



# Bird's Eye View of COVID-19, Mobility, and Labor Market Outcomes Across the US

Péter Füleky<sup>1</sup> · István Szapudi<sup>2</sup>

Received: 30 December 2021 / Accepted: 25 March 2022  
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2022

## Abstract

COVID-19 dealt a formidable blow to the US economy. We present a joint analysis of the epidemiological and labor market outcomes across US states. We focus on the relationship across relevant indicators in the pre-vaccination era. As expected, we find strong correlation between changes in economic conditions and mobility. However, mobility fluctuations tend to be uncorrelated with local epidemics and occur simultaneously across most states. The magnitude of the mobility response is highly correlated with the rural vs. urban character of the area. Employment losses are most strongly associated with high population density and concentration of the leisure and hospitality industry. The relationship between job losses and the case fatality ratio is affected by the timing of the most severe COVID-19 waves.

**Keywords** COVID-19 · Case fatality ratio · Mobility · Labor market · Urban character

**JEL Classification** I18 · J28 · O18 · Q54 · R12

## Introduction

The COVID-19 epidemic is a once-in-a-generation shock to the world, including the United States. Its effects on the economy and public health will linger for a long time. The non-pharmaceutical pandemic responses of different states were diverse, and so was the outcome. We analyze epidemic, mobility, and labor-market data to shed light on factors affecting the balance between the economy and health.

The spread and impact of infectious diseases is typically analyzed via structural models that rely on rational actors balancing complex economic and health risks. In particular,

---

✉ Péter Füleky  
fuleky@hawaii.edu

István Szapudi  
istvan@hawaii.edu

<sup>1</sup> UHERO and Department of Economics, University of Hawai'i at Mānoa, 2424 Maile Way, Honolulu, HI 96816, USA

<sup>2</sup> Institute for Astronomy, University of Hawai'i at Mānoa, 2680 Woodlawn Drive, Honolulu, HI 96822, USA

disease spread is usually modeled by variations of the Susceptible–Infected–Recovered (SIR) model originally proposed by Kermack and McKendrick (1927) (see also Zhao et al. 2020; Sanche et al. 2020; Flaxman et al. 2020; Ziff and Ziff 2020; Lloyd-Smith et al. 2005; Adam et al. 2020; Hébert-Dufresne et al. 2020; Allen 2017; Mkhatshwa and Mummert 2010; Szapudi 2020; Yang et al. 2020b). In a departure from this practice, we opt for a data-centric approach that does not require potentially unsubstantiated assumptions. The exponential nature of disease spread and the tendency of COVID-19 to create localized clusters through super-spreading motivates a closer examination of correlations between epidemiology, mobility, and economic data. In particular, we look at the state and county level heterogeneity in the evolution and intensity of COVID-19 along with the mobility and labor market response. Our aim is a general understanding of the basic relationship among these processes.

We take COVID-19 deaths as our primary variable to quantify disease progression, occasionally supplementing it with confirmed cases and hospitalizations. Confirmed cases are sensitive to the number of tests performed, varying state to state and increasing in time as tests became more readily available. The the case fatality ratio (CFR), which is the proportion of deaths relative to cases, decreased over time as health care providers became more effective in treating the disease and testing increased. Both death and hospitalization rates are sensitive to the patients' demography, as the probability of a fatal outcome or hospitalization increases steeply with age. None of the variables are perfect for characterizing the epidemiology, and, especially in smaller counties and earlier times, the data can be lumpy. With these caveats, we use all three variables emphasizing death rates whenever possible.

For characterizing economic conditions, we use state and county-level economic variables, primarily focusing on changes in jobs and employment. The decline in employment directly captures the economic pain of the pandemic-stricken population, and at the same time, it is available at a higher frequency and finer spatial resolution than other indicators characterizing the overall state of the economy.

The critical variable at the intersection of economic conditions and non-pharmaceutical public health intervention is mobility: it is the best quantitative proxy for behavioral response to the pandemic. Note that while behavior is partly a response to local policy, mobility measures what people actually do as opposed to what they are told to do. A decrease in mobility in and of itself will not lower COVID-19 case counts; rather it is the number of encounters favorable to a transmission that matters. Thus, the number of transmissions could vary between urban and rural locations and by lifestyles. Note that mobility is an endogenous variable: a decrease in mobility suppresses the economy, but conversely, lower economic activity results in less mobility. Therefore, we focus on correlations rather than a causal relationship between mobility and the economy.

As the caveats outlined above suggest, the three classes of variables tracking the unfolding of the COVID-19 crisis are fraught with a lot of statistical and systematic uncertainty. Nevertheless, we hope to find a deeper understanding of the relationship between pandemic spread and economic outcomes. In addition to looking at correlations, we also use regression to analyze the dependence of employment fluctuations on labor market composition, epidemiological, and demographic factors.

Our paper contributes to a growing literature documenting the evolution of the COVID-19 pandemic and its impact on the economic downturn, in particular the labor market. Early research on the impact of the COVID-19 pandemic on economic conditions includes Chetty et al. (2020a) and Chetty et al. (2020b). The interplay between mobility and the economy has also been documented widely. Fuleky (2021) uses high frequency mobility

indicators to nowcast economic conditions during the COVID-19 pandemic. A substantial amount of research on labor market outcomes was surveyed by Handwerker et al. (2020). Dreger and Gros (2021) showed a strong relationship between mobility restrictions and unemployment insurance claims, with tightening measures having a 50% greater impact than their easing. The greater vulnerability of dense cities to economic losses was recognized by Cho et al. (2020), who found that larger U.S. cities fared worse than smaller communities. What distinguishes our study from others is that we try to provide a high-level description of public health and economic outcomes, rather than pursuing a more narrowly focused analysis.

