## 2.1 Introduction

Under the dual background of the accelerated global urbanization process and the reconstruction of the geopolitical landscape, the real estate market, as the core hub connecting macroeconomic stability and micro people's livelihood well-being, its fluctuation prediction has gone beyond the traditional economic category and has become an interdisciplinary proposition involving urban sociology, environmental humanism and digital ethics. According to the International Monetary Fund, real estate assets account for approximately 65% of global household wealth. The fluctuations in their prices not only affect the vulnerability of the financial system but also profoundly shape the trajectories of population migration, community governance models, and even changes in cultural identity. The reconstruction of housing supply and demand in Eastern Europe triggered by the Ukraine war in 2022 has confirmed the humanistic attribute of housing resources as a "social stabilizer". (Marcin Bas,2024) However, the traditional prediction paradigm relies on lagging macro indicators and static measurement models, making it difficult to cope with the nonlinear shocks of new risk factors such as climate change and sudden geopolitical crises. This led to decision-making failures in more than 37% of major cities worldwide from 2019 to 2023, with a housing price prediction error rate exceeding 20% (Chuan Zhao &Fuxi Liu, 2023). In this context, the rise of big data-driven technology marks a paradigm shift in real estate research from "experience-dependent" to "intelligent responsive". The urgency of this transformation stems from three realistic challenges: First, the exponential growth of multi-source heterogeneous data (satellite remote sensing, social media public opinion, IoT sensors, etc.) requires the construction of a new information extraction framework; Secondly, the integration of machine learning and spatial econometric technology enables researchers to decode complex correlations that traditional methods cannot capture (such as the spatial coupling effect between air pollution and housing price discounts). Thirdly, the prediction model needs to take into account both accuracy and transparency to respond to the majority of humanistic and ethical demands.

## 2.2 Definition of Research Influencing Factors

Therefore, we need to determine based on literature, ai and search engines which factors affect the real estate market and conduct systematic analysis as variables. After investigation, it was found that there are approximately a dozen variable factors influencing the real estate market. Among them, factors such as the economy, the stock market, and national policies almost cover all fields around the world. Regional issues are widespread but disorderly and complex, while factors like earthquakes and wars are accidental and cannot be generalized. Therefore, the author classified the following factors into three categories and conducted research: 1. Economic factors 2. Social factors 3. Regional factors. (Maria et al,2025)

## 2.3 Theoretical Background

The prediction of price fluctuations in the real estate market is essentially a complex and systematic issue involving multiple disciplines. Its theoretical foundation needs to integrate the cutting-edge achievements of economics, data science, spatial geography and policy research. The following is the construction of the analytical framework from the core theoretical level:

### Core theoretical framework

#### 2.3.1.1 Spatial Economics and Risk Spillover Theory

Based on Krugman's new economic geography, fluctuations in the real estate market have significant spatial dependence and network spillover effects. Geographically adjacent areas (such as within urban agglomerations) form risk contagion channels through capital flows, population migrations and information diffusion. For instance, the air pollution spillover in the Yangtze River Delta urban agglomeration of China, through the "environmental quality discount - Capital reallocation" mechanism (Yi Fang et al,2024), causes the real estate systemic risk index (RPI) of adjacent cities to be linked, with a spatial elasticity coefficient as high as 1.24. This theory provides the basis for constructing the "spatial weight matrix" for cross-regional risk modeling and explains the geographically sensitive gradient phenomenon that the decline in house prices in border cities of Poland (-34%) during the Ukraine war was significantly higher than that in the central area (-12%). (Marcin Bas, 2024)

#### 2.3.1.2 Theory of Explainability in Machine Learning

The interpretability of artificial intelligence (XAI) is the key to cracking the "black box" of predictive models. The SHAP value (Shapley Additive Explanations) based on cooperative game theory reveals the nonlinear interaction effect in house price prediction by quantifying the contribution degree of features. For instance, in the valuation of townhouses in Virginia, the United States, the interaction contribution rate of the SHAP value between "subway distance ≤500 meters" and "quality school district" reached 27% (Byeonghwa Park & Jae Kwon Bae,2015), and this finding challenged the assumption of the linear sum of traditional features. Meanwhile, the permutation Feature Importance (PFI) technique can screen core variables. For instance, in Reference One, 12 key contamination indicators (accounting for 27.9% of the original variable set) were identified through PFI, reducing the model complexity by 40% while maintaining an accuracy of over 95%.

#### 2.3.1.3 Dynamic Adaptation Theory of Policy Tools

According to Tinbergen's economic policy principles, policy tools need to be dynamically matched with policy goals. One of the literatures proposed the three-dimensional framework of "money - tax - macroprudential", indicating that the short-term elasticity of monetary policy on house prices through the interest rate channel (0.78) is significantly higher than that of tax policy (0.42), but the latter has long-term sustainability in suppressing speculative demand. The combination of this theory and spatial economics has given rise to the "policy tool - market maturity" matching model. For example, the housing price suppression effect of property tax policies in developed countries (-15%) is 60% stronger than that in emerging markets (-9.4%), revealing the moderating role of the institutional environment on policy effectiveness. (Chuan Zhao & Fuxi Liu, 2023)

