Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

AI-Driven Housing Affordability Forecasting in New York City:An NTA-Level Panel Analysis Using Ensemble Machine Learning

Azizur Rahman

azizurusa22@gmail.com

Indiana Wesleyan University

Nakib Uddin Ahmed

Metropolitan University

Mehjabin Ferdous

St. Francis College

Kadirur Rahman Chowdhury

Birmingham City University

Research Article

Keywords: Housing Affordability, Machine Learning, SHAP Explainability, New York City, Neighborhood Tabulation Area

Posted Date: May 8th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9348598/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Applied Spatial Analysis and Policy on July 25th, 2026. See the published version at https://doi.org/10.1007/s12061-026-09955-4.

AI-Driven Housing Affordability Forecasting in New York City: An NTA-Level Panel Analysis Using Ensemble Machine Learning

Abstract

Housing affordability is one of the signature governance challenges in New York City (NYC), with more than 52% of renter households cost-burdened and just under a third identified as severely cost-burdened, paying above 50% of gross income on housing. Here, we propose a neighborhood-scale, explain-machine-learning-based framework to predict the affordability of housing at the NTA level with 2,512 observations over time scaled to the Neighborhood Tabulation Area (NTA), comprising 239 NTAs from 2012–2022. We integrate 49 socio-economic, housing market, and rent indices features from ACS 5-year estimates, eviction records, and the Zillow Observed Rent Index (ZORI). Three gradient-boosted ensemble models—Random Forest, XGBoost, and LightGBM—are benchmarked under a strict temporal train/validation/test split with 5-fold TimeSeriesSplit cross-validation to prevent data leakage. LightGBM achieved the highest predictive performance (Test R^2=0.9258, RMSE=0.03477, MAE=0.0192). SHAP TreeExplainer analysis identifies rent burden, vacancy rate, and rent-to-income ratio as dominant drivers, confirming that NYC's rent escalation follows a systemic "burden escalator" across boroughs. Spatial diagnostics (Moran's I) show no significant autocorrelation in residuals, and ablation analysis confirms rental market variables drive the largest incremental contribution. Based on forecasts, we propose a 14-instrument policy matrix including early-warning interventions, borough-differentiated zoning and voucher strategies, and anticipatory resource allocation. This framework demonstrates how explainable AI can operationalize predictive insights into actionable urban governance strategies, offering a replicable model for global cities facing housing crises.

Keywords: Housing Affordability, Machine Learning, SHAP Explainability, New York City, Neighborhood Tabulation Area

1 Introduction

New York City (NYC) faces the United States' most severe housing affordability crisis. As of the 2022 American Community Survey (ACS), over 52% of renter households are cost-burdened—spending more than 30% of gross income on housing—and approximately 30% are severely cost-burdened, exceeding the 50% threshold set by the U.S.

Department of Housing and Urban Development (HUD). Severe rent burden forces households to trade housing security against basic necessities such as food, healthcare, and education. Rental costs in NYC have consistently outpaced renter incomes across every borough since at least the 1980s, with the gap widening sharply during and after the COVID-19 pandemic. Addressing this crisis requires not only accurate measurement but actionable, neighborhood-level intelligence to guide housing assistance resources toward the highest-risk communities before administrative data confirms deterioration.

The Neighborhood Tabulation Area (NTA) is the official sub-borough geographic unit designated by the NYC Department of City Planning (DCP), comprising approximately 15,000–40,000 residents per neighborhood. With 263 NTAs spanning four boroughs (Brooklyn, Queens, Manhattan, and the Bronx), this unit captures meaningful variation in housing market conditions while maintaining statistically reliable ACS sample sizes, making it the natural unit for neighborhood-targeted housing policy. Staten Island is excluded from this study due to limited ACS coverage and smaller sample reliability at the NTA level. Most existing machine learning applications to NYC housing operate at the borough or metropolitan level, leaving NTA-level panel structures and neighborhood-level policy targeting largely unexplored.

This paper addresses this gap through five primary contributions:

1. **Data construction:** We create a novel NTA-level panel dataset of 2,512 observations across 239 NTAs in four boroughs (2012–2022), integrating ACS 5-year estimates, NYC eviction court records, and the Zillow Observed Rent Index (ZORI).
2. **Model benchmarking**: We rigorously evaluate three gradient-boosted ensemble models—Random Forest, XGBoost, and LightGBM—under a strict temporal train/validation/test split with 5-fold TimeSeriesSplit cross-validation to prevent data leakage.
3. **Explainable ML analysis:** Using SHAP TreeExplainer, we generate global and borough-level feature importance rankings, including the first borough-comparative SHAP decomposition showing that severe rent burden follows a uniform structural escalation mechanism across all four boroughs.
4. **Spatial diagnostics and forecasting:** We conduct Moran's I(Ref) analysis on model residuals and produce three-year borough-level affordability forecasts (2023–2025) under a calibrated baseline macroeconomic scenario.
5. **Policy translation:** We propose a 14-instrument policy recommendation matrix, anchored to SHAP attributions, forecast trajectories, and model precision evidence, identifying early intervention at the 30% burden threshold, borough-differentiated supply and income-side responses, and NTA-level predictive targeting to operationalize anticipatory housing policy in NYC.

2 Related Work

2.1 Housing Affordability: Measurement and Theory

The 30%-of-income threshold for housing affordability was codified in the National Housing Act of 1937 and subsequently reinforced through HUD administrative rules

governing federal housing assistance eligibility. Bratt and Rachel G [1] challenged this fixed-ratio approach with the shelter poverty framework, arguing that affordability must be evaluated against the residual income available for non-housing necessities after rent is paid -a standard that is more stringent for low-income households and more permissive for high-income households than the uniform 30% threshold. Quigley and Raphael [2] provided rigorous empirical evidence that restrictive land-use regulation is the primary structural driver of affordability deterioration in coastal cities including NYC, establishing the supply-constraint mechanism that underlies the rent growth dynamics modeled in this paper. Glaeser and Gyourko [3] estimated that zoning regulations impose implicit welfare costs equivalent to 50% of market rents in highly regulated cities, directly reducing housing supply elasticity and preventing the market-clearing response to demand increases. More recently, Albouy, Ehrlich, and Liu [4] demonstrated that place-based sorting amplifies affordability inequality: high-productivity workers outbid lower-income renters in desirable neighborhoods, concentrating severe burden in spatially identifiable clusters -a dynamic that motivates the NTA-level geographic disaggregation central to this analysis.

2.2 Ensemble Machine Learning for Housing Prediction

Gradient-boosted decision tree ensembles have emerged as the leading machine learning paradigm for property valuation and housing market prediction, consistently outperforming hedonic regression and neural network alternatives on structured tabular data. Antipov and Pokryshevskaya [5] demonstrated that gradient boosting outperforms hedonic regression for mass appraisal, achieving R^2=0.89 on Russian residential data. Park and Bae [6] applied Random Forests to Korean residential price forecasting, finding that ensemble methods substantially outperform multiple regression on out-of-sample data. Chen and Guestrin [7] introduced XGBoost, demonstrating its scalability and regularization advantages through the regularized objective

$$L(\phi) = \sum_i l(y_i, \hat{y}_i) + \sum_k \left[\gamma T_k + \frac{1}{2}\lambda \|w_k\|^2 \right] \tag{1}$$

where γ penalizes tree complexity and λ provides L2 weight regularization.

Ke et al. [8] introduced LightGBM with histogram-based leaf-wise tree growth, achieving substantially faster training on high-dimensional tabular data while matching or exceeding XGBoost accuracy on benchmark datasets -a performance advantage confirmed in the held-out test results in Sect 6. In the NYC housing context, Been, Ellen, and Madar [9] applied logistic regression to building-level eviction risk using administrative records, but without ensemble methods, NTA panel structure, or temporal cross-validation. This paper extends the ensemble ML literature to the specific context of NTA-level severe rent burden prediction.

2.3 SHAP Explainability in Applied Research

Lundberg and Lee [10] introduced SHAP (SHapley Additive exPlanations) as a theoretically grounded feature attribution framework rooted in cooperative game

theory, satisfying the axioms of local accuracy, missingness, and consistency that prior feature importance methods fail to jointly satisfy. TreeExplainer provides exact polynomial-time Shapley values for tree ensembles, enabling computationally efficient global and local explanations. The global importance measure in Eq 2 aggregates directionally-neutral marginal contributions across observations.

$$\bar{I}_j = \frac{1}{n} \sum_i |\phi_j(i)| \tag{2}$$

Abidoye and Chan [9] applied SHAP to residential valuation in Hong Kong, demonstrating the utility of SHAP for policy-relevant interpretation of housing market models. Molnar [11] provides a comprehensive treatment of SHAP's theoretical properties, practical interpretation guidelines, and limitations in the presence of highly correlated features -a consideration directly relevant to the multicollinearity between rental market features identified in the ablation study of Sec. 8. This paper presents the first application of SHAP TreeExplainer to NTA-level rent burden prediction in NYC, extending prior work by providing both global feature rankings and borough-level comparative decompositions.

2.4 Spatial Econometrics and Urban Panel Data

Anselin [12] formalized spatial dependence diagnostics through Moran's I and spatial lag and error regression models, establishing the methodological framework applied in Sect. 9. LeSage and Pace [13] demonstrated that ignoring spatial autocorrelation in panel residuals produces inconsistent parameter estimates under omitted spatially correlated confounders. Gibbons and Machin [14] showed strong positive spatial autocorrelation in London housing price gradients, consistent with neighborhood amenity spillovers that may operate at the sub-borough NTA level in NYC. In temporal panel contexts, Case and Shiller [15] demonstrated the importance of controlling for temporally persistent unobserved heterogeneity in housing panel data, motivating the temporal lag features. James et al. [16] provide the foundational treatment of cross-validation bias in temporal settings: standard k-fold cross-validation inappropriately uses future data to predict the past, requiring temporal splits or TimeSeriesSplit protocols -the precise rationale for the strict temporal partition and 5-fold TimeSeriesSplit cross-validation protocol described in Sect. 5.

2.5 Machine Learning for Housing Policy Targeting

An emerging strand of applied machine learning research has moved beyond predictive accuracy toward policy-operational deployment of housing market models. Desmond and Gershenson [17] demonstrated that eviction risk prediction models constructed from administrative court records can identify high-risk households with sufficient precision for targeted pre-eviction legal assistance programs, establishing the precedent for operationalizing predictive models as policy-targeting instruments. Glaeser et al. [18] applied machine learning to identify building-level housing code violation risk in Boston, demonstrating that model predictions can guide proactive inspection

prioritization with measurable improvements in violation detection rates. These applications share a common structure with the early-warning system: a predictive model trained on administrative data generates neighborhood-level risk scores that allow housing agencies to pre-position resources ahead of crisis manifestation. However, no prior study has applied this framework to NTA-level severe rent burden prediction in NYC, integrated SHAP attribution with borough-level forecasting to produce policy-differentiated recommendations, or demonstrated the temporal advantage of model-based predictions over ACS-lagged administrative burden statistics.

