COVID, Unemployment and Labor Force in Bernalillo County

Jack Chen

Introduction

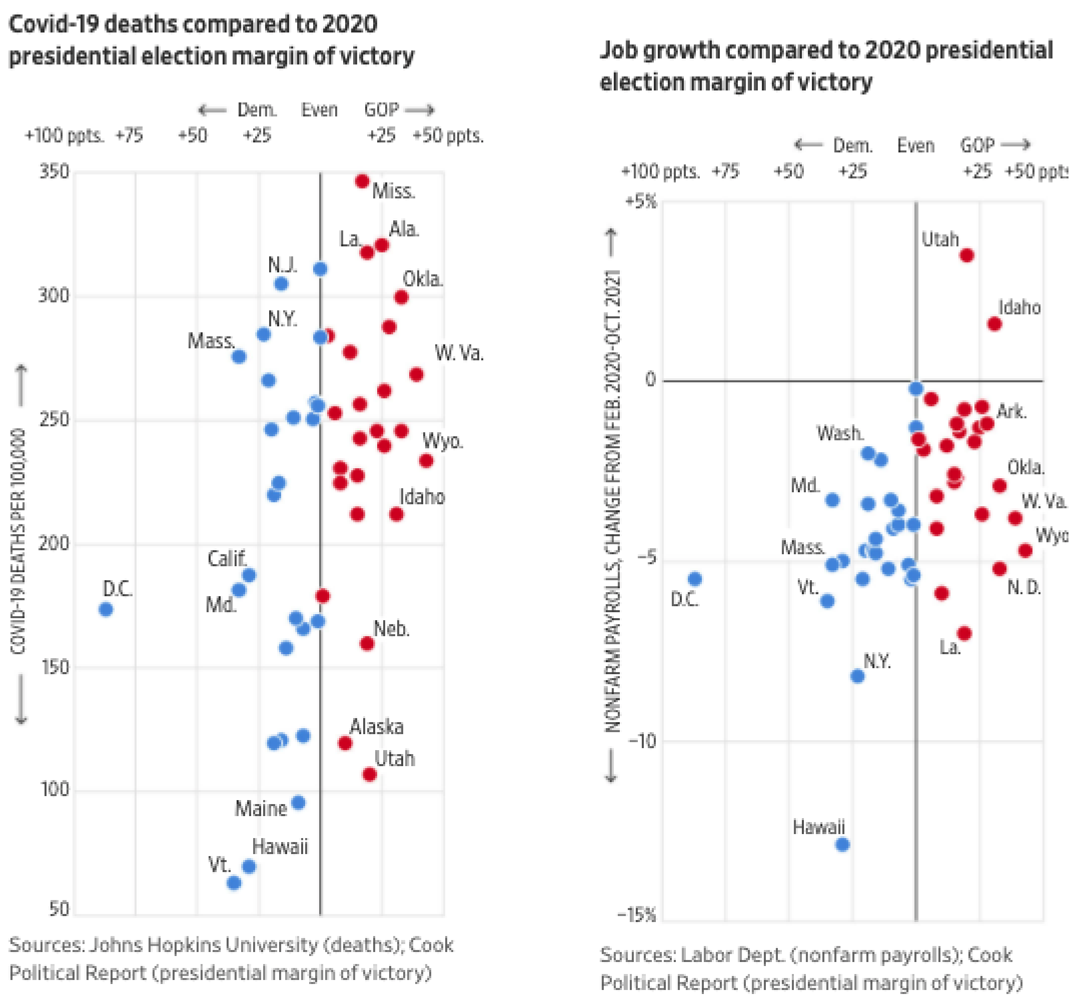
COVID-19 is a black swan event that put a pause on our everyday life. In the past 22 months, we have experienced countless unprecedented events. Countries across the globe are reacting to covid differently. Asian countries have "zero tolerance" policies and are willing to close their countries in exchange for a lower disease transmission rate. Western countries are attempting to learn to live with covid, and "back-to-normal" is the goal for liberal and conservative administrations.

Employment statistics are under the scrutiny of news reports and that is not surprising, considering the rapid changes in employment at the beginning of the pandemic and now in the midst of the Great Resignation. One example of this change is that the labor force participation rate nationally has dropped from 63.4% to 61.6%, according to the US Bureau of Labor Statistics. For Bernalillo county, the unemployment rate increased from 4.5% at the beginning of 2020 to 13%. The unemployment rate has since dropped gradually and now sits at 5.4%.