## Data

We take advantage of both panel and cross-sectional data to analyze the heterogeneity of the pandemic across the US. Data on COVID-19 cases and deaths are available from several sources, with journalists and researchers complementing the information-gathering carried out by public institutions. We use the data on cases and deaths reported at the daily frequency by the New York Times, mainly due to the ease of programmatic access to the data on GitHub (NYT 2021a). While this and other sources typically reflect the most current and accurate information, the data quality is still affected to a certain extent by a fragmented American public health system. Occasionally, previously unreported cases and deaths are reported as a lump sum, resulting in volatile time series.

During the COVID-19 pandemic, a considerable amount of research tracking economic conditions relied on mobility indicators (for example Fuleky 2021). Such analyses were partly facilitated by easier data accessibility as some private entities made their data sets freely available for non-commercial use. In this study, we use the Mobility and Engagement Index (denoted below as *Mobility*) constructed by the Federal Reserve Bank of Dallas (Atkinson et al. 2020), with the underlying data provided by SafeGraph (2021). The Mobility and Engagement Index summarizes the information in seven different variables based on geolocation data collected from a large sample of mobile devices.

To capture economic fluctuations, we use two labor market indicators released by the US Bureau of Labor Statistics (BLS): state-level payrolls (denoted *Payrolls*) and county-level employment (denoted *Employment*). Estimates of payrolls (or jobs) are based on a survey of businesses, while employment estimates are based on a survey of households. State-level estimates of payrolls are available with an approximately one-month publication delay through the Current Employment Statistics program (BLS 2021a). County estimates of payrolls are only available with a five-month delay through the Quarterly Census of Employment and Wages program (BLS 2021c), but overall employment numbers are available with just a one-month delay through the Local Area Unemployment Statistics program (BLS 2021b); hence we opt for the latter as a proxy for county-level economic conditions.

To fix the frame of reference, percentage changes in payrolls and employment are indexed to February 2020 levels. For example, the percentage change in employment between February 2020 and March 2021 is calculated as

$$\Delta Employment_{Feb20,Mar21} = \frac{Employment_{Mar21} - Employment_{Feb20}}{Employment_{Feb20}} \times 100 \quad (1)$$

Cumulative, or average, percentage changes in payrolls and employment are calculated by compounding the monthly percent deviations from February. For example, the cumulative percentage change in employment over the thirteen-month period between February 2020 and March 2021 is calculated as

$$\overline{\Delta Employment}_{Feb20,Mar21} = \left( \left( \prod_{i=Mar20}^{Mar21} \frac{Employment_i}{Employment_{Feb20}} \right)^{1/13} - 1 \right) \times 100 \quad (2)$$

The national aggregate of the Dallas Fed's Mobility and Engagement Index is scaled so that the national average of January-February is zero, and the lowest national weekly value (week ended April 11, 2020) is -100. Since the starting period is near zero, further conversion of the index to percentage changes is inappropriate. To capture fluctuations, we use simple differences

$$\Delta Mobility_{Feb20,Mar21} = Mobility_{Mar21} - Mobility_{Feb20}$$

and simple averages

$$\overline{\Delta Mobility}_{Feb20,Mar21} = 1/13 \sum_{i=Mar20}^{Mar21} (Mobility_i - Mobility_{Feb20}),$$

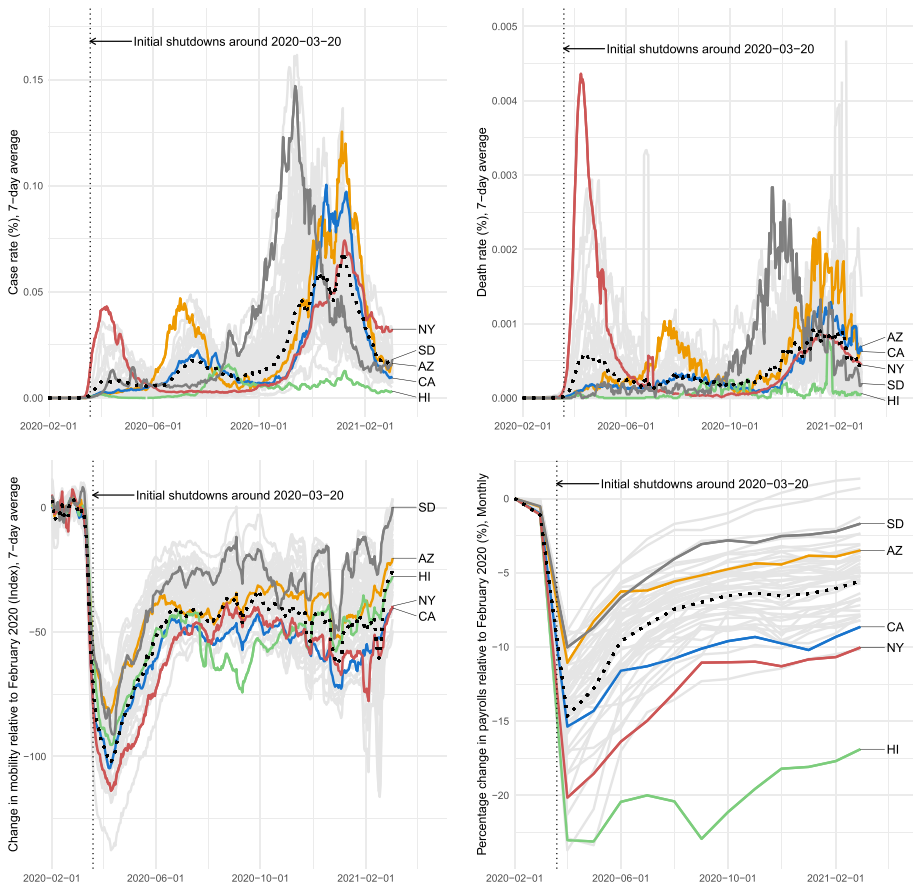
as counterparts of Eqs. (1) and (2), respectively. Case and death counts are recorded in cumulative terms; we use simple differences to capture changes in these epidemiological measures. In all situations, we scale case and death counts by population to obtain case and death rates (denoted  $\Delta Cases$  and  $\Delta Deaths$ , respectively).

