Mathematical and Computational Applications

MDPI

Article

Evaluating Regularization Estimators Under Severe Multicollinearity: A Simulation and Empirical Study on Housing Prices

Osman Ufuk Ekiz * and Meltem Ekiz

Department of Statistics, Faculty of Sciences, Gazi University, Ankara 06560, Turkey; ozmeltem@gazi.edu.tr
* Correspondence: ufukekiz@gazi.edu.tr

Abstract

Accurate housing price prediction is important for market efficiency and purchasing decisions. However, multicollinearity among independent variables remains a major challenge in linear regression, causing variance inflation and reducing the reliability of the ordinary least squares (OLS) estimator. Although regularization methods such as ridge regression, least absolute shrinkage and selection operator (LASSO), and elastic net (EN) are widely used, evidence regarding their variance behavior under controlled multicollinearity structures remains limited. This study addresses this gap through simulation experiments conducted under controlled correlation structures with sample sizes ranging from 100 to 2000, 5 to 70 independent variables, and correlation coefficients between 0.50 and 0.99. The findings are further validated using the California Housing Dataset, where mean squared prediction error (MSPE) is computed on the full dataset, while root mean squared error (RMSE), mean absolute error (MAE), and the coefficient of determination (R^2) are evaluated on a hold-out test set. Simulation results show that LASSO generally yields the lowest variance estimates under moderate multicollinearity, whereas EN becomes more competitive as multicollinearity and dimensionality increase. In the California Housing application, EN reduces MSPE by approximately 95.5% relative to OLS. These findings provide insight into the behavior of linear regression estimators and offer practical guidance for researchers in selecting appropriate models for housing price modelling.

Keywords: multicollinearity; regularization; mean squared prediction error; predictive accuracy

Academic Editors: Sandra Ferreira and Dário Ferreira

Received: 12 May 2026
Revised: 16 June 2026
Accepted: 17 June 2026
Published: 19 June 2026

1. Introduction

The housing sector is undergoing rapid transformation on a global scale due to digitalization, sustainability-oriented regulations, and evolving market dynamics. As a result, the structure of housing markets has become increasingly complex, making accurate price prediction a critical issue for market participants. Accurate pricing is essential not only for competitiveness but also for effective decision-making in areas such as investment planning, credit risk assessment, and policy development. However, the presence of highly correlated independent variables often leads to multicollinearity in regression models used for housing price prediction. In such cases, classical approaches based on the ordinary least squares (OLS) estimator tend to produce unstable parameter estimates, inflated variance, and overfitting [1–3].

In recent years, housing price prediction studies have increasingly employed machine learning techniques such as random forests, gradient boosting algorithms, and neural

 https://doi.org/10.3390/mca31030111

networks due to their strong predictive capabilities [4–6]. However, despite their predictive success, such models often suffer from limited interpretability and high computational complexity. Consequently, regularization-based linear models remain attractive alternatives because they provide interpretable coefficient estimates while effectively mitigating multicollinearity through shrinkage mechanisms [7–10].

The application of regularization techniques to housing price modeling has evolved from limited use to a well-established methodological approach. Early empirical studies demonstrated that ridge and least absolute shrinkage and selection operator (LASSO) estimators effectively address multicollinearity issues in datasets such as the Ames Housing Dataset, improving model stability and interpretability [11]. With the growing integration of machine learning, subsequent studies incorporated LASSO as a benchmark model alongside more complex algorithms, highlighting its continued relevance in predictive performance comparisons [12]. In parallel, time-series approaches emerged, where elastic net (EN), ridge regression, and LASSO were embedded within vector autoregression frameworks to capture regional spillover effects in housing markets [13]. More recent contributions have emphasized systematic comparisons of regularization techniques; for instance, several studies evaluated these models on the California Housing Dataset, generally finding that ridge regression provides stable performance while LASSO offers advantages in feature selection and model sparsity [14,15]. Additional empirical evidence indicates that EN improves predictive performance in residential construction demand forecasting by addressing multicollinearity and balancing the bias–variance trade-off [16], while LASSO enables effective dimensionality reduction and variable selection in macroeconomic housing price models, as demonstrated in the Chinese housing market context [17].

Although previous studies have demonstrated the effectiveness of ridge, LASSO, and EN estimators in handling multicollinearity [18–20], most existing research has primarily focused either on predictive performance or on specific empirical applications [21,22]. Existing studies suggest that the relative performance of ridge, LASSO, and EN is context-dependent, varying with data characteristics such as multicollinearity structure, sample size, and evaluation criteria [8,12]. Comparatively limited attention has been devoted to systematically investigating the variance behavior of these estimators under controlled multicollinearity structures while simultaneously validating the findings using real-world housing datasets within a unified analytical framework. Furthermore, many empirical studies rely on a single modelling setting and provide limited discussion regarding the stability–sparsity trade-off among regularization approaches under varying dimensionality and correlation levels.

Building on the existing literature, the present study provides a systematic comparison of OLS, ridge [18,23], LASSO [19,24,25], and EN [20,26] estimators under varying degrees of multicollinearity. Unlike many previous studies that focus solely on predictive accuracy or a single empirical application, this study combines controlled Monte Carlo simulation experiments with a real-world housing price application to evaluate both estimator variance and predictive performance within a unified framework. The simulation design considers different sample sizes, dimensional structures, and correlation levels in order to examine how multicollinearity affects estimator stability across diverse modelling scenarios. Subsequently, the California Housing Dataset [27] is used to empirically validate the simulation findings through prediction-based performance measures.

The main contributions of this study can be summarized as follows:

1. the variance behavior of OLS, ridge, LASSO, and EN estimators is systematically evaluated under controlled multicollinearity scenarios;
2. the effects of varying sample size, dimensionality, and correlation intensity on estimator performance are comparatively analyzed;

3. simulation findings are supported through an empirical application using the California Housing Dataset;
4. the study provides practical insights into the trade-off between sparsity, stability, and predictive accuracy in housing price modelling.

Accordingly, the study seeks to address the following research questions:

- How does increasing multicollinearity affect the variance behavior of OLS, ridge, LASSO, and EN estimators?
- Under which dimensional and correlation settings do regularization-based estimators outperform OLS in terms of stability and predictive performance?
- Which estimator provides the most suitable balance between sparsity, variance reduction, and predictive accuracy in housing price modelling?

The remainder of the paper is organized as follows. Section 2 presents the methodological framework, including the regression estimators, simulation design, and dataset description. Section 3 reports the simulation and empirical findings. Section 4 discusses the results in relation to the existing literature and practical implications. Finally, Section 5 concludes the study by summarizing the main findings, limitations, and directions for future research.

2. Materials and Methods

This section introduces the linear regression framework and the corresponding parameter estimators to establish the methodological foundation of the study. Multiple linear regression is used to model the relationship between a dependent variable Y and p independent variables as follows:

$$Y = X\beta + \varepsilon. \tag{1}$$

Y is an $n \times 1$ vector of observations of the dependent variable, X is an $n \times p$ matrix of independent variables, and β is a $p \times 1$ vector of unknown regression parameters. The term ε is an $n \times 1$ vector of unobservable random errors, which are assumed to be independent and normally distributed with mean zero and variance σ^2. In this study, we investigate four different estimators of the parameter vector β in Equation (1).

