How External Factors Affect the Pittsburgh Housing Market

Group 7

Nate Stevens, Hannah Harhai, Daniel Lucash

COMPUTATION IN INFORMATION SCI, University of Pittsburgh

***Abstract*— This project seeks to determine how or if external factors affect the Pittsburgh housing market. We will look into a variety of demographic metrics as well as how national events/crises can play an effect on the market.**

I. Introduction

The purpose of our project is to investigate housing prices in the City of Pittsburgh over the last 20 years. We plan on seeing if there is any correlation between housing prices and other factors, such as income, unemployment rates, poverty levels, and national crises. Housing prices and how they change is very important to the City of Pittsburgh's economy, since increased or decreased prices can alienate some people from living in the city limits. It is also important to do this research with the hopes of being able to predict when the housing market becomes unstable, so that the city government can intervene and mitigate the damage. At the time of researching this project, housing prices in Pittsburgh are at an all time high and we'd like to compare the recent changes in the market with changes that have happened in the past.

II. Project Goals

The main goal of our project is to find if there is a correlation between any form of overarching factors and the housing market of Pittsburgh, and with this information we want to show how the market has changed over a large period of time, stating possible reasons for each part. Another goal of ours is to use this info to make a theoretical prediction for the near future of the housing market, of whether it's going to increase or decrease and why that would happen based on our data gathered and how the understanding of current events plays a factor into that.

III. Project Methods

A. Methods and Tools

The main tools we will be using in this project consist of RStudio and its various libraries. Within RStudio we will utilize libraries for data importing, data manipulation, and data display. More specifically, we plan to use R's statistical and various graphing packages. We plan to use data from the Federal Reserve Bank of St. Louis which analyzes the quarterly average of Pittsburgh home prices since 2000, as well as data on the unemployment rates, annual personal income, and poverty levels. We plan to utilize ggplot and ggplot2 to create plots of data for how our individual demographic metrics have changed over time. We will then use scatter plots to evaluate the relationship between our metrics and housing prices. We will use linear regression to graphically measure the relationship as well as calculating correlation coefficients to statistically measure the relationship. In summary, our method will be to import our selected data sets, manipulate the data for our use, create plots, compile data together, and perform statistical calculations to represent all of the data we analyzed.

B. Miscellaneous Variables

While researching for this project, it was clear we were going to run into some variables that would cause issues within our data due to the fact that the housing market is extremely complex and requires multiple things to be working correctly in order for it to stay stable. These miscellaneous variables can

range from multiple recessions to a pandemic, and here we will seek to explain how they have affected the housing market. Before we even look at the data that affected our chart real-time, we need to look back at the 1940's. It was during this time that homeownership began to skyrocket thanks to The Servicemen's Readjustment Act of 1944, also known as the GI Bill of Rights, which gave many benefits to veterans, including a lower mortgage rate.

In the early 1970s, Freddie Mac was created to help fix the expanding mortgage industry, and Freddie Mac does this by buying mortgage loans, then packages them to be sold off to investors. With this, mortgages were more freely available, meaning more people could buy a home. Then soon after the Home Mortgage Disclosure Act of 1975 was created to help this cause even more.

Cut to early 2000s, the housing market experienced a huge boom due to the growing normality of subprime mortgages, and thanks to this people who had poor credit histories or other financial issues could now afford a loan, and now that the lenders were selling more loans than they had available, they started offering riskier loans instead, which were designed to be unaffordable later on down the line.

Then in 2008, this cycle of risk loaning/subprime lending and scooping them back up finally collapsed, causing the Foreclosure rate to skyrocket from 717,522 in 2006 to 2,330,483 in 2008. Thanks to this, the economy went into recession and housing stagnated until around 2010. Moving into the 2010's we see housing slowly creep up, because even though the housing crash lost trillions of dollars and millions of people had to foreclose their home, advertisements and idealism created the idea that in America the best idea was to own a home of your own, meaning that once the markets left stagnation, people wanted to get back in on purchasing a home.

Later on in 2020, Covid-19 struck. Thanks to that, interest rates collapsed, and once lockdown ended, people decided to start buying homes farther from cities. This along with the wave of newly-capable home buyers meant that the market had a bounceback with extremely competitive conditions.