All the raw data described above are time series reported at the state or county level. We use both dimensions of the data in panel regressions trying to predict employment fluctuations and Granger causality tests. In our cross-sectional analysis, we use cumulative values as of March 2021, the final month of the sample, and consider additional pre-pandemic characteristics of each location, including demographic measures and labor market composition. Since population and population density (denoted, *Population* and *Pop\_density*, respectively) exhibit an approximately log-normal distribution across counties, we apply a logarithmic transformation to these variables. Demographic indicators—such as the fraction of the population with a Bachelor of Arts or higher degree (denoted *BAddegrees*) or the prevalence of broadband subscription (denoted *Broadband*)—are sourced from the US Census Bureau's 2019 American Community Survey (Census 2019). Since political orientation appeared to affect individual behavior during the pandemic, we also consider electoral outcomes in 2020. We obtained state and county level democratic vote shares in 2020 (denoted *DemocratPct*) from the New York Times (NYT 2021b). We also consider the concentration of different industries at each location captured by the location quotient (denoted  $\cdot LQ$ ) published by the US Bureau of Labor Statistics in the Quarterly Census of Employment and Wages dataset (BLS 2021c). In very small counties, some proportions may be exaggerated or data may be suppressed to avoid disclosing confidential information; therefore we filtered out counties with population of less than 10,000.

## Analysis

The evolution of COVID-19 spread is easiest to describe by the case and death rate, displayed in the top left and right panel, respectively, in Fig. 1. Select states are highlighted with different colors. Both the case and death rates exhibit “wave”-like patterns, with death rates relatively higher during the first wave (the dotted line in the figure represents the US average). The relative moderation of death rates during later waves is likely the result of accumulated experience in treating COVID-19, better preparedness of the healthcare industry, and more extensive testing as resources grew. Overall the three waves rise and fall at an exponential rate, in line with disease spread predicted by the SIR model. However, there is visible heterogeneity in the evolution of these indicators across states, with significant differences in the timing and size of the peaks.

The bottom panels of Fig. 1 display the non-pharmaceutical response. It is remarkably synchronized across the states, with no visible variation in the timing of state-specific responses to the epidemiological waves. Both the overall decline in mobility (bottom left

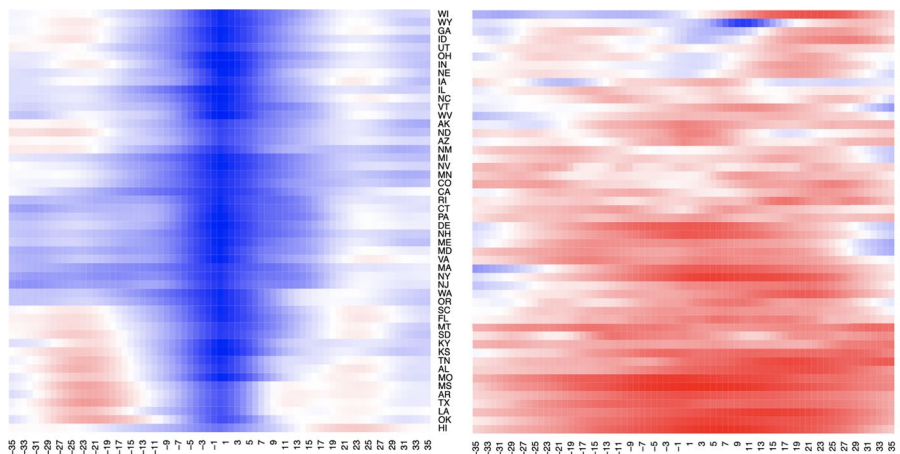


**Fig. 1** The top left (right) panel displays the case-rate (death-rate) per day averaged over a seven day period. The bottom left (right) panel displays the changes in mobility (payrolls) relative to the pre-COVID-19 levels. The dotted line in each figure is the national rate or level

panel) and the evolution of job losses (bottom right panel) varied across states. However, after the first wave, they did not mirror the cyclical pattern exhibited by the epidemiological indicators. Since much economic activity is associated with the movement of labor and goods, the strong correlation between mobility and economic damage—both in time and in degree—is not surprising. Seasonal effects, such as the Labor Day, Thanksgiving, and Christmas holidays, drive some high-frequency fluctuations in mobility; hence they are synchronized across all states. The degree of the mobility drop, as we show later, is related to the urban-rural character of the states rather than the strength of local epidemiological waves. A surprising consequence is that the local response was relatively insensitive to the heterogeneity and severity of the pandemic across states. The first dip occurred in the wake of nationwide shutdowns as cases in some states began to rise rapidly. After the initial shutdowns were lifted, mobility rebounded, and the labor market embarked on an almost monotonous albeit slowing path to recovery.

It is worth noting here that Hawaii's experience is unique among the US states. It is an outlier in both the epidemiological and the economic dimension; it had the lowest case and death rate in the nation, but at the same time it suffered the highest job losses. Hawaii also seemed to be more responsive to local waves of the virus than other states, with both mobility and jobs dipping during the worsening epidemiological conditions in late summer and then slowing again in the winter.

