An Analysis of the Impact of Rent Control on New York City Housing

Benjamin W. Schweitzer Robert Garrett Lydia Carter
Alison Tuiyott Karsten Maurer Thomas J. Fisher

Abstract

It is the concern of policy makers every year in New York whether or not the enacted rent control policy has a positive effect on the New York City rental market. In order to measure the efficacy of this policy, we aim to study the change in housing quality of people who live within these rent controlled homes. A housing quality index metric was created in order to study how it changes over time in relationship to rent controlled versus non rent controlled properties. The impact of rent control on housing quality will be analyzed, thus assessing policy effectiveness. The analysis indicates that rent controlled homes are associated with higher damage rates than non rent controlled homes perhaps indicating that the inverse of the intended effect is occurring.

Key Words: Multivariate Multiple Regression, Rent Control, Survey Analysis

1. Introduction

For the 2019 Statistical Computing Data Expo, participants were provided with New York City Housing and Vacancy Survey data from 1991 to 2017 along with a series of guided question that participants could pursue for the research project [4]. This project addresses a combination of the first two guided questions which can be seen below in order to study the relationship between rent control status and the quality of a house.

> 1. Create a housing quality index for the NYCHVS that enables a view of the housing conditions faced by residents. Contestants may consider the relative importance of different conditions now and/or how the prevalence of these issues has shifted over time.
>
> 2. For the last 50 years, part of the NYC rental stock has been subject to price controls. Currently, about half of the City's rental stock fall under rent control or rent stabilization. Contestants may choose to look at a variety of factors such as quality, housing costs, or population and the relation to rent control status.

Rent control policy in New York is an extremely influential and complex system that effects millions of homes in New York City. Rent control policy has taken on many forms throughout the years and is subject to a great deal of economic and political argument about how it is being implemented and the effectiveness of proposed implementations. The goal of the research project was to introduce a way to model the effect of New York City rent control policy on the quality of housing in New York City over the last twenty years.

The surveys were aggregated to generate three measures of damage to describe the condition of a house: external, utility, and pest. The damages were modeled using multiple predictor variables including an indicator of rent control in attempt

to isolate and measure the effect of these predictor variables on the housing damages. Through interpretation of model coefficients, we were able to observe the modeled relationships between predicted damage rates. The model indicated that homes under rent control while controlling for other factors were more likely to have instances of all three damage types than non rent controlled properties.

2. Data

2.1 New York City Housing and Vacancy Survey

The data provided for the competition was from the New York City Housing and Vacancy Survey from the years of 1991, 1993, and then every three years up until 2017. The United Census Bureau describes the information in the survey as follows.

> "Detailed data from the survey cover many characteristics of the City's housing market, including characteristics of the City's population, households, housing stock, and neighborhoods. The rental vacancy rate is the primary focus of the survey, because that value is crucial to the current rent regulation laws." [4]

The data consists of over 130 features and 14000 rows of coded survey responses for one unit to a variety of questions. The survey responses are coded according to a corresponding data dictionary for each year. The questions give information about the physical conditions of the unit, various costs of living, demographic information about persons living in the unit, rent control information about the unit, persons living in the unit's opinions on the unit, survey weights, and a vast amount of information about each unit surveyed.

2.2 Survey Aggregation

We identified the variables that would indicate a physical external damage to the home such as a broken window, hole in the floor, cracked stairwell, etc. By aggregating across all the variables that indicate a physical external damage, if any damage is deemed present the new indicator variable "External Damage" takes on the value of 1, or 0 if no damages are present. The same process was repeated for the new indicators "Utility Damage" and "Pest Damage" where we would identify any variables that would indicate a utility damage or a pest damage, and then aggregate across those variables to create two new indicator variables,"Utility Damage" and "Pest Damage". If any utility is deemed to have a damage or pest deemed present from the information gathered from the collection of variables that indicate these damages, the new indicator variable has a value of 1, or 0 if there is no damage across all indicators. This same process is used to determine whether or not a unit was under rent control. Since there are many variables that indicate whether or not a unit is under a specific kind of rent control policy, by aggregating across each of these features and checking if at least one version of rent control is present for a property you can generate a new cumulative "Rent Control" indicator. This allows for the reduction of multiple specific rent control indicators into one feature. Lastly the homes above the 75th percentile of monthly rent costs were removed since they are not the target of rent control policy, and then the remaining units were binned by quartiles of the monthly rent cost feature and a new variable was created that indicates which monthly rent quartile a unit belongs to.

