장에서는 기술통계, 패널 회귀, 설명력 검증, 비대칭 분석의 결과를 보고한다. 제6장에서는 연구의 요약, 시사점, 한계 및 향후 과제를 논의한다.

II. 선행연구 검토 및 본 연구의 차별성

1. 해외 선행연구 검토

주택시장에서는 연속 거래가 이루어지는 금융자산 시장과 달리, 탐색 마찰과 비정기적 거래에 의해 가격이 형성된다(Wheaton, 1990). 이러한 시장에서 공실, 거래량, 가격은 탐색-매칭(search-matching) 구조를 통해 결정되며, 거래 구조가 가격 수준과 변동성에 내생적으로 영향을 미친다(Díaz & Jerez, 2013; Genesove & Han, 2012). Wheaton(1990)은 매칭 함수에 의해 거래가 결정되고 이 과정에서 가격과 공실이 동태적으로 조정됨을 보였으며, Anenberg & Bayer(2020)는 거래량 충격이 가격뿐만 아니라 변동성의 변화를 내생적으로 유발할 수 있음을 보여주었다. 거래량이 외생적으로 주어지는 변수가 아니라, 잠재 매수-매도자의 진입과 탐색 기제에 의해 내생적으로 결정된다는 점을 밝힌 것이다.

금융자산 시장 연구들은 고빈도 자료를 활용하여, 관측 변동성을 기초(fundamental) 변동성과 미시구조 노이즈로 분해하는 연구를 활발히 진행해왔다(Ait-Sahalia et al., 2011; Bandi &

Russell, 2006). 이들은 호가 단위, 스프레드, 비동시성 등의 미시구조 요인이 고빈도 수익률의 분산을 부풀리는 메커니즘을 탐색하였다. 이러한 메커니즘을 주택시장에 직접 적용한 연구는 드물지만, Anenberg & Laufer(2017)는 매물가격 자료가 실거래 자료보다 빠르게 관측된다는 점을 이용해 기존 주택가격 지수보다 수개월 앞서 주택가격 흐름을 포착할 수 있음을 보였다. Hill et al.(2022)은 주간 단위와 같은 고빈도 지수에서는 시의성과 신뢰성 사이의 절충이 불가피하며, 실거래가격 자료의 공개 지연이 큰 경우 매물가격 자료를 활용할 필요가 있음을 지적하였다. 이 두 연구는 주택 거래자료가 얼마나 충분히 축적되는가의 문제를 지수 산정의 쟁점으로 다루었다.

거래량, 가격, 변동성 간 관계에 대해서도 이론 및 실증 연구가 풍부하게 진행되어 왔다. Stein(1995)은 가격과 거래량의 동시 변동을 레버리지 제약과 초기 지불액 효과의 관점에서 설명하였고, Piazzesi & Schneider(2009)는 가격 상승기의 거래 급증과 하락기의 거래 절벽·가격 조정 지연 패턴을 모멘텀 거래자의 행태를 반영하여 모형화하였다. Head et al.(2014)은 탐색 마찰과 거래 유동성이 주택가격과 신규 건설을 결정하는 메커니즘임을 보였다.

2. 국내 선행연구 검토

국내 선행연구들은 주택가격 변동성이 일정한 분산을 갖는 과정이 아니라 군집성과 지속성을 보이는 현상임을 확인해 왔으며, 아파트 시장에서는 비대칭 변동성이, 일부 주택유형에서는 대칭 변동성이 더 적합하게 나타난다는 점을 보였다(임재만, 2006). 또한 낮은 변동성 국면이 장기간 지속되는 가운데 변동성의 크기에 따라 설명변수의 유의성과 방향이 달라질 수 있음이 제시되었고(김대원·유정석, 2014), 기대인플레이션과 비기대인플레이션에 대한 반응을 통해 주택의 인플레이션 헤지 가능성도 논의되었다(심성훈, 2010). 반면 가격안정 정책은 변동성 완화에 일관된 효과를 보이지 않았으며, 일부 시기를 제외하면 안정화 효과가 제한적이라는 평가도 제시되었다(곽승준·이주석, 2006).

경기국면 관련 연구들은 주택가격 변동을 단일 평균 반응으로 설명하기 어렵다는 점을 보여주었다. 전국과 수도권, 비수도권 비교에서는 모든 지역에서 수축기 지속확률이 더 높았으나 수도권은 확장기의 지속성과 강도가 상대적으로 크게 나타났고(전해정, 2015), 서울의 매매, 전세, 월세시장에서는 불황기와 호황기에서 유의한 설명변수가 서로 다르게 나타났다(전해정, 2017). 수도권 내부 비교에서도 서울은 확장기가 더 길고 경기와 인천은 수축기가 더 길게 나타났으며, 전세가격, 금리, 통화량의 영향 역시 지역별로 차별적으로 작용하였다(이현미·전해정, 2020).

거시경제 변수와 거래량을 함께 다룬 연구들은 가격변동과 거래량변동이 동일한 충격에 동일하게 반응하지 않음을 보여주었다. 산업생산, 회사채수익률, 주택담보대출 등은 주택가격과 유의한 동태

적 상관관계를 보였고(전해정·박헌수, 2012), 금리, 대출, 보증, 전세가격, 인구 및 경기 변수는 장기와 단기에서 서로 다른 방식으로 작용하였다(윤성민 외, 2016). 또한 거래량과 가격 사이에는 상호 인과 관계가 존재하지만 거래량 증감이 언제나 동일한 가격반응으로 이어지지 않는 것으로 나타났다(임재만, 2011), 수요충격은 가격변동에, 공급충격은 거래량 변동에 더 크게 작용하고, 이자율 및 가계대출 충격은 가격 쪽에 더 민감하게 반응하는 것으로 나타났다(민선옥·이영수, 2019).

지역 간 파급효과 연구들은 변동성과 가격변화가 공간적으로 전이된다는 점을 확인하였다. 수도권에서는 공통요인의 설명력이 높고 서울 강남과 버블7 지역이 변동성 순수출 지역으로 나타났으며(박영준·김기호, 2017), 대도시 분석에서는 상승기와 하락기의 변동성 전이가 비대칭적이라는 결과가 제시되었다(김상배·이승아, 2021). 극단적 상승 국면에서 전이 강도가 더 커지고 강남이 대표적 전달자로 작동한다는 점도 확인되었으며(고희운·강상훈, 2023), 최근 장기 시계열 분석 역시 팬데믹 급등기와 2022년 이후 급락기에서 동조화가 강화되고 서울, 특히 강남의 정보 전달 기능이 두드러졌음을 보였다(윤성민, 2024). 한편 동일 시장에서도 지수의 산정방식에 따라 위험과 초과수익률의 관계가 달라졌다는 결과(이정근·전재범, 2020)는 관측된 변동성이 순수한 가치 변동만이 아니라 거래정보의 질과 측정방식에도 영향을 받는다는 점을 시사한다.

3. 본 연구의 차별성

본 연구의 차별성은 다음과 같다. 첫째, 주택시장에서 관측 변동성을 「가치 변동 성분」과 「거래 구조에서 오는 잡음 성분」으로 분해하는 새로운 지표를 구축한다. Anenberg & Laufer(2017)가 매물자료를 활용하여 계약일 기준 또는 실시간에 가까운 주택가격지수의 시의성 문제를 보정하고, Hill et al.(2022)이 헤도닉 기반 지수의 저거래량 기간 처리와 고빈도 지수의 안정성 문제를 다룬 데 비해, 본 연구는 낮은 거래량을 지수 산정상의 기술적 제약이 아니라 가격 신호의 불확실성을 발생시키는 요인으로 해석한다. 따라서 본 연구의 불확실성 지수(PUI)는 가격지수를 더 빨리 산정하거나 저거래량 구간을 보정하기 위한 방법이 아니라, 거래 희소성이 관측 변동성에 포함된 표본 잡음으로 전환되는 정도를 직접 측정하는 지표라는 점에서 차별성이 있다. 둘째, 한국 아파트 시장의 실거래 데이터를 활용한 패널 데이터를 통해 PUI가 RV에 대한 설명력을 가짐을 확인하였다. 셋째, 거래량-변동성 관계의 시장 국면별 비대칭성을 분석함에 있어, 동일한 거래량 변화에 대해 RV와 PUI가 보이는 반응의 차이를 비교하였다. 본 연구는 헤도닉 잔차의 흠어짐(산포, 퍼짐)과 거래 희소성을 결합한 새로운 지표(PUI)를 구축하고, 이를 통해 기존 관측 변동성 지표만으로는 식별하기 어려웠던 거래량과 변동성 사이의 국면별 비대칭 관계를 실증한다는 점에서 기존 연구와 차별된다.

III. 거래량 기반 가격 불확실성 지수의 도출

1. 주택가격 관측의 두 특성: 시점 희소성과 가격 잡음

주택가격은 주식이나 채권과 구별되는 두 가지 관측상의 특성을 가진다. 이 두 특성은 모두 주택 거래의 고유한 구조에서 비롯되지만, 서로 다른 차원의 문제이다. 첫 번째 차원은 시점 희소성(sparsity)이다. 주택가격은 거래가 성사되는 시점에만 관측된다. 연속적으로 호가가 형성되는 금융 자산과 달리, 주택의 가격은 매수자와 매도자가 탐색과 협상을 거쳐 거래가 성사되는 순간에만 외부로 드러난다. 따라서 분기·연 단위로 거래가 드문 지역에서는 해당 지역의 가격 수준이 소수의 거래로부터 터만 추론되어야 한다.

두 번째 차원은 가격 잡음(noise)이다. 거래 시점에 관측되는 개별 가격은 해당 주택의 잠재가치를 정확히 반영하는 것이 아니라 일정한 오차를 포함한다. 지역 i 에 위치한 주택 j 의 시점 t 에서의 잠재가치를 v_{jit}^* 라 하자. 이 잠재가치는 입지, 주택 고유의 특성, 거시경제 조건 등에 의해 연속적으로 결정되지만 시장에서는 이 가치가 직접적으로 관측되지는 않는다. 가격이 관측되는 것은 주택 j 가 실제로 거래될 때뿐이며, 이때 관측가격 p_{jit} 는 $p_{jit} = v_{jit}^* + \epsilon_{jit}$ 로 표현된다. 여기에서 ϵ_{jit} 는 평균이 0이고 분산이 $\sigma_{\epsilon_{jit}}^2$ 인 가격 잡음이다. 이 잡음은 매수자와 매도자 간 협상력의 차이, 거래 시급성(급매 여부), 개별 주택 내 미세한 상태 차이(리모델링, 하자 등) 및 기타 정보 비대칭성으로 인해 발생한다. 가격 잡음은 개별 거래에 대해 고유한 것으로서 $E[\epsilon_{jit}] = 0$ 이고 서로 다른 거래 간에 독립이라고 가정($\text{Cov}(\epsilon_j, \epsilon_k) = 0$ ($j \neq k$))한다.

시점 희소성과 가격 잡음은 서로 독립된 현상으로서, 한쪽이 다른 쪽의 원인이 되는 관계가 아니다. 시점 희소성은 거래 자체의 빈도에 관한 문제이고, 가격 잡음은 개별 거래가 잠재가치로부터 벗어나는 정도에 관한 문제이다. 다만 두 특성은 가격지수의 추정 정밀도를 통해 결합되며, 본 연구의 거래량 기반 가격 불확실성 지수(PUI)는 이 결합 구조를 직접 측정하기 위한 지표이다.

2. 가격지수의 추정 오차

주택시장 참여자 및 주택정책을 수립하는 공공기관은 개별 주택의 가격보다는 지역 i 의 시점 t 에서의 가격 수준(즉, 가격지수)에 더 큰 관심을 가진다. 이 가격지수는 주택 특성의 영향을 제거(통제)한 뒤의 표준적 시장 가격 수준을 말한다. 주택 특성을 통제하는 일반적인 방법은 헤도닉 가격모형으로서

〈식 1〉과 같다.

$$\ln p_{jit} = \mu_{it} + \mathbf{x}'_j \beta_i + \alpha_{d(j)} + \epsilon_{jit} \quad \langle \text{식 1} \rangle$$

여기에서 μ_{it} 는 지역 i 의 시점 t 에서의 로그가격지수이며, \mathbf{x}_j 는 주택 j 의 관측 가능한 특성 벡터(전용면적, 층수, 경과 연수 등), $\alpha_{d(j)}$ 는 주택 j 가 속한 단지 d 의 고정효과이다. 〈식 1〉을 정리하여, 시점 t 에 거래된 주택 j 에 대해 주택 특성과 단지 고정효과를 제거한 로그가격을 $\tilde{p}_{jit} = \ln p_{jit} - \mathbf{x}'_j \hat{\beta}_i - \hat{\alpha}_{d(j)}$ 로 정의하면 $\tilde{p}_{jit} = \mu_{it} + \epsilon_{jit}$ 이다. 로그가격지수 μ_{it} 의 추정치 $\hat{\mu}_{it}$ 는 지역 i , 시점 t 에서의 거래량 N_{it} 의 특성 제거 로그가격의 평균이므로 $\hat{\mu}_{it}$ 는 〈식 2〉와 같다.

$$\hat{\mu}_{it} = \frac{1}{N_{it}} \sum_{j=1}^{N_{it}} \tilde{p}_{jit} = \mu_{it} + \frac{1}{N_{it}} \sum_{j=1}^{N_{it}} \epsilon_{jit} \quad \langle \text{식 2} \rangle$$

로그가격지수 추정량 $\hat{\mu}_{it}$ 이 진정한 로그가격지수 μ_{it} 로부터 벗어나 있는 추정 오차는 $\hat{\mu}_{it} - \mu_{it}$ 이다. 이 추정 오차의 분산을 구하면, 잔차 간 독립 가정에 의해,

$$\text{Var}(\hat{\mu}_{it}) = \text{Var}\left(\frac{1}{N_{it}} \sum_{j=1}^{N_{it}} \epsilon_{jit}\right) = \frac{1}{N_{it}^2} \sum_{j=1}^{N_{it}} \text{Var}(\epsilon_{jit}) = \frac{1}{N_{it}^2} \cdot N_{it} \cdot \sigma_{\epsilon_{jit}}^2 = \frac{\sigma_{\epsilon_{jit}}^2}{N_{it}} \quad \langle \text{식 3} \rangle$$

이 된다. 실제 해도닉 모형에서 $\hat{\mu}_{it}$ 는 단순 평균이 아니라 다수의 통제변수를 포함한 OLS 추정의 시점 더미 계수이므로, 〈식 3〉은 근사값이다. 거래건수가 통제변수 개수보다 충분히 크면 〈식 4〉가 성립한다.¹⁾

$$\text{Var}(\hat{\mu}_{it}) \approx \sigma_{\epsilon_{jit}}^2 / N_{it} \quad \langle \text{식 4} \rangle$$

1) 엄밀하게는 〈식 4〉의 근사가 성립하기 위해서는 조건부 등분산성 및 잔차 독립성이 필요하고, 시점 더미가 기타 독립변수의 영향을 제거한 이후에도 충분한 변동을 가져야 한다. 단지 고정효과가 포함된 모형에서는 특정 단지의 거래가 일부 분기에 집중되어 있을 경우 자유도 손실이 발생하여 분산을 과소추정할 수 있다. 본 연구가 유효 거래건수를 30건 이상으로 적용(아래 〈식 7〉 설명 부분 참고)한 것은 이러한 자유도 손실의 영향을 완화하기 위한 조치이다.

이 추정 오차의 분산은 관측되는 가격지수가 진정한 가격 수준으로부터 얼마나 벗어나 있는지를 나타내므로, 추정 오차의 분산은 가격 정보의 불확실성 크기를 의미한다. <식 3>에 의하면 이 불확실성의 크기는 ① 개별 거래에 내재된 가격 잡음의 크기 $\sigma_{\epsilon_{it}}^2$ 가 클수록, ② 그 잡음을 평활화할 수 있는 거래건수 N_{it} 가 적을수록 커진다. 즉, 가격지수가 전달하는 가격 신호의 질은 가격 잡음의 크기와 거래건수에 의해 동시에 결정된다. 거래가 충분하면 개별 거래의 잡음이 상쇄되어 가격지수가 정밀해지고, 거래가 적으면 소수의 거래에 내재된 잡음이 지수에 그대로 투영된다. 따라서 주택시장 변동성이 주택 자산 가치의 실제 변동을 반영하는 것인지, 아니면 거래 부족에서 오는 잡음을 반영하는 것인지를 판별하기 위해서는 이 두 요인을 함께 고려하는 지표가 필요하다.

3. 거래량의 내생적 결정

<식 4>의 거래건수 N_{it} 는 외생적으로 주어지는 것이 아니다. 탐색 마찰(search friction)의 존재하에서 거래량은 시장 참여자의 진입 및 퇴장 결정에 의해 내생적으로 결정된다(Diamond, 1982; Wheaton, 1990). 탐색 마찰하에서 잠재 매수자 B_{it} 와 잠재 매도자 S_{it} 가 만나 합의에 도달할 때 거래가 발생하며, 이를 매칭 함수로 표현하면 <식 5>와 같다.

$$N_{it} = M(B_{it}, S_{it}) = \kappa B_{it}^\alpha S_{it}^{1-\alpha} \quad \langle \text{식 5} \rangle$$

여기에서 $\kappa > 0$ 은 시장의 효율성을 나타내는 매칭 기술(matching technology) 파라미터이고, $\alpha \in (0, 1)$ 는 매수자 탄력성이다(문외솔, 2023; Genesove & Han, 2012; Ngai & Tenreiro, 2014; Pissarides, 2000, pp. 6, 7). 여기에서 중요한 점은 잠재 매수자 수 B_{it} 가 시장의 불확실성에 의존한다는 것이다. 시장 참여를 고려하는 잠재 매수자는 현재 가격 정보의 신뢰도를 평가한다. 가격 정보가 거칠다면(즉, 관측 변동성이 높으면) 매수 후 자산 가치의 불확실성이 커지므로 시장 진입을 유보한다. 이를 수식으로 표현하면 <식 6>과 같다.

$$B_{it} = \bar{B}_i \cdot g(\text{Var}(\hat{\mu}_{it})), \quad g' < 0 \quad \langle \text{식 6} \rangle$$

여기에서 \bar{B}_i 는 잠재 수요의 기저 수준으로서 가격 불확실성과 무관하게 존재하는 기본적인 매수 잠재 수요량이다. $g(\cdot)$ 은 감소 함수로서 가격 불확실성이 클수록 실제 시장에 진입하는 매수자가 줄어든다는 것을 의미한다.

4. 거래량과 가격 변동성 간 피드백 구조

〈식 4〉~〈식 6〉을 결합하면 자기강화적 피드백(self-reinforcing feedback) 구조가 형성된다. 이를 단계별로 추적하면 다음과 같다. 1단계는 거래 감소이다. 어떤 외생적 충격으로 잠재 매수자 수가 줄어들면, 〈식 5〉에 의해 거래건수 N_{it} 가 감소한다. 그러면, 2단계 불확실성 증가로 진입한다. N_{it} 가 감소하면, 〈식 4〉에 의해 가격지수의 추정 분산 $\text{Var}(\hat{\mu}_{it})$ 가 증가한다. 즉, 시장 참여자가 관측하는 가격 신호가 거칠어진다. 3단계는 추가 거래 위축이다. 가격 불확실성이 증가하면, 〈식 6〉에 의해 잠재 매수자 수 B_{it+1} 이 추가적으로 감소한다. 이는 다시 1단계로 작용한다.

5. 가격 잡음의 시변성

가격 잡음 $\sigma_{\epsilon it}^2$ 은 시간에 따라 변한다. 시장 과열기에는 패닉 매수, 급매, 투기적 거래 등으로 인해 개별 거래의 가격 분산 자체가 확대될 수 있다. 반대로 안정기에는 거래 쌍방이 충분한 정보를 가지고 협상하므로 가격 잡음이 줄어든다. 이것은 관측 변동성의 분해를 더 풍부하게 만든다. 거래가 줄어드는 시기에는 N_{it} 의 감소와 $\sigma_{\epsilon it}^2$ 의 변화가 동시에 발생할 수 있으며, 두 효과가 같은 방향으로 작용하면 관측 변동성의 증폭이 커진다. 〈식 4〉에서 분자와 분모가 동시에 움직이는 것이다. 이러한 시변성을 반영하기 위해, 본 연구는 가격 잡음을 시/구×시점별로 추정하였다. 이를 위해 헤도닉 잔차의 이동 구간(rolling window) 표준편차를 사용하여 $\hat{\sigma}_{\epsilon it}$ 를 〈식 7〉과 같이 산출하였다.

$$\hat{\sigma}_{\epsilon it} = \left[\frac{1}{n_w - 1} \sum_{s \in W(t)} \hat{\epsilon}_s^2 \right]^{1/2} \quad \langle \text{식 7} \rangle$$

여기서 $W(t)$ 는 시점 t 를 포함하는 직전 8분기(2년) 이동 구간이고, n_w 는 해당 윈도우 내 거래건수이다. 이를 통해 가격 잡음의 크기가 시간에 따라 변하는 것을 반영할 수 있다. 본 연구에서는 추정의 안정성을 확보하기 위해 $n_w \geq 30$ 인 경우에만 추정하였다.

이동 구간을 8분기(2년)²⁾로 설정한 근거는 다음과 같다. 본 연구의 분석 표본은 시/구×분기별로 유효 거래건수 30을 충족하므로, 8분기 윈도우 내 잔차 수는 최소 240개 이상이 확보되어 표본 표준편

2) Asadov et al.(2025)은 17개 OECD 국가의 분기 주택가격 자료에 대해 분기 자료의 특성을 고려하여 8분기 이동 표준편차를 변동성 측정치로 채택하였다.

차 추정의 통계적 안정성이 확보된다. 4분기 이동 구간은 표본 안정성이 부족할 수 있으며, 12분기 이상의 이동 구간은 시변성의 포착을 과도하게 약화시킬 것으로 판단된다.

여기에서 본 연구가 다루는 변동성 개념을 다음과 같이 명확히 하고자 한다. 첫째, 횡단면적 가격 잡음 분산은 동일 시점·동일 지역에 거래된 유사 주택들 사이에서 헤도닉 모형 잔차의 표준편차로 측정되며, 이는 가격 협상력 차이나 거래 시급성 등 거래 고유의 미세 특징에서 발생한다. 둘째, 시계열 변동성은 품질 조정된 지역 가격지수가 분기 간에 변동하는 정도를 측정하며, 거시·지역 충격에 따른 가격 수준 자체의 변화를 반영한다. 본 연구의 관측 변동성 지표(RV)는 후자에 해당하며, 헤도닉 잔차의 이동 구간 표준편차는 전자, 즉 횡단면적 가격 잡음 분산을 직전 8분기 윈도우를 활용하여 추정한 값이다. 횡단면 분산을 시계열 풀링(pooling)으로 추정하는 것은 분기별 표본만으로는 통계적 안정성을 확보하기 어렵다는 자료상의 제약을 보완하기 위함이다. 따라서 본 연구의 잔차 표준편차는 엄밀히 말하면 해당 시점을 포함하는 직전 8분기 윈도우 내 평균적 횡단면 분산의 추정치로 해석된다. 거래량 기반 가격 불확실성 지수(PUI)는 이 잔차 표준편차를 분자로, 해당 시점의 거래건수 제곱근을 분모로 결합하여 가격지수 추정의 불확실성을 측정한다.

6. 거래량 기반 가격 불확실성 지수(Price Uncertainty Index)의 도출

이상의 논의로부터 본 연구의 가설 지표인 거래량 기반 가격 불확실성 지수(PUI)를 도출할 수 있다. <식 4>, <식 7>에 의해, PUI는 <식 8>과 같이 정의된다.

$$PUI_{it} = \frac{\hat{\sigma}_{\epsilon_{it}}}{\sqrt{N_{it}}} \quad \langle \text{식 8} \rangle$$

PUI는 미래 가격의 불확실성이 아니라 관측되는 가격 신호의 질을 측정하는 지표이다. 가격 신호의 질이 낮다는 것은 수학적 관점에서 가격지수 추정량의 표준오차가 크다는 것을 의미한다. PUI의 분자 $\hat{\sigma}_{\epsilon_{it}}$ 는 해당 시기 개별 거래에 내재된 가격 잡음의 크기이며, 분모 $\sqrt{N_{it}}$ 는 그 잡음을 평활화할 수 있는 거래건수의 제곱근이다.

한편, 기존의 관측 변동성 지표중에서 대표적인 지표로는 가격지수 수익률의 이동 구간 표준편차 지표(RV)가 있다(<식 9>).

$$RV_{it} = \left[\frac{1}{L-1} \sum_{k=1}^L (r_{it-k} - \bar{r}_i)^2 \right]^{1/2} \quad \langle \text{식 9} \rangle$$

여기에서 $r_{it} = \hat{\delta}_{it} - \hat{\delta}_{it-1}$ 은 분기 수익률이고, L 은 이동 구간(rolling window)이다. 이 지표는 가격이 얼마나 크게 움직였는지를 측정하지만, 그 움직임이 주택 자산 가치의 실제 변동에서 온 것인지, 거래 부족에서 오는 추정 잡음에서 온 것인지를 구분하지 못한다.

이에 반해, PUI의 해석은 명확하다. 같은 수준의 관측 변동성(RV)이라 하더라도, PUI가 높으면 거래가 적어서 거칠게 보이는 것(가격지수 추정의 정밀도가 낮음)이고, PUI가 낮으면 주택 자산 가치가 실제로 변동하는 것이다. 즉, PUI는 관측 변동성 RV 중 거래 구조에 기인하는 부분을 분리해내는 역할을 한다. 본 연구에서 RV는 PUI의 역할을 부각하기 위한 비교 기준(benchmark)으로 활용되었다.

PUI는 RV를 완벽히 대체하는 지표는 아니다. PUI는 <식 4>의 가격지수 추정 분산의 제곱근, 즉 가격지수 추정량의 표준오차이며, RV는 그렇게 추정된 가격지수의 시계열 변동이다. 따라서 두 지표는 동일한 양을 다르게 측정한 것이 아니라, 관측 변동성(RV) 안에 포함되어 있는 추정 잡음 성분의 크기 (PUI)를 분리해 측정하는 분해 구조³⁾를 가진다. 본 연구가 RV와 PUI를 병렬적으로 보고하는 것은 두 변동성 지표를 비교하기 위함이라, 관측 변동성을 그 구성 성분에 따라 진단하기 위함이다.

