Research on Soaring Home Value in U.S. with Regional Locality Analysis

Principal Investigator: Yunbo Ni yni33@wisc.edu

Introduction

Since the pandemic, home values have been constantly increasing to an extravagant extent. This project aims to uncover possible factors that may have contributed to the soaring prices with a focus on the demand side. In doing so, economic indicators such as Fixed Rate Mortgage Rate, Unemployment Rate, Personal Income, and Covid-10 Cases will be put into consideration. Moreover, an analysis of the localities will be performed over Wisconsin and Illinois, which would compare and contrast the changes in home value among counties with different population densities.
The goal of the project is to determine possible relationships so that stakeholders could take precautionary measures against the soaring home values with these indicators.

Data Description and Methodology

All the data are used in Section 1 are retrieved as or converted to monthly data.
The prime variable in this research is the Home Value Growth Rate. The reason to use percentage change instead of level data is because the growth rate could provide more insight into the future trend of the Home Value.

The primary dependent variables are 30-Year Fixed Rate Mortgage Rate and 15-Year Fixed Rate Mortgage Rate. The secondary dependent variables are Real Disposable Personal Income, Unemployment Rate, and the Covid-19 Case Growth Rate. The reasoning behind Covid-10 Case is similar to why Home Value is analyzed as Growth Rate. The Growth Rate is calculated based on the monthly Covid-19 cases summed over daily data.

In the *Linear Correlation* section, plotting and OLS-Analysis will be conducted over the variables to determine a possible linear correlation.

In the *Locality* section, choropleth maps will be used to visualize the difference in home value change and population density in different Wisconsin and Illinois counties.

Linear Correlation

1.1 Plotting

1.1.1 General Plotting

In this part, the target variable is plotted against all the independent variables to see if there's visible linear relationship among them. The time period is from 2020-01 to 2022-02, which is from the start of pandemic to whichever the last updated time for the variables.

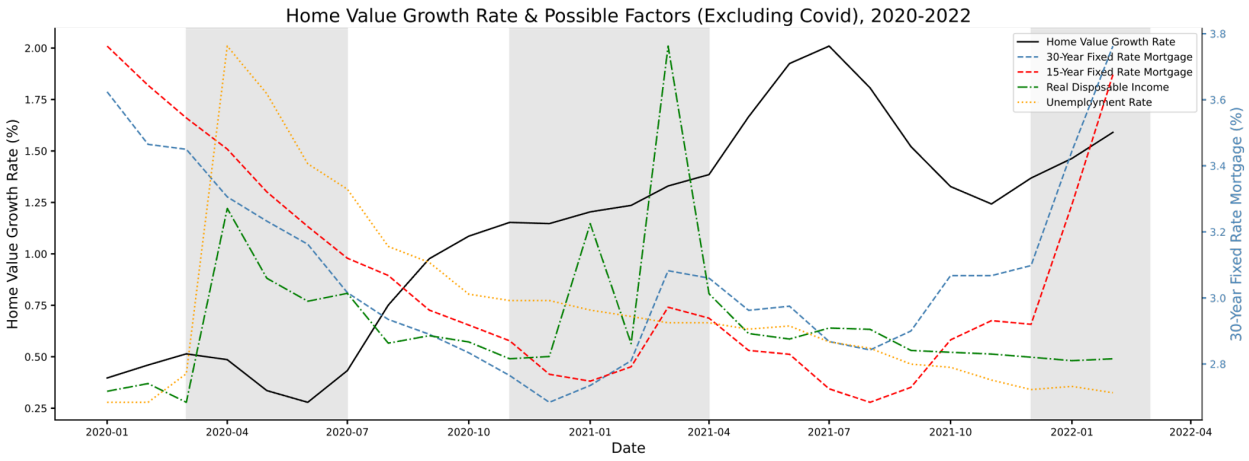


Figure 1

From *Figure 1,* there appears to exist certain negative linear relationship between Home Value Growth Rate and each of our independent variables. However, in the grey boxes, certain variables have shown some odd variations. From 2020-03 to 2020-07, both the Real Disposable Income and the Unemployment Rate have increased significantly. From 2020-11 to 2021-04 the Real Disposable Income have experienced a series of peaks and downs, each to a relatively large extent. From 2021-12 to 2022-2, both Fixed Rate Mortgage Rate have quickly climbed up. Despite these regions, the variations in the independent variables align approximately with the target variable. Further analysis over each variable will be conducted in the following parts.

