

시계열모형을 활용한 준공 후 미분양주택 예측

Forecasting Unsold Houses After Completion Using Time Series Models

강 태 영 (Kang, Tae-Young)*

이 상 군 (Lee, Sang-Gun)**

< Abstract >

This study, based on a time-series model, quantitatively predicts the future trend of houses that remain unsold after completion of construction, which is a serious risk in the housing market. A complex combination of failure to forecast demand due to the pre-sale contract system, macroeconomic and financial market changes, and supply strategy failures contributes to houses remaining unsold after construction completion.

In this context, this study explains a time lag in demand and supply and a structure of accumulated inventory based on the theoretical framework of the DW four-quadrant model, empirically analyzing them based on a time series model. Considering monthly data from January 2011 to March 2025, this study sets the number of unsold houses after construction completion in the private sector as the dependent variable to predict. It then compares the predictive power of the ARIMA, ARIMAX, SARIMA, SARIMAX models

This study identifies the models with the ACF and the PACF to set candidate parameters, estimates the ones with the best parameters based on the AIC through grid search, and compares the prediction performance with the MAPE and the RMSE to determine the final model. The results of the empirical analysis suggest that the ARIMAX model with the exogenous variables performed the best, and the residuals passed a white noise test.

After predicting unsold houses after construction completion for the subsequent years, based on the final ARIMAX model, this study discovered a continued uptrend in the number of the unsold houses without a clear resolution in the short term, which further highlights the importance of preemptive responses and structural remedies.

Keyword : Unsold Houses After Completion, Time Series Analysis, Forecasting Model Comparison, ARIMAX, Unsold Housing Trend Forecast

I. 서론

1. 연구의 배경 및 목적

악성 미분양으로 일컫는 준공 후 미분양주택은 2025년 3월 말 기준, 25,117호로 집계되어, 2013년 11월 이후 약 11년 만에 최대치를 기록했다. 준공 후

미분양주택은 공사가 완료되고 사용검사 이후에도 분양이 되지 않아 시행사 또는 건설사가 보유하고 있는 주택으로, 시장의 비효율성과 수요공급의 괴리를 가장 명확하게 보여주는 지표이며 주택경기 침체의 실질적인 위험을 반영한다. 이는 단순한 미분양과는 달리, 공사가 완료되었음에도 불구하고 실수요자나 투자자의 선택을 받지 못한 주택으로, PF(Project Financing) 실패, 지역경제 침체 등과 직결된다.

* 서강대학교 일반대학원 부동산학협동과정 박사과정, kang891121@hanmail.net, 주저자

** 본 학회 정회원, 서강대학교 경영대학 교수, slee1028@sogang.ac.kr, 교신저자

통상 부동산 PF는 브릿지론과 본PF로 나누어지는데, 브릿지론은 부지 매입이나 사업계획승인 등 인허가 절차가 진행 중인 사업 초반에 일으키고, 착공 시점에는 본PF를 통해 브릿지론을 상환하는 구조이다. 이후 입주자모집을 통해 수분양자로부터 받은 분양수입금으로 본PF 대출금을 상환하고 공사비 등 대금도 정산하며 사업장을 청산하게 된다. 그런데 준공 후 미분양이 발생하게 되면 이와 같은 프로세스에 문제가 발생하게 된다. 특히 PF를 기반으로 한 사업장의 미분양 누적은 시행사의 자금 압박과 시공사의 연대보증으로 인한 PF 우발채무 가능성을 증가시킴으로 건설업체의 연쇄 부도를 초래한다(이형, 2008). 즉 시행주체의 자금 경색 → 금융비용 부담 증가 → 연체 및 조기상환 요구 → 신규 착공 중단으로 이어지며 이는 준공 후 미분양이 단지 주택시장에만 국한된 현상이 아니라, 건설사, 대출금융기관 등의 연쇄적 부실로 이어질 수 있다는 점이다. 실제로 최근 몇 년간 미분양 누적이 장기화되면서 중소형 건설사를 중심으로 자금난이 심화되었고, PF 대출의 조기상환 요구와 유동성 경색이 겹치면서 다수 건설사의 법정관리 신청 등 구조적 부실이 다수 관측되고 있다. 건설업체 부도는 분양계약자를 비롯한 여러 이해관계자들에게 경제적 손실을 발생시킬 뿐 아니라 금융기관들의 동반 부실화를 초래하게 된다.(이성근·전광섭, 2009)

이와 같은 현상은 2011년 저축은행 사태에서도 유사하게 관측된 바 있다. 당시 부동산 PF 대출이 과도하게 누적된 가운데 준공 후 미분양 물량이 적체되면서 저축은행의 유동성 위기와 도산, 중소형 건설사 연쇄 부도, 지역경제 침체로 이어졌다. 이러한 위기 재현 가능성에 대응하기 위해 정부는 2025년 3월 준공 후 미분양주택 매입 사업을 발표하며 지방 소재 미분양 물량 3천 호를 선별 매입하겠다고 밝혔다.(국토교통부, 2025) 이는 정책적으로 선제적인 대응이지만, 전체 미분양 재고대비 약 15%에 불과한 소규모 개입으로, 영향을 미치기에는 매우 제한적인 조치이다. 즉, 미분양 리스크가 주택시장·건설사·금융권 전반으로 확산되는 구조적 성격을 고려할 때, 보다 예측 기반의 접근이 필요하다.

그간의 선행연구는 대부분 미분양주택 수를 주요 변수로 설정하여 금리, 경기, 물가 등 거시경제 요인과의 상관관계 또는 영향요인을 분석하여 미분양의 원인을 도출하거나 정책적 해법을 모색하는 데 초점을 두었

고, 일부 연구에서는 준공 시점에 따라 준공 전과 준공 후 미분양을 구분하여 영향을 비교하였으나, 준공 후 미분양주택 수를 독립된 변수로 설정하고, 이를 외생 변수와 함께 예측모형에 적용한 실증 연구는 전무한 상태이다.

이에 본 연구는 준공 후 미분양 문제에 대해 보다 구조적으로 진단하고 예측하는 것을 목적으로 한다. 특히, 준공 후 미분양이 건설사 부도 및 금융기관 리스크 확산의 선행 지표로 작용하고 있음에도 불구하고, 이에 대해 예측 기반의 연구는 제한적이라는 점에서 기존 연구와 방법론적 차별성을 바탕으로 기존 연구의 한계를 보완하고자 한다.

2. 연구의 범위 및 방법

1) 연구의 범위

본 연구는 2011년 1월부터 2025년 3월까지의 월간 데이터를 대상으로, 준공 후 미분양주택 수의 시계열적 변동과 그에 영향을 미치는 주요 거시경제 변수 간의 영향을 고려한 모형을 통해 향후 단기간 발생 추이를 예측하는 것을 본 연구의 시간적 범위로 한다. 2011년 미분양에 따른 중소형 건설사 연쇄 부도 및 저축은행 사태부터, 다시 급증한 최근 상황까지의 중장기적 흐름을 포괄하여 분석함으로써, 미분양 추세를 실증적으로 예측하는 데 분석의 초점을 둔다.

2) 연구의 방법

먼저 국내·외 학술논문, 연구보고서 등을 종합적으로 검토하여 미분양주택이 발생하는 이론적 배경과 발생 요인에 대해 검토하였다. 이를 토대로 기존 연구에서는 주로 미분양주택 수와 거시경제 변수 간의 인과관계나 미분양주택에 미치는 거시경제 요인을 규명하는 데 집중해 왔고 정량적인 예측 기반 연구는 미비한 점을 파악하였다. 또한, 본 연구에서 수행할 시계열분석의 방법론적 배경을 조사하여 본 연구에 적용하는 것이 적합한지 검토하였고 이론적 배경과 관련 문헌을 토대로 외생변수를 선정하였다.

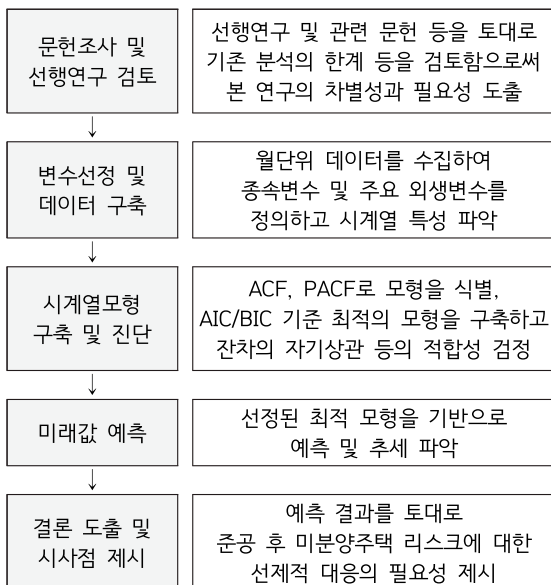
모든 변수는 ADF, PP 단위근 검정을 통해 정상성을 확인하고 필요시 1차 차분하여 안정화된 시계열로 변환하였다. 모델의 파라미터인 시차(order)와 계절성은 ACF/PACF로 모형을 식별하여 후보 파라미터를 설정하였고, Grid Search를 통해 AIC/BIC 기준 최적의

파라미터로 ARIMA, ARIMAX, SARIMA, SARIMAX 시계열모형을 모형을 구축하였다.

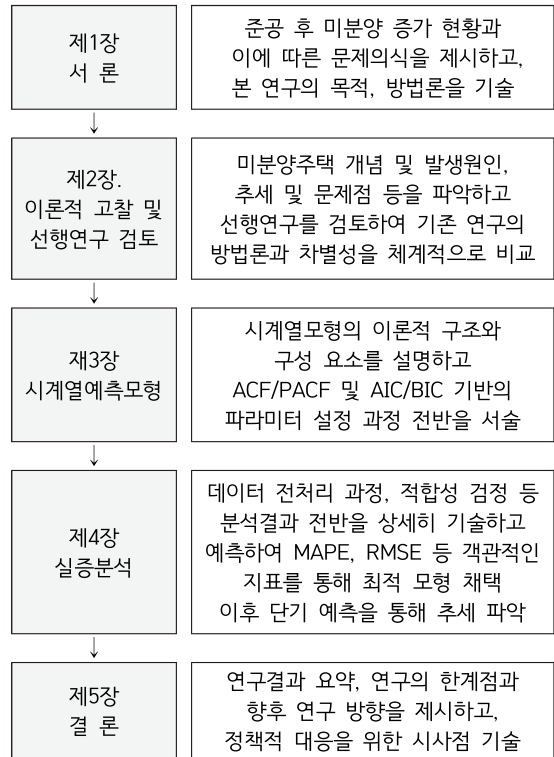
이후 모형이 실제 시계열 흐름을 얼마나 정확히 반영하고 있는지를 확인하고, 예측의 신뢰도를 사전에 검증하기 위해 전체 시계열 데이터를 기반으로 2011년 1월에서 2023년 12월까지 훈련용 데이터(training data)로 모형을 적합하고, 2024년 1월에서 2025년 2월까지 테스트 데이터(test data)로 결과값을 사후적으로 예측 후 실제 관측값과 MAPE(Mean Absolute Percentage Error), RMSE(Rooted Mean Squared Error) 등의 객관적인 지표로 비교하여 최적의 모형을 선정하였다. 최종 모형 채택 이후에는 연구 기간 전체 데이터를 기준으로 Grid Search를 다시 수행하여 파라미터를 재조정하였다. 또한 외생변수가 포함된 시계열모형의 경우 외생변수의 미래값 역시 독립변수로 포함되어야 하기 때문에 외생변수의 미래값은 단순 추세값이 아닌 각각 `auto.arma()` 함수를 활용하여 예측 기간에 해당하는 미래값을 산출하였다. 이를 기반으로 모형에 투입하여 2025년 4월부터 2026년 4월까지의 준공 후 미분양주택을 예측하였다. 예측결과를 바탕으로 단기적 증감 추세를 파악하고, 이를 통해 정책적 대응 방향에 대한 시사점을 도출하였다.

위의 기술한 연구의 범위와 방법을 바탕으로, 본 연구의 구체적인 흐름과 구성은 <표 1>, <표 2>로 나눌 수 있다.

<표 1> 연구의 흐름



<표 2> 연구의 구성



II. 이론적 고찰 및 선행연구 검토

1. 미분양주택 개념 및 발생 원인

미분양주택이란 일반적으로 주택건설사업자가 지자체로부터 사업계획승인을 득하고 「주택공급에 관한 규칙」 제20조, 제21조에 따른 입주자모집 승인 및 공고에 따른 입주자의 계약일이 지난 이후에도 분양계약이 체결되지 아니하여 선착순의 방법으로 공급하는 주택을 말하며, 준공 후 미분양주택은 사용검사 후에도 분양되지 않은 주택을 말한다.(국토교통부, 2021)

특히 준공 후 미분양은 매각되지 않아 잔여로 남아 있는 주택으로, 특히 금융기관과 건설사 입장에서서는 자금 회수가 지연되며 부실위험으로 직결되기 때문에 악성 미분양이라 표현하고 있다. 또한, 주택도시시기금 투입 여부 따라 공공부문, 민간부문 주택으로 구분된다. 공공부문은 주택도시시기금을 지원받아 건설하는 주택으로 주로 국민임대주택, 행복주택 등 정부 또는 공공기관이 건설해서 장기간 임대하거나 분양하는 주택

이고, 민간부문은 순수 민간지분으로 건설하는 주택을 말한다. 본 연구에서는 민간부문 준공 후 미분양주택에 대해 다루고자 한다.

