PhD in Business Economics

"A Machine Learning and Explainable AI Approach to Predict and Explain Belgian Residential Real Estate Rent."

Ian Dave Lenaers

Promotor: prof. dr. Lieven De Moor

Summary

This dissertation develops an interpretable, data-driven framework for analyzing residential rental prices in Belgium by combining machine learning (ML) models with explainable artificial intelligence (XAI) techniques. Studying the dynamics of rental prices is becoming increasingly important due to concerns about housing affordability, declining homeownership rates and the growing policy focus on energy efficiency. Traditional hedonic models, such as Ordinary Least Squares regression and related econometric methods, offer interpretability, but without explicit model specification they are limited in their ability to capture nonlinearities, interaction effects and spatial dependencies that characterize real estate markets. ML models, on the other hand, offer better predictive performance but have a black-box nature. This dissertation deals with this trade-off by combining ML and XAI to improve both predictive performance and interpretability.

The research pursues three main objectives. First, ML models for predicting rental prices are compared. Second, the study shows how XAI techniques can be used to examine complex black-box models. Third, the role of energy efficiency in the Belgian rental market is analyzed, including its development over time.

Using large-scale datasets of Belgian residential rental listings, the first part of the dissertation evaluates 28 predictive models. The results show that, among others, tree-based ML models consistently outperform linear models. However, predictive performance varies across the distribution of rental prices, with larger errors observed in the bottom and top deciles.

In the second part, multiple XAI techniques are applied and compared, including shapley additive explanations, permutation feature importance, partial dependence plots and accumulated local effects plots. The research confirms that rental prices are driven by complex, nonlinear relationships and interactions among structural characteristics and location factors. Furthermore, the results show that different XAI techniques deliver complementary knowledge of model behavior, which promotes their combined use for a comprehensive interpretation.

The third part focuses on energy efficiency. A comparison of classical econometric models shows that the estimated effects of energy efficiency are highly sensitive to model choices. In an extension of this analysis, ML and XAI techniques are used to examine temporal changes in energy efficiency effects between 2014 and 2023. The results show that the importance of energy efficiency has increased over time, although its impact on rental prices remains modest compared to other determinants.

This dissertation contributes to the literature in three ways. First, it offers a systematic comparison of ML models to predict residential rents. Second, it shows how XAI can be used not only as a transparency tool but also as an analytical tool to uncover market dynamics. Third, it includes a temporal analysis of the drivers, particularly energy efficiency, behind rental prices using ML and XAI, thus enhancing the policy relevance of the research.

These results are relevant for policy makers, investors and real estate professionals who wish to understand the dynamics of the rental market and they demonstrate the added value of combining ML and XAI to develop accurate, interpretable and policy-relevant models for real estate analysis.