2.1. Estimators

2.1.1. OLS Estimator

The OLS estimator of the parameter β in Equation (1) is obtained by minimizing the residual sum of squares (RSS), given by

$$\underset{\beta}{\operatorname{argmin}} \sum_{i=1}^{n} \left(Y_i - \sum_{j=1}^{p} \beta_j X_{ij} \right)^2, \tag{2}$$

which yields the closed-form solution [28,29]

$$\hat{\beta}^{OLS} = \left(X'X \right)^{-1} X'Y. \tag{3}$$

In the presence of multicollinearity, the determinant of the matrix $X'X$ becomes close to zero, which in turn inflates the variance of the OLS estimator in Equations (2) and (3). A commonly used diagnostic measure for detecting multicollinearity is the variance inflation factor (VIF). The VIF for the jth independent variable is defined as

$$\text{VIF}_{(j)} = \frac{1}{1 - R_j^2}, \tag{4}$$

$j = 1, 2, \ldots, p-1$, where R_j^2 denotes the coefficient of determination obtained by regressing this variable on the remaining independent variables [30]. A VIF value greater than 5 indicates the presence of severe multicollinearity in the data. To evaluate the model's predictive accuracy, the mean squared prediction error (MSPE) is employed as a performance metric defined as the average of squared differences between observed and predicted values and is given by:

$$\text{MSPE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2, \tag{5}$$

providing an overall measure of prediction error that assigns greater weight to larger deviations. Here, y_i represents the observed value, whereas $\hat{y}_i$ represents the predicted value for the ith observation.

2.1.2. Ridge Estimator

The ridge estimator of the parameter β is obtained by minimizing the following objective function, which is formed by adding an $L_2 = \lambda\sum_{j=1}^{p}\beta_j^2$ penalty term consisting of the sum of squared parameters to the RSS:

$$\underset{\beta}{\text{argmin}}\left\{\sum_{i=1}^{n}\left(Y_i - \sum_{j=1}^{p}\beta_j X_{ij}\right)^2 + \lambda\sum_{j=1}^{p}\beta_j^2\right\}. \tag{6}$$

Minimizing this objective function gives the ridge estimator as

$$\hat{\beta}^{Ridge} = \left(X'X + \lambda I\right)^{-1}X'Y. \tag{7}$$

In Equations (6) and (7), $\lambda \geq 0$ denotes the regularization parameter, and its optimal value is typically determined through cross-validation [31,32]. The ridge estimator approaches zero as $\lambda \to \infty$, and converges to the OLS estimator as $\lambda \to 0$. Moreover, ridge regression keeps all independent variables in the model but applies a penalty that shrinks the coefficient magnitudes [33]. The primary strength of this estimator is its ability to produce lower-variance estimates than OLS under high multicollinearity, which improves the stability and reliability of the predictions [34].

It is important to note that the OLS and ridge estimators admit explicit closed-form solutions, as their objective functions are convex and differentiable. In the case of OLS, strict convexity holds under full column rank conditions, whereas the inclusion of the L_2 penalty in ridge regression ensures strict convexity and a unique solution. This property allows the corresponding parameter estimates to be obtained analytically. In contrast to OLS and ridge regression, which admit closed-form analytical solutions, the inclusion of the non-differentiable $L_1 = \lambda\sum_{j=1}^{p}|\beta_j|$ penalty term in the LASSO and EN objective functions means that these estimators do not admit closed-form solutions. Consequently, their parameter estimates are obtained via numerical optimization procedures, most commonly the coordinate descent algorithm.

2.1.3. LASSO Estimator

The LASSO estimator differs from other estimators by incorporating an L_1 penalty term, defined as the sum of the absolute values of the parameters, into the objective function presented in Equation (6). Minimization of this objective function via the coordinate descent algorithm yields the LASSO estimator $\hat{\beta}^{LASSO}$ [35,36]

$$\hat{\beta}^{LASSO} = \underset{\beta}{\text{argmin}}\left\{\sum_{i=1}^{n}\left(Y_i - \sum_{j=1}^{p}\beta_j X_{ij}\right)^2 + \lambda\sum_{j=1}^{p}|\beta_j|\right\}. \tag{8}$$

By shrinking some coefficients exactly to zero, the L_1 penalty simplifies the model and improves its generalization performance, particularly in datasets containing many irrelevant or noisy independent variables. In addition, it helps address the multicollinearity problem, as it tends to retain only one predictor from a group of highly correlated independent variables [37–39].

2.1.4. EN Estimator

The EN estimator combines L_1 and L_2 penalties, resulting in both variable selection and improved stability of the estimated parameters. The EN estimator is obtained by minimizing the following objective function using the coordinate descent algorithm:

$$\hat{\beta}^{EN} = \underset{\beta}{\text{argmin}} \left\{ \frac{1}{2}\sum_{i=1}^{n} \left(Y_i - \sum_{j=1}^{p} \beta_j X_{ij} \right)^2 + \left[\frac{1}{2}(1-\alpha)\lambda \sum_{j=1}^{p} \beta_j^2 + \alpha\lambda \sum_{j=1}^{p} |\beta_j| \right] \right\}, \quad (9)$$

where $0 \leq \alpha \leq 1$. When α = 1 the estimator reduces to the LASSO estimator, whereas α = 0 corresponds to the ridge estimator. The hyperparameters λ and α are generally determined using grid search or k-fold cross-validation procedures. EN is particularly advantageous in datasets where the number of independent variables is close to or exceeds the number of observations and strong correlations exist among them. By regularizing and shrinking the coefficients, it simplifies the model and promotes an optimal balance [39].

A comprehensive simulation study is provided in the following section to evaluate which of these estimators performs best under severe multicollinearity and high-dimensional conditions.

2.2. Simulation Design

A simulation study is conducted in MATLAB R2023b to investigate the predictive behavior of the estimators under controlled experimental conditions. Within this framework, datasets are generated according to the regression model specified in Equation (1), where the independent variables are drawn from a multivariate normal distribution and the error vector ε consists of independent and identically distributed random variables with $\varepsilon_i \sim$ N (0, 1). To systematically control the degree of multicollinearity among independent variables, the variance–covariance matrix is defined as

$$\sum_{ij} = G^{|i-j|}, \cdots i,\ j\ =\ 1,\ 2,\ \ldots,\ p,\ \ 0 < G < 1. \quad (10)$$