Another factor that we need to pay attention to when considering the outcome of our data is the issue of demographics. Based on this, many factors for the housing market can change. The ways that demographics can affect this is by the fact that if there's more people looking to move to Pittsburgh or people looking to move out of Pittsburgh can cause the price of buying a house in the area to increase or decrease. Along with this, depending on what kind of buyers the market is currently placating to, the housing market will have to deal with a change in scenario. For example, if there's more people looking to buy apartments than houses, the prices for housing may decrease, and the apartment prices may increase. Alongside this, if the buyers are looking to rent instead of owning a home, this can also change how the housing market is affected, since there may be less houses on the market to buy, making the prices of those houses rise.

The final variable to point out that could change the housing market for our data is the government policies / subsidies. This includes taxes, subsidies, and such. An example of how this could affect the market is from 2009, where after the crash the US Government instituted the first-time homebuyer's tax, which was an incentive program to help revitalize the housing market. This may decrease the price of housing outside of our variables, giving us a slightly skewed data set during that time frame until 2010.

IV. Results/Data

We first decided to create a scatterplot of the average housing price in Pittsburgh as this is the data we want to compare our various metrics against. We then created scatterplots for each of the metrics including, unemployment rate, per capita personal income, and the amount of people in poverty. All these metrics were evaluated over the same time interval starting in 2000. With the use of scatter plots we were able to visualize the trends in data over the past two decades. We then compared the average housing price data to unemployment rate, per capita personal income, and people in poverty to see what kind of relationship, if any, exists between each of the variables. Plots of the data are all shown below in figures along with more detailed descriptions and results for each metric and analysis.

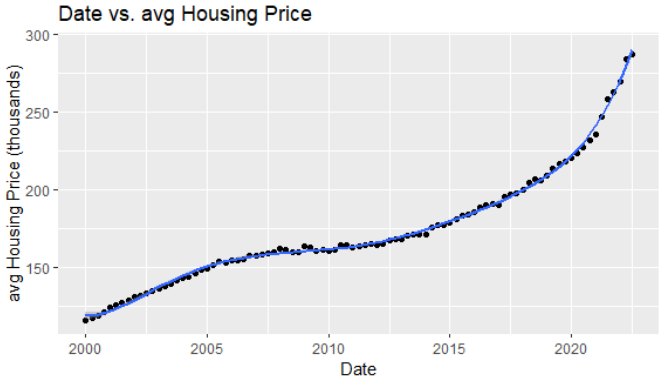

Fig. 1. Date vs. avg Housing Price

Figure 1 shows us how the average housing price in Pittsburgh has changed over the past two decades. We can visually see how the average price has continually increased since the year 2000. This is the data we will be comparing for our external factors below.

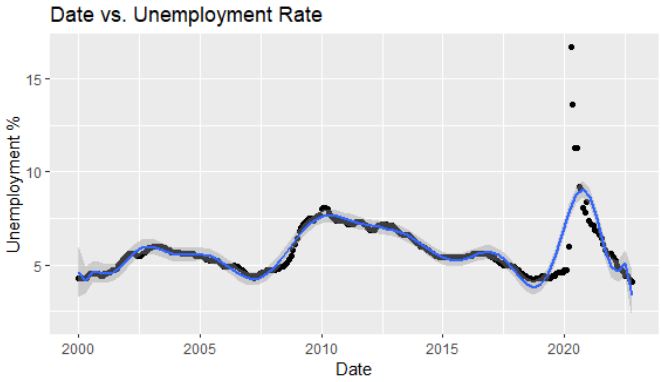

Fig. 2. Date vs. Unemployment Rate

Fig. 3. Unemployment Rate vs. Housing Price

The first metric we thought might affect the housing market was the unemployment rate. Again, we used data provided by the Federal Reserve Bank of St. Louis [2]. Looking at Figure 2, we can see the trend of unemployment slowly declining until a sudden rise, later going back to repeat the cycle. Several factors can be attributed to this, including but not limited to: large changes in the market like Covid and the Great Recession, as well as changes in unemployment policy following Presidential elections. However, this graph does not trend well when comparing it to the average housing price (Figure 1). This becomes more obvious when plotting both unemployment rates and average housing price as a function of time. The resulting graph, Figure 3, as we can see has no clear or predictable trend. Therefore, we can conclude that unemployment rates in Pittsburgh have no effect on the average price of a house.