2.3 Exploratory Data Analysis

This data set can be explored visually to identify trends in the data to aid the modeling stage. In Figure 1, the median sub-borough utility damage rate of homes in Brooklyn over time is displayed, along with bands that represent the middle 50 percent of sub-borough damage rates that were used to calculate the borough wide proportion. There exists a clear gap between the rent controlled and non rent controlled utility damage rates each year, indicating there may a correlation between rent control status and higher utility damage rates.

In Figure 2, the proportion of homes with an external damage partitioned by borough, rent quartile, and whether or not the home was under rent control is displayed. It is clear that in almost all boroughs, for each rent quartile, the proportion of rent controlled homes with an external damage is higher than the proportion of homes not under rent control. This trend is seen in most years in all three damage types, reinforcing the trend we started to see in Figure 1, indicating a potential positive correlation between rent control status and higher damage rates in the three categories.

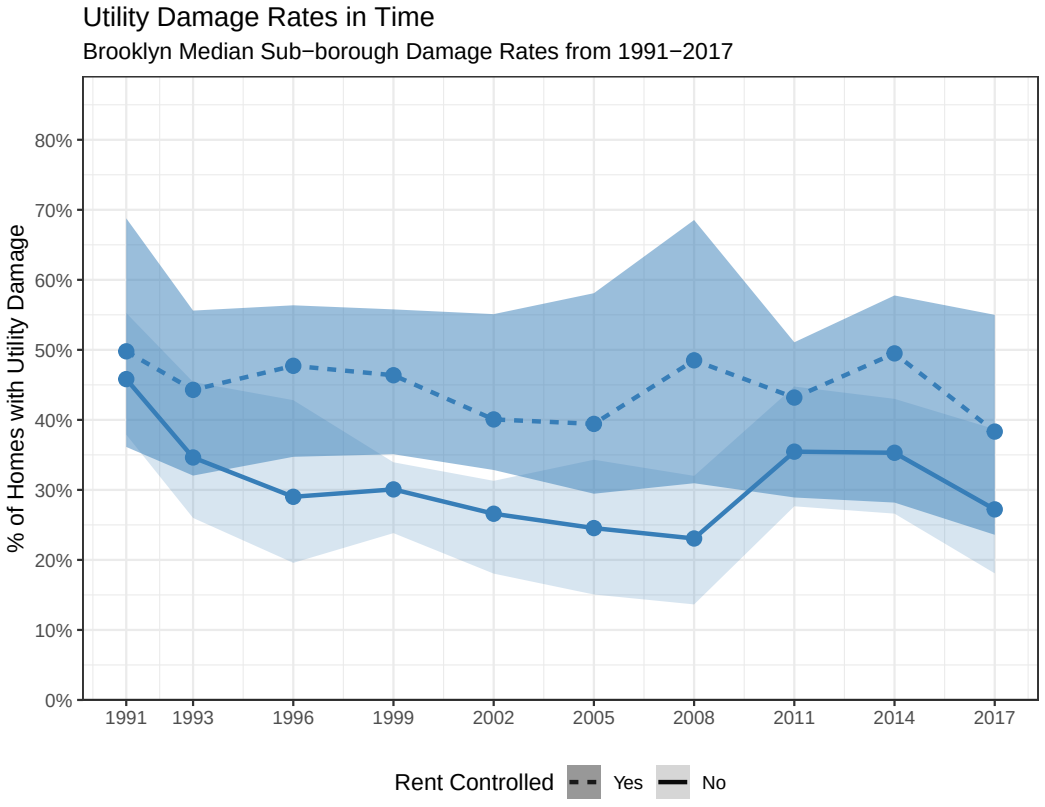

Figure 1: Brooklyn median utility damage rate from 1991-2017 grouped by rent control status. The bands represent the inter-quartile range of the utility damage rate. The rent controlled homes had higher median utility damage rates every year during this time period, and for stretches the 25 percentile for the rent controlled utility damage rate was higher than the non rent controlled homes' 75 percentile.