It might seem obvious that COVID has caused unemployment and dropping out of the labor force. However, can we use statistics to confirm that? This is the main question that I will be examining in the paper. Another area I explored is whether masking policies lowered covid transmission.

Background/Related Work

A Wall Street Journal article pointed out that COVID-19 has carved a partisan divide through the US. Democratic-leaning states have gone further in protective measures but paid for it with much weaker job recovery. The initial message that the only way to "reopen the economy" is to stop the spread of the coronavirus has been proven false. Infections and economic activity appear to be inversely related.



(graphs from WSJ article)

Several researchers used Granger Causality tests to determine the relationship between GDP, unemployment and inflation. This research paper noted that in Japan and the UK, GDP granger causes unemployment. In the US, there is a two-way relationship between unemployment and GDP, i.e. unemployment granger causes GDP drop and vice versa. Granger Causality test is one of the most popular methods for time series analysis and it will also be the tool I use for this project.

Methodology

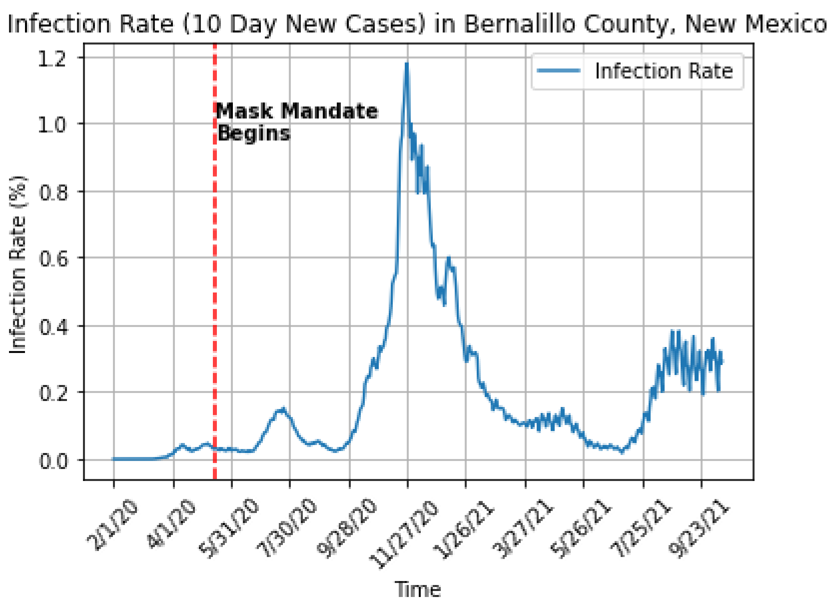
One of the most important theories in statistics is that correlation does not imply causation. Without empirical experiment, it is difficult to establish causation. Granger Causality test is the closest method to establish a causal relationship in statistics with observational data.

According to [Wikipedia](), Granger Causality test is a statistical hypothesis test for determining whether one time series is useful for forecasting another. “A time series X is said to Granger-cause Y if it can be shown, usually through a series of t-tests and F-tests on lagged values of X (and with lagged values of Y also included), that those X values provide statistically significant information about future values of Y.”

I chose the Granger Causality test as it is a useful and common hypothesis testing method in economics and time series analysis. One advantage to using the Granger Causality test is that it is not known to generate ethical concerns. However, Granger Causality tests present their findings in terms of p values and the interpretation of p values is often erroneous in the scientific realm. Alas, all models are wrong but some are useful. To quote Warren Buffett, it is better to be approximately right than precisely wrong. Approximately right is the aim for my research.

Findings

Masking Effectiveness



This graph shows the covid infection rate for Bernalillo County, New Mexico from 2/1/2020 to 10/15/2021. In Bernalillo County, the mask mandate started on May 15th, 2020 and is still in effect as of today. The x axis is time and the y axis is the infection rate. Infection rate is defined as the number of infections over the number of those at risk, according to [Wikipedia]().

In my analysis, I assumed that new covid patients would have a 10-day recovery window, during which they could spread the disease. This assumption is naïve, as patients could already be spreading covid prior to their diagnosis. However, given the data limitations, I believe this assumption to start the recovery window at diagnosis is a fair one.

