MODELING RENTAL HOUSING PRICES IN THE SEATTLE AREA

By

Lucas Rassbach and Billy Beggs

Washington State University

Economics 490 Capstone Research Paper

April 24, 2020

**Chapter 1: Introduction**

Modeling rental prices is essential for making educated decisions regarding investment, personal purchasing or renting, and policy creation with respect to the real estate market. Being able to forecast how rental prices change will not only help renters, but also investors and policy makers alike. The ability to understand why past price changes have occurred is crucial to making accurate predictions about the future of the real estate market. After the subprime mortgage crisis of 2008, more attention than ever has been placed on the real estate market and its importance to the aggregate economy. People around the country and the globe are now taking a closer look at real estate and what causes its prices to shift because of how widespread the effects of that crisis were.

Everyone at some point will face a housing decision, making this topic one of great importance for all people in modern society. The differences in rental rates around the United States are astounding. For example, a one bedroom apartment in the Seattle market is $1,857 on average which is considered high for the national average, but when compared to San Fransisco’s average rate of $3,483 for the same one bedroom apartment, then Seattle’s rate doesn’t seem high at all. With the rise in income inequality that is being experienced in the United States, it is more important than ever to understand rental markets and creates a need to forecast the prices people pay to rent homes month in and month out.

In this paper we uncover several economic factors that could potentially be used to forecast rental housing prices in Seattle. We explored many explanatory variables for our model including per capita personal income, unemployment rates, consumer price indexes, population growth, permits for new housing and total gross domestic product. Moreover, we found things like building regulations and the availability of land to build new apartment complexes and homes to be important factors in setting rental prices but were not quantitatively feasible to include in our model. Ultimately, our best model settled on per capita personal income, unemployment rate, gross domestic product, population and new permits for housing units as the variables that produced the best forecasting model. Though the other variables mentioned above have some effect on rental housing prices, the ones chosen had quantitative data to support thorough examination and had the most importance in modeling the change in the rental housing price index.

Through the process of arriving at our best model, lots of information was uncovered regarding the housing market and what factors, both economic and geographical, make rental housing prices move. In building our model, we reviewed literature covering topics like housing speculation, or non-owner-occupied home purchases, housing vouchers and their effect on consumer rental decisions, as well as much more. This is an undeniably complex market and must be examined carefully when generating a model that is accurate and thorough.

**Chapter 2: Literature Review**

The first literature that we reviewed was titled “Housing Vouchers and the Price of Rental Housing” by Michael D. Eriksen and Amanda Ross. This paper explores if and how housing subsidies for low income residents affect the price of rentals. This study found that an increase in vouchers did not affect the overall price of rental housing but instead found that recipients were renting more expensive units after receiving the subsidy. This is important for our paper because we are referencing neoclassical consumption theory, which suggests that as personal income increases, people will seek more expensive housing as they look to maximize their utility of consumption. Also, one of our independent variables is per capita personal income and our hypothesis is that an increase in personal income will result in higher rental prices. This contradicts our hypothesis and is a good counterargument, allowing us to examine different perspectives.

Our next literature review examines how strictly land use is regulated in the United States. This paper, titled “The Local Residential Land Use Regulatory Environment Across U.S. Housing Markets: Evidence from a New Wharton Index”, is coauthored by Joseph Gyourko, Jonathan Hartley and Jacob Krimmel. The authors discovered several important items surrounding the regulation of land use in the US. First, they discovered that 9 out of the top 10 markets are located along the northeast coast from Boston down to Washington D.C. or the west coast, being Seattle, Portland, San Fransisco, and Los Angeles. This is important as it relates to our model because we are examining one of the most highly regulated real estate and land development markets in the US. We included data on residential building permits issued in Seattle-Tacoma-Bellevue in an earlier model, but it was omitted because it lowered our r^2 value. The second item of importance in this paper was the discovery that the regulatory environment has not changed since before the Great Recession. The housing bust in 2008 neither led to new regulations nor dismissed older regulations. This is relevant to our paper because we can assume that our data (which goes back to 2010) is not affected by new regulations.