IV. 연구 방법

본 연구의 실증 분석은 세 단계로 구성된다. 첫째, 거래량-변동성 관계(<식 11>)에서는 RV를 종속 변수로 하여 거래량의 설명력이 가격 잡음의 시변성을 경유하는 구조를 확인한다. 둘째, 설명력 검증(<식 12>~<식 14>)에서는 PUI를 RV에 추가한 모형의 추가 설명력을 평가한다. 셋째, 호황기-경색기 비대칭 분석에서는 RV와 PUI를 각각 종속변수로 두어, 거래량 변화에 대한 두 지표의 반응이 시장 국면에 따라 어떻게 다른지를 비교한다.

1. 데이터

본 연구의 주된 데이터는 국토교통부 아파트 실거래가 자료이다. 분석 기간은 2006년부터 2025년

3) 추정된 가격지수를 $\hat{p}_{it} = p_{it} + \eta_{it}$ 라고 하면, η_{it} 는 평균 0, 분산 $\sigma_{\epsilon}^2/N_{it}$ 의 추정 잡음이며 $PUI_{it} = \sqrt{Var(\eta_{it})}$ 에 해당한다. 분기 차분 $\Delta \hat{p}_{it} = \Delta p_{it} + \Delta \eta_{it}$ 의 분산은 추정 잡음이 시점 간 독립이라는 가정하에 $Var(\Delta \hat{p}_{it}) \approx Var(\Delta p_{it}) + 2\sigma_{\epsilon}^2/N$ 으로 분해되며, 따라서 $RV = \sqrt{Var(\Delta p_{it}) + 2 \cdot PUI^2}$ 의 관계가 성립한다. 여기에서 2배 계수는 연속된 두 분기의 추정 잡음 분산이 차분 과정에서 합산되기 때문에 발생한다. 즉, RV는 진정한 가치 변동의 분산 $Var(\Delta p_{it})$ 와 추정 잡음 성분 $2 \cdot PUI^2$ 의 합의 제곱근이며, PUI는 그 중에서 추정 잡음 성분이 분리되어 측정되는 지표이다.

까지이며, 총 관측 수는 약 741만 건이다. 각 거래 건에 대해 매매 거래금액(만 원), 전용면적(m^2), 층수, 건축 연도, 아파트 단지명, 지번 주소(시도+시/구+동+번지), 거래 시점(년, 분기, 월) 정보가 포함되어 있다. 분석 범위는 유의미한 아파트 거래건수가 획득되는 지역으로서 수도권(서울특별시, 인천광역시, 경기도)과 지방 5대 광역시(부산, 대구, 울산, 광주, 대전)로 한정하였다. 경기도 내 일부 지역(강화군, 옹진군, 연천군, 포천시, 가평군, 양평군, 여주시, 동두천시)과 광역시 소속 군 지역(기장군, 달성군, 군위군, 울주군)은 아파트 거래량이 분석에 불충분하여 제외하였다. 최종적으로 93개 시/구가 분석 대상으로 선정되었다.

헤도닉 가격모형에서 단지 고정효과를 사용하기 위해서는 각 거래가 어떤 단지에 속하는지를 정확하게 식별해야 한다. 이를 위해 국토교통부 건축물대장 총괄표제부(2026년 1월 기준) 원시자료를 결합하였다. 결합 키는 지번 주소를 활용하였다. 총괄표제부의 대지위치 필드에서 “번지” 문자열을 제거하면 실거래가 자료의 지번 주소와 동일한 형태가 되어 직접 결합이 가능하다. 결합 과정에서 동일 지번에 복수 건축물이 등재된 경우(예: 같은 대지에 복수 단지가 공존하는 경우)는 거래를 특정 건물에 귀속시킬 수 없으므로 분석에서 제외하였고, 비주거 건축물(세대수=0)도 제외하였다. 최종 결합 매칭률은 80.7%로 나타났으며, 미매칭 건의 주된 원인은 복수 단지 공존 지번의 거래이다. 매칭률은 연도별로 약 80% 내외로 안정적이었다.⁴⁾ 이 과정을 거쳐 약 598만 건의 아파트 거래를 헤도닉 모형 추정기 초 표본으로 확보하였으며, 이 매칭된 표본의 분기별 거래건수가 유효 거래건수(N_{it})가 된다.

시/구별 분기 거래량 변수로는 한국부동산원 부동산통계정보시스템(R-ONE)의 (월) 행정구역별 아파트매매거래현황 자료를 사용하여, 월별 거래건수를 분기별로 합산하여 시/구×분기 단위의 공식 거래량 변수($totalN_{it}$)를 구축하였다. 실거래가 자료에서 직접 집계한 건수와 달리, 공식 통계는 해제 건 처리 및 전수 포함 여부에서 일관성이 보장되므로 분석 변수로서 더 적합하다.

분석 단위는 시/구(i)×분기(t) 패널이다. 헤도닉 모형 추정에는 2006년 2025년 전체 기간의 거래 데이터를 사용하였다. 다만, 잔차 표준편차의 이동 구간 추정에 직전 8분기(2년) 윈도우를 사용하므로 초기 연도의 추정값은 불안정하다. 이를 고려하여 이후 패널 모형에 의한 실증 분석 기간은 초기 연도들을 제외하고 2009년 1분기부터로 설정하였다. 2006~2008년은 이동 구간이 안정화되는 시드(seed) 기간으로 활용된 것이다. 분석 단위를 분기로 설정한 이유는, 월별 구성 시 거래건수가 통계적 추정에 불충분할 정도로 적어지는 문제를 피하기 위함이다.

4) 매칭되지 않은 자료는 거래 가격이나 면적 등 관측 가능한 특성에 따른 체계적 선택이 아니라 동일 지번에 복수 건축물이 등재된 경우의 식별 모호성이라는 원인에 기인하므로, 매칭 표본의 특성이 전체 분포로부터 체계적으로 벗어났을 가능성은 제한적이다.

2. 헤도닉 가격모형

헤도닉 가격모형은 ① 주택 특성을 통제된 뒤의 분기 고정효과로부터 시/구별 가격지수와 수익률을 추출하여 RV를 산출하기 위해, ② 추정 잔차로부터 가격 잡음의 크기를 측정하여 PUI의 분자를 구성하기 위해 추정한다. 각 시/구별 헤도닉 가격모형은 <식 10>과 같다.

$$\ln P_{jit} = \delta_t + \beta_1 \ln A_j + \beta_2 F_j + \beta_3 F_j^2 + \beta_4 Age_{jt} + \beta_5 Age_{jt}^2 + \alpha_{d(j)} + \epsilon_{jit} \quad \langle \text{식 10} \rangle$$

여기에서 P_{jit} 는 아파트 j 의 시/구 i , 시점 t 에서의 거래가격(만 원), δ_t 는 분기 고정효과로서 시/구별 로그 가격지수, A_j 는 전용면적(m^2), F_j 는 층수, Age_{jt} 는 경과 연수, $\alpha_{d(j)}$ 는 주택 j 가 속한 단지 d 의 고정효과(fixed effects), ϵ_{jit} 는 잔차이다. 단지 고정효과는 동일 지번 주소와 동일 단지명을 공유하는 거래를 하나의 그룹으로 정의하여, 입지, 브랜드, 조망, 주차 환경 등 관측 불가능한 단지 특성을 통제한다.⁵⁾ 헤도닉 가격모형에서 단지의 평균 용적률 및 단지 규모를 통제하는 것이 중요하나, 본 모형에서는 아파트 단지 고정효과를 통제하므로, 단지의 물리적 특성을 설명하는 변수를 별도로 포함할 필요가 없다. 분기 고정효과는 해당 시/구의 분기별 가격 수준을 나타내며, 분기간 가격 변화율을 통해 가격지수를 추출하는 데 활용된다. 경과 연수의 2차항은 감가상각의 비선형성을, 층수의 2차항은 중간층 프리미엄의 역U자형 패턴을 반영하기 위해 추가되었다. 다수의 고정효과를 효율적으로 추정하기 위해 Stata reghdfe 패키지를 활용하였다.

3. 시/구×분기 패널자료 구축

헤도닉 가격모형으로부터의 추정값들 및 R-ONE 공식 거래량 자료를 결합하여 시/구×분기 단위의 패널데이터를 구축하였다. 패널데이터의 주요 변수는 다음과 같다(<표 1>). PUI 계산의 분모는

5) 단지 고정효과는 개별 단지가 갖는 시간불변적 입지, 브랜드, 주거환경, 학군, 교통 접근성 등 관측되지 않은 고정 특성을 통제하기 위한 것이다. 다만 단지 고정효과가 시간에 따라 변화하는 모든 국지적 요인을 완전히 통제하는 것은 아니다. 예컨대 교통망 확충, 학교 신설, 정비사업 기대 변화, 상권 변화 등 특정 단지 또는 생활권에 시기별로 다르게 작용하는 요인은 잔차에 일부 포함될 수 있다. 따라서 본 연구의 잔차 표준편차는 모든 품질 및 입지 요인을 완전히 제거한 후 남은 순수한 무작위 오차라기보다, 단지 고정효과, 주택 특성 및 분기별 평균 가격수준을 통제할 이후에도 개별 거래가격이 해당 지역·분기의 표준적 가격수준으로부터 흩어져 있는 정도로 해석하는 것이 적절하다. 이에 따라 PUI 역시 관측 가격신호의 흩어짐과 거래 희소성이 결합되어 나타나는 가격 정보의 불확실성을 측정하는 지표이며, 관측되지 않은 시간가변적 국지 요인이 일부 반영될 수는 있다.

〈표 1〉 패널 분석 변수의 정의 및 구성

변수	정의
$PUI_{it}(= \hat{\sigma}_{\epsilon it}/\sqrt{N_{it}})$	가격 불확실성 지수
RV_{it}	관측 변동성(분기 수익률의 이동 구간 표준편차)
$\hat{\sigma}_{\epsilon it}$	헤도닉 잔차의 이동 구간 표준편차
N_{it}	유효 거래건수(헤도닉 추정 표본)
$totaln_{it}$	공식 거래량(R-ONE, 분기 합산)
$r_{it}(= \hat{\delta}_{it} - \hat{\delta}_{it-1})$	분기 수익률

헤도닉 함수 추정 표본의 거래건수 N_{it} 의 제공근을 사용하고, 패널 회귀의 설명변수로서의 거래량은 공식 통계치 $totaln_{it}$ 를 사용한다. 이는 PUI의 분모가 잔차 평활화의 정도를 반영해야 하므로 헤도닉 함수로부터 추정된 거래건수의 제공근이어야 하는 반면, 패널 모형에 투입되는 거래량은 시장 상황을 반영하는 공식 통계가 더 적합하기 때문이다.⁶⁾

패널 데이터에 대하여 분석의 신뢰성을 확보하기 위해 두 단계의 표본 정제 작업을 진행하였다. 첫째, 유효 거래건수 N_{it} 가 30건 미만인 시/구×분기 관측치를 제외하였다. 제외된 관측치는 전체의 약 2%(125건)으로 나타났다. 둘째, 패널 분석에서 극단값의 과도한 영향을 방지하고자, PUI(×1000), RV, 수익률(r)에 대하여 상하 1% 윈저라이징(winsorizing)을 적용하였다. 표본 정제 후 최종 패널 데이터는 93개 시/구, 68분기의 6,324개 관측치에서 유효 거래 건수 30건 미만인 125개 관측치를 제외하여, 최종 6,199개 관측치로 구성된 비균형 패널(unbalanced panel) 형태로 구성되었다.

4. 패널 모형의 구축

1) 거래량과 변동성의 관계

시/구×분기 패널에서 거래량이 관측 변동성 RV를 구조적으로 설명하는지를 〈식 11〉의 설정을 통해 검증하였다.

$$RV_{it} = \alpha_i + \gamma_i + \beta \cdot f(N_{it}) + X'_{it}\Gamma + u_{it} \quad \langle \text{식 11} \rangle$$

6) N_{it} 와 $totaln_{it}$ 간 상관계수는 0.9197로 나타나, 두 변수의 출처 차이가 본 연구의 분석 결과에 미치는 실질적 영향은 제한적이라 판단된다.

여기에서 α_i 는 시/구 고정효과, γ_t 는 분기 고정효과, $f(N_{it})$ 은 로그(\ln)[모형 1] 또는 역수($1/N$) 변환모형 2이다. 시/구 고정효과는 지역의 시불변적 구조 차이(도시 규모, 시장 특성 등)를, 분기 고정효과는 전국 공통의 거시경제 추세(금리, 경기 순환 등)를 통제한다. 따라서 β 는 같은 시/구 내에서 거래량이 시간에 따라 변할 때 변동성이 어떻게 반응하는지를 식별한다. 또한 <식 11>에 잔차 표준편차 $\hat{\sigma}_{\epsilon it}$ 를 설명변수로 추가한 모형([모형 3])을 별도로 추정하여, 거래량 외에 가격 잡음도 RV를 설명하는지를 확인하였다. 모든 패널 회귀에서 시/구 군집 표준오차(clustered standard errors)를 적용하였다.

2) Price Uncertainty Index의 설명력 검증

기존 RV를 통제한 뒤에도 PUI가 미래 변수에 대해 추가적 설명 정보를 담고 있는지 검증하였다.⁷⁾ 여기에서는 두 가지 목표변수를 사용하였다. 첫째 목표변수는 미래 가격변동이다. h 분기 후 절대수익률 $|r_{i,t+h}|$ 을 RV만으로 설명한 모형(<식 12>)과 RV 및 PUI로 설명한 모형(<식 13>)을 비교하였다.⁸⁾

$$|r_{i,t+h}| = \alpha_i + \gamma_t + \beta_1 RV_{it} + u_{i,t+h} \quad \langle \text{식 12} \rangle$$

$$|r_{i,t+h}| = \alpha_i + \gamma_t + \beta_1 RV_{it} + \beta_2 PUI_{it} + u_{i,t+h} \quad \langle \text{식 13} \rangle$$

$h = 1, 2$ (단기, 중기), 4 (장기)에 대해 추정하여, PUI의 계수 β_2 가 유의하고 within R^2 가 증가하면 PUI가 RV에 없는 정보를 담고 있다고 결론내릴 수 있다.

둘째 목표변수는 미래 거래량 변화이다. PUI가 높다는 것은 가격 신호가 거칠다는 뜻이며, 이것이 시장 참여자의 행태에 영향을 미친다면 향후 거래량 변화를 선행할 수 있다. $h = 1, 2$ (단기, 중기)에 대하여 <식 14>를 추정하였다.

$$\Delta \ln N_{i,t+h} = \alpha_i + \gamma_t + \beta_1 RV_{it} + \beta_2 PUI_{it} + u_{i,t+h} \quad \langle \text{식 14} \rangle$$

7) 본 절의 설명력 검증은 PUI가 미래 가격이나 거래량을 인과적으로 결정한다는 가설을 입증하기 위한 것이 아니라, PUI가 RV에 없는 정보를 담고 있는지 확인하는 검증 절차이다. 따라서 단독 모형이 아닌 「RV 단독 vs RV+PUI」 비교 구조를 사용하며, 판단 기준은 PUI 계수의 유의성과 결정계수의 증가이다. 또한 PUI가 미래 거래량 변화를 설명하는 결과는 거래량의 평균 회귀 과정을 PUI가 포착함을 의미하며, 이는 거래 회소성이 극단적인 시기를 선행적으로 식별하는 측정상의 유용성으로 해석된다.

8) 절대수익률이 아니라 수익률을 그대로 쓰면 양수(상승)와 음수(하락)가 상쇄되어, 가격이 얼마나 크게 움직였는가를 측정할 수 없다. 설명하고 싶은 것은 미래에 가격이 오를지, 내릴지가 아니라 미래에 가격이 얼마나 크게 변동할지(즉, 미래 가격 변동의 크기)이다.

3) 호황기·경색기 비대칭 분석

호황기·경색기 비대칭 분석의 목적은 거래량이 변동성에 미치는 영향이 시장 국면에 따라 다른지를 검증하는 것이다. 앞의 이론적 논의에서 도출된 피드백 구조에 의하면, 거래가 이미 부족한 상태에서 추가적인 거래 감소는 가격 신호의 질을 급격히 악화시키지만, 거래가 충분한 상태에서는 추가 거래가 가격 신호의 질에 미치는 추가적 기여는 크지 않을 수 있다. 만약 이러한 비대칭성이 실증적으로 확인된다면, 이는 두 가지 중요한 함의를 갖는다. 첫째, 동일한 강도의 거래 억제 정책이라 하더라도 시장이 호황기인지 경색기인지에 따라 변동성에 미치는 영향이 상이하므로, 거래 규제의 강도와 시행 시점을 시장 국면에 연동하는 정책 설계⁹⁾가 필요하다. 둘째, 이 비대칭적 반응이 RV보다 PUI에서 더 뚜렷하게 나타난다면, PUI가 거래 구조의 변화에 더 민감한 지표임을 추가적으로 입증하는 증거가 된다.

호황기와 경색기의 구분을 위해 다음과 세 가지 기준을 적용하고 결과의 강건성을 비교하였다. [방법 1]은 해당 시/구의 해당 분기 공식 거래량이 직전 분기보다 증가하면 호황기, 감소하면 경색기로 분류하는 것이다. 가장 직관적인 기준이지만, 계절성을 반영하지 못한다는 한계가 있다. 통상 1분기는 거래가 적고 2~3분기는 많으므로, 단순 전분기 증감이 계절 패턴을 호황 또는 경색으로 오분류할 수 있다.

[방법 2]는 각 분기의 전국 중위 거래량을 산출한 뒤, 해당 시/구의 거래량이 이 중위값을 초과하면 호황기, 이하이면 경색기로 분류하는 것이다. 동일 분기 내에서의 상대적 위치로 판단하므로 [방법 1]의 계절성 문제를 완화할 수 있다. 다만 시/구 간 시장 규모 차이를 반영하지 못하여, 대규모 시/구는 대부분 호황기로, 소규모 시/구는 대부분 경색기로 분류되는 경향이 있다.

[방법 3]은 해당 시/구의 전체 분석 기간 평균 거래량을 기준으로, 이를 초과하는 분기를 호황기, 이하인 분기를 경색기로 분류하는 것이다. 시/구 고유의 시장 규모를 기준으로 삼으므로 지역 간 규모 차이와 계절성을 동시에 통제할 수 있다.

이러한 방법들에 대하여 <식 11>의 종속변수를 RV와 PUI로 각각 대체하여 별도로 추정하였다. 검증하고자 하는 대상은 두 가지이다. 첫째, 경색기의 거래량 계수 절대값이 호황기보다 큰지, 둘째, 이 비대칭성이 RV보다 PUI에서 더 뚜렷한지 여부이다.

9) 거래를 억제하는 정책(양도세 증과, LTV·DTI 강화, 다주택자 주택 추가 취득 제한 등)의 시행 시점과 강도를 시장 국면에 따라 차등화하는 정책 설계를 의미한다. 구체적으로, 거래가 활발한 호황기에는 이러한 규제를 강화하더라도 변동성에 미치는 부작용이 크지 않지만, 거래가 이미 위축된 경색기에는 동일한 규제가 가격 신호의 질을 급격히 악화시킬 수 있다. 다만 이는 거래 억제 정책의 본래 목표와 정보 효율성 비용을 함께 고려하여 차등 적용의 가능성을 시사하는 것이며, 어떤 단일 방향의 정책 처방으로 편향되는 것이 옳다는 의미는 아니다.

V. 실증 분석 결과

1. 기술 통계

〈표 2〉는 원저라이징을 통해 정제된 표본의 주요 변수에 대한 기술 통계를 나타낸 것이다. 최종 표본은 93개 시/구, 68분기(2009년 1분기~2025년 4분기)의 6,199개 시/구×분기 관측치로 구성되었다. PUI(×1000)는 평균 5.422, 중위값 4.340으로 오른쪽 꼬리가 긴 분포를 보이고 있다(왜도 1.86).¹⁰⁾ 이는 소규모 시/구에서 거래가 적은 분기에 PUI가 크게 상승하는 특성을 반영하는 것으로 사료된다. 잔차 표준편차($\hat{\sigma}_\epsilon$)는 평균 0.111, 표준편차 0.028로 비교적 안정적인 분포를 보이는데, 이는 PUI의 변동이 분모(거래건수)에 주로 기인함을 나타낸다.

다음으로 PUI와 RV의 상관관계를 검토하였다(〈표 3〉). PUI와 RV의 상관계수는 0.3819로, 두 지표가 약 62% 수준에서 독립적 정보를 담고 있음을 보여준다. 이는 PUI가 RV에 없는 차별적 정보를 포함하고 있을 가능성을 시사하며, 새로운 지표를 구축하는 근거가 된다. PUI와 거래량의 상관은 -0.566로, RV와 거래량의 상관(-0.163)보다 약 3.5배 강하다. PUI가 거래 구조의 변화에 RV보다 훨씬 민감하게 반응하는 지표임을 확인할 수 있다.

〈표 2〉 패널 모형 변수의 기술 통계

변수	평균	표준편차	중위값	5백분위수	95백분위수
PUI(×1000)	5.422	3.402	4.340	2.139	12.538
RV	0.038	0.028	0.030	0.010	0.096
$\hat{\sigma}_\epsilon$	0.111	0.028	0.107	0.075	0.162
N_{it}	814	833	585	90	2,288
total n_{it}	1,061	941	801	164	2,844
r_{it}	0.012	0.043	0.011	-0.052	0.081

10) PUI는 헤도닉 잔차 표준편차를 거래건수 제곱근으로 나눈 값이므로 실제 값의 크기가 매우 작게 나타난다. 따라서 이후 기술통계표와 회귀분석표에서는 해석과 가독성을 높이기 위해 PUI에 1,000을 곱한 PUI(×1000)를 보고하였다. 이는 단순한 단위 변환에 불과하므로 변수의 순위, 상관관계, 통계적 유의성 및 실증 결과의 해석에는 영향을 미치지 않는다. 표에 제시된 PUI(×1000)의 값을 원지수 기준으로 해석하려면 해당 값을 1,000으로 나누면 된다.

〈표 3〉 PUI와 RV 간 상관관계

	PUI(×1000)	RV	totaln _{it}
PUI(×1000)	1.0000		
RV	0.3819	1.0000	
totaln _{it}	-0.5664	-0.1630	1.0000

PUI, price uncertainty index; RV, realized volatility.

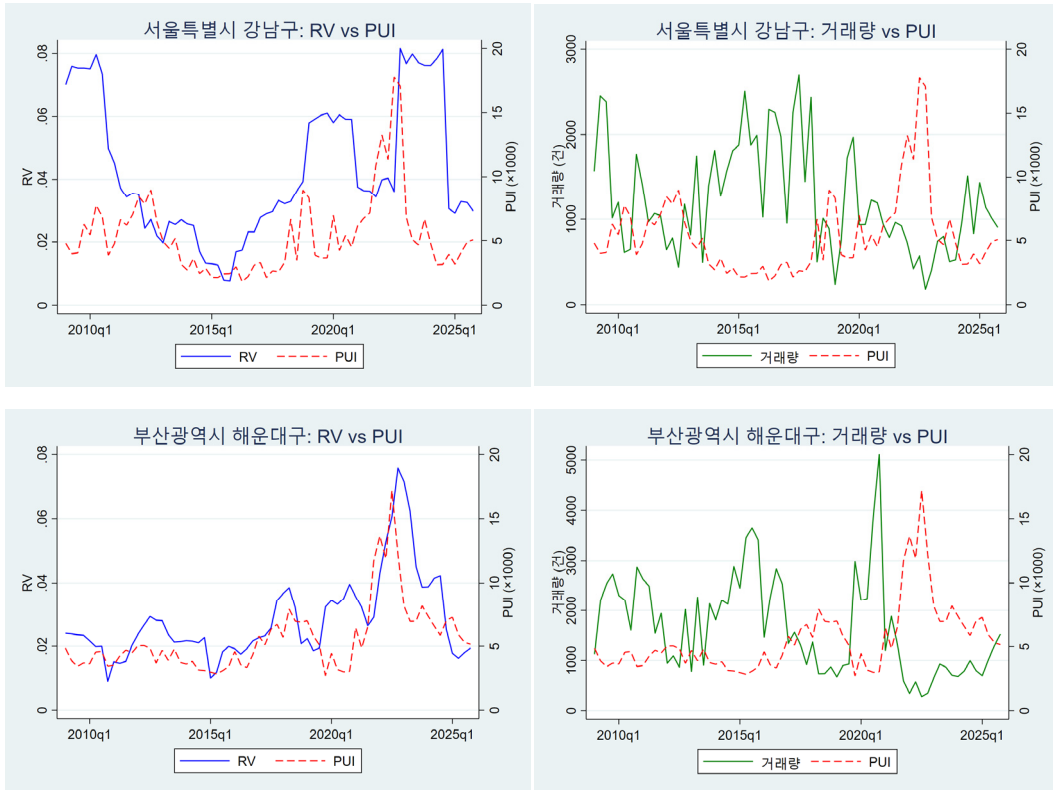
2. Realized Volatility와 Price Uncertainty Index의 시계열 비교

〈그림 1〉은 대표적인 시/구에서 RV와 PUI의 시계열을 비교한 것이다. 시계열 그래프에서 다음과 같은 세 가지 특징이 발견되었다. 첫째, RV와 PUI가 갈라지는 구간이 분명하게 존재한다. 서울 강남구의 경우 2010~2012년 기간에 RV는 높지만(0.07~0.08) PUI는 그보다 상당히 낮은 값(4~7)을 형성하고 있다. 또한 2024~2025년에도 RV는 0.08 수준의 높은 값을 유지하는 반면 PUI는 3~5 수준으로 낮아져 있어, 두 지표의 갈라짐이 다시 두드러진다. 이 시기는 거래가 활발하여 가격 신호의 질이 양호했으나(낮은 PUI), 가격 자체는 크게 변동하던 구간이다(높은 RV). RV만으로는 변동성이 큰 시장으로 판단되지만, PUI로 보면 정보가 충분한 구간이었다.

둘째, 2022년 전후 거래 절벽 시기¹¹⁾에 두 지표가 동시에 급등하며 동조화된다. 이 시기는 가격도 크게 변동하고(RV 상승), 동시에 거래 부족으로 가격 신호의 질도 나빠진(PUI 상승) 이중의 불확실성 구간이다. 이러한 이중 불확실성 구간에서는 기존의 RV만으로 시장 상황을 진단할 경우, 변동성 상승의 원인이 가치 변동인지 거래 부족인지를 구분할 수 없다.