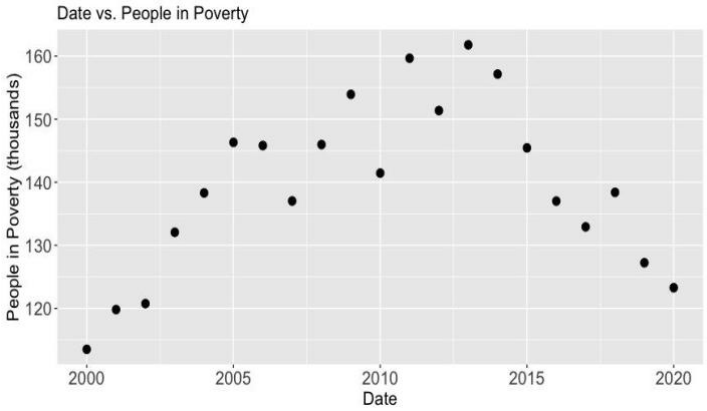

Fig. 4. Date vs. People in Poverty

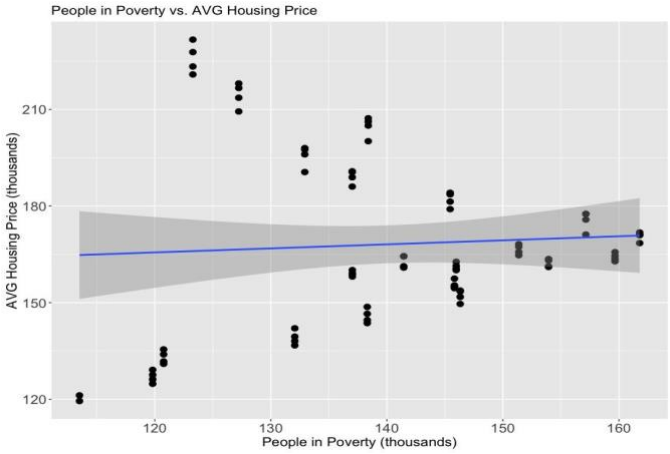

Fig. 5. People in Poverty vs. Housing Price

Next, we wanted to see if the number of people in poverty had any correlation with the average housing price in Pittsburgh. Poverty is measured by "The United States Census Bureau that uses a set of money

income thresholds that vary by family size and composition to determine who is in poverty. If a family's total income is less than the family's threshold, then that family and every individual in it is considered in poverty. The official poverty thresholds do not vary geographically, but they are updated for inflation using the Consumer Price Index (CPI-U). The official poverty definition uses money income before taxes and does not include capital gains or noncash benefits (such as public housing, Medicaid, and food stamps)" [10]. The data of people in poverty examines all people in Allegheny County as there was no dataset exclusive to just the City of Pittsburgh. However, Allegheny County comprises the City of Pittsburgh as well as nearby surrounding areas, so we concluded that the dataset would still be representative.

Here in Figure 4 we can see how the number of people in poverty has changed over the past two decades. The plot steadily increases from the year 2000 to around 2012-13, then from there it begins to decline towards the year 2020. Figure 5 displays the relationship between the number of people in poverty and the average housing price in Pittsburgh across the time interval of 2000-2020. The blue line across the plot is the linear regression line to model the relationship between the two variables. By simply examining the graph you can tell that the regression line indicates a very slight positive relationship. The visibly small slope of the regression line also indicates a weak relationship between the two variables of interest. To further statistically measure the strength of the relationship, we calculated the correlation coefficient in R which yielded 0.0603. Since the correlation coefficient is so close to zero it further concludes that there is very little to no relationship between the number of people in poverty and the average housing price in Pittsburgh.