The correlograms in Fig. 2 illustrate the contemporaneous changes in mobility across states (left) and the lack of a coincidence between changes in local deaths and local mobility (right). These plots are based on the gradients of the variables depicted in the top right and bottom left panels of Fig. 1: the cross-correlations are calculated using the 4-week *change* in the death rate and mobility. To avoid the dominant impact of the initial shutdowns, the data underlying the correlograms is restricted to the post-shutdown period between June 2020 and March 2021. The dark blue color in the left panel indicates high positive correlation between the national average and local mobility gradients. The narrow



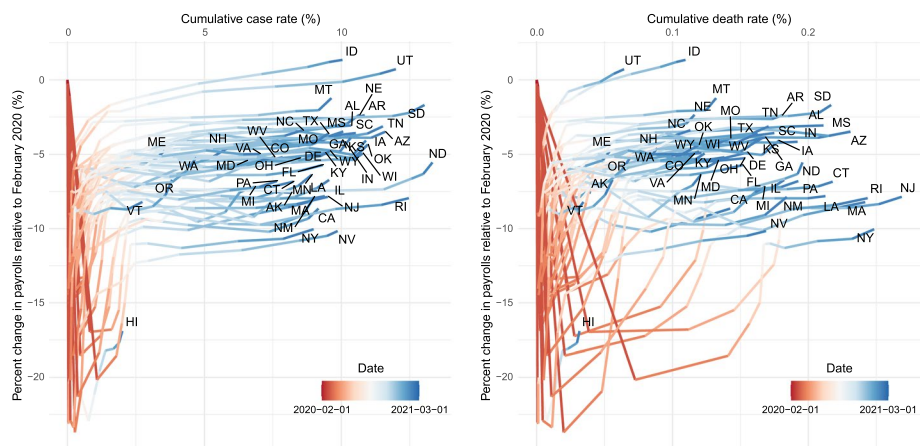
**Fig. 2** Left panel: cross-correlation between nationwide mobility and local mobility gradients, right panel: cross-correlation between local death rate and local mobility gradients (blue = positive correlation, red = negative correlation). The gradients are 4-week changes in the underlying variables displayed in Fig. 1. For negative lags, the first variable leads the second one by the indicated number of days. For positive lags, the first variable lags the second one by the indicated number of days. A measure of pattern similarity across states is the average pairwise correlation of the rows, which is 0.89 in the left and 0.28 in the right panel



width of the dark blue vertical band around zero lag indicates near-contemporaneous change of mobility across states. The tight relationship across states is also evidenced by a 0.89 average pairwise correlation of the rows in the left panel.

In contrast, there is no single lag that dominates the correlations across states in the right panel. Red color denotes negative cross-correlation between the death rate and mobility gradients: 1) for negative lags, it indicates that an increase in deaths is typically followed by a decrease in mobility or vice versa, 2) for positive lags, it indicates that a decrease in mobility is followed by an increase in deaths or vice versa. The positive near-contemporaneous correlation between local deaths and mobility in blue color at the top of the right panel indicates that in those states an increase in mobility is associated with an increase in deaths or vice versa. While one can come up with explanations for these patterns, they tend to be weak and heterogeneous across states (the average pairwise correlation of the rows is only 0.28). The prevalence of negative correlation implies a tradeoff between deaths and mobility, but it is impossible to pinpoint a typical lag between changes in deaths and changes in mobility. The relationship between local outbreaks and mobility appears to be idiosyncratic across US states. The conclusions are the same when we look at the cross-correlations between the case rate and mobility gradients.

Figure 3 depicts the relationship between job losses (approximating economic conditions) and the epidemiological impact of the COVID-19 virus (approximated by cumulative case rates and death rates, where the latter are less sensitive to testing and other local idiosyncrasies). When the virus first began spreading in the US, both case and death rates were affected by the novelty of the disease: testing capability only ramped up slowly. Hence, the case detection rate was lower initially, and effective treatments were only discovered gradually, resulting in a disproportionately large number of deaths (or case fatality ratio) initially. The lines track each state’s path in this two-dimensional space from February 2020 (origin) to March 2021 (endpoint marked with state label). The more to the right a state is, the less successful it has been on the epidemiological front. On the vertical axis, economic success and failure correspond to staying near the top and sinking to the bottom,



**Fig. 3** Scatterplot of the relationship between job losses (approximating economic conditions) and the epidemiological impact of the COVID-19 virus (approximated by cumulative case rates in the left and death rates in the right panel). The lines track each state’s path at monthly steps in this two-dimensional space from February 2020 (origin) to March 2021 (endpoint marked with state label)

respectively. A trade-off between the disease spread and economic growth would suggest a scatter plot extending from the bottom left to the top right. As we noted above, Hawaii is an outlier. Removing it from the sample leaves us with no clear relationship between the case or death rate and the time dimension of the scatterplots. Instead, it seems that every state went through two distinct phases, emphasized by a color-coded time dimension. While containing the spread of the disease in most states, the initial shutdowns inflicted the most significant economic damage and led to large job losses (vertical drop in red). After a subsequent partial recovery, the labor market stagnated, and the virus continued to spread (lateral move in blue).

In March 2021, most states had about 3% to 10% fewer jobs than before the recession. However, a closer inspection reveals that the economic damage across states is associated with their urban-rural character. This is a consequence of the relatively coarse non-pharmaceutical measures, and the tight correlation of mobility with both economic conditions and the urban-rural character of a state. Before quantifying these relationships, we construct an index describing the urban-rural character of states. Specifically, we extract the first principal component (denoted *PC1*) from five related variables: the share of the population with BA or higher degree, the prevalence of broadband access, the share of the vote for the Democratic Party in the 2020 election, the size of the population and population density. All these variables tend to be strongly associated with the urban-rural dimension. The first principal component summarizes that dimension well, capturing nearly 60% of the variance in the underlying variables. Table 1 summarizes the pairwise correlations between the first principal component and the underlying variables as well as the pandemic outcomes: decline in payrolls and in mobility, the cumulative number of cases and deaths. The correlations are obtained from the cross-section of states. The shaded lower triangle presents the—very similar—results of a robustness check that excludes from the sample Hawaii, Idaho and Utah, the states with the most extreme labor market outcomes (see also Fig. 3).