1.1.2 Primary Research Variables

From *Figure 2*, it is evident that there's a strong correlation between the 30-Year Fixed Rate Mortgage and 15-Year Fixed Rate Mortgage. Hence, only the 30-Year rate is retained to avoid multicollinearity in later analysis. From the graph, we can also tell that there exists a negative linear relationship between Home Value and Mortgage Rates.

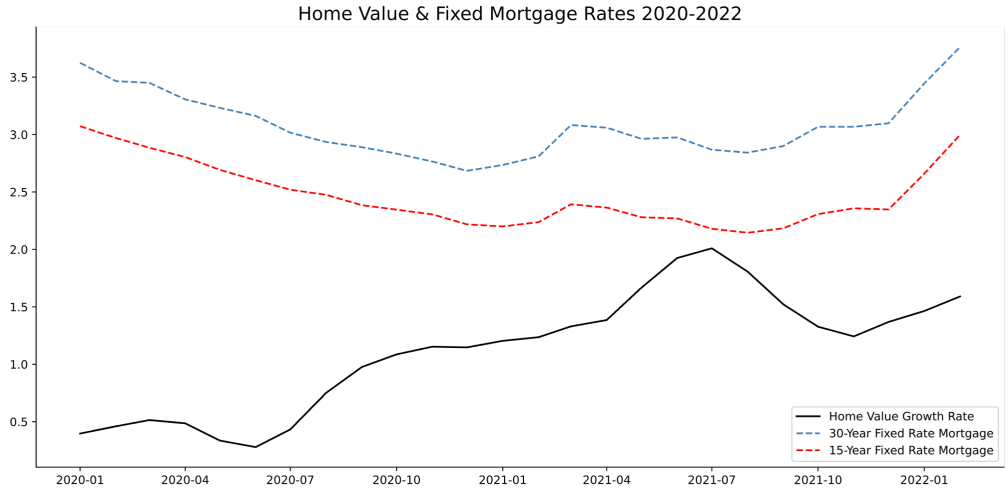


Figure 2

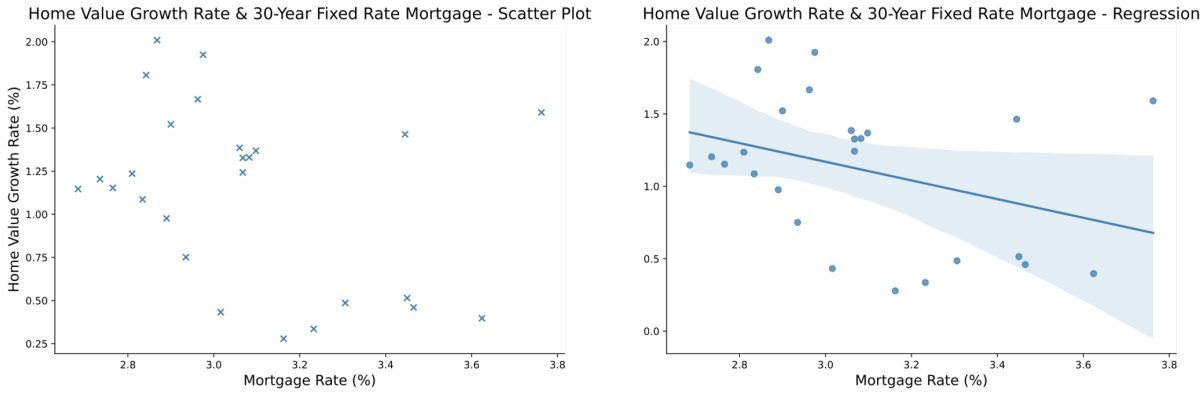


Figure 3

From *Figure 3*, the Home Value Growth Rate turns out to have a possible strong negative relationship with the 30-Year Fixed Rate Mortgage Rate. Statistical Testing will be conducted later.

1.1.3 Secondary Research Variables

This part will focus on the plotting of three secondary variables.