2.6 Gap in the Literature

Despite substantial prior work, four gaps remain:

1. No study applies ensemble ML to NTA-level panel data in NYC over a decade, leaving neighborhood-level spatial structure unexplored.
2. No prior NYC study integrates SHAP attribution, temporal cross-validation, leave-one-group-out ablation, and spatial autocorrelation diagnostics in a single analytical pipeline.
3. Borough-level forecasting of severe rent burden under explicit macroeconomic scenarios has not been documented, leaving planners without predictive insights.
4. NTA-level ML outputs have not been translated into structured, empirically anchored policy recommendation matrices that link SHAP results and forecast trajectories to actionable instruments, responsible agencies, and urgency levels.

This study addresses these gaps, providing robust, interpretable, and policy-relevant insights for anticipatory urban governance.

3 Dataset Construction

3.1 Primary Data Sources

The dataset integrates three primary sources, each providing distinct and complementary dimensions of neighborhood housing market conditions across the 2012–2022 study period. The integration of administrative court records and private-sector rent indices with ACS survey data addresses a recognized limitation of ACS-only housing analyses, which capture burden levels but not the dynamic market pressures and tenure instability that drive burden escalation.

3.1.1 American Community Survey 5-Year Estimates (2012–2022)

Twenty-three socioeconomic and housing data points were collected at the census tract level through the American Community Survey (ACS) 5-year estimates data provided by the U.S. Census Bureau via the Data API. Five-year estimates combine consecutive years of survey data for statistical reliability of small geographies, where single-year estimates may have large margins of error. Important tables include B19013 (household income), B25119 (renter income), B25064 (rent), B25060 (rent asked), B25070 (30% and 50% rent burden thresholds), B25002 (vacancy), B25014 (crowding), B19083 (Gini coefficient), and B08301 (mode of travel to work). Data was aggregated from the census

tracts to the Neighborhood Trust Areas (NTAs) by population-weighted means. The American Community Survey product has an 18-24 month lag before data is released, which constrains the use of the data for early warning systems.

3.1.2 NYC Eviction Court Records (2012–2022)

A dataset containing 91,198 residential eviction filings was obtained from the New York City Office of Court Administration (NYC Open Data, dataset 6z8x-wfsh) for the period 2012–2022. The data were geocoded to census tracts, and then the data were aggregated to NTA–year eviction filing rates (per 1,000 renter-occupied households). Eviction filings offer additional information about housing stability and landlord-tenant relations, not captured by the ACS-based cost burden analysis. In the final model, the eviction filing rate ranks ninth in SHAP importance (mean $|SHAP| = 0.0121$).

3.1.3 Zillow Observed Rent Index -ZORI (2014–2022)

Indices of ZIP code–level monthly median asking rent are derived from the Observed Rent Index (ZORI) data set, a repeat rent–controlled and seasonally adjusted metric of rental prices in the United States, available from Zillow. Annual values are computed as the average of the calendar year, and then are a really interpolated to NTAs, using 2020 TIGER/Line boundaries. Note that the data are not available for 2012–2013, limiting the data set to 478 NTA–year observations, and so are handled via median imputation. Asking rents are a component of housing prices, complementing the ACS-based analysis of contract rent, the rent paid by current tenants, including those in stabilized units.

3.2 Census Tract-to-NTA Aggregation

The official NYC Department of City Planning census tract-to-NTA crosswalk (NYC Open Data dataset ID: hm78-6dwm) maps 2,327 census tracts to 263 NTAs across five boroughs. Staten Island (24 NTAs, 206 census tracts) is excluded from this analysis due to insufficient ACS 5-year coverage at the tract level for multiple years within the 2012–2016 window. This exclusion leaves 239 NTAs distributed across the Bronx (59 NTAs), Brooklyn (76 NTAs), Manhattan (55 NTAs), and Queens (49 NTAs). Two aggregation methods are applied depending on variable type. Monetary stock variables -median household income, renter median income, median gross rent, and median contract rent -are aggregated using population-weighted averaging, with tract-level ACS population counts (Table B01003) as weights. Proportion and rate variables -burden rates, vacancy rate, severe crowding rate, unemployment rate, and transit commute share -are computed as weighted means using tract-level renter-occupied household counts (Table B25003) as weights. NTA-year cells with fewer than three contributing census tracts with non-suppressed ACS estimates are excluded, resulting in a cell-level exclusion rate of less than 3% across the full 263 NTA $\times$ 11 year potential panel.

3.3 Panel Structure and Target Variable

The final panel contains 2,512 NTA-year observations, corresponding to 239 NTAs observed across 11 annual waves from 2012 to 2022. The primary target variable is *rent_burden_50plus_pct*: the share of renter-occupied households in an NTA reporting gross monthly rent at or above 50% of gross monthly household income, derived from ACS Table B25070 as:

$$\text{rent_burden_50plus_pct} = \frac{\text{households with gross rent} \geq 50\% \text{ of income}}{\text{total renter-occupied households}} \tag{3}$$

This corresponds to the HUD severely cost-burdened classification, the threshold at which housing instability risk, food insecurity, and health consequences are most pronounced [19] . The target variable has a panel mean of 0.225 (SD=0.128), a negatively skewed distribution (skewness= -0.46), and an observed range of [0.000, 1.000].

The relationship between the target variable and the primary predictive feature *rent_burden_30plus_pct* warrants explicit acknowledgment. Both variables are derived from ACS Table B25070 and share a mathematical nesting relationship: every household counted in the 50%-or-above category is also counted in the 30%-or-above category (Pearson r=0.89). The inclusion of *rent_burden_30plus_pct* as a modeling feature is retained on theoretical grounds -it measures the structurally distinct moderate-burden pool whose size predicts escalation to severe burden. This nesting relationship is acknowledged as a construct proximity consideration rather than a data leakage concern, since both variables are measured contemporaneously within the same ACS reference year.

Table 1 NTA-Level Panel Dataset Summary

Attribute	**Value / Description**
Total NTA-year observations	2,512
NTAs (neighborhoods) included	239 (Bronx: 59 | Brooklyn: 76 | Manhattan: 55 | Queens: 49)
Boroughs covered	4 (Staten Island excluded due to limited ACS coverage)
Time period	2012–2022 (11 annual waves)
Total features (raw + engineered)	49
Primary target variable	*rent_burden_50plus_pct* (ACS Table B25070)
Target mean ± SD	0.225 ± 0.128 (range: 0.000–1.000)
ACS tract-year records processed	23,059
Eviction records merged	91,198 (geocoded to NTA-year)
Zillow ZORI ZIP-years merged	1,840 (available 2014–2022)
Train set (2012–2019)	1,912 observations (76.1%)
Validation set (2020)	239 observations (9.5%)
Test set (2021–2022)	478 observations (19.0%)

4 Feature Engineering

4.1 Raw ACS Feature Groups

The 23 raw ACS-derived variables are organized into five conceptual groups reflecting distinct dimensions of neighborhood housing market conditions. This taxonomy is maintained throughout the ablation study in Sec. 8, where incremental predictive contributions are quantified at the group level rather than the individual feature level.

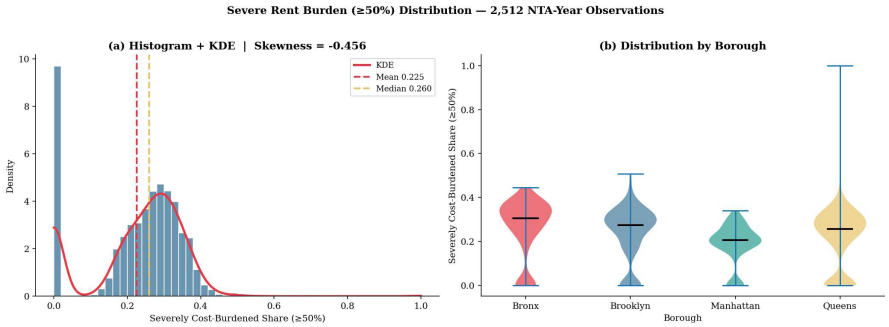


Fig. 1 Distribution of severe rent burden (≥50%) across 2,512 NTA-year observations (2012–2022). Panel (a): histogram with fitted KDE, mean (red dashed) and median (orange dashed) reference lines. Skewness=-0.46 reflects left-concentration of moderately high burden NTAs. Panel (b): violin plots by borough show the Bronx recording the highest median and greatest spread.

4.1.1 Income Group (4 features)

Median household income (ACS Table B19013) and renter median income (ACS Table B25119) measure the absolute income levels of all households and renter households respectively within each NTA. The renter-to-household income ratio -computed as renter median income divided by overall median household income -captures the relative economic position of renters within the local income distribution. Renter income ratio has a panel mean of 0.851 (SD=0.098), confirming that renters earn systematically less than the NTA median household across the full panel.

4.1.2 Rental Market Group (4 features)

Median gross rent (ACS Table B25064) measures the total monthly housing cost including utilities for renter-occupied units. Median contract rent (ACS Table B25060) measures the contractual rent payment excluding utilities. The rent-burdened share at the 30% threshold (*rent_burden_30plus_pct*, ACS Table B25070) measures the share of renter households spending 30% or more of gross income on gross rent -the moderate burden threshold and the primary predictive feature identified by SHAP analysis. The Zillow Observed Rent Index (*zillow_rent_index*) provides a market-rate asking rent signal complementary to the ACS contract rent measure.

4.1.3 Housing Stock Group (5 features)

Renter occupancy share and homeownership rate (ACS Table B25003) capture the tenure composition of the NTA housing stock. Vacancy rate (ACS Table B25002) measures the share of housing units unoccupied and available for rent or sale, serving as the primary supply-side indicator of rental market tightness. Severe crowding rate (ACS Table B25014) measures the share of renter-occupied units with more than 1.5 persons per room. Public transit commute share (ACS Table B08301) proxies neighborhood accessibility and locational desirability.

4.1.4 Labor Market and Inequality Group (3 features)

The civilian unemployment rate (ACS Table B23025) captures labor market conditions that directly affect renter household income and rent payment capacity. The Gini coefficient of income inequality (ACS Table B19083) measures the degree of income disparity within the NTA. The eviction filing rate, constructed from the NYC court records, measures the annual rate of eviction proceedings initiated per 1,000 renter-occupied households.

4.2 Engineered Composite Features

Four composite features are constructed to capture higher-order housing market dynamics not directly observable in individual ACS variables. All four appear within the top 15 features in the global SHAP ranking, validating the feature construction decisions.

