Open Access Library Journal
2024, Volume 11, e12078
ISSN Online: 2333-9721
ISSN Print: 2333-9705

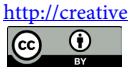
Housing Market Analysis: Supply-Demand Dynamics a Non-Parametric Approach

Protais Lekelem Dongmo

Department of Mathematics, University of Yaounde 1, Yaounde, Cameroon
Email: dongmo20gilles@gmail.com

How to cite this paper: Dongmo, P.L. (2024) Housing Market Analysis: Supply-Demand Dynamics a Non-Parametric Approach. *Open Access Library Journal*, **11**: e12078.
https://doi.org/10.4236/oalib.1112078

Received: August 7, 2024
Accepted: September 21, 2024
Published: September 24, 2024



Abstract

Homeownership is part of the "American Dream" and a key tool for households to build wealth. However, increasing house prices have made homeownership less attainable recently. In order to determine the factors affecting home supply, we used non-parametric approaches. Some of them include the Nadaraya-Watson estimator, Local Polynomial estimator, B-spline, P-spline, and the Adapted He approach. The latter is particularly useful when dealing with outliers or non-standard conditional distributions in the data. However, when the functions of the covariates are not easily specified in a parametric manner, a nonparametric regression technique is often employed. One such technique is using B-splines, a nonparametric approach used to estimate the parameters of the unspecified functions in the model. According to the root mean square error, B-splines were identified as the appropriate model. But this model is very liable to overfitting when the number of knots is increased and also becomes less efficient in the presence of outliers, thus the need for a more robust nonparametric regression model to overcome this, the Adapted-He approach in time-varying coefficient model was applied. The estimation procedure involves minimizing the quantile loss function using an LP-Problem technique. These methods were all applied to the US housing data. The study results indicated that interest rates and consumer sentiment had positive and negative effects on the monthly supply when using all of the above-mentioned models. This means that an increase in consumer sentiment will cause an increase in demand, which, in return, will cause a decrease in supply. Whereas an increase in the interest rates of the houses will cause an increase in supply.

Subject Areas

Mathematical Economics, Mathematical Statistics

Keywords

Homeownership, Quantile Regression, Adapted-He Approach, Local Polynomial

Estimator, B-Spline, S-Spline, P-Spline, Nadaraya-Watson Estimator

1. Introduction

Homeownership is part of the "American Dream" and a vital tool for households to build wealth. However, fewer Americans have purchased homes in the years after the recession, and rising home prices have made homeownership less accessible. It is uncertain what this will signify in the long run for the economy [1]. The percentage of households that own a home has historically been positively connected with the cost of purchasing a property. According to [1], from 1996 to 2006 in the United States, both the price of houses and the homeownership rate increased. This increasing trend ended abruptly with the global financial crisis, which saw house prices plunge and drove homeownership rates to historically low levels. If it became less attractive in the wake of the financial crisis, we might expect both prices and homeownership to decrease. Similarly, if the current increase in house prices were driven by people buying homes to live in, we might expect the homeownership rate to increase along with prices. However, recent evidence shows that house prices and homeownership are diverging. Therefore, several factors could be driving the decoupling of house prices and the homeownership rate.

It was then necessary to question the factors influencing house ownership. Our main objective was determining the factors influencing US house ownership. To achieve this, we, specifically in the first place, decided to visualize the density plots of some variables, assuming an unknown distribution for each of these variables. Next, we use non-parametric regression to establish a relationship between home supply and each of the explanatory variables.

This project shall be divided into four main parts. We shall start with a description of the dataset, and next shall be a brief discussion on the methods we will apply to achieve our goal. In the third section, results shall be displayed and discussed. We will then end our report with a conclusion.

2. US Housing Data

This data set is based on the US housing data taken from https://www.kaggle.com/datasets/utkarshx27/factors-influence-house-price-in-us. It is categorized into two sets of data: demand and supply datasets. These datasets contain supply-demand factors that influence US home prices.

3. Methodology

This section describes the methods used to determine the possible factors affecting house ownership.

3.1. Non-Parametric Density

Consider the problem of estimating the density function $f(x)$ of a scalar,

continuously-distributed i.i.d. Sequence x_i at a particular point x: If the density f is in a known parametric family (e.g., Gaussian), estimation of the density reduces to an estimation of the finite-dimensional parameters that characterize that particular density in the parametric family. Without a parametric assumption, though, estimating the density f overall points in its support would involve estimating an infinite number of parameters, known in statistics as a non-parametric estimation problem. There are various methods of non-parametric density estimation (Kernel density estimation, Nearest neighbor estimation, and the histogram).