기존 연구에 따르면 미분양주택 발생 요인은 크게 제도적 요인, 거시경제 및 금융시장 변화, 공급 전략 실패의 세 가지로 구분할 수 있다. 첫째, 제도적 요인으로 이중희(1995), 임덕호(2009)는 미분양주택이 국내 주택시장의 특수한 선분양제도의 구조적 한계에서 비롯된다고 보았다. 선분양제도는 「주택공급에 관한 규칙」 제15조, 16조에 따라 대지의 소유권을 확보하고, 해당 기관¹⁾으로부터 분양보증을 받은 경우, 착공과 분양을 할 수 있도록 규정하고 있다. 이를 통해 주택건설 사업자는 공사비 등 각종 사업비를 분양수입금과 연동하여 지급하는 구조로 자금을 조달하기 때문에 초기에 자금조달이 용이할 수 있지만 예상했던 분양률이 저조할 경우 착공 이후에도 미분양 리스크가 해소되지 않고 누적되는 구조적 한계를 나타낸다.

둘째, 거시경제 및 금융시장 변화이다. 준공 후 미분양은 수요공급 간의 불균형 외 거시경제 및 금융시장 악화에도 발생 원인이 있다. 금리 인상, 주택담보대출 규제 강화는 수요자의 구매 여력을 감소시켜 주택 구입을 보류하게 만들고, 경기침체로 인한 인플레이션 등의 원인으로 공사비가 상승하여 분양가격이 높아지는 경우 수요를 저하시키는 요인으로 작용하여 미분양 주택이 발생하게 된다. 김상기·이상효·김재준(2010)은 주택가격 하락이 수요 위축으로 이어져 미분양을 증가시키고, 황관석·김지혜·이운상(2023)은 고금리와 주택시장 침체가 미분양 누적을 심화시켰음을 강조하였다. 즉 수요 측면에서의 충격이 기존 공급 계획에 따라 건설되고 있던 주택 물량을 미분양 상태로 남아있게 만든다고 볼 수 있다.(정재형, 2000)

셋째, 공급전략의 실패는 수요를 정확하게 예측하지 못한 채 무리하게 공급을 진행하는 데서 비롯된다. 특히 실수요자에 대한 분석이 부족하거나, 일시적인 시장 호황에 따른 낙관적인 전망에 따라 수요보다 과도하게 공급되는 경우 미분양으로 이어진다. 즉 수요자의 소득수준과 선호를 고려하지 않은 중대형·고가 주택 위주의 공급, 교통인프라 및 생활편의시설 부족,

수도권과 지방의 상이한 주택수요 등을 고려하지 못한 입지의 선정은 장기적 미분양을 초래하는 주요 요인으로 나타난다.(남희용·김태섭, 2001; 장세웅, 2011)

2. 미분양주택 발생 현황

국토교통부로부터 자료 구축이 시작된 2007년부터 최근까지 통계자료를 살펴보면 미분양주택의 증감은 국내외 경제 여건, 금융시장 변화, 정부의 부동산 정책에 따라 시기별로 큰 차이를 나타냈다. 특히 경기침체, 금융위기, 공급 과잉 등의 요인이 중첩되는 시기에는 미분양이 급증하는 양상을 보였고, 반대로 금리 인하나 규제 완화 등의 정책적 대응이 이루어질 경우 일시적으로 감소세를 나타내기도 하였다.

2007년부터 2009년까지 글로벌 금융위기의 영향으로 주택 수요가 급감하고, 건설 경기는 위축되어 미분양주택이 2008년 역대 최고치를 기록하였다. 정부는 2008년 6월 미분양 대책²⁾을 발표하며 취득세 감면 등 각종 세제혜택과 LTV상향 등 미분양주택 해소를 위한 다양한 지원책을 마련하였다. 이러한 정책으로 미분양은 일부 감소하였지만, 준공 후 미분양은 시장에서 장기간 해소되지 못하며 건설사와 금융기관의 유동성 위기를 심화시켜 PF(Project Financing)에 집중되었던 저축은행의 부실화로 직결되었다.(이형, 2008)

2010년부터 2014년까지는 앞서 발생한 대규모 미분양과 이에 따른 저축은행 사태의 후속 조치가 본격화된 시기이다. 정부는 PF대출 부실채권 정리를 추진하였으며, 이에 따라 다수의 저축은행이 연쇄적으로 영업정지 및 구조조정에 돌입하는 사태가 발생되었다. 더불어 정부는 미분양주택 해소를 위해 보금자리주택³⁾ 150만호 건설계획 발표³⁾와 주택담보대출 규제 강화⁴⁾ 등 다양한 대책을 시행하였고, PF대출 건전성 기준을 강화하여 공급을 억제하였다. 이로 인해 무리한 청약과 분양을 감소시키는 영향이 발생하여 미분양 및 준공 후 미분양 물량이 모두 감소하는 시기로 나타났다.

2015년부터 2019년까지는 경기 회복세와 저금리 기조, 정부의 공급 확대 정책⁵⁾에 힘입어 전국 미분양 주택 수는 지속적으로 감소하였다. 그러나 준공 후 미

1) 주택도시보증공사(HUG) 또는 보증보험을 영위하며 국토교통부장관이 지정하는 보험회사이다.

2) 국토해양부는 2008년 6월 11일 보도자료를 통해 지방 미분양 대책을 발표하였다.

3) 국토해양부는 2009년 5월 11일 보도자료를 통해 보금자리주택 시범지구 추진 계획을 제시하였다.

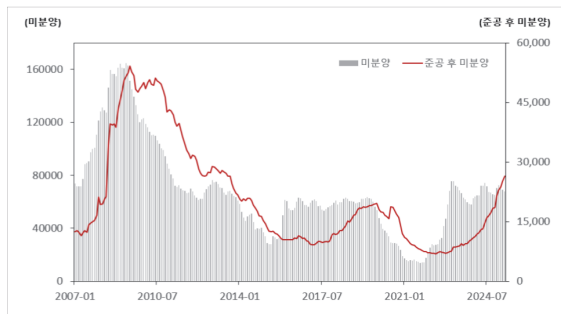
4) 금융위원회는 2011년 6월 29일 가계부채 연착륙 종합대책을 발표하였다.

5) 국토교통부는 2015년 6월 29일 보도자료를 통해 2015년 주택종합계획을 발표하였다.

분양주택 수는 이와는 반대로 꾸준히 증가하는 이중적인 양상을 보였다. 2015년 10,518호였던 준공 후 미분양은 2019년 말에는 18,000호를 넘어섰다. 주택공급이 확대되었음에도 일부 지역에서의 수요 부족과 분양시장 침체 등이 복합적으로 작용하였다.

2020년부터 2024년까지의 시기는 코로나19 팬데믹에 따른 글로벌 인플레이션과 고금리 기조로 2022년 미분양은 급증하였다. 준공 후 미분양주택 역시 2023년부터 급격하게 증가하였는데, 이는 경기침체와 투자 수요 감소가 복합적으로 작용하였다.(황관석·김지혜·이운상, 2023) 이후 2024년 12월 준공 후 미분양주택은 21,480호에 이르며, 2013년 이후 11년 만에 역대 최고치를 기록하였다.

<그림 1> 미분양주택 추이



출처: 「미분양주택현황보고」, 국토교통부 / 단위 : 호

<표 3> 연도별 미분양주택 현황

연도	미분양주택	준공 후 미분양주택	
		민간부문	공공부문
2007	112,254	110,715	1,539
2008	165,599	164,293	1,306
2009	123,297	122,962	335
2010	88,706	88,706	-
2011	69,807	69,807	-
2012	74,835	74,835	-
2013	61,091	61,091	-
2014	40,379	40,379	-
2015	61,512	61,512	-
2016	56,413	56,413	-
2017	57,330	57,330	-
2018	58,838	58,838	-
2019	47,797	47,797	-
2020	19,005	19,005	-
2021	17,710	17,710	-
2022	68,107	68,107	-
2023	62,489	62,489	-
2024	70,173	70,173	-

출처: 「미분양주택현황보고」, 국토교통부 / 단위 : 호

3. 준공 후 미분양주택 증가에 따른 사회적 효과

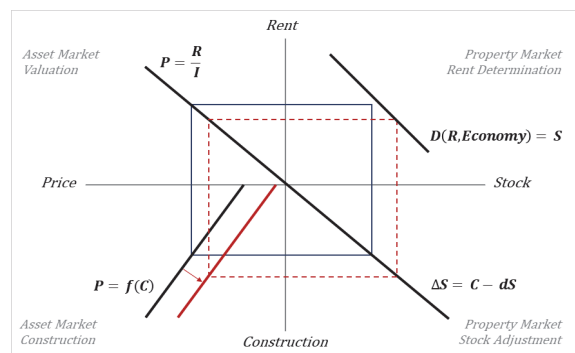
준공 후 미분양주택의 누적은 건설사, 금융권, 주택시장, 그리고 지역경제 전반에 걸쳐 부정적인 파급효과를 야기한다. 장기간 해소되지 못한 미분양은 건설사의 자금 회수 지연과 유동성 악화를 초래하여 부도위험을 높이고(이형, 2008; 이성근·전광섭, 2009), 민간부문의 주택공급 위축으로 국민경제에 부정적 영향을 미칠 수 있다.(황관석·김지혜·이운상, 2023) 특히 PF(Project Financing)에 의존하는 국내 주택공급구조는 미분양 발생 시 금융기관으로 리스크가 전이되는 구조를 형성하여 금융기관에 재정적 부담을 초래한다.(임덕호, 2009; 이형, 2008)

4. 미분양주택의 이론적 근거

미분양주택의 발생 과정을 분석하는데 이론적 근거로서 DiPasquale and Wheaton(1992)의 공간-자산시장이론 바탕으로 이루어진 DiPasquale-Wheaton 4사분면 모형으로 나타낼 수 있다. 해당 모형은 부동산시장의 구조를 자산시장과 공간시장으로 구분하여 각 상호작용을 통해 복합적으로 작동한다고 설명한다.

<그림 2>는 3사분면 건설시장에서 공급곡선이 오른쪽으로 이동한 경우로 공급을 확대했으나 수요예측 실패로 인해 초과공급이 발생한 상황을 나타낸다. 초과공급은 공간시장에서 공실률을 상승 및 임대료를 하락시키며 결국 4사분면에서 미분양이 누적되는 결과로 나타난다. 이는 이후의 신규 건설량을 줄이는 효과를 가져올 수 있고 장기적으로 균형 상태를 이루고자 할 것이다.(이의준·김경민, 2011)

<그림 2> 초과공급에 따른 주택시장 변화



출처: DiPasquale and Wheaton(1992) 재구성

5. 선행연구

주택시장에 대해 다수의 실증분석 연구가 있지만, 준공 후 미분양이 연구의 대상이 된 경우는 드물고 특히 시계열모형에 적용하여 실증적 예측을 한 문헌은 전무한 상태이다. 기존 연구들은 미분양주택이 시장에 미치는 영향을 규명하고, 결정요인을 분석하거나, 정책적 해결방안을 제시하는 데 초점이 맞춰져 있다. 본 연구에서는 기존 선행연구를 미분양 발생 원인과 결정요인 분석, 미분양주택이 주택시장에 미치는 영향, 미분양 해소 및 대응을 위한 정책적 접근의 세 가지로 분류하여 검토함으로써 본 연구의 차별성을 도출하였다.

1) 미분양 발생 원인과 결정요인 분석

강병주(1995)는 대전광역시 주택 미분양 실태를 파악하여 미분양의 원인을 수요에 비해 과다한 주택공급과 부동산 투기억제책으로 인한 수요 감소를 원인으로 보았다. 이를 해소하기 위해 주택금융정책 개선, 택지 공급가격 인하 및 토지사용 승인절차 간소화, 공급자 및 수요자의 조세제도 완화 등을 제시하였다.

정재형(2000)은 미분양률이 높은 대도시 3개 지역을 표본으로 공급자, 수요자 면접조사를 통해 미분양 원인을 파악하였다. 발생 원인으로 선분양제도와 평형 규제, 분양가상한제 등 각종 규제와 주택금융 미비를 제시하였고 이를 해소하기 위해 무이자금융 또는 분양대금 납부조건 완화를 제시하였다.

정창무·김지순(2005)은 주택정책과 거시경제변수가 미분양에 어떤 영향을 미치는지 서울특별시를 중심으로 벡터오차수정모형을 통하여 실증분석하였다. 미분양아파트는 투자자 입장에서 매입 시기를 판단하는 지표로 활용된다는 것을 규명하였다. 또한, 금리나 환율과 같은 거시경제지표는 미분양아파트 재고량을 변동시키며, 정부 정책 역시 미분양 재고량을 변화시킨다는 것을 나타내었다. 다만 미분양이 주택시장에 어떤 영향을 미치는지, 주택착공량과 미분양 관계에 관한 추가 연구의 필요성을 언급하였다.

서범준·이효중·정창무(2010)는 시도별 데이터를 활용하여 패널모형을 구축하고 연립GMM 방식을 활용하여 지역별 미분양 결정요인과 조정속도에 대해 실증 분석하였다. 분석결과 전기 미분양과 신규 분양은 미분양 증가에 영향을 미치고 가격, 대출, 인구 증가는 미분양 감소에 영향을 나타내었다. 특히 조정속도는

수도권이 지방에 비해 2배 이상 빠른 것으로 나타나 지역 간 차이를 나타냈고, 일부 지역에서는 미분양이 장기화됨을 규명하였다.

김상기·이상효·김재준(2010)은 미분양주택과 주택매매가격 및 주택전세가격과의 관계를 벡터오차수정모형을 통해 동태적 관계를 규명하였다. 매매가격의 영향이 매매가격 변동의 영향보다 상대적으로 미치는 영향이 큰 것으로 나타났고, 전세가격과 미분양주택은 서로 미치는 영향이 미미한 것으로 나타났다.