Not surprisingly, *PC1* has a strong correlation with the underlying variables describing the level of urbanization across states. *PC1* in turn is highly correlated with changes

**Table 1** Correlation matrix

	$\overline{\Delta Payr}$	$\overline{\Delta Mobi}$	$\Delta Case$	$\Delta Deat$	<i>PC1</i>	<i>BAdeg</i>	<i>Broadb</i>	<i>Democ</i>	<i>Popul</i>	<i>PopDen</i>
$\overline{\Delta Payr}$	1.00	0.55	0.51	0.01	-0.46	-0.31	-0.27	-0.64	-0.06	-0.34
$\overline{\Delta Mobi}$	0.57	1.00	0.23	-0.24	-0.86	-0.72	-0.54	-0.80	-0.39	-0.70
$\Delta Case$	0.34	0.17	1.00	0.61	-0.40	-0.36	-0.30	-0.55	0.03	-0.15
$\Delta Deat$	-0.10	-0.26	0.64	1.00	0.01	-0.10	-0.33	-0.05	0.29	0.38
<i>PC1</i>	-0.50	-0.86	-0.40	0.02	1.00	0.88	0.74	0.88	0.45	0.70
<i>BAdeg</i>	-0.37	-0.73	-0.39	-0.09	0.88	1.00	0.78	0.75	0.14	0.40
<i>Broadb</i>	-0.38	-0.59	-0.34	-0.28	0.77	0.79	1.00	0.55	0.08	0.19
<i>Democ</i>	-0.61	-0.78	-0.50	-0.04	0.90	0.78	0.61	1.00	0.27	0.58
<i>Popul</i>	-0.14	-0.41	-0.02	0.25	0.47	0.14	0.11	0.29	1.00	0.57
<i>PopDen</i>	-0.31	-0.68	-0.10	0.41	0.69	0.40	0.21	0.56	0.58	1.00

The correlations in shaded cells are based on a sample excluding Hawaii, Idaho, and Utah. The variables  $\overline{\Delta Payrolls}$  and  $\overline{\Delta Mobility}$  represent the average percentage change in the respective indicator relative to February 2020 during the period March 2020-March 2021. The variables  $\Delta Cases$  and  $\Delta Deaths$  are the cumulative levels of these indicators (equivalent to the difference between February 2020 and March 2021). *PC1* is an index capturing the level of urbanization across states; it is the first principal component in a data matrix consisting of *BAdegrees*, *Broadband*, *DemocratPct*, *Population* and *PopDensity*.

in mobility. The correlation between mobility and jobs is slightly weaker but—as expected—positive. However, while case and death rates have a fairly strong positive



relationship across states, they appear to have relatively weak and opposite relationships with the non-epidemiological indicators. Cumulative case rates are positively, while cumulative death rates are negatively or not at all correlated with changes in jobs and mobility. Finally, the correlation between the urban-rural character and the cumulative case and death rates is negative and zero, respectively. To summarize, states experiencing high death rates have typically seen the greatest reduction in mobility, while states with high case rates were typically rural and have seen fewer losses in jobs and mobility.

In addition to inspecting pairwise correlations, we also estimate a linear regression with county-level data to shed light on the marginal relationships between labor market outcomes and underlying conditions. The county-level analysis affords a larger sample and greater detail in the predictor space. The explanatory variables control for epidemiological conditions, demography, mobility, and labor market composition. In the regression equation

$$\begin{aligned} \overline{\Delta Employment}_{Feb20,Mar21,c} = & \alpha_0 + \alpha_1 \overline{\Delta Mobility}_{Feb20,Mar21,c} \\ & + \alpha_2 \Delta Cases_{Feb20,Mar21,c} + \alpha_3 \Delta Deaths_{Feb20,Mar21,c} \\ & + \alpha_4 DemocratPct_{2020,c} + \alpha_5 Population_{2019,c} + \alpha_6 PopDensity_{2019,c} \\ & + \alpha_7 BAdegrees_{2019,c} + \alpha_8 Broadband_{2019,c} + \alpha_9 MiningLQ_{2019,c} \\ & + \alpha_{10} ConstructionLQ_{2019,c} + \alpha_{11} ManufacturingLQ_{2019,c} \\ & + \alpha_{12} RetailLQ_{2019,c} + \alpha_{13} InformationLQ_{2019,c} + \alpha_{14} FinanceLQ_{2019,c} \\ & + \alpha_{15} BusinessServicesLQ_{2019,c} + \alpha_{16} EduHealthLQ_{2019,c} \\ & + \alpha_{17} LeisureHospLQ_{2019,c} + \alpha_{18} OtherServicesLQ_{2019,c} + \epsilon_c, \end{aligned} \quad (3)$$

the subscript  $c$  is a county index,  $\Delta$  and  $\overline{\Delta(\cdot)}$  indicate temporal differences described in “Data” section, and  $\epsilon$  is assumed to be an independently and identically distributed random error. Again, the  $\cdot LQ_{2019,c}$  variables are location quotients capturing the concentration of jobs in industry  $\cdot$  in year 2019 in county  $c$ . All variables are standardized with respect to their distribution across counties, and therefore the coefficients can be interpreted as the employment change (measured in standard deviations) associated with a one standard deviation change in a particular predictor. Note, this conditional analysis does not imply causality, but it helps to uncover the strength of association between the predictors and the outcome.

The regression results are presented in Table 2. Although conditioning on multiple explanatory variables simultaneously, the county-level regression results lead to similar conclusions as the state-level bivariate correlations. Changes in employment are positively associated with changes in mobility and cases, but negatively with deaths. Of the variables related to urbanization, population density has the strongest relationship with job losses. As seen in Table 1, *BAdegrees* and *Broadband* are approaching collinearity, and along with *Population*, their direct correlations with job losses are weak, hence their insignificant coefficient estimates are not surprising. The location quotient coefficient estimates indicate that counties with high concentrations of the leisure and hospitality, mining, and construction industries tended to be the most vulnerable to large job losses. Conditionally on the other variables in the model, a one standard deviation difference in *PopDensity*, *LeisureHospLQ*, and  $\Delta Deaths$  is associated with 0.39, 0.23, and 0.18 standard deviations of job losses, respectively, across the counties in the sample. While these results are suggestive, the coefficient of determination signals that 73% of the variation in the dependent variable is *not* explained by the model. Yet, diagnostic tests do not reveal further systematic variation in the residuals.