3.1.1. Histogram

It is the simplest method to estimate the density $f = f_x(x)$ from an i.i.d sample $X_1, \cdots, X_n$. The estimation function formula is given in two cases: the discrete case and the continuous case.

- The discrete case: The function is given by:

$$\hat{f}_x(x) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}(X_i = x) \tag{1}$$

where $\mathbb{I}(X_i = x)$ is an indicator function.

- The continuous case: The function $f_x(x)$ is estimated by averaging X_i in a specific interval $x \pm \frac{h}{2}$, where h is the interval length. Thus, the function is given by

$$\hat{f}_x(x) = \frac{1}{nh}\sum_{i=1}^{n}\mathbb{I}\left(-\frac{1}{2} \le \psi_i \le \frac{1}{2}\right) \tag{2}$$

where $\psi_i = \frac{X_i - x}{h}$.

However, using the histogram to estimate the density has some limitations, among which: the choice of bins causing the discontinuities of the estimate, the estimation of midpoints of some specific intervals, the estimated function being rougher, and lastly, the function $f(x)$ is not continued, it jumps at points $x_i \pm h/2$ with its derivative equals to zero. Therefore, to overcome this, ([2]) replaces the indicator function with a kernel function which is continuous and differentiable.

3.1.2. Kernel Density Estimation

As said in the previous section, [2] proposes another way of estimating the probability density function of a random variable by replacing the indicator function with a kernel function that satisfies the following relation:

$$\int_{-\infty}^{\infty} K(u)\mathrm{d}u = 1 \tag{3}$$

Therefore, according to ([2]), Equation (2) becomes:

$$\hat{f}_x(x) = \frac{1}{nh}\sum_{i=1}^{n}K\left(\frac{X_i - x}{h}\right) \tag{4}$$

Where h is the bandwidth or smoothing parameter, moreover, several kernel functions exist in the literature. Still, according to [3], selecting the smoothing parameter (bandwidth) is much more important than choosing the kernel function. In other words, the kernel function is less critical when using an optimal bandwidth. Hence in this project, we shall use one of the optimal bandwidth criteria (Scott bandwidth, Silverman's rule of thumb bandwidth, Silverman's Long-tailed distribution, Unbaised cross-validation bandwidth) depending on the structure of our data. However, the two previous techniques are based on using fixed bandwidths which tend to be under-smooth in some parts of the function and over-smooth in some of the density functions. To overcome this, Fix and Hodges proposed another method of estimation called the K-nearest-neighborhood (K-NN).

3.1.3. K-Nearest Neighbors

This estimation method proposes a bandwidth h which involves the distance x to the data ($d(x_i, x) = |x - x_i|$. Looking at the I-nearest distance denoted by $d_k(x)$, $k = \sqrt{n}$, then the estimator obtained is called the Kth nearest-neighborhood (K-NN). The estimated function is given by:

$$\hat{f}(x) = \frac{1}{2nd_k(x)} \sum_{i=1}^{n} K\left(\frac{X_i - x}{2d_k(x)}\right) \tag{5}$$

Having an idea of the density function of the variables, we can then look at the relationship between them by using the nonparametric regression techniques presented in the next section.

3.2. Non-Parametric Regression Techniques

The primary goal in nonparametric regression is to construct an estimate $\hat{f}$ of f, from i.i.d. samples $(x_1, y_1), \cdots, (x_n, y_n) \in \mathbb{R}^d \times \mathbb{R}$ that have the same joint distribution as (X, Y). We often call X the input, predictor, feature, etc., and Y the output, outcome, response, etc. Importantly, we do not assume a specific parametric form for f in nonparametric regression. In other words, in nonparametric regression, given the function $f(x) = m(x) + \epsilon$, we assume that $m(x)$ is not known. Thus, we try to estimate it by a more smooth and flexible $\hat{m}(x)$ through nonparametric methods. Many techniques have been developed in the literature, but we shall briefly discuss the strategies we used in this project.