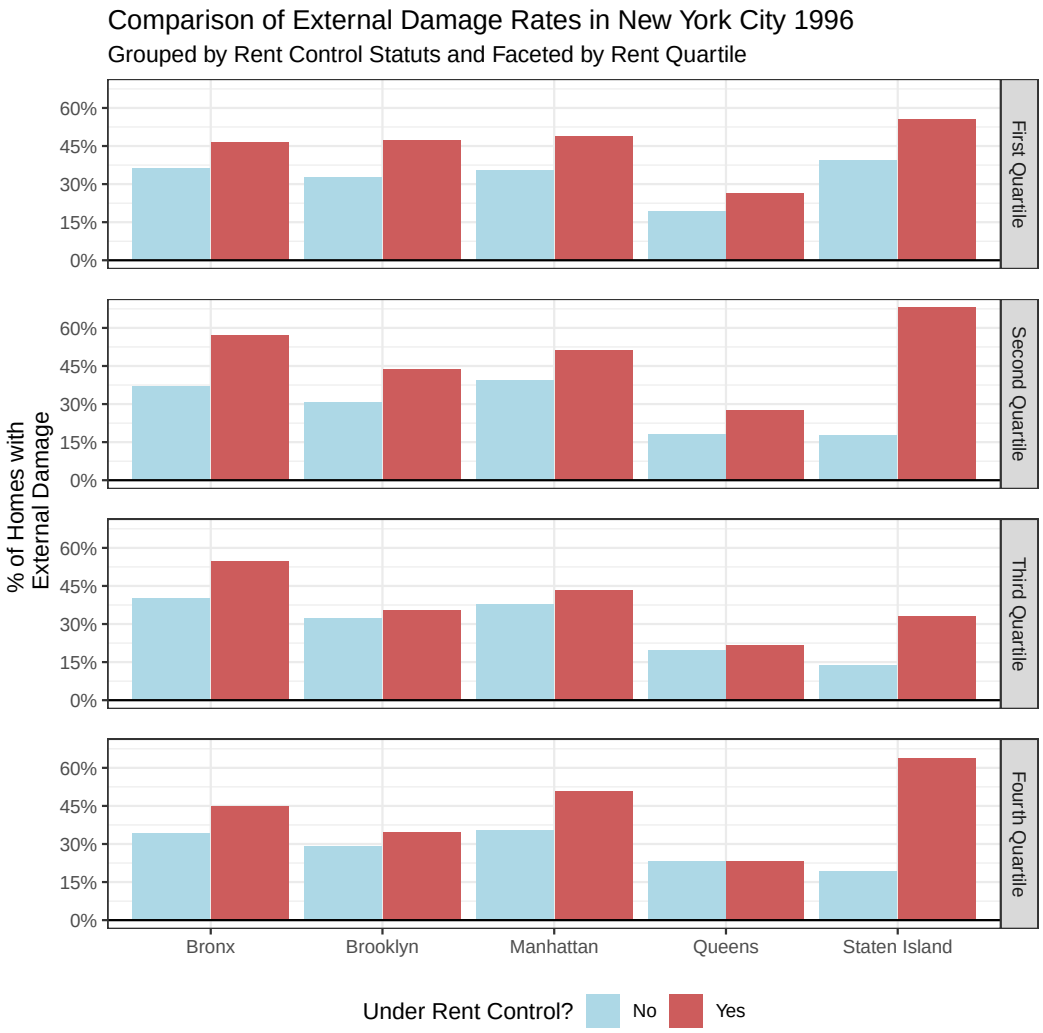

Figure 2: Proportion of homes with a external damage rates for each of the five NYC boroughs in 1996 faceted by first, second, third, and fourth rent quartile grouped by whether or not the property is under rent control. In almost all comparisons of damage rates, the proportion of rent controlled homes with damages is higher than the non rent controlled homes.