The next literature that we reviewed is a paper titled “Economic Consequences of Housing Speculation” authored by Gao, Sockin and Xiong. The term housing speculation refers to non-owner-occupied homes for vacation or investment. The study found that as housing speculation rates increased, both the boom period of 2004-2006 and the recession period of 2007-2009 were both more severe because of high speculation rates. This is important to our model because we are attempting to forecast rental prices in the Seattle area. These speculators will commonly be landlords looking to rent their properties and this makes this an important area of discussion as speculators are involved in setting the rental rates for their respective properties. Specifically, our variable ‘new permits for housing units’ related directly to this speculation because a lot of those new units are a result of housing speculation by investors looking to build new non-owner-occupied units.

Our final literature review, titled “Urban Growth and it’s Aggregate Implications”, investigates the effect of urban growth on aggregate income, agglomeration benefits and urban costs. The main benefits describe human capital spillover, entrepreneurship, and learning, with the main costs being higher commute times and increased congestion. This is especially relevant to our model because we are including the population of Seattle as an input in our forecast. Population growth has many assorted benefits and costs included and it will be interesting to see the effect it has on our forecast.

**Chapter 3: Research Question and Hypothesis**

The research question of our project is: Can we accurately forecast rental housing prices in Seattle to predict future values using economic variables? Specifically, we narrowed to: Can we model future rental housing prices in Seattle using per capita personal income, unemployment rate, gross domestic product, population, and new permits for housing units as explanatory variables? Neoclassical consumption theory suggests that as personal income increases in an area, people will seek more expensive housing to maximize utility. The same utility maximization is expected to be true for landlords when they lose tenants in a recession causing them to attempt to recover costs by raising rental rates on their remaining tenants. Furthermore, basic supply and demand theory suggest that as the demand rises, indicated by population growth, so will prices and that as supply rises, indicated by new permits for housing units, prices will drop.

**Chapter 4: Methodology**

For our model we chose to use an OLS Regression using the Python Interpreter to run the OLS model. Our code is as follows:

"""

Created on Tue Feb 18 13:02:05 2020

@author: lrass

"""

from statsmodels.stats.outliers\_influence import variance\_inflation\_factor

import matplotlib.pyplot as plt

import numpy as np

import statsmodels.api as sm

import pandas as pd

def main():

file\_name = 'SeaTacBel\_Comp\_Data.csv'

dat = pd.read\_csv(file\_name)

df = pd.DataFrame(dat)

print(df)

x = df[["Rent\_PI\_lag2","Rent\_PI\_lag3","PCPI","PCPI\_lag1","PCPI\_lag2","UER","UER\_lag1","UER\_lag2","NPHU","NPHU\_lag1","NPHU\_lag2","GDP","GDP\_lag1","GDP\_lag2"]]

y = df["Rent\_PI\_dif"]

x = sm.add\_constant(x)

model = sm.OLS(y,x).fit()

predictions = model.predict(x)

summary = model.summary()

print('Dependent variables included:')

for i in x:

print(i)

print(summary)

print('------------------')

adf,pval,usedlag,nobs,cv,icbest = sm.tsa.stattools.adfuller(y, maxlag=None, regression='ct', autolag='AIC', store=False, regresults=False)

print('adf=%f' % adf)

print('pvalue=%f' % pval)

print('num of lags=%d' % usedlag)

print('num of obs=%d' % nobs)

print('crit values=',cv)

print('-----------------')

This code interpreted the csv data and presented an OLS printout sheet as well as produced an Augmented Dickey-Fuller test.

Our model is defined as Y1 = b0 + b1x + b2x +b3x + b4x +b5x. b0 is our constant. b1 represents Per capita Personal Income, measured in dollars. We chose to include this in our model in order to see if increases in the Seattle area’s income can forecast rental prices. b2 represents the unemployment rate in Seattle-Tacoma-Bellevue. Including this variable in our model shows how rental prices’ response to a recession (high unemployment rate) and during economic booms (low unemployment rate) can be forecasted. b3 represents the total gross domestic product for the Seattle-Tacoma-Bellevue area. We included this in our model because GDP is one of the most commonly referenced and measured economic indicators and is relevant to any economic or financial forecast. b4 represents population of the area. We think that population is an important factor to include because as population increases, demand for housing in the area will increase as well. b5 represents the new housing permits issued in Seattle-Tacoma-Bellevue. This is important to our model because an increase in housing permits is a good representation of an increase in the overall supply of housing. Additionally, each variable has had a lag introduced, creating a new variable lagged for one period in order to help forecast future values of the rental price index. We chose to include this in order to have the rate of change for these variables represented in our forecast.