Real Disposable Income:

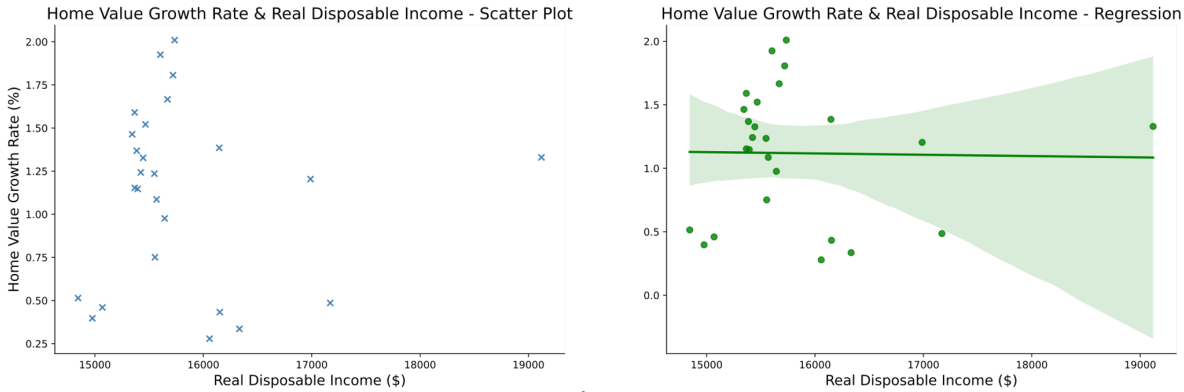


Figure 4

From *Figure 4,* we can't see a clear linear relationship between two variables in concern. The scatter plots mostly vertically cluster at the left side of the graph.

Unemployment Rate:

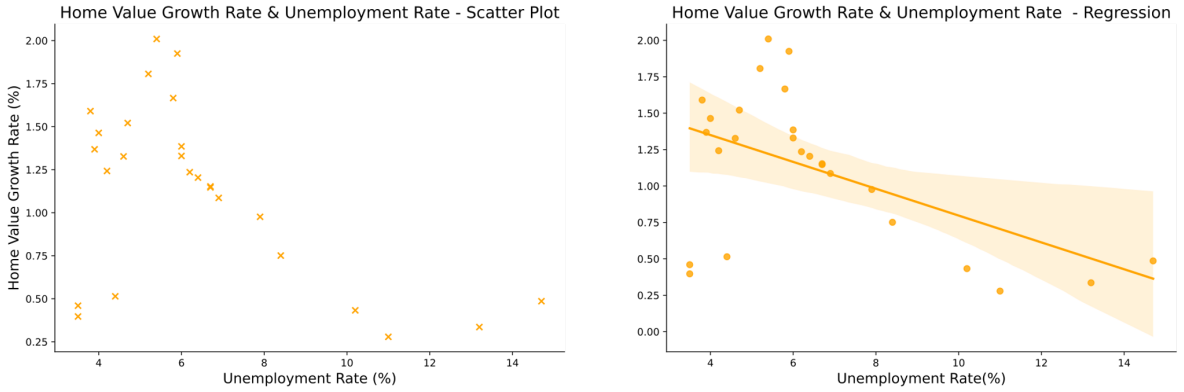


Figure 5

From *Figure 5*, Unemployment Rate seem to have a strong negative relationship.

Covid-19 Case Growth Rate:

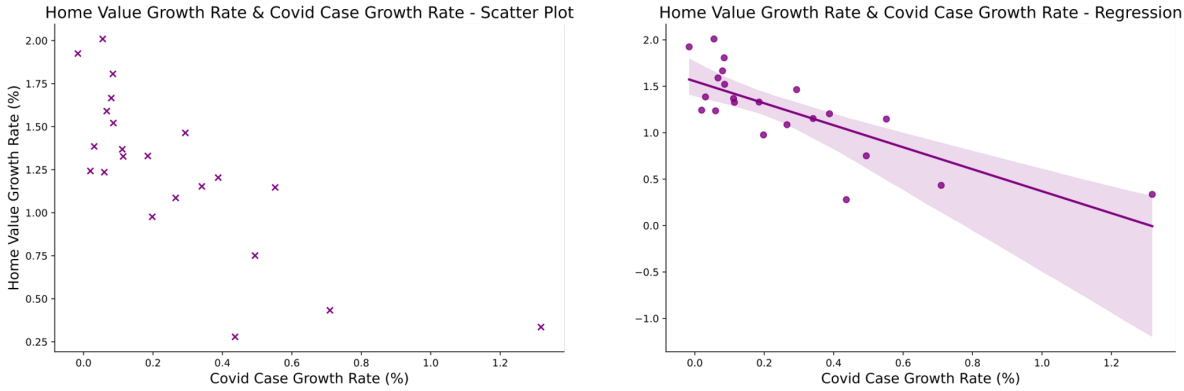


Figure 6

From *Figure 6*, the Home Value Growth Rate also seems strongly correlated with the Covid-19 Case Growth Rate. It needs to be pointed out that due to the exponential growth of Covid-19 case in the first quarter of 2020, the data points start from 2020-05.