G represents the correlation parameter controlling the severity of multicollinearity among p independent variables [40]. For each estimator, the total variance of $\hat{\beta}$ is obtained as the trace of its covariance matrix,

$$\text{trace}\left(\hat{\Sigma}_{\hat{\beta}}\right) = \sum_{j=1}^{p} \hat{\sigma}^2_{\hat{\beta}_j}, \quad (11)$$

which corresponds to the sum of the variances of the parameter estimates [41]. This criterion is used because multicollinearity primarily inflates estimator variance and destabilizes coefficient estimates in classical regression models. The parameter vector β is constructed as a $1 \times p$ vector, where the first $\lfloor p/2 + 1 \rfloor$ components are set to 1 and the remaining components are set to 0. This sparse coefficient structure is selected to reflect variable-selection settings frequently encountered in regularization-based regression studies. For the estimators defined in Equations (6)–(9), the optimal regularization parameter λ is selected through 10-fold cross-validation over the range 10^{-4} to 10^4 by minimizing the mean squared prediction error (MSPE) given in Equation (5). Accordingly, the optimal λ values are obtained as 4.0091 for ridge regression, 3.8269 for LASSO, and 0.0966 for EN. For the EN estimator, the mixing parameter α is fixed at 0.50. To evaluate the performance of the estimators, 10,000 simulation replications are carried out. In each replication, the synthetic data set is regenerated and all estimators are re-estimated. The simulation framework considers sample sizes n = {100, 500, 1000, 2000}, numbers of predictors p = {5, 10, 30, 50, 70} and correlation coefficients G = {0.50, 0.70, 0.90, 0.95, 0.99} obtained from Equation (10). The resulting variance measures of the $\hat{\beta}^{OLS}$, $\hat{\beta}^{Ridge}$, $\hat{\beta}^{LASSO}$ and $\hat{\beta}^{EN}$ estimators, obtained from Equation (11), are presented in Tables 1–5, where all values are scaled by 10^3.

2.3. California Housing Dataset and Data Preprocessing

A real data application is conducted to validate the findings of the simulation study. The analysis is based on the California Housing Dataset [27], which comprises 20,440 observations from California census districts and includes demographic and housing-related variables. The median house value within a census block group is considered the dependent variable, while the independent variables used in the analysis are defined as follows:

- X_1: A measure of how far west a house is. Higher values indicate locations farther west.
- X_2: A measure of how far north a house is. Higher values indicate locations farther north.
- X_3: Median age of houses within a block. Lower values indicate newer buildings.
- X_4: Total number of bedrooms within a block.
- X_5: Total number of people residing within a block.
- X_6: Total number of households, a group of people residing within a home unit, for a block.
- X_7: Median income for households within a block of houses.

The categorical variable representing the proximity of a housing block group to the ocean is excluded due to the linear regression framework with continuous independent variables. This dataset is selected because housing-price data commonly exhibit strong correlations among independent variables, making the dataset suitable for evaluating the effects of multicollinearity on regression estimators. To assess the degree of multicollinearity, VIF values are computed for all independent variables using Equation (4). The resulting VIF values are 7.8979, 7.8539, 128.4585, 156.6527, 5.7425, 5.7597, and 1.0696, respectively, indicating the presence of severe multicollinearity for most predictors except the seventh variable. These findings suggest that the OLS estimator may produce unstable parameter estimates with inflated variances under this dataset structure.

Prior to model estimation, the dataset is preprocessed to improve model stability and ensure comparability among the estimators. Missing values are replaced with the median of the corresponding variable. All independent variables are standardized using z-score normalization before applying the ridge, LASSO, and EN estimators. Subsequently, the dataset is randomly divided into training and testing subsets using an 80:20 split ratio. Optimal regularization parameters for ridge, LASSO, and EN estimators are determined separately using 10-fold cross-validation over the range 10^{-4} to 10^4. Accordingly, the optimal values of λ are obtained as 0.0869 for ridge regression, 3237.4575 for LASSO, and 0.1265 for EN. For the EN estimator, the mixing parameter α is fixed at 0.50. Model performance is evaluated using MSPE, root mean squared error (RMSE), mean absolute error (MAE), and coefficient of determination (R^2). These criteria collectively assess predictive accuracy, robustness, and explanatory capability.

3. Results

3.1. Simulation Results

This section presents the results obtained from the simulation study conducted under different combinations of sample sizes, numbers of independent variables, and multicollinearity levels. The performance of the OLS, ridge, LASSO, and EN estimators is evaluated in terms of the sums of variance estimates and predictive accuracy measures.

Tables 1–5 report the sums of the estimated variances of the regression coefficients under different simulation scenarios, with all values reported in units of 10^3 for comparability of estimator stability. Overall, the results indicate that estimator performance is jointly influenced by sample size, the number of explanatory variables, and the degree of multicollinearity. As the sample size increases, the total variance estimates decrease substantially for all methods, indicating improved stability, whereas increasing the number of predictors and the correlation level generally leads to higher variance values, particularly for OLS. When the correlation coefficient is below $G = 0.90$, performance differences among estimators are relatively modest, especially in low-dimensional settings (Tables 1 and 2). However, as multicollinearity strengthens ($G \geq 0.90$, particularly at $G = 0.95$ and $G = 0.99$), the instability of OLS becomes increasingly pronounced, with substantially larger variance values, especially in high-dimensional cases (Tables 3–5). This behavior is consistent with the theoretical sensitivity of OLS to near-linear dependencies among predictors. In contrast, regularized methods (ridge, LASSO, and EN) reduce coefficient variability by shrinking regression parameters, thereby improving stability. Among these, LASSO frequently yields the smallest total variance estimates due to the sparsity induced by the L_1 penalty, while ridge regression and EN also achieve substantial reductions relative to OLS.

Table 1. Sum of the variance estimates ($\sum_{j=1}^{p} \hat{\sigma}^2_{\hat{\beta}_j}$) for $G = 0.50$.

p	$n = 100$				$n = 500$			
	OLS	**Ridge**	**LASSO**	**EN**	**OLS**	**Ridge**	**LASSO**	**EN**
5	0.0485	0.0485	0.0483	0.0484	0.0094	0.0093	0.0093	0.0093
10	0.1196	0.1193	0.1176	0.1185	0.0223	0.0222	0.0215	0.0219
30	0.5107	0.4322	0.1979	0.2892	0.0762	0.0688	0.0168	0.0331
50	1.2200	0.0832	$<10^{-4}$	0.0021	0.1351	0.0186	$<10^{-4}$	0.0006
70	2.9067	0.0004	$<10^{-4}$	$<10^{-4}$	0.1990	0.0001	$<10^{-4}$	$<10^{-4}$
p	$n = 1000$				$n = 2000$			
	OLS	**Ridge**	**LASSO**	**EN**	**OLS**	**Ridge**	**LASSO**	**EN**
5	0.0046	0.0046	0.0045	0.0046	0.0024	0.0024	0.0023	0.0023
10	0.0110	0.0110	0.0105	0.0107	0.0055	0.0055	0.0051	0.0053
30	0.0368	0.0334	0.0050	0.0121	0.0181	0.0165	0.0014	0.0040
50	0.0641	0.0094	$<10^{-4}$	0.0004	0.0312	0.0047	$<10^{-4}$	0.0002
70	0.0924	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$	0.0444	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$

Table 2. Sum of the variance estimates ($\sum_{j=1}^{p} \hat{\sigma}^2_{\hat{\beta}_j}$) for $G = 0.70$.