3.2.1. Nadaraya-Watson Estimator

Nadaraya and Watson proposed and estimator $\hat{m}(.)$ of $m(.)$, which is given by:

$$\hat{m}(x) = \frac{\sum_{i=1}^{n} K_h(X_i - x) Y_i}{\sum_{i=1}^{n} K_h(X_i - x)} = \sum_{i=1}^{n} W_i(x) Y_i \tag{6}$$

The bandwidth h determines the degree of smoothing. An extensive h increases the width of the bins, increasing the smoothness of $\hat{m}(x)$. A small h decreases the width of the containers, producing a less smooth $\hat{m}(x)$ [4]. The Nadaraya-Watson estimator can be seen as a weighted average of $Y_1, \cdots, Y_n$ using the set of

weights $W_i(x)_{i=1}^n$ (they add to one). The set of varying weights depends on the evaluation point *x*. The Nadaraya-Watson estimator is a local mean of $Y_1,\cdots,Y_n$ about $X=x$. Moreover, The Nadaraya-Watson estimator can be seen as a particular case of a broader class of nonparametric estimators, the local polynomial estimators. Precisely, Nadaraya-Watson corresponds to performing a regular local fit; this is so because it approximates $m(x)$ locally by a continuous $m(x)\approx m_0$ via the method the local least square method, hence commonly known as the local least square estimator. This estimator has, however, certain limitations, among which; it yields a poor approximation in the presence of genuinely linear data, as it yields a non-linear output. Moreover, it is inconsistent at the boundaries; if $m(x)$ is positively sloped, the Nadaraya-Watson estimator will be Upward biased. The following section gives a broader class of nonparametric estimators and their advantages concerning the Nadaraya-Watson estimator.

3.2.2. Local Polynomial Estimator

We motivated the NW estimator at *x* as an average of the y_i for observations in a neighborhood of *x*: A local constant approximation. Instead, we can do OLS in the same neighborhood. If we use a weighting function (kernel function), giving each weight on every point *x* locally, this is called the local polynomial (LPE) estimator. The idea is to fit the local model

$$Y_i=\beta_0+\beta_1(X_i-x)+\beta_2(X_i-x)^2+\cdots+\beta_p(X_i-x)^p+\epsilon_i \tag{7}$$

The estimator is obtained via:

$$\sum_{i=1}^{n}\left(Y_i-\sum_{k=0}^{p}\left(\beta_k(X_i-x)^k\right)\right)^2 K_h(X_i-x)=\min_{\beta_0,\cdots,\beta_p}! \tag{8}$$

Which yields the coefficients below:

$$\hat{\beta}=\left(X^{\mathrm{T}}WX\right)^{-1}X^{\mathrm{T}}WY \tag{9}$$

where $W=W_{ii}$ with $W_{ii}=K_h(X_i-x)$, $Y=(Y_1,\cdots,Y_n)^{\mathrm{T}}$.

Moreover, the local polynomial estimator is preferable to the Nadaraya estimator as it preserves the linearity of data and behaves better at boundaries. However, it uses a single polynomial to give a fit to the data, which is disadvantageous in the presence of non-linear data as it makes the estimator less flexible in such situations.

3.2.3. B-Splines

To overcome the problem of flexibility in the presence of non-linear data, the use of piecewise polynomial functions with local support concerning a given degree and domain of partition was proposed [5]. in other words, in the case of local polynomial estimation, the relationship between the dependent and independent variables was the same for their entire domain. Splines, in contrast, can split a problem into multiple local solutions, which can all be combined to produce a useful global solution. The general equation

$$f(X)=\beta_0+\beta_1B_1(X_1)+\beta_2B_2(X_2)+\cdots+\beta_mB_m(X_m) \tag{10}$$

where B_j are the basis function of a known degree ν and knot, it is defined as follows:

$$B_j(x;\nu) = \frac{x - t_j}{t_{j+\nu} - t_j} B_j(x;\nu-1) + \left(1 - \frac{x - t_{j+1}}{t_{j+\nu} - t_{j+1}} B_{j+1}(x;\nu-1)\right) \tag{11}$$

with

$$B_j(x;0) = \begin{cases} 1 & t_j \le x \le t_{j+1} \\ 0 & \text{elsewhere} \end{cases} \tag{12}$$

with t_j being the knots, β_j are the coefficient vector of the basis B-splines.