박재룡·유정석(2010)은 미분양 주택과 주택시장의 관계성을 분석하기 위해 거시경제 변수를 독립변수로 최소자승법(OLS)을 적용하였다. 분석결과 가격과 대출금리가 상승할수록 미분양주택이 늘어났고, 실질소득이 늘어날수록 구매력이 높아져 미분양주택은 감소하는 것으로 나타났다. 또한, 미분양주택의 적체 원인을 수요억제정책, 건설업체의 과다한 공급, 수요자의 구매력 저하에서 비롯된다고 보았다.

장세웅(2011)은 벡터오차수정모형을 통해 주택시장에서의 미분양 발생 요인을 체계적으로 분석하였다. 분석결과 주택가격 하락, 공사비 상승, 대출금 감소는 미분양을 증가시키는 요인으로 나타났다. 또한, 미분양 매입 정책의 가시적 효과는 시행 시점이 앞설수록 효과가 증대될 것으로 보았으며, 공급측면 정책보다 수요측면 지원 정책이 더 효과적일 것으로 주장하였다.

전해정(2014)은 벡터오차수정모형을 이용해 거시경제변수가 미분양주택 발생에 얼마나 영향을 미치는지 실증적으로 분석하였다. 미분양주택 발생 시 양의 방향으로 지속적으로 증가하며 가격에 크게 영향을 받는 것으로 나타났다. 다만 전국단위로 분석되어 지역별 세분화하지 못하였고, 다양한 거시경제변수를 통한 분석을 수행할 필요성을 제시하였다.

김대원·유정석(2014)은 미분양주택을 효과적으로 산정하기 위해 거시경제 요인, 경제 심리지수, 정책 규제지수 등의 설명변수를 각각 GMM 방식으로 예측한 후 고정효과모형에 LSDV 방식으로 추정하였다. 대출금리, 가격 증감률, 재고대비 분양률은 1기 시차를 두고 미분양 발생에 영향을 미치는 것으로 나타났으며, 경제심리 및 정책규제, 글로벌 경기침체는 시차 없이 미분양 결정에 영향을 미치는 것으로 나타났다.

정상철·성주한(2019)은 창원시를 대상으로 인구, 가격, 금리, 인허가실적 등의 자료를 토대로 벡터오차수정모형을 통하여 분석하였다. 2016년 이전 매매가

격 변화율, 2016년 이후 인허가실적이 미분양에 가장 큰 영향을 주는 것으로 나타났고 전세가격 변화, 금리, 인구변화율이 영향을 미치는 것으로 나타났다.

양영준·유선종(2019)은 제주도 주택 미분양에 영향을 미치는 요인을 다중회귀분석을 통해 분석하였다. 데이터 마이닝 기법으로 15개월간 분양한 59개 단지 와 269개 공급타입을 기준으로 모집공고일로부터 12 개월 동안 미분양을 조사하였다. 모형을 통한 분석결과 아파트가 단독주택 및 다세대주택에 비해 미분양률이 낮았고, 층수가 높을수록, 도심에서 멀어질수록 미분양률이 높은 것으로 나타났다.

박정수·정상철·성주한(2020)은 창원시 구별 패널 데이터를 활용하여 고정효과 모형을 구축하여 미분양 발생 요인을 분석하였다. 그 결과 2016년 이전 가격변화율이 큰 영향을 미쳤고, 2016년 이후 기존 인허가 물량이 공급으로 전환되면서 미분양 발생에 가장 큰 영향을 미친 변수로 나타났다. 또한, 소득수준을 비교하였을 때 40, 50대 인구가 늘어나는 것이 미분양을 감소시키는 요인으로 작용하였다. 다만, 공급 중심 변수에 초점을 맞춘 것과 한정된 지역을 한계로 언급했다.

양지영(2025)은 벡터오차수정모형을 통해 미분양 발생 요인에 대해 거시적으로 규명하였고, 미시적으로 지역별 패널데이터로 고정효과 모형을 통해 미분양 발생 요인을 실증적으로 분석하였다. VECM 분석결과 장기적으로 매매가격 변동률과 공사비지수가 미분양 주택에 유의미한 영향을 미치는 것으로 나타났으며, 패널분석 결과 수도권 및 광역시는 공급 조절, 기타 지역은 경제 활성화와 소비심리 개선이 핵심 요인으로 나타났다.

<표 4> 미분양 발생 원인과 결정요인 분석

연구자	연구내용 및 방법론	연구결과 및 한계점
강병주 (1995)	문헌조사 및 통계자료 활용하여 미분양 원인을 과잉공급, 투기억제로 인한 수요 감소로 제시	미분양 완화방안 제시 범위가 한정되었고 실증적 분석의 한계
정재형 (2000)	공급, 수요자 측면 설문 및 문헌 조사를 통해 지역별 미분양 발생 원인을 파악	가격할인이 효과적인 해소 방안으로 제시 특정 지역 대상 및 설문 수를 충분히 확보 못함
정창무, 김지순 (2005)	VECM을 활용하여 주택정책과 경제변수가 미분양에 어떤 영향을 미치는지 실증분석	정부정책, 금리, 환율 미분양 재고 변동 원인 착공량과 미분양 관계 등 추가연구 필요

연구자	연구내용 및 방법론	연구결과 및 한계점
서범준, 이효중, 정창무 (2010)	패널분석을 통해 지역별 미분양아파트 결정요인과 조정속도에 대해 실증적으로 분석	수도권이 지방에 비해 조정속도가 빠르기에 차별화된 정책 필요 연간 데이터 사용으로 정교한 추정의 한계
김상기, 이상효, 김재준 (2010)	VECM을 활용하여 미분양주택과 주택매매가격 및 주택전세가격 사이의 동태적 관계 규명	매매가격의 영향이 크고 전세가격은 미약함 정책변화, 경제변수를 포함하여 종합적 분석 필요성 제시
박재룡, 유정석 (2010)	글로벌 금융위기 전후로 통상적인 최소자승법(OLS)을 적용하여 분석	가격, 금리는 양(+), 실질소득은 부(-)의 관계로 나타남 단기적 분석 및 변수를 충분히 고려하지 못함
장세웅 (2011)	VECM을 통해 미분양 발생 요인을 체계적으로 분석하고 이를 정량적으로 설명	가격, 공사비, 대출금 등의 거시경제변수와 동태적 관계 입증 정책변수가 제외되어 추가적인 연구 필요
전해정 (2014)	VECM을 통해 거시경제변수가 미분양주택 발생에 미치는 영향을 분석	미분양 발생 시 지속적 증가하며 가격에 크게 영향을 미침 지역별 세분화, 다양한 거시경제변수 포함 필요
김대원, 유정석 (2014)	패널데이터를 활용하여 미분양 결정모형 도출 GMM방식으로 예측, LSDV으로 추정하여 미분양 감소 방안 제시	정책규제와 심리지수가 영향을 미침 짧은 시계열 자료로 정교한 분석의 한계
정상철, 성주한 (2019)	2016년 전후 창원시 대상 VECM을 통해 미분양에 미치는 요인 실증분석	16년 이전 가격 변화율, 16년 이후 인허가실적이 미분양에 영향을 미침 창원시에 한정되어 전국단위 분석 필요
양영준, 유선종 (2019)	59개 단지에 대해 다중회귀분석을 통해 제주도 주택 미분양에 미치는 요인을 분석	미분양률은 아파트가 낮았고, 고층이거나, 도심에서 멀수록 높아짐 소표본 일반화 한계
박정수, 정상철, 성주한 (2020)	통합 창원시 구별 데이터를 활용하여 패널분석을 통해 미분양 아파트 발생 요인을 분석	인허가 물량이 가장 큰 영향을 미치고 40~50대 인구 증가가 미분양 감소 요인으로 작용 공급 변수에 초점, 한정된 분석의 한계
양지영 (2025)	VECM을 통해 경제변수와 상호작용을 거시적으로 분석, 지역별 패널데이터로 고정효과 모형을 통해 미시적으로 분석	가격과 공사비가 유의미한 영향을 미침 수도권은 공급 조절, 지방은 경제 활성화와 소비심리 개선이 핵심 정책적 요인을 계량화한 연구 필요성 제시

2) 미분양주택이 주택시장에 미치는 영향 분석

박천규·이영(2010)은 주택시장 체감지표가 주택 가격, 미분양주택 등의 시장지표를 예측하는데 얼마나 유용한지를 분석하기 위해 해당 변수들의 시간적 추세와 상관관계를 ARX(1) 모형을 통하여 분석하고, 체감지표가 포함된 확장모형과 기본모형을 RMSE를 통해 비교하였다. 체감지표로의 추가로 전반적인 예측력이 개선되었으며, 미분양주택이 신규 주택 수요를 분석하는 중요한 변수로 활용될 수 있고, 건설업경기BSI, 전세수요BSI가 미분양주택을 예측하는 효과적인 지표로 작용하였다.

이의준·김경민(2011)은 수도권 지역 준공 전, 준공 후 미분양이 주택시장에 미치는 영향을 파악하기 위해 고정효과모형과 확률효과모형을 통해 두 유형의 차이를 규명하였다. 준공 전 미분양과 준공 후 미분양은 핫스팟이 다르며, 준공 전 미분양은 주택시장 변화에 따라 지역적으로 이동하는 경향을 보인 반면, 준공 후 미분양은 특정 지역에 고착화되는 특징을 나타내었다. 또한, 단기이자율의 상승은 준공 후 미분양을 증가시키는 요인이고, GDP의 상승은 준공 전 미분양을 증가시키는 요인으로 준공 후 미분양은 현재의 주택시장을 반영하고 준공 전 미분양은 미래의 주택시장을 반영한 결과로 나타났다.

최재규·유승규·김재준(2013)은 미분양주택과 건설업체 부실화가 상호 밀접한 관계가 있을 것이라고 예상하며 건설업체 부실화 정도를 KMV모형의 예상 부도확률로 측정하고 측정된 예상부도확률과 미분양주택 간의 관계성을 벡터오차수정모형을 활용하여 동태적 관계를 규명하였다. 주택건설업체 부실화는 주택 가격의 변화보다 시간이 흐름에 따른 미분양주택 발생이 큰 영향을 받는 것으로 나타났고 건설사 부도는 자기 자신의 위험에 의해서도 발생되기 때문에 미분양주택 등 다른 변수에 대한 부도확률 설명력은 매우 낮은 것으로 나타났다.

성주한·정상철(2019)은 지역별 부동산가격이 다른 상황에 정부의 가격정책이 어떻게 적용할지에 대해 서울과 전국 16개 시도의 패널데이터로 고정효과 모형을 구축하여 가격에 영향을 미치는 요인이 무엇인지 분석하였다. 전세가격과 인구의 증가는 매매가격을 상승시킨 요인으로 나타났고, 금리 상승과 인허가실적 및 미분양의 증가, 실업률이 높은 경우 가격을 하락시킨 것으로 나타났다.

김리영·서원석(2020)은 2019년 미분양이 가장 많이 누적되었던 경남과 미분양이 크게 줄어든 경기도를 대상으로 미분양에 미치는 요인을 비교분석 하기 위해 그랜저인과분석과 벡터오차수정모형을 이용해 규명하였다. 미분양의 원인으로 경남은 지역경제 기반 경제적 요인, 경기도는 가격이나 소비자 심리, 거시경제 등이 유의미한 원인으로 나타나 지역이 가진 여건에 따라 미치는 영향이 상이한 것을 규명하였다. 추후 분양시장과 재고시장을 심층적으로 다뤄야 할 필요성을 제시하였다.

<표 5> 미분양주택이 주택시장에 미치는 영향 분석

연구자	연구내용 및 방법론	연구결과 및 한계점
박천규, 이영 (2010)	ARX(1) 모형을 통하여 주택시장 체감지표가 미분양주택 등의 시장지표를 예측하는데 유용한지 분석	체감지표 추가로 예측력이 개선되었으며, 미분양주택 예측에 효과적인 지표로 작용 시계열 데이터 부족
이의준, 김경민 (2011)	준공 전, 준공 후 미분양이 주택시장에 미치는 영향을 파악하기 위해 패널분석을 통해 분석하고, 차이를 규명	준공 전, 준공 후 미분양 핫스팟이 다르며 주택시장에 미치는 영향이 다름을 규명 시계열 데이터 부족, 지역 이질성 미반영
최재규, 유승규, 김재준 (2013)	건설업체 부실화 정도를 KMV로 측정하고 측정된 예상부도확률과 미분양주택 사이의 관계성을 VECM 활용하여 실증분석	미분양주택과 건설업체 부실화의 동태적 관계를 실증적으로 규명 상장사 26개 대상으로 한정된 기업에 국한됨
성주한, 정상철 (2019)	서울과 전국 16개 시도의 패널데이터를 활용하여 고정효과 모형을 통해 지역별 부동산 가격에 영향을 미치는 요인 실증분석	매매가격에 공통적으로 미치는 요인 규명 시도별 특성이 나타나지 않아 지역변수 포함하여 추가연구 필요성 제시
김리영, 서원석 (2020)	미분양이 가장 누적되었던 경남과 크게 줄어든 경기도 대상으로 VECM을 활용하여 거시경제 변수와 상호작용을 분석	경남은 경제적 요인, 경기도는 거시경제 요인이 유의미한 원인으로 나타나 지역별 영향이 상이함을 규명 분양·재고시장 심층 분석 필요

3) 미분양 해소 및 대응을 위한 정책적 접근

김태섭·김은덕(1996)은 다단계 다중회귀분석을 통해 미분양 발생 원인을 추정하였다. 공급 관련 변화요인이 수요 관련 변화요인보다 증감에 약 2배 이상의 영향을 미친다고 보아 공급 유발 정부정책이 미분양 발생에 더 큰 요인으로 작용한다고 밝혔다.

남희용·김태섭(2001)은 경기도 20개 단지 대상으로 다중회귀분석을 통해 가격이 미분양에 가장 큰 영향을 미치는 변수로 나타났다. 또한, 세대별 특성을 고려한 분양가 차등화, 특화설계, 교통 편리성, 주거 쾌적성 등이 미분양을 감소시키는 원인으로 나타났다.