p	$n = 100$				$n = 500$			
	OLS	**Ridge**	**LASSO**	**EN**	**OLS**	**Ridge**	**LASSO**	**EN**
5	0.0545	0.0544	0.0540	0.0542	0.0105	0.0105	0.0103	0.0104
10	0.1405	0.1398	0.1364	0.1381	0.0258	0.0257	0.0243	0.0250
30	0.6077	0.4525	0.1294	0.2350	0.0905	0.0750	0.0082	0.0217
50	1.4576	0.0541	$<10^{-4}$	0.0035	0.1610	0.0118	$<10^{-4}$	0.0007
70	3.4996	0.0004	$<10^{-4}$	$<10^{-4}$	0.2375	0.0001	$<10^{-4}$	$<10^{-4}$
p	$n = 1000$				$n = 2000$			
	OLS	**Ridge**	**LASSO**	**EN**	**OLS**	**Ridge**	**LASSO**	**EN**
5	0.0052	0.0052	0.0051	0.0051	0.0027	0.0027	0.0026	0.0026
10	0.0127	0.0126	0.0117	0.0121	0.0063	0.0063	0.0056	0.0060
30	0.0437	0.0365	0.0024	0.0068	0.0214	0.0180	0.0007	0.0021
50	0.0762	0.0059	$<10^{-4}$	0.0004	0.0371	0.0030	$<10^{-4}$	0.0002
70	0.1097	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$	0.0528	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$

Table 3. Sum of the variance estimates ($\sum_{j=1}^{p} \hat{\sigma}^2_{\hat{\beta}_j}$) for G = 0.90.

p	n = 100				n = 500			
	OLS	Ridge	LASSO	EN	OLS	Ridge	LASSO	EN
5	0.0637	0.0633	0.0619	0.0626	0.0121	0.0120	0.0114	0.0117
10	0.1685	0.1654	0.1536	0.1594	0.0309	0.0304	0.0257	0.0280
30	0.7408	0.3402	0.0254	0.0692	0.1099	0.0634	0.0025	0.0053
50	1.7694	0.0181	0.0016	0.0029	0.1949	0.0038	0.0003	0.0006
70	4.2257	0.0003	$<10^{-4}$	$<10^{-4}$	0.2871	0.0001	$<10^{-4}$	$<10^{-4}$
p	n = 1000				n = 2000			
	OLS	Ridge	LASSO	EN	OLS	Ridge	LASSO	EN
5	0.0060	0.0060	0.0055	0.0058	0.0030	0.0030	0.0027	0.0029
10	0.0153	0.0151	0.0118	0.0133	0.0076	0.0074	0.0052	0.0062
30	0.0528	0.0311	0.0010	0.0022	0.0261	0.0155	0.0005	0.0011
50	0.0923	0.0019	0.0001	0.0003	0.0448	0.0010	$<10^{-4}$	0.0001
70	0.1331	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$	0.0642	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$

Table 4. Sum of the variance estimates ($\sum_{j=1}^{p} \hat{\sigma}^2_{\hat{\beta}_j}$) for G = 0.95.

p	n = 100				n = 500			
	OLS	Ridge	LASSO	EN	OLS	Ridge	LASSO	EN
5	0.0654	0.0644	0.0616	0.0630	0.0127	0.0125	0.0112	0.0118
10	0.1779	0.1713	0.1479	0.1592	0.0322	0.0312	0.0221	0.0263
30	0.7779	0.2228	0.0133	0.0258	0.1147	0.0443	0.0021	0.0041
50	1.8631	0.0089	0.0030	0.0022	0.2051	0.0018	0.0009	0.0004
70	4.4477	0.0002	$<10^{-4}$	$<10^{-4}$	0.3028	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$
p	n = 1000				n = 2000			
	OLS	Ridge	LASSO	EN	OLS	Ridge	LASSO	EN
5	0.0063	0.0062	0.0053	0.0057	0.0031	0.0031	0.0024	0.0027
10	0.0159	0.0153	0.0094	0.0120	0.0079	0.0076	0.0038	0.0054
30	0.0556	0.0221	0.0010	0.0022	0.0273	0.0111	0.0004	0.0011
50	0.0970	0.0009	0.0005	0.0002	0.0471	0.0005	0.0003	0.0001
70	0.1394	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$	0.0673	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$

Table 5. Sum of the variance estimates ($\sum_{j=1}^{p} \hat{\sigma}^2_{\hat{\beta}_j}$) for G = 0.99.

p	n = 100				n = 500			
	OLS	Ridge	LASSO	EN	OLS	Ridge	LASSO	EN
5	0.0688	0.0630	0.0502	0.0563	0.0132	0.0117	0.0068	0.0089
10	0.1842	0.1526	0.0736	0.1063	0.0331	0.0276	0.0058	0.0121
30	0.8097	0.0522	0.0107	0.0120	0.1189	0.0108	0.0019	0.0028
50	1.9571	0.0018	0.0040	0.0010	0.2137	0.0004	0.0010	0.0002
70	4.6381	0.0001	$<10^{-4}$	$<10^{-4}$	0.3153	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$
p	n = 1000				n = 2000			
	OLS	Ridge	LASSO	EN	OLS	Ridge	LASSO	EN
5	0.0065	0.0056	0.0026	0.0038	0.0032	0.0026	0.0010	0.0015
10	0.0164	0.0134	0.0019	0.0043	0.0082	0.0065	0.0007	0.0015
30	0.0577	0.0055	0.0009	0.0015	0.0285	0.0028	0.0004	0.0008
50	0.1008	0.0002	0.0005	0.0001	0.0488	0.0001	0.0003	0.0001
70	0.1458	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$	0.0702	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$

Among the regularized estimators, EN exhibits the most stable overall behavior under severe multicollinearity and high-dimensional settings. The combined L_1 and L_2 penalties allow EN to balance coefficient shrinkage and variable selection, leading to improved stability and predictive performance. Ridge regression also provides considerable variance reduction compared with OLS; however, because it retains all predictors in the model, its predictive performance can be relatively weaker than EN in highly correlated scenarios. It should also be noted that some variance values reported for LASSO and EN may appear extremely small or rounded to zero; however, these values do not indicate exact zero variance but rather reflect the strong shrinkage effect induced by regularization under certain multicollinearity conditions.

Table 6 reports the estimated regression coefficients for a representative simulation setting with $G = 0.99$, $n = 100$, and $p = 10$. In this particular realization, the OLS estimator produces coefficient estimates that are relatively close to the true parameter values and yields the smallest total bias, $\sum_{j=1}^{p}|\hat{\beta}_j - \beta_j|$. The penalized estimators, on the other hand, produce coefficient estimates that are generally closer to zero, resulting in larger deviations from several true coefficients and consequently higher total bias values. Among the penalized methods, LASSO achieves the smallest total bias, whereas ridge regression produces the largest. Since the results correspond to a single simulation run, they should be interpreted as an illustrative example of coefficient estimation under extreme multicollinearity rather than as definitive evidence of the overall superiority of any estimator.

Table 6. Estimated coefficients and total bias values.