However, B-splines tend to overfit when the number of knots increases ([6]). Most of the techniques proposed above are less performant in sparse data. Thus, the need for a more robust regression model.

3.2.4. Individual P-Spline Quantile Regression in Varying Coefficient Model

In the classical multiple linear regression, the coefficients are assumed to be constant and tend to be less efficient in sparse data. Thus, indicating the necessity of a more flexible regression model which considers the coefficients as an unknown smooth function of another variable. Here the regression coefficient is estimated by the means of P-splines. The equation characterizing this model is given as follows:

$$\begin{aligned} q_\tau\left(Y(t) \mid X(t), t\right) &= \beta_0(t) X^{(0)}(t) + \beta_1(t) X^{(1)}(t) + \cdots + \beta_p X^p(t) + \epsilon(t) \\ &= \beta_0(t) + \sum_{k=1}^{p} X^{(k)}(t) \beta_k(t) \end{aligned} \tag{13}$$

where $X(t) = \left(X^{(0)}(t), \cdots, X^p(t)\right)^{\mathrm{T}}$ and $\beta(t) = \left(\beta_0(t), \cdots, \beta_p(t)\right)^{\mathrm{T}}$ is the vector of unknown regression coefficient functions at time t, with $\beta_0(t)$ the baseline effect [6]. The main issue of this model is crossings [7]. Thus, the need for a model to overcome this is known as Adapted He approach [7] [8].

3.2.5. Adapted He Approach

The adapted He approach is more appropriate for the other methods that exist to deal with the crossings [7], as it can be adapted to our varying coefficient model. The model is formulated as follows:

$$\begin{aligned} Y(t) &= X^{\mathrm{T}}(t)\beta(t) + V(t)\epsilon(t) \\ q_{\tau h}\left(Y(t) \mid X(t), t\right) &= X^{\mathrm{T}}(t)\beta(t) + V(t) a^{\tau h}(t) \end{aligned} \tag{14}$$

it relies on two assumptions:

- **H1**: The (conditional) median quantile of the error term $\epsilon(t)$ equals zero: $q_{0.5}\left(\epsilon(t)\right) = 0$.
- **H2**: The (conditional) median quantile of the absolute value of the error term $\epsilon(t)$ equals one: $q_{0.5}\left(\left|\epsilon(t)\right|\right) = 1$.

The algorithm of the adapted He approach operates in three steps:

- **Step 1** Estimate $\beta(t)$

$$q_{\tau h}\left(Y(t)\mid X(t),t\right)=X^{\mathrm{T}}(t)\beta(t)+V(t)a^{\tau h}$$

- **Step 2** Estimate $V(t)$

$$q_{0.5}\left(\left|Y(t)-X^{\mathrm{T}}(t)\beta(t)\right|\right)=V(t)$$

- **Step 3** Rewrite:

$$\left|Y(t)-X^{\mathrm{T}}(t)\beta(t)\right|=V(t)\epsilon(t)$$

Estimate the unknown τ-th (conditional) quantile of the error term, $a^{\tau h}(t)$
The estimated quantile function is then given by:

$$\hat{q}_{\tau h}\left(Y(t)\mid X(t),t\right)=X^{\mathrm{T}}(t)\hat{\beta}(t)+\hat{V}(t)\hat{a}^{\tau h}$$

The optimal choice of the knots will be based on https://github.com/AlbertoRodrigues/estimating_knots_regression_splines_model/tree/main/estimation_number_knots .

The results of the various methods discussed are presented in the next session.

4. Results

This section discusses the results obtained from applying the various non-parametric approaches discussed above on our dataset (house supply data).

4.1. Non-Parametric Density Estimation Results

This section presents the exploration of our data through the use of density of the different non-parametric functions discussed above to study how these functions behave on this data. Moreover, this section aims to show that the different variables needed for our study are not normally distributed. Therefore, this section shall end with the test Shapiro to significantly show that these variables are not normally distributed.