My numerator of the infection rate is calculated as the sum of new cases in the past 10 days, including the day of. This also effectively smooths the data, and takes away day of the week fluctuations.

For the number of those at risk, ideally I would use the number of close contacts. However, with the given data, I use the unaffected population as a proxy. Assuming that people who have already gotten covid would have immunity for the duration of the analysis, we would calculate the denominator of the infection rate as the entire population minus all the confirmed cases. As the denominator did not fluctuate significantly, the infection rate curve looks similar to the new cases curve.

Additionally, the compliance of the mask mandate is fairly high in Bernalillo County, with ~90% of the population frequently or always wearing masks, according to a New York Times dataset.

Looking at my graph, I do not believe I have enough data to conclude if the mask mandate is effective. Infection rate, without actually using the number of close contacts, closely tracks the new cases rate. After the mask mandate started, the rate of infection increased drastically, and the first derivative of the infection rate also increased, as we can see from the graph. I caution that there are many confounding variables, including seasonality, that increase the slope of the infection rate. Also, without the mask mandate, the infection rate could be growing exponentially. One can argue that absent the mask mandate the infection rate would be significantly higher than what we have observed. From my analysis, I cannot reach a conclusion on the effectiveness of the mask mandate.

Causation

	Cause: Labor Force	Cause: Unemployment	Cause: Infection Rate
Response: Labor Force	1.0	0.0	0.15
Response: Unemployment	0.0	1.0	0.61
Response: Infection Rate	0.0	0.01	1.0

The table above represents the p values of pairwise Granger Causality tests.

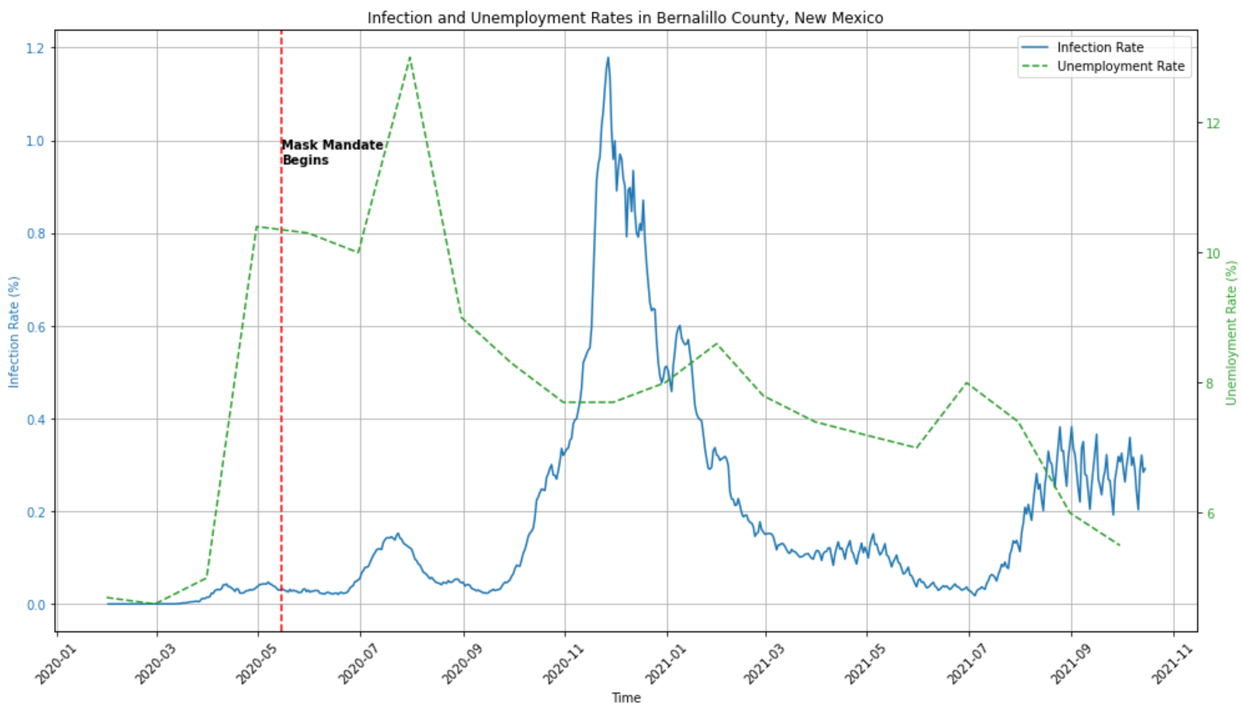
The p values in the yellow boxes suggests that we do not have enough evidence to reject the null hypothesis that covid infection rate does not granger cause higher unemployment or labor force reduction. In other words, I cannot state that infection rate granger causes unemployment rate spikes.