3. Methods

3.1 Bayes Regularization

During the aggregation process, we ran into a challenge with some of the sub-borough level proportions. In some cases, there would only be a few homes to represent the rent controlled or non-rent controlled homes within a rent quartile for a sub-borough. So, if there are only a few houses and they all have or do not have a damage present, the estimated proportion for the whole sub-borough becomes 100% or 0%. This is most likely incorrect, so we must do some form of correction to the proportions to ensure the sub-boroughs proportions are being accurately represented. Through Bayes regularization using a Beta conjugate with shape parameters chosen from groupings that didn't result in a 0% or 100%, we were able to 'regularize' the proportions to ensure the groups with few houses to represent the entire group have proportions we can use in the modeling stage and this change can be seen in Figure 3 [1].

$$\mathbf{p} = \frac{n}{n+\alpha+\beta} \times \hat{p} + \frac{\alpha+\beta}{n+\alpha+\beta} \times \frac{\alpha}{\alpha+\beta}$$

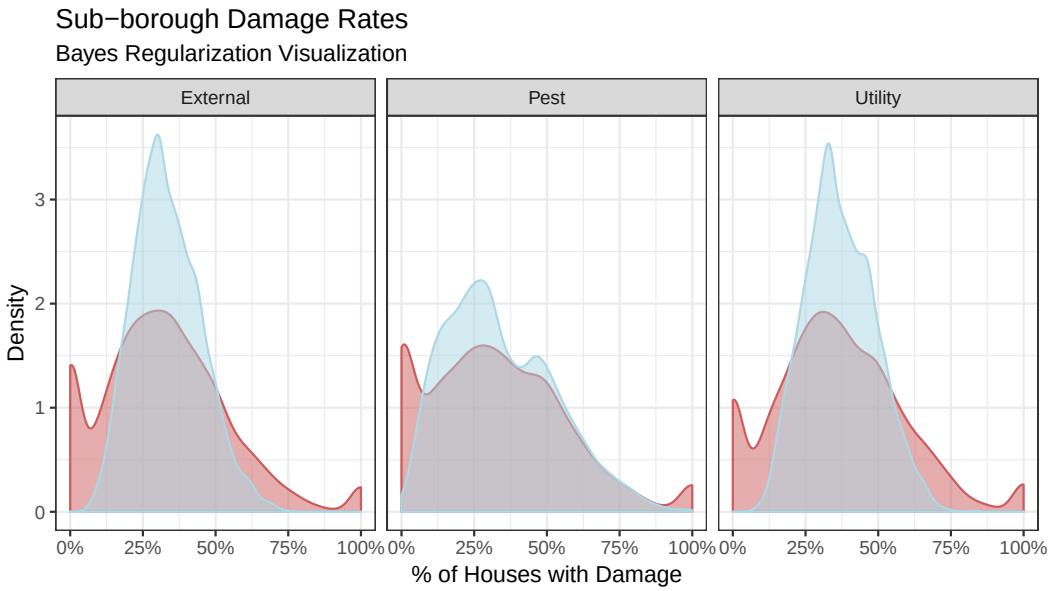

Figure 3: Density plots of the three damage rates before and after the Bayes regularization. The responses before regularization (red) have peaks at 100% and 0% which creates issues during the modeling stage. Post regularization (blue) the responses have been normalized and the peaks are no longer present, allowing us to move forward with modeling.

3.2 Model

The model is a multiple response multivariate model with the responses being the transformed log odds ratio of the damage proportions, and predictor features being rent quartile, rent control, borough, interaction of rent control and time, and interaction of rent control and rent quartile. The response is of the form:

$$\mathbf{Y} = \begin{bmatrix} log\left(\frac{External}{1-External}\right) \\ log\left(\frac{Utility}{1-Utility}\right) \\ log\left(\frac{Pest}{1-Pest}\right) \end{bmatrix}$$