**Table 2** Linear regression results

Dependent variable: $\overline{\Delta Employment}$		
Coefficient	Estimate	p-value
$\overline{\Delta Mobility}$	<b>0.145</b>	<b>0.000</b>
$\Delta Cases$	<b>0.130</b>	<b>0.000</b>
$\Delta Deaths$	<b>-0.183</b>	<b>0.000</b>
$BAdegrees$	0.026	0.506
$Broadband$	0.013	0.700
$DemocratPct$	<b>-0.120</b>	<b>0.000</b>
$Population$	0.024	0.562
$PopDensity$	<b>-0.388</b>	<b>0.000</b>
$MiningLQ$	<b>-0.171</b>	<b>0.000</b>
$ConstructionLQ$	<b>-0.131</b>	<b>0.000</b>
$ManufacturLQ$	-0.064	0.232
$InformationLQ$	0.037	0.082
$FinanceLQ$	0.008	0.711
$BusinessServLQ$	-0.003	0.941
$EduHealthLQ$	-0.016	0.646
$LeisureHospLQ$	<b>-0.231</b>	<b>0.000</b>
$OtherServicesLQ$	0.024	0.301
$RetailLQ$	<b>0.060</b>	<b>0.037</b>

Number of observations: 2027,  $R^2 = 0.27$

The dependent variable,  $\overline{\Delta Employment}$ , is the average percentage change in employment relative to February 2020 during the period March 2020-March 2021. The shaded cells contain estimates of industry importance. All variables are scaled to zero mean and unit standard deviation. Estimates that are significant at the 5% level of marginal significance are printed in bold font. The sample consists of 2027 counties with data available in each of the predictors. The coefficient of determination implies that the estimated Eq. (3) explains 27% of the variation in the dependent variable

The purely cross-sectional regression in Eq. 3 explains cumulative employment outcomes after the first year of the pandemic, but the two-dimensional structure of employment, mobility and case rates allowed us to also explore some relationships via panel regressions. (Since deaths exhibited only limited fluctuation in many counties, and therefore were collinear with the fixed effect, we used cases as the epidemiological variable for this analysis.) The robust Hausman test described in Section 10.7.3. of (Wooldridge 2010) rejected the consistency of the random effects model, and the Lagrange multiplier test for the error components model (Baltagi et al. 1992) suggested to use the two-ways fixed effect panel regression

$$\Delta Employment_{t,c} = month_t + county_c + \beta_{mob} \Delta Mobility_{t,c} + \beta_{cas} \Delta Cases_{t,c} + \epsilon_{t,c} \quad (4)$$

where the fixed effects  $month_t$  and  $county_c$  capture time effects constant across counties and county specific characteristics constant over time, respectively. The dependent variable is

the month-to-month percent change in employment, and the explanatory variables are the month-to-month changes in mobility and case rates.

Since the panel variables are not standardized, the interpretation of the coefficient estimates is tied to the units of the underlying variables rather than providing a ranking of effects by magnitude. So instead of looking at the numerical values of the coefficients, let us just note that in the panel regression Eq. (4)  $\Delta Mobility_{t,c}$  is a significant predictor of  $\Delta Employment_{t,c}$ , but  $\Delta Cases_{t,c}$  is not. That is, conditionally on the other explanatory variables in the model,  $\Delta Cases_{t,c}$  does not have significant relationship with  $\Delta Employment_{t,c}$ , further underlining the lack of employment response to waves of the pandemic. We confirmed the robustness of this result by allowing full model heterogeneity across counties (Pesaran 2006). The mean group estimates provided the same qualitative conclusion that, conditionally on mobility, case rates did not have a significant relationship with employment.

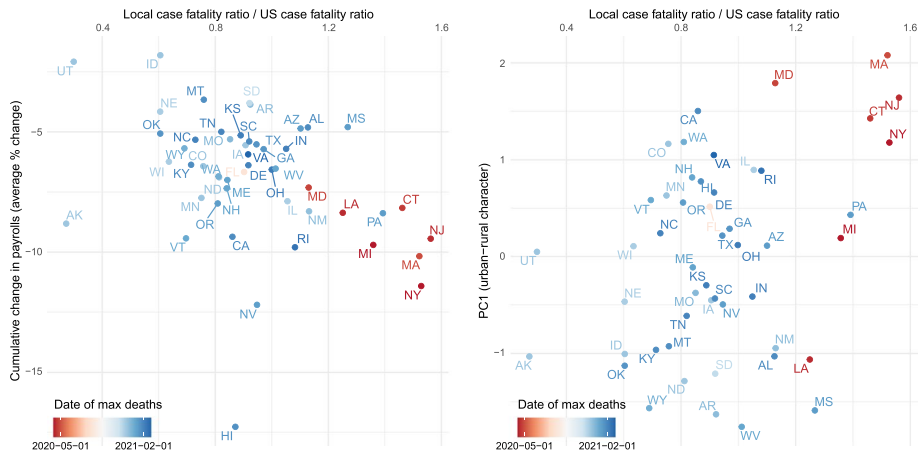
Finally, we make use of the two-dimensional structure of these variables to carry out pairwise Granger causality tests (Dumitrescu and Hurlin 2012). The panel version of the test determines whether a time series is useful for forecasting another *in at least one county*, which turns out to be the case for all pairwise combinations of  $\Delta Employment_{t,c}$ ,  $\Delta Mobility_{t,c}$ ,  $\Delta Cases_{t,c}$ . For our purposes it is much more revealing to see the proportion of counties where the lagged predictor has significant explanatory power. By looking at Table 3 we can draw two conclusions. First, there are very few counties in which changes in the case rate Granger cause changes in either employment or mobility. This again underscores the decoupling of economic response from the waves of the pandemic. Second, the proportion of counties with significant results is very similar for the direct and the reverse tests. Thus, there is no evidence for a directional causal relationship. Instead employment and mobility are endogenously fluctuating together and there seems to be a very weak relationship between these and case rates.