True β	$\hat{\beta}^{OLS}$	$\hat{\beta}^{Ridge}$	$\hat{\beta}^{LASSO}$	$\hat{\beta}^{EN}$
1	0.9709	0.9556	0.9326	0.9521
1	1.0144	0.9584	1.0322	0.9729
1	1.0161	0.9357	1.0409	0.9619
1	1.0179	0.8854	0.9748	0.8942
1	0.9696	0.7123	0.7046	0.7028
0	−0.0298	0.2593	0.1263	0.2176
0	0.0427	0.1491	0.0859	0.1292
0	0.0045	0.0742	0.0472	0.0712
0	−0.0069	0.0497	0.0367	0.0644
Total bias	0.0728	0.4486	0.3493	0.4210

Figures 1 and 2 provide graphical illustrations of the variance behavior of the estimators under moderate and severe multicollinearity levels, respectively. The variance of the OLS estimator is highly sensitive to both the correlation level and the value of p. In particular, for moderate values of p, OLS exhibits pronounced spikes in variance, which become even more severe under high correlation ($G = 0.99$). In contrast, LASSO and EN estimators exhibit values that are very close to each other across almost all p values, indicating a much more stable pattern. As p increases, and especially as the sample size grows, the variance of all estimators tends to decrease, and the differences between them become less pronounced. To enhance visual clarity and facilitate comparison among the estimators, the vertical axis is limited to the range [0, 0.6]. Therefore, a small number of variance values exceeding 0.6 fall outside the displayed range.

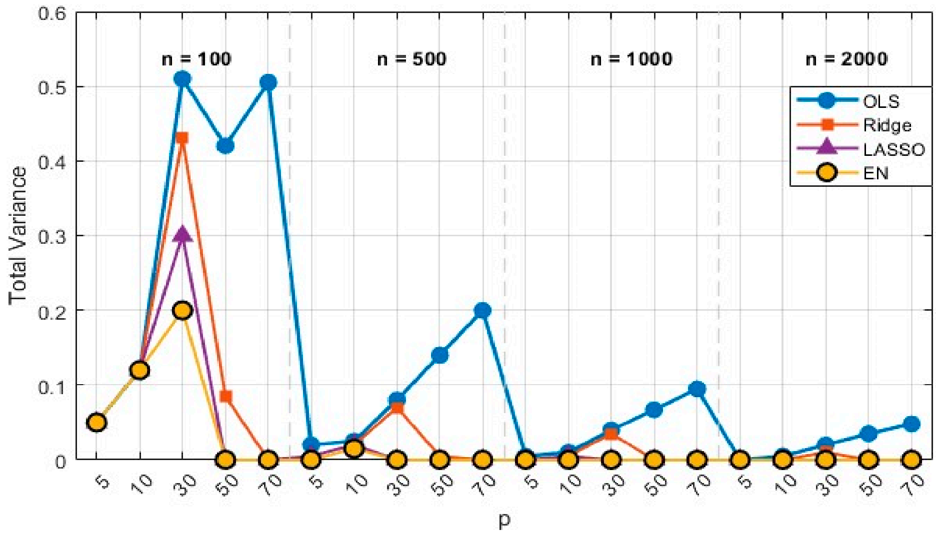

Figure 1. Variance of estimators for $G = 0.50$ across different n and p values.

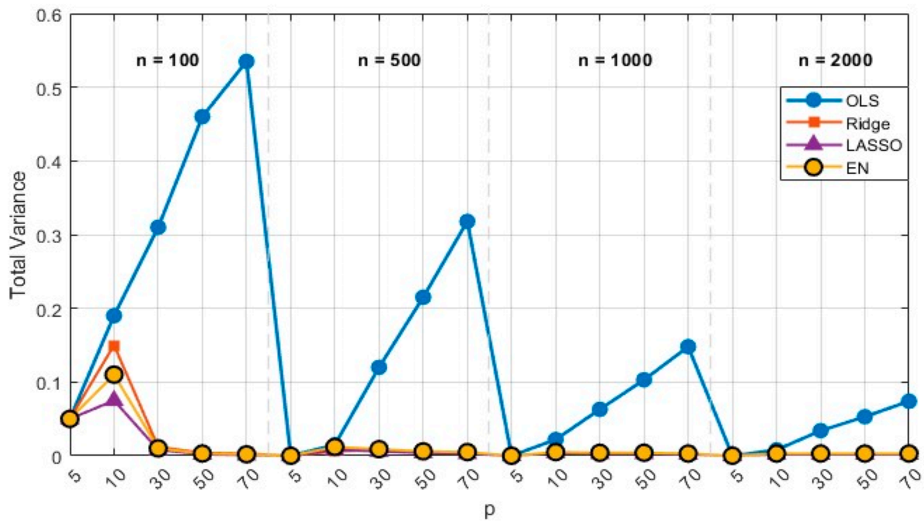

Figure 2. Variance of estimators for $G = 0.99$ across different n and p values.

3.2. Real Data Application

The predictive performances of the OLS, ridge, LASSO, and EN estimators are evaluated using MSPE, RMSE, MAE, and R^2 measures. MSPE is computed using the entire dataset, whereas the other performance metrics are evaluated on the test set. The number of observations in the test dataset is denoted by m. RMSE is defined as $\mathrm{RMSE} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2}$ and measures the average magnitude of prediction errors while assigning greater weight to large deviations. MAE is calculated as $\mathrm{MAE} = \frac{1}{m}\sum_{i=1}^{m}|y_i - \hat{y}_i|$ and represents the average absolute difference between observed and predicted values. The coefficient of determination is given by $R^2 = 1 - \frac{\sum_{i=1}^{m}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{m}(y_i - \overline{y})^2}$ where $\overline{y}$ denotes the mean of the observed values, and it indicates the proportion of variability in the dependent variable explained by the model. Lower RMSE and MAE values and higher R^2 values indicate better predictive performance.

Table 7 presents the coefficient estimates obtained from the California Housing Dataset. While the estimated coefficients generally retain the same direction of effect across the estimators, their magnitudes differ considerably, reflecting the varying degrees of regularization imposed by ridge regression, LASSO, and EN.

Table 7. Parameter estimates for the California Housing Dataset.

Parameter Estimates	$\hat{\beta}^{OLS}$ $(\times 10^6)$	$\hat{\beta}^{Ridge}$ $(\times 10^3)$	$\hat{\beta}^{LASSO}$ $(\times 10^4)$	$\hat{\beta}^{EN}$ $(\times 10^3)$
$\hat{\beta}_0$	4.3317	0.0000	0.0000	0.0000
$\hat{\beta}_1$	−0.0527	−0.9008	−5.2670	−1.8833
$\hat{\beta}_2$	−0.0528	−1.8109	−5.2783	−3.1812
$\hat{\beta}_3$	0.0210	0.2002	0.0697	0.3157
$\hat{\beta}_4$	−0.0042	0.2002	0.0000	0.3157
$\hat{\beta}_5$	−0.0000	−0.0007	−0.0046	−0.0015
$\hat{\beta}_6$	0.0001	0.0055	0.0140	0.0103
$\hat{\beta}_7$	0.0000	0.0007	0.0003	0.0011

As shown in Table 8, the LASSO estimator exhibits superior predictive performance according to the test-set evaluation metrics. In particular, it produces the lowest RMSE and MAE values, while also yielding the highest R^2 value. In contrast, OLS shows the weakest performance. These findings suggest that regularization-based methods, particularly LASSO, appear to be more effective in improving predictive accuracy and stability than the classical OLS estimator. Furthermore, although the EN estimator yields the lowest MSPE value, this does not substantially alter the overall comparative assessment, as MSPE reflects performance over the entire dataset, whereas RMSE, MAE, and R^2 are computed on the test set. This difference accounts for the variation in ranking across evaluation criteria.