Summary

The selected variables exhibit linear relationship with the Home Value Growth Rate to different extent. Statistical testing will be performed in the next part to determine if the correlation is statistically significant.

1.2 OLS Analysis

1.2.1 30-Year Fixed Rate Mortgage Rate

The test result (*Appendix A.1*) shows that the Home Value Growth Rate is in fact positively dependent on the Fixed Rate Mortgage Rate, which has a P-Value equal to 0.000 and a R-Squared Value equal to 0.800. The positive coefficient seems counter intuitive towards the plotting. This may due to the huge concurrent decrease from 2020-03 to 2020-07 and the huge concurrent increase from 2021-12 to 2022-02. Due to the limited data points, these two changes may dominate the test result.

1.2.2 Real Disposable Income

The test result (*Appendix A.2*) shows that the Real Disposable Income also positively affects the value of Home Value Growth Rate. The P-Value is also equal to 0.000 and the R-Squared Value is 0.827. However, the coefficient is only 0.00007, which doesn't provide us with an economically significant inference. Again, this may due to the lack of data points and the strong volatility of this variable in the 2020-11 – 2021-04 period.

1.2.3 Unemployment Rate

The test result (*Appendix A.3*) shows that the Unemployment Rate also has a positive correlation with the target variable with a P-Value of 0.000 and R Squared Value of 0.561, which is lower than the previous two independent variables.

1.2.4 Covid-19 Case Growth Rate

The test result (*Appendix A.4*) indicates a positive correlation with the target variable with a P-value of 0.044 and R Squared Value of 0.180.

1.2.5 Multivariate Regression

The test results (*Appendix A.5*) shows the multivariate regressions results of all the independent variables. In this case, the P-Value for the 30-Year Fixed Rate Mortgage Rate is no longer significant, with a P-Value of 0.770, and the P-Value for Covid-19 Growth Rate is 0.230, which is also not statistically significant. It is worth notice that the coefficient for both Covid-19 Growth Rate and Unemployment Rate turn to negative under multivariate condition. The test also points out that there may exist strong multicollinearity between the independent variables.

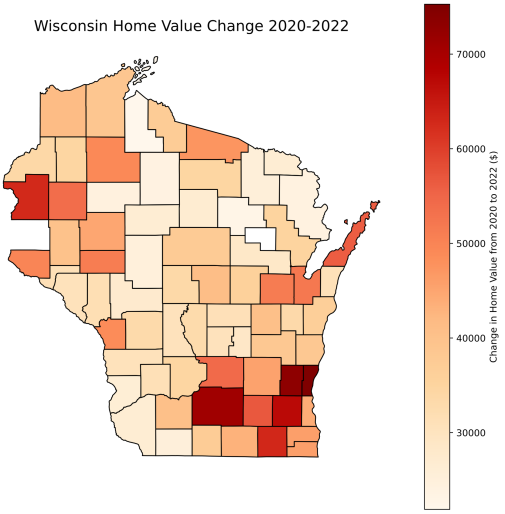
Summary

The OLS analysis results seem to contradict the results from plotting. Given the limited data points, we cannot conclude any strong statistical inference from the variables. However, these independent variables do correlate with the Home Value Growth to different extent. We could conduct more thorough research on the variables to determine a stronger pattern. For the time being, stakeholders could keep an eye on these variables as a possible indicator for the future pattern of Home Value Growth.

Locality

In this part, a simple comparison among regional differences in Wisconsin and Illinois is conducted as a preliminary approach to uncover the possible locality of variations in Home Value Growth Rate.

2.1 Wisconsin



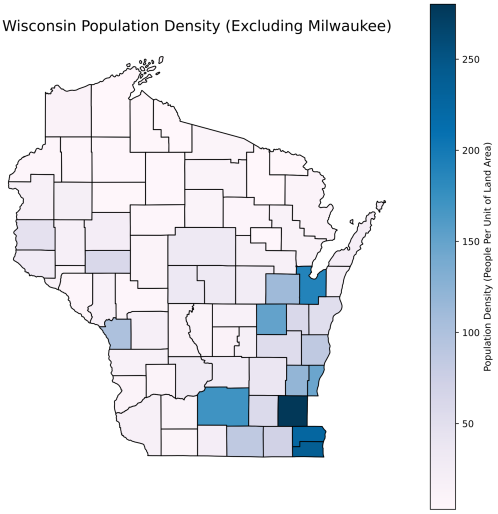


Figure 7 **Figure 8**

Figure 7 shows the change in Home Value from 2020 to 2022, while *Figure 8* shows the population density for Wisconsin Counties (Note that Milwaukee County is excluded because of its extremely higher density compared to all other counties.) From two figures, we can see that there are certain correlations between the density and the value change. In more dense areas, the home value tends to increase to a larger amount. However, there also exists evident outliers such as counties in the north west regions. Even though they have relatively lower density, they mostly experienced a median level increase in home value from 2020 to 2022.