김현아(2007)는 전국 및 지역별 미분양주택, 분양 및 입주 물량 등의 주택시장 지표를 활용하여 과거 미분양이 급증한 외환위기 시점과 비교하여 유사점과 차이점을 도출하였다. 또한, 미분양의 장기화를 예측하여 중소 건설업체의 도산, PF부실화, 지역경제 침체 등 파급효과를 언급하였고 해결방안으로 금융지원, 규제완화, 임대주택 활성화 등의 필요성을 제시하였다.

허재완·김은경(2009)은 미분양 발생이 주기적으로 반복되고 있음을 확인하여 3개 시기로 나누어 지역적, 규모별, 부문별 특성과 시기별로 어떠한 요인이 미분양주택 증가에 영향을 미치는지 연구하였다. 그 결과 미분양주택은 지역별·시기별 차이가 있고, 시간이 지날수록 준공 후 미분양의 비중이 높아져 이는 지방 미분양 증가와 밀접한 관련이 있는 것으로 보았다.

김덕례·황은정(2009)은 2009년 준공 후 미분양주택이 과거 외환위기보다 2배 이상 늘어남에 따라 과거 정부 정책의 효과를 통계자료 및 문헌분석, 심층면접조사, 분산분석을 통해 향후 정책 방향을 제시하였다. 과거 미분양 해소 정책이 수요나 공급 단일에 맞춰진 정책보다 복합적인 정책이 더 효율적이었고 주택수요자에 대한 금융지원, 각종 세제 완화, 주택담보대출의 확대 등이 효과가 있다고 보았다.

임덕호(2009)는 2008년 금융위기로 인한 전국적인 미분양 확산을 왜곡된 분양시장을 정상화 할 수 있는 기회라 보아 각종 규제의 실태를 파악하고 분양시장 정상화 방안을 제시하였다. 정상화를 위해서는 공공과 민간의 역할을 명확히 구분하고, 민간택지 분양가 자율화, 공공택지 후분양제 도입 및 원가 공개 의무화 등이 필요하다고 보았다.

이성근·전광섭(2009)은 미분양아파트 증가가 시행사 부도, 금융시장 불안 등 경제에 미치는 악영향이

크다고 진단하였고 해소 방안으로 매입대금 등의 확정채권, 분양대금 등의 장래 채권으로 구분하여 다양한 구조의 유동화 방안을 제안하였다. 다만 이런 구조에서 주된 위험은 부동산 가치하락에 있기 때문에 강구책이 사전에 있어야 할 것이고, 적시 매각의 불확실성으로 현금화에 어려움이 있으며, 세제혜택이 없는 점을 문제로 언급하였다.

정성훈·김주영(2010)은 미분양을 해소화하기 위한 방식 중 하나로 건설업체의 유동성 확보를 위한 마스터신탁 금융구조화에 대해 분석하였다. 부동산신탁펀드, 미분양 펀드, 준공 전 미분양 리츠(CR리츠) 등 기존 방식의 한계를 파악하고 이를 해소하고자 마스터신탁 구조의 도입 필요성을 제시하였다. 마스터신탁 사업방식은 자산을 하나의 큰 신탁으로 묶고 ABS를 발행하는 방식으로 새로운 증권을 발행하는데 관련 비용이 낮고 절차상 유연한 측면에서 효율적이기 때문에 자금 유동성에 큰 도움이 될 것으로 보았다.

김주영·김형순(2010)은 미분양주택 발생구조를 지역적으로 분석하고자 다층(위계적)모형을 활용하여 실증분석하였다. 도시차원 인구 30만 이상이면서 제조업 중심 도시들의 미분양이 가장 높게 나타났으며, 공급량 증가와 미분양은 정의 관계를 나타내는 반면, 소득수준과 인구성장률의 경우 단기에 미분양 결정요인이 되지 못하는 것으로 나타났다.

허재완·손성민(2013)은 정부의 미분양 정책 효과성에 대해 정책 더미변수와, 거시경제 변수를 통제변수로 포함한 벡터오차수정모형을 이용해 실증적으로 분석하였다. 전국적으로 DTI 규제가 강한 효과를 보였고, 수도권에서는 세제완화, 지방에서는 환매조건부 미분양 매입, 건설사 지원이 가장 효과적인 대책으로 나타났다.

황관석·김지혜·이운상(2023)은 2022년 급격한 금리상승과 긴축정책으로 거래량이 감소하고 미분양주택이 증가하자 1차적 PF 부실로 이어지고, 2차적 수분양자 및 국민경제에 영향을 미칠 것으로 보아 지역별 패널분석과 과거 사례 비교분석을 통해 선제적 대응방안을 강구하였다. 지방을 중심으로 미분양주택은 빠르게 증가하고 있으며 금리, M2와 같은 통화량 증가에 영향을 미치는 것으로 나타났다. 해결방안으로 위험 단계를 구분하여 대응이 필요하고 중소형 주택이 상대적으로 많기 때문에 실수요자 수요지원 정책, 미분양 매입임대 사업 등을 추진할 필요가 있다고 보았다.

<표 6> 미분양 해소 및 대응을 위한 정책적 접근

연구자	연구내용 및 방법론	연구결과 및 한계점
김태섭, 김은덕 (1996)	다단계 다중회귀분석을 통해 수요와 공급 관련 어느 요인이 민감하게 반응하는지 분석	수요보다 공급 요인에 미분양이 민감하게 반응함을 규명 분석 방법의 한계 존재
남희용, 김태섭 (2001)	회귀분석과, T-test를 통하여 미분양아파트 특성 요소 파악 및 해소 방안 제시	적정 가격과 특화설계, 입지가 미분양 최소화 요인이라고 규명 연구범위 한정 및 시계열 변화 미반영
김현아 (2007)	주택시장 지표를 활용하여 외환위기 시점과 유사점과 차이점을 도출	미분양의 장기화에 따른 PF부실, 지역경제 침체 예상하여 해소책 제시 정량적 분석의 한계
허재완, 김은경 (2009)	시계열데이터 비교로 미분양주택 특성과, 시기별로 어떠한 요인이 영향을 미치는지 파악	지역별·시기별 차이가 있고, 시간이 지날수록 준공 후 미분양 증가 실증적 분석의 한계
김덕례, 황은정 (2009)	정부정책의 효과를 문헌조사 및 ANOVA를 통해 분석하여 문제점 진단 및 해소 방안 제시	할인분양, 미분양주택 매입 등의 지원 필요 계량모형 기반의 인과관계 분석 필요
임덕호 (2009)	금융위기로 인한 전국적인 미분양 및 각종 규제를 파악하여 정상화 방안 제시	분양가 자율화, 후분양제 등 제도적 개선 필요 이론적 접근의 한계
이성근, 전광섭 (2009)	미분양아파트의 고유 특성을 분석하여 다양한 금융구조의 유동화 방안을 제안	환매조건부 매입방식 미분양주택을 해소하는 방식의 도입 필요 현금화 및 세제 혜택이 없는 점을 문제로 제기 증권 발행비용이 낮고 절차상 유연한 측면에서 적극적 활용 필요 효율적 관리 및 체계화 어려움
정성훈, 김주영 (2010)	미분양 해소 방안으로 마스터신탁 사업방식 도입 필요성 제시	미분양주택 해소 관점에서 지역 중심의 정책이 필요 지역별 다양한 특성을 고려하지 못한 한계
김주영, 김형순 (2010)	지역별 특성 변수로 다층성장모델 기반 실증분석 후 정책적 시사점 제시	DTI 완화 세제완화, 미분양 매입 또는 건설사 지원이 효과적 정책 효과를 정확히 파악하지 못한 한계
허재완, 손성민 (2013)	VECM을 통해 미분양 정책 효과성에 대해 실증적으로 분석	
황관석, 김지혜, 이운상 (2023)	확률효과 모형을 통해 지역별 패널분석과 과거 사례 비교분석을 통해 선제적 대응방안 제시	수요지원 정책, 미분양 매입임대 등을 추진하고 지역별 공급관리 정책 병행이 필요

4) 선행연구와의 차별성

초기의 연구는 주로 경제이론을 바탕으로 설문조사, 주택시장 지표를 활용하여 정책적 시사점을 제시한 연구가 대부분이며, 이후 연구는 미분양주택이 주택시장에 미치는 영향을 파악하거나 결정요인을 규명하여 해결방안을 제시하는데 초점이 맞추어져 있어 정량적인 예측에 관한 연구는 전무하다.

본 연구는 준공 후 미분양주택에 대해 다루어 리스크 확산의 선행 지표로서 갖는 의미를 규명하고, 이를 독립적으로 분석함으로써 기존 연구와 차별화된 연구를 시도하였다. 또한, 보다 정교한 예측을 위해 여러 시계열모형을 활용하여 최적의 모형을 선정하고 이를 토대로 단기적 추세를 예측하여 대응방안을 제시하고 실질적인 기여를 제공함에 기존 연구와 차별이 있다.

6. DW모형과 선행연구를 토대로 외생변수 선정

DW모형 공간시장에서 가격은 수요와 공급의 결정 요인이며 가격이 증가하면 자산시장에서 부동산 가치는 상승하게 된다. 가치평가는 기대수익률(cap rate)을 통해 산정하기에 금리가 주요 요인이며, 건설시장에서는 자산 가치 상승이 주택공급으로 이어지는지를 설명하기 때문에 공급 조절에 영향을 미치는 건설업 GDP, 공사비, 착공실적, 주택금융보증잔액 등이 주요 요인으로 도출할 수 있다.

또한 선행연구를 통해 미분양주택 증감에 유의미하게 영향을 미치는 변수를 참고하여 외생변수를 설정함으로써 이론적 기반과 실증적 근거를 함께 반영하였다. 선정된 외생변수들의 정의 및 전처리 과정은 제4장에서 구체적으로 서술하였다.

<표 7> 외생변수 선정

변수명	내용	참고 선행연구
<i>price</i>	아파트 매매가격지수	(김상기·이상호·김재준, 2010; 전해정, 2014; 양지영, 2025)
<i>C_index</i>	건설공사비지수	(장세웅, 2011; 양지영, 2025)
<i>begin</i>	아파트착공실적	(김주영·김형순, 2010; 박정수·정상철·성주한, 2020)
<i>C_gdp</i>	건설업 GDP	(박재룡·유정석, 2010; 이의준·김경민, 2011)
<i>rate</i>	금리 (회사채 AA-)	(정창무·김지순, 2005; 김대원·유정석, 2014)
<i>hf</i>	주택금융보증 잔액	(장세웅, 2011; 김덕례·황은정, 2009)

<표 8> 실증분석 기반 선행연구 방법론 비교

연구자	방법론	다중 회귀분석	다층 성장모형	ARX(1)	ANOVA	ECM	VECM	패널분석	시계열예측
김태섭·김은덕(1996)		●							
남희용·김태섭(2001)		●							
정창무·김지순(2005)						●			
김덕례·황은정(2009)					●				
서범준·이효중·정창무(2010)								●	
김상기·이상효·김재준(2010)							●		
박재룡·유정석(2010)		●							
박천구·이영(2010)				●					
김주영·김형순(2010)			●						
장세웅(2011)							●		
이의준·김경민(2011)								●	
최재규·유승규·김재준(2013)							●		
허재완·손성민(2013)							●		
전해정(2014)							●		
김대원·유정석(2014)								●	
정상철·성주한(2019)							●		
양영준·유선종(2019)		●							
성주한·정상철(2019)								●	
박정수·정상철·성주한(2020)								●	
김리영·서원석(2020)							●		
황관석·김지혜·이운상(2023)								●	
양지영(2025)							●	●	
본 연구(2025)									●
합계(23)		4	1	1	1	1	8	7	1

III. 시계열예측모형

1. 시계열예측모형

이 장에서는 본 연구에서 준공 후 미분양주택 수를 예측하기 위하여 사용하는 여러 시계열모형을 설명한다. 시계열모형은 종속변수의 시계열적 특성을 포함하는 대표적 모형인 ARIMA 모형을 기초로 계절성을 추가하여 계절적 패턴을 반영하거나, 외생변수를 포함하여 종속변수에 미치는 영향을 반영한 모형을 차례대로 설명한다.