For our research question of ‘Can we accurately model past and future rental prices in Seattle using economic variables?’, these indicators will help forecast effects of recessions, increasing income, and both increased demand and supply on rental prices. This data was sourced from the Federal Reserve Economic Data website <https://fred.stlouisfed.org/>.

A relevant economic theory that motivates the use of this particular model is neoclassical consumption theory. For Per Capita Personal Income, as personal income rises, consumers will want to maximize their utility as allowed by their income. Consumers will move to nicer apartments as their income rises. For unemployment, as an economy heads into a recession, landlords will lose tenants and revenues as people lose their jobs or get their hours cut. The landlords in turn will raise their prices to account for lost revenues.

**Table V** has the descriptive statistics for all of the variables included in our final model.

**Table V**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Variable** | **Count** | **Mean** | **Stdev** | **Min** | **Max** |
| Rent\_PI | 105 | 298.74 | 40.04 | 247.2 | 379.69 |
| Rent\_PI\_dif | 105 | -1.18 | 1.46 | -7.91 | 7.5 |
| Rent\_PI\_lag1 | 105 | 299.92 | 40.6 | 247.2 | 380.24 |
| Rent\_PI\_lag2 | 105 | 301.14 | 41.13 | 247.2 | 382.12 |
| Rent\_PI\_lag3 | 105 | 302.37 | 41.63 | 247.2 | 383.16 |
| PCPI | 105 | 61011.8 | 7954.56 | 48796 | 74620 |
| PCPI\_lag1 | 105 | 60765.9 | 7928.95 | 48796 | 74620 |
| UER | 105 | 5.97 | 1.98 | 3.5 | 10.3 |
| UER\_lag1 | 105 | 6.03 | 2.01 | 3.5 | 10.3 |
| GDP | 105 | 315120.4 | 47620.95 | 242081.9 | 392036.9 |
| GDP\_lag1 | 105 | 313683.4 | 47554.09 | 241152.4 | 392036.9 |
| Pop | 105 | 674823 | 44266.53 | 610639 | 744955 |
| Pop\_lag1 | 105 | 673543.9 | 44160.98 | 610639 | 744955 |
| NPHU | 105 | 1665.53 | 611.42 | 541.08 | 3412.68 |
| NPHU\_lag1 | 105 | 1650.56 | 612.64 | 541.08 | 3412.68 |

**Chapter 5: Results**

Table X below shows the output result of the autoregressive model that was originally chosen, however this changed when we discovered a unit root issue with the Dickey Fuller test. During the process many autoregressive models were tested and compared based on adjusted R 2 values, the Dickey Fuller test and AIC values. Table W is a summary of the models ran originally and the criteria which we evaluated them upon.