셋째, 서울 강남구, 부산 해운대구의 [거래량 vs PUI] 그래프에서 거래량과 PUI가 반대로 움직인다. 거래량이 급감하는 분기에 PUI가 치솟고, 거래가 급증하면 PUI가 바닥으로 내려간다. 이 패턴은 PUI의 분모인 거래건수가 PUI의 단기 변동을 지배하고 있음을 시각적으로 보여주고 있다. 특히 2022년 거래 절벽기에 PUI가 두 지역에서 모두 17 수준까지 급등하는 매우 강한 동조 패턴이 관찰되

11) 2022년 거래 절벽의 주된 원인은 세 가지로 요약된다. 첫째, 금리 급등이다. 한국은행이 2021년 8월 0.50%였던 기준금리를 2022년 말 3.25%, 2023년 초 3.50%까지 인상하였다. 주택담보대출 금리가 연 2%대에서 6%~7%대로 급등하면서 매수자의 구매력이 급격히 위축되었다. 둘째, 누적된 규제의 영향이다. 다주택자 양도세 증가, LTV·DTI 강화, 종합부동산세 인상 등이 2017년부터 누적적으로 시행되어 매도·매수 양측의 거래 유인을 억제하고 있었다. 특히 양도세 증가로 다주택자가 매도를 유보하는 잠김 효과(lock-in effect)가 강하게 작용하였다. 셋째, 가격 급등 이후의 심리 반전이다. 2020~2021년 유동성 장세에서 서울 아파트 가격이 40%~50% 급등한 뒤, 금리 인상과 함께 가격 하락 기대가 형성되면서 매수자는 관망하고 매도자는 고점 가격을 고수하여 호가 갭이 확대되었다. 매수·매도 간 가격 괴리가 커지면서 거래 성사 자체가 어려워진 것이다.



RV, realized volatility; PUI, price uncertainty index.

〈그림 1〉 RV와 PUI의 시계열 비교

며, 2018년 전후에도 거래량 감소와 함께 PUI가 부분적으로 상승하는 양상이 동일한 메커니즘의 약한 발현으로 해석된다.

첫 번째 특징과 두 번째 특징을 결합하면, PUI의 유용성이 명확해진다. 2010~2012년 시기처럼 RV는 높지만 PUI가 낮은 구간에서, RV만 관찰하면 변동성이 큰 시장으로 판단되지만, PUI를 함께 보면 거래가 활발하여 가격 정보는 충분하나 가치 자체가 크게 움직이는 시장으로 진단이 달라진다. 반대로 검색기에 RV는 낮지만 PUI만 높은 구간이 나타난다면, 이는 가격이 안정적으로 보이지만 실제로는 거래 부족으로 가격 정보 자체를 신뢰할 수 없는 시장이다. RV만으로는 이러한 상황을 포착할 수 없다. 2022년과 같이 두 지표가 동조화되는 것은 가치 변동과 거래 부족이 동시에 작용하는 위기 국면의 특수한 경우이며, 평상시에는 두 지표가 갈라지는 것이 오히려 일반적이다. PUI의 유용성은 이처럼 두 지표가 갈라지는 구간에서 발휘되며, RV와 PUI를 병행하여 관찰할 때 관측 변동성의 요인을 온전히 식별할 수 있다.

3. 패널 회귀: 거래량과 변동성의 관계

〈표 4〉는 시/구×분기 고정효과 패널 모형에서 거래량이 관측 변동성(RV)을 설명하는지를 검증한 결과이다. [모형 1]에서 거래량의 로그값($\ln N$)의 계수는 -0.0043 으로 1% 수준에서 유의하다. 거래가 많아지면 변동성이 줄어든다는 전망¹²⁾과 일치한다. [모형 2]에서 거래량의 역수($1/N$) 역시 유의한 양의 계수를 보이고 있다.¹³⁾ 더 중요한 결과는 [모형 3]이 담고 있다. 잔차 표준편차($\hat{\sigma}_\epsilon$)를 추가하면 Within R^2 가 0.404에서 0.553으로 증가한다. 동시에 거래량($\ln N$)의 계수가 0.0007로 줄어들며 유의성이 완전히 사라진다. 이는 거래량이 RV에 기여하는 변동의 상당 부분이 가격 잡음($\hat{\sigma}_\epsilon$)의 시변성에 의해 흡수됨을 의미한다. PUI는 가격 잡음 $\hat{\sigma}_\epsilon$ 을 분자에 담고 있다. 결국, PUI는 RV의 변동을 유발하는 성분을 포함하고 있다는 것이다.

〈표 4〉 거래량과 관측 변동성(RV) 패널 회귀 분석 결과

변수/통계량	모형 1	모형 2	모형 3
$\ln N_{it}$	-0.0043^{**} (0.0015)		0.0007 (0.0012)
$1/N_{it}$		0.696* (0.349)	
$\hat{\sigma}_{\epsilon it}$			0.485*** (0.054)
시/구 고정효과	○	○	○
분기 고정효과	○	○	○
Within R^2	0.404	0.402	0.553
N	6,199	6,199	6,199

주: 종속변수=RV. 괄호 안은 시/구 군집 표준오차. * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$.
RV, realized volatility.

- 12) 본 연구의 RV는 동일 시점·동일 지역의 개별 거래 가격이 평균에서 떨어진 정도(횡단면 분산)가 아니라, 분기별로 추정된 가격지수의 시계열 변동을 측정하는 지표이다. 분기 가격지수는 해당 분기 거래로부터 추정된 평균 추정량이므로, 표본 평균의 표준오차가 표본 크기의 제곱근에 반비례한다는 성질에 따라 거래량이 클수록 추정이 정밀해지고, 그 결과 시계열 변동에 섞여 있던 추정 잡음 성분이 줄어드는 경향이 있다. 따라서 거래량과 RV의 음의 관계는 모집단 가격 분산 자체가 거래량에 따라 변한다는 의미가 아니라, 거래량 증가가 가격지수 추정 정밀도를 높여 관측되는 시계열 변동에서 추정 잡음 성분을 축소시킨다는 의미로 이해할 수 있다.
- 13) [모형 1]의 로그 변환은 거래량-변동성의 일반적 단조 관계를 점검하기 위한 것이며, [모형 2]의 역수 변환은 식 (4)에서 도출된 가격지수 추정 분산의 $1/N$ 비례 구조에 부합하는 함수 형태로, 거래량 부족이 가격 신호의 불확실성을 비선형적으로 확대하는지를 확인하는 역할을 한다.

4. Price Uncertainty Index의 설명력 검증

1) 미래 가격변동 설명력

〈표 5〉는 RV만으로 미래 절대수익률을 설명한 모형과 PUI를 추가한 모형을 비교한 결과이다. 세 가지 시계 모두에서 PUI는 통계적으로 유의한 음(-)의 계수를 보인다. 1분기 후에서 PUI의 계수는 -0.0013으로 유의수준 0.1%에서 매우 유의하며, 2분기 후 -0.0008, 4분기 후 -0.0009로 모두 유의수준 5%에서 유의하다. PUI가 높으면 이후 분기의 가격변동이 오히려 작아진다는 일관된 패턴이 관찰되는데, 이는 PUI가 높은 시기(거래가 부족한 시기)에는 가격을 크게 움직일 거래 자체가 부족하기 때문으로 해석된다. PUI의 음(-)의 계수가 단기뿐 아니라 중기·장기에서도 유지된다는 점은 거래 희소성이 가격 신호의 질을 악화시키는 효과가 단순한 단기 변동이 아니라 비교적 지속적인 영향임을 시사한다. 다만 PUI의 주된 기여는 가격변동 설명보다는 다음 〈표 6〉에 기술된 미래 거래량의 설명에서 더 뚜렷하게 나타난다.

RV의 시간 구조는 다음과 같이 해석된다. h=1에서 양의 계수(0.140)는 현재 가격이 크게 움직이는 시/구에서 다음 분기에도 큰 가격 움직임이 이어지는 지속성을 의미한다. 한편 h=4에서는 음의 계수(-0.031)가 관찰되나 $p=0.147$ 로 통계적으로 유의하지 않아 1년 시계의 패턴에 대해 단정적으로 해석하기는 어렵다. 다만 h=1에서의 지속성 효과가 중기 이상으로 강하게 이어지지 않는다는 점은 단기적 가격변동 급등에 과도하게 반응하는 정책 개입이 불필요할 수 있음을 시사한다.

〈표 5〉 미래 가격변동의 설명: RV 단독 vs RV+PUI

변수/통계량	h=1 RV만	h=1 RV+PUI	h=2 RV만	h=2 RV+PUI	h=4 RV만	h=4 RV+PUI
RV	0.140** (0.024)	0.180*** (0.027)	0.026 (0.025)	0.049 (0.030)	-0.031 (0.021)	-0.006 (0.025)
PUI		-0.0013*** (0.0003)		-0.0008* (0.0004)		-0.0009* (0.0004)
Within R^2	0.283	0.289	0.272	0.274	0.276	0.279
N	6,106	6,106	6,014	6,014	5,830	5,830

주: 종속변수는 $|r_{it+h}|$. 시/구·분기 고정효과 포함. 괄호 안은 시/구 군집 표준오차. * $p<0.05$, *** $p<0.001$.

RV, realized volatility; PUI, price uncertainty index.

〈표 6〉 미래 거래량의 설명: RV 단독 vs RV+PUI

변수/통계량	h=1 RV만	h=1 RV+PUI	h=2 RV만	h=2 RV+PUI
RV	-0.373** (0.130)	-0.858*** (0.172)	-0.216 (0.222)	-1.422*** (0.334)
PUI		0.0165*** (0.0016)		0.0411*** (0.0040)
Within R^2	0.616	0.620	0.569	0.589
N	6,106	6,106	6,014	6,014

주: 종속변수는 $\Delta \ln N_{it+h}$. 시/구·분기 고정효과 포함. 괄호 안은 시/구 군집 표준오차. ** $p < 0.01$, *** $p < 0.001$.

RV, realized volatility; PUI, price uncertainty index.

2) 미래 거래량 설명

〈표 6〉은 RV와 PUI가 미래 거래량 변화를 설명하는지를 검증한 결과이다. PUI의 계수가 1분기 후 0.0165, 2분기 후 0.0411로 모두 유의수준 0.1%에서 매우 강하게 유의하였다. PUI가 높은 시/구에서 이후 거래량이 회복되는 패턴이며, 이는 거래가 위축되어 PUI가 높아진 시/구에서 시장이 정상 수준으로 회귀하는 동학을 PUI가 포착하고 있음을 시사한다. RV만으로는 2분기 후 거래량 변화를 설명하지 못하지만(-0.216, $p=0.331$), PUI를 추가하면 RV도 유의수준 0.1%에서 강하게 유의해지며(계수 -1.422***), R^2 가 0.569에서 0.589로 증가한다. PUI가 임계 수준¹⁴⁾을 넘어서는 시/구는 현재 가격 정보의 질이 낮아 정상화가 필요한 시장으로 판단할 수 있다.

5. 호황기·경색기 비대칭 분석

본 절은 세 번째 연구문제에서 명시한 두 가지 질문, ① 거래량-변동성 관계의 호황기·경색기 비대칭성, ② 이러한 비대칭적 반응이 RV와 PUI에서 어떻게 다르게 나타나는가에 답하기 위한 분석이다. 이에 따라 종속변수로 RV와 PUI를 각각 사용하여 두 지표가 거래량 변화에 어떻게 다르게 반응하는지를 비교한다. 세 가지 국면 구분 방법에 따른 호황기·경색기 표본 구성은 〈표 7〉과 같다.¹⁵⁾ 모든 방법에

14) 본 연구는 이러한 PUI의 임계 수준에 대한 객관적 기준을 제시하지는 않는다. 다만 〈표 2〉에 나타난 바와 같이 PUI($\times 1000$)의 95백분위수가 12.538이므로, 이를 초과하는 시/구 \times 분기는 전체 관측치 중 가격 정보의 질이 가장 낮은 상위 5%에 해당한다. 예컨대 가격 정보의 질이 가장 낮은 5%의 시장을 모니터링 대상으로 삼고자 한다면, PUI($\times 1000$)의 임계 수준을 12.538로 설정할 수 있을 것이다.

〈표 7〉 국면 구분 방법별 호황기·경색기 표본 구성

국면 구분 방법	그룹	N	평균 거래량	평균 PUI (×1000)	평균 RV
방법 1: 전분기 증감	경색기	2,924	892.8	5.953	0.038
	호황기	3,182	1,224.3	4.888	0.038
방법 2: 전국 분기 중위	경색기	3,113	492.6	7.060	0.039
	호황기	3,086	1,633.7	3.770	0.037
방법 3: 시/구별 장기 평균	경색기	3,454	697.7	6.579	0.043
	호황기	2,745	1,517.4	3.966	0.033

주: 방법 1의 N=6,106인 것은 시/구별 첫 분기 관측치가 전분기 차분 계산 과정에서 손실되었기 때문이다. 방법 2와 3의 N=6,199로 전체 표본과 일치한다.

PUI, price uncertainty index; RV, realized volatility.

서 호황기는 거래량이 많고 PUI가 현저히 낮은 공통 패턴을 보여, PUI의 개념이 국면 구분과 일관되게 작동함을 확인할 수 있다. 다만 세 방법은 호황기와 경색기의 분류 양상이 상이하다.

[방법 1] (전분기 증감)은 호황기와 경색기 평균 거래량의 차이가 약 1.4배(1,224 vs 893)로 가장 작다. 이는 계절성 패턴이 호황/경색 분류에 일부 흡수되어 거래량 차이를 좁히는 경향이 있음을 시사한다. [방법 2] (전국 중위)는 호황기 평균 거래량이 경색기의 약 3.3배(1,634 vs 493)에 달하여, 대규모 시/구가 체계적으로 호황기로, 소규모 시/구가 경색기로 분류되는 경향이 정량적으로 확인된다. [방법 3] (시/구 평균)은 호황기와 경색기 평균 거래량 차이가 약 2.2배(1,517 vs 698)로 중간 수준이며, 시/구 간 규모 차이는 통제되고 시/구 내 시간 변동만이 분류에 반영된다는 점에서 세 방법 중 가장 균형 잡힌 분류로 평가할 수 있다.

한편 RV의 평균값은 세 방법 모두에서 호황기와 경색기 간 차이가 미미하거나 거의 없는 반면(방법 1, 2에서는 차이 없음, [방법 3]에서만 0.043 vs 0.033으로 약간 차이), PUI는 모든 방법에서 호황기와 경색기 간 약 18%~47% 수준의 차이를 보인다. 이는 PUI가 RV보다 거래 구조의 변화에 민감하게 반응하는 지표라는 본 연구의 주장이 표본 평균 수준에서 일부 드러나는 것으로 보인다.

15) 본 연구에서 호황기와 경색기는 주택가격의 상승기와 하락기를 의미하는 것이 아니라, 거래량을 기준으로 정의한 거래 유동성 확대기와 거래 유동성 경색기를 의미한다. 이는 본 연구의 관심이 가격 순환 자체가 아니라 거래 희소성이 가격 신호의 정밀도에 미치는 영향에 있기 때문이다. 따라서 본 연구의 비대칭 분석 결과는 가격 상승기와 하락기 사이의 비대칭성이 아니라, 거래 유동성 확대기와 경색기 사이의 비대칭성으로 해석하는 것이 적절하다. 다만 PUI는 거래건수를 분모로 포함하는 지표이므로, 거래량 기준의 국면 구분이 경색기에서 PUI의 거래량 민감도를 상대적으로 강화할 가능성이 있다. 이에 본 연구는 전분기 대비 거래량 증감, 동일 분기 전국 중위 거래량, 해당 시/구 평균 거래량이라는 세 가지 기준을 병행하여 특정 기준에 대한 의존성을 완화하고자 하였다.

〈표 8〉은 호황기와 경색기를 구분하여 거래량이 RV와 PUI에 미치는 영향의 비대칭성을 분석한 결과이다. 세 가지 국면 구분 기준 모두에서 PUI에 대해서는 일관된 비대칭성이 확인되었으며, RV에 대해서는 개별 계수의 유의성 패턴은 유사하나 동일성 검정에서 그 비대칭성의 통계적 견고성은 제한적임이 드러난다. 첫째, RV를 종속변수로 한 모형에서 호황기 거래량 계수는 세 가지 방법 모두에서 통계적으로 유의미하지 않은 것으로 나타났다. 반면, 경색기에서는 유의한 음의 계수가 나타난다. 이는 분리 추정에서 보면 거래가 충분한 시기에는 추가적인 거래가 관측 변동성에 영향을 미치지 않고 거래가 부족해질 때 거래량과 변동성의 관계가 나타나는 것으로 보이지만, 동일성 검정에서 그 비대칭이 통계적으로 강하게 입증되지는 않는다는 점에 유의해야 한다. 특히 [방법 3]에서 호황기 계수가 통계적으로 유의미하지 않아 사실상 0이라는 결과는, 거래가 시/구 고유의 장기 평균을 상회하는 상태에서는 거래량의 변화가 변동성에 미치는 추가적인 기여점이 매우 제한적임을 보여준다.

둘째, PUI를 종속변수로 설정한 모형에서는 RV 모형과는 양상이 뚜렷하게 다르다. 호황기에서도 거래량 계수가 세 가지 방법 모두에서 강하게 유의하며, 경색기에서는 그 절대값이 더욱 커진다는 것을 확인할 수 있다. 즉, PUI는 시장 국면과 무관하게 거래량 변화에 민감하게 반응하되, 경색기에서 그 민감도가 증폭되는 패턴을 보이고 있다.

두 지표의 차이는 호황기와 경색기 계수의 동일성 검정에서 더욱 명확히 드러난다(〈표 9〉). PUI를

〈표 8〉 호황기·경색기 비대칭 분석: ln(N)의 계수

	방법 1: 전분기 증감		방법 2: 전국 중위		방법 3: 시/구 평균	
	호황기	경색기	호황기	경색기	호황기	경색기
종속변수: RV						
lnN	-0.003 (0.0017)	-0.007*** (0.0020)	-0.004 (0.0023)	-0.005* (0.0024)	-0.001 (0.0015)	-0.006* (0.0026)
Within R ²	0.442	0.407	0.477	0.357	0.403	0.408
N	3,182	2,924	3,086	3,113	2,745	3,454
종속변수: PUI(×1000)						
lnN	-2.919*** (0.199)	-3.941*** (0.239)	-2.044*** (0.139)	-4.516*** (0.354)	-2.100*** (0.147)	-4.271*** (0.342)
Within R ²	0.624	0.725	0.714	0.717	0.307	0.67
N	3,182	2,924	3,086	3,113	2,745	3,454

주: 1) 시/구-분기 고정효과 포함. 괄호 안은 시/구 군집 표준오차. * p<0.05, *** p<0.001.

2) 방법 1: 전분기 대비 거래량 증감, 방법 2: 전국 분기별 중위 거래량 대비, 방법 3: 시/구별 장기 평균 대비.

RV, realized volatility; PUI, price uncertainty index.

〈표 9〉 호황기·경색기 거래량 계수 동일성 검정

	방법 1: 전분기 증감	방법 2: 전국 중위	방법 3: 시/구 평균
종속변수: RV			
β (호황기)	-0.0027(0.0017)	-0.0045(0.0023)	-0.0008(0.0015)
β (경색기)	-0.0073(0.0020)	-0.0054(0.0024)	-0.0064(0.0026)
z 통계량	1.77	0.29	1.85
p-value	0.077	0.770	0.064
종속변수: PUI(×1000)			
β (호황기)	-2.9190(0.1992)	-2.0438(0.1394)	-2.0995(0.1475)
β (경색기)	-3.9411(0.2394)	-4.5162(0.3540)	-4.2711(0.3424)
z 통계량	3.28**	6.50***	5.83***
p-value	0.001	<0.001	<0.001

주: 괄호 안은 표준오차. ** $p < 0.01$, *** $p < 0.001$.
 RV, realized volatility; PUI, price uncertainty index.

종속변수로 한 모형에서는 세 가지 국면 구분 방법 모두에서 호황기와 경색기 계수의 차이가 최소 1% 이상의 수준에서 강하게 유의한 것으로 나타나([방법 1]: $z=3.28$, $p=0.001$; [방법 2]: $z=6.50$, $p < 0.001$; [방법 3]: $z=5.83$, $p < 0.001$), PUI의 비대칭성이 통계적으로 견고함이 확인되었다. 반면 RV를 종속변수로 한 모형에서는 세 방법 모두에서 호황기·경색기 계수 차이가 5% 수준의 유의성에 도달하지 못하여, RV의 비대칭성은 분리 추정에서 한 계수만 유의한 것처럼 보이지만 두 계수의 차이 자체는 통계적으로 강하게 입증되지 않음이 확인되었다. 이는 RV가 거래 구조의 변화를 경색기라는 극단적 국면에서 부분적으로 반영하는 것처럼 보이나 그 비대칭성이 통계적으로 견고하지는 않은 반면, PUI는 어떤 국면에서든 거래 구조의 변화를 일관되게 포착하며 그 비대칭성이 통계적으로 매우 강하게 입증되는 지표임을 보여준다.

VI. 결론

1. 연구 요약

주택가격 변동성을 측정하는 기존의 지표는 가치의 실제 변동과 거래 희소성에서 오는 표본 잡음을 구분하지 못한다. 주택은 주식이나 채권과 달리 가격이 거래 시점에만 관측되므로, 거래가 부족한

시장에서는 동일한 가치 변동에서도 더 거친 가격 신호가 나타난다. 본 연구는 이 문제에 주목하여, 거래 구조가 관측 변동성에 내생적으로 기여하는 메커니즘을 이론적으로 전개하고, 이를 직접 측정하는 새로운 지표인 거래량 기반 가격 불확실성 지수를 구축하였다.

실증 분석 결과, PUI와 기존 관측 변동성(RV)의 상관관계는 0.38에 불과하여 두 지표가 상당히 다른 정보를 담고 있음을 확인하였다. 거래량-변동성 관계는 가격잡음의 시변성을 경유하여 작동하며, PUI는 RV가 포착하지 못하는 미래 거래량 변화를 강하게 설명하였으며, 미래 가격 변동에 대해서도 단기·중기·장기에 걸쳐 일관되게 추가 설명력을 가짐을 확인하였다. 호황기·경색기 비대칭 분석에서는, 거래가 충분한 시기에 추가 거래가 변동성에 영향을 주지 않지만 거래가 부족해지면 변동성 증폭효과가 강하게 나타남을 확인하였으며, 이 비대칭성은 RV에서는 동일성 검정 기준으로 통계적 견고성이 제한적이었던 반면 PUI에서는 모든 국면 구분 방법에 걸쳐 강하고 유의하게 관찰되었다.

2. 연구의 시사점

1) 주택시장에 제공하는 시사점

본 연구의 결과는 주택시장 참여자에게 실질적 시사점을 제공한다. PUI는 주택시장 분석 기관에게 관측 변동성의 해석을 정교화하는 도구로 활용될 수 있다. 같은 수준의 RV라 하더라도, PUI가 높으면 거래가 적어서 거칠게 보이는 것이고, PUI가 낮으면 실제로 가치가 흔들리는 것이다. 이 구분은 시장 평가의 불확실성 범위를 설정하거나, 시장 보고서에서 변동성의 원인을 진단하는 데 활용될 수 있다. 지금까지는 변동성이 크면 일률적으로 위험한 시장으로 분류되었지만, 실제로는 정보가 부족한 시장과 가치 변동이 실제로 큰 시장을 구분하여 의사결정의 질을 높일 수 있다. 투자자에게 PUI는 시장 진입 시점의 판단에 대해 유용한 정보를 제공할 수 있다. 본 연구에서 PUI가 높은 지역 주택시장에서 이후 거래량이 회복되는 패턴이 확인되었는데, 이는 PUI가 시장 경색의 정점을 선행적으로 포착하는 지표로 기능할 수 있음을 의미한다. PUI가 극단적으로 높아진 시점은 가격 정보의 신뢰도가 낮아 매수·매도 판단이 어려운 시기이지만, 동시에 경색이 곧 완화될 가능성이 있는 시기이기도 하다. 투자자는 RV만 보았을 때와 PUI를 함께 보았을 때 시장 상황에 대한 판단을 달리할 수 있다. 금융기관에게, PUI는 주택담보대출의 담보가치 평가와 리스크 관리에 대한 시사점을 제공한다. 거래가 뜸한 지역의 주택에 대한 담보가치 평가는 본질적으로 불확실성이 크며, PUI는 이 불확실성의 크기를 측정하는 데 유용한 판단 기준을 제시할 수 있다. 금융기관이 LTV 비율을 설정하거나 포트폴리오의 지역별 리스크를 평가할 때, 단순한 가격 변동성(RV)에 더해 가격정보의 질(PUI)을 함께 고려하면

보다 정교한 리스크 관리가 가능할 것이다.

2) 정책적 시사점

본 연구가 주택정책 수립 기관에 제시하는 시사점은, 거래 활성화가 가격 수준을 안정시키는 것이 아니라 가격 신호의 질을 높인다는 점이다. 거래가 충분하면 시장 참여자가 현재 가격 수준을 신뢰할 수 있고, 그 위에서 합리적 의사결정이 가능하다. 거래가 부족하면 가격 신호 자체를 온전히 신뢰하기 어려우므로 시장이 경색된다. 이와 관련하여 거래를 억제하는 정책(양도세 증과, 대출 규제 강화, 다주택자 규제 등)은 가격 안정, 투기 억제, 가계부채 관리 등 정책 목표를 달성하는 효과와 별개로, 시장의 정보 효율성에 부수적 비용을 발생시킬 수 있다는 가능성에 대해서도 함께 고려할 필요가 있다. 비대칭 분석 결과가 보여주듯, 이 비용은 거래가 이미 적은 경색 국면에서 특히 크다. 거래가 충분한 호황기에는 추가적 거래 억제가 변동성에 큰 영향을 미치지 않지만, 거래가 부족한 경색기에는 소폭의 추가 감소만으로도 가격 신호의 질이 악화될 수 있다. 이는 동일한 규제 강도라도 시장 국면에 따라 비용이 비대칭적임을 의미하므로 이러한 국면을 고려한 정책 설계가 필요하다. PUI를 시장 모니터링 지표로 활용하면, 단순히 가격 상승·하락의 문제를 넘어 현재 관측되는 가격 정보가 얼마나 신뢰할 만한지를 함께 판단할 수 있다. 정책 수립 기관이 지역 주택시장별 PUI를 정기적으로 산출하여 모니터링한다면, 특정 지역의 시장 정보가 열악해지는 시점을 조기에 포착하여, 필요 시 거래 활성화를 위한 선제적 조치를 적시에 시행할 수 있을 것이다.