1) ARIMA

Box and Jenkins(1970)가 제안한 ARIMA 모형은 시계열데이터를 예측하는데 일반적으로 많이 사용되었던 모형으로 자기회귀(AutoRegressive, AR)는 종속변수의 과거 값으로 현재의 시계열을 설명하고, 자기회귀이동평균(Moving Average, MA)은 백색잡음의 과거 값으로 이를 포함하여 모형화 한다. 또한 일반적

으로 시계열 데이터는 비정상적(nonstationary)인 경우가 많은 것을 고려하여 d 차 차분(Integrated)을 적용하여 정상 시계열로 변환 후 적합 시킨 모형을 말한다. 시계열 Y_t 가 평균이 μ 인 $ARIMA(p, d, q)$ 의 모형의 경우 식 (1)로 표현할 수 있다.

$$\phi_p(B)(1-B)^d(Y_t - \mu) = \theta_q(B)\epsilon_t \quad (1)$$

where $\epsilon_t = WN(0, \sigma^2)$

Y_t 는 t 시점에서의 관측된 값을 나타내고, B 는 후진 연산자(back-shift operator)를 의미한다. p 는 자기회귀(AR)차수, q 는 이동평균(MA)차수, d 는 차분차수를 의미하며 $\phi_p(B)$ 는 식 (2), $\theta_q(B)$ 는 식 (3)과 같이 각각 p, q 차 다항식으로 나타낼 수 있다.

$$\phi_p(B) = 1 - \phi_1 B - \dots - \phi_p B^p \quad (2)$$

$$\theta_q(B) = 1 - \theta_1 B - \dots - \theta_q B^q \quad (3)$$

2) SARIMA

SARIMA(Seasonal ARIMA)모형은 ARIMA모형에 계절성 및 반복된 주기의 시계열적 패턴이 포함되어 확장된 모형으로 Box and Jenkins(1994)에 의해 체계적으로 정리되었다. SARIMA의 특성은 계절적 자기회귀(SAR)와 계절적 이동평균(SMA) 요소가 포함되어 주기적인 변동성을 모델링 할 수 있다는 점으로 비 계절적 요소와 계절적 요소를 동시에 고려하여 복합적인 시계열 패턴을 세밀하게 분석할 수 있다. 이에 따라 SARIMA 모형은 패턴이 있는 시계열데이터의 분석 정확성을 높이고 예측 오류를 줄여 보다 현실적인 분석을 수행할 수 있다. 시계열 Y_t 가 평균이 μ 인 SARIMA(p, d, q)(P, D, Q)_s의 모형의 경우 식 (4)로 표현할 수 있다.

$$\begin{aligned} \phi_p(B)\Phi_P(B^S)(1-B)^d(1-B^S)^D(Y_t-\mu) = \\ \theta_q(B)\Theta_Q(B^S)\epsilon_t \end{aligned} \quad (4)$$

where $\epsilon_t = WN(0, \sigma^2)$

앞서 ARIMA모형과 같이 Y_t 는 t 시점에서의 관측된 값을 나타내고, B 는 후진 연산자, p 는 자기회귀(AR)차수, q 는 이동평균(MA)차수, d 는 차분차수를 의미한다. SARIMA모형에서 추가로 포함된 P 는 계절 자기회귀(SAR)차수, Q 는 계절 이동평균(SMA)차수이며, S 는 계절주기로 본 연구에는 월별데이터를 사용하여 계절 주기는 12로 설정하였다. $\phi_p(B)$, $\theta_q(B)$ 는 식 (2), 식 (3)과 같으며 $\Phi_P(B^S)$ 는 식 (5), $\Theta_Q(B^S)$ 는 식 (6)과 같이 각각 P , Q 차 다항식으로 나타낼 수 있다.

$$\Phi_P(B^S) = 1 - \Phi_1 B^S - \dots - \Phi_P B^{PS} \quad (5)$$

$$\Theta_Q(B^S) = 1 - \Theta_1 B^S - \dots - \Theta_Q B^{QS} \quad (6)$$

3) ARIMAX

ARIMAX(ARIMA with eXogeneous variable)는 ARIMA의 확장된 형태로 기본 모형에 외생변수가 추가된 모형을 말한다. 거시경제변수와 같은 다양한 외생변수의 포함은 모델의 복잡성을 증가시킬 수 있지만 해당 요인이 종속변수에 미치는 영향을 분석하여 예측 정확성을 높이는 데 도움을 줄 수 있는 장점이 있다.

특히 ARIMAX 모형은 외생변수의 선정과 모형의 파라미터 설정이 예측 성능에 큰 영향을 미칠 수 있으므로 외생변수의 선정과 AIC, BIC 등의 통계량을 활용하여 모형에 적합한 시차를 설정하여 예측의 신뢰성을 높이는 절차가 중요하게 작용한다. 이러한 특성으로 시계열데이터를 분석하는데 장점을 가지며 정교한 예측을 가능하게 한다. 시계열 Y_t 가 평균이 μ 이고 외생변수 x_{it} 가 k 개인 ARIMAX(p, d, q)의 모형의 경우 식 (7)로 표현할 수 있다.

$$\begin{aligned} \phi_p(B)(1-B)^d(Y_t-\mu) = \theta_q(B)\epsilon_t + \sum_{i=1}^k \gamma_i x_{it} \end{aligned} \quad (7)$$

where $\epsilon_t = WN(0, \sigma^2)$

Y_t 는 t 시점에서의 관측된 값으로 p, q 차 다항식인 $\phi_p(B)$, $\theta_q(B)$ 는 식(2), 식(3)에서 설명한 수식과 동일하며 γ_i 는 외생변수 x_{it} 의 계수를 나타낸다.

4) SARIMAX

SARIMAX(SARIMA with eXogeneous variable)는 SARIMA 모형의 확장된 형태로 ARIMAX와 같이 외생변수가 포함되어 있는 모형을 말한다. 이 모형의 특징은 계절적 패턴과 외부 요인을 동시에 고려한 종합적인 시계열모형으로 데이터의 복합적인 구조를 파악하고 예측의 정확성을 높일 수 있는 특성을 갖고 있다. 시계열 Y_t 가 평균이 μ 이고 외생변수 x_{it} 가 k 개인 SARIMAX(p, d, q)(P, D, Q)_s의 모형의 경우 식 (8)로 표현할 수 있다.

$$\begin{aligned} \phi_p(B)\Phi_P(B^S)(1-B)^d(1-B^S)^D(Y_t-\mu) = \\ \theta_q(B)\Theta_Q(B^S)\epsilon_t + \sum_{i=1}^k \gamma_i x_{it} \end{aligned} \quad (8)$$

where $\epsilon_t = WN(0, \sigma^2)$

Y_t 는 t 시점에서의 관측된 값으로 p, q, P, Q 차 다항식인 $\phi_p(B)$, $\theta_q(B)$, $\Phi_P(B^S)$, $\Theta_Q(B^S)$ 는 식 (2), 식 (3), 식 (5), 식 (6)에서 설명한 수식과 동일하며 γ_i 는 외생변수 x_{it} 의 계수를 나타낸다.

2. Box-Jenkins 절차

Box-Jenkins 방법은 ARIMA기반 시계열분석에서 데이터에 가장 적합하고 최적의 모형을 찾는 과정으로 3단계로 나누어지며 각 단계는 모형 식별, 추정, 진단 및 예측으로 분류할 수 있다.

1) 모형 식별

모형 식별 단계는 Box-Jenkins절차에서 핵심적인 과정으로 먼저 분석하고자 하는 데이터가 평균과 분산이 일정한 정상 시계열인지 판단하고 비정상적인 시계열일 경우 차분(differencing)이나 로그변환을 통해 정상성을 확보한다. 정상성 여부는 시계열 플롯으로 확인할 수 있지만 ADF(Augmented Dickey-Fuller test)와 같은 통계적 검정을 통해 단위근(unit root) 존재 유무를 확인하여 정상성 여부를 확인한다.

정상 시계열을 확인하였다면 이후 자기상관함수(ACF)와 편자기상관함수(PACF)로 모형을 식별한다. ACF는 시차에 따라 데이터의 자기상관 정도를 나타내고, PACF는 특정 시차의 자기상관을 나타낸다. 이를 통해 AR() 모형과 MA() 모형의 차수를 결정할 수 있다. 즉 ACF와 PACF의 패턴을 분석하여 데이터에 적합한 ARIMA() 모형의 파라미터를 설정할 수 있다. PACF가 특정 시차에서 절단되는 형태를 보이고 이후 시차에서 0에 가까운 값을 갖는다면 절단되는 시차를 적용한 AR모형이 적합할 가능성이 있고, ACF에서는 특정 시차에서 절단되는 시차를 적용한 MA모형이 적합할 가능성이 있다.

2) 모형 추정

최대우도법(Maximum Likelihood Estimation)을 통해 모형의 파라미터를 추정하고 모형 선별 기준으로 알려진 AIC(Akaike Information Criterion) 또는 BIC(Baysian Information Criterion) 통계량을 고려하여 적합한 모형을 결정한다. AIC와 BIC 값은 식 (9)로 나타낼 수 있다.(Hyndman and Athanasopoulos, 2018)

$$\begin{aligned} AIC &= T \cdot \log\left(\frac{SSE}{N}\right) + 2(k+2) \\ BIC &= T \cdot \log\left(\frac{SSE}{N}\right) + 2(k+2)\log(N) \end{aligned} \quad (9)$$

위 식에서 T 은 관측치의 수이며, k 는 모수의 개수이다. AIC와 BIC값이 작을수록 모형이 적합함을 의미하므로 작은 IC 값을 갖는 차수를 찾는 것이 중요하다. 즉 AIC와 BIC가 유사한 수준의 작은 값을 갖는다면, 차수의 합이 간단한 모형을 선정한다.(이근철·최성훈, 2020)

3) 진단 및 예측

추정된 모형의 잔차가 백색잡음(White Noise) 여부를 판단하기 위해 Ljung and Box(1978)이 제안한 Ljung-Box Q 통계량을 이용하여 잔차의 자기상관 여부를 검정하며 검정 통계량은 식 (10)으로 나타낼 수 있다

$$Q(h) = n(n+2) \sum_{k=1}^h \frac{\hat{\rho}_k^2}{n-k} \quad (10)$$

n 은 표본의 크기이고 $\hat{\rho}_k^2$ 는 lag k 에서의 표본 자기상관, h 는 검정할 lag의 수를 나타내며 자기상관 계수가 동시에 0과 같은 값을 가진다는 귀무가설을 검정한다. 귀무가설을 기각하여 모형이 적합하지 않다면, 식별 단계로 돌아가 모형을 재조정하는 과정을 다시 수행한다.

위와 같은 과정을 통해 앞서 설명한 각 시계열모형이 적절히 진단되고 선택되었다면, 이를 사용하여 예측을 수행한다. 예측의 성능 지표는 MAPE, RMSE를 통해 평가하여 비교하였고 식 (11)과 같다.

$$\begin{aligned} MAPE &= \frac{1}{n} \sum_{t=1}^n \left| \frac{Y_t - F_t}{Y_t} \right| \times 100(\%) \\ RMSE &= \sqrt{\frac{1}{n} \sum_{t=1}^n (Y_t - F_t)^2} \end{aligned} \quad (11)$$

n 은 예측에 사용한 데이터의 수, Y_t 의 경우 t 시점에서의 실제값, F_t 의 경우 t 시점에서의 예측값을 의미하여 값이 작을수록 우수한 모형인 것을 알 수 있다.

IV. 실증분석

1. 시계열자료 준비 및 변환

본 논문에서 예측에 사용된 종속변수는 국토교통부에서 매월 공표하는 민간부문 준공 후 미분양주택 자료로 월별로 수집된 데이터이다. 2011년 저축은행 사태로 인한 금융위기를 기점으로 이후 저금리 기조, 코로나 팬데믹 등 증감을 반복하다 다시 급증한 최근 2025년까지의 중장기적 흐름을 포괄적으로 분석하기 위해 2011년 1월부터 2025년 3월까지 기간의 데이터를 사용하였다. 이 중 2011년 1월부터 2023년 12월까지 훈련용 데이터(training data)로 모형을 적합하고, 2024년 1월부터 2025년 3월까지를 테스트 데이터(test data)로 모형의 성능을 평가하는데 사용하였다.

훈련용 데이터 설정 기간은 다양한 변동성을 포착할 수 있다. 이 기간 경제상황의 변화, 정부의 부동산 정책, 금리 변동 등 여러 외부 요인이 부동산시장에 영향을 미쳤으며, 이를 토대로 예측 모델이 다양한 상황에서의 패턴을 학습할 수 있다. 반면, 테스트 기간 데이터는 모델이 훈련되지 않은 새로운 예측값으로써, 실제 관측값과 비교하여 모델의 예측력을 평가하였다.

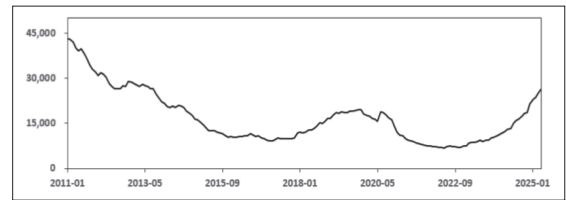
모형을 적합하기 위해 주어진 정상성(stationarity)을 확인하고, 비정상성을 띄고 있으면 적절하게 변환을 해줄 필요가 있다. <그림3>은 관찰된 준공 후 미분양주택의 월별 추이를 나타낸 그래프이다. 그래프에서 나타나듯이 준공 후 미분양주택은 과거 13년간 감소와 증가를 반복해왔고 2022년 12월 이후 점차 다시 증가 추세를 보이고 있다. 이러한 불규칙한 추세가 있는 시계열데이터는 정상성 여부를 의심해볼 필요가 있는데, 검정 방법 중 널리 이용되고 있는 ADF(Augmented Dickey-Fuller) 검정과 PP(Phillips-Perron) 검정을 수행하였다. 검정결과 ADF 통계량은 -1.901, PP통계량은 -1.953로 1% 임계값 보다 높게 나타나 유의수준 1% 에서 정상성을 따르지 않는다는 귀무가설을 채택하게 된다. 이에 따라 본 연구에서는 로그변환을 통해 분산을 안정화시키고, 1차 차분을 통해 추세를 제거하였다. 즉 미분양주택 수를 증가율로 변환하였고 <그림4>은 증가율 그래프이다. 변환된 데이터의 정상성 테스트 결과 ADF 통계량은 -5.039, PP통계량은 -8.132로 1% 임계값 보다 낮게 나타나 정상성을 따르지 않는다는 귀무가설을 기각하여 정상성을 확보하였다.