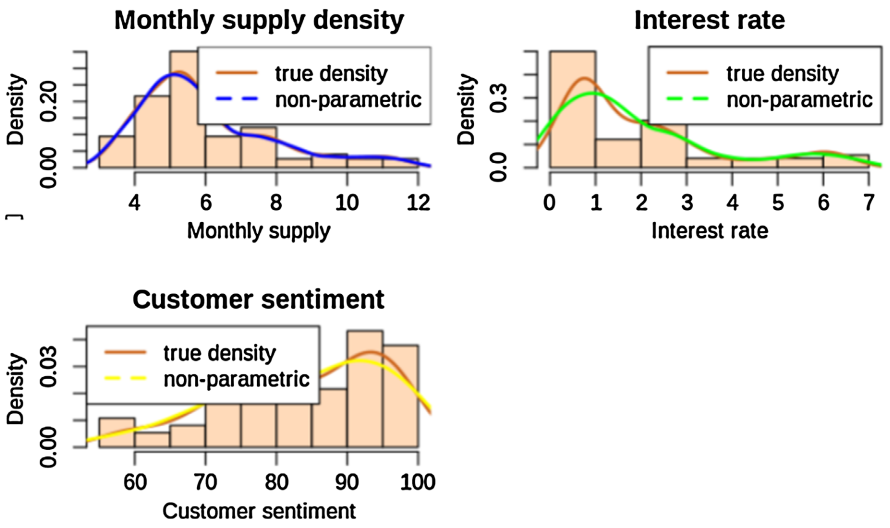


Figure 1. The density plot of the variables considered.

In Figure 1, we notice that the three variables (Monthly supply, Interest rate, and Customer sentiment) are not normally distributed. Moreover, Moreover, using the optimized bandwidths for each of the variables, that is, Silverman(nrd0) optimal bandwidth for monthly supply and customer sentiment (unimodal but not symmetric), Silverman's Long-tailed distribution (Silverman-LT) for interest rate (skewed and long-tailed distribution), it was observed that the kernel density was very close to the true normal density indicating that, the choice of the kernel function does not matter when using optimal bandwidth. The idea of non-normality can be confirmed by looking at the results of the test of Shapiro below:

Table 1. Shapiro test of normality.

Variables	P-values
Monthly supply	<0.0001
Interest rate	<0.0001
Customer sentiment	<0.0001

In Table 1, we can conclude that at a significance level of 5%, these variables are not normally distributed. Thus, a non-parametric regression technique is needed to study the relationship between these variables.

4.2. Non-Parametric Regression Techniques Results

In this section, we study the relationship between monthly supply and interest rate as well as customer sentiment by using the Nadaraya-Watson estimator, Local polynomial estimator, B-splines, and smooth splines with the number of knots equal to 10 and degree 2. The figure below visualizes the relationship between these variables.

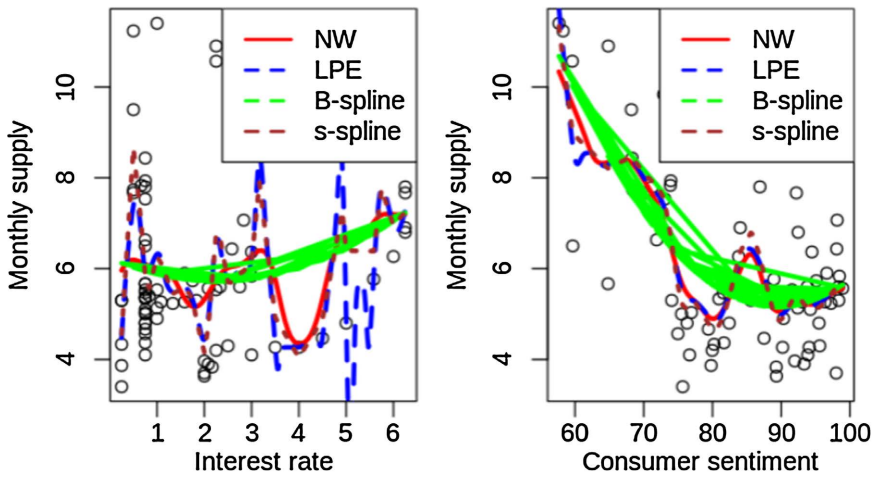


Figure 2. Non-parametric regression.

Figure 2 illustrates the relationships between the variables using various regression techniques. We notice that the Nadaraya-Watson estimator captures the data

well, which is not surprising as it is more flexible in the presence of non-linear data. However, the B-splines do not capture the data very well compared to the other models. On the other hand, the s-splines which penalize the B-spline tend to fit the data more appropriately. The LPE, on the contrary, tends to overfit the model as it tries to capture outlying data. This is so because the LPE uses one polynomial function to give a fit to this non-linear data. Furthermore, the Root mean square in the table below gives an idea of the best model.

