The Effects of COVID-19 Related State-Level Restrictions on National and State Unemployment

Chrystine Tadler, Clayton Perry, Richard Griffiths[1]
[1]mGTI Statistics, Columbia, MD

Abstract
The COVID-19 pandemic has affected the American population in numerous ways beyond specific health outcomes. One of the largest impacts is on the rate of unemployment after many businesses and corporations were unexpectedly forced to close or restructure. This research investigates a relationship between unemployment levels, State-level COVID-19 related travel or mobility restrictions, and the severity of the pandemic over time through the use of the ACS 2019 1-year estimates, the Census Household Pulse Survey, and State-level COVID-19 restriction and case data. A customizable multi-level model is presented which can simulate a variety of restriction levels to illustrate the effects on unemployment levels. We also introduce a visualization tool that allows users to explore the relationships between restrictions and unemployment levels.

Key Words: Interactive Visualization, Unemployment, COVID-19 Restrictions

1. Summary and Study Objectives

This research studied the effect of the coronavirus pandemic on unemployment rates. In particular, we were not only interested in the effect of the virus, but the associated effect of state-imposed activity restrictions on unemployment and the labor force. We constructed two different forms of regression models to examine the effect of COVID-19 case counts and state restrictions on unemployment (UE) counts and participation in the civilian labor force (CLF) which were combined to construct an estimated unemployment rate (UER).

We explored these relationships using publicly-available data and developed a dashboard visualization with the capability to illustrate these relationships. The modeling and visualization, in turn, provide a knowledge base and opportunity for simulation that could be used to plot a strategy for future pandemics, including COVID-19 variants.

To structure our research, we developed a set of study objectives and supporting research questions. The study objectives were threefold:

- Use existing data to explore relationships between the pandemic, demographics, COVID-19 case counts and UERs through modeling;
- Build a visualization tool to easily compare results;
- Simulate alternate state-imposed restriction implementations to understand their effect on the UER and how this could affect planning strategies for future pandemics.

Two sets of research questions were established to support the study objectives and further guide our examination. The first set related to our first study objective and accessible existing data:

- How did the patterns of the virus and the state-imposed restrictions affect COVID-19 case counts and the UER?
    - Specifically, we focused on unemployment that was directly related to the pandemic. For example, unemployment because of temporary workplace closure.
    - While not a primary focus, we did investigate any effects by demographics.

The second set of research questions related to our third study objective, the simulation of alternate scenarios:

- What might have been the effects on UERs and COVID-19 case counts if state restrictions had been implemented differently?
- For future pandemics or variants, what can we learn about how these state-imposed restrictions can be best implemented?
- What is the best strategy for managing state-imposed restrictions with respect to containment of the virus and the employment situation?

Of particular interest through this research is the determination of an optimal strategy for managing state-imposed restrictions in relation to public health and the economy. We hoped to develop a tool that, through simulation, identifies the level of state-imposed restrictions that give the best results in terms of containment of the virus and the employment situation.

2. Methodology

The first component of our methodology was to predict UER due to COVID-19 related reasons. Using the Census Bureau's Household Pulse Survey microdata, we modeled the incidence of UE for pandemic-related reasons from daily COVID-19 case counts, state-imposed restrictions, and demographic information. Using the Census Bureau's Current Population Survey microdata, we modeled participation in the CLF from daily COVID-19 case counts, state-imposed restrictions, and demographics. Both models were fit using logistic regression. The modeled UE and CLF estimates were used to calculate a pandemic-related UER.

Because of variations in daily reporting of new COVID-19 cases, most media sources relied on seven-day averages. For similar reasons, we chose to smooth the daily case counts using a generalized linear model that modeled the COVID-19 case counts as a function of time and the impact of state-imposed restrictions. The smoothed COVID-19 case counts were subsequently used in the logistic regression models for UE and CLF participation.

To structure the model for COVID-19 case counts, we needed to make an assumption about the underlying cyclical nature of the pandemic. It was difficult to determine how many waves of the COVID-19 pandemic occurred by the summer of 2021, since there had been a number of interventions that could have disrupted its natural cycle (Maragakis, 2020).

The effect of policy restrictions on new cases certainly affected the form of the pandemic curve over time, so what looked like a plateau may only be a temporary dropoff in new cases owing to a set of policy restrictions. Additionally, there was assuredly an effect from the vaccination process.

