

To The New York Realtors Association

By Monica Zhang

December 9, 2022

Analysis on change of housing prices for single-family residences and single-unit apartments or condos in Brooklyn between Q3 2020 and Q4 2020

Home purchase prices fluctuate as demand, interest rates, and other macroeconomic factors vary. This business memo analyzes whether there was a change in Brooklyn home purchase prices between Q3 2020 and Q4 2020. Our analysis finds there was no significant difference to prices between the two time periods. However, while our model shows no significant difference, there may still have been an upward trend in prices seen on specific zip codes or a specific type of houses in Brooklyn.

To determine if there was any change in housing prices, a linear regression model was built using data from the City of New York for purchases in Brooklyn during years 2016 - 2020. The analysis focuses on single-family residences and single-unit apartments or condos with purchase prices above \$100k and below \$5.5M.

$$\text{sqrt}(\text{price}) = \beta_0 + \beta_1 \text{grosssqft} + \beta_2 \text{zip} + \beta_3 \text{yrquarter}$$

The above linear regression uses total area of the property in gross square feet, zip code, and year along with fiscal quarter to predict *prices*. *Grosssqft* is a numerical variable, while *zip* and *yrquarter* are categorical variables.

The model has an adjusted R^2 of 0.625, meaning it has over 62% explanatory power and the predicted power RMSE is ~450k making our model reliable for analysis. 54 out of 57 predictors are also highly significant to the 0.01% level, meaning the variables included were all necessary to predict prices.

To understand price changes for the third and fourth quarter, we focus on the coefficient β_3 for Q3 and Q4 of years 2017 - 2020. In all Q3 and Q4 for years 2017-2020, β_3 is significant to the 0.1% level. When looking at the table below we notice that:

- When looking at the 95% confidence interval for the coefficients, the intervals for *yrquarter* overlap for years 2017 - 2020.
- The overlap is minimal in 2020 compared to years 2017-2019.
 - This suggests there may still have been a change in prices between Q3 and Q4 of 2020 but, since we still see an overlap, we cannot say that the change was statistically significant.

- We also notice a negative difference in 2018 and 2019, while the difference is positive in 2020.

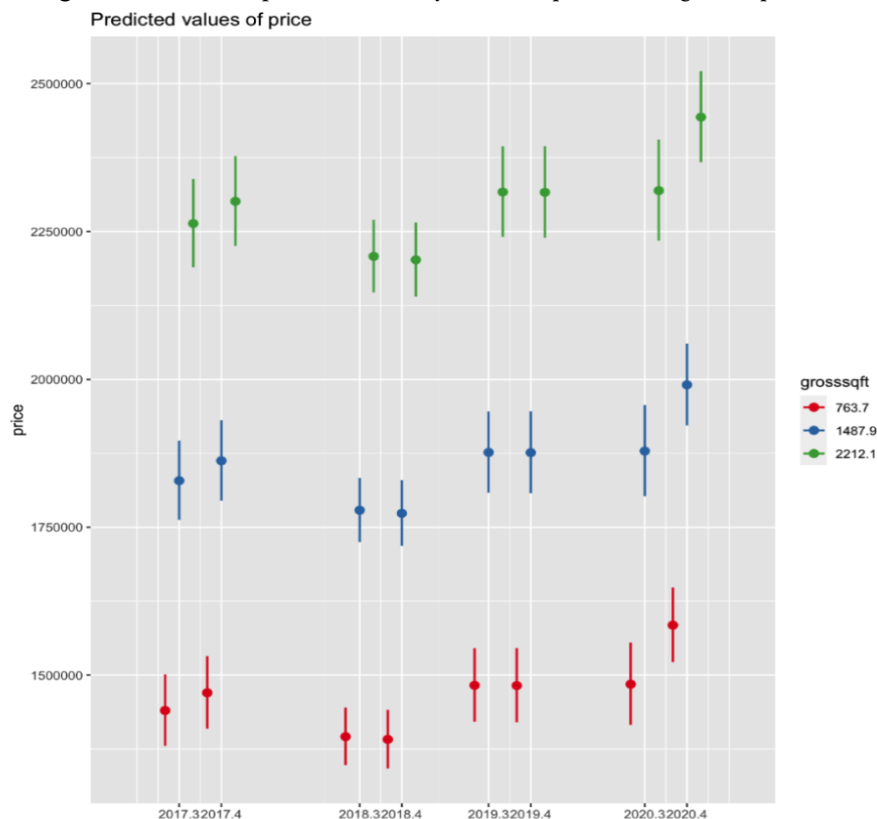
Table 1: 95% Confidence Interval for β_3