The predictors variables we used to model the response are as follows. Rent Control (RC) has a value 1 if home is under rent control policy, 0 if otherwise. I^{th} Rent Quartile (Qi) is 1 if home falls into the i^{th} rent quartile, 0 if otherwise and if all Qi's are 0 the home is in the 1st rent quartile. Year (T) has a value of 1990 subtracted from the year of the observation. $Brooklyn$ has a value 1 if home is in Brooklyn, 0 if otherwise. $Manhattan$, $Queens$, $StatenIsland$ use the same logic, and if all borough variables have a value of 0, then the home is in the Bronx.

$$\begin{aligned}\mathbf{Y} =& \boldsymbol{\beta_0} + \\ & \boldsymbol{\beta_1} \times T + \\ & \boldsymbol{\beta_2} \times Q2 + \boldsymbol{\beta_3} \times Q3 + \boldsymbol{\beta_4} \times Q4 + \\ & \boldsymbol{\beta_5} \times RC + \\ & \boldsymbol{\beta_6} \times Brooklyn + \boldsymbol{\beta_7} \times Manhattan + \boldsymbol{\beta_8} \times Queens + \boldsymbol{\beta_9} \times StatenIsland + \\ & \boldsymbol{\beta_{10}} \times RC \times T + \\ & \boldsymbol{\beta_{11}} \times RC \times Q2 + \boldsymbol{\beta_{12}} \times RC \times Q3 + \boldsymbol{\beta_{13}} \times RC \times Q4\end{aligned}$$

4. Results

The fitted model estimate coefficient confidence intervals for each response can be seen in Figure 4, allowing for interpretation of model coefficients.

This allows for a closer look at the modeled relationship between rent control and the three damages. The rent control coefficient confidence interval for the external damage is (0.1963, 0.3472), for utility damage is (0.2269, 0.3719), for pest present (0.0917, 0.3631). All three of these are positive, significant predictors which indicates that homes under rent control are correlated with larger log odds of all three damage rates, indicating that homes under rent control are correlated with higher damage rates in all three cases.

A few other interesting modeled effects can be seen by looking at the significantly non-zero coefficients. It appears that the aforementioned higher damage rates for rent controlled homes is larger for homes with higher monthly rent values. Perhaps indicating that rent controlled homes that charge higher monthly rent are correlated with a larger gap in the damage rates between rent controlled and non rent controlled homes. Houses from Queens and Staten Island have lower damage rates than houses from the Bronx, Brooklyn, and Manhattan, indicating there is some relationship between borough and the damage rates for all three categories. The proportion of homes with pests present has risen over time, perhaps due to survey changes, or better reporting of pests. The proportion of homes with external and utility damage has fallen over time, leading to the conclusion that damage rates across all of New York have been decreasing over time.

