

## Econ 388 Data Assignment 3

### *COVID-19 Mortality Rates and Housing Prices: An Empirical Analysis*

#### **Introduction**

COVID-19 has swept the world into a series of lockdowns, shortages, and crises. The impact of the virus has been felt by every household from 2020 and to this day. The effects of this pandemic extend from the workplace to the home. This analysis attempts to determine the effect of county-level per-capita COVID-19 mortality rates in 2020 on housing prices in 2021. This is an important research question because if higher mortality rates impact housing prices, then political officials need to pay attention to the impact of these events. Finding a relationship between deaths from the illness and the economy is a significant finding that can impact future policies in real estate and guide future research on both short-term and long-term effects on the housing market.

#### **Data Sources**

##### *New York Times*

The first step in collecting the data for this analysis was to find COVID-19 mortality rates by county for the year 2020. The New York Times has compiled easy access to this data that can be downloaded from github and includes daily death rates and FIPS codes which are unique identifiers for each county in the United States. After the mortality rates by county were collected, the analysis needed population data to convert the mortality rates to a per-capita basis.

##### *U.S. Census*

The next data file collected for this analysis was from the U.S. Census of 2020. This data was accessed as a table from the U.S. Census website but had to be modified to include population at a county level. After some small modifications, this data was ready to merge with the other files based on the unique identifier of the FIPS code. Now that the mortality rates were intact, the data needed housing price information.

##### *FHFA*

The FHFA website provided the data used to track the housing price index changes from 2020 to 2021. The problem with this data is that it uses the three-digit zip code system to identify the housing price index. This required another data set to connect the three-digit zip codes to the FIPS codes.

##### *U.S. Department of Housing*

The U.S. Department of Housing has a website that provides conversion for zip codes to county FIPS codes based on a selected time period. This is useful because there are several zip codes in one county, and it allows the zip-level data to be aggregated on a county-level. By

converting five-digit zip codes to three-digit zip codes, the housing price index information was able to merge with this data and subsequently with the other data sets based on the FIPS codes. This completed the base data set, but a couple items were added to the analysis from one more source.

### *ERS*

To allow for controls of changes in the unemployment rate and median household income by county, the Economic Research Service website was used to download data. This data included employment statistics for the last 20 years and the current median household income, all on a county-level. This additional source allows for further analysis on the question introduced.

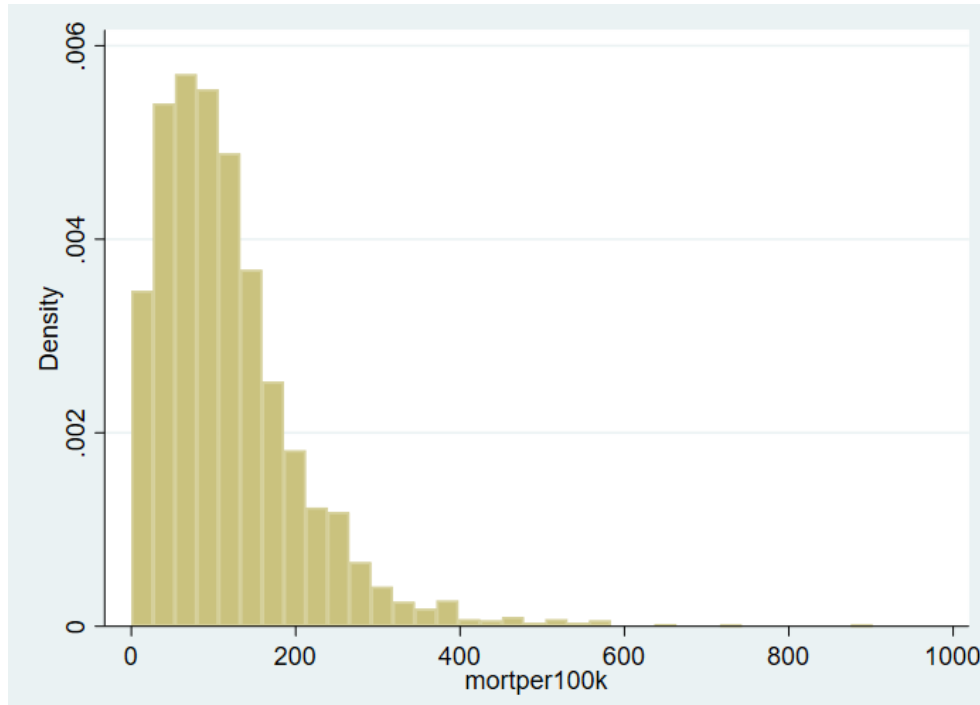
### **Data Summary**

	<b>Average</b>	<b>Minimum</b>	<b>Maximum</b>
<i>HPI Change</i>	7.053	-4.119	18.978
<i>Median Household Income</i>	\$57,623	\$22,901	\$160,305
<i>Unemployment Rate Change</i>	48.04%	-32.79%	225.00%
<i>Mortality Rate per 100K</i>	118	0	902

The table above shows summary statistics for the variables analyzed in the data including the percentage change in the housing price index (HPI) from 2020 to 2021, the median household income for 2021, the unemployment rate change from 2020 to 2021, and the mortality rate per 100,000 people due to COVID-19. The summary statistics provide some interesting information.

1. The housing price index, on average, increased for counties at a rate of 7 points.
2. Most counties saw an increase in unemployment from 2020 to 2021.
3. The mortality rate for COVID-19 had a minimum of 0 deaths for a county in 2020 and this occurred in 89 counties.
4. The maximum mortality rate of 902 deaths per 100,000 people occurred in Jerauld County, South Dakota.