3. 연구의 한계 및 향후 연구 과제

본 연구에는 한계점은 다음과 같다. 첫째, 거래 구조에서 변동성으로의 인과 방향을 엄밀하게 식별하지 못하였다. 본 연구는 패널 고정효과 모형과 비대칭 분석을 통해 강한 연관성을 보였으나, 외생적 충격을 활용한 인과 식별에는 사전 추세의 한계로 인해 분석에 한계가 있었다. 둘째, 본 연구가 서술한 거래량-변동성 피드백 구조에 대해 매칭 함수와 양측 행태에 명시적 함수형을 부여하여 균형의 존재, 유일성, 안정성을 엄밀히 분석하는 이론 모형을 개발하는 것은 후속 과제로 남긴다. 셋째, 호황기·경색기 구분과 관련하여, 향후 연구에서는 가격 기반 국면전환 모형, 미분양, 금리, 대출여건 등 외부 지표를 활용한 대안적 국면 구분을 병행하여, 거래 유동성 국면과 가격 순환 국면이 PUI 및 관측 변동성에 미치는 영향을 보다 정교하게 분리할 필요가 있다. 넷째, 본 연구의 변동성 측정은 8분기 이동 구간을 활용하였는데, GARCH 계열의 시간가변 계량모형을 적용한 결과와 비교가능할 것이다.

다만, GARCH류의 모형이 추정하는 대상은 수익률의 동적 시계열 분산이라는 점에서 본 연구가 주목하는 가격 잡음과는 다른 차원의 변동성을 다룬다는 한계점은 있다. 다섯째, 본 연구의 헤도닉 모형은 분기 고정효과를 통해 시/구×분기 단위의 평균적 유동성 효과를 흡수하나, 개별 거래 수준에서 거래량 또는 유동성이 가격 형성에 미치는 영향을 직접 통제하지는 못한다는 한계가 있다. 시/구×분기 거래량 변수는 동일 시/구×분기 내 모든 거래에 대해 같은 값을 가져 분기 고정효과와 완전한 공선성을 가지므로 시/구별 헤도닉 모형에 직접 포함할 수 없었다. 이로 인해 본 연구의 진차 표준편차는 순수한 가격 잡음만이 아니라 거래의 유동성 효과를 일부 포함할 가능성이 있다.

향후 연구에서는 다음과 같은 확장이 가능하다. 첫째, 주택 매매거래를 급감시키는 사건을 외생적 충격으로 활용하여, 「거래구조 → 변동성」의 인과 방향을 보다 엄밀하게 식별할 수 있을 것이다. 둘째, 분석대상을 아파트에서 비아파트로 확장하면, 거래 빈도가 더 낮은 시장에서 PUI의 유용성을 검증할 수 있을 것이다. 셋째, PUI의 임계값을 설정하여 시장 정보 경보 시스템을 설계하는 정책 연구가 가능하다. 이러한 후속 연구를 통해, PUI가 주택시장의 정보 효율성을 진단하고 정책 의사결정의 질을 높이는 실용적 도구로 기능할 수 있기를 기대한다.

참고문헌

- 고희운, 강상훈. (2023). 한국주택시장에서의 전이효과 분석: 분위별 전이지수 방법 활용. *주택금융연구*, 7(1), 131-152.
- 곽승준, 이주석. (2006). 부동산정책이 주택가격의 변동성 변화에 미치는 영향: 주택가격의 변동성 변화 시점을 중심으로. *주택연구*, 14(2), 175-194.
- 김대원, 유정석. (2014). 마코프 국면전환 AR 모형을 이용한 주택 가격 지수 변동성 결정 요인 분석. *주택연구*, 22(3), 69-99.
- 김상배, 이승아. (2021). 대도시 아파트매매가격 변동성의 전이효과: 상승기와 하락기를 중심으로. *주택연구*, 29(2), 107-133.
- 문외솔. (2023). 주택가격과 주택거래량: 주택시장 매칭모형을 중심으로. *주택도시금융연구*, 8(2), 25-54.
- 민선옥, 이영수. (2019). 부호제약을 이용한 한국 주택시장의 변동요인 분석. *부동산분석*, 5(1), 1-13.
- 박영준, 김기호. (2017). 수도권 주택가격 변동의 동조화와 변동성 전이. *부동산학보*, 69, 131-145.
- 심성훈. (2010). 주택가격 변동성의 특성 및 인플레이션과의 관계. *부동산학보*, 41, 225-240.
- 윤성민. (2024). 주택가격 변동의 지역간 파급효과 분석. *주택금융연구*, 8(2), 71-103.
- 윤성민, 손승화, 이정인. (2016). 지역주택가격 변동의 장단기 결정요인에 관한 실증분석. *부동산학보*, 67, 198-211.
- 이경근, 전재범. (2020). 아파트 매매가격지수와 실거래가격지수를 활용한 위험수익률 특성 비교: 서울특별시 강남구를 중심으로. *감정평가학논집*, 19(3), 83-116.
- 이현미, 전해정. (2020). 수도권 지역 주택가격 순환주기 변동의 영향요인 분석. *부동산경영*, 22, 307-331.
- 임재만. (2006). 주택매매가격의 변동성에 관한 연구. *주택연구*, 14(2), 65-84.
- 임재만. (2011). 주택거래량은 주택가격 변동을 설명할 수 있는가? *국토연구*, 69, 3-18.
- 전해정. (2015). 마코프 국면전환모형을 이용한 주택시장 경기국면 변동 분석에 관한 연구. *부동산학보*, 63, 119-129.
- 전해정. (2017). 주택매매, 전세, 월세가격 변동성 결정요인 분석에 관한 실증연구: 마코프 국면전환 모형을 이용하여. *부동산학보*, 70, 178-191.
- 전해정, 박헌수. (2012). 주택시장과 거시경제변수 요인들간의 동태적 상관관계 분석. *주택연구*,

202), 125-147.

- Aït-Sahalia, Y., Mykland, P. A., & Zhang, L. (2011). Ultra high frequency volatility estimation with dependent microstructure noise. *Journal of Econometrics*, 160(1), 160-175.
- Anenberg, E., & Bayer, P. (2020). Endogenous sources of volatility in housing markets: The joint buyer-seller problem. *International Economic Review*, 61(3), 1195-1228.
- Anenberg, E., & Laufer, S. (2017). A more timely house price index. *The Review of Economics and Statistics*, 99(4), 722-734.
- Asadov, A. I., Ibrahim, M. H., & Yildirim, R. (2025). Impact of house price on economic stability: Some lessons from OECD countries. *The Journal of Real Estate Finance and Economics*, 71(2), 254-284.
- Bandi, F. M., & Russell, J. R. (2006). Separating microstructure noise from volatility. *Journal of Financial Economics*, 79(3), 655-692.
- Diamond, P. A. (1982). Aggregate demand management in search equilibrium. *Journal of Political Economy*, 90(5), 881-894.
- Díaz, A., & Jerez, B. (2013). House prices, sales, and time on the market: A search-theoretic framework. *International Economic Review*, 54(3), 837-872.
- Genesove, D., & Han, L. (2012). Search and matching in the housing market. *Journal of Urban Economics*, 72(1), 31-45.
- Head, A., Lloyd-Ellis, H., & Sun, H. (2014). Search, liquidity, and the dynamics of house prices and construction. *American Economic Review*, 104(4), 1172-1210.
- Hill, R. J., Scholz, M., Shimizu, C., & Steurer, M. (2022). Rolling-time-dummy house price indexes: Window length, linking and options for dealing with low transaction volume. *Journal of Official Statistics*, 38(1), 127-151.
- Ngai, L. R., & Tenreyro, S. (2014). Hot and cold seasons in the housing market. *American Economic Review*, 104(12), 3991-4026.
- Piazzesi, M., & Schneider, M. (2009). Momentum traders in the housing market: Survey evidence and a search model. *American Economic Review*, 99(2), 406-411.
- Pissarides, C. A. (2000). *Equilibrium unemployment theory* (2nd ed.). Cambridge, MA: MIT Press.

- Stein, J. C. (1995). Prices and trading volume in the housing market: A model with down-payment effects. *The Quarterly Journal of Economics*, 110(2), 379-406.
- Wheaton, W. C. (1990). Vacancy, search, and prices in a housing market matching model. *Journal of Political Economy*, 98(6), 1270-1292.

(논문 접수일: 2026.03.09. 수정논문 접수일: 2026.05.26. 논문 채택일: 2026.06.05.)

Apartment Transaction Volume and Price Volatility: Construction and Application of a Transaction–Volume–Based Price Uncertainty Index

Chunil Kim*

Abstract

Observed housing price volatility comprises both fundamental value changes and sampling noise arising from transaction scarcity; however, conventional indicators fail to distinguish between these two components. This study aimed to address this problem by constructing a novel transaction- volume-based price uncertainty index (PUI), defined as the rolling standard deviation of hedonic model residuals divided by the square root of the number of transactions. Using apartment transaction records and building registry data, a quarterly panel of 93 districts across the Seoul metropolitan area and five major metropolitan cities was constructed. The correlation between the PUI and realized volatility (RV) was only 0.38, confirming that the two indicators captured distinct information. The explanatory power of transaction volume on volatility operated through time-varying pricing noise, and the PUI significantly predicted future transaction volume changes that RV failed to capture. Asymmetric analysis revealed that additional transactions had no effect on volatility during boom periods, whereas volatility amplification was pronounced during market contractions; this asymmetry was more clearly observed in the PUI than in RV.

Keywords : Housing Price Volatility, Price Uncertainty Index, Transaction Volume, Hedonic Price Model, Transaction Volume–Volatility Asymmetry

* Chunil Kim, Corresponding author, Assistant Professor, Division of Real Estate and Construction, Kangnam University, ckim@kangnam.ac.kr

© Copyright 2026 Housing Finance Research Institute. 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.

주거 점유 형태 전이 행렬을 통해 본 주거 사다리의 분절 현상 분석

정경채* · 박정은**

요약

본 연구는 최근 5년간의 주택금융 및 보급자리론 실태조사 데이터를 활용하여 마르코프 연쇄(Markov chain) 기반 전이 행렬 모델을 이용해 한국 주거 사다리의 구조적 변동과 비대칭성을 분석하였다. 실증 분석 결과, 지난 5년간 주택 시장의 기저 이동 질서가 급격히 재편(Frobenius norm 0.5877)됨에 따라 주거 사다리의 상하위 계층 간 이동 속도 격차가 심화된 것으로 나타났다. 특히 상위 구간(전세→자가)의 주거 상향 시간은 단축(MFPT[mean first passage time] 1.96단계)된 반면, 하위 구간(보증금 있는 월세→자가)은 도달 시간이 지연(MFPT 5.02단계)되는 비대칭성이 확인되었다. 나아가 구조적 전이 민감도 분석 결과, 현실적 대안인 전세 기반 순차 진입의 파급 효과가 5년 사이 약 9배(0.026→0.231) 급증하며 새로운 핵심 병목으로 부상한 반면, 최하위 가구의 상향 이동 민감도는 사실상 0에 수렴하여 정책 사각지대가 존재함이 입증되었다. 장기 에르고딕 시뮬레이션 결과, 향후 전세 비중이 7.3%까지 급감하며 주거 시장이 자가와 월세로 양극화될 우려가 크다. 이에 본 연구는 최하위층을 위한 직접 복지와 중간층을 위한 징검다리 금융을 병행하는 이원화된 정책 패러다임을 제언한다.

핵심어 : 주거 사다리, 마르코프 연쇄, 이동 속도의 비대칭성, 평균 도달 시간(Mean First Passage Time, MFPT), 구조적 전이 민감도

* 정경채, 주저자, 한양대학교 도시공학과 연구교수, 서울대학교 경제학부 강사, staffj@snu.ac.kr

** 박정은, 교신저자, 한양대학교 도시대학원 및 부동산융합대학원 조교수, jep2104@hanyang.ac.kr

© Copyright 2026 Housing Finance Research Institute. 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.

1. 서론

1. 연구의 배경: 사회적 기제로서의 주거 사다리와 그 균열

한국 경제의 고속 성장기 동안 '주거 사다리(housing ladder)'는 단순한 주거지 선택의 문제를 넘어, 가계의 자산 형성 및 사회적 계층 이동을 가능케 했던 핵심적인 경제적·사회적 기제였다. 전통적인 한국형 주거 사다리는 가계의 생애주기와 소득 성장에 발맞추어 '보증금 없는 월세 → 보증금 있는 월세 → 전세 → 자가'로 이어지는 선형적 상향 이동 경로를 의미했다. 특히 한국 특유의 전세 제도는 임차 주거비 부담을 낮추는 동시에 최근 전세 자금 대출이 보편화되기 전까지는 전세 보증금이라는 강제 저축 수단을 제공함으로써, 자가 취득을 위한 레버리지(leverage) 역할을 충실히 수행해 왔다.

정부 또한 이러한 사다리의 선순환을 유도하기 위해 생애 최초 LTV(loan-to-value) 80% 완화, 디딤돌 대출, 버팀목 전세 자금 대출 등 다양한 정책 금융을 투입하며 가계의 상향 이동을 지원해 왔다. 그러나 최근 주택 시장을 둘러싼 거시 경제적 환경과 임대차 시장의 구조적 변화는 기존 주거 이동 경로의 작동 원리에 심각한 균열을 내고 있다.

우선적으로 과거의 전세 보증금은 임차인이 근로 소득을 꾸준히 모아 축적한 순수 자기 자본으로서 훌륭한 강제 저축의 기능을 담당했다. 하지만 최근 전세 자금 대출의 보편화로 인한 전세의 금융화 현상은 보증금의 질적 성격을 순자산에서 부채로 변질시켰으며, 이로 인해 전세 제도가 지녔던 전통적인 자산 축적 및 징검다리 기능이 근본적으로 약화되었다.

이처럼 전세 제도가 부채에 크게 의존하게 된 상황에서 최근 고금리 기조의 장기화는 임대차 시장에 치명적인 왜곡을 초래했다. 과거 저금리 시기에는 전세 자금 대출 이자가 월세 비용보다 저렴하여 임차인이 대출을 활용해 전세로 진입하는 것이 경제적으로 유리했다. 그러나 대출 이자가 월세 비용을 상회하는 이른바 주거비 역전 현상이 발생하면서 임차인들의 자발적인 전세 진입 유인이 크게 저하되고 있다.

이에 더하여 과도한 레버리지로 확대된 전세 보증금이 무리한 갭 투자에 악용되면서 임차인의 자산을 지켜주던 안전성마저 무너져 내렸다. 과거 고성장기에 전세 보증금은 다음 단계인 자가 취득을 위한 안전한 자산 보관처이자 종자돈의 역할을 충실히 수행했다. 그러나 최근 깡통 전세 및 대규모 전세 사기 사태 등으로 보증금 미반환 위험이 광범위하게 현실화되면서 전세는 자산 축적의 발판이 아닌 자본 상실의 뇌관으로 작용할 위험성이 커졌다.

결정적으로 이 모든 중간 단계를 거치더라도 최종 목적지인 자가 취득의 문턱은 극복 불가능한 수준으로 높아졌다. 과거에는 주택 가격 상승률이 가계의 근로 소득 증가율과 어느 정도 보조를 맞추었기에 임차인들은 예측 가능한 저축을 통해 다음 주거 단계로 도약할 수 있었다. 그러나 최근 막대한

유동성과 함께 주택 가격이 급등하여 근로 소득과 자산 가격이 완전히 탈동조화되면서 전세라는 징검다리를 밟더라도 개인의 저축만으로는 도달할 수 없는 거대한 자본의 장벽이 형성되었다.

결국 주거 사다리의 중추를 담당하던 전세 제도의 강제 저축 기능, 경제적 이점, 자산의 안전성, 그리고 예측 가능한 상향 동력이 이처럼 동시 다발적으로 흔들리면서 현재 가계의 주거 이동이 단순한 일시적 정체를 겪고 있는 것인지 아니면 상향 이동의 경로 자체가 구조적으로 단절되는 위기에 처해 있는지에 대한 면밀한 학술적 진단이 불가피해졌다.

2. 문제 제기: 구조적 변동의 계량화와 정책의 역설

최근 주거 사다리의 단절 위기가 고조됨에 따라 정부는 무주택 가구의 주거 상향을 지원하기 위해 생애 최초 LTV 완화, 특례 보증자리 론, 디딤돌-버팀목 대출 등 연간 100조 원 안팎에 달하는 막대한 규모의 정책 금융을 주택 시장에 투입하고 있다. 이러한 용자 중심의 정책적 개입은 가계의 자금 조달 문턱을 낮추어 주면 자연스럽게 월세 → 전세 → 자가로 이어지는 상향 이동 경로가 복원될 것이라는 낙관적인 기대에 기반하고 있다.

그러나 최근 심화되는 주택 시장의 구조적 변동 양상을 고려할 때 과연 이러한 획일적인 금융 지원책이 의도한 대로 주택 시장 참여자 모두에게 긍정적으로 작동하여 실질적이고 보편적인 주거 사다리 복원에 기여하고 있는지에 대해서는 근본적인 의문이 제기된다.

이러한 문제 의식은 한국 특유의 주택 금융 구조를 미국의 사례와 비교할 때 더욱 뚜렷해진다. 미국의 경우 주거 복지 및 금융 정책이 저소득층을 위한 직접적인 임대료 바우처 지원(Section 8 등)이나 초기 자본이 부족한 층을 위한 파격적인 모기지 보증(FHA loan) 등 주로 월세 하위 계층의 주거 안정이나 직접적인 자가 진입을 돕는 데 초점이 맞추어져 있다. 반면 한국의 정책 금융은 주로 전세 보증금에 대한 대출 지원이나 자가 구입을 위한 LTV 완화 등 자산이나 신용도가 이미 일정 수준 이상 형성된 계층, 즉 사다리의 중간 이상을 점유한 가구에 혜택이 집중되는 구조적 한계를 띤다. 따라서 문제의 핵심은 정부의 금융 지원 효과가 가계가 처한 현재의 점유 형태 및 자산 수준에 따라 균등하게 나타나고 있는가 하는 점이다.

만약 정부의 개입이 자본과 신용도를 이미 일정 수준 갖춘 상위 구간의 이동만을 촉진하고 하위 구간의 진입 장벽은 낮추지 못한다면, 이는 정책이 의도치 않게 주거 이동의 비대칭성을 심화시키는 역설적인 결과를 초래하게 된다. 사다리 하단에 위치한 취약 계층이 상위 단계로 도약하지 못한 채 월세 시장에 장기적으로 고립되는 이른바 월세 함정에 빠진다면, 이는 단순한 주거 불안을 넘어 사회적 계층 이동의 완전한 단절을 의미한다. 결국 본 연구는 정부의 획일적인 금융 지원책이 의도치 않게 주거 사다리의

허리를 절단하고 상위 구간의 이동은 활발해지는 반면 하위 구간의 병목은 심화되는 상하위 계층 간 주거 이동 속도의 비대칭성을 오히려 고착화하고 있지 않은가라는 비판적 문제 의식에서 출발한다.

따라서 현재의 주거 사다리가 직면한 위기를 진단하기 위해서는 단순히 거시적인 자가 점유율의 증감만을 살펴서는 안 된다. 실제 주거 시장의 기저에서 가계의 이동 메커니즘이 어떻게 재편되고 있는지 그리고 정책 금융의 효과가 각 점유 형태별 구간마다 얼마나 차별적으로 작동하고 있는지에 대한 동태적이고 정밀한 계량 분석이 시급히 요구된다.

3. 연구의 목적 및 차별성

본 연구는 최근 5년간 주거 이동의 양상이 소득 및 자산 계층에 따라 어떻게 비대칭적으로 재편되고 있는지 마르코프 연쇄(Markov chain) 기반의 전이 행렬 분석을 통해 실증적으로 점검하고자 한다. 기존 연구들이 특정 시점의 단면적인 주거 실태 분석에 치중했다면 본 연구는 다음과 같은 세 가지 측면에서 차별성을 확보하고자 한다.

첫째, 전이 행렬을 활용하여 점유 형태 간 이동 확률을 시계열적으로 추적함으로써 주거 이동의 동태적 메커니즘을 미시적으로 분석한다. 둘째, 평균 도달 시간 및 기대 체류 기간 등의 지표를 통해 소득 계층 및 점유 형태별 주거 정책의 심각성을 수치화하고 정책이 시장에 미치는 영향력을 계량적으로 산출한다. 셋째, 분석 결과를 바탕으로 단순한 자가 점유율 제고를 넘어 분절된 사다리를 다시 연결하기 위한 징검다리 금융으로의 패러다임 전환을 제언함에 있어서 징검다리 금융의 당위성을 강조하는 데 그치지 않고 실무적인 작동 기제와 위험 관리 방안을 심층적으로 다루어 정책 당국에 구체적인 로드맵을 제시한다.

이를 통해 본 연구는 주거 안정이 자산 가치의 보존을 넘어 사회적 통합과 이동성 회복의 토대가 될 수 있는 정책적 근거를 마련하고자 한다.

II. 선행연구 검토

주거 사다리에 관한 기존 연구들은 과거 가계의 소득과 생애 주기에 따른 자연스러운 주거 이동 패턴을 분석하는 데 머물렀으나, 최근 주택 시장이 구조적 변동을 겪으면서 거시 경제적 충격, 임대차 시장의 구조적 붕괴, 그리고 자산 불평등이 어떻게 맞물려 가계의 상향 이동을 가로막고 있는지를 다각도에서 규명하고 있다. 본 장에서는 최근 시장의 구조적 변동과 양극화를 다룬 실증 문헌들을 고찰하고 본 연구의 차별성을 도출한다.

1. 주거 사다리의 균열과 자산 대물림의 고착화

한국의 전통적인 주거 사다리는 개인의 저축과 근로 소득을 바탕으로 한 상향 이동을 전제로 하였다. 방송희(2017)는 계층별 주거 트렌드 분석을 통해 소득 수준 및 생애 주기에 맞춘 주거 안정 지원의 패러다임 전환 필요성을 일찍이 제기한 바 있다. 그러나 최근의 실증 연구들은 자산 불평등이 심화됨에 따라 개인의 노력만으로는 주거 상황이 점차 불가능해지고 있음을 지적한다. 이소영·이창무(2019)는 부모의 경제력이 청년층 임차 가구의 주거 사다리 상향 이동에 미치는 영향을 분석하여 부의 이전 없이는 자가 진입이 사실상 차단된 현실을 규명하였다.

특히 최근 주택 가격의 급등과 맞물려 심화된 가계 양극화 현상은 다수의 연구에서 확인된다. 최종호 외(2024)는 포트폴리오 이론 관점에서 주택 시장의 자산 증식 기능이 유주택자와 무주택자 간의 불평등을 비가역적으로 확대하고 있음을 실증하였다. 특히 심승규·오수현(2022)은 무주택 가구의 생애 첫 주택 진입 시기와 거주 기간을 분석한 결과 초기 자산 규모 등 사회경제적 지위의 대물림이 주택 마련에 지배적인 영향을 미치며 결국 주거 사다리가 불평등을 고착화하는 기제로 작용하고 있음을 실증하였다. 이러한 주거 사다리 단절에 대한 청년층의 불안감은 2030세대의 무리한 주택 매입, 이른바 영끌 현상으로 가시화되었음이 실증 분석을 통해 확인된다(홍정훈·임재만, 2024).

2. 전세제도의 금융화와 갭 투자의 구조적 리스크

최근 주거 사다리의 구조적 붕괴를 설명하는 핵심 기제로 전세 제도의 금융화에 주목하는 연구들이 비중 있게 다루어지고 있다. 이후빈·홍다솜(2022)은 한국 주거 자본주의의 특성을 분석하며 전세 자금 대출의 급격한 팽창이 임차인의 거주 안정성 제고라는 당초 취지를 넘어 주택 투자의 금융화를 초래했음을 실증하였다. 이들은 완화된 전세 대출이 임차인의 레버리지 확대를 부추기는 동시에 결과적으로 임대인의 무자본 갭투자를 용이하게 만드는 자금 조달 경로로 작동하여 주택 시장 전체의 위험을 증폭시켰음을 지적했다.

나아가 이후빈(2024)은 전세 제도의 금융화가 무주택 임차 가구를 주택 소유로부터 더욱 배제시키고 민간 임대 시장의 불안정을 촉진하는 근본 원인임을 규명하였다. 해당 연구는 임차인의 전세 자금 대출을 지렛대 삼은 임대인의 갭 투자가 주택 가격 상승에 따른 이익은 사유화하는 반면, 보증금 미반환과 같은 막대한 위험은 하위 임차인과 공적 보증 기관에 전가하는 한국형 주거 불평등의 고착화 과정임을 강도 높게 비판하였다. 이러한 선행 연구들은 과거 순수 자기 자본에 기반한 강제 저축 수단이었던 전세 제도가 이제는 부채 기반의 고위험 구조로 변질되었으며 결과적으로 상향 이동을 돕던 주거 사다

리의 중간 단계를 스스로 파괴하는 구조적 모순에 직면했음을 학술적으로 뒷받침하고 있다.

3. 전세의 월세화

최근 하위 임차 계층의 주거 이동을 가로막는 가장 치명적인 요인은 전세 시장의 불안정과 붕괴에 따른 비자발적 월세화 현상이다. 정대성(2022)은 아파트 매매-전세-월세 시장 간의 수익률 전이 효과를 계량적으로 분석하여 전세 시장의 불확실성과 가격 변동성이 하위 시장인 월세 시장으로 즉각 전이되어 세입자의 부담을 가중시키는 메커니즘을 확인하였다. 이에 더해 선수봉·최민섭(2023)은 정부의 주택 정책 및 규제가 오히려 임대인의 수익성을 악화시켜 조세 전가 행위를 유발하고 수도권 아파트 전세의 월세화를 가속화했음을 지적하였다.

이러한 전세의 월세화는 단기적인 현상을 넘어 장기적인 추세로 굳어지고 있다. 최근 기계 학습을 활용하여 아파트 월세 지수를 산정한 김민성 외(2024)의 연구는 전세 계약 비중이 급감하고 보증부 월세가 임대차 시장의 주력으로 부상하고 있는 시장의 구조적 팽창을 시사한다. 이는 과거 방승희(2017)가 선제적으로 지적했던 계층별 주거 트렌드의 이질화가 현실화된 것으로 하위 가구의 주거비 부담 증가와 자산 축적 기회의 상실을 의미한다.

4. 비아파트 시장의 붕괴

더욱 심각한 문제는 사다리의 하부를 지탱하던 비아파트 전세 시장의 붕괴이다. 이상영·서정렬(2023)은 주거 복지를 명분으로 확대된 전세 대출과 보증 제도가 무자본 갭 투기의 수단으로 악용되면서 대규모 전세 사기를 촉발했음을 비판하였다. 실제로 민병철(2023)은 강동 전세 아파트를 중심으로 보증금 미반환 위험을 추정하여 전세 제도의 구조적 취약성을 경고하였고, 안선영·이상엽(2025)은 수도권 보증 사고 실거래 데이터를 통해 미반환 리스크가 다세대, 연립 등 비아파트에 압도적으로 집중되어 있음을 실증하였다.