2.2 Illinois

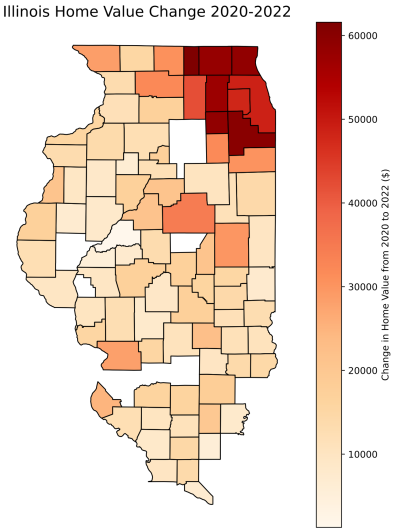


Figure 9

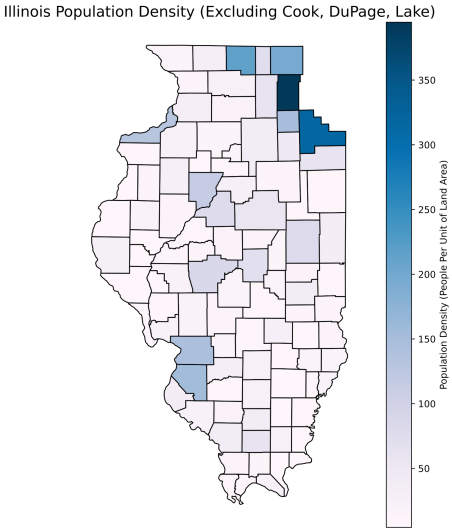


Figure 10

Figure 9 shows the change in Home Value from 2020 to 2022, while *Figure10* shows the population density for Illinois Counties (Note that Cook, DuPage, and Lake County is excluded because of its extremely higher density compared to all other counties.) From the figures, we can see that counties in Illinois exhibit a stronger relationship between population density and home value change. The denser the area, the more changes in the home value. Illinois has less to none outliers in comparison with Wisconsin. It is also worth notice that there are three counties excluded from *Figure10*, which are the densest areas. However, from *Figure9,* we can see that they are not in fact the regions experience the most changes in home values.

Summary

The preliminary locality analysis has produced some promising results. We can infer that the Home Value would experience a larger change when the region is denser with populations. We also notice that the region surrounds the densest regions could also have a more intense change. This may due to the fact that people may move out to the surrounding areas to avoid the already higher Home Value in denser areas. These results are based on Wisconsin and

Illinois. More research could be further conducted to discuss this promising factor that affects Home Value.

Reference

The New York Times. (2021). Coronavirus (Covid-19) Data in the United States. Retrieved [2022.05.06], from https://github.com/nytimes/covid-19-data.

Zillow. (2022). Zillow Home Value Index (ZHVI). Retrieved[2022.05.06], from https://www.zillow.com/research/data/.

US Census Bureau. (2022). Average Household Size and Population Density – County. Retrieved[2022.05.06], from https://covid19.census.gov/datasets/

Appendix A.
1. OLS Regression Results 30-Year Fixed Rate Mortgage Rate

OLS Regression Results

Dep. Variable:	home_value_growth_rate	**R-squared (uncentered):**	0.800
Model:	OLS	**Adj. R-squared (uncentered):**	0.792
Method:	Least Squares	**F-statistic:**	100.2
Date:	Fri, 06 May 2022	**Prob (F-statistic):**	3.14e-10
Time:	19:28:25	**Log-Likelihood:**	-21.283
No. Observations:	26	**AIC:**	44.57
Df Residuals:	25	**BIC:**	45.83
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
30_year_mortgage	0.3554	0.035	10.011	0.000	0.282	0.429

Omnibus:	1.957	**Durbin-Watson:**	0.087
Prob(Omnibus):	0.376	**Jarque-Bera (JB):**	1.309
Skew:	-0.291	**Prob(JB):**	0.520
Kurtosis:	2.067	**Cond. No.**	1.00

Notes:
[1] R^2 is computed without centering (uncentered) since the model does not contain a constant.
[2] Standard Errors assume that the covariance matrix of the errors is correctly specified.