95% Confidence Interval for β_3	Q3	Q4	Difference from Q3 to Q4
2017	(60.0, 104.7)	(72.2, 117.3)	(12.3, 12.6)
2018	(44.5, 83.0)	(42.5, 81.1)	(-2.0, -1.9)
2019	(77.3, 122.6)	(76.8, 122.7)	(-0.5, 0.1)
2020	(74.8, 126.7)	(118.7, 163.3)	(43.9, 36.6)

We further look at our model by plotting houses of different square footage and their predicted price movement between Q3 and Q4 in years 2017-2020. As seen in the graph below, while our estimate of the difference is not statistically significant, we seem to notice an upward trend in prices for houses of the same square footage between Q3 and Q4 of 2020. The biggest difference was observed in the homes with larger square footage. A similar trend was not observed for the previous 3 years.

The trend observed in Q3 and Q4 of 2020 below may be attributed to historically low interest rates or the low supply of houses as everyone increased the time they spent at home and a high demand.

Figure 1: Predicted prices based on year with quarter and gross square footage



While the plot above shows prices may have increased between Q3 and Q4 of 2020, we're uncertain on the movement of the actual prices.

A limitation of our model is that we do not explore how interaction terms play out. Further research could explore the interaction of gross square footage and year quarter, or zip code and year quarter.

- The interaction with zip codes will help us understand if the increase was seen in only specific areas of Brooklyn. This is worth analyzing because demand may have increased in residential areas and decreased in business and office areas as a consequence of the covid lockdown.
- On the other hand, the interaction between year quarter with gross square footage will analyze an effect between the time the home was bought and the size of the property. As the population was forced to work from home because of the pandemic, the increase in prices may have been more steep in larger properties.

Our model is also not ideal because the residuals show heteroskedasticity and they are not normally distributed. Heteroskedasticity may be present given the large range of our observed variables and because outside factors greatly affect prices across the years. Also, while we limited prices to be above \$150k and below \$5.5M, we still see outliers in our data.

Figure 2: Scatter plot of residuals

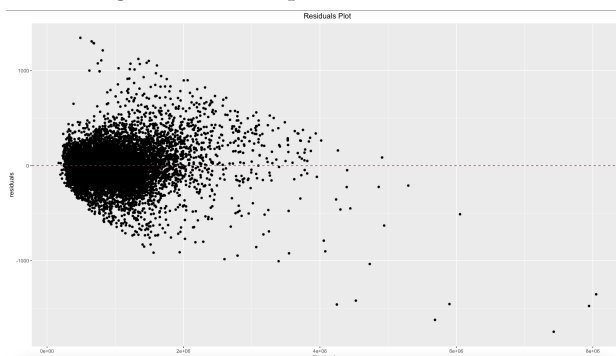
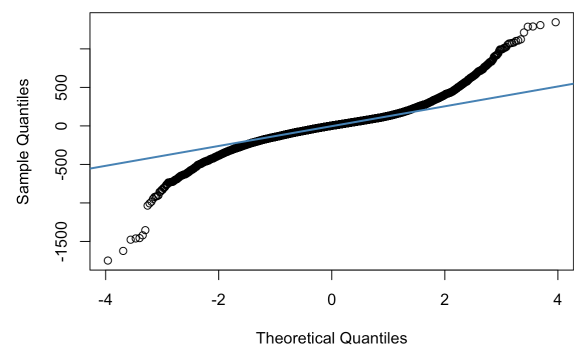


Figure 3: Normal Q-Q plot



In conclusion, we cannot say with certainty if housing prices in Brooklyn changed between QQ3 to Q4 of 2020. We believe that there may have been an increase in certain neighborhoods of Brooklyn and affecting larger houses, but these hypotheses need to be further tested with improvements to the model.