The p values in the red box shows that unemployment rate and labor force granger cause covid infection rate. Having the historical unemployment statistics, according to the p value, can allow us to predict the future covid infection rate. That is a very counterintuitive finding that I will discuss in the next section.

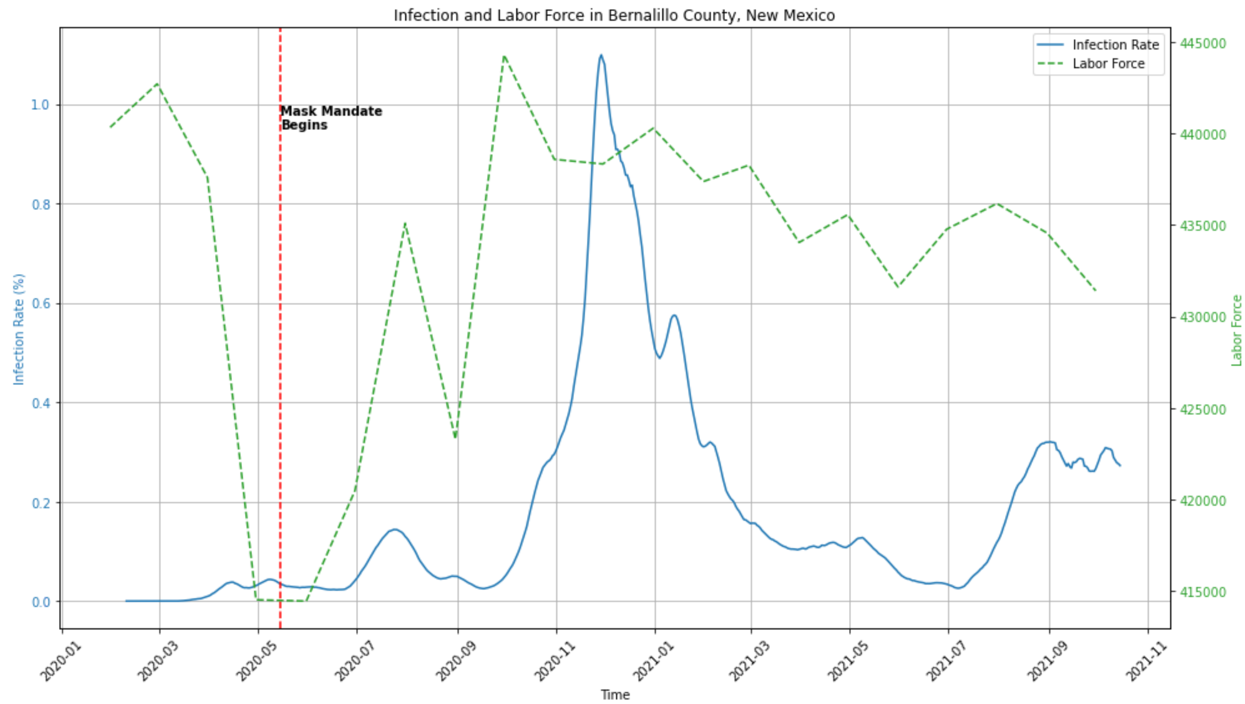
The Granger Causality test implied that unemployment rate and the labor force granger caused infection rate, while the reverse is not true.

Discussion/Implications

My finding with the relationship between covid infection rate and unemployment as well as the labor force is counterintuitive. If covid infection rate indeed was caused by unemployment and people dropping out of the labor force, we can reduce covid transmission by providing everyone a job. The CDC would be glad that we found a cure for COVID-19. This phenomenon can be explained by the fact that the unemployment and labor force trends precede the covid infection rate, insteading of lagging.



The infection and unemployment rate graph shows that the unemployment rate peaked in summer 2020 and decreased since. However, covid infection rate did not peak until the winter last year. We can tell from the graph that the movement in the unemployment data might have preceded the movement in covid infection rate. That also holds true for the labor force trend, depicted below.