As established above, the heterogeneity of the COVID-19 response—in terms of mobility and economic damage—is highly correlated with the urban-rural character ( $PC1$ ) of a location. Figure 4 connects these variables with yet another measure of epidemiological outcomes. It shows the scatterplot of job losses (left panel) and  $PC1$  (right panel) against excess CFR, defined as the ratio of the local CFR and the national CFR. Excess CFR is lowered by both higher than average testing and better than average medical care, but we are not able to distinguish these two effects with our data. As Fig. 4 illustrates, lower CFR

**Table 3** Granger causality results

y	x	Significant proportion
$\Delta Employment$	$\Delta Mobility$	0.671
$\Delta Mobility$	$\Delta Employment$	0.490
$\Delta Employment$	$\Delta Cases$	0.010
$\Delta Cases$	$\Delta Employment$	0.012
$\Delta Mobility$	$\Delta Cases$	0.009
$\Delta Cases$	$\Delta Mobility$	0.012

The Granger causality test is based on the regression  $y_{t,c} = \alpha_0 y_{t-1,c} + \beta_0 x_{t-1,c} + \epsilon_{t,c}$  carried out for each county  $c$ . The sample ranges from February 2020 to March 2021. The rows of the table indicate results for various combinations of the  $y$  and  $x$  variables. The third column lists the proportion of counties with significant Granger causality tests



**Fig. 4** The local to national CFR ratio related to average job losses on the left and urban-rural character of states on the right. The job losses between February 2020 and March 2021 are calculated according to Eq. 2. Higher values of PC1 indicate stronger urban character of states. The color of observations represents the date of the maximum 7-day average death rate. The correlation coefficient is  $-0.38$  and  $0.35$  between the variables in the left and right figure, respectively

tends to be associated with lower economic damage, but the timing of the worst COVID-19 waves explains part of the relationship. The first wave, hitting highly urbanized North-Eastern states, caught the medical community off guard: many cases went unnoticed due to the lack of tests, but deaths soared. On the other hand, some rural states were spared from the brunt of the pandemic until the fall. By then, greater awareness, understanding of the disease, availability of tests and more resources targeted towards medical care had a mitigating effect on CFR. This underscores the value of early non-pharmaceutical responses to both flatten the curve, allowing for better use of scarce medical resources, and to delay the crest of the various COVID waves, allowing for time to develop better pharmaceutical interventions. The colors in Fig. 4 highlight the importance of timing: they indicate the date of maximum deaths in both panels. All the states hit by COVID-19 early have positive excess CFR. They also tend to be more urban and tend to have worse economic outcomes.

## Conclusion

We explored broad relationships between epidemiology and economy during the first year of the COVID-19 pandemic. Although the magnitude of the mobility response varied highly across states, the timing of the fluctuations was remarkably uniform. Instead of a flexible local response, the US states moved in apparent unison. Following the initial shutdowns, mobility recovered partially, and subsequently it exhibited only modest contemporaneous swings across states. Given the exponential nature of disease spread, and the difference in timing and degree of local outbreaks, such uniform mobility response is surprising; the behavior across states was likely influenced by the importance of national news. Although, the local mobility response does not appear to be a function of local outbreaks, urban regions took a more extensive hit both in terms of mobility and labor market outcomes. The high correlation between mobility and labor market conditions is intuitively clear since most economic activity requires movement

of people and goods. The urban-rural divergence stems in part from the different lifestyles and the different amount of mobility associated with economic activity in these environments.

While the exponential nature of COVID-19 spread explains its wave-like pattern, it is worth noting that most mitigation tools, such as testing capacity, contact tracing, number of intensive care unit beds and ventilators, and even daily vaccination capacity, are linear. This mismatch between COVID-19 spread and available mitigation tools, the short time scale for decision making, and the unfamiliarity of the public and political leaders with exponential growth hindered the emergence of nimble local responses. Although the lack of localized response is surprising, we must add that local response has a limited power to save the part of the local economy that is interconnected to other regions, even if it saves lives. The clearest example is Hawaii with its exceptionally low case and death rate but suffering the worst economic damage among all US states. Hawaii's economy started to recover only when the global environment improved enough to support travel and tourism.

What can we conclude based on our results? We only see limited evidence for a tradeoff between public health and the economy across US states. Tradeoff would imply a (much stronger) positive correlation in the scatterplot of states in Fig. 3. If the correlation were stronger, we could argue about the optimal response along that relationship based on ethical, economical, even political and life style grounds. Due to the lack of such correlation, we can only speculate that many of the deaths could have been avoided without further economic damage, if only a more nimble and localized policy had been followed. The "two-size fits all" urban-rural policy response with global timing appears to be inadequate. A caveat to the above conclusion is that job losses are associated with CFR. While it makes common sense that higher than average mortality impedes economic activity, this ratio depends on the quality and availability of health care and on the amount of testing. These dependencies are difficult to quantify and we cannot speculate which, if any, dominates the observed correlation. However, we found that the timing of outbreaks had a significant effect: communities hit by COVID-19 early on had both worse CFR and greater economic losses.

If these data are any guide to the future, they show that we need to improve localized response. The exponential nature of disease spread is likely to be true for any future pandemic. Extensive surveillance coupled with swift and decisive local response, restrictions on mobility and travel immediately after discovery but only for limited time and at the source of outbreaks, would require much more coordination and collaboration between the public and local authorities. Consistent and clear communication, quick reaction to misinformation, and familiarizing the public with exponential growth to motivate timely interventions would avoid lackluster response (see for example Yang et al. 2020a). But even then, supply chain disruptions could reverberate for a while despite limited and ultra-local interventions.

**Funding** The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

**Data Availability** The data analyzed in the current study are publicly available from the sources indicated in "Data" section.