**Table W**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **DepVars Included** | **Lags** | **Adj R^2** | **Aug D-F** | **ADF p-value** | **AIC** |
| Rent\_PI lag1 | Rent\_PI: 1, OtherDepVars: 0 | 0.999 | -2.539051 | 0.308721 | 362.9 |
| Rent\_PI lag1, Rent\_PI lag2 | Rent\_PI: 2, OtherDepVars: 0 | 0.999 | -2.539051 | 0.308721 | 364.4 |
| Rent\_PI lag1, Rent\_PI lag2, Rent\_PI lag3 | Rent\_PI: 3, OtherDepVars: 0 | 0.999 | -2.539051 | 0.308721 | 364.9 |
| Rent\_PI lag1, Rent\_PI lag2, Rent\_PI lag3, PCPI, UER, GDP, Pop, NPHU | Rent\_PI: 3, OtherDepVars: 0 | 0.999 | -2.539051 | 0.308721 | 361.5 |
| Rent\_PI lag1, Rent\_PI lag2, Rent\_PI lag3, PCPI, PCPI lag1, UER, UER lag1, GDP, GDP lag1, Pop, Pop lag1, NPHU, NPHU lag1 | Rent\_PI: 3, OtherDepVars: 1 | 0.999 | -2.539051 | 0.308721 | 360.3 |
| Rent\_PI lag1, Rent\_PI lag2, Rent\_PI lag3, PCPI, PCPI lag1, PCPI lag2, UER, UER lag1, UER lag2, GDP, GDP lag1, GDP lag2, Pop, Pop lag1, Pop lag2, NPHU, NPHU lag1, NPHU lag2 | Rent\_PI: 3, OtherDepVars: 2 | 0.999 | -2.539051 | 0.308721 | 364.5 |
| PCPI, PCPI lag1, PCPI lag2, UER, UER lag1, UER lag2, GDP, GDP lag1, GDP lag2, Pop, Pop lag1, Pop lag2, NPHU, NPHU lag1, NPHU lag2 | Rent\_PI: 0, OtherDepVars: 2 | 0.994 | -2.539051 | 0.308721 | 557.0 |

As shown in Table W, very similar adjusted R 2 values were produced for all the models. All but one of the models produced a value of 0.999 and the last model, which excluded lags on Rent\_PI, produced a value of 0.994. This narrowed our choices down to just the models with Rent\_PI lags in them. Likewise, the Dickey Fuller test produced the same value and p-value for each of the models. This left just AIC to evaluate the models upon. With AIC evaluation, the lowest value indicated the model that performed best in the AIC test. This led us to the decision that the model with 360.4 as the AIC value being the best model for prediction. It should be noted that BIC is a similar test that penalizes more for the degrees of freedom present, or complexity of the model, meaning that BIC focuses more on the model's closeness with the true model for the data. Since this project is aimed at prediction, we chose to focus on AIC because it is more useful in prediction where complexity of the model does not play as large of a role. The equation for this model is as follows: Rent\_PI = 0.8696 Rent\_PI\_lag1 + 0.1086 Rent\_PI\_lag2 + 0.0168 Rent\_PI\_lag3 + 0.00007144 PCPI + 0.00004058 PCPI\_lag1 – 0.45 UER + 0.9304 UER\_lag1 – 0.0005 GDP + 0.0004 GDP\_lag1 + 0.00003378 Pop + 0.00004529 Pop\_lag1 + 0.0004 NPHU + 0.0004 NPHU\_lag1.

**Table X**

Dependent variables included:

const

Rent\_PI\_lag1

Rent\_PI\_lag2

Rent\_PI\_lag3

PCPI

PCPI\_lag1

UER

UER\_lag1

GDP

GDP\_lag1

Pop

Pop\_lag1

NPHU

NPHU\_lag1

OLS Regression Results

==============================================================================

Dep. Variable: Rent\_PI R-squared: 0.999

Model: OLS Adj. R-squared: 0.999

Method: Least Squares F-statistic: 8008.

Date: Tue, 21 Apr 2020 Prob (F-statistic): 4.43e-133

Time: 14:04:44 Log-Likelihood: -166.17

No. Observations: 105 AIC: 360.3

Df Residuals: 91 BIC: 397.5

Df Model: 13

Covariance Type: nonrobust

================================================================================

coef std err t P>|t| [0.025 0.975]