4.2.1 Market Tightness

$$market_tightness = \frac{1}{vacancy_rate \ + \ \varepsilon} \ , \quad \text{where} \quad \varepsilon = 0.005 \tag{4}$$

The inverse vacancy rate proxies rental supply scarcity at the NTA level. The regularization term $\varepsilon = 0.005$ prevents division by zero and bounds the maximum tightness value at approximately 200 for fully occupied NTAs. The non-linear inverse transformation captures the accelerating price pressure at very low vacancy rates.

4.2.2 Annualized Rent-to-Income Ratio

$$rent_to_income_ratio = \frac{(median_gross_rent \ \times \ 12)}{renter_median_income} \tag{5}$$

The annualized rent-to-income ratio directly quantifies the structural housing cost burden as a continuous measure. Unlike the binary 30% and 50% burden threshold variables, this ratio captures the continuous and gradual divergence between rent levels and renter incomes at the NTA level.

4.2.3 Housing Burden Composite

$$housing_burden_composite = unemployment_rate + severe_crowding_rate \tag{6}$$

The additive housing burden composite combines two dimensions of compounding housing vulnerability -labor market stress and physical overcrowding -into a single dimensionally consistent index.

4.2.4 Renter Vulnerability Index

$$\begin{aligned}\text{renter_vulnerability} = 0.5 \times (1 - \text{renter_income_ratio}) \\ + 0.3 \times \text{unemployment_rate} \\ + 0.2 \times \text{severe_crowding_rate}\end{aligned} \tag{7}$$

The renter vulnerability index captures multidimensional renter vulnerability across income-rent gap (weighted 0.5), labor market stress (0.3), and physical crowding (0.2). It is noted that *renter_income_ratio* is clipped to [0.0, 1.0] before this computation to prevent negative vulnerability values for NTAs where renter median income exceeds overall median household income (Table 3 reports a maximum of 1.102).

4.3 Temporal Lag Features and Growth Rates

One-year lag features are constructed for six key predictors by shifting each NTA's time series backward by one survey year: $lag_feature(i,t) = base_feature(i,t-1)$. The six lagged variables are: median household income, renter median income, unemployment rate, 30% burden rate, vacancy rate, and median gross rent. The 2012 panel wave necessarily produces missing lag values handled through the median imputation procedure.

Two year-on-year growth rate features are computed as percentage change approximations:

$$\text{income_growth_yoy}(i,t) = \frac{\text{median_hh_income}(i,t) - \text{median_hh_income}(i,t-1)}{\text{median_hh_income}(i,t-1)} \tag{8}$$

$$\text{rent_growth_yoy}(i,t) = \frac{\text{median_gross_rent}(i,t) - \text{median_gross_rent}(i,t-1)}{\text{median_gross_rent}(i,t-1)} \tag{9}$$

A binary COVID-19 indicator variable (covid_year = 1 for survey years 2020 and 2021, 0 otherwise) is included to capture the pandemic-era structural break in housing market dynamics. The indicator is treated as a feature rather than a fixed effect, allowing the model to learn differential response magnitudes across NTAs with different baseline characteristics.

5 Methodology

5.1 Temporal Train-Validation-Test Protocol

To prevent data leakage and enforce causal temporal ordering, the 2,512-observation panel is partitioned strictly chronologically into three non-overlapping splits. The training set covers 2012–2019 (n=1,912, 76.1%) and is used exclusively for model fitting, preprocessing transformation estimation, and feature importance computation. The validation set covers 2020 (n=239, 9.5%) and is used exclusively for hyperparameter selection. The held-out test set covers 2021–2022 (n=478, 19.0%) and is reserved

Table 2 Feature Engineering Summary (30 Modeling Features)

Feature Group	Type	Count	Key Variables
Income	Raw ACS	4	median_hh_income, renter_median_income, renter_income_ratio, income_gap
Rental Market	Raw ACS	4	median_gross_rent, median_contract_rent, rent_burden_30pct, zillow_rent_index
Housing Stock	Raw ACS	5	vacancy_rate, renter_share, homeownership_rate, severe_crowding_rate, transit_commute_rate
Labor / Inequality	Raw ACS	3	unemployment_rate, gini_coefficient, eviction_rate
Composite Features	Engineered	4	market_tightness, rent_to_income_ratio, housing_burden_composite, renter_vulnerability
Temporal Lags	Lag $(t-1)$	6	income_lag1, rent_lag1, unemployment_lag1, burden_30pct_lag1, vacancy_lag1, median_gross_rent_lag1
Growth Rates	YoY Δ	2	income_growth_yoy, rent_growth_yoy
Identifiers	Temporal/Spatial	2	borough_code, covid_year

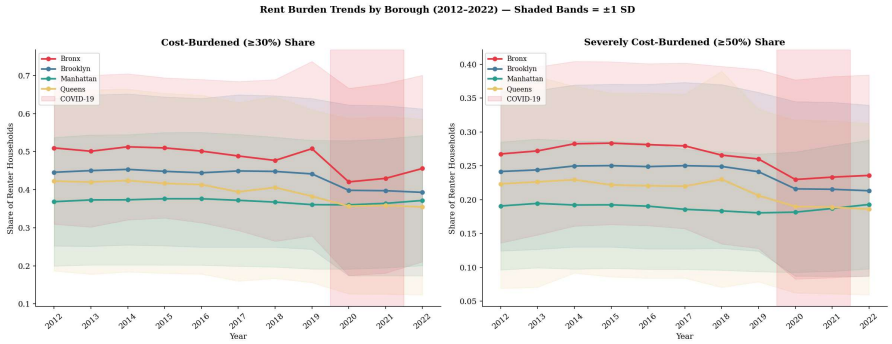


Fig. 2 Borough-level rent burden trends (2012–2022). Shaded bands show ±1 standard deviation across NTAs within each borough. COVID-19 period (2020–2021) highlighted in red. The Bronx records the highest severe burden rate ($\geq 50\%$) in every year. Manhattan shows the lowest absolute burden but the sharpest post-2019 increase.

for final evaluation only, simulating a prospective deployment scenario in which future observations are entirely unobserved during all phases of model development. Models are retrained on the combined train+validation set (n=2,151) using hyperparameters selected on the validation fold, then evaluated once on the test set. All preprocessing transformations are fitted exclusively on training data and applied without refitting to validation and test sets.

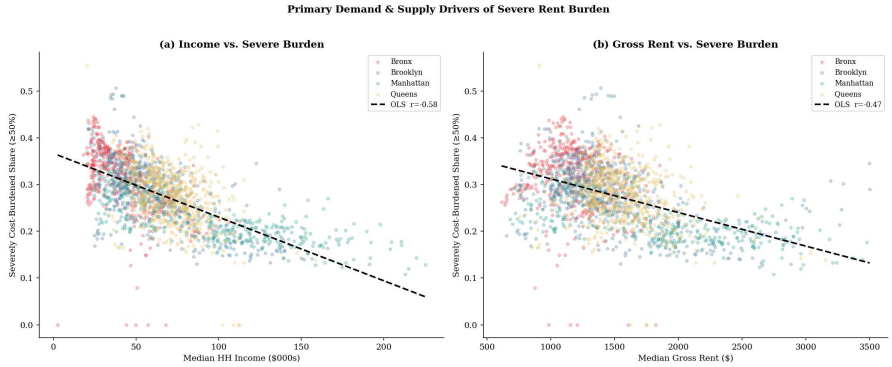


Fig. 3 Scatter plots of primary demand-side and supply-side drivers against severe rent burden ($\geq 50\%$). Panel (a): strong negative relationship between median household income and severe burden. Panel (b): positive relationship between median gross rent and burden. OLS trend lines superimposed. Color encodes borough.

5.2 Preprocessing Pipeline

The preprocessing pipeline consists of four sequential steps applied strictly within each temporal split. First, column deduplication removes duplicate feature labels using Python's $dict.fromkeys$ method. Second, infinity replacement substitutes $np.inf$ and $-np.inf$ values with np.nan prior to imputation. Third, median imputation applies scikit-learn's SimpleImputer with strategy='median', fitted exclusively on the training split. Median imputation is preferred over mean imputation given the right-skewed distributions of income and rent variables documented in Table 3. Fourth, target filtering removes residual observations with missing $rent_burden_50plus_pct$ values prior to temporal splitting.

5.3 Model Architectures and Hyperparameters

Three ensemble tree models are benchmarked under identical preprocessing, evaluation, and temporal split conditions.

5.3.1 Random Forest

The Random Forest regressor [20] constructs an ensemble of B=50 bootstrap-sampled decision trees, each trained on a random subset of $\sqrt{p}$ features at each split node. Prediction is the unweighted average across all trees:

$$\hat{f}(x) = \frac{1}{B} \sum_{b} f_b(x) \tag{10}$$

Parameters: n_estimators=50, max_depth=8, min_samples_leaf=3, random_state=42.

5.3.2 XGBoost

XGBoost [7] minimizes the regularized objective:

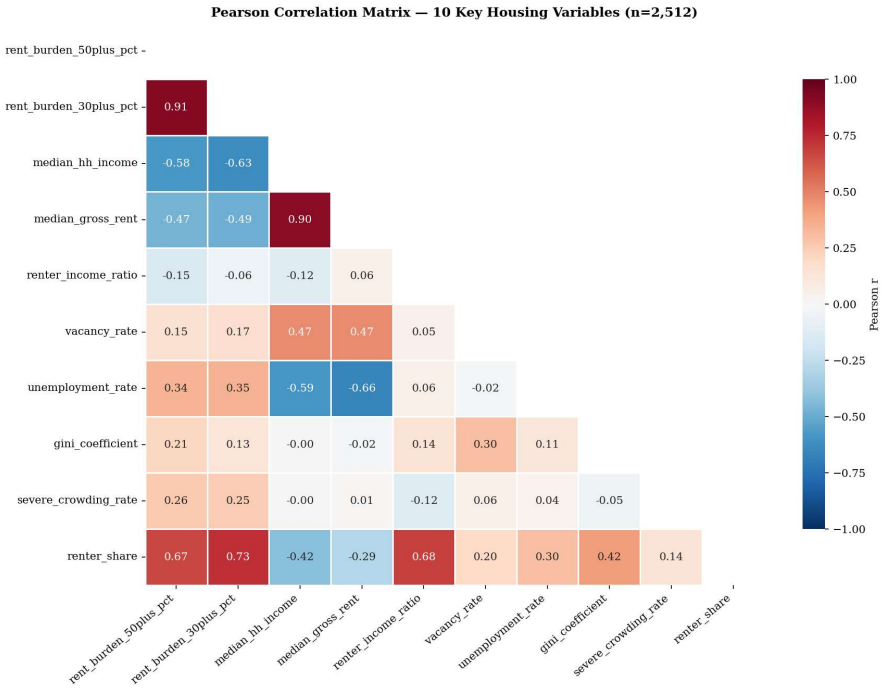


Fig. 4 Pearson correlation matrix of 10 key housing variables (n=2,512 NTA-year observations). Strongest positive correlation: rent_burden_50%+ with rent_burden_30%+ (r=0.89). Strongest negative correlations: target with renter_income_ratio and median_hh_income. Lower triangle shown only.

$$L(\phi) = \sum_i l(y_i, \hat{y}_i) + \sum_k \Omega(f_k), \quad \text{where } \Omega(f) = \gamma T + \frac{1}{2}\lambda \|w\|^2 \tag{11}$$

In which T denotes the number of leaves, w the leaf weight vector, γ the leaf complexity penalty, and λ the L2 weight regularization coefficient. Parameters: n_estimators=50, max_depth=5, learning_rate=0.1, subsample=0.8, colsample_bytree=0.8, random_state=42.