Table 8. Performance metrics for the California house dataset.

Estimators	MSPE	RMSE	MAE	R^2
OLS	1.0728×10^{14}	35,147.59	29,644.24	0.7898
Ridge	1.1376×10^{13}	33,136.49	28,320.83	0.8132
LASSO	7.7350×10^{15}	32,976.55	28,008.11	0.8150
EN	4.8240×10^{12}	33,363.16	28,501.92	0.8106

To complement the tabular analysis, Figure 3 presents a graphical illustration of the results, highlighting differences among the estimators in terms of coefficient behavior and predictive accuracy. Figure 3a compares the coefficient estimates obtained from the OLS, ridge, LASSO, and EN estimators, excluding the intercept term for improved visualization. Although the estimated coefficients generally preserve the same direction of effect across the estimators, noticeable differences are observed in their magnitudes due to the varying degrees of regularization. In particular, the coefficients associated with $\hat{\beta}_1$ and $\hat{\beta}_2$ exhibit the largest absolute values across all methods, indicating that these predictors have the strongest influence on housing prices. Ridge regression shrinks all coefficients toward zero while retaining every predictor in the model. LASSO applies stronger shrinkage and sets the coefficient $\hat{\beta}_4$ exactly to zero, thereby performing variable selection. EN produces coefficient estimates that generally lie between those of ridge regression and LASSO, reflecting its hybrid regularization structure. Overall, the figure illustrates that regularization reduces coefficient variability while preserving the main patterns of association between the independent variables and the dependent variable. Figure 3b, on the other hand, compares the estimators in terms of MSPE values, providing an evaluation of their predictive performance. A logarithmic scale is used to appropriately accommodate the large differences in magnitude among the MSPE values across the considered estimators.

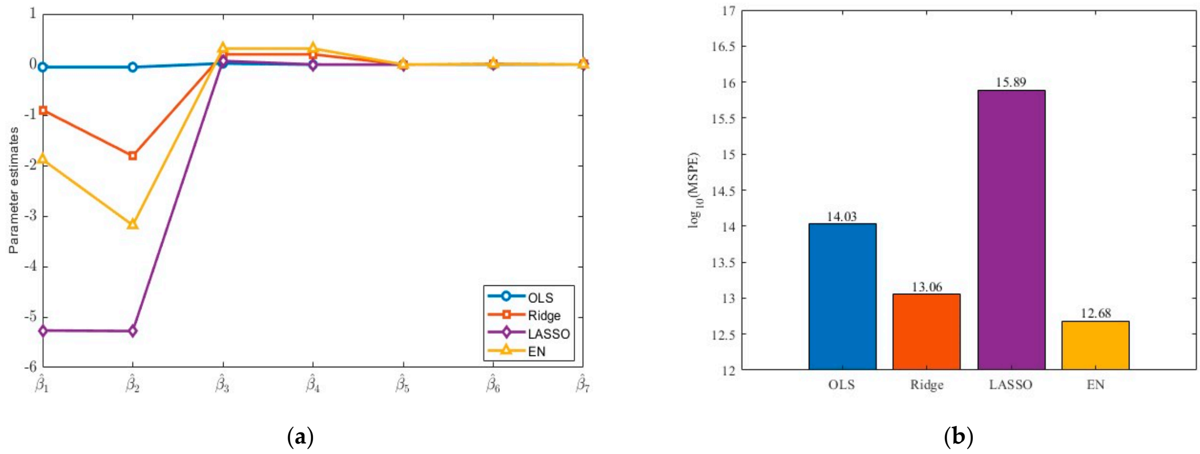

Figure 3. Comparative results on the California Housing Dataset: (**a**) Estimated regression coefficients. (**b**) Log MSPEs across regression estimators.

The results indicate that the EN estimator achieves the lowest prediction error among the considered estimators. Ridge regression also demonstrates strong, stable performance with low MSPE values, whereas OLS yields substantially higher errors. Although LASSO effectively performs variable selection, it suffers from severe underfitting in this setting, leading to weaker predictive accuracy. In contrast, by retaining more information through coefficient shrinkage, ridge regression and EN achieve superior results. Specifically, EN provides a well-balanced trade-off between model flexibility and regularization, proving its effectiveness for this dataset.

Figure 4 compares the observed house values with the predicted values obtained from the regression estimators. In each panel, the horizontal axis represents the observed values (y) and the vertical axis represents the predicted values ($\hat{y}$). Compared with OLS, the penalized regression estimators produce more stable prediction patterns, while the OLS estimator exhibits greater dispersion, suggesting a stronger influence of multicollinearity on prediction accuracy.

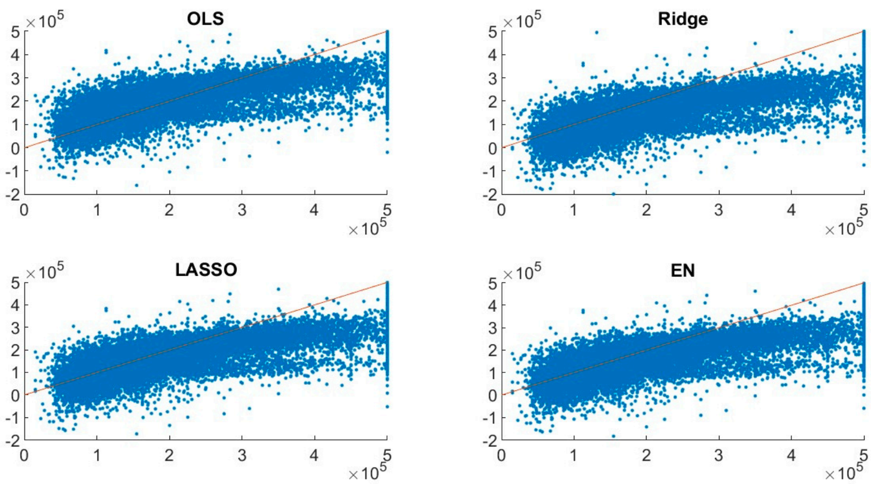

Figure 4. Comparison between observed and predicted house values for the competing regression estimators.

4. Discussion

In empirical housing data, independent variables are often highly correlated, leading to multicollinearity problems that adversely affect the stability and predictive reliability of conventional regression models. The findings of this study demonstrate that regularization-based estimators provide substantial advantages over OLS under such conditions. Both the simulation study and the California Housing Dataset analysis indicate that ridge regression, LASSO, and EN achieve more stable parameter estimates and improved predictive performance when strong correlations exist among independent variables. Furthermore, the simulation results show that estimator performance depends on both the severity of multicollinearity and the dimensionality of the dataset. While the differences among the estimators remain relatively small under moderate correlation levels, more pronounced distinctions emerge as the correlation structure becomes stronger and the number of independent variables increases.