--------------------------------------------------------------------------------

const -35.4048 13.989 -2.531 0.013 -63.193 -7.617

Rent\_PI\_lag1 0.8696 0.103 8.405 0.000 0.664 1.075

Rent\_PI\_lag2 0.1086 0.136 0.796 0.428 -0.162 0.379

Rent\_PI\_lag3 0.0168 0.106 0.159 0.874 -0.193 0.226

PCPI 7.144e-05 0.000 0.282 0.779 -0.000 0.001

PCPI\_lag1 4.058e-05 0.000 0.151 0.880 -0.000 0.001

UER -0.4500 0.454 -0.991 0.324 -1.352 0.452

UER\_lag1 0.9304 0.445 2.091 0.039 0.047 1.814

GDP -0.0005 0.000 -1.746 0.084 -0.001 6.93e-05

GDP\_lag1 0.0004 0.000 1.323 0.189 -0.000 0.001

Pop 3.378e-05 4.9e-05 0.690 0.492 -6.35e-05 0.000

Pop\_lag1 4.529e-05 5.32e-05 0.851 0.397 -6.05e-05 0.000

NPHU 0.0004 0.000 0.987 0.326 -0.000 0.001

NPHU\_lag1 0.0004 0.000 1.203 0.232 -0.000 0.001

==============================================================================

Omnibus: 41.896 Durbin-Watson: 2.002

Prob(Omnibus): 0.000 Jarque-Bera (JB): 1405.589

Skew: 0.121 Prob(JB): 6.03e-306

Kurtosis: 20.923 Cond. No. 1.20e+08

==============================================================================

Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

[2] The condition number is large, 1.2e+08. This might indicate that there are

strong multicollinearity or other numerical problems.

------------------

adf=-2.539051

pvalue=0.308721

num of lags=6

num of obs=98

crit values= {'1%': -4.054251125423931, '5%': -3.4562790670553936, '10%': -3.153866135708761}

-----------------

Upon further evaluation of the Augmented Dickey-Fuller test results, we concluded that the p-value of 0.308721 indicated we had a unit root issue in these models. This led to us exploring models with the difference in Rent\_PI from the last period to the current period as our new dependent variable. This produced a much better result in the Augmented Dickey-Fuller tests and gave values close to 0.000 but reduced our adjusted R 2 values greatly. This is actually a good thing, because the presence of the unit root issue inflates adjusted R 2 values and we believe that is what was occurring in our previous models. With this new dependent variable that fixed the unit root issue, our focus returned to maximizing the adjusted R 2 as our main criteria for model selection. Below is Table Y with the selection criteria values from the models run with the new dependent variable of Rent\_PI – Rent\_PI\_lag1.

**Table Y**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **DepVars Included** | **Lags** | **Adj R^2** | **Aug D-F** | **ADF p-value** | **AIC** |
| Rent\_PI lag2, Rent\_PI\_lag3, PCPI, UER, NPHU, GDP | Rent\_PI: 2, OtherDepVars: 0 | 0.207 | -6.471378 | 0.000000 | 360.2 |
| Rent\_PI lag2, Rent\_PI lag3, PCPI, PCPI\_lag1, UER, UER\_lag1, NPHU, NPHU\_lag1, GDP, GDP\_lag1 | Rent\_PI: 2, OtherDepVars: 1 | 0.233 | -6.471378 | 0.000000 | 360.3 |
| Rent\_PI lag2, Rent\_PI lag3, PCPI, PCPI\_lag1, PCPI\_lag2, UER, UER\_lag1, UER\_lag2, NPHU, NPHU\_lag1, NPHU\_lag2, GDP, GDP\_lag1, GDP\_lag2 | Rent\_PI: 3, OtherDepVars: 0 | 0.210 | -6.471378 | 0.000000 | 366.9 |

Based on the outputs from the models with the new dependent variable, the best model created was the middle one with an adjusted R 2 value of 0.233. This leaves us with the equation for our final model as: Rent\_PI – Rent\_PI\_lag1 = 0.0151 Rent\_PI\_lag2 + 0.0132 Rent\_PI\_lag3 + 0.00009762 PCPI + 0.0003 PCPI\_lag1 – 0.4705 UER + 0.8090 UER\_lag1 + 0.0002 NPHU + 0.0004 NPHU\_lag1 – 0.0002 GDP + 0.00006247 GDP\_lag1

The regression output for our final model is located below in Table Z.