In addition to a table summary, a histogram is provided to summarize the data. The histogram below highlights the key variable of mortality rates per 100,000 people across counties. This highlights the mean of 118 deaths with the minimum of 0 and the maximum of 902.



## Empirical Analysis

The individual factor of COVID-19 mortality rates cannot account alone for the change in the housing price index from 2020 to 2021. For this reason, median household income and the change in unemployment rates have been included as variables to help control for other factors that have contributed to the change in the housing market. An ordinary least squares regression model has been with the independent variables mentioned to determine the strength of the relationship. The specific regression model used in the empirical analysis is as follows;

$$\text{housechange}_i = \beta_0 + \beta_1 \text{mortper100k}_i + \beta_2 \text{medianincome}_i + \beta_s \text{unempchange}_i + u_i$$

...where *housechange* is the change in the housing price index in points from 2020 to 2021, *mortper100k* is the mortality rate of COVID-19 deaths in the year 2020 for every 100,000 people, *medianincome* is the median household income, and *unempchange* is the percentage change in the unemployment rate from 2020 to 2021.

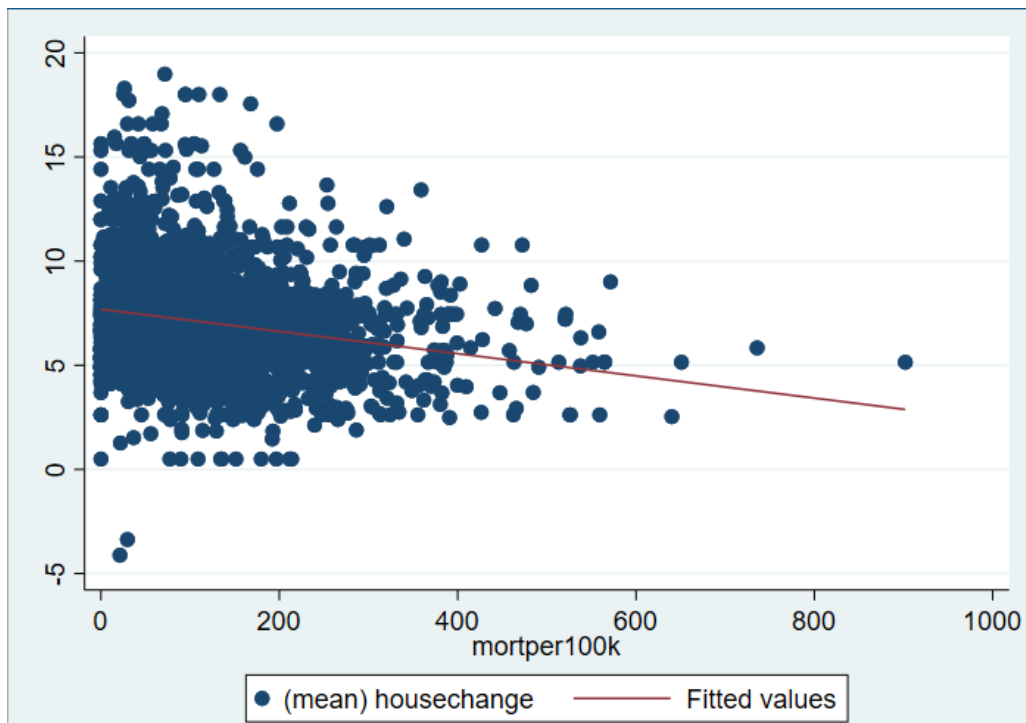
This model relies on a couple of assumptions. First, it assumes that the median household income controls for the effect of wealth difference across counties. Second, it assumes that the unemployment rate change controls economic factors of changing employment levels that may affect housing. Due to limited data and time, change in median household income was not able to be used as a measure to control for the differences in time. Additionally, median household income and unemployment rate change are not all-inclusive controls; more factors could be added such as inflation, median population age, and changes in median household income.

Running the analysis described above produces the following results:

Variable	Coefficient	Std. Error	t-score	P> t
<i>mortper100k</i>	-0.005	0.0004	-11.09	0.000
<i>medianincome</i>	-9.00e-06	2.94e-06	-3.06	0.002
<i>unempchange</i>	0.802	0.157	5.11	0.000
<i>_cons</i>	7.822	0.202	38.64	0.000
<b>Observations:</b> 3,129				<b>R-squared:</b> 0.051

The coefficients listed in the table show the ceteris paribus effect of each independent variable on the change in HPI from 2020 to 2021. The variable *mortper100k* shows that one death decreases the HPI change by 0.005 points holding all else constant. The variable *medianincome* shows that a one dollar increase in median income decreases the HPI change by 0.000009 points holding all else constant. The variable *unempchange* shows that a 100% increase in the unemployment rate over a year increases the HPI change by 0.8 points holding all else constant.

The analysis of the p-value shows that all results are statistically significant at the 10%, 5%, and 1% levels. The t-scores represent the same conclusion. It should be noted that the statistical significance is not the same as economic significance. A more in-depth analysis could provide insight on if the magnitude of these variables is important. Given the statistical (and likely economic) significance of all variables included, correlation could be breaching causality. The regression analysis shows that more COVID-19 deaths per capita likely leads to a lower HPI change. Below is a fitted line showing the relationship between *housechange* and *mortper100k*.



## Conclusion

In conclusion, there is likely a causal relationship between HPI changes and COVID-19 mortality rates controlling for household income and unemployment rate changes. Limitations to the data include lack of controlling variables such as inflation, median population age, and changes in median household income. Further analysis should be done to analyze the short-term and long-term effects of the pandemic on the housing market, the impacts of COVID-19 on HPI changes on a global scale, and policies that could be put in place as a result of these findings. This empirical analysis serves as a start for changing the future.