## Declarations

**Competing interests** The authors have no relevant financial or non-financial interests to disclose.

## References

- Adam D, Wu P, Wong J, Lau E, Tsang T, Cauchemez S, Leung G, Cowling B (2020) Clustering and superspreading potential of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infections in Hong Kong
- Allen LJ (2017) A primer on stochastic epidemic models: Formulation, numerical simulation, and analysis. *Infect Dis Model* 2(2):128–142
- Atkinson T, Dolmas J, Koch C, Koenig E, Mertens K, Murphy A, Yi K-M (2020) Dallas fed mobility and engagement index gives insight into COVID-19's economic impact
- Baltagi BH, Chang Y-J, Li Q (1992) Monte carlo results on several new and existing tests for the error component model. *J Econ* 54(1-3):95–120
- BLS (2021a) Current employment statistics. <https://download.bls.gov/pub/time.series/sm/>. Accessed 30 May 2021
- BLS (2021b) Local area unemployment statistics. <https://download.bls.gov/pub/time.series/la/>. Accessed 30 May 2021
- BLS (2021c) Quarterly census of employment and wages. <https://www.bls.gov/cew/>. Accessed 30 May 2021
- Census (2019) American community survey. <https://www.census.gov/programs-surveys/acs>. Accessed 30 May 2021
- Chetty R, Friedman JN, Hendren N, Stepner M et al (2020a) How did COVID-19 and stabilization policies affect spending and employment? a new real-time economic tracker based on private sector data. Technical report, National Bureau of Economic Research
- Chetty R, Friedman JN, Hendren N, Stepner M, Team TOI (2020b) The economic impacts of COVID-19: Evidence from a new public database built using private sector data. Working Paper 27431, National Bureau of Economic Research
- Cho SJ, Lee JY, Winters JV (2020) Employment impacts of the COVID-19 pandemic across metropolitan status and size. Technical report Bonn, Germany: IZA—Institute of Labor Economics
- Dreger C, Gros D (2021) Lockdowns and the us unemployment crisis. *Econ Disasters Climate Chang* 5(3):449–463
- Dumitrescu E-I, Hurlin C (2012) Testing for granger non-causality in heterogeneous panels. *Econ Model* 29(4):1450–1460
- Flaxman S, Mishra S, Gandy A, Unwin H, Coupland H, Mellan T, Zhu H, Berah T, Eaton J, Perez Guzman P, Schmit N, Cilloni L, Ainslie K, Baguelin M, Blake I, Boonyasiri A, Boyd O, Cattarino L, Ciavarella C, Cooper L, Cucunuba Perez Z, Cuomo-Dannenburg G, Dighe A, Djaafara A, Dorigatti I, Van Elsland S, Fitzjohn R, Fu H, Gaythorpe K, Geidelberg L, Grassly N, Green W, Hallett T, Hamlet A, Hinsley W, Jeffrey B, Jorgensen D, Knock E, Laydon D, Nedjati Gilani G, Nouvellet P, Parag K, Siveroni I, Thompson H, Verity R, Volz E, Walters C, Wang H, Wang Y, Watson O, Winskill P, Xi X, Whitaker C, Walker P, Ghani A, Donnelly C, Riley S, Okell L, Vollmer M, Ferguson N, Bhatt S (2020) Report 13: Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries
- Fuleky P (2021) Nowcasting the trajectory of the COVID-19 recovery. *Applied Economics Letters*, forthcoming
- Handwerker EW, Meyer PB, Piacentini J, Schultz S, Sveikauskas L (2020) Employment recovery in the wake of the COVID-19 pandemic. Technical report, U.S. Bureau of Labor Statistics
- Hébert-Dufresne L, Althouse BM, Scarpino SV, Allard A (2020) Beyond  $r_0$ : Heterogeneity in secondary infections and probabilistic epidemic forecasting
- Kermack WO, McKendrick AG (1927) A contribution to the mathematical theory of epidemics. *Proc R Soc Lond Ser A, Containing Pap Math Phys Character* 115(772):700–721
- Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM (2005) Superspreading and the effect of individual variation on disease emergence. *Nature* 438 (7066):355–359
- Mkhatshwa T, Mummert A (2010) Modeling super-spreading events for infectious diseases. Case study sars
- NYT (2021a) Coronavirus (COVID-19) data in the united states. <https://github.com/nytimes/covid-19-data>. Accessed 30 May 2021
- NYT (2021b) Presidential election results 2020. [https://raw.githubusercontent.com/MEDSL/election\\_night\\_2020/main/nyt\\_counties\\_final12042020.csv](https://raw.githubusercontent.com/MEDSL/election_night_2020/main/nyt_counties_final12042020.csv). Accessed 30 May 2021
- Pesaran MH (2006) Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica* 74(4):967–1012
- SafeGraph (2021). <https://www.safegraph.com/academics>. Accessed 30 May 2021
- Sanche S, Lin YT, Xu C, Romero-Severson E, Hengartner N, Ke R (2020) High contagiousness and rapid spread of severe acute respiratory syndrome coronavirus 2. *Emerg Infect Dis* 26(7):1470–1477



- Szapudi I (2020) Heterogeneity in SIR epidemics modeling: superspreaders and herd immunity. *Appl Netw Sci* 5(1)
- Wooldridge JM (2010) *Econometric analysis of cross section and panel data*. MIT Press, Cambridge
- Yang M-J, Gaulin M, Seegert N, Looney A (2020a) What drives the effectiveness of social distancing in combating COVID-19 across U.S. States? Technical report, University of Utah
- Yang M-J, Seegert N, Gaulin M, Looney A, Orleans B, Andrew P, Stratford K, Samore M (2020b) What is the active prevalence of COVID-19? Technical report, University of Utah
- Zhao S, Lin Q, Ran J, Musa SS, Yang G, Wang W, Lou Y, Gao D, Yang L, He D, Wang MH (2020) Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in china, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak. *Int J Infect Dis* 92:214–217
- Ziff AL, Ziff RM (2020) Fractal kinetics of COVID-19 pandemic

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.