<표 9> 정상성 테스트 결과

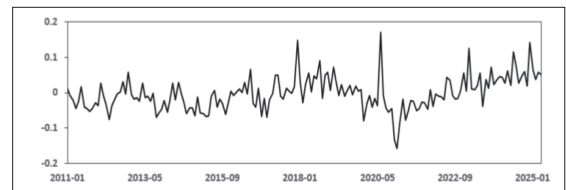
구분	ADF		PP	
	tau-stat	1% critical value	z-stat	1% critical value
Unsold (origin)	-1.901	-3.460	-1.953	-3.470
Unsold	-5.039***	-3.460	-8.132***	-3.470

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

<그림 3> 준공 후 미분양주택 원 시계열 (2011.01~2025.03)



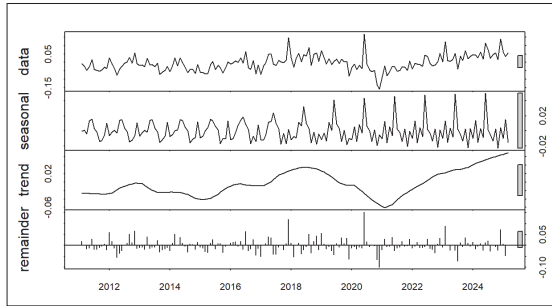
<그림 4> 준공 후 미분양주택 증가율 (2011.02~2025.03)



원 시계열데이터에서는 뚜렷한 계절성이 관찰되지 않았으나, 증가율로 변환 후 계절성을 의심할 수 있다. 이는 데이터의 기본 추세와 변동성을 제거 후 드러난 패턴으로 모델링 과정에서 중요한 고려사항이 된다. 따라서 Cleveland et al.,(1990)이 제안한 STL 방법으로 시계열 데이터를 Seasonal(계절성), Trend(추세), Residual(잔차) 등 세 요소로 분해하였고 그 결과는 <그림5>과 같다.

STL 분해를 통해 Seasonal(계절성)을 살펴보면 차분된 데이터에서 시각적으로 계절성이 존재한다고 판단되어 QS(quality control statistics) 검정을 수행하였다. 해당 통계량은 계절성이 없다는 귀무가설을 검정하며 계절 차분에 해당하는 자기상관 값이 통계적으로 유의한지를 검정 해준다(U.S. Census, 2013). 검정결과 p-value가 0.05보다 큰 것으로 나타나 유의미하게 계절성이 없다는 결과를 나타냈다. 다만 QS검정은 비주기적 계절성을 포착하는데 한계가 있기 때문에 계절 차수를 포함하거나, 포함하지 않는 방법을 모두 시도해 보고 예측 성능을 비교하여 최적의 모델을 구축하고자 한다.

<그림 5> 준공 후 미분양주택 증가율 STL 분해



2. 외생변수 정의 및 전처리

앞서 DW모형과 선행연구를 참고하여 선정된 외생 변수 중 건설업 GDP의 경우 한국은행에서 분기별로 제공되고 있는데, 종속변수 포함하여 모든 변수는 월별 데이터로써 분석을 위해 변환이 필요하다. 이를 해결하기 위해 선형보간법(linear interpolation)을 적용하여 분기별 자료를 월별 자료로 변환하였다.

또한 각기 다른 단위와 비정상성이 확인되어 금리를 제외한 나머지 변수의 데이터를 로그변환을 통해 분산을 안정화시키고, 모든 변수의 1차 차분을 통해 추세를 제거하여 안정성을 확보하였다.

<표 10> 외생변수 선정 및 설명

변수명	내용	단위	기간	출처
<i>price</i>	아파트매매가격지수 (2017년=100)	지수	월간자료	한국 부동산원
<i>C_index</i>	건설공사비지수 (2020년=100)	지수	월간자료	한국건설 기술연구원
<i>begin</i>	아파트착공실적	호	월간자료	국토교통부
<i>C_gdp</i>	건설업GDP (계절조정, 실질)	십억원	분기자료	한국은행
<i>rate</i>	금리 (회사채 3년, AA-)	%	월간자료	한국은행
<i>hf</i>	주택금융보증 잔액	억원	월간자료	한국주택 금융공사

<표 11> 외생변수 전처리 과정

변수명	전처리
<i>price</i>	$\ln price_t - \ln price_{t-1}$
<i>C_index</i>	$\ln C_index_t - \ln C_index_{t-1}$
<i>begin</i>	$\ln begin_t - \ln begin_{t-1}$
<i>C_gdp</i>	$\ln C_gdp_t - \ln C_gdp_{t-1}$
<i>rate</i>	$rate_t - rate_{t-1}$
<i>hf</i>	$\ln hf_t - \ln hf_{t-1}$

<표 12> 외생변수 정상성 테스트 결과

구분	ADF		PP	
	tau-stat	1% critical value	z-stat	1% critical value
<i>price</i>	-3.849***	-3.460	-4.025***	-3.470
<i>C_index</i>	-8.236***	-3.460	-11.810***	-3.470
<i>begin</i>	-14.392***	-3.460	-25.625***	-3.470
<i>C_gdp</i>	-6.877***	-3.460	-5.894***	-3.470
<i>rate</i>	-7.046***	-3.460	-7.979***	-3.470
<i>hf</i>	-3.313**	-3.460	-4.392***	-3.470

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

3. 기초통계량

원 데이터는 절대적 수준을 평가하는 데 사용되며 2011년 1월부터 2025년 3월까지 준공 후 미분양주택은 평균 16,849호로 나타났으며 표준편차가 8,612로 변동성이 큰 것을 알 수 있다. 또한, 최소값과 최대값의 범위가 6,830호에서 43,207호 사이로 넓은 분포를 나타내고 있다. 이를 해결하기 위해 본 연구에서는 전처리 과정을 통해 데이터의 변동성을 조정하였다. 증가율로 변환한 데이터는 변동성을 줄이고, 평균과 분산의 일정성을 유지하고 정상성을 확보하여 추세와 계절성을 보다 정확하게 평가할 수 있다.

<표 13> 기초통계량 (n = 171)

구분	Mean	StdDev	Min	Max	Median
<i>Unsold(origin)</i>	16,849	8,612	6,830	43,207	14,856
<i>price(origin)</i>	104	18	81	145	100
<i>C_index(origin)</i>	98	18	76	130	93
<i>begin(origin)</i>	28,595	16,751	4,332	98,622	25,321
<i>C_gdp(origin)</i>	25,755	2,534	21,199	29,183	27,165
<i>rate(origin)</i>	3.0	1.0	1.6	5.5	2.8
<i>hf(origin)</i>	771,214	361,792	194,444	1,326,419	716,705
<i>Unsold</i>	-0.31	0.91	-2.33	2.04	-0.29
<i>price</i>	0.05	1.03	-2.33	3.02	0.08
<i>C_index</i>	-0.07	0.99	-2.25	2.14	-0.10
<i>begin</i>	-0.07	0.96	-2.57	2.26	-0.06
<i>C_gdp</i>	-0.11	0.99	-2.09	2.76	-0.15
<i>rate</i>	0.05	1.04	-2.60	2.11	0.03
<i>hf</i>	-0.08	1.05	-2.69	2.40	0.01

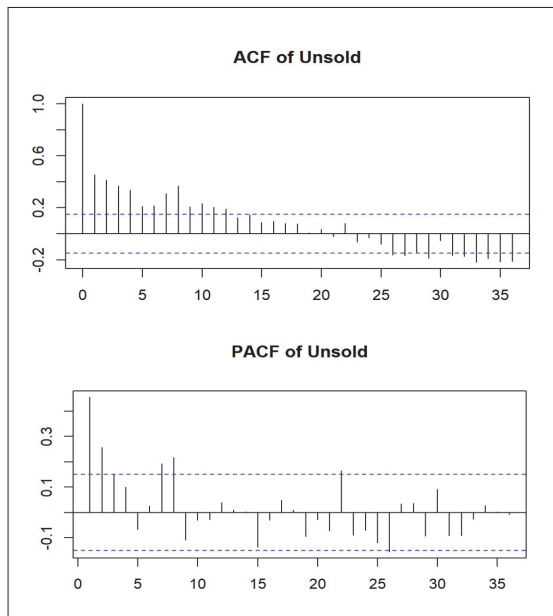
4. 모형 식별

<그림6>은 전처리 과정을 통해 정상성을 확보한 시계열 데이터의 ACF(Autocorrelation Function) 및 PACF(Partial Autocorrelation Function)를 시각화하여, 자기회귀(AR) 차수, 이동평균(MA) 차수, 계절 자기회귀 차수, 계절 이동평균 차수의 예상 파라미터를 추정하였다.

ACF 분석결과, lag 1에서 가장 높은 스파이크가 나타났고 이후 lag 5까지는 점차 감소하는 경향을 보였다. 그러나 lag 5 이후에는 다시 증가하여 한 차례 반등하였다가 이후 다시 감소 추세를 보였다. 이러한 패턴을 고려하여 MA 차수는 1~5범위 내 후보로 검토하였고, lag 12, 24 등 계절 시차에서 유의미한 스파이크는 관측되지 않았다. PACF 분석결과 lag 1, 2, 7, 8에서 유의수준을 벗어나는 스파이크가 관측되었고 이후 시차에서는 빠르게 감소하거나 유의수준 내 머무르는 패턴을 나타냈다. 이러한 PACF의 구조는 자기회귀(AR)의 존재 가능성을 제시하여 AR 차수는 1~5 범위 내 후보로 검토하였다.

전체적으로 ACF 및 PACF 그래프의 해석 결과 AR 및 MA 항을 모두 포함하는 모형을 기본으로 설정하고, 추후 모형 추정 단계에서 AIC 및 BIC 통계량을 통해 각 최적의 차수를 결정하여 모형을 선정하고자 한다.

<그림 6> ACF, PACF 그래프



5. 모형 추정

앞서 ACF 및 PACF 분석을 통해 시계열의 구조적 특성을 진단한 결과를 토대로 각 모형에 대해 최대우도법(Maximum Likelihood Estimation) 기반으로 모수를 추정하고, 정보 기준인 AIC 및 BIC을 활용하여 최적의 파라미터를 설정하였다. 파라미터 탐색은 사전에 후보로 설정한 범위를 Grid Search로 전수조사하여 ARIMA, ARIMAX 모형에는 총 72가지, SARIMA, SARIMAX 모형은 비계절 차수와 계절 차수를 각각 0~3 범위로 설정하여 총 512가지 조합의 모형을 추정하였다. 각 조합에 대해 AIC와 BIC를 계산한 후, 가장 낮은 값을 갖는 모형을 최적 모형으로 판단하였고 <표 14>와 같이 나타낼 수 있다.

<표 14> 모형별 최적 파라미터 설정

AIC 기준	(p, d, q)	(P, D, Q) _s	AIC	BIC
ARIMA	(4, 1, 3)	-	-564.407	-537.016
SARIMA	(3, 1, 3)	(1, 1, 1)	-492.228	-465.562
ARIMAX	(4, 1, 2)	-	-581.188	-538.580
SARIMAX	(3, 1, 2)	(0, 1, 1)	-510.460	-471.943
BIC 기준	(p, d, q)	(P, D, Q) _s	AIC	BIC
ARIMA	(1, 1, 1)	-	-554.727	-542.553
SARIMA	(1, 1, 1)	(0, 1, 1)	-482.809	-470.958
ARIMAX	(4, 1, 2)	-	-581.188	-538.580
SARIMAX	(3, 1, 2)	(0, 1, 1)	-510.460	-471.943

AIC 통계량 기준 ARIMA(4,1,3), ARIMAX(4,1,2), SARIMA(3,1,3)(1,1,1)_[12], SARIMAX(3,1,2)(0,1,1)_[12] 차수의 조합이 가장 낮은 AIC 값을 기록하였다. 한편 외생변수가 포함되지 않은 모형의 경우 BIC 기준에서는 복잡도에 대한 패널티가 상대적으로 크게 적용되며 단순한 구조의 모형이 최적 차수로 식별되었다. BIC는 모형의 자유도에 대한 패널티를 AIC보다 더 크게 부여하기 때문에, 복잡도가 높은 모형은 상대적으로 불리하게 평가되는 경향이 있다. 반면 AIC는 예측력을 더 중시하는 경향이 있으며, 모형의 설명력과 데이터 적합도를 종합적으로 고려하여 적절한 복잡도를 허용한다.(Hyndman and Athanasopoulos, 2018) 따라서 미래값을 예측하는 것이 목표인 경우, AIC 기준으로 모형을 선택하는 것이 더 타당하기 때문에 본 연구에서는 AIC가 가장 낮은 파라미터 조합을 선택하였다.

6. 진단 및 예측

<표15>는 앞서 수행한 모형 식별 및 추정 과정을 통해 도출된 ARIMA, ARIMAX, SARIMA, SARIMAX 각각의 계수 추정 결과와 예측 정확도를 비교하여 모형의 적합성을 종합적으로 분석한 결과이다. 테스트 기간의 실제 관측값과 모형을 통한 예측값의 평가를 위한 주요 지표로 평균절대백분오차(MAPE), 평균제곱근오차(RMSE)를 활용하였으며 이를 통해 각 모형의 성능 차이를 정량적으로 비교하였다. 추가로 잔차에 대한 Ljung-Box 검정을 수행하여 모형에서 산출된 잔차가 백색잡음(white noise)의 특성을 만족하는지 확인하였으며, 외생변수가 포함된 모형의 경우 분산팽창지수(VIF)를 통해 다중공선성 문제를 점검하여 모형의 적합성을 제시하였다.

ARIMA는 외생변수를 포함하지 않은 단일 시계열 모형으로서 예측오차(MAPE)는 3.79%로 비교적 높은 값을 나타냈다. 잔차에 대한 Ljung-Box 검정 결과 p값은 0.587으로 귀무가설을 기각할 수 없어 통계적으로 백색잡음의 특성을 가진다고 판단할 수 있다. 다만 계수 추정 결과 일부 AR(2)와 AR(4), MA(2) 항이 통계적으로 유의하여 설명력이 제한적이고 예측 성능은 MAPE는 3.79%, RMSE는 0.454로 가장 낮은 수준에 머물렀다.