**Table Z**

Dependent variables included:

const

Rent\_PI\_lag2

Rent\_PI\_lag3

PCPI

PCPI\_lag1

UER

UER\_lag1

NPHU

NPHU\_lag1

GDP

GDP\_lag1

OLS Regression Results

==============================================================================

Dep. Variable: Rent\_PI\_dif R-squared: 0.307

Model: OLS Adj. R-squared: 0.233

Method: Least Squares F-statistic: 4.159

Date: Thu, 23 Apr 2020 Prob (F-statistic): 8.97e-05

Time: 15:12:32 Log-Likelihood: -169.15

No. Observations: 105 AIC: 360.3

Df Residuals: 94 BIC: 389.5

Df Model: 10

Covariance Type: nonrobust

================================================================================

coef std err t P>|t| [0.025 0.975]

--------------------------------------------------------------------------------

const -6.5212 4.354 -1.498 0.138 -15.167 2.125

Rent\_PI\_lag2 0.0151 0.100 0.151 0.880 -0.183 0.213

Rent\_PI\_lag3 0.0132 0.106 0.124 0.902 -0.198 0.225

PCPI 9.762e-05 0.000 0.617 0.539 -0.000 0.000

PCPI\_lag1 0.0003 0.000 1.613 0.110 -5.94e-05 0.001

UER -0.4705 0.451 -1.042 0.300 -1.367 0.426

UER\_lag1 0.8090 0.442 1.829 0.071 -0.069 1.687

NPHU 0.0002 0.000 0.633 0.528 -0.000 0.001

NPHU\_lag1 0.0004 0.000 1.245 0.216 -0.000 0.001

GDP -0.0002 0.000 -0.623 0.535 -0.001 0.000

GDP\_lag1 6.247e-05 0.000 0.230 0.819 -0.000 0.001

==============================================================================

Omnibus: 44.488 Durbin-Watson: 2.126

Prob(Omnibus): 0.000 Jarque-Bera (JB): 1610.211

Skew: 0.292 Prob(JB): 0.00

Kurtosis: 22.176 Cond. No. 1.60e+07

==============================================================================

Warnings:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

[2] The condition number is large, 1.6e+07. This might indicate that there are

strong multicollinearity or other numerical problems.

------------------

adf=-6.471378

pvalue=0.000000

num of lags=5

num of obs=99

crit values= {'1%': -4.053254236405479, '5%': -3.455806184392646, '10%': -3.1535907061122397}

-----------------

**Chapter 6: Discussion**

Our original model aimed to accurately forecast future values of the rental prices in Seattle, represented by the Rent\_PI dependent variable. In predictive models the goodness of fit, indicated by adjusted R 2  , is the most important indicator. The R 2 value of 0.999 implies that our original model explains 99.9% of the variation in the Rental Price Index dataset. This high value suggests that with previous data from the Rental Price Index, using lags, and our five independent variables, our model maximizes the accuracy when predicting the rental price index value of the future period given these variables. However, after running Augmented Dickey-Fuller tests, we chose to instead use our final model that included all variables, each with one lag, and used the rate of change of the rent price index as the dependent variable.

As testing is concerned, the VIF values and P-values present in our model are not important because the model is one of predictive nature. We chose to instead include the Augmented Dickey-Fuller test. The Augmented Dickey-Fuller test (ADF) tests the null hypothesis that the dependent variable’s data set is a unit root, or that the lagged variable has a coefficient of one. The ADF returned a p-value of .309 across all models, indicating that there is a 30.9% chance that the Rent\_PI variable is a unit root. This indicates that our data is non-stationary, or that statistical values will change over time. With this knowledge, we can assume that the models we tested with R2 values of 0.999 are inflated due to this unit root issue. Our final model’s ADF test returned a p-value of 0.000, ensuring that the unit root issue was not present and that the R2 values are accurate.

Even though our new model produced adjusted R 2 values far lower than our values in previous models, it is a much more accurate representation of the ability of our model to predict future values. The elimination of the unit root issue and inflated R 2 values makes this model much more realistic. However, because the new value is so low it will likely not be as accurate a prediction as the previous models indicated.

Changing our dependent variable to the difference between the current rent price index and the rent price index of the previous period addressed the unit root issue and our final model has an R2 value of 0.223. While this is much lower than the previous model, there is no question of its integrity on the basis of using stationary vs. non-stationary data. This is not a bad thing, but eludes to the conclusion that this model should not be considered in large financial decisions or policy making decisions. The suggestion of this paper is that there should be additional research on the topic.