이러한 위험은 김성찬·성주한(2025)이 분석한 DSR(debt service ratio) 규제 강화에 따른 임차권 등기 명령 건수 급증 현상으로 표출되며 세입자의 주거 불안을 극대화하였다. 그 결과 국토연구원(2024)과 한국개발연구원(2024)이 경고한 바와 같이 비아파트 임대차 시장의 기반이 크게 훼손되었고, 윤성진 외(2026)의 연구는 청년층을 중심으로 확산되는 비아파트 소유 및 전세 기피 현상이 단순한 선호 변화가 아닌 생존을 위한 위험 회피 기제임을 지적하였다. 특히 이지연·이정란(2026)의 실증 분석에 따르면 보증 가입 요건 강화 등의 제도 변화가 외려 서민들을 비아파트 월세 시장으로 대거

밀어내며 하위 계층이 자산 형성이 불가능한 월세 함정에 강제로 갇히는 결과를 낳았음이 확인된다.

5. 주택 가격의 급등과 극복 불가능해진 자산 격차

최근의 선행 연구들은 폭등한 주택 가격이 단순한 주거 비용의 증가를 넘어 노동 소득으로는 결코 극복할 수 없는 구조적 자산 격차를 고착화하는 핵심 기제임을 공통적으로 지적한다. 박정재 외(2024)는 실증 분석을 통해 주택 가격 상승기에 자가 보유 가구와 무주택 가구 간의 순자산 격차가 극심하게 확대되었으며, 이것이 전체 가계의 경제적 불평등을 견인하는 주된 원인임을 규명하였다. 국토연구원의 오민준(2022)과 이형찬 외(2020)의 연구 역시 가계 총자산 불평등에 대한 거주 주택 자산의 기여도가 압도적이며 주택 소유 여부 및 가격 상승에 따른 막대한 자본 차익이 근로 소득의 축적 속도를 상회하면서 자산 불평등을 비가역적으로 심화시켰음을 시사한다. 이는 과거 성실한 임금 소득의 저축을 통해 주택을 마련하고 다음 주거 단계로 도약하던 전통적인 상향 이동의 기반이 붕괴되었음을 시사한다.

이처럼 극단적으로 벌어진 부동산 자산 격차는 무주택 가구에 극복 불가능한 진입 장벽으로 작용하며 궁극적으로 한국 사회 전반의 역동성을 파괴하는 계층 고착화를 초래하고 있다. 서재원·진장익(2023)은 자산 가치의 공간적 불평등이 사람들의 계층 이동 가능성 인식에 미치는 영향을 계량 분석하여 폭등한 주택 가격이 유발한 부동산 격차가 일반 시민들과 청년층으로 하여금 개인의 노력을 통한 미래 계층 상승 기대감을 유의미하게 꺾고 심리적 좌절감을 심화시킨다는 점을 실증하였다. 결국 근로 소득만으로는 결코 좁힐 수 없을 만큼 벌어진 주택 가격의 초격차는 하위 임차 가구를 영구적인 주거 불안정과 월세 함정에 가두고 나아가 세대 간 부의 대물림을 통해 경제적 불평등을 구조적으로 재생산하는 치명적인 주거 단절의 원인이 되고 있다.

6. 선행 연구의 한계 및 본 연구의 차별성

이상의 선행 연구들은 비아파트 시장의 붕괴, 비자발적 구조적 월세화, 세대 간 부의 대물림, 그리고 주택 금융 정책이 야기하는 레버리지의 계층 편중 현상 등 개별적인 주택 시장의 모순을 훌륭하게 실증해 냈다. 그러나 기존 문헌들은 주로 횡단면 데이터에 기반하여 특정 시점의 현상이나 거시 변수 간의 인과 관계를 규명하는 데 그쳤으며, 이러한 충격들이 개별 가구의 점유 형태 간 연쇄적 이동 궤적을 어떻게 단절시키고 특정 구간의 체류를 고착화하는지를 동태적으로 추적한 연구는 상대적으로 부족하다.

이에 본 연구는 거시 경제 충격과 정책 변화가 중첩된 2021~2025년의 가계 반복 횡단면 데이터를 대상으로 마르코프 연쇄(Markov chain) 전이 행렬 모형을 선도적으로 도입한다. 각 점유 형태별 자가 도달 평균 소요 시간의 계층 간 격차를 시계열적으로 산출하며, 특히 최종호 외(2024) 등이 질적으로 비판했던 정책 금융의 편중 부작용을 보다 엄밀하게 입증하기 위해 편미분 기반의 구조적 전이 민감도 지표를 산출한다. 이를 통해 현행 주택 금융 정책이 자산 축적의 마지막 단계(전세 → 자가)에만 유효하게 작동하고 사다리의 최하단(월세 → 전세)에는 전혀 닿지 못하는 정책 사각지대의 크기를 정량적으로 도출한다는 점에서 기존 문헌들과 확고한 학술적 차별성을 확보한다.

III. 데이터 및 분석 방법

본 연구는 주거 점유 형태 간의 동태적 이동을 분석하기 위해 마르코프 연쇄(Markov chain) 모델을 기본 골격으로 하며, 시장의 구조적 변화와 정책적 사각지대를 계량화하기 위해 행렬 거리 및 민감도 분석을 병행한다.

1. 자료원 및 주요 변수

본 연구는 주택금융통계시스템에서 제공하는 주택금융 및 보금자리론 실태조사를 활용하였다. 해당 데이터베이스에서는 2021년부터 2025년까지 연간 5,000가구를 대상으로 한다. 주택금융 및 보금자리론 실태조사는 매년 새로운 가구를 추출하는 횡단면 조사이나 응답 시점에 조사 대상 가구의 현재 점유 형태뿐만 아니라 직전 거주지의 과거 점유 형태를 동시에 묻는 회고적 문항을 포함하고 있다. 이는 계량 경제학 및 사회학 연구에서 널리 활용되는 방법으로 개별 가구 단위에서 1기의 상태 변화를 완벽하게 추적할 수 있는 미시적 의사 패널 데이터를 구축할 수 있게 해 준다. 즉, 2021년도에 조사된 5,000가구의 횡단면 표본은 그 자체로 2020년에서 2021년 사이의 점유 형태 전이 이력을 온전히 담고 있다. 따라서 이러한 전이행렬은 본 연구의 주요 목적인 자산 축적 단계별 이동을 분석하는데 적합한 자료라고 볼 수 있다(주택금융연구원, 2026). 나아가 표본 규모의 통계적 신뢰성 및 우연성(noise) 개입 여부를 투명하게 검토하기 위해, 연간 5,000가구 표본을 기준으로 전년도와 당해 연도 점유 형태 간의 실제 점유형태 이동 가구 수(transition counts) 기초 통계표를 산출하여 <부록>에 별도로 제시하였다.

분석 모형 설정에 앞서, 주거 이동의 다차원적 특성과 본 연구의 방법론적 전제를 명확히 할 필요가 있다. 주거 이동은 입지, 주택 유형 등과 같은 주거 효용과 점유 형태 등과 같은 자산 축적의 다차원적

결합이다. 예를 들어 비아파트 전세에서 아파트 월세로의 이동이 주거 수준의 상향일 수는 있으나, 자산 축적의 징검다리 관점에서는 매몰 비용이 증가하는 구조이다. 본 연구의 분석 목적은 주거의 질적 향상이 아닌 다음 주거 단계로 도약하기 위한 금융 레버리지 및 종자돈 형성 경로로서의 기능에 국한되므로, 본 연구는 재무적 관점에서의 점유 형태 중심의 분석을 수행한다. 분석의 정밀도를 높이기 위해 주거 점유 상태 공간(state space, S)을 자산 축적의 사다리 위계를 반영하여 다음과 같이 5가지 범주로 정의하고 기호화한다.

O (Owned): 자가

J (Jeonsei): 전세

R_1 (Deposit Rental): 보증금 있는 월세

R_2 (No-Deposit Rental): 보증금 없는 월세, 사글세, 연세 등

R_3 (Free): 무상

본 연구의 마르코프 연쇄 모델은 위 5가지 상태를 모두 포괄하여 전체 주택 시장의 기저 질서를 온전히 모델링한다. 다만, 실증 분석(제4장) 및 정책적 시사점 도출 과정에서는 주거 사다리 상향 이동의 핵심 연결 고리이자 주택금융 정책의 주 타겟인 전세(J), 보증금 있는 월세(R_1), 보증금 없는 월세(R_2) 구간의 전이 양상 및 평균 도달 시간(mean first passage time, MFPT)에 논의의 초점을 맞춘다.

2. 분석 기법: 마르코프 연쇄(Markov Chain) 모델

Grinstead & Snell(1997)에서 소개된 대로 가계의 주거 이동을 1차 이산시간 마르코프 과정(first-order discrete-time Markov process)으로 가정한다. 시점 t 에서 $t+1$ 로 넘어갈 때, 주거 상태 i 에서 상태 j 로 이동할 전이 확률(transition probability) P_{ij} 는 다음과 같은 조건부 확률로 정의된다.

$$P_{ij} = \Pr(X_{t+1} = j | X_t = i) \quad \langle \text{식 1} \rangle$$

여기서 X_t 는 시점 t 에서의 주거 점유 상태이며, 전체 전이 확률을 원소로 하는 5×5 행렬 P 는 모든 행의 합이 1이 되는 확률 행렬(stochastic matrix)의 성질을 갖는다.

$$P_{ij} \geq 0, \quad \sum_{j \in S} P_{ij} = 1, \quad (\forall i \in S) \quad \langle \text{식 2} \rangle$$

본 모형을 통해 장기 정상 상태(steady-state) 분포를 시뮬레이션하기 위해서는 전이 행렬의 에르고딕(ergodic) 성질이 확인되어야 한다. 한국 주택 시장은 비자가 → 자가로의 강한 일방향성을 띠고 있어, 경제학적 직관으로는 자가(O) 점유 형태가 한 번 진입하면 빠져나올 수 없는 완벽한 흡수 상태(absorbing state, $P_{OO} = 1$)로 오인될 소지가 있다. 만약 자기가 완벽한 흡수 상태라면 시스템은 에르고딕성을 상실하고, 장기 분포를 논하는 것 자체가 수학적 오류가 될 수 있다.

그러나 실증 패널 데이터(〈부록 표 1〉)에 따르면, 자가 가구의 자가 체류 확률(P_{OO})은 연도별로 0.917~0.935 수준에 머물며 완벽한 1.0이 아님이 증명된다. 실제로 매년 약 6.5%~8.3%의 가구가 자가에서 전월세 시장으로 하향 및 수평 이동하고 있다. 이러한 자가 이탈 현상은 단순한 파산 등 비자발적 몰락에만 기인하지 않는다. 현실 주택시장에서는 1) 은퇴 고령층이 생활 자금 확보를 위해 자택을 매각하거나 임대하고 소형 임차로 이동하는 자산 유동화(downsizing), 2) 직장 발령이나 자녀 교육 등의 사유로 본인 소유 주택을 전세로 임대하고 타지역에서 전월세로 거주하는 한국 특유의 전략적 주거 소비 행태가 지속적으로 관찰되기 때문이다.

결과적으로 본 모형의 상태 공간 내 모든 점유 형태 간에는 0이 아닌(non-zero) 직간접 확률 이동 경로가 상존하여 마르코프 연쇄의 기약성(irreducibility)을 충족하며, 특정한 주기로만 상태가 반복되지 않으므로 비주기성(aperiodicity) 요건도 만족한다. 따라서 엄밀한 수학적 의미에서 에르고딕(ergodic) 가정이 성립하며, 거시적 장기 균형점 산출의 이론적 타당성이 확보된다. 결국 본 연구에서 산출된 장기 정상 상태는 완벽한 흡수 상태로 인한 수학적 쏠림 현상이 아니라, 막대한 상향 진입 압력과 생애주기에 따른 유의미한 하향 이탈 압력이 장기적으로 맞서며 도달하는 준-흡수 상태(quasi-absorbing state)의 동태적 균형점으로서 학술적 타당성을 지닌다. 따라서, 비록 5년의 단기 데이터를 기반으로 장기적 에르고딕성을 완벽히 담보하기에는 무리가 있으나, 장기 거시 균형점의 상대적 변화를 파악하기 위해 본 가정을 채용하였다.

3. 주거 사다리의 동태적 지표 산출

Kemeny & Snell(1960)에서 소개된 평균도달시간 및 흡수 상태 분석의 방법론을 주거 사다리를 분석하는 데 적용하고자 한다.

1) 기대 체류 기간(Expected Duration)

특정 주거 상태 i 에 진입한 가구가 다른 상태로 이탈하기 전까지 해당 상태에 연속적으로 머무는

것으로 기대되는 시간(단위: 년) $\mathbb{E}[D_i]$ 는 전이 행렬의 대각 원소(P_{ii})를 이용하여 다음과 같이 계산된다.

$$\mathbb{E}[D_i] = \frac{1}{1 - P_{ii}} \quad \langle \text{식 3} \rangle$$

이 지표가 높을수록 상향 이동이 지연되고 해당 주거 단계(예: 월세 시장)를 탈출하지 못하는 '고착화(lock-in)' 현상이 심화되었음을 의미한다.

2) 평균 도달 시간(Mean First Passage Time)

상태 i 에서 출발하여 최종 목표 상태인 '자가(O)'에 처음으로 도달하는 데 걸리는 평균 소요 단계(년)를 m_{iO} 라 할 때, 이는 다음과 같은 재귀적(recursive) 관계식을 통해 산출된다.

$$m_{iO} = 1 + \sum_{k \neq O} P_{ik} m_{kO} \quad (\text{단, } m_{OO} = 0) \quad \langle \text{식 4} \rangle$$

여기서 m_{iO} 는 출발 상태 i 에서 최종 목적지인 자가(O)에 도달하기까지 소요되는 단순한 단일 1회성 이동 기간이 아니다. 예를 들어 무보증 월세(R_2)에서 자가(O)로의 평균 소요 단계(m_{R_2O})는 무보증 월세에서 자가로 직행하는 극소수의 확률 경로뿐만 아니라, 보증부 월세를 거치는 경로($R_2 \rightarrow R_1 \rightarrow O$), 보증부 월세와 전세를 모두 거치는 경로($R_2 \rightarrow R_1 \rightarrow J \rightarrow O$) 등 상태 공간 내에서 발생 가능한 모든 다단계 중간 전이 경로의 확률 가중치와 각 단계별 체류 시간 기댓값을 재귀적으로 연쇄 합산하여 산출된다. 이동 표본이 적더라도 시스템 전체의 유기적 확률을 반영하므로 특정 계층의 구조적 지연 시차를 계량화하는 데 유효하다.

본 연구에서는 특히 '무보증 월세에서 자가로의 도달 시간($m_{R_2,O}$)'과 '전세에서 자가로의 도달 시간($m_{J,O}$)' 간의 격차를 산출하여 주거 사다리 하위 구간의 계층별 분절 정도를 정량화한다.

3) 정상 상태 분포(Steady State Distribution)

관측된 전이 행렬이 에르고딕(ergodic) 성질을 가진다고 가정할 때, 현재의 주거 이동 매커니즘(전이 확률 행렬 P)이 장기적으로 지속될 경우 도달하게 될 거시적 균형 상태의 주거 점유 분포

$\pi = [\pi_O, \pi_J, \pi_{R_1}, \pi_{R_2}, \pi_F]$ 는 다음의 선형 연립방정식을 만족하는 좌측 고유벡터(left eigenvector)로 도출된다.

$$\pi P = \pi, \quad \sum_{i \in S} \pi_i = 1 \quad \langle \text{식 5} \rangle$$

4. 구조적 변화 및 구조적 전이 민감도 분석

1) 행렬 거리 분석(Frobenius Norm)

2021년의 주거 전이 구조(P_{2021})와 2025년의 전이 구조(P_{2025}) 간의 총체적이고 거시적인 변화 크기를 측정하기 위해 유클리드 공간에서의 행렬 놈(norm)인 Frobenius norm을 활용한다 (Meyer, 2000).

$$\|P_{2025} - P_{2021}\|_F = \sqrt{\sum_{i \in S} \sum_{j \in S} (P_{2025,ij} - P_{2021,ij})^2} \quad \langle \text{식 6} \rangle$$

이 값이 클수록 지난 5년간 거시경제적 충격(고금리 등)으로 인해 주택 시장의 기저 이동 질서 자체가 근본적이고 급격하게 재편되었음을 의미한다.

2) 이동성 지수(Shorrocks Index)

주거 사다리 시스템 전반의 유동성과 활력을 측정하기 위해 Shorrocks(1978)의 이동성 지수 $M(P)$ 를 산출한다.

$$M(P) = \frac{n - \text{tr}(P)}{n - 1} \quad \langle \text{식 7} \rangle$$

여기서 n 은 상태의 개수(본 연구에서는 5)이며, $\text{tr}(P)$ 는 전이 행렬의 대각합(trace)이다. 지수가 0에 가까울수록 계층 고착화가 완벽히 진행되었음을, 1에 가까울수록 계층 간 주거 이동이 활발함을 나타낸다.

3) 구조적 전이 민감도 분석(Sensitivity Analysis)

기존의 정책 민감도라는 용어는 특정 점유 형태의 이동이 전적으로 주택금융 정책에 기인한다는 과도한 전제로 비칠 수 있어, 본 연구에서는 시스템의 거시적 탄력성을 의미하는 구조적 전이 민감도로 명명한다. 이는 특정 구간의 전이 확률(P_{ij})이 미세하게 증가할 때, 장기적인 최종 자가 점유율(π_o)이 얼마나 탄력적으로 변화하는지 그 한계 효과(marginal effect)를 수치 미분(numerical differentiation) 방식으로 산출한다.

$$S_{ij} \approx \frac{\pi_o(P_{ij} + \epsilon) - \pi_o(P_{ij})}{\epsilon} \quad \langle \text{식 8} \rangle$$

본 연구의 핵심적인 방법론적 차별성은 위 식의 전이 목적지(j)를 유연하게 설정하여 다양한 상향 이동 경로의 파급력을 입체적으로 비교·평가할 수 있다는 점이다. 수식의 최종 결과값은 거시 주택시장의 궁극적 목표인 최종 자가 점유율(π_o)의 상승 폭을 나타내지만, 그 원인이 되는 전이 구간(P_{ij})을 설정함에 있어 두 가지 경로를 분리한다.

첫째는 하위 임차 계층이 최종 목적지인 자가로 단번에 건너뛰는 자가 직행 경로($j = O$)이며, 둘째는 사다리의 바로 다음 징검다리로 한 칸 상향 이동하는 순차 진입 경로(예: 보증금 있는 월세 → 전세 등, $j \neq O$)이다. 이를 통해 특정 계층의 이동 병목을 해소하고자 할 때, 막대한 자본이 요구되는 자가 직행을 무리하게 유도하는 것과 현실적인 중간 징검다리를 밟게 하는 것 중 어느 방식이 거시적 주거 안정(자가율 제고)에 수리적으로 더 효율적인지 명확히 보일 수 있다.

산출된 한계 효과 값이 극히 낮게 도출되는 전이 구간은, 어떠한 정책적 지원으로도 전체 주거 사다리 생태계에 긍정적 파급을 미치지 못하는 정책 체감의 사각지대에 고립되어 있음을 시사한다.

4) 시스템 수렴 속도(Convergence Rate)

Stewart(2009)의 책에서 소개된 시스템 수렴 속도 분석을 해당 연구에서 사용하고자 한다. 좀 더 구체적으로 말하면, 현재의 전이 행렬이 정상 상태로 얼마나 빠르게 수렴하는지 측정하기 위해 두 번째로 큰 고윳값의 절댓값($|\lambda_2|$)을 산출한다. 이 값이 작을수록 장기 균형 상태로의 진입 속도가 가속화됨을 의미한다.

IV. 실증 분석 결과

본 장에서는 2021년부터 2025년까지의 전이 행렬 데이터를 바탕으로 주거 사다리의 구조적 변동성, 이동성 지수, 그리고 점유 형태별 고착화 현상을 동태적으로 분석한 결과를 제시한다.

1. 주택 시장의 구조적 변동성: Frobenius Norm 분석

2021년의 전이 행렬(P_{2021})과 2025년의 전이 행렬(P_{2025}) 간의 유클리드 거리를 측정한 결과, Frobenius norm 값은 0.5877로 산출되었다. 일반적으로 마르코프 전이 행렬 간 거리가 0.5를 상회하는 것은 단순한 표본 오차나 단기적 추세 변화를 넘어, 시장의 기저 이동 질서가 근본적으로 재편되었음을 의미한다.

이는 금리 급등기와 전세 시장 불안을 거치며 촉발된 '전세의 월세화'와 '자가 취득 문턱의 급격한 상승'이 가계의 주거 이동 경로 자체를 물리적으로 변모시켰음을 시사한다. 즉, 과거의 선형적 이동 패턴을 전제로 한 정책 설계가 더 이상 유효하지 않을 수 있다는 강력한 통계적 경고 신호이다.

2. 주거 이동성 지수 및 성분 분해(Shorrocks Index)

주거 사다리의 전반적인 활력을 나타내는 Shorrocks 이동성 지수($M(P)$)는 2021년 기준 0.716에서 2025년 0.655로 하락하였다. 이는 지난 5년간 주택 시장의 기저에서 계층 간 주거 이동의 역동성이 전반적으로 위축되며 특정 주거 형태에 갇히는 고착화 현상이 심화되고 있음을 시사한다.

이러한 전체 이동성의 위축을 상향 이동성과 하향 이동성으로 분해하여 살펴보면 시장의 구조적 악화 추세가 더욱 명확히 드러난다.

첫째, 자산 형성을 동반하는 상향 이동성 지표는 2021년 0.659에서 2025년 0.596로 감소하였다. 비록 전체 이동에서 차지하는 비중은 여전히 높으나, 그 절대적인 수치와 비중(92.1% → 90.9%)이 동반 하락하고 있다는 점은 과거 활발하게 작동하던 주거 상향의 동력이 급격히 상실되고 있음을 뒷받침한다.

둘째, 이와 대조적으로 하향 이동성 지표는 같은 기간 0.057에서 0.060으로 상승하였다. 전체 이동성에서 하향 이동이 차지하는 비중 역시 7.9%에서 9.1%로 눈에 띄게 확대되었다. 이는 고금리 충격과 전세 시장 불안으로 인해 굳건했던 자산 보유 계층의 주거 하방 경직성(downward rigidity)

에 균열이 발생하고 있으며 비자발적인 계층 하락 위험이 점차 현실화 되고 있음을 경고한다.

결과적으로 이 두 가지 변화 양상을 종합하면, 주거 사다리 전반의 이동 관성이 둔화되는 가운데 더 나은 주거 환경으로 도약하는 상향 이동의 문은 좁아지고, 반대로 하위 계층으로 추락하는 하향 이탈의 위험은 커지는 매우 부정적인 방향으로의 구조적 재편이 진행되고 있음을 명확히 보여준다.

3. 점유 형태별 고착화 및 자가 도달 시간(Mean First Passage Time) 분석

본 연구의 핵심 발견인 점유 형태별 '체류 기간'과 '자가 도달 시간(MFPT)'의 계층 간 격차, 그리고 시계열적 변화 추이는 사다리의 단절을 극명하게 보여준다(〈표 1〉).¹⁾

1) 사다리의 시차와 주거 이동의 비대칭성(Mean First Passage Time 분석)

2021년 기준 전세 가구가 자가에 도달하는 데 평균 2.18단계가 소요된 반면, 보증금 있는 월세 가구는 2.71단계가 소요되어 약 1.24배의 시차가 존재했다. 더욱 심각한 것은 2025년의 변화이다. 전세 계층의 도달 시간은 1.96으로 단축(가속)된 반면, 월세 계층의 도달 시간은 5.02로써 급격히 길어지며 약 2.56배로 시차가 벌어지며 시간적 격차의 비대칭성이 확대되고 있다. 즉, 자본력에 따른 자산 형성 속도의 양극화가 뚜렷해졌다.

보증금 있는 월세의 기대 체류 기간(1.91 → 3.73)이 상위 단계이자 징검다리인 전세(1.68)보다 점점 더 길게 역전되어 나타나는 현상은, 하위 임차 계층이 '보증금 마련'이라는 중간 칸을 돌파하지

〈표 1〉 점유 형태별 동태적 지표 및 구조적 전이 민감도(2021 vs 2025)

주거형태	기대 체류 기간 ($\mathbb{E}[D_i]$)	자가 도달 평균 소요 단계 (m_{i0})	구조적 전이 민감도 (S_{ij})	구조적 전이 민감도 (S_{i0})
전세(J)	1.68 → 1.68	2.18 → 1.96	0.119 → 0.116 (전세→자가)	
보증금 있는 월세(R_1)	1.91 → 3.73	2.71 → 5.02	0.026 → 0.232 (보증월세→전세)	0.135 → 0.381 (보증월세 → 자가)
보증금 없는 월세 등(R_2)	1.31 → 1.25	3.16 → 3.74	0.002 → -0.004 (무보증월세 → 보증월세)	0.016 → 0.011 (무보증월세 → 자가)

1) 기대 체류 기간 및 도달 시간 수치는 마르코프 모델 조건 하의 수학적 기댓값(단위: 전이 단계/step)으로, 절대적인 물리적 시간(years)보다는 계층 간 이동 속도의 '상대적 격차'를 비교하는 데 목적이 있다.

못하고 월세 시장에 장기 정체되는 고착화가 시작되는, 즉 월세 함정(rental trap) 현상이 고착화되고 있음을 뒷받침한다. 물론 부록 기초 통계에서 확인되듯 해당 최하위 구간에서 실제 상향 이동에 성공한 표본 가구 수는 극히 희소하여 산출된 수치에 일부 통계적 우연성이 개입될 한계는 존재한다. 그러나 역설적으로 이처럼 극단적인 표본의 부재 자체가 이들 계층의 상향 궤적이 물리적으로 철저히 차단되어 있음을 강력히 방증하며 도출된 지표들은 이들이 주택 시장 궤도 밖으로 밀려나 있음을 보여주는 징후적 결과로 해석되어야 한다.

특히 주목할 점은 보증금 없는 월세의 기대 체류 기간이 1.31단계에서 1.25단계로 소폭 감소한 현상이다. 표면적으로는 해당 구간을 빠르게 탈출한 것으로 보일 수 있으나, 이들의 자가 도달 소요 시간이 3.16에서 3.74로 크게 지연되었다는 점을 결합하여 해석하면 이는 주거 상향이 아닌 극단적 주거 불안정성의 발현을 시사한다. 즉, 최하위 임차 계층이 월세 비용조차 감당하지 못하여 안정적인 거주를 유지하지 못하고 빈번하게 단기 이사를 반복하거나 연세(혹은 사글세) 및 무상 거주 등 더 열악한 기타 점유 형태로 밀려나는 비자발적이고 생존형인 이탈 현상이 심화되었음을 방증하는 결과로 해석될 수 있다.