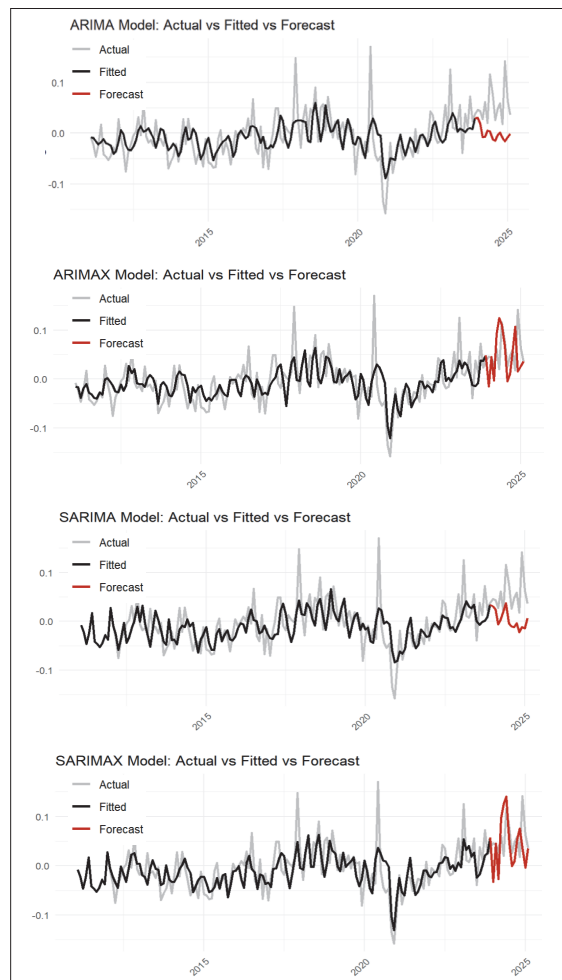
ARIMAX는 외생변수가 포함됨으로써 AR(1)~AR(3), MA(1)~MA(2) 계수가 모두 1% 수준에서 유의하게 추정되었다. 또한 MAPE는 0.60%, RMSE는 0.065로 예측력이 크게 향상되어 네 가지 모형 중 가장 낮은 값을 기록하였고, 잔차에 대한 Ljung-Box 검정결과 0.576로 백색잡음을 나타내었다. 외생변수 중 가격은 5% 유의수준, 공사비지수, 주택착공실적, 건설업GDP는 1% 유의수준에서 유의한 음(-)의 계수를 보여, 해당 지표가 상승할 경우 준공 후 미분양 주택 증가율을 감소시킨다고 나타났다. 또한 외생변수의 VIF는 1.20 이하로 다중공선성 문제는 없는 것으로 나타났다.

SARIMA는 ARIMA에 계절성이 포함된 모형으로서 AR(1)~AR(3), MA(1)~MA(3) 계수가 모두 1% 유의수준에서 유의성을 보였으나 SAR(1)은 유의하지 않아 계절성은 미미한 것으로 나타났다. MAPE가 3.31%, RMSE는 0.405로 ARIMA보다 개선되었지만 비교적 높은 값을 나타냈고 Ljung-Box 검정결과 p값은 0.249로 잔차가 백색잡음을 나타냈다.

SARIMAX는 외생변수와 계절성 모두 반영된 모형으로서 AR(1)~AR(3), MA(1)~MA(2), 계절 이동평균항 SMA(1)이 모두 1% 유의수준에서 유의하였다. MAPE가 0.78%, RMSE는 0.095로 단일 시계열 모형보다 우수한 성능을 보였고, Ljung-Box 검정결과 p값은 0.391로, 잔차가 백색잡음이라는 점에서 안정성과 적합성 모두 확보하였다. 다만 QS검정결과 계절성이 없고 실제 계절효과가 있더라도 영향력이 미미하기 때문에 ARIMAX보다 예측력은 낮게 나타났다.

결과적으로 ARIMAX 모형의 예측력과 잔차 진단(Ljung-Box) 등을 종합적으로 고려하였을 때 가장 우수한 성과를 보여 최적의 모델로 판단된다. <그림 7>은 각 모형에 대해 실제값(Actual), 적합값(Fitted), 예측값(Forecast)을 시각화한 결과이다.

<그림 7> 모형별 적합 및 예측 결과



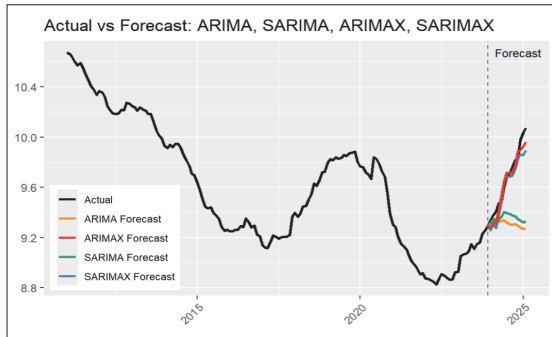
〈표 15〉 모형별 추정계수 및 예측성능 비교 결과

Model	MAPE/ RMSE	Box Test (p_value)	VIF	Coefficients			
				Param	Coef.	Std Err	p-value
ARIMA(4, 1, 3)	3.79%/ 0.454	0.587	-	ϕ_1	0.265	0.344	0.441
				ϕ_2	-0.524**	0.238	0.028
				ϕ_3	0.004	0.325	0.988
				ϕ_4	0.351**	0.154	0.023
				θ_1	-0.006	0.362	0.986
				θ_2	0.856***	0.144	0.000
				θ_3	0.374	0.366	0.306
ARIMAX(4, 1, 2)	0.60%/ 0.065	0.576	-	ϕ_1	0.589***	0.084	0.000
				ϕ_2	-0.811***	0.093	0.000
				ϕ_3	0.347***	0.091	0.000
				ϕ_4	0.090	0.089	0.310
				θ_1	-0.437***	0.025	0.000
				θ_2	0.996***	0.033	0.000
			1.110	price	-0.878**	0.406	0.031
			1.203	C_index	-1.558***	0.484	0.001
			1.161	begin	-0.009***	0.004	0.026
			1.021	C_gdp	-1.415***	0.445	0.001
1.042	rate	-0.023	0.016	0.143			
1.056	hf	-0.498*	0.295	0.091			
SARIMA(3, 1, 3)(1, 1, 1) _[12]	3.31%/ 0.405	0.249	-	ϕ_1	0.985***	0.117	0.000
				ϕ_2	-1.043***	0.061	0.000
				ϕ_3	0.708***	0.098	0.000
				θ_1	-0.733***	0.159	0.000
				θ_2	1.134***	0.069	0.000
				θ_3	-0.348**	0.161	0.031
				Φ_1	-0.036	0.097	0.708
				Θ_1	-0.999***	0.152	0.000
SARIMAX(3, 1, 2)(0, 1, 1) _[12]	0.78%/ 0.095	0.391	-	ϕ_1	0.667***	0.079	0.000
				ϕ_2	-0.894***	0.056	0.000
				ϕ_3	0.454***	0.082	0.000
				θ_1	-0.457***	0.025	0.000
				θ_2	0.999***	0.029	0.000
				Θ_1	-0.995***	0.136	0.000
			1.110	price	-0.774*	0.425	0.069
			1.203	C_index	-0.898	0.668	0.179
			1.161	begin	-0.018***	0.005	0.000
			1.021	C_gdp	-1.442***	0.450	0.001
			1.042	rate	-0.029*	0.017	0.098
			1.056	hf	-0.476	0.292	0.103

*** p < 0.01, ** p < 0.05, * p < 0.1

<그림 8>은 차분 예측값을 누적 합을 통해 복원한 결과를 나타낸 그래프이다. 검정 실선은 실제 관측된 미분양주택 수이고, 예측 구간 각 모형의 예측값을 포함하여 비교하였다. 앞서 지표에도 나타나듯이 ARIMAX 모형이 예측력이 가장 높은 결과를 나타냈다.

<그림 8> 차분 복원을 통한 모형별 예측값 비교

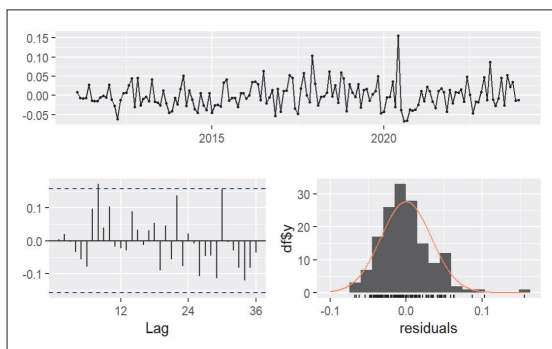


<그림 9>는 ARIMAX 모형의 잔차 진단결과를 나타낸 것으로 일정한 평균을 중심으로 무작위 분포하여 안정적인 모습을 보인다. 또한 ARIMAX 잔차에 대한 조건부 이분산성 여부를 확인하기 위하여 ARCH 효과 검정을 수행한 결과 5% 유의수준에서 귀무가설을 기각할 수 없어, ARCH 또는 GARCH 형태의 이분산성이 존재하지 않는 것으로 나타났다.

<표 16> ARCH-LM 테스트 결과

모형	lags	LM statistic	p-value
ARIMAX	3	2.3109	0.5104
ARIMAX	6	3.8587	0.6958
ARIMAX	9	5.0617	0.8289
ARIMAX	12	5.6682	0.9319

<그림 9> ARIMAX 모형의 잔차 진단 결과



7. 최적모형 기반 미래예측

예측력 평가를 통해 최종적으로 선정된 ARIMAX 모형을 토대로 모형의 구조는 유지하되 훈련데이터를 전체 기간으로 확장하고 Grid Search를 통해 파라미터를 재조정하여 향후 12개월(2025년 4월 ~ 2026년 4월) 준공 후 미분양주택 추이를 예측하였다.

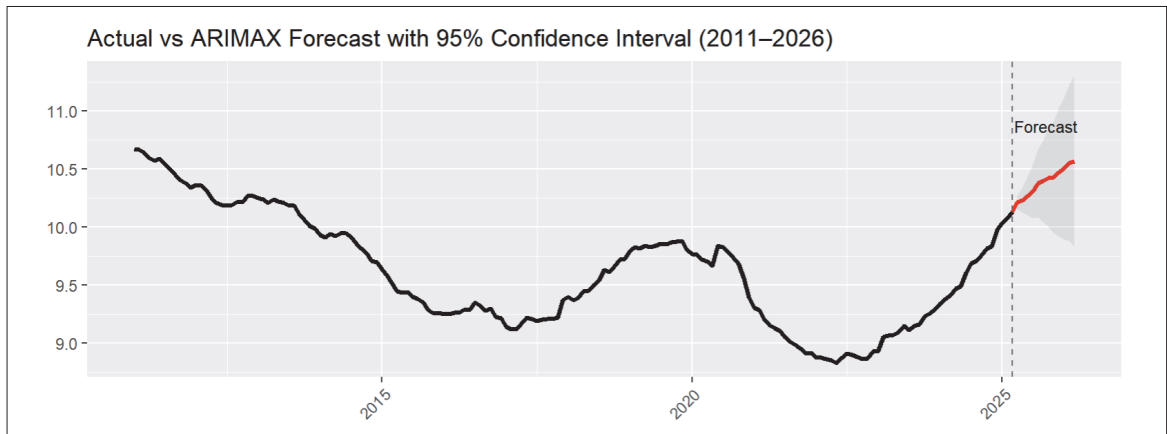
ARIMAX 모형은 외생변수를 반영하여 시계열을 예측하는 구조로, 미래값 예측을 수행하기 위해서는 외생변수의 미래값이 사전에 정의되어야 한다. 이에 따라 본 연구에서는 총 6개의 외생변수에 대해 각각 auto.arima() 함수를 활용하여 예측기간에 해당하는 미래값을 산출하였다. 이 함수는 다양한 모형을 추정하고 AIC 값이 가장 낮은 모형을 최적의 모형으로 결정해 준다. <표17>은 외생변수별 모형을 추정한 결과로 계절성을 포함하거나 포함하지 않은 다양한 구조의 모형이 선택되었다. AIC와 BIC 기준으로도 적절한 적합도를 나타냈고 잔차의 1차 자기상관은 0.37 이하로 낮게 나타나, 잔차의 독립성도 확보되었다.

<표 17> 외생변수별 적용모형

구분	model	AIC	BIC	ACF(1)
price	ARIMA(1, 1, 0)	-1323.16	-1316.89	-0.001
C_index	SARIMA(1, 1, 1)(0, 1, 1) _[12]	-1302.92	-1290.67	0.022
begin	SARIMA(2, 1, 1)(2, 0, 0) _[12]	195.01	213.83	0.026
C_gdp	SARIMA(4, 1, 0)(2, 0, 1) _[12]	-1354.73	-1329.64	0.037
rate	ARIMA(2, 1, 0)	-178.26	-178.11	-0.004
hf	SARIMA(1, 1, 1)(2, 0, 0) _[12]	-1187.09	-1171.41	-0.016

<그림 10>은 ARIMAX 모형을 기반으로 예측된 준공 후 미분양주택의 시계열 그래프와 95% 신뢰구간을 나타낸 그래프이다. 검정색 실선은 2011년 1월부터 2025년 3월까지 실제값을, 빨간색 실선은 2025년 4월부터 2026년 4월까지 예측값을 나타내며, 회색 음영은 95% 신뢰구간을 의미한다. 예측 결과 준공 후 미분양주택 수는 전반적으로 최근 몇 년간의 상승세가 지속될 가능성을 시사한다. 예측 구간의 폭은 시간이 지남에 따라 점차 확대되며, 이는 시계열 예측의 일반적인 불확실성 증가 현상으로 볼 수 있다. 또한 Box-Ljung 검정 결과 p-value 0.5766으로 잔차가 백색잡음으로 나타나, ARIMAX 모형이 적절하게 시계열의 구조를 설명하고 있음을 나타냈다.