I believe that the unemployment peak is caused by the stay-at-home order and limited testing capabilities at the beginning of covid transmission in the US. When the service industry shut down at the onset of stay-at-home order, restaurants, airlines and many non-essential businesses sent their workers home. The *expectation* of covid transmission caused the unemployment rate. As the infection rate increased last winter, government policies and the outgoing administration chose to focus on economic recovery, rather than disease control. “Back to normal” is the goal for both the Trump and the Biden administrations, in their own ways. Therefore, unemployment did not increase while or shortly before covid infection skyrocketed, largely due to the lax covid government policies.

Limitations

My research is not without limitations and caveats. The first limitation is the imprecision of determining the number of close contacts (explained in the findings section) and the infrequency of economic data. Unemployment rate and labor force participation are measured monthly, and since the beginning of covid, it has been less than 22 months. Therefore, to conduct this analysis, I had only 20 months of economic data available for the analysis. The data scarcity also begets parameter tuning difficulties for Granger Causality test. As the test looks at the number of lag periods, different numbers of lagged periods of X can produce different results. The maximum lag period my model can utilize is 5 months, as any more months would not leave enough periods for the hypothesis tests to render test results. This limitation is the main bottleneck for my analysis.

Another caveat is the model assumption. One of the Granger Causality test's assumptions is that the datasets need to be stationary. As the covid infection rate isn't stationary, I utilized the change in infection rate from month to month. This would fulfill the assumption of stationarity. However, I would caution that the population's ideological shift and social political issues might come up and the infection rate pattern can persist for a period of time, which would skew the results.

My study also uses the number of unemployed people as a proxy for the unemployment rate. It makes sense to treat the difference of number of the unemployed for stationarity than the rate itself. Furthermore, the unemployment rate is defined as the number of unemployed people divided by the labor force. Theoretically, the labor force can decrease significantly due to residents moving out of the county, or because of people dropping out of the workforce for covid related reasons. However, the labor force decreased by only 2.0% since January 2020 in Bernalillo county. Therefore, the denominator of the unemployment rate did not change significantly and I believe that it is valid to use the unemployment rate as a proxy.

Conclusion

I do not have enough data to conclude if the mask mandate is effective and that is due to missing the number of close contacts. Additionally, this paper does not find any indication that the covid infection rate granger caused a labor force exodus or unemployment spikes. On the contrary, the study shows that the unemployment rate and the labor force granger caused the covid infection rate. That might be counterintuitive. However, due to the polarization of political views, the US observes an inverse relationship of economy and covid recovery. Unemployment peaked at the beginning of covid but the lax covid response is the compromise to achieve economic recovery. Labor force fluctuations happened before significant covid infection spikes. Since the summer of 2020, the fluctuations in infection rate did not affect employment status drastically in Bernalillo county, NM.

References

1. US Bureau of Labor Statistics Labor Force Participation: https://www.bls.gov/charts/employment-situation/civilian-labor-force-participation-rate.htm
2. WSJ Omicron Could Widen Blue-Red Economic Divide: https://www.wsj.com/articles/omicron-could-widen-blue-red-economic-divide-11638367200
3. The Unemployment Rate Amid the COVID-19 Pandemic: Propose the Best Practices Policy to Maintain Labor Market Stability: https://jurnal.ugm.ac.id/jsp/article/view/56450
4. Granger Causality Test Wikipedia: https://en.wikipedia.org/wiki/Granger_causality
5. Wikipedia Infection Rate: https://en.wikipedia.org/wiki/Infection_rate

Data Sources

1. Kaggle repository of John Hopkins University COVID-19 data: https://www.kaggle.com/antgoldbloom/covid19-data-from-john-hopkins-university?select=RAW_us_confirmed_cases.csv
2. CDC Masking Mandate by State: https://data.cdc.gov/Policy-Surveillance/U-S-State-and-Territorial-Public-Mask-Mandates-Fro/62d6-pm5i
3. The New York Times mask compliance survey data: https://github.com/nytimes/covid-19-data/tree/master/mask-use
4. US Bureau of Labor Statistics: https://data.bls.gov/timeseries/LAUMT351074000000004?amp%253bdata_tool=XGtable&output_view=data&include_graphs=true
5. Granger Causality python code: https://rishi-a.github.io/2020/05/25/granger-causality.html