Table 2. Root mean square error for interest rate and month supply.

Model	Rmse (SNR)
NW	2.13 (0.65)
LPE	2.29 (0.58)
B-spline	1.85 (0.87)
S-spline	2.47 (0.48)

Table 2 shows each model's root mean square error; it indicates that the B-spline is a more appropriate model for this relationship. Furthermore, looking at this diagram regarding the b-splines, we can see that as the number of house supplies slightly increases, the interest rate Furthermore, we look at the root mean square error for monthly supply and consumer sentiment; we have the following;

Table 3. Root mean square error for consumer sentiment and month supply.

Models	rmse (SNR)
NW	2.60 (20)
LPE	2.54 (21.19)
B-spline	1.38 (66.3)
S-spline	2.47 (20.8)

From **Table 2**, we observe that B-splines still have the smallest root mean square error, and regarding the signal-to-noise ratio, it has the most significant signal-to-noise ratio; thus, it is better than others.

Global Interpretation of the Models

The monthly supply of houses is the number of homes available for sale divided by the number of homes sold each month. A low monthly collection indicates more buyers than sellers, which can lead to higher prices. However, we observe that an increase in the interest rate from 1% to 3% will lead to a very mild decrease in monthly supply, but it will increase slowly with an increasing interest rate. Price inflation might be a possible reason for this change. Furthermore, looking at the graph of monthly supply versus consumer sentiment, we observe that an increase in consumer sentiment will lead to a decrease in supply. This is so because Consumer sentiment measures how confident consumers are about the economy. A

high level of consumer sentiment indicates that consumers are more likely to buy homes, which can lead to higher demand and higher prices, and, consequently, a decrease in monthly supply.

4.3. Robust Non-Parametric Regression

Though highly adapted to non-linear data, the models discussed above remain less efficient for sparse data, as seen in the figures above. Thus, more robust techniques must be implemented in the presence of outliers and leverage points. The figures below present the graphs of individual quantile regression in varying coefficient models and the adapted He approach.

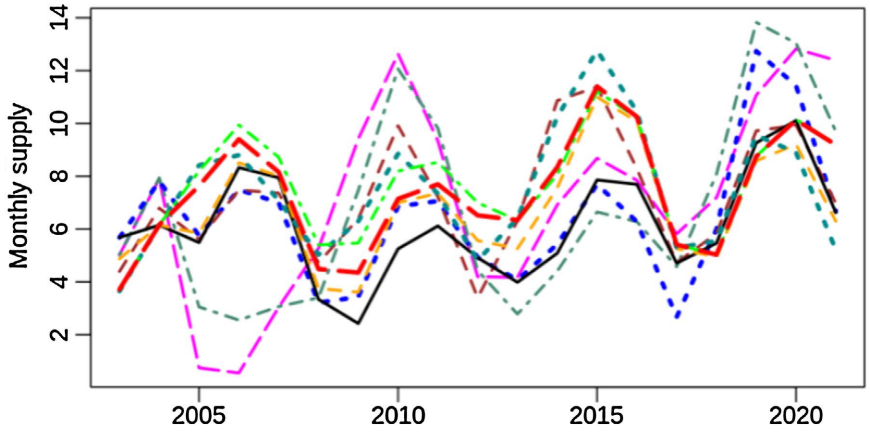


Figure 3. Individual quantile regression in VCM.

Figure 3 shows the monthly evolution over time. However, we observed that the estimated quantile curves cross each other. To overcome this, the crossings Adapted He Approach was applied. And we had the following Figure.

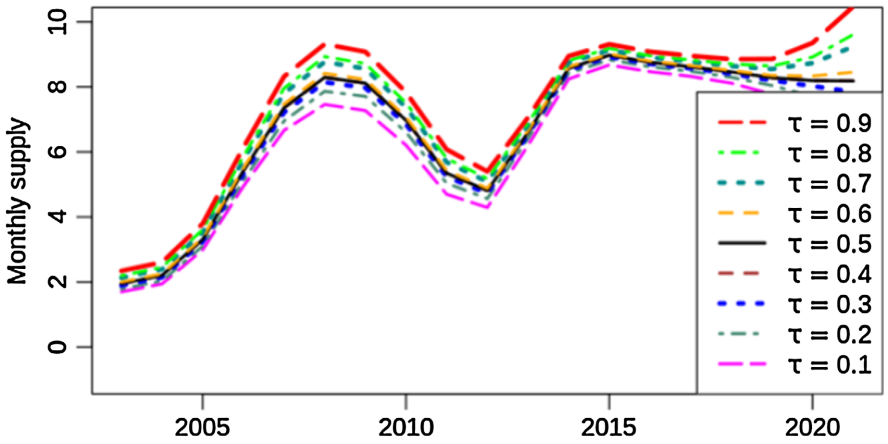


Figure 4. Adapted He quantile regression in VCM.