Among the competing regularization methods, LASSO performs particularly well in several simulation scenarios due to its ability to reduce variance through sparse coefficient estimation. By shrinking some coefficients exactly to zero, LASSO simplifies the model structure and performs automatic variable selection. However, the empirical results suggest that aggressive shrinkage may remove informative predictors in highly correlated datasets, leading to reduced predictive accuracy. Ridge regression, on the other hand, stabilizes coefficient estimates by shrinking all regression coefficients toward zero without excluding variables from the model. This characteristic improves estimation stability while preserving the contribution of all independent variables. Moreover, the simulation design assumes a sparse coefficient structure, where only a subset of the independent variables is associated with nonzero regression coefficients. Under such settings, methods capable of variable selection may have an advantage, which helps explain the strong performance of LASSO observed across many scenarios.

The empirical findings further indicate that EN provides the strongest overall predictive performance. By combining the L_1 and L_2 penalties, EN balances coefficient shrinkage and variable selection, thereby achieving improved prediction accuracy while maintaining model stability.

The graphical analyses support the numerical findings reported in the Results section. The coefficient comparison plots highlight the differences in coefficient magnitudes across the competing estimators and illustrate the shrinkage effect induced by regularization. Furthermore, the observed-versus-predicted plots provide a visual assessment of predictive performance across the competing estimators and complement the quantitative evaluation based on MSPE.

The findings obtained in this study are generally consistent with previous research emphasizing the benefits of regularization methods in the presence of multicollinearity. Earlier studies reported that ridge regression improves estimation stability by reducing variance [18], whereas LASSO performs effective variable selection through coefficient shrinkage [19]. More recent studies have shown that EN is particularly effective when groups of highly correlated independent variables are present [20]. The results from both the simulation study and the California Housing Dataset are consistent with these observations, suggesting that EN may exhibit a balanced performance between predictive accuracy and model stability.

Overall, the simulation and empirical analyses suggest that no single estimator is universally optimal under all conditions. Rather, the choice of estimator should depend on the degree of multicollinearity, the dimensionality of the dataset, and the performance criterion of primary interest. In the simulation study, LASSO generally exhibits lower total variance across many scenarios, whereas EN becomes more competitive under stronger multicollinearity. In the California Housing application, LASSO performs better according to the test-set evaluation metrics (RMSE, MAE, and R^2), while EN achieves the lowest overall MSPE. These differences arise from the fact that RMSE, MAE, and R^2 are computed

https://doi.org/10.3390/mca31030111

on the test set, whereas MSPE is based on the full dataset and therefore reflects a different aspect of generalization performance.

5. Conclusions

This study investigates the predictive performance of OLS, ridge, LASSO, and EN estimators in the presence of multicollinearity using a comprehensive simulation framework and the California Housing Dataset as a real-world application. The simulation design systematically varies sample size, dimensionality, and correlation structure to assess estimator behavior under controlled and increasingly complex conditions, while the empirical analysis provides evidence from a practical high-dimensional regression setting.

The results consistently demonstrate that multicollinearity undermines the stability and predictive reliability of the OLS estimator, with effects becoming more pronounced in high-dimensional and highly correlated environments. In contrast, regularization-based methods substantially mitigate these issues by imposing coefficient shrinkage, thereby improving both estimation stability and out-of-sample predictive performance.

Comparative evidence further reveals that estimator performance is inherently data-dependent and driven by the interaction between dimensionality and correlation strength. In the empirical application, LASSO achieves the most favorable test-set performance according to RMSE, MAE, and R^2, whereas EN yields the lowest overall MSPE. These findings indicate that the relative performance of the estimators depends on the evaluation criterion considered. The results further suggest that the relative advantages of LASSO and EN vary according to the correlation structure and the performance measure of interest.

Taken together, the simulation and empirical results provide strong empirical support for the use of regularization techniques in regression settings characterized by multicollinearity. Overall, the results demonstrate that regularization-based estimators offer effective alternatives to OLS, with the choice of estimator depending on the characteristics of the data and the performance criterion of primary interest.

Although the findings provide useful insights into the behavior of OLS, ridge, LASSO, and EN estimators under multicollinearity, several limitations should be acknowledged. First, the empirical analysis is based solely on the California Housing Dataset. Although this dataset is widely used in regression studies, the results may not fully generalize to datasets with different characteristics, distributions, or correlation structures. Second, the study focuses on classical and regularization-based linear regression methods and does not include comparisons with modern machine learning approaches such as Random Forest, XGBoost, LightGBM, or Neural Networks. Therefore, the conclusions should be interpreted within the scope of linear predictive modeling. Third, the simulation results depend on the selected experimental settings, including the sample sizes, numbers of independent variables, and correlation levels considered in the study. Different simulation designs may lead to different relative performances of the estimators. In particular, the use of a sparse coefficient structure in the simulation design may influence the relative performance of shrinkage-based estimators, and alternative dense settings may yield different comparative results, especially for ridge regression. Finally, although multiple evaluation metrics are employed, additional performance measures, alternative validation strategies, and formal statistical inference could provide further insights into model behavior. Future research may extend the analysis to other real-world datasets, alternative correlation structures, and more advanced predictive modeling techniques, as well as the selection of the EN mixing parameter through cross-validation or grid search procedures.

Author Contributions: Both authors contributed equally to all stages of the study, including conceptualization, data analysis, and manuscript writing. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in this study are included in the article. The datasets analyzed in this study are available in Kaggle at https://www.kaggle.com/datasets/dhirajnirne/california-housing-data (accessed on 12 May 2026).

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

OLS	Ordinary least squares
LASSO	Least absolute shrinkage and selection operator
EN	Elastic net
RSS	Residual sum of squares
VIF	Variance inflation factor
MSPE	Mean square prediction error
RMSE	Root mean squared error
MAE	Mean absolute error