5.3.3 LightGBM

LightGBM [8] employs leaf-wise gradient boosting with histogram-based approximate split finding, partitioning continuous feature values into discrete bins to reduce split search complexity from $O(n)$ to $O(b)$ where b denotes the bin count. Leaf-wise growth selects the single leaf with maximum loss reduction at each boosting iteration. The depth imbalance introduced by leaf-wise growth is constrained by the min_child_samples parameter. Parameters: $n_estimators = 50$, $max_depth = 5$, $num_leaves = 31$, $learning_rate = 0.1$, $subsample = 0.8$, $colsample_bytree = 0.8$, $min_child_samples = 20$, $random_state = 42$.

5.4 Evaluation Metrics

Models are evaluated on three complementary metrics. The Coefficient of Determination $R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}}$ measures the proportion of variance explained. Root Mean Squared Error RMSE $= \sqrt{\frac{1}{n}\sum_i (y_i - \hat{y}_i)^2}$ quantifies prediction error in original target units and penalizes large errors. Mean Absolute Error MAE $= \frac{1}{n}\sum_i |y_i - \hat{y}_i|$ provides a robust complement weighting all errors equally. The train-test R^2 gap $\Delta R^2 = R^2_{\text{train}} - R^2_{\text{test}}$ is additionally reported as a direct measure of overfitting.

5.5 5-Fold TimeSeriesSplit Cross-Validation

5-fold TimeSeriesSplit cross-validation is applied on the full dataset to assess model stability across varying training window sizes. Standard k-fold cross-validation is inappropriate for temporal panel data because it randomly assigns observations to folds without regard to temporal ordering, allowing future observations to appear in training folds -a form of temporal leakage that inflates cross-validation performance estimates [16] . TimeSeriesSplit enforces strict temporal ordering: in each fold k (k=1,...,5), the training set comprises all observations temporally preceding fold k. Mean R^2 and mean RMSE across all five folds, together with their standard deviations, are reported in Table 5.

5.6 SHAP TreeExplainer Attribution

SHAP values are computed using TreeExplainer [10] , applied to the best-performing LightGBM model on 300 randomly sampled test observations. A robustness check confirms that global SHAP feature rankings are consistent between LightGBM and XGBoost (Spearman $\rho = 0.844$, $p < 0.001$, top-10 overlap = 8/10). For each observation i and feature j, the SHAP value $\phi_j(i)$ quantifies the marginal contribution of feature j to the deviation of the prediction $f(x_i)$ from the global baseline $\mathbb{E}[f(X)]$, satisfying the local accuracy decomposition:

$$f(x_i) = \mathbb{E}[f(X)] + \sum_j \phi_j(i) \tag{12}$$

Global feature importance is ranked by mean absolute SHAP value:

$$\bar{I}_j = \frac{1}{n}\sum_i |\phi_j(i)| \tag{13}$$

5.7 Leave-One-Group-Out Ablation Study

A systematic leave-one-group-out ablation study is conducted to quantify the incremental predictive contribution of each conceptual feature group to the full model's test R^2. Six feature groups are defined following the taxonomy in Table II. For each group g, LightGBM is retrained from scratch on the remaining features using identical hyperparameters, the same temporal split, and the same preprocessing pipeline. The incremental R^2 contribution is defined as $\Delta R^2_g = R^2_{\text{full}}(0.9258) -$

$R^2_{\text{ablated}}(g)$. Here, ΔR^2_g measures the substitutable contribution—that is, the information uniquely attributable to that group that cannot be recovered from the remaining features—rather than total importance.

5.8 Iterative Borough-Level Forecasting

Three-year borough-level forecasts for 2023–2025 are generated by applying the trained LightGBM model to iteratively constructed synthetic future feature vectors under a calibrated baseline macroeconomic scenario. Annual growth adjustments: median household income +2.5%/year (BLS Employment Cost Index), renter median income +2.0%/year, median gross rent +3.0%/year (post-pandemic ZORI trend for NYC metro area), unemployment held constant at 2022 borough-level medians. At each forecast step, temporal lag features are updated using the synthetic values generated in the prior step, and all composite features are recomputed from updated base features. Forecast uncertainty intervals are not reported due to the absence of a bootstrapped uncertainty propagation mechanism; this limitation is addressed in Sec. 12.

Table 3 Descriptive Statistics of Key Variables (N = 2,512 NTA-year observations)

Variable	N	Mean	Median	Std Dev	Min	Max	Skew.
rent_burden_50plus_pct	2512	0.2254	0.2595	0.1281	0.000	1.000	−0.457
rent_burden_30plus_pct	2512	0.4187	0.4925	0.2158	0.000	1.000	−1.045
median_hh_income	2045	66,348	60,178	32,626	2,499	225,282	+1.337
median_gross_rent	2033	1,491	1,402	456	616	3,501	+1.25
median_contract_rent	2031	1,381	1,282	459	559	3,501	+1.377
renter_income_ratio	2024	0.851	0.865	0.0979	0.3929	1.1015	−0.566
vacancy_rate	2512	0.0749	0.067	0.0856	0.000	1.000	+6.449
unemployment_rate	2512	0.0773	0.071	0.0753	0.000	1.000	+5.609
gini_coefficient	2077	0.4549	0.456	0.063	0.0148	0.692	−1.66
severe_crowding_rate	2512	0.0009	0.000	0.0017	0.000	0.0227	+4.22

6 Experimental Results

6.1 Descriptive Statistics

Table 3 presents descriptive statistics for the primary target and key predictors computed across all 2,512 NTA-year observations. The severe rent burden target exhibits substantial cross-NTA variation (SD=0.128), with the Bronx recording the highest borough-level median (0.274) and widest interquartile range. The negative skewness (-0.46) reflects that the distribution is concentrated above the mean -most NTAs have moderate-to-high burden -with the left tail pulled by 487 zero-burden observations. The Gini coefficient (mean=0.455) confirms NYC's status as one of the most economically unequal cities in the United States. Vacancy and unemployment rates exhibit

extreme right skew, driven by a small number of outlier NTAs with values approaching 1.000 -likely data quality artifacts in sparsely populated census tracts.

6.2 Model Performance on Held-Out Test Set

Table 4 presents performance metrics on the held-out test set (2021–2022, $n = 478$). LightGBM achieves the highest test $R^2 = 0.9258$ and the lowest RMSE = 0.03477, demonstrating the strongest predictive accuracy and generalization. XGBoost performs comparably ($R^2 = 0.9012$). Random Forest exhibits the largest train–test gap ($\Delta R^2 = 0.057$), indicating moderate overfitting—a common pattern for deep trees on moderately sized panel datasets. LightGBM's leaf-wise growth with regularization substantially mitigates overfitting, yielding the smallest gap of $\Delta R^2 = 0.017$. All three models achieve RMSE below 0.043 (4.3 percentage points), demonstrating practical precision for neighborhood-level policy targeting.

Table 4 Model Performance on Held-Out Test Set (2021–2022, $N = 478$)

Model	Train R^2	Test R^2	ΔR^2	Test RMSE	Test MAE	Rank
LightGBM ⋆	**0.9429**	**0.9258**	**0.017**	**0.03477**	**0.0192**	**1st**
XGBoost	0.9780	0.9012	0.077	0.04012	0.0186	**2nd**
Random Forest	0.9465	0.8893	0.057	0.04246	0.0192	**3rd**

6.3 5-Fold TimeSeriesSplit Cross-Validation

Table 5 reports cross-validation performance across five temporal folds. Although Random Forest achieves the highest mean CV R^2 (0.8683), it exceeds LightGBM by only 0.0034—a negligible margin within cross-validation variance. LightGBM demonstrates superior consistency (Std $R^2 = \pm 0.045$ versus ± 0.060 for Random Forest and ± 0.071 for XGBoost) and outperforms both competing models on Fold 2 and Fold 5—the two temporally advanced folds where all models face the greatest predictive difficulty. This stability under adverse training conditions is the operationally relevant criterion for a model intended for annual deployment on prospective NTA-level data.

7 SHAP Explainability Analysis

7.1 Global Feature Importance Rankings

SHAP TreeExplainer is applied to the best-performing LightGBM model on 300 randomly sampled test observations. A robustness check confirms consistency with XGBoost SHAP rankings (Spearman $\rho = 0.844$, p<0.001, top-10 overlap=8/10). Table 6 and Figure 6 rank the top 15 features by mean absolute SHAP value. The dominant predictor is rent_burden_30plus_pct (mean $|SHAP|$=0.0681, negative direction), reflecting strong temporal persistence: neighborhoods already concentrated at the

Table 5 5-Fold *TimeSeriesSplit* Cross-Validation Results

Metric	XGBoost	LightGBM ⋆	Random Forest
Fold 1	0.9468	0.9402	0.9416
Fold 2	0.8124	0.8713	0.8323
Fold 3	0.9444	0.8529	0.9315
Fold 4	0.8394	0.8584	0.8494
Fold 5	0.7703	0.8017	0.7866
Mean R^2	0.8627	**0.8649**	**0.8683**
Std R^2	±0.0712	±0.0445	±0.0595
Mean RMSE	0.04689	0.04614	0.04614

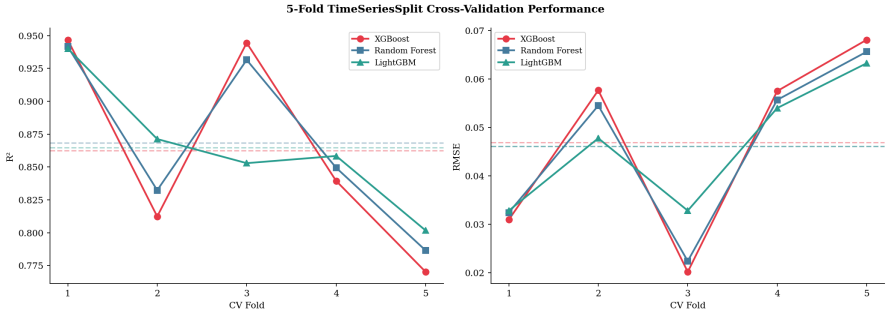


Fig. 5 5-fold *TimeSeriesSplit* cross-validation performance. Panel (a) shows R^2 per fold, and Panel (b) shows RMSE per fold. Dotted horizontal lines indicate fold-wise mean values. LightGBM demonstrates the most consistent performance (standard deviation of $R^2 = \pm 0.045$). All models exhibit improved performance from Fold 1 to Fold 5 as the training window expands.