Sen et al. (2021) suggest there were three COVID-19 waves in 2020, with peaks spaced at approximately four months. This agrees in broad terms with the 1918 Spanish flu pandemic cycle. It is generally accepted there were three waves of the Spanish Flu pandemic, which occurred approximately four months apart. (See Barry, 2017 and CDC, 2019.) Overlaying that structure on the COVID-19 pandemic, we hypothesized that, without policy restrictions and health guidelines, there would have been COVID-19 case count peaks in July/August of 2020 and again in December 2020. A third peak in April of 2021 would then be expected had policy restrictions and the vaccination campaign not intervened.

2.1 Data Sources

To develop our models for UE, CLF and COVID-19 case counts, we used the following public data sources:

- 2019 ACS 1-year estimates Data: This data provided state population totals and was used in calculating proportions of the population reported as COVID-19 cases.
- Monthly CPS Data: This microdata provided monthly employment status and was used in modeling changes in the CLF.
- Census Household Pulse Survey: This microdata provided unemployment status for reasons related to the COVID-19 pandemic. Question 13 of the Household Pulse Survey asks respondents to provide a reason that they were not working, if they were off work. Several response options specified COVID-19 related reasons. This data was used in the logistic regression modeling of pandemic-related unemployment.
- Johns Hopkins University COVID-19 Data Repository: This data provided daily, state-level estimates of reported COVID-19 cases and was used to develop smoothed estimates of the counts of daily and monthly COVID-19 cases. These estimates were used as independent variables in models for pandemic-related UE and CLF.
- Oxford COVID-19 Government Response Tracker: This data provided timing and intensity level of state-imposed restrictions and was used as independent variables in modeling pandemic-related UE, changes in the CLF, and COVID-19 case counts.

2.2 Model Prototypes

The statistical models presented below are prototypes of a set of models that, with further refinement, should be capable of assisting researchers understand how variations in COVID-19 cases and state-imposed restrictions impact the CLF and UE.

The first model constructed estimates of the number of people unemployed for reasons related to the pandemic using Household Pulse Survey microdata. The dependent variable in the model was an indicator of whether a sample person was unemployed for COVID-19 related reasons. The independent variables were COVID-19 case counts and demographics including gender, age, education level and race. The logistic regression model had the following form:

$$g(I_{msi}) = \beta_{0s} + \alpha_{s1} \cdot C_{ms} + \alpha_{s2} \cdot EG_{msi} + \alpha_{s3} \cdot EA_{msi} + \alpha_{s4} \cdot EE_{msi} + \alpha_{s5} \cdot ER_{msi}$$

where

- I_{msi}: indicator of Covid-19 related unemployment (RSNNOWORK in (8,9,10,11)), for sample respondent *i* in state s, month *m*
- C_{ms}: estimated new COVID-19 case count for month *m* in state *s*
- EG_{msi}: gender for sample individual *i* in state s, month *m*
- EA_{msi}: age for sample individual *i* in state s, month *m*
- EE_{msi}: education level for sample individual *i* in state s, month *m*
- ER_{msi}: race for sample individual *i* in state s, month *m*

The second model used a similar logistic regression form and estimated the number of people in the CLF using CPS microdata. The dependent variable was an indicator of whether a person was in the CLF while independent variables were COVID-19 case counts and demographics including gender, age, education, and race. This logistic regression model had the form:

$$g(CLF_{msi}) = \beta_{0s} + \alpha_{s1} \cdot C_{ms} + \alpha_{s2} \cdot EG_{msi} + \alpha_{s3} \cdot EA_{msi} + \alpha_{s4} \cdot EE_{msi} + \alpha_{s5} \cdot ER_{msi}$$

where

- CLF_{msi}: indicator of whether CPS sample individual *i* was in the labor force in survey month *m*, state *s*

The pandemic-related UER was calculated as the number of people unemployed for reasons related to the pandemic divided by the number of people in the civilian labor force.

Using a gamma-based generalized linear model, we smoothed the estimated COVID-19 case counts prior to use in the previously defined models. In this model, the daily proportion of a state's population reported as new COVID-19 cases was the dependent variable:

$$p_{ds} = \frac{C_{ds}}{N_s}$$

where

- C_{ds}: the daily COVID-19 case counts from the Johns Hopkins University COVID-19 Data Repository, state *s* and day *d*
- N_s: the population of state *s* from the 2019 ACS 1-year estimates

The independent variables consisted of indicators for the various state-imposed restrictions and their respective intensity levels. These indicators were captured at the daily level and the ultimate gamma-based generalized linear model incorporated these indicators, a day enumeration, and the shape and pattern of the 1918 Spanish Flu pandemic through a sinusoidal function. The smoothed COVID-19 case counts were aggregated to the monthly level and fed back into the UE and CLF models as independent variables.