**Chapter 7: Conclusion**

The final model created displays a forecast using lags on Rent\_PI, as well as our independent variables, per capita personal income, unemployment rate, gross domestic product, population and new permits for housing units. Our dependent variable is the difference in the previous period’s rental price index and the current period’s index. Our model uses data from personal income, unemployment rate, gross domestic product and population, as well as these variables in a 1 period lag format, and the rent price index in multiple lagged formats in order to forecast the change in the rental price index. All the independent variables were chosen to forecast the change in future rental price index values based on neoclassical consumption theory as well as supply and demand theory. Consumption theory pertains to this model because we are assuming that as consumers increase their wealth, they will also increase their utility of consumption. The supply and demand theory pertains as well because as supply of housing increases, we expect to see a drop in predicted rental prices, while as population increases, we expect to see an increase in demand that would increase rental prices. Since the model has such a mediocre adjusted R 2 value, we can conclude that it will provide some kind of prediction about future movements in the rental price index, given past data for our independent variables. We recommend that this model be used for educational purposes only and that further research be done on the topic.

**Bibliography**

BLS Data Viewer. (n.d.). Retrieved from <https://beta.bls.gov/dataViewer/view;jsessionid=516F7555E54D52366A700E69FF2D1DDF>

Bronshtein, A. (2019, December 27). Simple and Multiple Linear Regression in Python. Retrieved from <https://towardsdatascience.com/simple-and-multiple-linear-regression-in-python-c928425168f9>

Brownlee, J. (2019, October 10). Probabilistic Model Selection with AIC, BIC, and MDL. Retrieved from https://machinelearningmastery.com/probabilistic-model-selection-measures/

Consumer Price Index for All Urban Consumers: Rent of Primary Residence in Seattle-Tacoma-Bellevue WA (CBSA). (2020, February 13). Retrieved from <https://fred.stlouisfed.org/series/CUURA423SEHA>

Eriksen, Michael D., and Amanda Ross. (2015). Housing Vouchers and the Price of Rental Housing. American Economic Journal: Economic Policy, 7 (3): 154-76.

Gao, Z., Sockin, M., & Xiong, W. (2019). Economic Consequences of Housing Speculation. NBER Working Paper Series. Retrieved from <https://www.nber.org/papers/w26457.pdf>

Gyourko, J., Hartley, J., & Krimmel, J. (2019). The Local Residential Land Use Regulatory Environment Across U.S. Housing Markets: Evidence from a New Wharton Index. doi: 10.3386/w26573

McNamara, N. (2019, May 23). Seattle Grew By About 15,000 People In 2018: Census Data. Retrieved March 26, 2020, from https://patch.com/washington/seattle/seattle-grew-15-000-people-2018-census-data

Megbolugbe, I., Marks, A., & Schwartz, M. (1991). The Economic Theory of Housing Demand: A Critical Review. *The Journal of Real Estate Research,* *6*(3), 381-393. Retrieved March 27, 2020, from www.jstor.org/stable/44095553

Per Capita Personal Income in Seattle-Tacoma-Bellevue, WA (MSA). (2019, November 20). Retrieved from <https://fred.stlouisfed.org/series/SEAT653PCPI>

Prabhakaran, S. (2019, November 2). Augmented Dickey-Fuller (ADF) Test - Must Read Guide - ML. Retrieved from https://www.machinelearningplus.com/time-series/augmented-dickey-fuller-test/

Regression Plots. (n.d.). Retrieved from <https://www.statsmodels.org/dev/examples/notebooks/generated/regression_plots.html>

Shanelynn. (2019, October 6). Summarising, Aggregating, and Grouping data in Python Pandas. Retrieved from https://www.shanelynn.ie/summarising-aggregation-and-grouping-data-in-python-pandas/

Tavares, E. (n.d.). Variance Inflation Factor (VIF) Explained. Retrieved from <https://etav.github.io/python/vif_factor_python.html>

Total Gross Domestic Product for Seattle-Tacoma-Bellevue, WA (MSA). (2019, December 12). Retrieved from <https://fred.stlouisfed.org/series/NGMP42660>

Unemployment Rate in Seattle-Tacoma-Bellevue, WA (MSA). (2020, February 5). Retrieved from <https://fred.stlouisfed.org/series/SEAT653URN>