2. OLS Regression Results Real Disposable Income

OLS Regression Results

Dep. Variable:	home_value_growth_rate	**R-squared (uncentered):**	0.827
Model:	OLS	**Adj. R-squared (uncentered):**	0.820
Method:	Least Squares	**F-statistic:**	119.8
Date:	Fri, 06 May 2022	**Prob (F-statistic):**	5.01e-11
Time:	19:28:38	**Log-Likelihood:**	-19.393
No. Observations:	26	**AIC:**	40.79
Df Residuals:	25	**BIC:**	42.04
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
real_disposable_income	7.054e-05	6.44e-06	10.946	0.000	5.73e-05	8.38e-05

Omnibus:	2.139	**Durbin-Watson:**	0.113
Prob(Omnibus):	0.343	**Jarque-Bera (JB):**	1.202
Skew:	-0.159	**Prob(JB):**	0.548
Kurtosis:	1.996	**Cond. No.**	1.00

Notes:
[1] R^2 is computed without centering (uncentered) since the model does not contain a constant.
[2] Standard Errors assume that the covariance matrix of the errors is correctly specified.

---

3. OLS Regression Results Unemployment Rate

OLS Regression Results

Dep. Variable:	home_value_growth_rate	**R-squared (uncentered):**	0.561
Model:	OLS	**Adj. R-squared (uncentered):**	0.543
Method:	Least Squares	**F-statistic:**	31.95
Date:	Fri, 06 May 2022	**Prob (F-statistic):**	6.97e-06
Time:	19:28:47	**Log-Likelihood:**	-31.527
No. Observations:	26	**AIC:**	65.05
Df Residuals:	25	**BIC:**	66.31
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
unemployment_rate	0.1295	0.023	5.652	0.000	0.082	0.177

Omnibus:	3.933	**Durbin-Watson:**	0.164
Prob(Omnibus):	0.140	**Jarque-Bera (JB):**	3.175
Skew:	-0.852	**Prob(JB):**	0.204
Kurtosis:	2.832	**Cond. No.**	1.00

4. OLS Regression Results Covid-19 Cases growth Rate

OLS Regression Results

Dep. Variable:	home_value_growth_rate	**R-squared (uncentered):**	0.180
Model:	OLS	**Adj. R-squared (uncentered):**	0.141
Method:	Least Squares	**F-statistic:**	4.612
Date:	Fri, 06 May 2022	**Prob (F-statistic):**	0.0436
Time:	19:29:13	**Log-Likelihood:**	-35.139
No. Observations:	22	**AIC:**	72.28
Df Residuals:	21	**BIC:**	73.37
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
covid_growth_rate	1.3975	0.651	2.148	0.044	0.044	2.751

Omnibus:	8.615	**Durbin-Watson:**	0.106
Prob(Omnibus):	0.013	**Jarque-Bera (JB):**	6.235
Skew:	-1.178	**Prob(JB):**	0.0443
Kurtosis:	4.118	**Cond. No.**	1.00

Notes:
[1] R^2 is computed without centering (uncentered) since the model does not contain a constant.
[2] Standard Errors assume that the covariance matrix of the errors is correctly specified.

5. OLS Regression Results Multivariate Regression

OLS Regression Results

Dep. Variable:	home_value_growth_rate	**R-squared (uncentered):**	0.961
Model:	OLS	**Adj. R-squared (uncentered):**	0.952
Method:	Least Squares	**F-statistic:**	110.9
Date:	Fri, 06 May 2022	**Prob (F-statistic):**	2.01e-12
Time:	19:48:15	**Log-Likelihood:**	-1.6366
No. Observations:	22	**AIC:**	11.27
Df Residuals:	18	**BIC:**	15.64
Df Model:	4		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
30_year_mortgage	0.0649	0.219	0.297	0.770	-0.395	0.524
real_disposable_income	0.0001	4.67e-05	2.593	0.018	2.3e-05	0.000
unemployment_rate	-0.1164	0.049	-2.392	0.028	-0.219	-0.014
covid_growth_rate	-0.4645	0.374	-1.242	0.230	-1.250	0.321

Omnibus:	1.911	**Durbin-Watson:**	0.841
Prob(Omnibus):	0.385	**Jarque-Bera (JB):**	1.447
Skew:	0.611	**Prob(JB):**	0.485
Kurtosis:	2.708	**Cond. No.**	9.77e+04