3. Model Results

Unemployment rates were calculated from the modeled UE and CLF population estimates for each month of our study timeframe. Figure 1 displays these rates and COVID-19 case counts for the state of Virginia under the actual state-imposed restriction implementations. UERs were high early in the pandemic, but decreased substantially throughout the lifetime of the pandemic with slight increases around the time of COVID-19 case count spikes during winter months. This trend was also seen in other states. It is hypothesized that this initially high UER might have been due to newly implemented restrictions, but that the existence of restrictions over time helped to keep the UERs low. Generally, our estimated UERs are higher than expected, which is discussed further in the limitations section.

Two alternate scenarios were looked at in the simulation. These showed a consistent story about the ultimate effect of the restrictions on the unemployment rate. This is discussed more in the Simulation section below.

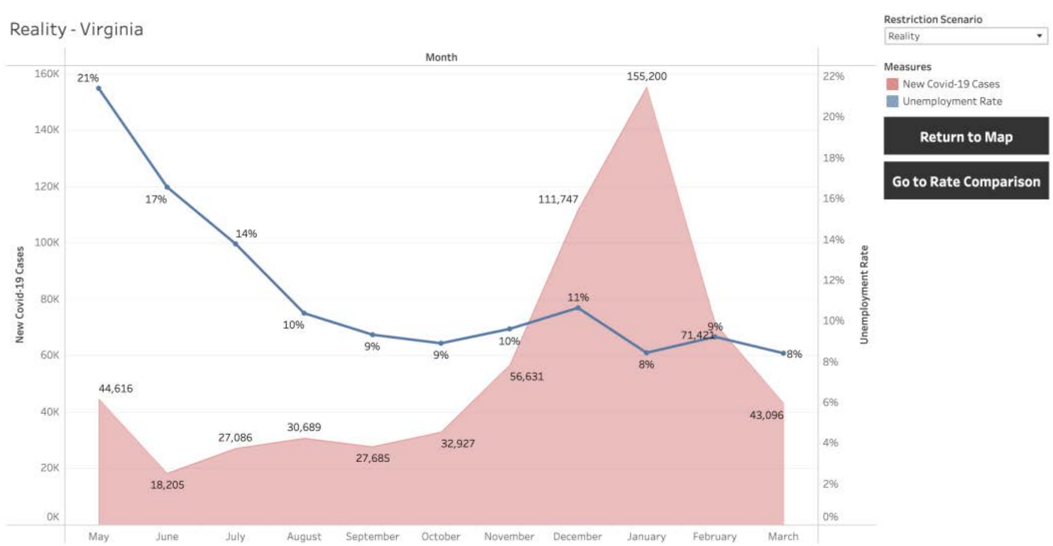

Figure 1: Unemployment Rates and COVID-19 case counts - Virginia

One of the documented effects of the pandemic has been the differential impact on UE and CLF participation by gender. The Household Pulse Survey and CPS data show evidence to support this theme. Our logistic regression models showed that the pandemic and state-imposed restrictions had a much greater effect on both women's participation in the labor force and on their UER. Figure 2 illustrates this with a graph of the estimated odds of unemployment, males versus females, from our unemployment logistic regression model.