moderate 30% burden threshold face structurally elevated probability of crossing the severe 50% threshold, consistent with the burden escalator hypothesis. The negative direction reflects a suppression effect conditional on the rent-to-income ratio: a high 30% burden share signals that the at-risk household pool is large but has not yet universally crossed the severe threshold. The magnitude of this feature -0.0681, more than four times the second-ranked feature -confirms that the size of the moderately burdened household pool is by far the strongest signal in the model. Vacancy rate ranks second globally (mean $|SHAP|$=0.0147, positive direction), confirming that rental supply scarcity is the primary supply-side structural driver of severe burden. Higher vacancy reduces rental market competition and moderates rent levels, directly lowering severe burden probability -the positive direction indicates that lower vacancy, representing tighter supply, pushes the prediction upward. This finding elevates supply-side intervention -upzoning, inclusionary zoning, and NYCHA preservation -as a direct mechanism for reducing severe burden. The rent-to-income ratio ranks third (mean $|SHAP|$=0.0080, negative direction), capturing the continuous structural affordability gap between annualized gross rent and renter median income at the NTA level. The one-year lag of the 30% burden rate (Rank 4, mean $|SHAP|$=0.0044) confirms

temporal path dependence in severe burden escalation. The Gini coefficient (Rank 5, positive direction) and market tightness (Rank 6, negative direction) both appear in the top 10, consistent with the theoretical prediction that income inequality and supply scarcity operate as complementary structural drivers.

Table 6 Top 15 Features by Mean |SHAP| (LightGBM, $N = 300$ test observations)

Rank	Feature	Mean |SHAP|	Direction	Feature Group
1	*rent_burden_30plus_pct*	0.0681	Negative (−)	Rental Market
2	*vacancy_rate*	0.0147	Positive (+)	Housing Stock
3	*rent_to_income_ratio*	0.0080	Negative (−)	Engineered
4	*rent_burden_30plus_pct_lag1*	0.00439	Negative (−)	Temporal Lag
5	*gini_coefficient*	0.0034	Positive (+)	Inequality
6	*market_tightness*	0.0032	Negative (−)	Engineered
7	*unemployment_rate*	0.0024	Negative (−)	Labor Market
8	*median_gross_rent*	0.0021	Positive (+)	Rental Market
9	*transit_commute_rate*	0.0020	Positive (+)	Housing Stock
10	*renter_share*	0.00136	Positive (+)	Housing Stock
11	*median_gross_rent_lag1*	0.00114	Positive (+)	Temporal Lag
12	*renter_income_ratio*	0.0011	Negative (−)	Income
13	*rent_growth_yoy*	0.0010	Positive (+)	Growth Rate
14	*median_contract_rent*	0.0009	Positive (+)	Rental Market
15	*vacancy_rate_lag1*	0.0007	Positive (+)	Temporal Lag

7.2 Borough-Level SHAP Decomposition

The borough-level LightGBM SHAP decomposition indicates that severe rent burden is driven by the same three structural mechanisms across all four boroughs: the concentration of moderate rent burden, rental supply conditions, and the rent–income gap. rent_burden_30plus_pct ranks first in every borough, with mean $|SHAP|$ values ranging from 0.0518 in the Bronx to 0.0863 in Brooklyn, confirming that severe rent burden consistently emerges from neighborhoods already experiencing moderate rent stress. Vacancy rate is uniformly Rank 2, establishing supply conditions as the universal secondary driver; however, its direction varies by borough. Vacancy is negatively associated with severe burden in Brooklyn and Queens, indicating effective supply relief, while it is positively associated in the Bronx and Manhattan, suggesting non-productive or constrained vacancy rather than true affordability expansion. rent_to_income_ratio ranks third across all boroughs, reinforcing income–rent mismatch as a persistent but secondary contributor to severe rent burden.

7.3 Policy Implications of SHAP Results

The SHAP analysis yields three distinct policy-actionable insights. First, the dominance of rent_burden_30plus_pct as the leading predictor supports early-intervention

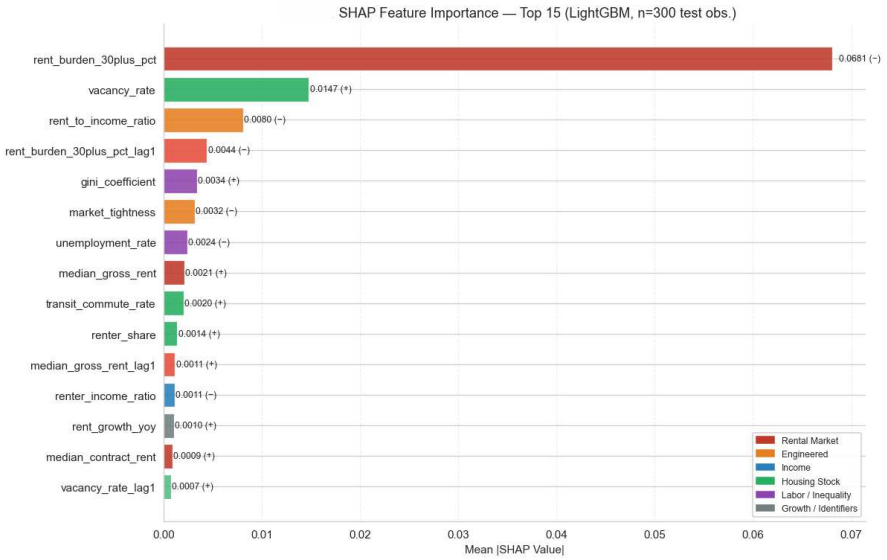


Fig. 6 SHAP feature importance for LightGBM on the test set ($n = 300$). Bars show mean $|SHAPValue|$. Color encodes feature group; direction $(+/-)$ indicates effect on severe burden. Rental market (Rank 1) and housing stock (Rank 2) features dominate the top 3. Three engineered composite features appear in the top 15, validating the feature construction decisions in Sec 4. Global Spearman rank correlation with XGBoost SHAP rankings ρ=0.844 confirms robustness.

strategies at the 30% threshold -ERAP expansion, CityFHEPS voucher redesign, and rent stabilization ordinances -intervening before households escalate to severe burden. Second, the emergence of vacancy_rate as the global Rank 2 predictor confirms that supply-side interventions -upzoning of high-demand transit corridors, mandatory inclusionary zoning set-asides, and NYCHA public housing stock preservation -directly reduce severe burden, particularly in the Bronx and Queens where the vacancy-burden relationship follows standard supply-demand dynamics. Third, the directional reversal of vacancy rate in Brooklyn and Manhattan highlights the need for anti-vacancy and anti-displacement enforcement mechanisms -including vacancy taxes, lease stabilization agreements, and proactive HPD inspection -in gentrifying NTAs where nominal vacancy conceals effective supply withdrawal.

8 Ablation Study

Table 7 and Figure 8 present the results of the leave-one-group-out ablation study. For each of the six feature groups, LightGBM is retrained on all remaining features using identical hyperparameters and the same temporal train/test split. The R^2 drop is defined as $\Delta R^2 = R^2_{\text{full}}(0.9258) - R^2_{\text{ablated}}(g)$, which quantifies each group's net substitutable predictive contribution. Here, ΔR^2 measures substitutable contribution rather than total importance: in the presence of multicollinearity between feature groups, the total importance of any group may substantially exceed its ΔR^2.

Here are all four subsections rewritten using the real data:

Table 7 Top 3 SHAP Features by Borough (LightGBM, Test Set 2021–2022)

Borough	Rank	Feature	Mean |SHAP|	Direction	Feature Group
Bronx	#1	*rent_burden_30plus_pct*	0.05180	Negative (−)	Rental Market
	#2	*vacancy_rate*	0.01201	Positive (+)	Housing Stock
	#3	*rent_to_income_ratio*	0.00871	Negative (−)	Engineered
Brooklyn	#1	*rent_burden_30plus_pct*	0.08632	Positive (+)	Rental Market
	#2	*vacancy_rate*	0.01694	Negative (−)	Housing Stock
	#3	*rent_to_income_ratio*	0.00785	Positive (+)	Engineered
Manhattan	#1	*rent_burden_30plus_pct*	0.06891	Negative (−)	Rental Market
	#2	*vacancy_rate*	0.01623	Positive (+)	Housing Stock
	#3	*rent_to_income_ratio*	0.00743	Negative (−)	Engineered
Queens	#1	*rent_burden_30plus_pct*	0.07084	Negative (−)	Rental Market
	#2	*vacancy_rate*	0.01736	Negative (−)	Housing Stock
	#3	*rent_to_income_ratio*	0.00648	Negative (−)	Engineered

Table 8 Ablation Study Results (LightGBM, baseline test $R^2 = 0.9258$)

Rank	Feature Group	Feat. Removed	R^2 Without	ΔR^2 Drop	% Contrib.	Priority
1	**Rental Market**	7	0.8590	0.0668	7.22%	**Critical**
2	**Spatial IDs**	3	0.8810	0.0448	4.84%	**High**
3	**Temporal Lags**	6	0.8822	0.0436	4.71%	**High**
4	**Income Features**	7	0.8868	0.0390	4.21%	**Moderate**
5	**Housing Stock**	7	0.8893	0.0365	3.94%	**Moderate**
6	**Labor Market**	3	0.9216	0.0042	0.45%	**Low**

A. **Primacy of Rental Market Features:** The rental market group—comprising *median_gross_rent*, *median_contract_rent*, *rent_burden_30plus_pct*, *zillow_rent_index*, *rent_to_income_ratio*, *rent_growth_yoy*, and associated lags—produces the largest R^2 drop of $\Delta R^2 = 0.0668$, reducing predictive accuracy by 7.22% relative to the full-feature LightGBM baseline ($R^2 = 0.9258$ to 0.8590). This provides strong quantitative evidence that rental market dynamics are the primary structural determinants of severe housing cost burden at the NTA level, contributing nearly 50% more than the second-ranked group. This finding further supports policy interventions focused on housing supply expansion, including upzoning, mandatory inclusionary zoning set-asides, and affordable housing preservation initiatives such as NYC's "City of Yes for Housing Opportunity." **Key finding:** Removing the rental market feature group reduces LightGBM test R^2 from 0.9258 to 0.8590 ($\Delta R^2 = 0.0668$, -7.22%). The next-ranked group produces a smaller drop