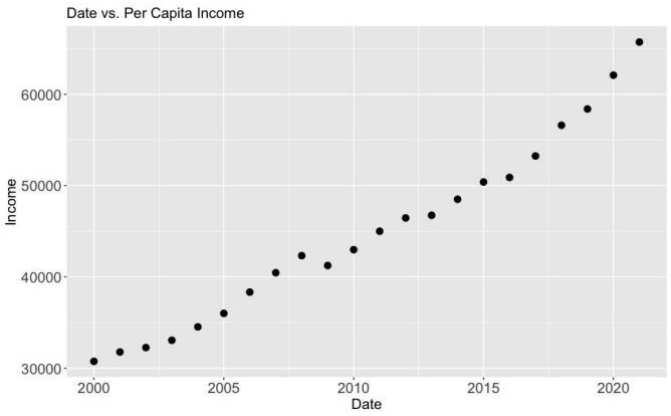

Fig. 6. Per Capita Income vs. Date

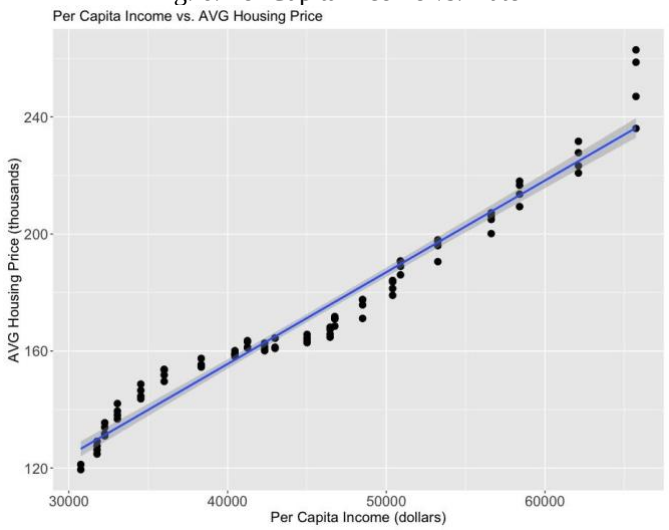

Fig. 7. Per Capita Income vs. Housing Price

Our last factor we wanted to examine was per capita income. We rounded up data from the past 20 years and from Figure 6 it can be seen how the per capita income has continually increased since the year 2000. When selecting our factors we believed this one would most likely have the strongest correlation/relationship to housing prices. This was further proven when we examined the relationship between the two variables graphically in Figure 7. As shown by the graph per capita income and housing price have a strong positive linear relationship. Statistically, when calculating the correlation coefficient it yielded 0.9775 which is very close to 1. Further, meaning there is a strong relationship/correlation between the two variables. This makes sense because as incomes go up over the years so do markets/prices.

IV. Conclusion

In conclusion, the external factors, unemployment, per capita income, and poverty rates were the ones we thought would have the biggest impact on the housing market. As seen through the data, unemployment and poverty rates did not have much of a relationship on the housing market at all. However, per capita income and average housing prices have continually increased over the past two decades so it's no surprise that they have a strong correlation. We believe the housing prices will continue to go up as long as per capita income levels also increase. If someone wanted to expand upon this data it would be interesting to see if or how any other external factors/variables play a role on the market.

References

[1] Federal Reserve Bank of St. Louis, *All-Transactions House Price Index for Pittsburgh, PA (MSA)*, US Federal Housing Finance Agency, 1976-2002.

[2] Federal Reserve Bank of St. Louis, *Unemployment Rate in Pittsburgh, PA (MSA)*, US Federal Housing Finance Agency, 1976-2002.

[3] Federal Reserve Bank of St. Louis, *Per Capita Personal Income in Pittsburgh, PA (MSA)*, US Federal Housing Finance Agency, 1976-2002.

[4] Federal Reserve Bank of St. Louis, Resident Population in Pittsburgh, US Federal Housing Finance Agency, 1976-2002.

[5] Pew Research Center, Foreclosures in the U.S. In 2008, May 12, 2009

[6] United States Census Bureau, Zillow and Census Bureau Data Show Pandemic's Impact on Housing Market, October 4, 2021

[7] The Washington Post, 10 years later: How the housing market has changed since the crash, October 4th, 2018

[8] Federal Reserve Bank of St. Louis, Estimate of People of All Ages in Poverty in Allegheny County, PA, US Federal Housing Finance Agency, 1976-2002.

[9] Investopedia, 4 Key Factors That Drive the Real Estate Market, December 5th, 2021

[10] United States Census Bureau, How the Census Bureau Measures Poverty, September 22, 2022