As shown in Figure 4, there is little or no variation in the distance between the quantile curves. Generally, the quantile curves regarding the monthly supply tend to rise from before 2005 till they peak around 2008 due to the financial crisis as a

possible contributing factor with cheap credit and lax lending standards that fueled a housing bubble. After that, they decreased slightly some years later before rising to 2015, and a mild decrease was observed up to 2020, the COVID-19 pandemic period. However, other factors like interest rates and consumer sentiment could still influence the monthly supply. Below are the graphs of their corresponding coefficients over time.

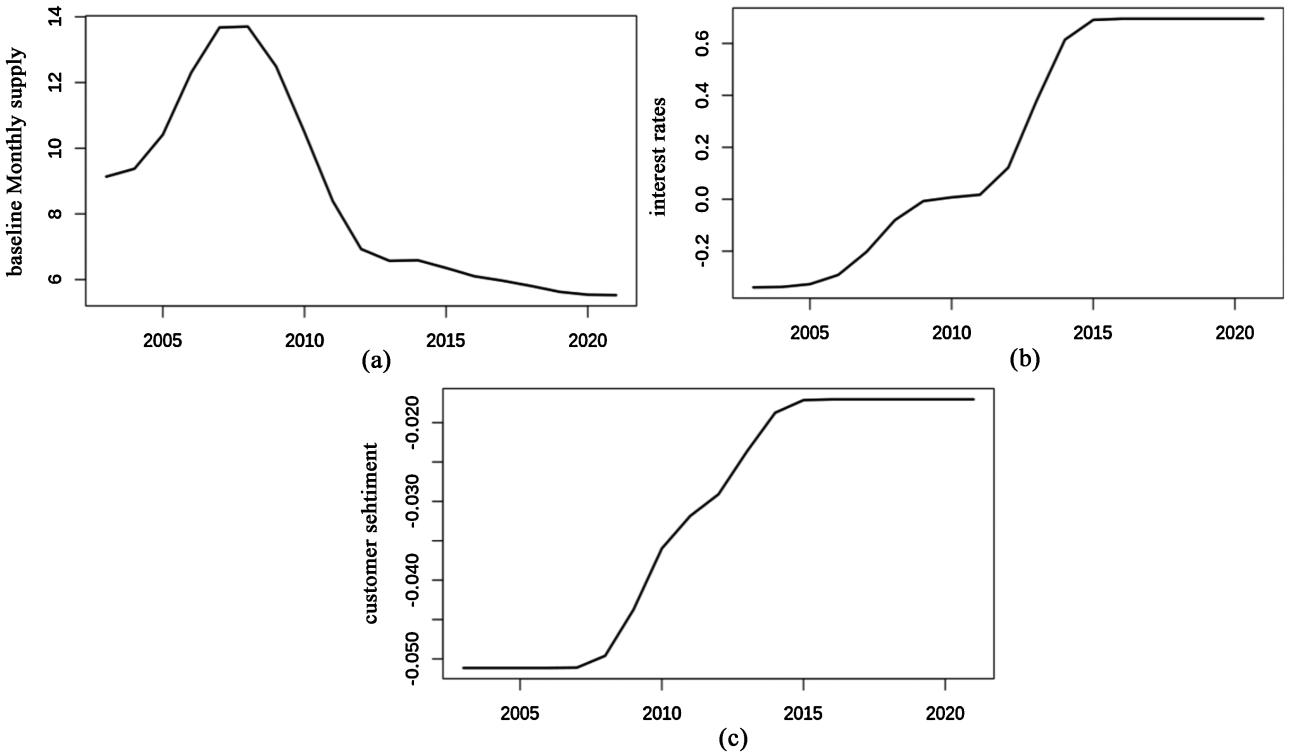


Figure 5. Evolution of the multiple regression parameters over time. (a) baseline; (b) interest rates; (c) consumer sentiment.