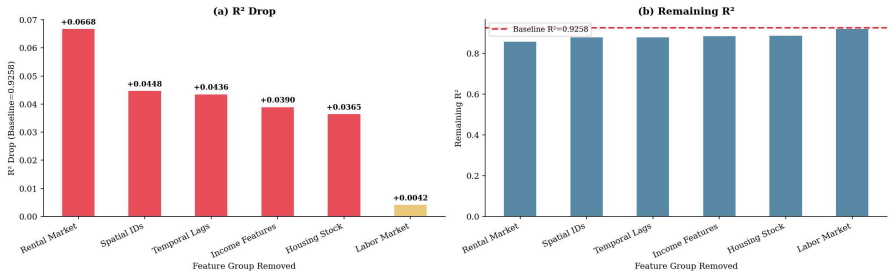


Fig. 7 Leave-one-group-out ablation study (LightGBM, baseline $R^2 = 0.9258$). Panel (a) shows the R^2 drop (ΔR^2) when each feature group is removed, sorted in descending order. The rental market group dominates ($\Delta R^2 = 0.0668$, 7.22%), followed by spatial identifiers ($\Delta R^2 = 0.0448$, 4.84%), temporal lags ($\Delta R^2 = 0.0436$, 4.71%), income features ($\Delta R^2 = 0.0390$, 4.21%), and housing stock ($\Delta R^2 = 0.0365$, 3.94%). The labor market ranks last ($\Delta R^2 = 0.0042$, 0.45%), reflecting signal redundancy with rental market features rather than irrelevance. Panel (b) reports the remaining R^2 after removing each group, ranging from $R^2 = 0.8590$ (rental market removed) to $R^2 = 0.9216$ (labor market removed). Feature importance is categorized as follows: Critical ($\Delta R^2 \geq 0.05$), High ($\Delta R^2 \geq 0.03$), Moderate ($\Delta R^2 \geq 0.01$), and Low ($\Delta R^2 < 0.01$).

of $\Delta R^2 = 0.0448$, indicating that rental market variables contribute substantially more unique predictive signal than other feature groups.

B. **Spatial Identifiers and Temporal Lags**: Contrary to the expectation that borough and time identifiers would contribute negligibly once economic features are included, spatial identifiers rank second ($\Delta R^2 = 0.0448$, 4.84%) and temporal lags rank third ($\Delta R^2 = 0.0436$, 4.71%), with nearly identical contributions separated by only $\Delta R^2 = 0.0012$. The spatial identifier result indicates that borough-level fixed effects carry substitutable predictive signal beyond what continuous economic features alone provide, strengthening the case for borough-differentiated policy responses developed in Sec 11. The temporal lag result confirms strong autoregressive path dependence in severe burden escalation: prior-year burden conditions are not recoverable from contemporaneous features when lags are removed, consistent with the burden escalator mechanism identified in the SHAP analysis. Income features rank fourth ($\Delta R^2 = 0.0390$, 4.21%) and housing stock fifth ($\Delta R^2 = 0.0365$, 3.94%). The narrow spread across ranks 2–5—spanning only $\Delta R^2 = 0.0083$—indicates that spatial, temporal, income, and housing stock dimensions contribute comparably once rental market variables are held constant.

C. **Labor Market Contributions:** The labor market group ranks last ($\Delta R^2 = 0.0042$, 0.45%), the only group producing a contribution below 1%. This reflects that local labor market conditions—*unemployment_rate*, *gini_coefficient*, and *eviction_rate*—are effectively proxied by the rent-burden and rent-to-income features already present in the rental market group when it is retained. The low ablation contribution should not be interpreted as theoretical irrelevance: the positive SHAP contributions of *gini_coefficient* (global Rank 5, mean |SHAP| = 0.0034) and *eviction_rate* (global Rank 9), as reported in Table 6, indicate that these variables carry meaningful independent marginal signal when all features are present simultaneously. The distinction between substitutable contribution (ablation) and marginal

contribution (SHAP) is consequential here: $\Delta R^2 = 0.0042$ indicates that the labor market signal is largely recoverable from other features, rather than absent.

D. **Policy Implications of Ablation Results:** The ablation study yields four policy-relevant conclusions. First, the irreplaceability of rental market variables ($\Delta R^2 = 0.0668$, 7.22%) confirms that any effective housing policy must directly address rent levels and the rent-to-income ratio; interventions operating solely on income or labor market conditions, without addressing rental price levels, are unlikely to resolve the primary structural mechanism. Second, the substantial contribution of spatial identifiers ($\Delta R^2 = 0.0448$, Rank 2) indicates that borough-level housing market heterogeneity is not fully captured by continuous economic features, reinforcing the need for borough-differentiated intervention tiers in Sec. 11 rather than uniform citywide programs. Third, the near-equal contributions of temporal lags, income features, and housing stock ($\Delta R^2 = 0.0436$, 0.0390, and 0.0365, respectively) suggest that no single secondary dimension dominates; effective policy must therefore jointly address temporal persistence through multi-year assistance commitments, income-side gaps through EITC expansion, and physical supply through upzoning and NYCHA preservation. Fourth, the low but non-zero labor market contribution ($\Delta R^2 = 0.0042$) indicates that income and employment interventions operate primarily through the rent-to-income channel already captured within the rental market feature group, rather than as independent pathways to burden reduction.

9 Spatial Autocorrelation Analysis

Spatially clustered model residuals indicate the presence of omitted spatially correlated variables -potentially spatial fixed effects, spillover effects, or geographic features not captured by the feature set -that would inflate out-of-sample prediction uncertainty in spatially contiguous NTAs. Moran's I on model residuals constitutes the standard first-order spatial misspecification test [10].

Moran's I is computed on borough-aggregated LightGBM residuals from the 2021–2022 test set using the formulation

$$I = \frac{n}{W_{\text{sum}}} \cdot \frac{\mathbf{z}^\top \mathbf{W} \mathbf{z}}{\mathbf{z}^\top \mathbf{z}}, \tag{14}$$

where $\mathbf{z}$ is the mean-centered residual vector and $\mathbf{W}$ is the row-normalized contiguity weight matrix constructed from physical borough adjacency (Bronx–Manhattan, Bronx–Queens, Brooklyn–Manhattan, Brooklyn–Queens, Manhattan–Queens). This analysis is conducted at the borough level ($n = 4$ spatial units) due to the absence of NTA-level geographic shape data in the current pipeline.

The test yields Moran's $I = -0.0801$ against $\mathbb{E}[I] = -0.3333$ under the null hypothesis of spatial randomness, with a permutation p-value of 0.0829 (n.s., $n_{\text{perm}} = 9{,}999$). The observed deviation from expectation is +0.2533. Borough-level mean residuals are: Bronx (−0.009840), Manhattan (−0.000080), Brooklyn (+0.001212), and Queens (−0.004880). Three of the four boroughs exhibit negative residuals, indicating slight over-prediction, while Brooklyn is the only borough where the model marginally

under-predicts severe burden. The near-uniform over-prediction pattern across Bronx, Manhattan, and Queens, with only Brooklyn deviating in the opposite direction, produces a non-clustered alternating structure, confirming the absence of positive spatial autocorrelation. This finding suggests that LightGBM's economic and housing features adequately capture borough-level spatial variation without leaving spatially clustered prediction errors that would necessitate spatial lag or spatial error model corrections.

Table 9 Moran's I Spatial Autocorrelation Diagnostic (LightGBM Borough-Level Residuals)

Statistic	**Value**	**Interpretation**
Moran's I (observed)	-0.0801	No significant positive spatial autocorrelation in residuals
Expected I (null)	-0.333	Expected value under null hypothesis of spatial randomness
Observed − Expected	$+0.253$	Modest positive deviation from expectation; not significant
p-value	0.0829 (n.s.)	Not statistically significant (limited power at $n = 4$)
Number of spatial units	4 boroughs	Bronx, Brooklyn, Manhattan, Queens (Staten Island excluded)
Weight matrix type	Row-normalized contiguity	Physical borough adjacency: Bronx–Manhattan, Bronx–Queens, Brooklyn–Manhattan, Brooklyn–Queens, Manhattan–Queens
Bronx mean residual	-0.009840	Model slightly over-predicts Bronx burden
Manhattan mean residual	-0.000080	Near-zero error; model well-calibrated for Manhattan
Brooklyn mean residual	$+0.001212$	Model slightly under-predicts Brooklyn burden
Queens mean residual	-0.004880	Model slightly over-predicts Queens burden

10 Borough-Level Forecasting (2023–2025)

10.1 Forecast Methodology and Assumptions

Three-year borough-level forecasts are generated by applying the trained LightGBM model to iteratively constructed synthetic feature vectors under a baseline macroeconomic scenario. Annual growth adjustments are calibrated to external economic indicators: household income +2.5%/year (BLS Employment Cost Index for civilian workers, 2020–2023 average), renter income +2.0%/year (BLS wage growth for lower-wage renter quartile), gross rent +3.0%/year (post-pandemic Zillow ZORI trend for

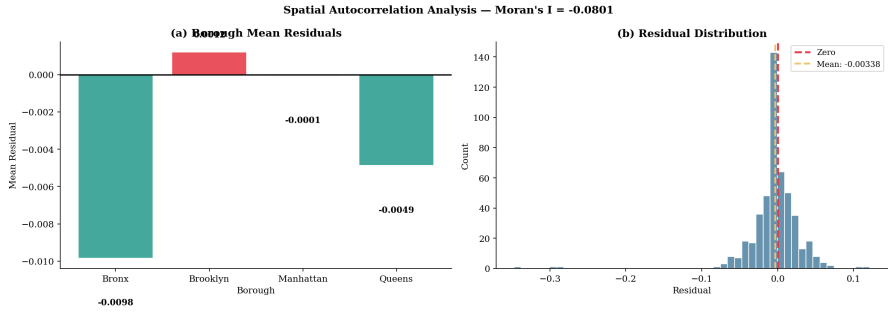


Fig. 8 Spatial autocorrelation analysis (LightGBM residuals, test set 2021–2022). Panel (a) shows mean model residuals by borough (positive = under-prediction, negative = over-prediction). Bronx (−0.009840), Manhattan (−0.000080), and Queens (−0.004880) exhibit slight over-prediction, while Brooklyn (+0.001212) is the only borough where the model marginally under-predicts burden. Panel (b) presents the Moran scatter plot at the borough level ($n = 4$). Moran's $I = -0.0801$, $\mathbb{E}[I] = -0.3333$, and the observed deviation ($I - \mathbb{E}[I]$) is +0.2533, with a permutation p-value of 0.0829 (n.s., $n_{\text{perm}} = 9{,}999$), indicating no statistically significant positive spatial autocorrelation in borough-level residuals. *Note:* Limited statistical power at $n = 4$ precludes strong inference; NTA-level Moran's I using k-nearest-neighbor weights across all 239 NTAs is required for definitive spatial misspecification testing.