<그림 10> 준공 후 미분양 주택 예측



V. 결론

1. 연구 요약 및 시사점

본 연구는 주택시장의 비효율성과 수요공급의 괴리를 가장 명확하게 보여주는 지표인 준공 후 미분양주택을 중심으로 거시경제 요인과 시계열 특성을 고려하여 향후 추세를 정량적으로 분석하였다.

2011년 1월부터 2025년 3월까지 구축된 월별 데이터를 기준으로 준공 후 미분양주택을 예측 대상인 종속변수로 설정하였고, DW 모형과 선행연구에서 미분양주택 발생에 유의미하게 영향을 미치는 요소를 포함하여 ARIMA, ARIMAX, SARIMA, SARIMAX 모형을 구축하고 예측력을 MAPE, RMSE 등의 객관적인 지표로 평가하였다.

모든 시계열 데이터는 로그변환, 차분 등의 전처리 과정을 거쳐 단위근을 제거하여 정상성을 확보하였고, 모델의 파라미터는 ACF/PACF 그래프를 토대로 후보범위를 설정하고 Grid Search로 모든 경우의 수를 전수조사 하여 AIC 기준 최적의 파라미터를 추정하였다. 또한 외생변수의 미래값은 단순 추세값이 아닌, 각 변수별 `auto.arima()` 함수를 활용하여 산출하였다.

시계열모형의 예측력 평가를 통해 가장 예측 오차가 낮은 ARIMAX 모형을 최종적으로 선정하여 이 모형을 기반으로 향후 12개월 미분양주택 추이를 예측한 결과, 단기간 내 뚜렷한 해소 없이 증가세가 지속될 것으로 나타났다. 이러한 결과는 기 착공된 주택의 준공

전환, 지역 간 불균형, 부동산시장 변동성과 같은 다양한 요인이 복합적으로 작용하였다고 볼 수 있다. Box test, ARCH-LM test 등의 모형의 진단 측면에서도 자기상관 및 이분산성이 존재하지 않아 모형의 적합성을 나타냈다.

이와 같은 결과로 단기적 조치보다 근본적이고 구조적인 해결방안이 필요하다. 특히, 준공 후 미분양 물량이 지속적으로 누적될 경우 시장 전반에 연쇄적인 위험으로 작용할 수 있다는 점에서 사전 대응이 중요하다. 정부는 2025년 3월 3천호의 준공 후 미분양주택 매입을 추진하였으나, 이는 전체 물량의 약 15%에 불과하여 증가추세를 고려할 때 효과는 제한적일 수밖에 없다. 따라서 본 연구에서 제시한 시계열 기반 예측모형은 미분양 문제를 조기에 진단하고 해소 방안을 마련하는 데 유용하게 활용될 수 있을 것이다.

2. 연구의 한계 및 향후 과제

본 연구는 시계열 예측모형을 통해 준공 후 미분양주택의 향후 추이를 실증적으로 분석하였으나, 적용범위 등 몇 가지 측면에서 한계가 존재한다. 먼저 자료의 구성으로 외생변수로 활용된 거시경제 변수가 전국단위로 수집된 통계자료이기 때문에, 지역 특성을 고려한 분석은 수행되지 못하였다. 미분양은 특히 지역 편차가 크고, 지방 중소도시에서 집중적으로 나타나는 경향이 있음에도 전국 기준으로 분석이 이루어졌다. 또한 본 연구에 사용된 시계열모형은 선형 기반의 예

측모형이기 때문에 정책 변화나 수요가 급변하는 외부 충격에 대해서는 즉각적으로 반응하지 못하는 한계가 있다.

이와 같은 한계에도 본 연구는 시장에 미치는 영향이 큰 준공 후 미분양에 초점을 맞추어 실증적으로 예측하였다는 점에서 의의가 있으며, 단기 예측을 통해 선제적 대응의 필요성을 제시하였다. 향후 본 연구에서 언급한 한계점을 보완하여 지역별 특성을 반영한 분석이 필요하고, 비선형을 고려한 모형을 통해 예측 정확도를 고려하여 후속연구가 확장될 필요가 있다.

논문접수일 : 2025년 8월 4일

논문심사일 : 2025년 9월 8일

게재확정일 : 2025년 10월 1일

참고문헌

- 강병주, “대전지역 주택 미분양 실태와 대책”, 「주택연구」, 3(1), 한국주택학회, 1995, pp. 5-24
- 국토교통부, “지방 준공 후 미분양 주택 3천호 매입 개시”, 국토교통부 보도자료, 2025. 03. 21
- 국토교통부, 「미분양주택현황보고 : 통계정보보고서」, 국토교통부, 2021
- 김대원·유정석, “패널분석을 이용한 지역별 미분양률 결정모형 개발 및 적용 연구”, 「주택연구」 제22권 제1호, 한국주택학회, 2014, pp. 149-178
- 김덕례·황은정, “미분양 해소를 위한 정부정책의 영향 분석과 향후 정책방향 연구”, 「주택산업연구원 연구보고서」 2009-2, 주택산업연구원, 2009, pp. 1-74
- 김리영·서원석, “벡터오차수정모형을 이용한 하위시장별 주택 미분양 영향요인 분석: 경기도와 경상남도의 분양 및 재고 시장을 중심으로”, 「감정평가학논집」 제19권 제1호, 한국감정평가학회, 2020, pp. 75-100
- 김상기·이상호·김재준, “주택매매가격 및 전세가격과 미분양주택량의 관계성 분석”, 「대한건축학회논문집」 제26권 제1호, 대한건축학회, 2010, pp. 278-285
- 김주영·김형순, “미분양 아파트 결정구조에 관한 연구”, 한국지역개발학회 세미나 논문집, 한국지역개발학회, 2010-3, pp. 91-104
- 김태섭·김은덕, “주택 미분양 발생요인과 도시별 특성분석”, 「주택포럼」 제1호, 주택산업연구원, 1996, pp. 109-114
- 김현아, “최근 민간 주택 미분양 증가의 원인과 해결방안”, 「한국건설산업연구원 연구보고서」 제21호, 한국건설산업연구원, 2007, pp. 1-34
- 남희용·김태섭, “미분양 아파트 특성 분석”, 「주택산업연구원 연구보고서」 제7호, 주택산업연구원, 2001, pp. 1-56
- 박재룡·유정석, “미분양 주택의 적체완화를 위한 정책개선방안에 관한 연구”, 「부동산연구」 제20권 제1호, 한국부동산연구원, 2010, pp. 263-280
- 박정수·정상철·성주한, “고정효과모형을 이용한 통합창원시 미분양 아파트의 결정요인 변화에 관한 연구”, 「부동산학보」 제80집, 한국부동산학회, 2020, pp. 52-67
- 박천규·이영, “주택시장 체감지표의 주택시장지표 예측력 분석”, 「부동산학연구」 제16권 제1호, 한국부동산분석학회, 2010, pp. 131-146
- 서범준·이효중·정창무, “지역별 미분양 아파트 규모의 결정요인과 조정속도에 대한 실증분석”, 「국토계획」 제45권 제7호, 국토연구원, 2010, pp. 97-117
- 성주한·정상철, “서울과 지방의 아파트 매매가격에 미치는 영향 요인에 관한 연구”, 「부동산학보」 제78집, 한국부동산학회, 2019, pp. 92-103
- 양영준·유선종, “제주특별자치도 주택 미분양률의 결정요인에 관한 연구”, 「부동산학보」 제76집, 한국부동산학회, 2019, pp. 92-106
- 양지영, “미분양주택 결정요인 분석”, 건국대학교 박사학위논문, 건국대학교 대학원, 2025
- 이근철·최성훈, “수요감소 요인 외생변수를 갖는 SARIMAX 모형을 이용한 관광수요 예측”, 「한국산업경영시스템학회지」 제43권 제4호, 한국산업경영시스템학회, 2020, pp. 59-66
- 이성근·전광섭, “주택건설업체 부도위험 연구”, 「부동산학보」 제39집, 한국부동산학회, 2009, pp. 302-315
- 이의준·김경민, “준공 전 또는 준공 후 미분양에 영향을 미치는 요인과 차이점: 수도권 시·군·구를 중심으로”, 「국토연구」 제70집, 국토연구원, 2011, pp. 95-110
- 이중희, “신축주택의 분양시기에 관한 효과분석 및 개선방안”, 「주택연구」 제3권 제2호, 한국주택학회, 1995, pp. 5-28
- 이형, “미분양 적체와 주택건설사의 유동성 문제”, 「HOUSING ISSUE」 제35호, 주택산업연구원, 2008, pp. 1-2
- 임덕호, “주택시장 정상화 방안”, 「주택연구」 제17권 제2호, 한국주택학회, 2009, pp. 257-282
- 장세웅, “미분양아파트 발생요인 분석에 관한 연구”, 한양대학교 박사학위논문, 한양대학교 대학원, 2011
- 전해정, “미분양주택량 결정 요인에 관한 실증분석”, 「부동산학보」 제56집, 한국부동산학회, 2014, pp. 110-122
- 정상철·성주한, “미분양 아파트 결정모형에 관한 연구: 통합창원시 사례를 중심으로”, 「부동산학보」 제76집, 한국부동산학회, 2019, pp. 61-77
- 정성훈·김주영, “미분양주택 해소를 위한 금융구조화 연구”, 「부동산학보」 제41집, 한국부동산학회, 2010, pp. 58-68
- 정재형, “신규주택 미분양에 관한 연구”, 서울대학교 환경대학원 석사학위논문, 서울대학교 환경대학원, 2000
- 정창무·김지순, “주택시장에서의 미분양 아파트의 역할에 대한 실증분석”, 「국토계획」 제40권 제2호, 국토연구원, 2005, pp. 81-91
- 최재규·유승규·김재준, “미분양주택 발생과 건설업체 부실화와 관계성 분석”, 「한국건축학회논문집」 제29권 제6호, 한국건축학회, 2013, pp. 89-96
- 허재완·김은경, “미분양주택의 시기별 분포특성 및 발생배경 비교”, 「부동산연구」 제19권 제2호, 한국부동산연구원, 2009, pp. 259-278
- 허재완·손성민, “미분양주택 정책의 지역별 효과에 관한 실증분석”, 「국토계획」 제48권 제1호, 국토연구원, 2013, pp. 131-147
- 황관석·김지혜·이운상, “미분양주택 변동원인과 대응방향 연구”, 「국토연구원 수시연구보고서」 수시 23-07, 국토연구원, 2023
- Box, George E. P and Gwilym M. Jenkins, Time Series Analysis: Forecasting and Control, Holden-Day, 1970
- Box, George E. P, Gwilym M. Jenkins, and Gregory C. Reinsel, Time Series Analysis: Forecasting and Control (3rd ed.), Prentice Hall, 1994
- Cleveland, Robert B and William S. Cleveland, Jean E. McRae and Irma Terpenning, “STL: A Seasonal-Trend Decomposition Procedure Based on Loess,” Journal of Official Statistics, Vol. 6 No. 1, 1990, pp. 3-73

38. DiPasquale, D. and Wheaton, W. C., "The markets for real estate assets and space: A conceptual framework", Journal of the American Real Estate and Urban Economics Association, 20(1), American Real Estate and Urban Economics Association, 1992, pp.181-197
39. Hyndman, Rob J and George Athanasopoulos, Forecasting: Principles and Practice (2nd ed.), OTexts, 2018
40. LJUNG, Greta M and BOX, George EP. On a measure of lack of fit in time series models. Biometrika, 1978, 65(2), pp. 297-303
41. U.S. Census Bureau, X-13-ARIMA/SEATS Reference Manual Version 1.1, Statistical Research Division, U.S. Census Bureau, 2013

<국문요약>

시계열모형을 활용한 준공 후 미분양주택 예측

강 태 영 (Kang, Tae-Young)
이 상 근 (Lee, Sang-Gun)

본 연구는 현재 주택시장에서 심각한 리스크로 부각되고 있는 준공 후 미분양 주택의 향후 추이를 시계열모형을 통해 계량적으로 예측하였다. 준공 후 미분양은 선분양제도로 인한 수요예측 실패, 거시경제 및 금융시장 변화, 공급전략 실패 등이 복합적으로 작용해 발생하며 특히 국내 주택건설 사업은 PF(Project Financing) 기반으로 추진되어 시행주체의 자금 경색과 금융권 부실로 직결되는 구조적 한계를 갖고 있다.

이에 본 연구는 DiPasquale-Wheaton(DW) 4분면 모형의 이론적 틀을 기반으로 수요·공급의 시차와 재고 누적 구조를 설명하고 이를 시계열모형을 통해 실증적으로 분석하였다. 2011년 1월부터 2025년 3월까지의 월별 데이터를 대상으로, 민간부문 준공 후 미분양 주택 수를 예측 대상인 종속변수로 설정하였고, 기존 연구에서 미분양주택 발생에 영향을 미치는 요소로 나타난 외생변수를 포함하여 ARIMA, ARIMAX, SARIMA, SARIMAX 모형의 예측력을 비교하였다.

ACF/PACF로 모형을 식별하여 후보 파라미터를 설정하였고, Grid Search를 통해 AIC 기준 최적의 파라미터로 모형을 추정 후 MAPE와 RMSE로 예측 성능을 비교하여 최종 모형을 결정하였다. 실증분석 결과 외생 변수를 포함한 ARIMAX 모형이 가장 우수한 성능을 보였으며, 잔차 진단결과 백색잡음을 충족하였고, ARCH 또는 GARCH 형태의 이분산성도 존재하지 않는 것으로 나타났다.

최종적으로 선정된 ARIMAX 모형을 기반으로 향후 1년간 준공 후 미분양 추이를 예측한 결과, 단기간 내 뚜렷한 해소 없이 지속적인 증가세를 보일 것으로 나타나 선제적 대응과 구조적 처방의 필요성이 더욱 강조된다.

주 제 어 : 준공 후 미분양주택, 시계열분석, 예측모형 비교, ARIMAX, 미분양 예측