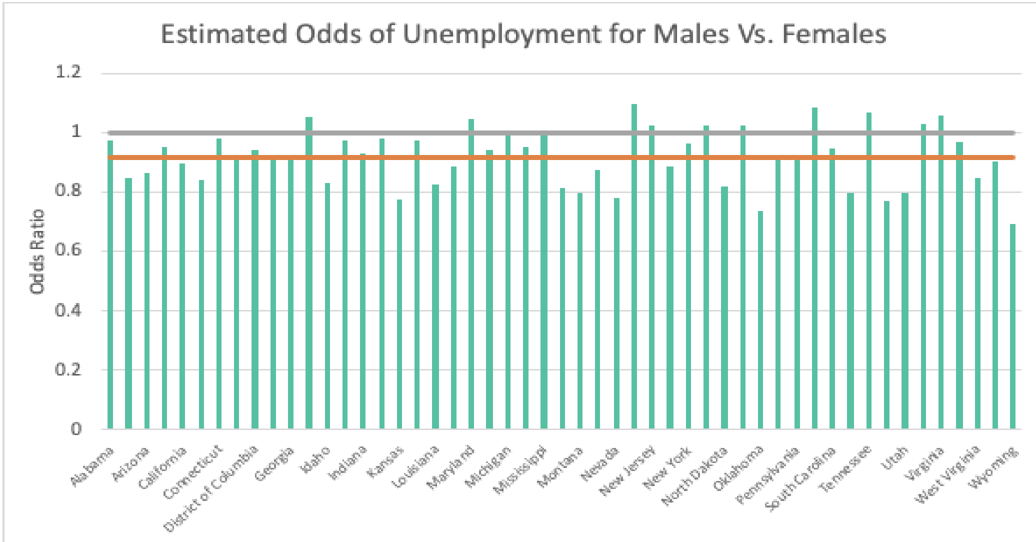

Figure 2: Odds of Unemployment by Gender

4. Visualization

The second component of our methodology was to build a visualization tool to display the relationships found in the models. The visualization allows users to see and better understand the association among COVID-19 case counts, time, and state-imposed restrictions, while accounting for demographics. Beyond this, the tool also allowed us to take a first look at a method of simulating hypothetical alternate implementation scenarios. The primary objective in developing this tool was to provide an easy way to generate a knowledge base that could be applied to other pandemics and COVID-19 variants.

We used Tableau to design our interactive dashboard. When accessing the tool, the user is initially presented with a choropleth map of COVID-19 related UERs by state under the true implementation scenario. See Figure 3. Users have the ability to progress through each month of our analysis timeframe as the visualization updates the unemployment rates. When users hover over a particular state, the unemployment rate is displayed.

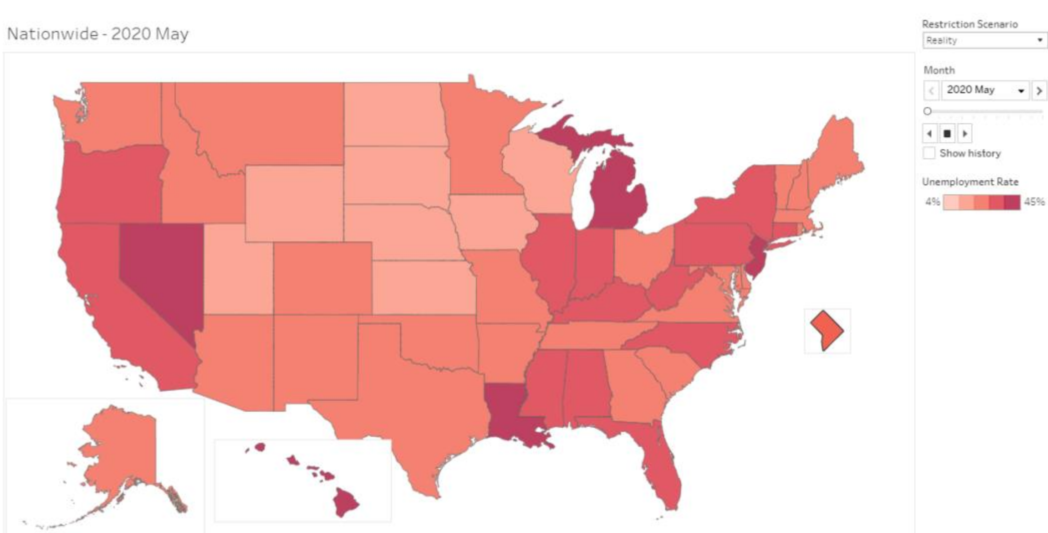

Figure 3: Unemployment Rates by State

From the choropleth map, users can click directly on a state and be directed to the second part of our visualization: a drill-down page that provides users with more detail for the chosen state. This page juxtaposes employment statistics and COVID-19 case counts. See Figure 4.