2) 구조적 병목과 사각지대의 계량화(전이 민감도 분석)

시간의 흐름에 따른 구조적 전이 민감도(한계 효과)의 두 갈래 경로(순차 진입 vs 직행 진입) 분석은 현재 주택 시장의 병목 현상과 사각지대를 입체적으로 계량화한다. 우선, 과거 상태가 보증금 있는 월세(R1)인 가구가 현재(혹은 미래)의 최종 목적지인 자가(O)로 직접 이동하는 전이 경로에 마찰력이 제거될 경우 시스템 전체의 자가율 상승에 미치는 파급 효과는 2025년 기준 0.381으로 구조적으로 가장 강력하다. 그러나 앞서 살펴본 바와 같이 이 계층의 자가 도달 소요 시간(MFPT)은 5.02단계로 지연되었으며, 주택 가격 급등으로 막대한 초기 자본이 요구되는 현실을 감안할 때 자가 직행 경로는 자본력을 갖춘 소수에게만 한정될 수밖에 없다.

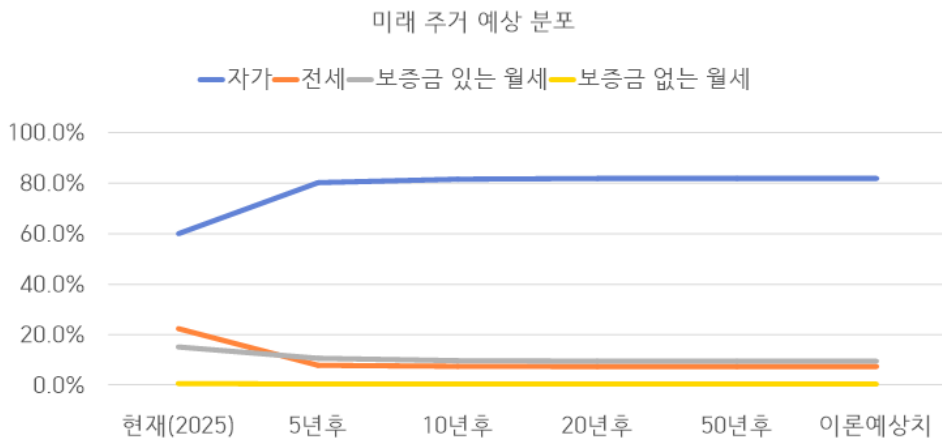
이에 본 연구가 주목하는 지점은 바로 '징검다리 경로(순차 진입)'의 폭발적인 파급력 상승이다. 과거 상태가 보증금 있는 월세(R1)인 가구가 현재의 상위 징검다리인 전세(J)로 이동하는 특정 전이 경로를 지원할 때 시스템 전체에 미치는 한계 효과는 과거 0.026(2021년)에서 2025년 0.232로 무려 약 8.9배 가까이 폭증하였다. 이는 전세 시장 불안 여파로 이 구간의 체류 기간이 두 배(1.91 → 3.73) 늘어난 현상과 맞물려, 도달하기 어려운 자가 직행만을 고집하기보다 단절된 허리(전세 진입)의 병목을 해소하는 것이 꼭 막힌 주거 사다리의 이동성을 복원하는 가장 현실적이고 효율적인 대안임을 수리적으로 입증한다.

반면, 최하위 계층인 무보증 월세 가구의 경우, 자가로 직행하는 파급 효과(0.011)나 상위 단계인 보증부 월세로 한 칸 순차 진입하는 파급 효과(-0.004) 모두 사실상 0에 수렴하였다. 이들 계층은 단순히 어떠한 경로의 상향 이동을 촉진하더라도 궁극적인 거시 시스템의 자가율 상승이나 시장 안정으로 파급되지 못할 만큼 전체 주거 사다리 생태계에서 정책 효과가 미치기 어려운 한계가 관찰된다. 결과적으로 자가 취득 중심의 기존 용자 체계 온기가 사다리 최하단에는 도달하지 못하는 광범위한 정책 체감의 사각지대가 존재함을 강력히 방증한다.

4. 미래 주거 분포 시뮬레이션 및 장기 예측

2025년의 전이 행렬 구조가 장기적으로 지속된다는 에르고딕(ergodic) 가정하에 균형점인 정상 상태 분포(π)를 시뮬레이션한 결과는 구조적 양극화를 경고한다. 이론적 정상 상태에서 자가 비율은 81.8% 수준으로 높게 수렴할 것으로 예측된다. 이는 현실의 주거 이동 경로 상에서 '자가' 점유 형태에 한 번 진입하면 하향 이탈할 확률이 극히 낮은 '준-흡수 상태(quasi-absorbing state)'의 특성을 강하게 띠기 때문이다.

자가를 제외하면 전체적으로 줄어든 것으로 예상되는데, 전세의 경우에는 22.2%에서 7.3%로, 보증금 있는 월세의 경우에는 14.9%에서 9.2%로, 보증금 없는 월세의 경우에는 0.4%에서 0.1%로 될 것으로 예측된다. 즉, 중간 사다리 역할을 하던 전세는 현재의 3분의 1수준으로 급감할 것으로 보인다(<그림 1>).



<그림 1> 미래 주거 예상 분포. 2025년의 전이 행렬이 그대로 유지된다고 가정했을 때 예측 미래 주거 분포 비율

결국, 자산 축적의 완충 지대 역할을 하던 전세가 점진적으로 축소·소멸하면서, 향후 한국의 주거 시장은 '자가(자산 보유자)' 대 '월세(자산 소외자)' 라는 양극단 구조로 재편될 가능성이 크며, 이는 가계 간 자산 불평등이 고착화할 위험이 있다.

5. 소결: 상하위 계층 간 이동의 비대칭성과 시스템의 수렴 속도

실증 분석 결과를 종합하면, 최근 한국의 주거 사다리는 자본력과 정책금융을 지렛대 삼은 "상위 구간(전세 → 자가)의 이동 가속"과 자본의 한계 및 정책 소외에 부딪힌 "하위 구간(월세 → 전세)의 이동 병목"이 동시에 발생하는 상하위 계층 간 주거 이동 속도의 뚜렷한 비대칭성 양상을 실증적으로 보이고 있다.

특히, 정상 상태로 향하는 시스템의 수렴 속도(convergence rate) 지표가 2021년의 0.469에서 2025년에는 0.287로 더 낮게 도출된 점이 시사하는 바가 크다. 마르코프 모형에서 이 값이 작아졌다는 것은 시스템이 장기 균형점(자가 81.8%, 전세 7.3%의 이분법적 고착화 상태)으로 빨려 들어가는 속도가 과거보다 훨씬 가속화되었음을 의미한다. 즉, 시장을 방치할 경우 중간 사다리가 붕괴된 이분법적 구조로 빠르게 고착화될 우려가 커졌음을 시사한다. 따라서 사다리 하단부의 병목을 강제로 뚫어줄 수 있는 외부의 강력하고 핀셋화된 정책적 개입이 절실히 요구된다.

V. 논의 및 시사점

1. 연구 결과의 요약

본 연구는 2021~2025년 가구 반복 횡단면 데이터 기반의 주거 점유 형태 전이 행렬 분석을 통해 최근 한국 주거 사다리의 상하위 계층 간 주거 이동 속도의 비대칭성 실태를 동태적으로 규명하였다. 실증 분석 결과 자본을 보유한 전세 계층은 정책 금융을 발판 삼아 자가 도달 소요 단계를 단축(MFPT 2.18단계에서 1.96단계 수준으로 가속화)하고 있는 반면, 보증금 있는 월세 계층은 정책적 사각 지대에 고립되며 자가 도달 소요 시간(MFPT 2.71단계에서 5.02단계 수준으로 가속화)이 심각하게 지연되는 비대칭성을 실증적으로 확인하였다. 특히 장기 에르고딕(ergodic) 시뮬레이션 결과 주거 사다리의 허리 역할을 하던 전세 비중이 향후 7.3%까지 급감하며 자가 대 월세의 이분법적 구조로 재편될 것이 수치로 입증되어 자산 불평등의 구조적 고착화 우려가 감지되었다.

2. 주거 열차의 이격과 정책의 역설

주거 시장의 병목을 타개하기 위해 투입된 막대한 정책 금융의 실효성에 대해서도 비판적 논의가 활발하다. 최종호 외(2024)는 포트폴리오 이론 관점에서 주택 시장의 투기적 성격과 완화적인 주택 금융 환경이 가계 양극화에 미치는 영향을 분석하였다. 이들은 주택이 거액의 초기 자본을 요하는 자산이기 때문에 획일적인 용자 중심의 지원책은 유동성 제약에 갇힌 취약 계층에게는 활용되지 못한 채 도리어 자본을 갖춘 상위 계층의 자산 증식용 레버리지로 편중 작동하여 가계 간 순자산 격차를 극심하게 벌렸음을 수학적으로 확인하였다.

유사한 맥락에서 홍정훈·임재만(2024)은 2030세대의 이른바 영끌 주택 매수 실태를 실증 분석하여 LTV 완화 등 정부의 대출 지원을 통해 자가 마련에 성공한 청년층 대다수가 이미 본인 소득이 높거나 부모의 지원이 가능한 상급 계층에 국한되었음을 밝혔다. 이는 주거 사다리 복원을 명분으로 한 현행 금융 정책이 하위 구간의 병목을 전혀 뚫지 못하고 오히려 주거 이동의 비대칭성을 가속화하는 현행 제도가 의도치 않은 계층 간 정책 효과의 비대칭성을 유발할 수 있음을 지표의 거시적 흐름을 통해 추론할 수 있다.

본 연구의 분석 결과는 한국 주거 시장이 마치 속도가 각기 다른 칸들로 구성되어 연결 고리가 끊어진 '이격된 열차'와 같음을 보여준다. 현행 정책금융이라는 가속 장치는 이미 궤도에 올라 자본을 축적한 '앞차(전세 및 보증금 있는 월세 가구)'에는 효과적으로 작동하여 자가 점유를 돕고 있으나, '뒷차(보증금 없는 월세 등 최하위 가구)'를 견인할 장치는 마련하지 못했다. 특히 자산 계층을 향한 수요 측면의 금융 지원이 이들의 자가 전환을 가속화하며 상위 주택 시장의 가격 하방을 지지한 반면, 이는 역설적으로 자본이 없는 하위 임차 가구에게는 주거 상향에 필요한 비용과 문턱을 기하급수적으로 높이는 '풍선 효과의 역설(paradox of balloon effect)'을 초래하였다.

3. 정책적 제언: 주택 금융의 본질적 한계 인식과 정책 패러다임의 이원화

1) 정책 금융의 구조적 한계와 거시 경제적 비효율성

본 연구의 실증 분석 결과에서 나타난 하위 임차 계층의 극단적으로 낮은 전이 민감도 현상을 현행 정책 금융 자체의 일차원적인 실패로만 해석하는 것은 경계해야 한다. 정책 금융은 본질적으로 대출금 회수와 재정 건전성을 전제로 하므로 담보력과 상환 능력이 현저히 떨어지는 취약 계층에게 전세 보증금이라는 거액의 여신을 무리하게 창출하여 직접적인 주거 상향을 견인하려는 시도는 필연적으

로 대규모 부실 위험과 기관 건전성 악화라는 역효과를 초래할 수밖에 없다. 따라서 상환 능력이 부족한 무보증 월세 가구 등 사다리 최하단 계층에게 대출 중심의 금융을 무리하게 투입하는 등 주택 시장의 분절을 타개하기 위해 금융이라는 단일 도구의 기능 확대에만 의존하는 것은 한계가 있다.

오히려 후생 경제학적 관점에서 볼 때 디딤돌 대출 등 획일적인 용자 중심의 지원책이나 주택 담보 대출에 대한 이자 지원은 거시 경제에 심각한 자원 배분의 비효율성을 초래할 우려가 있다. 미시 경제학의 한계 원리에 따르면 이러한 용자 및 조세 지원은 가구가 직면하는 주택 소비의 사적 한계 비용을 사회적 한계 비용보다 인위적으로 낮추는 결과를 낳는다. 이는 가구의 한계 편익을 초과하는 수준의 주택 과다 소비를 유발하며, 궁극적으로 공장, 기계, 학교 등 사회 전체의 생산적 자본으로 향해야 할 자원을 주택 시장으로 매몰시키는 사회적 후생 손실을 발생시킨다.

2) 용자 중심 지원의 역진성과 부채 주도형 투기 조장 경계

더욱 심각한 것은 이러한 정책이 주택 가격 상승이라는 맹목적 기대 심리와 결합될 때 나타나는 위험성이다. 주택 담보 대출 이자 상환액 소득 공제와 같은 조세 지원 제도는 누진세 체계 하에서 오히려 고소득층에게 더 큰 절세 혜택을 주며, 현행 여신 심사 구조 역시 소득과 자산이 높을수록 대출 한도가 크게 산출되므로 이러한 정부 지원의 실질적 편익은 상위 계층에 집중되는 역진성을 띤다. 즉, 무주택 서민의 내 집 마련 지원이라는 명분하에 막대한 유동성을 공급하는 현행 정책은 하위 계층의 주거 안정보다는 자본을 갖춘 계층에게 과도한 부채를 쥐여주며 자산 시장으로 유인하는 과도한 유동성 공급은 자산 불균형을 심화시키는 거시 건전성 리스크로 전이될 우려가 존재한다. 상환 능력이 부족한 계층에게 자력 도약이 불가능한 수준의 자산 취득을 강권하며 무리하게 대출을 쥐여주는 것은 가계 부실이라는 또 다른 거시 건전성 리스크를 잉태할 뿐이다.

3) 분절 타개를 위한 패러다임 이원화: 주거 복지와 초정밀 핀셋 금융

따라서 본 연구에서 확인된 사다리 하단의 극단적인 병목 현상을 더 이상 주택 금융 대출 한도 확대나 금리 인하의 틀 안에서만 해결하려 해서는 안 된다. 결국 분절된 주거 사다리를 복원하기 위해서는 정책의 패러다임을 이원화해야 한다.

먼저 사다리 최하단 계층의 병목을 해소하기 위해서는 용자 중심의 정책 금융 패러다임을 고집할 것이 아니라 직접적인 주거비 바우처 지급이나 저소득층 공공 임대 주택 공급 확대 등 타 주거 보조 프로그램과의 구조적 연계성을 강화하는 방향으로 정책의 초점을 맞춰야 한다. 정책 금융은 이러한

복지 시스템 내에서 가계가 일정 수준 이상의 재무적 펀더멘털을 갖춘 이후에 다음 단계로 도약하도록 돕는 보완적 견인 장치로 작동해야 한다.

동시에 정책 금융은 주거 사다리의 허리를 강화하기 위해 자산 보유 가구 중에서도 정부의 개입이 절실히 요구되는 대상을 엄밀하게 걸러내는 초정밀 선별 지원 방식을 강화하여 투기적 유동성 공급을 차단해야 한다. 상대적으로 상환 능력이 충분한 상위 계층에 대한 획일적인 DSR 완화 등은 오히려 시장의 유동성을 자극해 주택 가격 불평등을 가중시키므로, 본 연구의 지표를 조기 경보 시스템으로 활용하여 이동 병목이 발생하기 시작하는 특정 임계 계층에 정책 재원을 핀셋처럼 집중해야 한다. 주택 금융이 닿지 않는 본질적 사각 지대를 주거 복지로 메우는 명확한 역할 분담과 강력한 연계만이 고착화된 주거 이동의 분절을 타개하는 현실적인 해법이 될 것이다.

VI. 결론

본 연구는 마르코프 연쇄 메커니즘을 미시적 주택 정책 평가에 선도적으로 도입하여, 특정 계층이 겪는 '주거 정체'와 구조적 단절을 동태적으로 계량화했다는 점에서 학술적·정책적 의의를 지닌다. 특히 '구조적 전이 민감도'라는 편미분 지표를 통해 현행 주택금융 정책의 온기가 소득 및 자산 계층별로 얼마나 비대칭적으로 작용하는지를 확인하였다.

본 연구는 마르코프 연쇄를 주택 시장의 동태적 이동성 평가에 선도적으로 도입했으나, 모형의 추상화 및 데이터 측면에서 다음과 같은 한계를 내포하고 있다.

첫째, 본 연구는 거시적 자산 축적 경로를 모델링하기 위해 상태 공간을 점유 형태 중심으로 단순화하였다. 그러나 현실의 주거 이동은 주택 유형이나 입지에 따른 질적 이질성이 크다. 따라서 최근 고가 월세 시장의 등장이나 아파트/비아파트의 주택 유형에 따라 월세 → 전세 → 자가가 반드시 상향 이동을 의미하지 않을 수 있는데, 이를 온전히 통제하지 못한 것은 본 연구의 한계이므로 향후 다차원적 후속 연구가 필요하다.

둘째, 보증부 월세의 경우 보증금 비중에 따른 자산 규모의 이질성이 매우 큼에도 불구하고 이를 단일 유형으로 처리한 과도한 단순화의 한계가 존재한다.

셋째, 표본 규모에 따른 통계적 엄밀성 문제이다. 연간 5,000가구의 양질의 연도별 횡단면 데이터를 활용하였으나, 실제 최하위 무보증 월세 구간 등에서 상향 이동을 경험한 가구의 절대적인 표본 수는 한정적이다. 이는 그만큼 사다리 최하단의 상향 이동 자체가 현실에서 희소하다는 것을 방증하는 결과이기도 하나 동시에 해당 구간의 전이 확률 및 전이 민감도 추정치에 통계적 우연성(noise)이 크게 개입되었을 가능성을 배제할 수 없다. 따라서 해당 수치들을 절대적인 물리적 지표로 해석하기보

다는 계층 간 이동성의 상대적 단절을 보여주는 징후적 지표로 제한적인 해석이 요구된다.

넷째, 장기 에르고딕 시뮬레이션의 정태적(static) 해석 제약이다. 본 연구가 제시한 장기 미래 예상 분포(전세 비중 7.3% 수렴)는 최근 5년의 경직된 단기 전이 확률이 영구히 고정된다는 강한 수학적 가정하에 산출되었다. 따라서 이를 실제 미래를 확정하는 예측치(forecast)라기보다는, 구조적 충격이 방지될 경우 주택 시장이 직면할 수 있는 이분법적 양극화 위험을 경고하는 일종의 '이론적 스트레스 테스트(stress test)' 지표로 신중하게 접근해야 한다.

그러나 이러한 수학적 제약들은 모형의 결함이라기보다 분석을 위한 통제 조건에 가깝다. 오히려 본 연구는 이처럼 가장 정태적이고 보수적인 조건 하에서도 하위 임차 계층의 구조적 전이 민감도가 0에 수렴(-0.004)한다는 사실을 입증함으로써, 현실 주택 시장의 구조적 병목과 정책적 사각지대의 심각성이 수리적 결과치보다 훨씬 더 치명적일 수 있음을 역설적으로 증명하고 있다. 향후 거시경제 변수를 결합한 동태적 패널 벡터자기회귀(panel VAR) 모형이나 가구의 과거 이력을 추적하는 고차(higher-order) 마르코프 모형을 활용한 후속 연구가 이어진다면 분석의 정교함을 한층 높일 수 있을 것이다.

참고문헌

- 국토연구원. (2024). 주택 임대차시장 안정성을 높이고 세입자 보호와 지원강화 필요. 세종: 국토연구원.
- 김성찬, 성주한. (2025). DSR(Debt Service Ratio) 규제가 임차권등기명령 건수에 미치는 영향에 관한 연구. *주택금융연구*, 9(2), 103-129.
- 민병철. (2023). 보증금 미반환 위험의 추정: 강동전세 아파트를 중심으로. *주택금융리서치*, 28, 24-35.
- 박정재, 이승훈, 탁은명. (2024). 부동산 가격 변동과 경제불평등과의 관계: 자산과 소득 불평등을 중심으로. *한국경제학보*, 31(1), 59-90.
- 방송희. (2017). 계층별 주거TREND 분석을 통한 주거안정지원 강화방안. *주택금융연구*, 1, 88-141.
- 서재원, 진장익. (2023). 자산가치의 공간적 불평등이 행복수준과 계층이동 가능성 인식에 미치는 영향. *국토계획*, 58(4), 187-202.
- 선수봉, 최민섭(2023), 정부의 주택 정책이 수도권 아파트 전세의 월세화에 미치는 영향 연구: 임대인을 중심으로. *주거환경*, 21(2), 60, 121-135.
- 심승규, 오수현. (2022). 생애 첫 주택 거주 기간에 관한 연구. *주택금융연구*, 6(1), 29-53.
- 안선영, 이상엽. (2025). 전세보증금 미반환에 영향을 미치는 주요요인 연구: 수도권지역 전세보증 사고를 중심으로. *주택금융연구*, 9(2), 47-68.
- 오민준. (2022). 자산 불평등도 결정요인 분석 연구: 자산 불평등도 기여도 및 영향 분석(WP 22-05). 세종: 국토연구원.
- 윤성진, 조윤지, 최경아, 이다은, 조정희. (2026). 비아파트 소유 기피 현상과 주거정책 과제: 청년 주거안정을 중심으로. *국토정책Brief*, 1052, 1-8.
- 이상영, 서정렬. (2023). 전세 사기의 원인 분석과 대안 탐색. *동향과 전망*, 118, 242-272.
- 이소영, 이창무. (2019). 부모의 경제력이 청년층 임차가구의 주거사다리 이동에 미치는 영향에 관한 연구. *부동산학연구*, 25(4), 85-102.
- 이지언, 이정란. (2026). 전세보증금 반환보증 제도 변화가 주택 임대차 시장 구조에 미친 영향: 서울시 임대차 실거래 자료를 중심으로. *주택도시금융연구*, 11(1), 61-79.
- 이형찬, 송하승, 오민준, 김지혜, 최수. (2020). 사회통합을 위한 부동산자산의 불평등 완화방안 연구(기본 20-20). 세종: 국토연구원.

- 이후빈. (2024). 주택 소유자 사회의 모순과 한국의 자가 점유율 정체: 주택의 불로소득화와 전세제도의 금융화. *공간과 사회*, 34(2), 293-338.
- 이후빈, 홍다솜. (2022). 한국 주거자본주의의 성격1 : 주택거주와 투자의 동시 금융화. *공간과 사회*, 32(3), 9-61.
- 정대성. (2022). 아파트 매매가격, 전세가격 및 월세가격 간의 수익률 전이효과. *주택금융연구*, 4(2), 123-142.
- 주택금융연구원 (2026). 실태조사 통계. Retrieved from <https://houstat.hf.go.kr/research/portal/compose/surveyResearchPage.do>
- 최종호, 최창원, 윤종원. (2024). 주택시장의 투기적 성격과 가계 양극화: 포트폴리오 이론의 관점에서. *주택금융연구*, 8(2), 39-69.
- 한국개발연구원. (2024). *비아파트 임대시장 변화 및 주택 인허가 데이터 분석 자료* 세종: 한국개발연구원.
- 홍정훈, 임재만. (2024). 20·30세대 '영끌'에 관한 실증분석. *부동산분석*, 10(1), 63-78.
- Grinstead, C. M., & Snell, J. L. (1997). *Introduction to probability*. Providence, RI: American Mathematical Society.
- Kemeny, J. G., & Snell, J. L. (1960). *Finite Markov chains*. Princeton, NJ: Van Nostrand.
- Meyer, C. D. (2000). *Matrix analysis and applied linear algebra*. Philadelphia, PA: SIAM.
- Shorrocks, A. F. (1978). The measurement of mobility. *Econometrica*, 46(5), 1013-1024.
- Stewart, W. J. (2009). *Probability, Markov chains, queues, and simulation*. Princeton, NJ: Princeton University Press.

(논문 접수일: 2026.04.05. 수정논문 접수일: 2026.05.12. 논문 채택일: 2026.06.05.)

The Paradox of Housing Finance Policy and the Broken Housing Ladder: Application of a Markov Chain Model and Structural Transition Sensitivity Analysis

Kyongchae Jung*, Jung Eun Park**

Abstract

This study aimed to analyze structural shifts and mobility asymmetry in the South Korean housing ladder using a Markov chain model. The results indicated significant market restructuring (Frobenius norm: 0.5877) and a widening mobility gap; the upper tier's transition to ownership accelerated (mean first passage time [MFPT]: 1.96), whereas the lower tier faced severe bottlenecks (MFPT: 5.02). Sensitivity analysis revealed that the marginal impact of the standard stepping-stone pathway (the Jeonse system) surged ninefold (0.026 to 0.232), emerging as a critical hurdle, while addressing issues concerning the lowest tier remains a policy blind spot. Simulations project a decline in the Jeonse system of 7.3%, giving warning of a dichotomous ownership-versus-rent market. We propose a two-track policy: direct welfare provision for the lowest tier and targeted bridging financing for the middle tier to restore the broken housing ladder.

Keywords : Housing Ladder, Markov Chain, Asymmetric Mobility, Mean First Passage Time (MFPT), Structural Transition Sensitivity

* Kyongchae Jung, First author, Research Professor, Department of Urban Planning & Engineering, Hanyang University; Lecturer, Department of Economics, Seoul National University, staffj@snu.ac.kr

** Jung Eun Park, Corresponding author, Assistant Professor, Graduate School of Urban Studies and Graduate School of Convergence Real Estate, Hanyang University, jep2104@hanyang.ac.kr

© Copyright 2026 Housing Finance Research Institute. 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.