References

1. Ahmad, M.H.; Adnan, R.; Adnan, N. A comparative study on some methods for handling multicollinearity problems. *Matematika* **2006**, *22*, 109–119.
2. Kroll, C.N.; Song, P. Impact of multicollinearity on small sample hydrologic regression models. *Water Resour. Res.* **2013**, *49*, 3756–3769. [CrossRef]
3. Stewart, G.W. Collinearity and least squares regression. *Stat. Sci.* **1987**, *2*, 68–84. [CrossRef]
4. Sharma, H.; Harsora, H.; Ogunleye, B. An optimal house price prediction algorithm: XGBoost. *Analytics* **2024**, *3*, 30–45. [CrossRef]
5. Mullainathan, S.; Spiess, J. Machine learning: An applied econometric approach. *J. Econ. Perspect.* **2017**, *31*, 87–106. [CrossRef]
6. Mora-Garcia, R.-T.; Cespedes-Lopez, M.-F.; Perez-Sanchez, V.R. Housing price prediction using machine learning algorithms in COVID-19 times. *Land* **2022**, *11*, 2100. [CrossRef]
7. Abdelwahab, M.M.; Abonazel, M.R.; Hammad, A.T.; El-Masry, A.M. Modified two-parameter Liu estimator for addressing multicollinearity in the poisson regression model. *Axioms* **2024**, *13*, 46. [CrossRef]
8. Nayem, H.M.; Aziz, S.; Kibria, B.M.G. Evaluating estimator performance under multicollinearity: A trade-off between MSE and accuracy in logistic, lasso, elastic net, and ridge regression with varying penalty parameters. *Stats* **2025**, *8*, 45. [CrossRef]
9. Alharthi, M.F.; Akhtar, N. Modified two-parameter ridge estimators for enhanced regression performance in the presence of multicollinearity: Simulations and medical data applications. *Axioms* **2025**, *14*, 527. [CrossRef]
10. An, J.; Park, J.; Yeon, S.; Oh, C.; Lee, B.; Jung, W.; Yun, J.; Oh, H.J. Machine learning-based prediction of Coagulant dosing in drinking water treatment plants using polynomial regression with lasso regularization. *Processes* **2025**, *13*, 3829. [CrossRef]
11. Xin, S.J.; Khalid, K. Modelling house price using ridge regression and lasso regression. *Int. J. Eng. Technol.* **2018**, *7*, 498–501. [CrossRef]
12. Nayem, H.M.; Aziz, S.; Kibria, B.M.G. Comparison among ordinary least squares, ridge, lasso, and elastic net estimators in the presence of outliers: Simulation and application. *Int. J. Stat. Sci.* **2024**, *24*, 25–48. [CrossRef]
13. Gabauer, D.; Gupta, R.; Marfatia, H.A.; Miller, S.M. Estimating US housing price network connectedness: Evidence from dynamic Elastic Net, Lasso, and ridge vector autoregressive models. *Int. Rev. Econ. Financ.* **2024**, *89*, 349–362. [CrossRef]
14. Preethi; Murthy, D.H.R.; Hiremani, V.; Devadas, R.M.; Sapna, R. Optimizing polynomial and regularization techniques for enhanced housing price prediction accuracy. *SN Comput. Sci.* **2025**, *6*, 96. [CrossRef]
15. Zeng, Z. Using Linear Regression, Ridge Regression, Lasso Regression, and Elastic Net Regression for Predicting Real Estate Price. In *Proceedings of the 2nd International Conference on Data Science and Engineering (ICDSE)*; SCITEPRESS: Setúbal Municipality, Portugal, 2025. [CrossRef]
16. Elkhidir, E.; Patel, T.; Rotimi, J.O.B. Predictive modelling for residential construction demands using ElasticNet Regression. *Buildings* **2025**, *15*, 1649. [CrossRef]
17. Xia, H. Lasso's application in Chinese housing prices. *Commun. Stat. Theory Methods* **2026**, *55*, 2465–2476. [CrossRef]
18. Hoerl, A.E.; Kennard, R.W. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* **1970**, *12*, 55–67. [CrossRef]
19. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B* **1996**, *58*, 267–288. [CrossRef]

20. Zou, H.; Hastie, T. Regularization and Variable Selection via the Elastic Net. *J. R. Stat. Soc. Ser. B* **2005**, *67*, 301–320. [CrossRef]
21. Dormann, C.F.; Elith, J.; Bacher, S.; Buchmann, C.; Carl, G.; Carré, G.; Marquéz, J.R.G.; Gruber, B.; Lafourcade, B.; Leitão, P.J.; et al. Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography* **2013**, *36*, 27–46. [CrossRef]
22. Hebiri, M.; Lederer, J. How correlations influence lasso prediction. *IEEE Trans. Inf. Theory* **2012**, *59*, 1846–1854. [CrossRef]
23. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning*, 2nd ed.; Springer: New York, NY, USA, 2009.
24. Fu, W.J. Penalized regressions: The bridge versus the lasso. *J. Comput. Graph. Stat.* **1998**, *7*, 397–416. [CrossRef]
25. Muthukrishnan, R.; Rohini, R. LASSO: A Feature Selection Technique in Predictive Modeling for Machine Learning. In Proceedings of the International Conference on Advances in Computer Applications (ICACA), Coimbatore, India, 24 October 2016. [CrossRef]
26. Hans, C. Elastic net regression modeling with the orthant normal prior. *J. Am. Stat. Assoc.* **2011**, *106*, 1383–1393. [CrossRef]
27. Kaggle. Available online: https://www.kaggle.com/datasets/dhirajnirne/california-housing-data (accessed on 5 February 2026).
28. Bun, M.J.G.; Harrison, T.D. OLS and IV estimation of regression models including endogenous interaction terms. *Econom. Rev.* **2019**, *38*, 814–827. [CrossRef]
29. Yang, H.; Chang, X. A new two-parameter estimator in linear regression. *Commun. Stat. Theory Methods* **2010**, *39*, 923–934. [CrossRef]
30. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning: With Applications in R*, 2nd ed.; Springer: New York, NY, USA, 2021. [CrossRef] [PubMed]
31. Algamal, Z.Y. Shrinkage parameter selection via modified cross-validation approach for ridge regression model. *Commun. Stat. Simul. Comput.* **2020**, *49*, 1922–1930. [CrossRef]
32. Nakatsu, R.T. Validation of machine learning ridge regression models using Monte Carlo, bootstrap, and variations in cross-validation. *J. Intell. Syst.* **2023**, *32*, 20220224. [CrossRef]
33. Hoerl, R.W. Ridge regression: A historical context. *Technometrics* **2020**, *62*, 420–425. [CrossRef]
34. Lukman, A.F.; Ali, E. Contraction ridge estimator: Simulation and application to economic data. *Comp. Econ.* **2025**, 1–25. [CrossRef]
35. Yi, C.; Huang, J. Semismooth newton coordinate descent algorithm for elastic-net penalized huber loss regression and quantile regression. *J. Comput. Graph. Stat.* **2017**, *26*, 547–557. [CrossRef]
36. Wright, S.J. Coordinate descent algorithms. *Math. Program.* **2015**, *151*, 3–34. [CrossRef]
37. Ranstam, J.; Cook, J.A. LASSO regression. *Br. J. Surg.* **2018**, *105*, 1348. [CrossRef]
38. Hastie, T.; Tibshirani, R.; Wainwright, M. *Statistical Learning with Sparsity: The Lasso and Generalizations*, 1st ed.; Chapman and Hall/CRC: Boca Raton, FL, USA, 2015. [CrossRef]
39. Xu, Q.F.; Ding, X.H.; Jiang, C.X.; Yu, K.M.; Shi, L. An elastic-net penalized expectile regression with applications. *J. Appl. Stat.* **2021**, *48*, 2205–2230. [CrossRef] [PubMed]
40. Raymaekers, J.; Rousseeuw, P.J. The cellwise minimum covariance determinant estimator. *J. Am. Stat. Assoc.* **2024**, *119*, 2610–2621. [CrossRef]
41. Steland, A. Shrinkage for covariance estimation: Asymptotics, confidence intervals, bounds and applications in sensor monitoring and finance. *Stat. Pap.* **2018**, *59*, 1441–1462. [CrossRef]