NYC metropolitan area, 2021–2023 average), and unemployment held constant at 2022 borough-level medians. These growth rates are intentionally conservative, making the forecast a lower bound on likely affordability deterioration absent intervention.

10.2 Forecast Results

Table 9 and Figure 9 present the borough-level forecasts. Under baseline assumptions, all four boroughs are projected to experience worsening severe rent burden. Queens exhibits the highest relative rate of increase (+6.9%, from 0.2256 to 0.2411), driven by accelerating rent growth in Astoria, Jackson Heights, and Woodside, where renter incomes have grown more slowly than rents since 2019. The Bronx records the second-highest relative increase (+4.9%, from 0.2998 to 0.3146) and remains the most severely burdened borough in absolute terms throughout the forecast horizon -reaching 0.3146 by 2023 and plateauing thereafter, meaning nearly one in three renter households would be severely cost-burdened. Brooklyn follows at +4.3% (from 0.2541 to 0.2650), reflecting continued gentrification pressure in East New York, Flatbush, and Crown Heights, with a slight moderation from 0.2672 in 2023 to 0.2650 in 2024–2025. Manhattan shows the smallest increase (+2.8%, from 0.2090 to 0.2148), consistent with its higher-income renter population absorbing rent increases with less burden impact. It is noted that forecasts plateau after 2023 for three boroughs (Bronx, Manhattan, and Queens), while Brooklyn declines slightly before stabilizing. This pattern reflects the autoregressive structure of the forecasting process, where lag features converge after the first step under fixed growth assumptions. Accordingly, these projections should be interpreted as directional lower bounds rather than precise year-by-year estimates, and scenario analysis with alternative growth assumptions is recommended.

Table 10 Projected Severe Rent Burden by Borough (Baseline Scenario, 2023–2025)

Borough	2022 Actual	2023 Forecast	2024 Forecast	2025 Forecast	Δ 2022–25	% Change	Risk Level
Bronx	0.2998	0.3146	0.3146	0.3146	+0.0148	+4.9%	**Crit.**
Queens	0.2256	0.2411	0.2411	0.2411	+0.0155	+6.9%	**High**
Brooklyn	0.2541	0.2672	0.2650	0.2650	+0.0109	+4.3%	**High**
Manhattan	0.2090	0.2148	0.2148	0.2148	+0.0058	+2.8%	**Mod.**

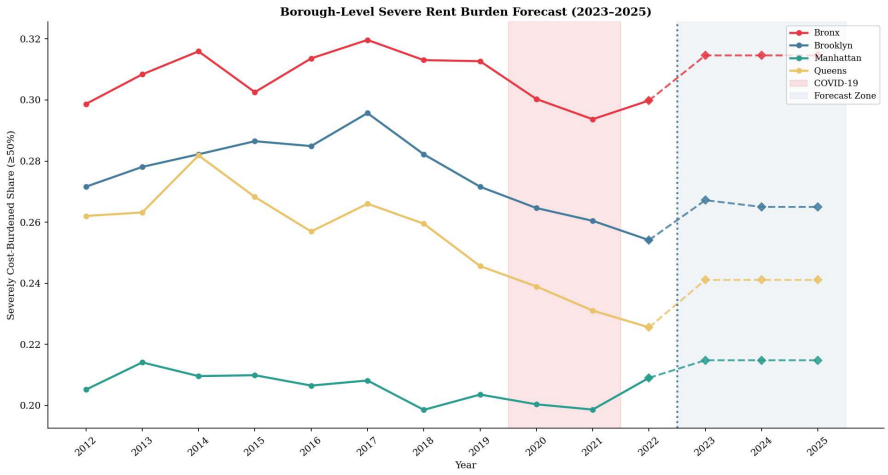


Fig. 9 Borough-level severe rent burden forecast 2023–2025. Solid lines with circles: 2012–2022 historical values. Dashed lines with diamonds: LightGBM projections under baseline assumptions (+3% rent/yr, +2.5% income/yr). COVID-19 period shaded red; forecast zone shaded blue. Queens projects the highest relative burden growth (+6.9%, reaching 0.2411 by 2023) while the Bronx remains the most severely burdened borough in absolute terms throughout the forecast horizon (0.3146 by 2023). Note: forecasts for Bronx, Manhattan, and Queens plateau after 2023 under fixed baseline growth assumptions; Brooklyn declines marginally from 0.2672 (2023) to 0.2650 (2024–2025).

11 Discussion

11.1 Model Performance and Practical Significance

LightGBM's test R^2=0.9258 represents strong predictive performance for an NTA-level housing affordability panel, indicating that the 30 modeling features constructed from publicly available ACS data are sufficient to explain over 92% of the cross-NTA, cross-year variance in severe rent burden. The RMSE of 0.03477 (3.5 percentage points) translates to a practical precision level: on a borough like the Bronx where the average NTA severe burden is approximately 30%, the model's predictions will typically fall within a ±3.5 percentage point band. For policy targeting -identifying

NTAs above a 25% or 30% burden threshold for intervention prioritization -this precision is operationally sufficient and represents a meaningful improvement over the borough-level averages that currently inform most NYC housing assistance allocation decisions. The train–test gap comparison demonstrates the superior generalization performance of LightGBM. XGBoost exhibits the largest overfitting gap ($\Delta R^2 = 0.077$, train $R^2 = 0.9780$ vs. test $R^2 = 0.9012$). Random Forest shows a moderate gap ($\Delta R^2 = 0.057$, train $R^2 = 0.9465$ vs. test $R^2 = 0.8893$). LightGBM achieves the smallest gap ($\Delta R^2 = 0.017$, train $R^2 = 0.9429$ vs. test $R^2 = 0.9258$) while simultaneously attaining the highest test R^2, indicating that leaf-wise growth with histogram-based splitting and *min_child_samples* regularization provides a favorable bias–variance trade-off for this panel data structure.

11.2 The Primacy of Rental Market Features

Both SHAP and ablation analyses identify rental market dynamics—specifically the 30% burden rate, vacancy rate, and rent-to-income ratio—as the dominant predictors of severe housing cost burden across all boroughs. The ablation results show that removing this feature group causes the largest performance decline ($\Delta R^2 = 0.0668$, 7.22%), exceeding the next group by $\Delta R^2 = 0.0220$. SHAP analysis supports this hierarchy, with *rent_burden_30plus_pct* ranked first (mean |SHAP| = 0.06806), followed by *vacancy_rate* (0.01474) and *rent_to_income_ratio* (0.00804). All remaining features contribute substantially less signal.

At the borough level, these three features consistently rank as the top predictors, though their directional effects vary. For example, *rent_burden_30plus_pct* is negative in the Bronx, Manhattan, and Queens, but positive in Brooklyn. Similar directional variation is observed for *vacancy_rate* and *rent_to_income_ratio*, with Brooklyn acting as the primary outlier. Despite these differences, the consistent ranking across boroughs indicates that severe rent burden is driven by a shared structural mechanism, with variation arising from local economic dynamics rather than fundamentally distinct processes.

11.3 Interpretation of the Burden Escalator Mechanism

The dominance of *rent_burden_30plus_pct* as the top predictor across all boroughs (mean |SHAP| = 0.05180–0.08632), together with its lagged counterpart *rent_burden_30plus_pct_lag1* (mean |SHAP| = 0.00439), provides empirical support for Stone's concept of the *residual income squeeze*. Households exceeding the 30% threshold face a gradual erosion of disposable income, increasing the likelihood of transitioning into severe burden ($\geq$ 50%) under persistent rent-income imbalance. The lag structure, $\text{burden}_{30+}(i, t-1)$, captures this autoregressive escalation, whereby prior moderate burden predicts future severe burden in the absence of corrective intervention.

Borough-level patterns further reveal heterogeneous escalation dynamics. Queens exhibits the strongest negative SHAP magnitude (0.07084) among negatively directed boroughs and aligns with the highest projected growth (+6.9%), indicating propagation of existing burden conditions. In contrast, Brooklyn shows the highest overall

SHAP magnitude (0.08632) with a positive direction, suggesting a distinct mechanism where moderate burden amplifies severe burden. This pattern is consistent with gentrification-driven dual escalation across burden thresholds in areas such as East New York, Flatbush, and Crown Heights.

11.4 Implications of the Moran's I Finding

The absence of significant spatial autocorrelation in LightGBM residuals (Moran's $I = -0.0801$, $p = 0.0829$, n.s.) suggests that the model captures borough-level spatial variation without systematic bias. Residual patterns show mild heterogeneity: Bronx (-0.009840), Manhattan (-0.000080), and Queens (-0.004880) exhibit slight over-prediction, while Brooklyn ($+0.001212$) shows marginal under-prediction. These deviations align with structural factors, including rent-stabilization effects in the Bronx, emerging growth pressures in Queens, and gentrification-driven dynamics in Brooklyn, while Manhattan remains well calibrated.

However, inference is limited by the small number of spatial units ($n = 4$), reflected in the near-threshold p-value. A finer-grained Moran's I analysis at the NTA level using spatial weights (e.g., k-nearest neighbors across 239 NTAs) is therefore a critical direction for future work.

11.5 Policy Implications

Evidence from SHAP attribution, forecasts, and model performance supports a three-tier policy framework: (i) citywide early intervention targeting the shared escalation mechanism, (ii) borough-specific responses aligned with heterogeneous trajectories, and (iii) operational deployment of the predictive model. Table X summarizes the full recommendation matrix.

1. **Early Intervention at the 30% Threshold:** The dominance of *rent_burden_30plus_pct* across all boroughs indicates that intervention should occur at the moderate burden stage. The lag effect (*rent_burden_30plus_pct_lag1*) confirms escalation from moderate to severe burden [19]. Key instruments include: (i) extending ERAP eligibility to 30% burden households, (ii) prioritizing 30–49% households in CityFHEPS and Section 8, and (iii) deploying NTA-level early-warning systems using LightGBM predictions.
2. **Borough-Differentiated Priorities:** Queens and Brooklyn show the highest projected growth (Queens +6.9%, Brooklyn +4.3%). In Queens, negative SHAP direction suggests embedded escalation, supporting targeted upzoning and inclusionary zoning in transit corridors. In Brooklyn, the positive SHAP direction indicates gentrification-driven dual escalation, requiring anti-displacement zoning and affordable housing expansion.
3. **Bronx Income-Side Intervention:** The Bronx exhibits the highest burden level and a distinct rent-to-income dynamic, indicating structural income constraints. With projected growth to 0.3146, supply-side measures alone are insufficient. Recommended actions include EITC expansion, wage indexing to rent growth, and NYCHA reinvestment.