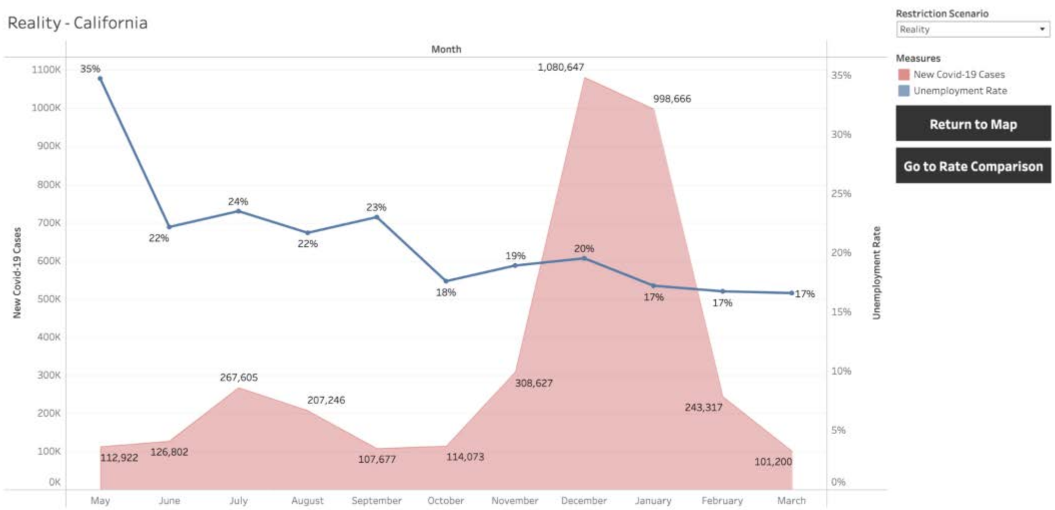

Figure 4: State Drill-down Page: California - Reality

With this page of the visualization, users have the ability to change the visualization, using the drop-down menus, to display output based on three different simulated scenarios. In Figure 3, modeled state unemployment rate estimates over time are super-imposed on new COVID-19 case counts for the state of California.

5. Simulation

Our third study objective focused on simulating the effects on UER and COVID-19 case counts had state-imposed restrictions been implemented differently. To do this, we re-purposed our models to predict UE, CLF, and COVID-19 case counts using different state-imposed restriction data which represented different implementation scenarios. For this research, we simulated two hypothetical alternate implementation scenarios:

- An alternative scenario in which no state restrictions were implemented. Under this scenario, the model supplied estimates of the COVID-19 case counts and UERs assuming no state-imposed restrictions were ever implemented.
- An alternative scenario in which all states implemented restrictions in a unified and consistently incremental pattern: no restrictions initially and, as time progressed, all restrictions increased in intensity level each month. This pattern continued until reaching the maximum intensity level for a given restriction and maintained this level through the end of the analysis time period.

These simulations were incorporated into the visualization tool to easily allow users to visualize and compare the consequences of different restriction strategies against that of reality.

Figure 5 gives an example of our first alternate scenario. Under this scenario for the state of New York, in which no restrictions were implemented, predicted COVID-19 case counts peak over the winter months and remain at fairly constant levels outside of the peak. Our models suggested in this scenario that the UER remained at an elevated level throughout.

Figure 5: State Drill-down Page: New York – Alternate Scenario 1

An example of the simulation under the second alternate scenario appears in Figure 6. In this scenario, the estimated COVID-19 case counts begin in about the same range as the first alternate scenario but drop to lower levels during the summer months. In the winter months, a peak in new COVID-19 case counts is experienced, similar to the first scenario and a subsequent drop follows. The pandemic-related UER in this scenario did not differ greatly from reality. Initially, the UER was high with the start of restriction implementation, but eventually decreased to levels lower than those of the first scenario.

Figure 6: State Drill-down Page: New York – Alternate Scenario 2

The final, current capability of our visualization is a comparison of COVID-19 related UERs across each of our simulated restriction implementation scenarios. Users have the chance to see the estimated UER across all months for each alternate scenario compared to that of reality. In Figure 7, we display the comparison for the state of Pennsylvania. The UERs were initially high in the two scenarios where restrictions were implemented due to restrictions and closures but by fall 2020, the UER for the scenario in which no restrictions were ever implemented was running higher than the two scenarios with restrictions (including reality). A straightforward interpretation of this result is that if the pandemic had been allowed to run its course, COVID cases would have continued to stay high and continued to have an effect on UE.