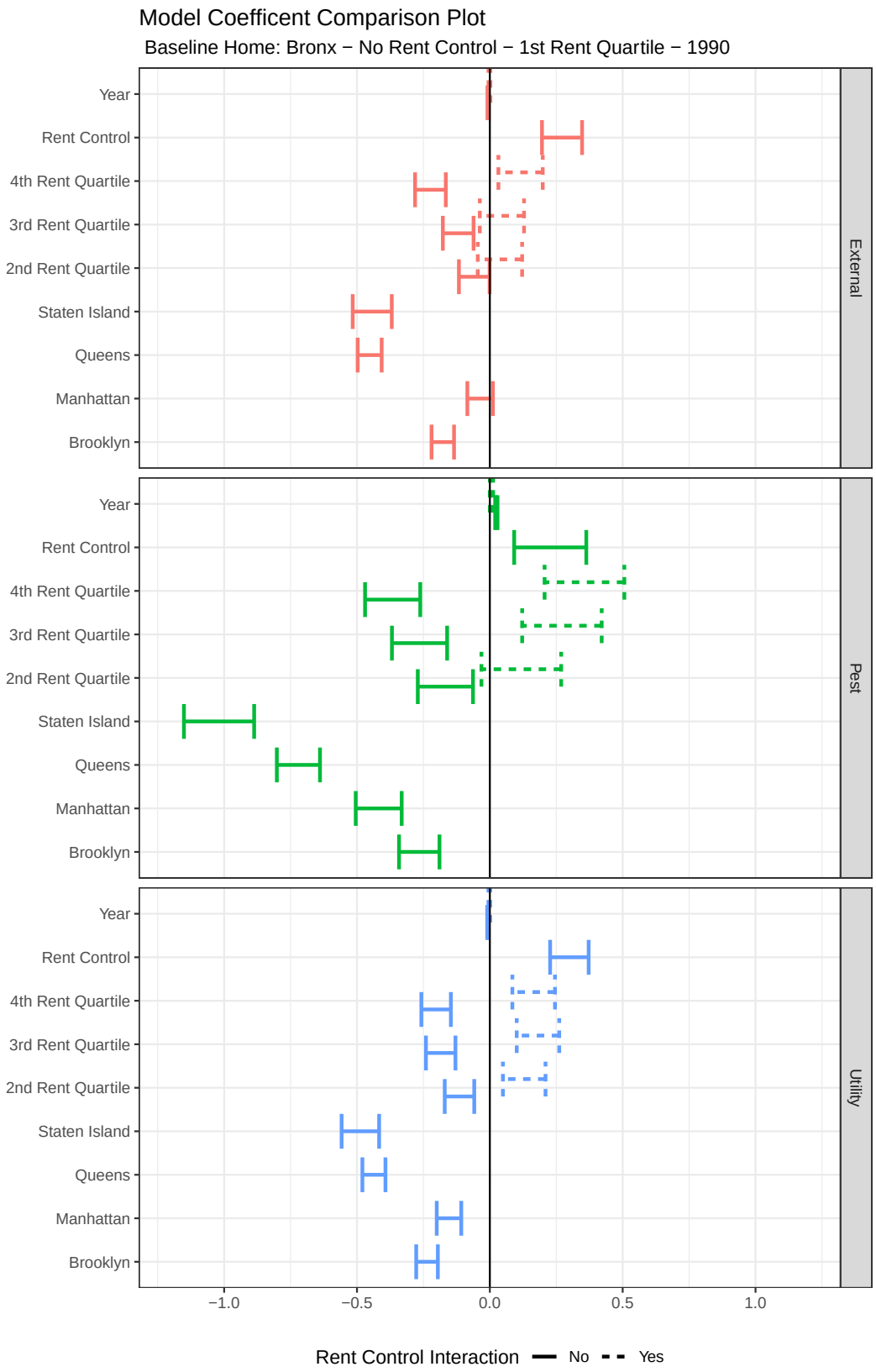

Figure 4: Visualizes the change in predicted log odds of the three responses if the unit were to have any changes in the predictor variables from the baseline case, a house in the Bronx under no rent control in the first rent quartile during 1990.

5. Conclusions

In conclusion, through this modeling process we measure the relationship between rent quartile, rent control, borough, interaction of rent control and time, and interaction of rent control and rent quartile and the three damage rates. This resulted in the isolation of a significant correlation between rent control status of a home and higher damage rates in all three measures of damage. While damage rates do tend to be decreasing over time, rent control appears to be having a negative impact on housing quality. Perhaps an explanation for this trend could be that landlords are less incentivized to improve/maintain a unit when the current occupants have a monthly rent cost that is fixed thus there is not an avenue to increase income from the property. The purpose of rent control is not just to control the costs of living and create affordable housing, the aim is to create affordable quality housing. In the current form, when controlling for other factors, a house under rent control is more likely to have damages which suggests rent control is not fulfilling the quality housing portion of its goal. This could indicate that rent control policy could use a change in methodology or implementation because while the current methods are creating affordable housing, the housing is not of the same quality of a similar house that is not under rent control policy. Identifying this trend creates the opportunity to change current methods and reverse this observed trend.

References

[1] Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A., Rubin, D. (2013). Bayesian Data Analysis. New York: Chapman and Hall/CRC, https://doi.org/10.1201/b16018

[2] Hadley Wickham (2016). ggplot2: Elegant Graphics for Data Analysis. New York: Springer-Verlag, http://ggplot2.org

[3] R Core Team (2018). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing, https://www.R-project.org/

[4] U.S. Census Bureau. (2018, October 17). New York City Housing and Vacancy Survey (NYCHVS): 1991 to 2017. Washington, DC: U.S. Government Printing Office.