Figure 5 shows that both the interest rates and consumer sentiment positively affect the monthly supply over time as they all increase monotonically over time. An increase in the interest rate and consumer sentiment from 2005 to 2015 caused an increase in the monthly supply. However, at baseline, the housing supply rose quickly and peaked around 2008 before declining monotonically over time. Moreover, these graphs are in the same direction as the simple non-parametric regression we did above 2. According to Fred, this observation could be explained by the fact that the wave of foreclosures during the recession period (financial crisis) may have made people warier about homeownership. Tighter credit conditions may have reduced access to mortgage credit, placing homeownership out of reach for many households. Real estate investors may buy properties to generate rental income, simultaneously bidding up the prices of homes while decreasing the supply available to potential homeowners.

5. Conclusion

This study aimed at analyzing the relationship between interest rates and

consumer sentiment. Non-parametric techniques to be used in this study were briefly discussed to achieve this objective. After that, we visualize the plots. Using optimized bandwidths while using the kernel density estimation, we check the non-normality of our different variables and confirm our results using the Shapiro test for normality. Moreover, the relationship between monthly supply and interest rate, monthly supply and consumer sentiments were then analyzed using non-parametric regression techniques such as the Nadaraya-Watson estimator, Local polynomial estimator, B-spline, and smooth(penalized) splines by using optimal bandwidths and knots respectively. Visually, the interest rate had little or no effect on the monthly supply, whereas a negative relationship between monthly supply and consumer sentiment was recorded. However, due to the inefficiency of these non-parametric techniques in the presence of outliers or no standard conditional distributions and to study the relationship between both consumer sentiment and interest rates, we went for more robust non-parametric regression techniques among which; individual quantile regression in varying coefficient models and Adapted He approach. The former tends to produce crossings in the estimate's quantile curves. The latter model was very efficient regarding the non-crossings of the estimated quantile regression curves. Diagrammatically we observed that the regression coefficients for interest rates and consumer sentiments were positive, which aligned with the conclusions drawn from the previous simple non-parametric regression as the baseline monthly supply estimate is pessimistic.

Conflicts of Interest

The author declares no conflicts of interest.

References

[1] The FRED (2024) Housing Recoveries Without Homeowners: National Trends. https://fredblog.stlouisfed.org/2018/03/a-housing-recovery-without-homeowners/?utm_medium=related_content&utm_source=series_page&utm_term=

[2] Rosenblatt, M. (1956) Remarks on Some Nonparametric Estimates of a Density Function. *The Annals of Mathematical Statistics,* **27**, 832-837. https://doi.org/10.1214/aoms/1177728190

[3] Hastie, T. and Tibshirani, R. (1987) Non-Parametric Logistic and Proportional Odds Regression. *Journal of the Royal Statistical Society*: *Series C* (*Applied Statistics*), **36**, 260-276. https://doi.org/10.2307/2347785

[4] Bierens, H.J. (1994) The Nadaraya-Watson Kernel Regression Function Estimator. *Topics in Advanced Econometrics*: *Estimation*, *Testing*, *and Specification of Cross-Section and Time Series Models*. Cambridge University Press, 212-247. https://doi.org/10.1017/CBO9780511599279.011

[5] De Boor, C. (2001) A Practical Guide to Splines (Applied Mathematical Sciences, 27). Springer.

[6] Tantular, B., Ruchjana, B.N., Andriyana, Y. and Verhasselt, A. (2023) Quantile Regression in Space-Time Varying Coefficient Model of Upper Respiratory Tract Infections Data. *Mathematics*, **11**, 855. https://doi.org/10.3390/math11040855

[7] Y Andriyana, Y., Gijbels, I. and Verhasselt, A. (2016) Quantile Regression in Varying-

Coefficient Models: Non-Crossing Quantile Curves and Heteroscedasticity. *Statistical Papers*, **59**, 1589-1621. https://doi.org/10.1007/s00362-016-0847-7

[8] Andriyana, Y., Gijbels, I. and Verhasselt, A. (2014) P-Splines Quantile Regression Estimation in Varying Coefficient Models. *TEST*, **23**, 153-194. https://doi.org/10.1007/s11749-013-0346-2