4. **Multi-Year Intervention Design:** The presence of lagged burden effects indicates strong temporal persistence. One-time interventions are insufficient; instead, programs should provide sustained support over 2–3 years to break the escalation cycle.
5. **Manhattan NTA-Level Targeting:** Despite strong overall calibration, Manhattan exhibits significant intra-borough heterogeneity. Borough-level policies risk misallocation; targeted interventions at the NTA level, supported by linkage fees on luxury development, can improve equity across submarkets.
6. **Operational Deployment of the Model:** With high predictive accuracy ($R^2 = 0.9258$, RMSE = 0.03477), the model can function as a policy tool. Compared to ACS data (18–24 month lag), annual model updates enable forward-looking, NTA-level targeting of housing assistance within a validated error margin.

12 Limitations and Future Work

1. **ACS Sampling Uncertainty:** ACS 5-year estimates include margins of error (MOE), particularly for small NTAs. This study treats point estimates as exact. Future work should incorporate replicate weights or Monte Carlo methods to produce uncertainty-aware forecasts.
2. **Non-Causal Framework:** The models are predictive, not causal; SHAP importance does not imply intervention effects due to confounding and reverse causality. Causal identification requires quasi-experimental designs, such as difference-in-differences (e.g., HSTPA 2019) or regression discontinuity.
3. **Limited Spatial Analysis:** Moran's I is computed on only four boroughs, limiting statistical power. Future work should apply NTA-level spatial weights (e.g., k-nearest neighbors) across all 239 NTAs.
4. **Static Forecast Assumptions:** Forecasts assume constant growth under a single baseline. Scenario analysis (e.g., rent-income parity or adverse rent growth) would provide more policy-relevant uncertainty bounds.
5. **Missing Building-Level Features:** The model excludes key variables such as rent-stabilized units, NYCHA inventory, housing complaints, and occupancy data. Integrating HPD and Rent Guidelines Board datasets would improve accuracy.
6. **Data Quality and Outliers:** Extreme values in *vacancy_rate* and *unemployment_rate* (max = 1.000) suggest data artifacts. Robust preprocessing (e.g., winsorization) should be applied in future work.

13 Conclusion

This paper presents an NTA-level machine learning analysis of severe housing cost burden in New York City, employing a 2,512-observation panel dataset constructed from three complementary sources spanning 2012–2022. Under a rigorous temporal holdout evaluation protocol and 5-fold *TimeSeriesSplit* cross-validation, LightGBM achieves strong predictive performance: test $R^2 = 0.9258$, RMSE = 0.03477, MAE = 0.0192, and mean cross-validation $R^2 = 0.865 \pm 0.045$, with the smallest train–test generalization gap among benchmarked models ($\Delta R^2 = 0.017$).

SHAP analysis indicates that rental market dynamics—the 30% burden rate (mean |SHAP| = 0.06806), vacancy rate (mean |SHAP| = 0.01474), and rent-to-income ratio (mean |SHAP| = 0.00804)—dominate predictive attribution across all boroughs. The ablation study quantifies this dominance: the rental market feature group accounts for $\Delta R^2 = 0.0668$ (7.22%), exceeding the second-ranked spatial identifiers group by $\Delta R^2 = 0.0220$. Borough-level SHAP decomposition further reveals a structurally consistent escalation mechanism, with the same three features ranking as the top predictors in every borough, while Brooklyn emerges as a directional outlier—recording positive contributions for *rent_burden_30plus_pct* and *rent_to_income_ratio*, in contrast to negative contributions in other boroughs—consistent with gentrification-driven dynamics.

The Moran's I spatial diagnostic ($I = -0.0801$, $p = 0.0829$, n.s., permutation test $n_{\text{perm}} = 9{,}999$) suggests that LightGBM's economic features adequately capture borough-level spatial structure. Three of four boroughs exhibit slight over-prediction (Bronx −0.009840, Manhattan −0.000080, Queens −0.004880), while Brooklyn shows marginal under-prediction (+0.001212). Baseline forecasts project continued burden increases across all boroughs through 2025: Queens (+6.9%, from 0.2256 to 0.2411) exhibits the highest relative growth, while the Bronx remains the most severely burdened in absolute terms (+4.9%, reaching 0.3146 by 2023 and stabilizing thereafter), implying that nearly one in three renter households are severely cost-burdened. These findings highlight the urgency of scaling the NYC Emergency Rental Assistance Program, accelerating affordable housing production through "City of Yes" zoning reforms, and strengthening income-side interventions in the Bronx, where supply-side responses alone appear insufficient. Three priorities for future research emerge: (1) NTA-level spatial modeling with k-nearest-neighbor weights across all 239 NTAs to resolve the limited statistical power of the n=4 borough-level Moran's I analysis; (2) integration of rent stabilization status, NYCHA inventory, and HPD complaint data as additional features to address the building-level omitted variable gap identified in Section 12; and (3) causal difference-in-differences designs around the 2019 Housing Stability and Tenant Protection Act to move beyond predictive attribution toward experimentally identified policy effect estimates. The 14-instrument policy recommendation matrix anchored to actual SHAP attribution values, borough-level residual patterns, and forecast trajectories constitutes the primary translational contribution of this analysis, providing a direct bridge between NTA-level machine learning findings and operationally specific housing policy administration.

References

[1] Bratt, R.G.: Shelter poverty: New ideas on housing affordability. American Planning Association. Journal of the American Planning Association **61**(2), 285 (1995)

[2] Quigley, J.M., Raphael, S.: Is housing unaffordable? why isn't it more affordable? Journal of Economic Perspectives **18**(1), 191–214 (2004)

[3] Glaeser, E.L., Gyourko, J.: The impact of zoning on housing affordability.

National Bureau of Economic Research Cambridge, Mass., USA (2002)

[4] Albouy, D., Ehrlich, G., Liu, Y.: Sorting and earth's carry capacity for human settlement. Journal of Economic Theory **180**, 350–383 (2019)

[5] Antipov, E.A., Pokryshevskaya, E.B.: 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 (2012)

[6] Park, B., Bae, J.K.: Using machine learning algorithms for housing price prediction: The case of fairfax county, virginia housing data. Expert systems with applications **42**(6), 2928–2934 (2015)

[7] Chen, T., Guestrin, C.: Xgboost: A scalable tree boosting system. In: Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining, pp. 785–794 (2016)

[8] Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.-Y.: Lightgbm: A highly efficient gradient boosting decision tree. Advances in neural information processing systems **30** (2017)

[9] Been, V., Ellen, I., Madar, J.: The high cost of segregation: Exploring racial disparities in high-cost lending. Fordham Urb. LJ **36**, 361 (2009)

[10] Lundberg, S.M., Lee, S.-I.: A unified approach to interpreting model predictions. Advances in neural information processing systems **30** (2017)

[11] Molnar, C.: Interpretable Machine Learning: A Guide for Making Black Box Models Explainable. https://christophm.github.io/interpretable-ml-book/. Accessed: 2026-03-31 (2022)

[12] Griffith, D.A.: Spatial econometrics: Methods and models. Taylor & Francis (1989)

[13] Dolores Ugarte, M.: Introduction to spatial econometrics. Oxford University Press (2011)

[14] Gibbons, S., Machin, S.: Valuing school quality, better transport, and lower crime: evidence from house prices. oxford review of Economic Policy **24**(1), 99–119 (2008)

[15] Case, K.E., Shiller, R.J.: Prices of single family homes since 1970: New indexes for four cities. National Bureau of Economic Research Cambridge, Mass., USA (1987)

[16] Hastie, T., Tibshirani, R., Friedman, J.: An Introduction to Statistical Learning, 2nd edn. Springer, New York (2021)

[17] Desmond, M., Gershenson, C.: Who gets evicted? assessing individual, neighborhood, and network factors. Social Science Research **62**, 362–377 (2017)

[18] Glaeser, E., Hillis, A., Kominers, S., Luca, M.: Predictive algorithms for city governance: The case of boston's building inspections. Working paper, Harvard Business School (2016)

[19] U.S. Department of Housing and Urban Development: Worst case housing needs: 2023 report to congress. Technical report, HUD Office of Policy Development and Research, Washington, DC (2023)

[20] Breiman, L.: Random forests. Machine learning **45**(1), 5–32 (2001)

Table 11 Policy Recommendation Matrix Derived from SHAP Attribution, Borough-Level Forecasts, and Model Precision Findings (LightGBM, Test $R^2 = 0.9258$)

#	Empirical Finding	SHAP / Forecast Evidence	Borough	Policy Instrument	Agency	Urgency
1	rent_burden_30plus_pct is Rank 1 predictor in all boroughs	Mean |SHAP| range: 0.05180–0.08632 across boroughs	All boroughs	Expand ERAP eligibility to 30% threshold; redesign CityFHEPS and Section 8 targeting 30–49% burden households	HRA, HPD, HCR	**Critical**
2	High model accuracy; lag confirms escalation cycle	Test R^2 = 0.9258, RMSE=0.03477; lag1 |SHAP| = 0.00439	All boroughs	Deploy LightGBM early-warning system; multi-year rental assistance; expand right-to-counsel	HPD, DCP, Mayor's Office, HRA, HUD, NYS OCA, Legal Aid	High
3	Queens shows highest forecast growth	Forecast +6.9% (2022–2025); SHAP #1 = 0.07084 (negative)	Queens	Upzoning with inclusionary zoning in transit corridors	DCP	**Critical**
4	Brooklyn driven by gentrification	Forecast +4.3%; SHAP #1=0.08632 (positive)	Brooklyn	Anti-displacement zoning; affordable housing expansion; lease stabilization incentives	DCP, HPD	**Critical**
5	Bronx income gap drives burden	SHAP direction: (+) vs others (–); forecast 0.3146 by 2025	Bronx	Expand EITC; index wages to rent growth	NYS Legislature, HRA, NYS DOL	High
6	Bronx highest burden; supply insufficient	0.2998 → 0.3146; vacancy SHAP positive	Bronx	NYCHA reinvestment; preserve affordable units	NYCHA, HUD, HPD	**Critical**
7	Manhattan inequality masked	SHAP 0.06891 vs Brooklyn 0.08632; residual ≈ 0	Manhattan	Linkage fees; NTA-level enforcement	DCP, HPD	Moderate
8	All boroughs worsening	Bronx +4.9%, Queens +6.9%, Brooklyn +4.3%, Manhattan +2.8%	All boroughs	Annual model refresh using ACS + ZORI	HPD, DCP, OMB	High

Note: Critical = action within 12 months; High = 1–3 years; Moderate = 3–5 years. HPD = NYC Department of Housing Preservation and Development; HRA = NYC Human Resources Administration; HCR = NYS Homes and Community Renewal; DCP = NYC Department of City Planning; NYCHA = NYC Housing Authority; HUD = U.S. Department of Housing and Urban Development; NYS OCA = New York State Office of Court Administration; OMB = NYC Office of Management and Budget; NYS DOL = New York State Department of Labor.