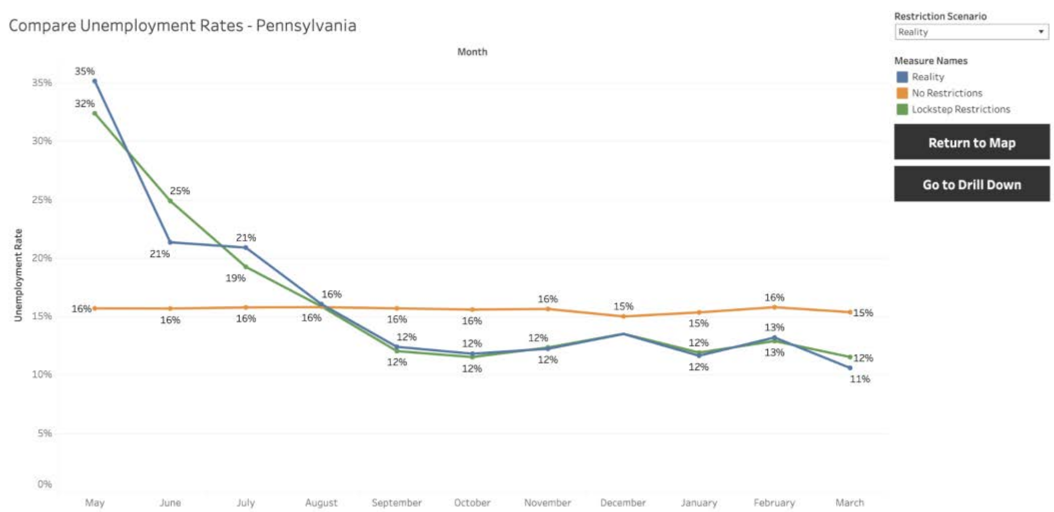

Figure 7: Comparison of Alternate Scenarios - Pennsylvania

6. Limitations and Future Work

Both the models and visualization discussed in the paper are works-in-progress. They have a number of current limitations and, along with those limitations, they carry hope of improvement in future versions.

Limitations include:

- The Household Pulse Survey estimates come with a large degree of uncertainty. We see an opportunity to test various methods of adjusting these estimates in the future to reduce the likelihood of over- or under- statement.
- The models presented in this paper are prototypes. We hope to expand their capabilities and those of the visualization in the future to allow for simulation of custom alternative scenarios. The current tool can be further developed to allow the user more individuality to specify a set of state-imposed restrictions and observe the impact of those restrictions on COVID-19 case counts and UER.
- In the future, we would like to account for effects of social distancing or vaccination administration.
- We would also like to investigate effects by demographics beyond gender.
- This research focused only on a subset of state-level restrictions, specifically those related to closures such as school or workplace closures. We would like to expand this to incorporate other forms of state-imposed restrictions.

References

Barry, J.M., 2017. How the Horrific 1918 Flu Spread Across America. Smithsonian Magazine. Accessed at How the Horrific 1918 Flu Spread Across America | History | Smithsonian Magazine

Centers for Disease Control and Prevention, 2019.1918 Pandemic (H1N1 virus). Accessed at 1918 Pandemic (H1N1 virus) | Pandemic Influenza (Flu) | CDC

COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Inf Dis. 20(5):533-534. doi: 10.1016/S1473-3099(20)30120-1. Accessed at https://github.com/CSSEGISandData/COVID-19.

U.S. Census Bureau. (2020). 2019 American Community Survey 1-year Public Use Microdata Samples [Data file]. Retrieved via tidycensus R package.

Centers for Disease Control and Prevention, 2019.1918 Pandemic (H1N1 virus). Accessed at 1918 Pandemic (H1N1 virus) | Pandemic Influenza (Flu) | CDC

Fields J.F., Hunter-Childs J., Tersine A., Sisson J., Parker E., Velkoff V., Logan C., and Shin H. Household Pulse Survey Data Tables, 2020 and 2021. U.S. Census Bureau. Accessed at https://www.census.gov/programs-surveys/household-pulse-survey/data.html

Maragakis, L.L., 2020. Coronavirus Second Wave? Why Cases Increase. Johsn Hopkins Medicine. Accessed at https://www.hopkinsmedicine.org/health/conditions-and-diseases/coronavirus/first-and-second-waves-of-coronavirus

Flood, S., King, M., Rodgers, R., Ruggles, S., Warren, R.J., Westberry, M. Integrated Public Use Microdata Series, Current Population Survey: Version 9.0 [dataset]. Minneapolis, MN: IPUMS, 2021. https://doi.org/10.18128/D030.V9.0

Sen, P.; T.K. Yamana; S. Kandula; M. Galanti; and J. Shaman, 2021. Burden and Characteristics of COVID-19 in the United States During 2020. Nature, published August 26, 2021. Accessed at Burden and characteristics of COVID-19 in the United States during 2020 | Nature

Thomas Hale, Noam Angrist, Rafael Goldszmidt, Beatriz Kira, Anna Petherick, Toby Phillips, Samuel Webster, Emily Cameron-Blake, Laura Hallas, Saptarshi Majumdar, and Helen Tatlow. (2021). "A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker)." Nature Human Behaviour. https://doi.org/10.1038/s41562-021-01079-8