부록

〈부록 표 1〉 연도별 주거 점유 형태 전이 확률 행렬(2021~2025)

구분	2021년	2022년	2023년	2024년	2025년
$O \rightarrow O$	0.935	0.929	0.919	0.917	0.932
$J \rightarrow O$	0.485	0.521	0.523	0.526	0.559
$R_1 \rightarrow O$	0.341	0.168	0.133	0.147	0.106
$R_2 \rightarrow O$	0.223	0.216	0.296	0.269	0.324
$R_3 \rightarrow O$	0.425	0.226	0.294	0.317	0.303
$O \rightarrow J$	0.029	0.034	0.036	0.044	0.031
$J \rightarrow J$	0.404	0.441	0.449	0.438	0.404
$R_1 \rightarrow J$	0.158	0.196	0.114	0.128	0.145
$R_2 \rightarrow J$	0.090	0.080	0.021	0.059	0.083
$R_3 \rightarrow J$	0.255	0.274	0.354	0.340	0.328
$O \rightarrow R_1$	0.022	0.030	0.024	0.024	0.023
$J \rightarrow R_1$	0.098	0.025	0.023	0.028	0.028
$R_1 \rightarrow R_1$	0.477	0.619	0.732	0.701	0.732
$R_2 \rightarrow R_1$	0.340	0.484	0.333	0.329	0.284
$R_3 \rightarrow R_1$	0.230	0.372	0.225	0.243	0.236
$O \rightarrow R_2$	0.004	0.003	0.003	0.002	0.002
$J \rightarrow R_2$	0.005	0.003	0.001	0.001	0.002
$R_1 \rightarrow R_2$	0.017	0.013	0.011	0.006	0.008
$R_2 \rightarrow R_2$	0.310	0.153	0.316	0.330	0.186
$R_3 \rightarrow R_2$	0.006	0.012	0.011	0.013	0.022
$O \rightarrow R_3$	0.010	0.004	0.017	0.012	0.012
$J \rightarrow R_3$	0.008	0.009	0.004	0.007	0.007
$R_1 \rightarrow R_3$	0.006	0.004	0.010	0.018	0.010
$R_2 \rightarrow R_3$	0.037	0.067	0.034	0.013	0.124
$R_3 \rightarrow R_3$	0.084	0.117	0.115	0.087	0.111

〈부록 표 2〉 연도별 주거 점유 형태 전이 가구 수 추정(5,000가구 표본 환산)

구분	2021년	2022년	2023년	2024년	2025년
$O \rightarrow O$	1,817	1,847	1,804	1,708	1,673
$J \rightarrow O$	698	879	1,033	999	1,071
$R_1 \rightarrow O$	322	125	75	89	73
$R_2 \rightarrow O$	16	22	20	11	18
$R_3 \rightarrow O$	257	110	127	187	164
$O \rightarrow J$	57	68	71	83	55
$J \rightarrow J$	582	744	886	833	774
$R_1 \rightarrow J$	150	147	64	78	100
$R_2 \rightarrow J$	6	7	2	2	5
$R_3 \rightarrow J$	154	133	153	200	178
$O \rightarrow R_1$	43	61	47	45	41
$J \rightarrow R_1$	141	43	45	53	53
$R_1 \rightarrow R_1$	451	462	411	426	506
$R_2 \rightarrow R_1$	31	42	21	12	16
$R_3 \rightarrow R_1$	138	181	97	144	128
$O \rightarrow R_2$	7	5	7	4	4
$J \rightarrow R_2$	7	6	3	2	3
$R_1 \rightarrow R_2$	16	10	6	4	6
$R_2 \rightarrow R_2$	15	13	22	13	10
$R_3 \rightarrow R_2$	3	6	5	8	12
$O \rightarrow R_3$	19	8	34	23	21
$J \rightarrow R_3$	11	16	9	13	14
$R_1 \rightarrow R_3$	6	3	6	11	7
$R_2 \rightarrow R_3$	2	5	2	1	7
$R_3 \rightarrow R_3$	51	57	50	51	61

위 표에서 확인되듯, 5,000가구 중 대다수는 전년도와 동일한 점유 형태를 유지(고착화)하고 있으며, 단기간 내 상하향 이동을 달성한 가구의 비율은 상대적으로 적다. 특히 하위 구간에서 자가로 직행하는 표본 규모는 제한적이므로, 도출된 단기 전이 확률에 불가피하게 표본의 우연성(noise)이 일부 개입될 수 있음을 고려하여 분석 결과를 해석한다.

〈부록 표 3〉 거주지역별 현황

연도	구분	자가(O)	전세(L)	보증금 있는 월세(R_1)	보증금 없는 월세 등(R_2)	무상(R_3)
2021년	서울	53.5	30.1	15.1	0.6	0.7
	경기	62.6	24.2	12.0	0.6	0.7
	광역시	62.7	15.6	18.8	0.7	2.2
	기타	66.6	10.9	17.8	1.7	2.9
2022년	서울	54.8	29.4	13.3	1.2	1.3
	경기	59.8	25.5	13.2	0.4	1.0
	광역시	60.5	21.1	16.2	0.3	1.9
	기타	61.8	15.5	18.9	1.2	2.6
2023년	서울	48.2	34.0	16.6	0.7	0.5
	경기	59.7	30.7	9.0	0.0	0.6
	광역시	60.8	22.5	15.0	0.8	1.0
	기타	70.4	12.4	10.6	1.6	4.9
2024년	서울	47.7	33.8	16.5	0.8	1.2
	경기	54.9	32.8	11.2	0.0	1.0
	광역시	64.3	19.8	13.7	0.2	1.9
	기타	67.6	14.2	13.7	1.3	3.2
2025년	서울	47.1	35.1	16.0	0.7	1.1
	경기	55.0	26.9	16.5	0.3	1.3
	광역시	66.0	16.5	14.7	0.4	2.3
	기타	66.9	15.4	13.1	1.2	3.4

〈부록 표 4〉 가구주 연령별 현황

연도	구분	자개(O)	전세(L)	보증금 있는 월세(R_1)	보증금 없는 월세 등(R_2)	무상(R_3)
2021년	30대 이하	22.1	37.1	35.6	1.7	3.5
	40대	60.6	22.9	14.6	0.5	1.4
	50대	74.5	14.5	9.0	0.9	1.2
	60대 이상	82.6	7.0	8.3	0.8	1.3
2022년	30대 이하	17.9	43.5	32.6	1.5	4.5
	40대	58.2	28.0	12.0	0.6	1.2
	50대	71.7	16.7	10.2	0.4	1.0
	60대 이상	81.2	7.0	10.3	0.7	0.8
2023년	30대 이하	23.3	46.2	27.6	0.7	2.3
	40대	56.3	32.4	9.9	0.6	0.9
	50대	74.5	15.5	8.1	0.7	1.2
	60대 이상	79.4	9.4	6.9	1.2	3.0
2024년	30대 이하	23.0	43.5	28.9	1.2	3.4
	40대	53.9	32.8	11.7	0.3	1.2
	50대	73.2	16.8	8.4	0.4	1.1
	60대 이상	79.0	10.6	7.9	0.5	2.0
2025년	30대 이하	19.2	41.8	34.6	0.7	3.7
	40대	54.4	32.3	11.5	0.6	1.2
	50대	72.8	16.4	9.1	0.4	1.3
	60대 이상	78.9	9.5	8.4	0.9	2.2

〈부록 표 5〉 가구주 연령별 현황

연도	구분	자개(O)	전세(J)	보증금 있는 월세(R_1)	보증금 없는 월세 등(R_2)	무상(R_3)
2021년	1분위	58.4	8.9	28.4	1.7	2.5
	2분위	52.8	13.6	29.7	1.6	2.3
	3분위	58.4	27.7	11.5	0.8	1.5
	4분위	68.9	21.3	7.5	0.5	1.8
	5분위	72.1	22.7	4.0	0.4	0.8
2022년	1분위	49.3	12.1	33.8	2.0	2.9
	2분위	53.6	21.4	23.0	1.0	1.0
	3분위	56.2	29.0	12.3	0.5	1.9
	4분위	64.7	26.0	7.1	0.3	1.9
	5분위	74.7	21.6	2.4	0.1	1.2
2023년	1분위	56.5	11.8	23.4	2.5	5.8
	2분위	52.1	22.7	22.4	0.8	1.9
	3분위	62.9	27.1	9.0	0.1	0.9
	4분위	62.0	32.1	4.6	0.5	0.9
	5분위	72.3	24.0	2.8	0.3	0.6
2024년	1분위	56.0	12.1	26.5	1.6	3.9
	2분위	48.5	23.6	24.3	0.8	2.8
	3분위	60.1	28.0	10.0	0.4	1.4
	4분위	66.9	27.0	4.9	0.0	1.2
	5분위	68.0	28.8	2.4	0.3	0.5
2025년	1분위	54.0	12.4	27.7	1.4	4.4
	2분위	47.8	21.5	28.1	1.0	1.6
	3분위	60.1	26.3	11.2	0.4	1.9
	4분위	65.3	26.7	6.0	0.4	1.6
	5분위	72.7	24.2	1.5	0.2	1.3

부록

- 학술지 편집 및 간행 기준
 - 논문 작성지침
 - 심사 판정기준
 - 논문 투고신청서
- 저작권 양도 동의서
- 연구윤리 서약서
- 연구윤리 자가점검표
- 논문 심사요청서
- 논문 심사의견서
- 심사위원 서약서
- 수정사항 답변서
- 심사결과 통보서
- 특수관계인 논문 공저 신고서
- 편집위원회

학술지 편집 및 간행 기준

제 정 2021. 1. 4.
개 정(1) 2021. 5.26.
개 정(2) 2022. 6.15.
전부개정 2023. 4.13.
개 정(3) 2025. 9.26.

제1장 총칙

제1조(목적)

이 기준은 한국주택금융공사 주택금융연구원이 정기적으로 간행하는 학술지「주택금융연구」의 편집과 간행에 필요한 사항을 규정함을 목적으로 한다.

제2조(적용범위)

주택금융연구원(이하 “연구원”이라 한다)이 간행하는 학술지「주택금융연구」(이하 “학술지”라 한다)는 이 기준에서 정하는 바에 따른다. 다만, 이 기준에서 정하지 않은 사항은 학술지 편집위원회(이하 “편집위원회”라 한다)가 따로 정할 수 있다.

제2장 학술지 편집위원회

제3조(편집위원회 목적)

학술지 편집과 간행 등에 관한 세부 사항을 주관하기 위해 편집위원회를 둔다.

제4조(구성 및 선임 등)

- ① 편집위원회는 편집위원장 1인, 편집위원 15인 이내, 간사 1인으로 구성한다. 다만, 필요시 간사를 대신하여 3인 이내의 간사단을 둘 수 있다.
- ② 편집위원은 다음 각 호의 요건을 충족하는 자 중에서 주택금융연구원장(이하 “연구원장”이라 한다)이 주택금융 관련 분야의 전공을 고려하여 균형 있게 위촉한다.
 1. 박사학위 소지자
 2. 최근 3년간 한국연구재단 등재지 또는 등재 후보지에 주저자 또는 교신저자로서 연구논문 3편 이상 게재

- ③ 편집위원장은 편집위원 중에서 학문적 업적과 학계 인지도 등을 종합적으로 고려하여 연구원장이 선임한다.
- ④ 연구원장은 편집위원회의 원활한 운영을 위하여 연구원 소속 직원 중에서 간사를 선임할 수 있다.

제5조(임기)

편집위원장 및 편집위원의 임기는 2년으로 하되 연임할 수 있다. (개정 2025.9.26.)

제6조(역할)

- ① 편집위원장은 논문 심사위원 위촉, 논문 심사의뢰, 학술지 편집과 간행에 따르는 제반업무와 행정 사항 총괄 등의 직무를 수행한다.
- ② 편집위원회의 직무는 다음 각 호와 같다.
 1. 논문 심사위원 추천
 2. 심사결과에 대한 판정
 3. 이의신청에 대한 심의
 4. 투고자와 심사위원 간 의견 중재
 5. 최종 게재 논문의 선정
 6. 게재 순서의 결정
 7. 투고자의 논문 사용에 대한 동의
 8. 학술지 간행 횟수 및 간행일 변경
 9. 그 밖에 학술지 간행에 필요한 사항
- ③ 간사는 편집위원회 회의 운영, 회의록 관리, 투고 논문 접수와 편집, 심사 등 학술지 간행에 따른 제반 실무를 수행한다.
- ④ 편집위원회는 학술지의 운영 및 질적 수준 향상을 위하여 노력해야 하며, 편집위원회 활동과 관련하여 취득한 투고자 및 심사자의 개인정보를 보호할 의무가 있다.

제7조(회의)

- ① 회의는 정기회의와 임시회의로 구분한다.
- ② 정기회의는 연 2회 개최하되, 필요시 임시회의를 개최할 수 있다. 임시회의는 편집위원장 또는 편집위원 3분의 1 이상의 요구로 소집한다.
- ③ 회의는 대면방식(구성원의 전부 또는 일부가 직접 회의에 출석하지 아니하고 모든 구성원이 음성을

동시에 송수신하는 원격통신수단에 의하여 의결에 참가하는 것을 포함)을 원칙으로 하되, 필요한 경우 서면회의로 진행할 수 있다. (개정 2025.9.26.)

제8조(회의 의결)

- ① 회의는 재적위원 과반수의 출석으로 개최하고, 출석위원 과반수의 찬성으로 의결한다. 다만, 가부 동수인 경우 편집위원장이 결정한다.
- ② 대면회의 및 서면회의 출석기준은 다음 각 호와 같다. (개정 2025.9.26.)
 1. 대면방식 회의의 경우 회의 출석자 및 원격통신수단 참가자 수 (개정 2025.9.26.)
 2. 서면회의의 경우 안전에 대한 회신서면(전자우편 회신을 포함한다) 수

제9조(회의비 및 자문료)

회의에 참석한 편집위원에게 예산 범위 내에서 회의비 및 자문료를 지급할 수 있다.

제3장 논문 심사위원

제10조(위촉)

편집위원장은 논문의 심사를 위해 전공분야가 적절한 심사위원을 편집위원회로부터 추천받아 논문 한 편당 3명을 위촉한다. 다만, 위촉된 심사위원은 심사할 논문이 심사위원의 전공분야가 아닐 경우, 편집위원회에게 통보하여 자격에 적합한 심사자를 다시 선정할 수 있도록 협력해야한다.

제11조(자격)

심사위원은 다음 각 호의 어느 하나에 해당하는 자여야 한다.

1. 주택·금융 관련분야 박사학위 소지자
2. 주택·금융 관련분야 석사학위 소지자로서 5년 이상의 실무 또는 연구경력을 갖춘 자
3. 대학 또는 전문대학 전임교원 이상의 직위를 가진 자
4. 그 밖에 위 각 호와 동등한 자격을 갖춘 것으로 편집위원장이 인정한 자

제12조(심사위원의 제척 등)

- ① 심사위원은 자신의 지위를 이용하여 개인이나 가족의 사적이익이나 경제적 이익을 추구하는 다양한 종류의 행위나 행동(이하 “이해상충”이라 한다)을 피해야 하며, 다음 각 호의 어느 하나에 해당하는 경우에는 당해 투고 논문의 심사위원이 될 수 없다. (개정 2025.9.26.)

1. 투고자 또는 공동저자와 민법 제779조에 따른 가족 관계가 있는 자 (신설 2025.9.26.)
 2. 투고자 또는 공동저자와 사제관계에 있거나 공동으로 용역, 자문, 연구 등을 수행하거나 하였던 자 (신설 2025.9.26.)
 3. 심사위원 본인이 투고한 논문(심사위원이 공동저자인 경우를 포함한다) 및 본인이 속한 기관의 구성원이 투고한 논문 등 기타 심사의 공정성을 해할 우려가 있다고 판단되는 자 (신설 2025.9.26.)
- ② 심사위원이 제1항 각 호의 어느 하나에 해당함을 인지한 경우에는 편집위원장에게 회피 신청을 해야 하며, 이로 인해 제외된 인원수만큼은 편집위원장이 심사위원을 추가 선임할 수 있다. (개정 2025.9.26.)
- ③ 편집위원장은 심사위원이 제1항 각 호의 어느 하나에 해당됨에도 불구하고 회피신청을 하지 않아 업무의 공정성이 침해될 우려가 있다고 판단하는 경우 해당 심사위원을 해촉할 수 있다. (신설 2025.9.26.)
- [제목개정 2025.9.26.]

제13조(심사료)

공정하고 적법하게 심사를 마친 심사위원에게 예산 범위 내에서 심사료를 지급할 수 있다. 다만, 공사 내부 심사위원에게는 심사료를 지급하지 않는다.

제4장 논문 투고 및 심사

제14조(논문투고 및 접수)

- ① 학술지에 게재 가능한 논문은 국내외 주택금융 관련 분야의 이론 및 통계, 객관적 실증 분석 연구논문으로서 독창성이 인정되는 정규논문이면서 「연구윤리규정」을 준수하여 작성된 논문만 학술지에 게재할 수 있다. (개정 2025.9.26.)
- ② 투고하고자 하는 논문은 투고일 현재 다른 간행물에 발표되지 않은 논문이어야 하며, 다른 간행물에 투고한 논문은 그 간행물에 게재하지 않기로 확정된 후에 투고할 수 있다.
- ③ 투고 논문은 공사의 논문투고시스템(이하 “시스템”이라 한다) 등 온라인으로 접수하는 것을 원칙으로 하며, 접수 시 저자의 인적사항을 삭제한 논문 원본과 다음 각 호의 서류를 함께 제출해야 한다.
 1. [별지 제1호]의 “논문 투고 신청서”
 2. [별지 제2호]의 “저작권 양도 동의서”
 3. [별지 제3호]의 “연구윤리 서약서”

4. [별지 제4호]의 “연구윤리 자가점검표”

5. KCI 문헌 유사도 검사 결과확인서

- ④ 논문은 홈페이지에 공지를 한 날부터 접수마감일까지 접수하며, 접수일은 시스템 등 온라인으로 논문이 접수된 날짜로 한다.
- ⑤ 간사는 투고 논문이 접수되면 2주 이내에 투고자에게 접수사실을 통보해야 한다.

제15조(논문심사)

- ① 논문 심사는 적격심사와 본심사의 순서로 하며, 적격심사를 통과해야 본심사를 할 수 있다.
- ② 편집위원회는 적격심사 단계에서 [별표 1] “논문 작성지침” 및 「연구윤리규정」 준수 여부를 확인하고, 이를 준수하지 않은 논문에 대해서는 편집위원회에서 수정을 요구할 수 있으며 이를 따르지 않는 경우 시스템을 통해 투고자에게 “부적격” 사실을 통보한다. (개정 2025.9.26.)
- ③ 편집위원회는 [별지 제5호]의 “논문 심사요청서”를 통하여 심사위원에게 본심사를 의뢰하고, 투고자의 인적사항이 삭제된 심사대상 논문을 심사위원에게 송부한다. 논문 심사위원은 특별한 사유가 없으면 의뢰받은 날부터 12일 이내에 심사를 하되, 재심은 7일 이내에 한다.
- ④ 심사위원은 기일에 맞춰 [별지 제6호]의 “논문 심사 의견서”와 [별지 제7호]의 “심사위원 서약서”를 편집위원회에 제출한다.
- ⑤ 심사위원은 접수 논문에 대해 ‘게재 가능’, ‘수정 후 게재’, ‘수정 후 재심’, ‘게재 불가’ 중 하나를 판정한다. 이 경우 심사위원 3인의 심사결과에 따른 판정 기준은 [별표 2] “심사 판정기준”에 따른다.
- ⑥ 심사기준은 다음 각 호와 같다.
 - 1. 연구목적의 명확성과 연구주제의 중요성
 - 2. 내용전개의 논리성과 일관성
 - 3. 연구방법의 적절성
 - 4. 연구자료의 신뢰성
 - 5. 분석 또는 평가의 타당성과 객관성
 - 6. 연구결과의 학술적 기여도
- ⑦ 논문심사 판정결과에 따라 ‘수정 후 게재’, ‘수정 후 재심’ 판정을 받은 투고자는 다음 각 호의 절차에 따라 논문을 수정하여 편집위원회에 제출해야 한다.
 - 1. 편집위원회는 ‘수정 후 게재’, ‘수정 후 재심’ 판정을 받은 논문의 투고자에게 심사위원 인적 사항을 삭제한 [별지 제6호] “논문 심사 의견서”를 첨부하여 수정을 요청
 - 2. 논문수정기간은 7일 이내를 원칙으로 하되, 수정의 난이도 및 간행일정 등을 고려해 기간의 가감 가능

3. 수정논문 제출 시 투고자는 심사의견에 대해 [별지 제8호]의 “수정사항 답변서”를 함께 제출
- ⑧ ‘수정 후 재심’ 판정에 따라 수정된 논문의 재심 판정은 다음 각 호에 따른다.
1. 논문수정을 요청받은 투고자가 논문 수정에 응하지 않거나, 기간 내 미제출 시 ‘게재 불가’로 처리
 2. 재심은 초심에서 ‘수정 후 재심’으로 평가한 심사위원이 실시하며, 재심 심사위원은 ‘게재 가능’, ‘수정 후 게재’, ‘게재 불가’ 중에서만 판정
 3. 재심은 1회에 한함
- ⑨ 편집위원회는 [별지 제6호]의 “논문 심사 의견서”를 첨부하여 [별지 제9호]의 “심사결과 통보서”를 투고자에게 전달해야 한다.
- ⑩ 최종 판정은 ‘게재 가능/수정 후 게재’ 이상의 판정을 받은 논문을 대상으로 편집위원회의 의결을 거쳐 결정한다.

제16조(이의제기)

- ① 투고자는 심사위원의 심사의견, 수정 요청내용, 평가결과에 대해 이의가 있는 경우 상당한 논거나 실증적 사례를 들어 편집위원회에 서면으로 이의를 제기할 수 있다.
- ② 편집위원장은 편집위원회의 심의 등을 통해 이의신청에 대해 최종 판단하며, 투고자는 최종 판정에 대해 다시 이의를 제기할 수 없다.

제17조(정보보호의무)

- ① 심사의뢰 시에는 투고자의 인적사항을 밝히지 않으며, 투고자에게도 심사위원의 신상 일체를 공개하지 않는다.
- ② 심사위원은 논문의 심사과정에서 획득한 정보를 누설하거나 남용해서는 안 된다.

제5장 논문 게재 및 간행

제18조(논문 게재)

- ① 제15조제9항에 따라 학술지에 논문을 게재하는 경우, 편집위원회는 투고자에게 게재 사실을 통보해야 한다.
- ② 게재하기로 결정된 논문을 해당 호에 게재하는 것이 불가능한 경우에는 편집위원회의 의결로 다른 호에 게재할 수 있으며 편집위원장은 그 사유를 논문 투고자에게 통보해야 한다.
- ③ 투고자는 논문이 게재될 경우 논문 게재를 위한 저작물의 복제·배포, DB구축 및 인터넷 등 정보통신망을 통한 서비스를 위한 저작물의 디지털화 및 전송 등에 동의해야 한다.

제19조(원고료)

게재된 논문에 대하여 투고자에게 예산범위 내에서 소정의 원고료를 지급할 수 있다. 이 경우, 저자가 복수일 때에는 [별지 제1호]의 “논문 투고 신청서”상 기재된 기여율대로 각 저자에게 나누어 지급한다.

제20조(간행 및 배포)

- ① 학술지는 연 2회 간행하고 간행 일자 6월 30일, 12월 31일로 하며, 간행 횟수 또는 간행 일자를 변경하고자 할 경우에는 편집위원회에서 결정한다.
- ② 학술지 간행일로부터 2주 이내에 게재 논문 및 논문별 디지털 식별 정보(Digital Object Identifier) 등을 한국학술지인용색인(KCI : Korea Citation Index)에 등록해야 한다.
- ③ 기타 학술지의 간행 부수, 배포 대상기관 및 배포 방법은 편집위원회에서 결정한다.

제6장 저작권 · 연구윤리

제21조(저작권의 귀속)

- ① 논문의 투고자는 게재된 논문 관련 판권을 포함한 저작권을 [별지 제2호]의 “저작권 양도 동의서” 양식에 의해 한국주택금융공사에 양도해야 하며, 학술지에 게재되는 모든 논문 및 원고에 대한 저작권은 한국주택금융공사에 귀속된다.
- ② 투고자가 논문을 타 학술지 게재 등 개인의 용도로 사용하고자 할 경우 편집위원회의 사전 동의를 얻어야 한다.

제22조(연구윤리)

- ① 투고자는 논문투고 시 투고 신청서에 한국학술지인용색인(KCI) 문헌 유사도 검사 결과를 기입하고 KCI 문헌 유사도 검사 결과확인서를 제출해야 한다. 이 경우, KCI 문헌 유사도 검사 결과 유사도가 15% 이상일 경우 편집위원장은 심사를 거부할 수 있다.
- ② 편집위원회는 연구부정행위를 판단할 때, 한국연구재단 지정 연구윤리정보센터 내 전문가위원 또는 한국저작권위원회의 자문을 구할 수 있다.
- ③ 연구부정행위는 「연구윤리규정」에 따라 판단하며, 부정행위에 해당할 경우 다음 각 호의 제재 조치를 취할 수 있다. (개정 2025.9.26.)
 1. 논문 목록에서 해당 논문 삭제
 2. 투고자의 향후 투고 금지(부정행위의 경중에 따라 최소 3년)
 3. 연구원 홈페이지 및 학술지 차기호에 고시 (개정 2025.9.26.)

4. 원고료 지급 거부 또는 환수

- ④ 특수관계인(만19세 미만의 미성년자 또는 4촌 이내의 친족)이 공동저자인 경우, [별지 제10호]의 “특수관계인 논문 공저 신고서”를 제출해야 하고, 특수관계인이 실질적으로 연구에 기여하였다고 편집위원장이 인정한 경우에만 논문 저자 자격이 주어진다.
- ⑤ 「연구윤리규정」에 따라 구성된 연구윤리위원회는 특수관계인이 연구 저자로 부정 등재([별지 제10호]의 “특수관계인 논문 공저 신고서” 미제출 또는 허위 기재 등)되어 해당 논문으로 입시·진학·취업 등에서 이익을 취한 경우 투고자 및 특수관계인의 소속기관 및 관련기관(입시·진학·취업 등 관련 학교, 연구기관, 기업체 등)에 해당 사실을 통보해야 한다. (개정 2025.9.26.)

부칙(전부개정)

제1조(시행일)

이 기준은 2023년 4월 14일부터 시행한다.

제2조(적용례)

개정규정은 시행일 이후 심사하는 투고 논문부터 적용한다.

부칙(1)

이 기준은 2025년 9월 26일부터 시행한다.

[별표 1] 논문 작성지침(제15조 제2항 관련) (개정 2025.9.26.)

논문 작성지침(제15조제2항 관련)

1. 논문작성 양식

- ① 원고는 한글(HWP) 또는 워드프로세서로 작성하며, 원고의 본문 분량은 A4용지 20페이지 내외로 한다.
- ② 원고는 표지, 본문, 참고문헌, 부록, 영문초록, 영문 핵심주제어 순으로 배열한다.
- ③ 표지에는 논문제목, 국문초록(500자 이내), 국문 핵심주제어(5개 내외)를 표기한다. 영문논문은 영문초록(15줄 이내), 영문 핵심주제어(5개 내외)를 표기한다. 저자명 및 소속 등은 별도의 논문 투고신청서에 작성한다.
- ④ 원고용지 편집은 다음에 의한다.

용지여백 (mm)		문단모양		글자모양 및 번호체계		
위쪽	10	좌우여백	0	제목	20pt	
아래쪽	10	간격	줄간격	160%	본문	12pt
왼쪽	20		문단위	0	글꼴	휴먼명조*
오른쪽	20		문단아래	0	도표/그림 제목	12pt
머리말	15	들여쓰기	10pt	각주	10pt	
꼬리말	15	정렬방식	양쪽혼합			
제본	0					

※ 워드(MS Word)에서 영어논문 작성시 Times New Roman 으로 작성

- ⑤ 본문에는 논문 저자명 등 투고자의 신분이 노출되지 않도록 주의하여야 한다.
- ⑥ 본문의 장절구분은 I, 1, 1), (1), ①, 가(영어논문 작성시, a)의 순서를 따르며, 본문 및 각주는 완전한 문장으로 작성한다.
- ⑦ 표와 그림의 제목에는 각각 일련번호(예: <표 1>, <그림 1>)를 매겨 각각 표와 그림의 상단 중앙에 배치하고, 자료 출처는 아랫부분에 배치한다.
- ⑧ 수식의 번호는 우측 정렬하여 일련번호(예: <식 1>)로 표기한다.