### 2.4 Predictive Models for Real Estate Market Analysis

Advanced prediction models are developed by using complex machine learning and deep learning, which effectively handle the massive data of real estate and the linear and nonlinear characteristics related to certain factors. For instance, the EMAPO model (Graham Squires a & Erwin Heurkens, 2015) can reveal the complexity of real estate development from multiple dimensions, but its accuracy is not stable. The vector auto-regression (VAR) model can achieve the same. Presenting multivariate dynamic relationships and predictions with time series data also requires the Granger causality test to ensure its practicability. (Amirouche Chelghoum et al,2025) The causal inference ability of the Difference-in-Differences model is slightly stronger and it is more suitable for project analysis. (Huadun Chen et al,2023) Machine learning is also a choice suitable for big data and complex features. Although its inference ability is slightly insufficient, its accuracy can still be guaranteed if multiple linear regression methods are combined. (Marcin Hernes et al,2024)

### 2.5Gaps in the Literature and Research Opportunities

Previous studies on data analysis in the real estate market have been tested and conducted in the literature. However, with further research and development scope, some limitations can be noted. (Mariia, 2021) First of all, only a very small number of literatures focus on the influence of multiple factors on real estate, and most studies are conducted on a single variable or in a single region. Although patterns can be inferred from it, it is not known whether they are applicable to other regions. Secondly, in multivariate testing, the accuracy rate cannot be guaranteed, and the differences existing among various models are not mentioned, which makes it difficult to ensure its practicality when put into the market. Therefore, it is crucial to determine a relatively simple and effective prediction method. Even if overly complex models can guarantee accuracy, they may not be put on the market to bring convenience to the people due to hardware requirements. The research in the literature provides a systematic framework for the academic comparison and policy practice of international real estate development, and offers a way for future research. However, it still needs to be carried forward to determine an effective research method to bring valuable predictions to investors, policymakers and the general public. (Petros et al, 2024)( Shu-hen Chiang & Chien-Fu Chen,2023)

**References**


1.Marcin Bas. (2024). The impact of the war in Ukraine on the residential real estate market on the example of Szczecin, Poland. Procedia Computer Science, 246,3004-3013. https://doi-org.ezproxy.utm.my/10.1016/j.procs.2024.09.371

2. Nafeesa Yunus.(2025). Effects of oil shocks on global securitized real estate markets. Finance Research Letters,80. https://doi-org.ezproxy.utm.my/10.1016/j.frl.2025.106871

3. Huaying Gu, Zhixue Liu, Yingliang Weng. (2017) Time-varying correlations in global real estate markets: A multivariate GARCH with spatial effects approach. Physica A: Statistical Mechanics and its Applications, Volume 471, 460-472. https://doi-org.ezproxy.utm.my/10.1016/j.physa.2016.12.056

4. Waheed Ullah Shah, Ijaz Younis, Ibtissem Missaoui, Xiyu Liu. (2025). Environmental transitions effect of renewable energy and fintech markets on Europe's real estate stock market. Renewable Energy, Volume 243, 122603. https://doi-org.ezproxy.utm.my/10.1016/j.renene.2025.122603

5. Yi Fang , Yanru Wang , Yan Yuan , Moyan Zhang. (2024). Urban air pollution and systemic risk of the real estate market in China. International Review of Economics & FinanceVolume 96, Part B,103626. https://doi-org.ezproxy.utm.my/10.1016/j.iref.2024.103626

6. Federico Dell’Anna. (2025). Machine learning framework for evaluating energy performance certificate (EPC) effectiveness in real estate: A case study of Turin’s private residential market. Energy Policy Volume 198, 114407. https://doi-org.ezproxy.utm.my/10.1016/j.enpol.2024.114407

7. Chuan Zhao , Fuxi Liu.(2023). Impact of housing policies on the real estate market - Systematic literature review. Heliyon, Volume 9, Issue 10, e20704. https://doi-org.ezproxy.utm.my/10.1016/j.heliyon.2023.e20704

8. Huthaifa Alqaralleh , Alessandra Canepa , Gazi Salah Uddin.(2023). Dynamic relations between housing Markets, stock Markets, and uncertainty in global Cities: A Time-Frequency approach. The North American Journal of Economicsand Finance Volume 68, 101950. https://doi-org.ezproxy.utm.my/10.1016/j.najef.2023.101950

9. Marcin Hernes , Piotr Tutak , Mateusz Siewiera.(2024). Prediction of residential real estate price on primary market using machine learning. Procedia Computer Science Volume 246,3142-3147. https://doi-org.ezproxy.utm.my/10.1016/j.procs.2024.09.358

10. İsmail Canöz, Hakan Kalkavan. (2024). Forecasting the dynamics of the Istanbul real estate market with the Bayesian time-varying VAR model regarding housing affordability Habitat International Volume 148, 103055. https://doi-org.ezproxy.utm.my/10.1016/j.habitatint.2024.103055

11. Michel Ferreira Cardia Haddad , Bo Sjö , David Stenvall, Gazi Salah Uddin, Anupam Dutta. (2024). Interconnectedness between real estate returns and sustainable investments: A cross-quantilogram and quantile coherency approach. Journal of Cleaner Production, Volume 479, 144085. https://doi-org.ezproxy.utm.my/10.1016/j.jclepro.2024.144085

12. Mohd Shahril Abdul Rahman Mariah Awang Zainab Toyin Jagun,(2024) Polycrisis: Factors, impacts, and responses in the housing market. Renewable and Sustainable Energy Reviews Volume 202, 114713. https://doi-org.ezproxy.utm.my/10.1016/j.rser.2024.114713

.