2. 본문 주와 참고문헌

- ① 인용·참고한 자료의 출처를 밝히는 참고주(Reference notes)는 본문 중에 괄호 속에 표기한다. 저자가 1명 또는 2명인 경우에는 모두 기입하고, 3명 이상인 경우에는 외 또는 et al.을 붙인다(예:홍길동, 2010; 홍길동-김철수, 2005; 홍길동 외, 2008; Hong, 2005; Hong & Kim, 2019; Hong et al., 2013).
- ② 본문의 내용에 설명을 부연하기 위한 내용주(Content notes)는 해당 부분의 오른쪽 위에 일련번호를 매기고(예 : 부동산업¹⁾은), 페이지 하단에 각주로 처리한다.
- ③ 참고문헌은 본문의 마지막에 장을 달리하여 기록하며 제1저자의 성을 기준으로 국내문헌은 가나다순으로, 외국문헌은 알파벳순으로 배열한다.
- ④ 같은 저자의 여러 문헌은 연도순으로 배치하되, 동일연도에 간행된 문헌이 둘 이상일 경우에는 문헌 제목의 가나다 혹은 알파벳 순서에 따라 간행연도에 a,b,c를 붙여 구분한다.
- ⑤ 참고문헌은 국문문헌, 영문문헌, 신문기사, 온라인 자료의 순서로 기재한다.
- ⑥ 참고문헌 작성은 다음에 의하며, 다음에 명기되지 않은 참고문헌의 작성은 APA Style (American Psychological Association, 미국심리학회) 6판을 따른다.

- 단행본

저자, 저자. (출판연도). *도서명*. 발행지: 출판사.

Author, A. A., Author, B. B., & Author, C. C. (year). *Title of book*(xx ed., Vol. xx). Location: Publisher.

Author, A. A. (year). Title of work. In A. Editor, & B. Editor (Eds.), *Title of book*(xx ed., Vol. xx, pp. xx-xx). Location: Publisher.

(예시)

강철수. (2013). *금융학개론*. 서울:와우패스.

Glaeser, E. L. (2012). *Urban public finance*. Cambridge, MA: National Bureau of Economic Research.

- 학회지 등 정기간행물

저자, 저자. (출판연도). 논문명. *학술지명*, 권(호), 쪽 수.

Author, A. A., Author, B. B., & Author, C. C. (year). Title of article. *Title of Periodical, volume*(issue), page range.

(예시)

방두완, 권혁신. (2020). 아파트가격 지역 간 전이효과 분석. *주택금융연구*, 4, 5-32.

Colenbrander, S., & Barau, A. (2019). Planning and financing urban development in the context of the climate crisis. *International Journal of Urban Sustainable Development*, 11(3), 237-244.

- 정부, 기업 등의 보고서

저자. (출판연도). *보고서명*(Report No. XXX). 지역: 발행처.

Author, A. A. (year). *Title of work* (Report No. XXX). Location: Publisher.

(예시)

서울특별시. (2008). *2008년 서울시통계연보*. 서울: 서울특별시.

World Bank. (1985). *Uganda - Report on urban finance and management* (Report No. 5322-UG). Washington, DC: World Bank.

- 학위논문

연구자. (수여연도). *논문명*(학위명). 수여기관명, 소재지.

Author, A. A. (Year of publication). *Title of thesis* (Type of thesis). Institution, Location.

(예시)

홍길동. (2010). *OO연구: OOO를 중심으로*(박사학위논문). OO대학교, 서울.

Jackson, J. B. (2003). *OO Study*(Ph.D. Dissertation). Yale University, New Haven, CT.

- 신문기사

기자명. (발행일자). 기사제목. *신문명*. Retrieved from URL

(예시)

홍길동. (2018년 12월 31일). 현정부의 부동산정책. *xx일보*. Retrieved from URL

- 온라인자료

저자명. (발행일자). 전자문서명. Retrieved from URL

Author, A. A. (Date of publication). Title of website. Retrieved from URL

(예시)

홍길동. (2018년 11월 20일). 도시재생뉴딜 시작. Retrieved from <http://blog.naver.com/ko/newdeal>

3. 제출방법 등

- ① 논문투고시스템 등 온라인으로 제출한다.
- ② 투고 논문은 다른 간행물에 발표되지 아니한 것이어야 하며, 또한 다른 간행물에 중복 심사 의뢰하여서는 안 된다. 단 다른 간행물에 심사의뢰하였으나 탈락이 확정된 것은 예외로 한다.
- ③ 논문 작성에 사용되는 언어는 한글을 원칙으로 한다.
- ④ 논문 투고시 심사비 및 게재료는 발생하지 않으며 게재 확정 시 소정의 원고료를 공사가 투고자에게 지급한다.
- ⑤ 투고된 원고는 반환하지 않으며, 저자(들)는 본 논문이 게재될 경우 논문 게재를 위한 저작물의 복제·배포, DB구축 및 인터넷 등 정보통신망을 통한 서비스를 위한 저작물의 디지털화 및 전송 등을 허락한다.
- ⑥ 위 목적을 위하여 필요한 범위 내에서의 편집·형식상의 변경을 허락한다.
- ⑦ 투고자는 한국주택금융공사의 내규인「연구윤리규정」제6장(연구윤리) 및「학술지 편집 및 간행 기준」을 준수한다.
- ⑧ 최종 판정은 심사결과를 종합하여 편집위원회에서 결정하므로, 본심사까지 통과하여 ‘게재 가능/수정 후 게재’ 이상의 등급을 부여받더라도 수정내용 확인 후 편집위원회에서 게재여부를 최종 판정한다.

[별표 2] 심사 판정기준(제15조제5항 관련)

심사 판정기준(제15조제5항 관련)

심사결과			판정
제 1 심사위원	제 2 심사위원	제 3 심사위원	
○	○	○	계재 가능/수정 후 계재
○	○	△	계재 가능/수정 후 계재
○	○	×	계재 가능/수정 후 계재
○	△	△	수정 후 재심
△	△	△	수정 후 재심
○	△	×	수정 후 재심
△	△	×	수정 후 재심
○	×	×	계재 불가
△	×	×	계재 불가
×	×	×	계재 불가

○ : 계재 가능/수정 후 계재

△ : 수정 후 재심

× : 계재 불가

[별지 1] 논문 투고 신청서 (개정 2025.9.26.)

논문 투고신청서

논문제목		한글:		
		영문:		
제 1 저 자	이름	한글: 한자: 영문:	교신저자 여부	(여, 부) %
	소속 및 직위			
	전화		이메일	
제 2 저 자	이름	한글: 한자: 영문:	교신저자 여부	(여, 부) %
	소속 및 직위			
	전화		이메일	
제 3 저 자	이름	한글: 한자: 영문:	교신저자 여부	(여, 부) %
	소속 및 직위			
	전화		이메일	
KCI문헌유사도 검사서비스결과		검사일자: 20 평균유사율: % 최고유사율: %	※ KCI문헌 유사도 검사 서비스결과 유사도가 15% 이상 일 경우 심사 거부를 결정할 수 있음	

본 논문의 저자(들)는 본 논문이 한국주택금융공사 학술지 「주택금융연구」에 게재되기를 희망합니다. 또한, 이 논문은 「주택금융연구」의 「연구윤리규정」 및 「학술지 편집 및 간행 기준」을 준수하여 작성하였음을 확인합니다.

제1저자 서명 (인)
제2저자 서명 (인)
제3저자 서명 (인)

년 월 일

한국주택금융공사 주택금융연구원 학술지편집위원회 귀중

저작권 양도 동의서

■ 논문제목(Title of Manuscript)

국문 :

영문 :

본 논문의 저자(들)는 한국주택금융공사 학술지「주택금융연구」에 논문 게재를 신청하면서 다음 사항에 대하여 확인 및 동의합니다.

1. 저자(들)는 본 논문이 학문적으로 독창성을 가지며, 다른 논문의 저작권을 침해하지 않았음을 확인합니다.
2. 저자(들)는 본 논문에 실질적인 공헌을 하였으며 논문의 내용에 대하여 책임을 함께 합니다.
3. 저자(들)는 본 논문이 「주택금융연구」에 게재된 후 연구부정행위가 있는 논문임이 확인된 경우 귀 공사의 조치(원고료반납, 투고 금지 등)에 따르겠습니다.

〈연구부정행위〉

- 1) 위조 : 존재하지 않는 자료나 연구결과를 만들어내고 이를 기록하거나 보고하는 행위
- 2) 변조 : 연구자료 또는 과정을 조작하거나 자료나 연구결과를 변경하나 생략하여 연구 기록이 진실에 부합하지 않게 하는 행위
- 3) 표절 : 정당한 권한 없이 타인의 아이디어, 과정, 결과 또는 기록을 도용하는 행위
- 4) 부당한 논문저자 표시 : 연구내용 또는 결과에 대하여 과학적·기술적 공헌 또는 기여를 한 사람에게 정당한 이유 없이 논문 저자 자격을 부여하지 않거나, 과학적·기술적 공헌 또는 기여를 하지 않은 사람에게 감사의 표시 또는 예우 등을 이유로 논문저자 자격을 부여하는 행위
- 5) 이중투고 : 동일 또는 실질적으로 유사한 논문을 2개 이상의 국내외 학술지에 투고하는 행위
- 6) 중복게재 : 투고자 본인이 이미 발표(게재)한 저작물과 동일 또는 실질적으로 유사한 저작물을 선행 저작물의 출간사실을 밝히지 않은 채 다시 투고(게재)하는 행위

4. 저자(들)는 「주택금융연구」의 간행인이 저자(들)나 「주택금융연구」간행인의 허락 없이 타인에 의해 이루어지는 저작권 침해에 대해서 이의를 제기할 권리가 있음을 동의합니다.
5. 저자(들)는 본 논문이 「주택금융연구」에 게재될 경우 본 논문에 따른 권리, 이익, 저작권 및 디지털 저작권에 대한 모든 권한행사(저작물의 복제·배포, DB구축 및 인터넷 등 정보통신망을 통한 서비스를 위한 저작물의 디지털화 및 전송 등)를 한국주택금융공사에 양도함을 확인합니다.
6. 위의 목적을 위하여 필요한 범위 내에서의 편집·형식상의 변경을 허락합니다.

저자	성명	소속	연락처	지필서명
제1저자				
제2저자				
제3저자				

년 월 일

한국주택금융공사 귀중

연구윤리 서약서

연구자는 학술지「주택금융연구」에 투고하는 논문에 대하여 아래의 윤리규정을 준수하여 작성하였음을 서약합니다.

1. 저자(들)는 학술지「주택금융연구」의 「연구윤리규정」 및 「학술지 편집 및 간행 기준」을 인지하고, 논문 투고 및 게재 시 준수한다.
2. 다음과 같은 부정행위가 있는 논문을 투고하지 않는다.
 - 1) 위조 : 존재하지 않는 자료나 연구결과를 만들어내고 이를 기록하거나 보고하는 행위
 - 2) 변조 : 연구자료 또는 과정을 조작하거나 자료나 연구결과를 변경하나 생략하여 연구 기록이 진실에 부합하지 않게 하는 행위
 - 3) 표절 : 정당한 권한 없이 타인의 아이디어, 과정, 결과 또는 기록을 도용하는 행위
 - 4) 부당한 논문저자 표시 : 연구내용 또는 결과에 대하여 과학적·기술적 공헌 또는 기여를 한 사람에게 정당한 이유 없이 논문 저자 자격을 부여하지 않거나, 과학적·기술적 공헌 또는 기여를 하지 않은 사람에게 감사의 표시 또는 예우 등을 이유로 논문저자 자격을 부여하는 행위
3. 투고 및 게재 논문은 원저에 한하며, 아래의 항목을 준수한다.
 - 1) 저자가 투고한 논문은 내용이나 기법상 다른 학술지에 발표하지 않은 독창성을 가져야 한다.
 - 2) 저자는 다른 학술지에 투고하여 심사 중인 논문과 동일한 내용을 담은 논문을 투고할 수 없다.
 - 3) 저자는 본인이 이미 발표(게재)한 저작물과 동일 또는 실질적으로 유사한 저작물을 선행 저작물의 출간사실을 밝히지 않은 채 다시 투고(게재)할 수 없다.
 - 4) 모든 자료는 출처를 명확히 밝혀야 한다.
 - 5) 연구의 과정과 결과에 참여하지 않은 사람은 공저자로 포함하지 않는다.
 - 6) 특수관계인(미성년자(만 19세 미만) 또는 친족(배우자, 자녀 등 4촌 이내))인 공동저자의 경우 실질적으로 연구에 기여하였다고 편집위원장이 인정한 경우에만 논문 저자 자격이 주어진다.
4. 연구 부정행위 발생의 개연성 방지를 위하여 이해상충* 소지를 인지한 경우에는 해당 사실을 구술 서면·전화·전자우편 등 가능한 모든 방법으로 편집위원회에 제보하겠으며, 구체적인 사실을 적시한다.

*한 개인이 자신의 지위를 이용하여 개인이나 가족의 사적 이익이나 경제적 이익을 추구하는 다양한 종류의 행위나 행동

	성명	소속/직위	연락처	자필서명
제1저자				
제2저자				
제3저자				

년 월 일

한국주택금융공사 주택금융연구원 학술지편집위원회 귀중

[별지 4] 연구윤리 자가점검표 (개정 2025.9.26.)

연구윤리 자가점검표

분류	점검 내용	점검 결과
전반적 사항	• 참여 연구자 전원이 연구 수행의 연구윤리 규정을 인지하였는가?	예/아니오
	• 참여 연구자 전원에게 연구윤리 준수에 대한 확답을 받았는가?	예/아니오
위조	• 면담이나 설문조사를 실행하지 않고 가상으로 구성하여 연구 결과를 허위로 제시한 경우가 없는가?	예/아니오
	• 설문조사, 실험, 관찰 등에서 나타나지 않은 데이터를 실재하는 것처럼 제시한 경우가 없는가?	예/아니오
	• 실험, 조사 등을 통해 얻은 자료의 통계학적인 유효성을 얻기 위해 허구의 자료를 추가한 경우가 없는가?	예/아니오
	• 연구계획서에 합치한다는 점을 제시하기 위해 연구 기록을 허위로 삽입한 경우가 없는가?	예/아니오
변조	• 연구 자료를 의도적으로 변경하여 결과의 상이함을 수정한 경우가 없는가?	예/아니오
	• 연구 자료의 통계분석 결과 명확하지 않은 부분을 고의 또는 중대한 과실로 그릇되게 설명한 경우가 없는가?	예/아니오
	• 통계학적 근거없이 연구 자료 일부를 선택적으로 생략, 삭제, 추가, 은폐한 경우가 없는가?	예/아니오
	• 연구 자료를 과장, 축소 또는 변형함으로써 왜곡된 연구 결과를 도출한 경우가 없는가?	예/아니오
표절	• 타인의 저작물이나 독창적인 아이디어를 활용하면서 출처를 표기하지 않은 경우가 없는가?	예/아니오
	• 타인의 저작물을 번역하여 활용하면서 출처를 표기하지 않은 경우가 없는가?	예/아니오
	• 재인용 표시를 하지 않고 직접 원문을 고찰한 것처럼 1차 문헌(원문)에 대한 출처를 표시한 경우가 없는가?	예/아니오
	• 출처표시를 정확하게 했으나 인용된 양 또는 질이 타인 저작물의 연구 독자성을 훼손할 정도로 적절한 범위를 넘는 경우가 없는가?(주중관계: 타인의 저작물이 주(主), 자신의 저작물이 종(從))	예/아니오
	• 타인의 저작물을 활용하면서 대상 전체가 아니라 일부에만 출처표시를 한 경우가 없는가?	예/아니오
	• 타인의 저작물을 그대로 인용하였음에도 불구하고 참조하였다고 표시한 경우가 없는가?	예/아니오
이해상충	• 연구의 부정행위 발생의 개연성 방지를 위하여 이해상충 소지를 인지한 경우 공사 또는 편집위원회에 보고하여야 함을 이해하였는가?	예/아니오
부당한 논문 저자표시	• 저자로서 정당한 자격을 갖춘 사람에게 저자 자격을 부여하지 않은 경우가 없는가?	예/아니오
	• 저자로서 정당한 자격을 갖추지 못한 사람에게 저자 자격을 부여한 경우가 없는가?	예/아니오
	• 연구에 기여하지 않은 특수관계인(만 19세 미만의 미성년자 또는 4촌 이내의 친족)이 공동저자에 포함되지는 않았는가?	예/아니오
중복게재	• 자신의 이전 저작물을 활용하면서 출처표시를 하지 않은 경우가 없는가?	예/아니오
	• 자신의 이전 저작물을 활용하면서 일부에만 또는 부정확하게 출처를 표시한 경우가 없는가?	예/아니오
	• 자신의 이전 저작물을 활용하면서 출처를 표시하였으나, 양적으로나 질적으로 적절한 범위를 넘어선 경우가 없는가?	예/아니오
IRB	• 주저자는 연구대상자로부터 얻은 자료 및 개인정보를 어떻게 수집·활용하고 처리하는가에 대해 설명하고 이에 대한 동의를 받았는가?	해당없음/ 예/아니오

년 월 일

제1저자 서명 (인)
제2저자 서명 (인)
제3저자 서명 (인)

한국주택금융공사 주택금융연구원 학술지편집위원회 귀중

[별지 5] 논문 심사요청서

논문 심사요청서

수 신 : 000 귀하

1. 000님의 건승하심을 기원합니다. 주택금융연구「00호 간행과 관련하여 다음의 논문이 투고되었습니다. “학술지 편집 및 간행기준”에 따라 심사하시어 심사의견을 회신해 주시기 바랍니다.

— 다 음 —

<input type="checkbox"/> 논문제목	
<input type="checkbox"/> 접수일자	년 월 일
<input type="checkbox"/> 심사의견서 제출기한	년 월 일

2. 귀하의 심사결과는 주택금융연구원의 “학술지 편집 및 간행기준”에 따라 처리됨을 알려드립니다.

- 붙 임 : 1. 투고논문 1부
2. 논문 심사의견서 양식 1부
3. 논문심사위원 서약서 1부.
4. 학술지 편집 및 간행기준 1부. 끝.

20

한국주택금융공사 주택금융연구원 학술지편집위원회

[별지 6] 논문 심사의견서

논문 심사의견서

제 목	국문:				
	영문:				
① 심사기준					
심 사 항 목	평 가				
	매우 우수(5)	우수(4)	보통(3)	미흡(2)	매우 미흡(1)
연구목적의 명확성과 연구주제의 중요성					
내용전개 논리성과 일관성					
연구방법의 적절성					
연구자료의 신뢰성					
분석 또는 평가의 타당성과 객관성					
연구결과의 학술적 기여도					
② 심사의견 및 수정보완 요구사항					
(논고에 대한 심사의견, 수정요구사항 등을 구체적으로 기술, 필요시 별지 사용)					
③ 심사결과*					
게재 가능		수정 후 게재		수정 후 재심	
* 재심은 '게재 가능', '수정 후 게재', '게재 불가'만 선택 가능					
④ 심사위원					
소속		심사 일자	년 월 일	성명	(인)
전화					

심사위원 서약서

본인은 한국주택금융공사 주택금융연구원 학술지「주택금융연구」논문을 심사하면서 다음의 사항을 준수할 것을 서약합니다.

— 다 음 —

1. 본인은 한국주택금융공사“학술지 편집 및 간행기준”에 따라 공정하고 성실하게 심사하겠습니다.
2. 심사할 논문이 본인의 전공분야가 아닐 경우, 편집위원회에게 통보하고 편집위원회의 의견을 따르겠습니다.
3. 논문심사 중 저자의 인적사항을 알게 되고, 사적인 친분관계*가 있는 등 이해충돌이 발생할 것으로 판단되는 경우 편집위원회에게 지체없이 그 사실을 통보하고 편집위원회의 의견을 따르겠습니다.
4. 논문의 심사과정에서 획득한 정보를 타인에게 누설하거나 남용하지 않겠습니다.
5. 본인은 각종 심사자료*를 생성형 AI 도구에 입력(업로드)하지 않겠습니다.

* 친분관계의 범위 : 5년 이내 공동연구 이력이 있는 자 중 하나 이상에 해당되는 경우, 현재 동일 소속, 배우자, 자녀 등 4촌 이내 친족

* 심사자료 : 투고 논문, 논문 심사 의견서, 수정사항 답변서 등 논문 심사와 관련된 제반 자료

년 월 일

심사위원: (인)

한국주택금융공사 주택금융연구원 학술지편집위원회 귀중

[별지 8] 수정사항 답변서

수정사항 답변서

[제1심사자]

수정요청사항	수정보완 내용 및 답변

[제2심사자]

수정요청사항	수정보완 내용 및 답변

[제3심사자]

수정요청사항	수정보완 내용 및 답변

년 월 일

제1저자 서명 (인)

제2저자 서명 (인)

제3저자 서명 (인)

한국주택금융공사 주택금융연구원 학술지편집위원회 귀중

심사결과 통보서

수 신 : 000 귀하

학술지「주택금융연구」에 귀하의 연구논문을 투고해 주셔서 감사드립니다. 투고된 논문은 본 연구원 학술지편집위원회에서 선정한 심사위원들이 엄정한 심사기준에 의해 면밀히 검토 하였습니다. 논문의 심사결과는 다음과 같으며, 최종판정은 본 연구원 학술지편집위원회에 서 확정되오니 양지하시기 바랍니다.

— 다 음 —

1. 심사결과 : 0000
2. 별 첨 : 심사의견서 1부

한국주택금융공사 주택금융연구원 학술지편집위원회

[별지 10] 특수관계인 논문 공저 신고서

특수관계인 논문 공저 신고서

□ 연구 개요

논문제목	
참여 연구원	※ 아래와 같은 양식으로 기재하여 주시기 바랍니다. - 참여연구원 A (성명/소속기관/부서명/직위) - 참여연구원 B (성명/소속기관/부서명/직위) - 특수관계인 (성명/소속기관/부서명/직위)

□ 특수관계인 유형 (해당하는 모든 란에 √ 표시)

친족(4촌 이내)			미성년자		
배우자	자녀	기타	지인 자녀	R&E 프로그램 참여자	기타

□ 특수관계인이 상기 논문의 저자로 포함되어야 하는 사유

※ 특수관계인 등이 상기 논문 성과 창출에 어떤 기여를 했는지 위주로 기술

□ 연구추진 단계별로 기여한 사항 기술(요약)

구분	연구기획 (연구 설계, 작업의 개념 정립 등)	연구수행 (연구데이터 수집/분석/해석, 기본 원고 작성 등)	원고초안작성 (중요한 내용을 초안하거나 비판적으로 개정)	최종원고확정 (논문 최종 검토 및 승인)
(저자 A)				
(저자 B)				
(특수관계인)				

※ 주의사항: 저자 요건을 갖춘 개인만 기재하시기 바랍니다. 즉 저자 표시 대상 개인과 참여연구원이 일치할 필요가 없습니다.

□ 저자별 기여율 산출 및 확인 내역 <저자 간 합의 사항>

구분	연구 기획	연구 수행	원고초안 작성	최종원고 확정	전체 기여도	확인 서명
(저자 A)	()%	()%	()%	()%	()%	
(저자 B)	()%	()%	()%	()%	()%	
(특수관계인)	()%	()%	()%	()%	()%	
계	30%(*)	40%(*)	20%(*)	10%(*)	100%(*)	

* 항목별(연구추진 단계별) 기여율 가중치는 연구과제의 특성에 따라 조정 가능함.

□ 연구윤리 확인내용 (예시)

항목	내용 (예시)
저자 임의변경과 부정확한 추가 방지	저자 일동은 저자 기여도 등을 감안하여 제1저자는 저자 A, 공저자는 저자 B, 저자 C(특수관계인)임을 확인합니다.
표절 및 이중게재 방지	표와 그림은 저자 B에 의해 주로 작성되었습니다. 표절 검사 프로그램 결과는 저자 C에 의해 작성되었습니다. 어떤 문장, 문단도 무단으로 표절한 부분이 없음을 확인합니다.
번역을 이용한 표절의 방지	논문 제목, 키워드, 주요 변수, 가설, 연구모형의 독창성을 보증합니다. 한글 초록과 영문 Abstract 모두 연구결과를 바탕으로 작성했습니다. 영문 번역 이전의 한글 원고로도 표절 검사 이상 없음을 확인합니다.
실증분석 연구윤리 위반 방지	실증분석에 사용된 원본 데이터는 잘 보관하고 있습니다. 표와 그림에 사용된 자료의 출처를 잘 표기하였습니다.
짜깁기 표절 방지	본 보고서와 가장 유사한 선행연구는 ()입니다. 아이디어 제안은 저자 A가 하였고, 이를 저자 C가 보완하였습니다.

년 월 일

위 내용이 모두 사실임을 확약합니다.

저자 구분	제1저자	제2저자	제3저자
저자 성명	(저자 A)	(저자 B)	(특수관계인)
확인 서명			

한국주택금융공사 주택금융연구원 학술지편집위원회 귀중

편집위원회

구분	이름	소속
편집위원장	유선종	건국대학교 부동산학과 교수
편집위원회 (가나다순)	김선주	경기대학교 부동산자산관리학과 교수
	김승희	강원대학교 부동산학과 교수
	김정섭	울산과학기술원 도시환경공학과 교수
	김종희	전북대학교 경제학부 교수
	김준형	명지대학교 부동산학과 교수
	남영우	나사렛대학교 국제금융부동산학과 교수
	서정렬	영산대학교 부동산학과 교수
	유승동	상명대학교 경제금융학부 교수
	진창하	한양대학교 경제학부 교수
	최선한	이화여자대학교 공과대학 교수
	최향미	충남대학교 경영학부 교수
	한상범	경기대학교 경제학부 교수
	편집간사	박종규
강영신		한국주택금융공사 주택금융연구원 연구원

주택금융연구

10권 1호

등록일자 | 2017년 5월 25일(부산남. 사00005)

발행일 | 2026년 6월 30일

발행인 | 김경환

편집인 | 박광길

발행처 | 한국주택금융공사 주택금융연구원(48400)

부산시 남구 문현금융로40(문현동)

부산국제금융센터

TEL : 051) 663-8178

편집·인쇄 | 거목문화사/거목인포 02-2277-3324
